cisco

ciscopress.com

IPv6在企业网络中的部署

IPv6 for Enterprise Networks

The practical guide to deploying IPv6 in campus, WAN/branch, data center, and virtualized environments

〔美〕 **Shannon McFarland,** CCIE #5245
Muninder Sambi, CCIE #13915
Nikhil Sharma, CCIE #21273
Sanjay Hooda, CCIE #11737
著

孙余强　孙剑　译

人民邮电出版社
POSTS & TELECOM PRESS

IPv6在企业网络中的部署

IPv6 for Enterprise Networks

〔美〕 **Shannon McFarland,** CCIE #5245
Muninder Sambi, CCIE #13915
Nikhil Sharma, CCIE #21273
Saniav Hooda, CCIE #11737
著

孙余强 孙剑 译

人 民 邮 电 出 版 社
北 京

图书在版编目（CIP）数据

IPv6在企业网络中的部署 /（美）麦克法兰德（McFarland, S.）等著；孙余强，孙剑译. —北京：人民邮电出版社，2012.1
ISBN 978-7-115-26836-5

Ⅰ. ①I… Ⅱ. ①麦… ②孙… ③孙… Ⅲ. ①企业－计算机网络－通信协议 Ⅳ. ①TP393.18

中国版本图书馆CIP数据核字(2011)第229450号

版 权 声 明

IPv6 在企业网络中的部署

◆ 著 [美] Shannon McFarland Muninder Sambi
Nikhil Sharma Sanjay Hooda

译 孙余强 孙 剑

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 电子邮件 315@ptpress com cn

网址 http //www ptpress com cn

北京昌平百善印刷厂印刷

◆ 开本：800×1000 1/16

印张：27 25

字数：553 千字 2012 年 1 月第 1 版

印数：1－3 000 册 2012 年 1 月北京第 1 次印刷

著作权合同登记号 图字：01-2011-4671 号

ISBN 978-7-115-26836-5

定价：69.00 元

读者服务热线：(010)67132705 印装质量热线：(010)67129223
反盗版热线：(010)67171154

内容提要

本书包含了在企业网内部署 IPv6 的所有知识，包含的主要内容有：启用 IPv6 的市场驱动力、IPv6 企业网络设计的层次化模型、企业网各区块内的 IPv6 部署方案、IPv6 网络管理，以及 IPv6 实验、试点网络环境的搭建等。

本书层次分明、阐述清晰、分析透彻、理论与实践并重，不仅适合准备 CCNA、CCNP 或 CCIE 认证考试的人员阅读，也是从事计算机网络设计、管理和运维工作的工程技术人员必不可少的参考资料。

关于作者

Shannon McFarland，CCIE #5245，在 Cisco 公司任企业咨询工程师一职，是企业网 IPv6 部署和数据中心设计方面的技术顾问，专注于应用程序部署和虚拟化桌面基础设施领域的研究。16 年以来，他从事过的工作包括大型企业园区网络、WAN/分支机构网络的设计；数据中心网络设计和微软操作系统以及服务器应用程序的优化；虚拟桌面基础设施的设计、部署和优化。最近 10 年，Shannon 经常参加各种全球性的 IPv6 活动（尤其是 Cisco Live[前身为 Networkers]）、IPv6 峰会以及业界的其他活动，并踊跃发言。他发表过多篇论文和 Cisco 验证设计书（Cisco Validated Designs [CVD]），所涉领域包括 IPv6、多播、Microsoft Exchange、VMware View，以及其他各类应用。他还是多部 Cisco Press 出版的图书的特约撰稿人。加盟 Cisco 之前，Shannon 曾在某增值经销商做过网络顾问，并在医疗行业当过网络工程师。目前，Shannon 与妻儿共居于科罗拉多州的罗克堡。

Muninder Sambi，CCIE #13915，Cisco Catalyst 4500/4900 系列平台产品营销经理。身为一名产品线经理，他负责为 Catalyst 4500 和 4900 系列平台——其中包括可供园区网用户和数据中心服务器接入的下一代新产品架构——制定产品战略决策。在接手该工作之前，Muninder 亦身居要职——为 Cisco 模块化交换平台制定长期软件和服务战略决策，工作的重点包括 IPv6 的技术革新。其中的一些新技术使得某些大企业和运营商成功部署了 IPv6/IPv4 双协议栈。Muninder 还是 Cisco IPv6 开发委员会的核心成员。Muninder 曾代表 Cisco 评审过大企业客户的多个网络设计框架。过去 12 年以来，Muninder 曾参与设计过多个企业园区网络、WAN 和数据中心网络。加盟 Cisco 以前，Muninder 在印度一家大型网络集成商担任网络顾问一职，并负责设计和实施 LAN、WAN 及托管数据中心网络。Muninder 现与妻儿生活在加州福里蒙特市。

Nikhil Sharma，CCIE #21273，在 Cisco 公司任技术营销工程师一职，负责为 Catalyst 4500 产品线的软硬件“雕琢”新特性。10 多年来，Nikhil 为不同的企业客户设计了多个大中型园区网络和数据中心网络，并参与其中的故障排除工作。

Sanjay Hooda，CCIE #11737，是 Cisco 公司的一名技术领导，专攻嵌入式系统，并协助定义新产品的体系结构。他当前关注的领域包括高可用性和大型分布式交换系统中的消息传递。过去 14 年以来，Sanjay 参与过的工作包括 SCADA（Supervisor 控制和数据获取）、许多大型软件项目，以及企业园区网网络、LAN、WAN 与数据中心网络的设计。

关于技术审稿人

Jim Bailey，CCIE #5275 (RS/SP 双料 CCIE)和 CCDE #20090008，在 Cisco 公司任 AS（高级服务）团队的技术领导，拥有 18 年网络技术相关工作经验。作为全球政府解决方案组高级服务团队成员，他专注于美国政府民用机构网络和军用网络的架构、设计和实施。过去 5 年来，他一直致力于在上述网络中集成 IPv6。

Ciprian P. Popoviciu 博士，Technodyne 公司企业服务团队 Cloud and Network3.0 practices 主管。之前，他曾在 Cisco 公司的多个岗位担任领导工作，在 Cisco 任职的最后 8 年里，他曾就 IPv6 协议和产品开发、IPv6 战略规划、IPv6 的启用，以及下一代网络架构和部署等，与各种标准化组织和全球性大客户有过密切合作。Ciprian 曾与人合著过两本热门的 IPv6 图书（由 Cisco Press 出版）、4 份 RFC，以及多篇涉及 IPv6 技术、战略和启用的论文。他还是 IEEE 资深会员、若干咨询研究委员会的成员，以及业界 IPv6 活动的踊跃发言者。

献　词

我要把本书献给 Linda、Zack 和 Carter。在生命中，能有你等相伴，我是何其幸运；我为儿女们的长大成人而感到自豪。感谢你们在我著书的那几个月中对我的宽容。我要感谢母亲为我所做的祈祷，以及对我无条件的爱；我要感谢父亲，是他让我知道学海无涯。感谢我的岳父、岳母大人，是你们培养出了 Linda，并让她融入了我的生命，Linda——你是天下无双。Bob（父亲），谢谢你，我的良师益友，你的言传身教使我明白了努力工作的真谛。

—Shannon McFarland

首先，我要把本书献给我的祖父（Gyani Gurcharan Singh），感谢你赐予我作为作家、诗人以及古典音乐家的灵感。我要感谢我的全家：爸爸（Surinder Singh Sambi）、妈妈（Sukhdev Kaur）、哥哥（Ravinder Singh Sambi 博士）、妻妹（Amrit Kaur）和老婆（Avnit Kaur），感谢你们在我写作期间对我无条件的支持。我还要把本书献给我的女儿（Japjot）、双胞胎儿子（Kabir Singh 和 Charan Kanwal Singh）和侄子（Kanwal 和 Bhanwra）。

—Muninder Singh Sambi

首先，我要感谢我的父母：爸爸（Satbir Singh）和妈妈（Indrawati），以及我的妻子（Suman），感谢你们在我写作期间对我的支持。我要把本书献给我的孩子：Pulkit 和 Apoorva。

—Sanjay Hooda

我要感谢我的妻子 Parul，感谢她在写作过程中给我的巨大支持。我要将本书献给我的女儿 Anshi，是你让我知道，真正的快乐往往是来源于生活中的琐事。

—Nikhil Sharma

致　谢

我要向传授我 IPv6 知识和经验、支持我“精修”IPv6（尤其在我对 IPv6 初窥门径之时），以及过去这么多年来为我提供帮助的一干人等表示感谢，我要感谢我的朋友们和我最忠实的支持者们，他们是：Freddie Tsao、Steve Pollock、Chris O'Brien 和 Mark Montanez。这些年来，能在这么多优秀的管理者麾下效力，我真是三生有幸，领导们对我总是包容有加，在工作方面也对我大开方便之门（尤其是 IPv6）。他们是（受篇幅所限，只能列出其中少数人的大名）：Todd Truitt、Vince Spina、Kumar Reddy、Mauricio“Mo”Arregoces、Dave Twinam 以及 Mark Webb。此外，我还要感谢为我提供直接或间接帮助的 Cisco（过去和现在）一干同仁，他们是 Patrick Grossetete、Chip Popoviciu、Eric Vyncke、Gunter Van de Velde、Tarey Treasure、Darlene Maillet、Angel Shimelish、Chris Jarvis、Gabe Dixon、Tim Szigeti、Mike Herbert、Neil Anderson、Dave West、Darrin Miller、Stephen Orr、Ralph Droms、Salman Asadullah、Yenu Gobena、Tony Hain、Benoit Lourdelet、Eric Levy-Abegnoli、Jim Bailey、Fred Baker 以及其他人等。最后，我要感谢 John Spence 和 Yurie Rich，感谢二位长期以来对 IPv6 部署的不懈努力，以及所分享的反馈经验。

—Shannon McFarland

首先，我要感谢本书的其他三位合著者：Sanjay Hooda、Nikhil Sharma 和 Shannon McFarland，感谢诸位在写作期间所奉献出的合作精神。特别要感谢 Shannon，不但要感谢你对我们写作积极性的调动，还要感谢你对 IPv6 难点的指导。

我要感谢我的良师益友 Sanjay Thyamagundalu，是你把我引入了网络这扇门，感谢你在本书写作过程中对我的力挺。

我还要感谢我的主管领导 Sachin Gupta，感谢你在我写作本书期间所提供的支持和动力。我要感谢本书的技术审稿人 Jim Bailey 和 Chip Popoviciu，感谢你们无私的分享 IPv6 方面的知识，感谢你们在审稿时对问题“打破砂锅，一追到底”的精神。

最后，我要感谢 Cisco Press 团队，特别要感谢 Brett Bartow 和 Dayna Isley，感谢二位从本书的初稿到审稿过程中保持的耐心和提供的指导意见。

—Muninder Singh Sambi

首先，我要感谢本书的三位合著者：Muninder、Shannon 和 Nikhil，感谢诸位在写作期间对我的关照。我还要感谢我的挚友 Sanjay Thyamagundalu 和上司 Vinay Parameswarannair，感谢二位在写作期间对我的支持。Sanjay Thyamagundalu 不但是本人创作灵感的源泉，而且还针对 IPv6 的各个领域，提供了自己独到的见解。

同样要感谢 Brett Bartow、Dayna Isley 和 Cisco press 的全体同仁，感谢你们容忍我一再推迟交稿。

—Sanjay Hooda

首先，我要感谢 Muninder Sambi，是你将我引入了五彩缤纷的网络世界，你对我可谓是亦师亦友。感谢 Sanjay Hooda，要是没有你搭建的试验环境，本书不可能完成。感谢 Shannon，感谢你对整个团队写作积极性的调动，每当我们觉得完成写作遥不可及时，你总会提醒我们终点就在前方。

我同样要感谢我的一干好友，他们总是有问必答：Amol Ramakant、Deepinder Babbar、Jagdeep Sagoo、Nitin Chopra，以及我电话上的 7×24 小时快速拨号按钮 1-800-Call-Manu。

—Nikhil Sharma

我们要对本书的技术审稿人 Chip Popoviciu 和 Jim Bailey 致以崇高的敬意，感谢你们在审校本书时所倾注的专业技能和心血。

最后，我们要感谢富有创造力的本书编辑 Brett Bartow 和 Dayna Isley，以及 Cisco Press 出版团队，感谢你们对我们的关照、容忍，以及为保证本书的品质所耗费的精力。

本书使用的图标

命令语法约定

本书命令语法遵循的惯例与 IOS 命令手册使用的惯例相同。命令手册对这些惯例的描述如下。

- 粗体字表示照原样输入的命令和关键字，在实际的设置和输出（非常规命令语法）中，粗体字表示由用户手动输入的命令（如 show 命令）。
- 斜体字表示用户应提供具体值的参数。
- 竖线（|）用于分隔可选的、互斥的选项。
- 方括号（[ ]）表示任选项。
- 花括号（{ }）表示必选项。
- 方括号中的花括号（[{}]）表示必须在任选项中选择一个。

前　言

Internet 协议版本 6（IPv6）是 IP 协议的下一个版本，可供 Internet 上所有类型的设备进行通信。IPv6 浮出水面已有不少年头，但在企业网中，也是最近几年才加快了对其部署的步伐。IPv6 尚处于不停地开发和完善之中，因此，在实际的部署中，无论是 IPv6 协议本身，还是该协议的部署方法，总会暴露出一些“纰漏”。

世界各地的企业能够与 IPv6 建立起联系，原因不外有二：其一，需要部署能够自动支持 IPv6 的操作系统或应用软件（有时，还是在不知情的情况下）；其二，为了满足额外的地址需求、拓展新兴的市场、应对兼并和收购所带来的难题，以及为了追求最新最尖端的技术和应用，而必须充分利用 IPv6 协议的新特性。无论出于哪种原因，对企业来说，不但要对 IPv6 所支持的各部署选项了然于胸，而且在规划和设计自己的 IPv6 部署方案时，既需勇于进取，亦需深思熟虑。

IP 无处不在，并早已与网络水乳交融。因此，在企业网中，为了正确地规划并部署 IPv6，IT 人员须首先针对自有网络中的各个区块，比如，企业园区网区块、数据中心区块以及 WAN 区块等，分别设计 IPv6 部署方案；然后，再对已部署了 IPv4 的网络的所有区块做统一规划。接下来，IT 人员还要根据商业和技术方面的驱动力，设法让 IPv6 和 IPv4 并肩运行。有时，会新辟区块部署 IPv6，在这样的区块中，无需运行 IPv4；但更为常见的是，在网络中，未必需要 IPv6 处处现身，但却要求 IPv4 无处不在。本书将企业网划分为各个不同的网络区块，并会向读者传授如何在相应的区块中设计和部署 IPv6。

目标和方法

企业有新项目上马时，在行政和业务论证方面，通常会遇到许多问题，但往往终止于技术方面的设计与实施，而且在设计和实施过程中还总离不开“边上马，边论证”的思维模式。本书的目标有两点：第一，为读者提供一种既实用而又已获得证明的 IPv6 部署方法，

该方法可按网络中的各区块，将大量的 IPv6 部署任务分解为一个个可量化的子任务；第二，展示了多个有效的配置案例，可供读者搭建 IPv6 实验室、IPv6 生产网络，以及 IPv6 生产网络的测试环境。

本书在内容组织方面具有一致性，介绍网络各区块 IPv6 的部署时，总是首先对部署情况做简要介绍，然后会示出网络拓扑结构示意图（只要适用），最后会给出各种网络设备的配置示例，以巩固读者对 IPv6 部署概念的理解。本书不但能够帮助读者掌握企业网中部署 IPv6 的各个选项，还能让读者了解实施上述部署选项的方法。

读者对象

本书适用于在企业 IT 部门供职的网络工程师，以及维护企业网络的系统集成和咨询工程师。读者应该了解 IPv6 的基本概念，其中包括：编址，IPv6 邻居之间、IPv6 主机和路由器之间的通信机制，IPv6 路由选择等。虽然本书若干章节包含了对某些 IPv6 原理和主题的介绍，但对于 IPv6 新手而言，由于这样的介绍既不够深入，也不涉及 IPv6 的基本原理，因此不足以作为 IPv6 基础知识的入门参考书。本书假定读者对网络技术、设计和部署均有深入的理解。本书并非 IPv6 和网络设计的入门书籍，而是介绍 Cisco 长期以来一直推崇的与二、三层网络设计相关的最佳做法。

本书的组织结构

本书设计灵活，读者既可从头到尾通读，也可根据工作需要，在章节之间自由翻阅。

本书 1～4 章是对 IPv6 部署的一般性介绍，包括以下主题。

- **第 1 章，“启用 IPv6 的市场驱动力”**：本章通过技术和商业两个方面，来讨论在企业网中部署 IPv6 的驱动力。本章亦提供了 IPv6 部署的增长趋势和常见的使用案例。
- **第 2 章，“层次化网络设计”**：本章简要介绍了业界公认的、业已成熟的层次化网络

设计模型，旨在让读者建立起网络设计原则的基本概念，本书的所有内容都是构建在这一基本概念之上。

- **第 3 章，“常见的 IPv6 和 IPv4 共存机制”**：本章将讨论企业网中最为常用的几种 IPv6 和 IPv4 共存机制（亦称为过渡机制）。本章所讨论的机制包括双栈、ISATAP、6to4 等。
- **第 4 章，“网络服务”**：本章将探究大多数 IPv4 部署中常用的网络服务，包括 IPv6 多播、服务质量（QoS）和路由协议等。本书的其他章节则会展示具体示例，以介绍如何部署上述服务。

本书 5～12 章侧重于企业网中 IPv6 的实际部署，专注于部署中的技术细节。

- **第 5 章，“IPv6 部署规划”**：本章将从一个更高的层次来展示 IPv6 部署前和部署阶段的注意事项。本章旨在以系统的方法来呈现 IPv6 的部署规划。
- **第 6 章，“园区网络中的 IPv6 部署”**：本章讨论在园区网络环境中通用的 IPv6 部署选项。本章既会深入探讨各种 IPv6 和 IPv4 共存机制，也会给出在园区网络中达成高可用性 IPv6 部署的配置案例。此外，还会讨论诸如 Cisco 虚拟交换系统之类的高级技术。
- **第 7 章，“部署虚拟化的 IPv6 网络”**：本章不但会讨论各种网络、设备、桌面和服务器虚拟化解决方案，而且还会提供其中一些上述解决方案的配置范例，其中包括 6PE 和 6VPE。
- **第 8 章，“WAN/分支网络的 IPv6 部署”**：本章将会向读者介绍网络中 WAN/分支网络区块的各种设计场景，还会针对不同类型的 WAN/分支网络及网络服务（包括有动态多点 VPN 和 Cisco ASA 安全服务）给出详尽的配置范例。
- **第 9 章，“数据中心网络中的 IPv6 部署”**：本章涵盖了数据中心网络中常用的技术、服务以及产品，亦会向读者展示一个通用的设计方案，并会给出与之相对应的各种设备的配置，读者可根据自己网络环境对其加以应用。本章将重点围绕数据中心网络中常用的各种产品，比如，Cisco Nexus 7000、1000v 以及 MDS 9000，展开讨论，并附带讨论了 Cisco NAM、ASA 以及其他产品和技术。
- **第 10 章，IPv6 远程访问 VPN 的部署**：本章讨论在远程访问 VPN 环境中启用 IPv6 的各种选项。本章既会给出利用不支持 IPv6 的产品传递 IPv6 VPN 流量的配置示例，也会给出在 IPv6 网络环境中使用 Cisco ASA 和 AnyConnect SSL VPN 解决方案的配置方法。
- **第 11 章，“管理 IPv6 网络”**：本章涵盖了企业网 IPv6 部署中常用的管理部件。这些

部件包括网络管理应用程序、工具、设备，以及通过 IPv6 传输的网管信息。

- **第 12 章，“按部就班：搭建 IPv6 实验网络，启动生产网络的试点工作”**：本章讨论 IPv6 专用实验室网络环境的需求和用途，以及在正式组建 IPv6 生产网络之前，开展试点工作的重要性。本章将从实用性和系统性的角度，向读者传授搭建 IPv6 实验网络环境、在其中进行 IPv6 应用测试，以及从 IPv6 实验环境过渡到生产网络的 IPv6 试点阶段等方面的经验。

目　录

CHAPTER 1

第 1 章　启用 IPv6 的市场驱动力

本章涵盖以下内容。

- **Internet 的发展及人们对 IPv6 的需求**：本部分内容将重点介绍：致使 Internet“增寿”的现有解决方案；比之其他解决方案，IPv6 所具备的优势。此外，本部分内容还简要罗列了 IPv6 的市场驱动力，并会回答与 IPv6 有关的常见问题及人们所关注的事宜。
- **IPv6 之于 IETF**：随着 IPv6 的大行其道，至关重要的是，要有类似于 IETF 这样的标准化机构出面，为这一遍布所有网络，并为所有计算机设备所共用的协议的诸多功能制定标准。
- **企业网 IPv6 的部署现状**：虽然许多企业目前还对启用 IPv6、制定 IPv6 部署规划持观望态度，但某些纵向联合型企业（some of the enterprise verticals）（比如零售型、制造型企业以及 Web2.0 和 IT 企业等）不但已率先在网络中启用了 IPv6，而且还采购了支持 IPv6 的计算机设备，并利用 IPv6 协议传输自己的商业应用流量。

Internet 已经从美国国防部内部使用的分布式计算系统，逐渐发展成为一个巨大的平台，在这个平台之上，各个企业不但能够不断创新自己的业务，而且还能以非常高效的方式为全球客户提供商品和服务。Internet 协议簇（TCP/IP）正是用来支撑这一平台，并为其提供通信手段的底层技术。

尽管 Internet 并未受到集中化的管制，但也有综合性的组织为其关键要素（比如 IP 地址空间及域名系统（DNS）等）制定（维护）规则、实施运维。上面提到的 Internet 关键要素由互联网名称和号码分配公司（ICANN）统一维护和管理，而互联网编号分配机构（IANA）则负责具体的实施操作。ICANN/IANA 会为出现在 Internet 上的资源分配唯一的标识符，包括域名、Internet 协议（IP）地址，及应用端口号等。

更多信息请见：

- ICANN: http://www.icann.org；
- IANA: http://www.iana.org。

IETF（Internet 工程任务组）（www.ietf.org）是由下属会员以松散的形式组织起来的国际性非赢利性机构，Internet 核心协议正是根据其下属会员专家们的意见来实施标准化。所有网络设备和产品都会使用各种 Internet 核心协议，来实现网络互联；设备制造商则会提供用户界面，供人们配置和使用上述协议。

IETF 在评估 Internet 使用方面的增长时，最为看重的是编址方面的问题[①]。该组织对上述问题做出了以下评估。

- **地址空间耗尽**：IETF、IANA 的业界参与会员、区域互联网注册管理机构（RIR）以及某些私营企业一致认为，公网 IPv4 地址空间将在 2011 年耗尽。
- **IP 路由表的膨胀**：根据类别来分类和分配 IP 地址的做法，已致使 Internet 骨干路由器的 IP 路由表达到了令人震惊的规模。

本章的第一节将会围绕 IPv4 地址空间消耗问题，以及由 IETF 制定地临时性应对措施展开深入讨论。读者还会了解到促使 IETF 痛下决心，开发 IPv6 的原因所在。

1.1 IP 地址耗尽及临时性的应对措施

公网 IP 地址不够，多台主机只能配置内网（私有）地址，然后再转换成一个或几个公网可路由 IP 地址，并以此来访问 Internet。NAT（网络地址转换）可让企业网内部使用本地私有地址（RFC 1918 地址）的多台设备与外部世界通信时，共用一个或多个公网 IP 地址。虽然 NAT 在一定程度上延缓了 IPv4 地址空间的消耗，但也致使一般性的应用程序在双向通信时更为复杂。上述临时性的应对措施还带来了如下问题。

- 架设网关、防火墙，以及开发应用程序时，需要专门的代码来处理 NAT/PAT（比如，使用 UDP 的 NAT 透明传输）。
- 将标准端口映射为非标准端口（端口转发）。
- 确立并使用与 NAT 有关的临时性解决方案（STUN、TURN 和 ICE 等）。

① 原文是“The IETF evaluated the growth of the Internet protocol with emphasis on addressing”。

- NAT/PAT 地址嵌套。
- 基础设施、应用程序以及安全方面的复杂性。
- 架设和管理多个地址池的复杂性。
- 耗费更多的时间、精力和金钱去编码并管理与 NAT 有关的临时性解决方案。
- 因为 IP 地址经过了转换，在单位的内网中，很难识别建立连接的设备[①]。

注意

对于 sensor[②]而言，即便部署于线内（inline），恐怕也不能完全丢弃攻击数据包。部署于线内的 sensor 开始丢弃匹配复合模式特征的数据包之前，部分攻击数据包可能已经被 IDS/IPS 放行了。相对而言，根据数据包的原子特征，执行丢弃动作要更为有效，其原因是 sensor 会在单个数据包的基础上，执行匹配动作。[③]

注意

广播和电视分别用了 40 年和 15 年时间，才将听（观）众人数发展到了 5 千万，而 Internet 用户达到这一数字仅用了 5 年时间。

设计 IPv6 的目的是要取代 IPv4。IPv6 的地址空间之大，超乎人们的想象，此外，引入 IPv6 不但可使网络管理更为简单，还能实现端到端的透明访问，并可改善网络的安全性及移动性，这些内容都会在以下几节讨论。

1.2 IPv6 的市场驱动力

IPv6 可让企业“激活”新的应用，外加将自己的业务推广到全球，从而获得经济上的收益。有以下 4 项主要因素驱使企业纷纷采用 IPv6，如图 1-1 所示。

- IPv4 地址方面的短板。
- 政府 IT 战略规划。

① 原文是“Inability to easily identify all connected devices on an organization's network”。

② 作者的行文方式不明确，不知此 sensor 是不是 IDS/IPS 上的 sensor。

③ 译者抓耳挠腮也闹不明白，为什么作者要在此处安插这段文字，现给出整段原文“Sensors, even inline, might not be completely successful at dropping packets of an attack. An attack could be on its way, if only partially, before even an inline sensor starts dropping packets matching a composite pattern signature. The drop action is much more effective for atomic signatures because the sensor makes a single packet match”。

- 基础设施的发展。
- 操作系统的支持。

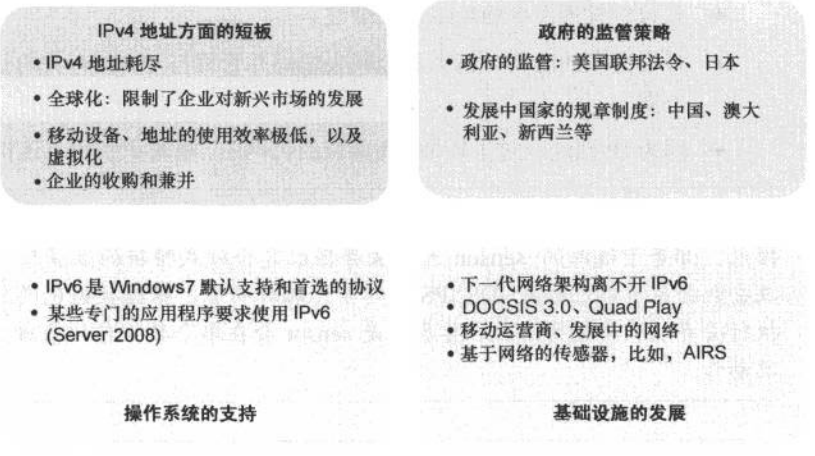


图 1-1 IPv6 的市场驱动力

本节的其他内容将会描述图 1-1 所示的 IPv6 主要市场驱动力。

1.2.1 IPv4 地址方面的短板

IPv6 的发展是拜 IPv4 编址体系所存在的以下几处短板所赐。

- **IP 地址耗尽问题**：Internet 应用程序和全球用户的数量都在飞速增长，加剧了 IP 地址的消耗。永远在线的设备（比如智能手机、Internet 工具、智能联网汽车、集成电话服务，以及多媒体中心等）[①] 在数量上也一直在保持增长态势。IPv4 只能提供 42 亿（4.294×10^9）个地址。对于当今的全球化和移动领域来说，IPv4 地址的耗尽只是时间问题。尽管 IPv4 地址耗尽的主要问题是最初 Internet 设计方面的“底子”问题（insufficient capacity of the original Internet infrastructure），可诸多商业驱动因素（包括全球化、移动设备的爆发性增长、虚拟化，以及商业方面的兼并和收购等）已“榨干”了 IPv4 技术的潜能，因此需要对如 IPv6 之类的技术进行评估，以充分利用其来支撑起 Internet。
- **全球化**：有了网络，所有企业的商务贸易才能更为活跃。企业在拓展新的市场，扩大业务的同时，其网络规模也会随之增长，势必需要更多的 IP 地址。

① 作者列举的后面 2 个东东似乎不能算作“设备”吧？

- **移动设备**：由于让手持式设备变为“计算机”的成本大幅下降，手机也逐渐跻身为 Internet 主机中的一员，从而导致 IP 地址需求量的增加。
- **地址使用效率低下**：对那些在 20 世纪八九十年代初便“攫取”了大量 IP 地址的单位而言，其对 IP 地址的实际需求要远低于其“攫取”量。比方说，某些大公司或大学都获得了一个可容纳 1600 万台主机的 A 类地址块。有很多地址要么从来都没有被使用过，要么上述占有大量地址的某些单位早已缩小了“编制”，反之，另一些单位则“吃着碗里的，拿着锅里的”（留着自己大量的地址块不用，通过兼并和收购又弄到了新的地址块）。
- **虚拟化**：如今，一个物理系统可当作多个虚拟系统来使用。每个虚拟系统可能都需要配置一个或多个 IP 地址。这方面的例子包括虚拟桌面基础设施（VDI）和托管式虚拟桌面的部署等。
- **兼并和收购（M&A）**：当公司被收购或与另一个公司合并时，网络中常会发生 RFC 1918 IPv4 私有地址冲突或重叠问题。举例来说，若甲公司网络所用的地址段为 10.x.x.x，而被其收购的乙公司网络也使用该地址段（如图 1-2 所示）。在收购期间，公司一般都会在网络中部署一个 NAT 重叠地址池，于是，甲、乙两公司网络便可使用该非重叠的地址空间（比如，使用 172.16.0.0/16）来彼此通信。这也是在对其中一个公司重新编址之前，让双方主机得以通信的临时性解决方法。

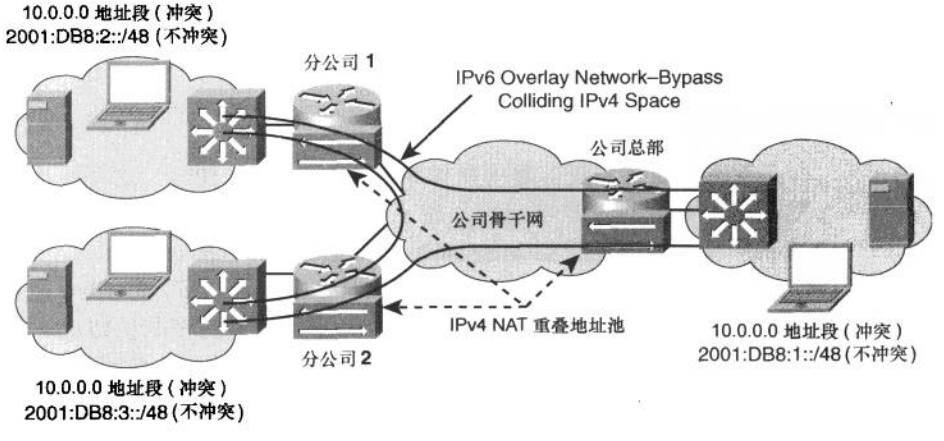


图 1-2 IPv6 覆盖模型——解决 M&A 地址冲突

对于上述情形，还可以使用 IPv6，在原有 IPv4 网络上“覆盖”一层 IPv6 网络，来减轻由 M&A 所造成的地址冲突影响。可在重要的系统和主机上启用 IPv6 功能，通过覆盖在 IPv4 网络上的 IPv6 网络，实现彼此通信。

1.2.2 政府 IT 战略规划

以大局（国家的 IT 战略规划和各国政府的要求）为重的理念，促使了许多企业和服务提供商去实施 IPv6，以更好地“迎合”政府（即私营企业与政府合作）。这样的例子有，2008 年 6 月 30 日，数家美国国防部的供应商联合行动，迅速规划和部署了 IPv6，以完成美国联邦政府下达的 IPv6 任务。上面提到的一部分企业不但已经与政府机构的网络对接（peer with federal agency networks），而且还会在政府需要 IPv6 之时提供基于 IPv6 的服务和产品。

1.2.3 基础设施的发展

随着用来支撑 Internet 的底层基础设施，以及某些新兴公共基础设施（比如能源管理、电力配送，以及新发展的其他公用事业）的不断完善，这些设施的规模也随之增长，并增长到了使用现有技术、产品和 IPv4 所能承受的临界点。诸多技术（比如智能电网、宽带电缆、移动运营商）的发展也使得越来越多的设备接入到 Internet。撇开所使用的技术和方法不谈，用来通信的手段只有一种，那就是 IP。IPv4 已渐渐力不从心，而 IPv6 则拥有无限的潜力。

1.2.4 操作系统的支持

几乎所有得到广泛部署的操作系统都默认支持 IPv6。这些操作系统都会默认启用 IPv6 地址，从而也加速了企业网内 IPv6 的启用进度。默认支持 IPv6 的主要操作系统包括 Microsoft Windows 7 和 Server 2008、Apple Mac OS X 以及 Linux 等。许多企业可能都会有这样一种认识，即 IPv6 是在自己不知情的情况下使用起来的，这是因为上述操作系统在默认情况下所启用的网络协议总是 IPv6 而非 IPv4。IT 人员也逐渐意识到，要想掌控 IPv6 的行为，除了必须以一种可控的方式去实施 IPv6 以外，还要对 IPv6 的各种特性谙熟于心。

1.2.5 部署 IPv6 的好处

通常，由企业以外的因素推动的市场计划和驱动力，可能会被迫受到业界因素的影响，而这些因素有时又归因于其他外部因素（比如 IPv4 地址耗尽），可大多数时候，企业还是会从中获益，当然，商务或技术方面的优势将会起决定性的作用。表 1-1 总结了部署 IPv6 所能给企业提供的若干好处。其中的某些内容在前文已有所论述，而许多其他内容将会在本书中随文展开讨论。

1.3 关于 IPv6 的常见问题

到目前为止，IPv6 的发展已超过了 10 个年头，可对世界上许多地方来说，该技术似乎还可有可无。如今，随着 IPv4 地址短缺问题日渐明晰，即便是非常保守的老学究们，对 IPv6 的兴趣和认识也在逐渐提高。

以下各节将回答或澄清多年积累的与 IPv6 有关的问题或谬论。

1.3.1 为了业务的发展，公司是否需要 IPv6

这是一个最常被问及的问题，尤其适用于尚未使用 IPv6 接入 Internet 的大多数公司。对一个公司来说，是否需要 IPv6，取决于以下三项主要因素。

- 要想保持企业业务方面的连续性，并将业务拓展到全球，就必须获取 IPv4 所无法提供的大块地址空间。
- IPv6 既会制造新的商机，也是一个革命性的平台。而 IPv4 根本不足以满足诸多网络应用的发展（比如，车载遥感系统等）的发展，因为 IPv4 绝对无法为数百万个安装在车上的网络传感器提供足够的地址。
- 诸如 Windows 7 和 Linux 之类的操作系统已经默认启用了 IPv6。

几乎可以肯定，拥有庞大人口基数且具备新兴技术能力的发展中国家，比如，印度和中国，将会直接迁移到 IPv6。对既想进军上述国家的市场，而又不想使用 IPv6 的企业，在竞争中将会处于劣势。

表 1-1 IPv6 的好处

IPv6 的技术优势	详细说明
充足的 IP 地址	与 IPv4 相比，这是 IPv6 所带来的最明显的好处。IPv6 地址由 128 位值，而非 IPv4 的 32 位值组成。因此，IPv6 能够提供无穷无尽的全球可路由地址
更为简单的地址部署	任何主机要获取网络资源，都必须为其分配 IP 地址。分配 IP 地址的传统方式是手工指定或 DHCP 自动获取。除了以上两种分配方式以外，IPv6 还内置有无状态地址自动配置特性（SLAAC），通过该特性可让主机自动配置地址，如此一来，在部署支持 IPv6 的端点时，可做到既便捷又简单。SLAAC 常用来配置无需供末端用户访问的设备。此类设备包括车载网络传感器、自动遥感设备，以及生产装备等 对于需让用户访问的主机，比如，桌面或服务器等，由于在路由器通告消息中不包含 DNS 信息，故而限制了 SLAAC 的部署。IETF 社团于是拿出了一份草案（RFC 5006），以扩展路由器通告消息（RA 消息），令其包含 DNS 信息。此外，还有一些标准社团中的积极分子为 RA 消息的扩展制定了标准，令其除包括 DNS 服务器信息以外，还包括 NTP、BOOTP，以及厂商专有的 DHCP 选项信息等。取决于主机操作系统的实现方式，配置了 IPv6 协议栈的网络适配器被激活时，网络适配器会根据自己的 MAC 地址和已知的 IPv6 前缀，自行分配一个 IPv6 地址。使用自动配置特性的新型主机，会依赖于一种叫做邻居发现（ND）的协议，从邻居路由器发出的有效信息中，获取自己的 IPv6 地址。该机制既不需要让网络管理员进行干涉，也不需要在网络中部署中央服务器以作地址分配之用——这要强于 IPv4——IPv4 的自动地址分配非得部署 DHCP 服务器不可
端到端网络连接的完整性	一旦部署了 IPv4 NAT，便出现了一个公网 IP 地址“一叶障目”（其“背后”隐藏了多个私有地址）的情形，这从根本上破坏了端到端连接的完整性。部署了 IPv6，IP 地址便不会再出现“僧多粥少”，也免去了网络地址转换之需
与 IPv4 相比，具有增强的安全能力	IPv6 内置有安全特性（天生支持 IPSec），可在两台或多台设备之间激活端到端的控制数据包（比如，用来建立路由协议邻接关系，或用作邻居发现的数据包等）的加密功能，但到目前为止，该功能还很少有人部署。对于数据平面 IPv6 数据流的加密，则要依赖于现有的 IPv4 机制，比如 IPSec
针对安全性、QoS 和加密功能，改进了扩展属性报头	IPv6 的扩展属性报头并不作为主数据包报头的一部分。这些扩展属性报头连同自己独一无二的数据包结构，会协助 IPv6 提供加密功能、移动功能、最优化的路由选择功能等。必要时，这些报头会被插在基本 IPv6 报头和有效载荷之间。基本 IPv6 报头会通过下一个报头字段，来标明扩展报头的存在。这既提高了路由器转发数据包的速度，也提高了其转发效率

续表

IPv6 的技术优势	详细说明
移动性的改善	当把主机从一个网段移动到另一个网段时，确保其原先所在网段的网关能够意识到这一情况，是开发移动 IP（MIP）功能的目标之一。起初，在使用移动 IP（基于 IPv4）的情况下，移动设备发送/接收的所有流量，都需途径其原先所在网段的网关（家庭网关[home gateway]），这也被称为三角式路由（triangular routing） 为了解决三角式路由效率低下的问题，在 IPv6 中，MIP 得到了进一步的扩展。一旦部署了 MIPv6，一台外部通信服务器会不停地汇报移动设备所在的网段，以及该移动中的设备转发流量所使用的网关。于是，大批量的数据流会直接在移动设备及其通信者之间流动，并不会途经其家庭网关。该过程也被称为直接路由。这不但降低了流量转发的成本，而且还极大地提升了网络的性能和可靠性
使用流标签来改进数据流的资源分配	IPv6 保留了来源于 IPv4 的所有区分服务和集成服务的 QoS 属性。此外，IPv6 报头还预留有 20 字节的流标签字段，可为终端应用程序所用，以针对特殊的服务或数据流类型，来提供数据流方面的资源分配。标准机构虽然已经制定了 IPv6 的流标签机制，但并没有多少企业倾向于去充分利用该机制

1.3.2 IPv6 将会完全取代 IPv4 吗

在整个 Internet 基础设施由 IPv6“大一统”之前的相当长一段时间内，IPv4 和 IPv6 将会并肩运转。企业和服务提供商不但在 IPv4 身上下了血本，而且也已对 IPv4 技术了如指掌。

由于使用 IPv6 的地方越来越多，企业也需要制定投资解决方案，让自己的传统 IPv4 网域能够无缝而又有效地与 IPv6 网域对接，并寄望以此来获得更好的投资回报。简而言之，希望采用 IPv6 的企业不仅无须放弃自己的 IPv4 基础设施，反而应该去充分利用过渡技术，令 IPv4 和 IPv6 在网络中共存。

1.3.3 与 IPv4 相比，IPv6 更加复杂并难于部署和管理吗

由 IPv6 所提供的超大 IP 地址空间，给网络架构师和管理员造成了 IPv6 比 IPv4 更加复杂的印象；这其实是一种偏见。IPv6 巨大的地址空间让网络架构师在设计网络时，不再因 IP 地址有限而捉襟见肘，从而降低了网络设计的难度[①]。

① 原文是“The vast address space equips architects to no longer reconfigure their limited address space, making network designs much easier”，直译过来是“……让网络架构师不再重配自己的地址空间……”，译者认为，这话不合逻辑，译文酌改。

IPv4和IPv6的所有附属协议（比如DNS等）在运作方式上几近相同。相较于IPv4，IPv6在自动配置特性和多播能力（支持嵌入式聚合点）的实施方面也更为简单。

IPv6还拥有某些新的附属协议，比如，多播侦听者发现和邻居发现协议，但在大多数情况下，这些新协议只是IPv4中类似机制的替代协议。除了以十六进制的形式来表示以外，IPv6地址分配起来（包括规划和部署）也更为简单，因为重点需要考虑的不再是主机号，而是基于地址块来分配的链路号或子网号。从许多方面来看，IPv6都只是版本号更高的IP协议。在执行地址分配规划时，IPv4和IPv6差别不大，都需要确立网络中的地址汇聚点。

对于一个成建制的IT部门来说（其编制包括网络工程师（架构师）、硬件维护工程师、存储架构工程师和管理员、软件工程师等），要想充分利用IPv6的种种特性，除了软硬件投资以外，还需针对这一新兴的技术，对部门员工展开培训。

1.3.4 引入了IPv6之后，还能以多宿主的方式连接到多家服务提供商吗

2007年以前，分配IPv6地址时，需执行严格的层次化地址分配策略；只允许企业从单一服务提供商获取网络地址，以避免全球路由表的重叠。

自2007年开始，情况有所改变，自那以后，企业可以采用类似于IPv4的独立于提供商（provider-independent，PI）的地址分配方案。

采用了PI地址分配方案之后，企业在接入提供商时，便能继续沿用类似于现有IPv4的设计，这是一种具备冗余性，而又非常可靠的解决方案。

但是，IPv6的发展还存在许多新的变数，对现有策略的调整也为业界所热议，上述因素都会对如何以多宿主的方式（使用IPv6）连接到多家服务提供商产生影响。如今，仍然存在许多与该主题相关的有待解决的问题，读者除了要对标准机构所出台的标准加以关注以外，还需与自己的服务提供商保持密切联系，以随时留意上述变化。

1.3.5 IPv6能提供更优的服务质量吗

只有IPv6报头的某些字段内置了QoS机制，预计将会使用这几个字段去区分隶属于不同类别的数据包，并把相关的数据包标识为数据流。设立这些IPv6报头字段的意图是，能够让路由器之类的网络设备识别出相关数据流和流量类型，并会针对二者执行快速查找。实战中，对上述报头字段的使用则完全可选，

这意味着除了对支持 IPv6 转发有最低程度的要求以外，大多数网络设备并不是非要支持 QoS 不可。

可是，IPv4 报头也有在功能上类似于 IPv6 的字段，在用法上也几近相同，因此要说 IPv6 在对 QoS 的支持方面胜过 IPv4，该论点是不能成立的。

1.3.6 在安全性方面，IPv6“自动”胜过 IPv4 吗

确切说来，IPv6 与 IPv4 谁都不比谁更安全，因为两者之间无法比较。IPSec 是融入 IPv6 的与安全相关的主要机制。任何基于 RFC 且与标准兼容的 IPv6 实现都必须支持 IPSec，但对该功能启用与否则没有硬性规定。这也给人们带来了一种误解：IPv6 默认就比 IPv4 安全。相反，要想在 IPv6 安全性方面有所建树，不但在实施上要精心规划，对系统和网络实施人员的素质也有极高的要求。

1.3.7 IPv6 不支持 NAT 会降低安全性吗

部署 NAT 就能提高安全性，这可算是一则笑话。部署 NAT 只是为了解决 IPv4 地址短缺问题，而 IPv6 地址充裕，因此 IPv6 无需 NAT。那些拿 NAT 当做安全特性的人们，会形成部署 IPv6 便会降低网络安全性的想法。可 NAT 在安全性方面并没有起到什么实际效果。通过隐匿来实现安全性（security through obscurity）的思路早已过时，其原因是：对于从 Internet 发起，入侵企业内网的常见攻击往往都不会针对第 3 层（针对可路由的 IP 地址发动攻击），而是针对第 4～7 层。IPv6 在设计时，便认为 NAT 是“多此一举”，RFC 4864 则刊载了使用 IPv6 时的本地网络（LNP）防护概念；因此，就安全性而言，IPv6 至少不逊于使用 NAT 的 IPv4。

1.4 IPv6 之于 IETF

1995 年以来，IETF 成立了若干工作组，积极制定了与 IPv6 有关的草案和 RFC。这些工作组涉及如下领域。

- 应用程序领域。
- Internet 领域。
- 运维和管理领域。
- 实时应用和基础设施领域。

- 路由领域。
- 安全领域。
- 传输领域。

与IPv6相关的大多数标准都与Internet、网络管理和运维以及路由领域有关。以上三个领域仍然在IPv6的部署、管理、过渡、安全性，以及围绕IPv6地址展开的标准制定方面发挥着非常积极的作用。IPv6的实现，以及与其体系结构配套且基于标准的组件，在确保不同厂商的互操作性方面起到了关键性的作用。

IETF草案和RFC数量众多，更新频繁。应付IPv6发生变化的最佳做法是，经常了解IETF和其他制定相关标准的组织的动向。欲知与IPv6有关的更多详情，请访问http://www.ietf.org。

除了IETF以外，IPv6论坛也在大力发展IPv6就绪性徽标计划（IPv6 Ready logo program[①]），该计划会针对与IPv6有关的网络基础设施（路由器、主机、操作系统以及协议栈），展开对标准的遵守程度以及互操作性测试工作（conformance and interoperability testing）。通过证明IPv6已经可用，且随时可用，来增强用户对IPv6的信心，是该计划的宗旨。IPv6就绪性徽标委员会负责定义标准的遵守程度和互操作性的测试规范，以协助不同的厂商证明自己的产品已为IPv6做好了准备。访问http://www.ipv6ready.org可获得IPv6就绪性徽标的更多细节，以及已获得徽标（经过认证）的产品列表。

1.5 企业网IPv6的部署现状

经过了标准化机构15年的“沉淀”，外加10余年的部署经验，如今，IPv6已被诸多服务提供商和大企业所采用。现在，IPv6已发展成为健壮而又成熟的协议，对新型应用程序来说，可谓是创新之举。

在许多纵向联合型企业中，正在紧锣密鼓地部署IPv6，如表1-2所示。

表1-2 IPv6在纵向联合型市场中的部署

纵向联合型市场	案例
高等教育和研究	楼宇传感器 多媒体业务 移动性

① 中文译法取自EMC china网站。

续表

纵向联合型市场	案例
制造业	嵌入式设备
	工业以太网
	支持 IP 的零部件
政府（联邦/公共部门）	国防部
	战术级作战人员信息网（WIN-T）
	未来作战系统（FCS）
	联合战术无线电系统（JTRS）
	全球信息网格带宽扩展系统（GIG-BE）
交通行业	信息通信
	交通管制
	热点
	运输服务
金融行业	兼并和收购——覆盖型网络
医疗行业	家庭护理
	无线资产跟踪
	成像
	移动性
客户	机顶盒
	网游
	家电
	语音/视频
	安防监控
公共事业	智能电网
	通过电线传递 IP 的业务

最初，只有在设备制造商和研究领域才能见到 IPv6 的身影，其间诞生了 IPv6 的首个实现。其后，IPv6 的部署则成雨后春笋之势逐渐在纵向联合型企业中蔓延开来，其中的一些具体使用案例包括传感器网络、机械手（robotic arms）、环境控制以及传感器等，而其他的使用案例，无论与纵向联合型企业有没有关系，在本质上都是相同的。

在 IPv6 的启用方面，绝大多数企业都要历经以下三个阶段，如图 1-3 所示。

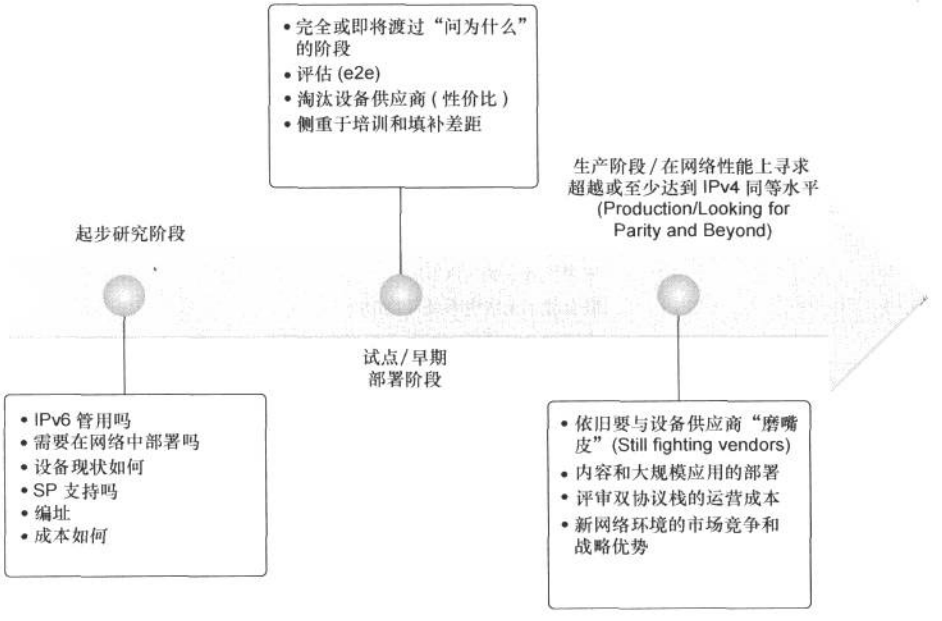


图 1-3 *启用 IPv6 的企业所要历经的三个阶段*

启用 IPv6 的企业所要历经的阶段一般被称为起步研究阶段。处在本阶段的企业都在研究 IPv6 是否管用、IPv6 所能带来的好处、IPv6 如何与自己的网络环境衔接、网络设备对 IPv6 支持还有什么空白，以及部署成本等。在本阶段，应通过在线资料向公司领导层灌输 IPv6 的实用性，即使用 IPv6 才能满足公司新业务发展的需求。对于公司的 IT 技术团队来说，在该阶段，则需要通过搭建实验环境、培训以及动手实验等环节，弄清 IPv6 的细节，以及 IPv6 与现有网络基础设施的瓜葛。许多处在该阶段的企业其实都并不清楚 IPv6 与自己有什么关系。

启用 IPv6 的企业所处的第二个阶段则名为试点/早期部署阶段。本阶段，在商业方面无论是否有上马 IPv6 的明确理由，都不应该再问“为什么要部署 IPv6”之类的幼稚问题，而企业也已做出了部署 IPv6 的决定。通常，在该阶段，考虑部署 IPv6 时，根本不需要有明确的商业理由，这与“先有客户再开店，还是先开店等客户”（getting our house ready for an unknown guest）的道理一样。那些曾维护过 VoIP 和 IP 电话的人们一定不会忘记，当上述技术的模式发生重大转变，且他们所维护的网络并没有及时开启高可用性（至少所开启的高可用性不足以满足语音的需求）

和 QoS 时，他们是多么的狼狈。在拥有明确的商业理由之前，为 IPv6 投入时间、精力和资金往往是企业在未雨绸缪，如果想让自己狼狈，上述一切投入当然是可以避免的。在本阶段，企业还应针对 IPv6 做出更为严肃的评估、开展培训，以及与不兼容 IPv6 的产品供应商进行正式的交涉。

启用 IPv6 的企业所要历经的最后一个阶段为生产阶段。在本阶段，企业正在寻求生产网络中高质量的 IPv6 部署。此时，即便不是全部，也至少是大半 IT 部件将要开始支撑 IPv6 网络，人们对 IPv6 的要求也逐渐等同于 IPv4，或至少要求 IPv6 部署之后，不会对业务造成影响。在业务上（传递业务流量时），可能仍会牵涉到不兼容 IPv6 的设备（产品）或厂商，但企业已淘汰掉了那些没有“进取心”的设备提供商，并将大半业务（流量）割接到了支持 IPv6 的平台上。企业业务照常开展，但侧重于使用支持 IPv6 的应用和服务，并以此来作为自己的竞争优势。

在启用 IPv6 的整个发展历程中，对企业而言，要持之以恒地向员工灌输 IPv6（无论是技术还是业务）的理念，在发展阶段中的每一步，可能都需要从相关厂商采购搭建 IPv6 平台的一切所需。

历史经验表明，企业在启用 IPv6 时，总会遇到部署方面的问题——鉴于只有少量产品支持 IPv6，而在自己的对等点部署 IPv6 的运营商也不多见，因此企业们纷纷打起了退堂鼓。服务提供商不部署 IPv6 的原因有二：其一，没有企业向其申请 IPv6 服务；其二，支持 IPv6 的产品风毛麟角。企业不生产支持 IPv6 的产品，是因为此类产品很少有企业或运营商去问津。于是，在某些情况下，便形成了一个恶性循环的怪圈，这一怪圈过去存在，现在仍然存在，因此必须要有敢于创新的企业出面，勇挑重担才能够打破这个怪圈。

从内容提供商的角度来看，Google 可算是 IPv6 部署的龙头老大，正是 Google 推出了“trusted adopter”计划[①]，链接为 http://www.google.com/ipv6。其他内容提供商和业界领先的网站也已针对自己的主机激活了 IPv6，用户可通过 IPv6 访问这些网站的主机。这些网站包括 Google (搜索和 Gmail)、YouTube、Netflix、Comcast 以及 Facebook。

与商贸杂志、blog、设备供应商以及某些怀疑者所持观点相反，许多企业都已经或正在开展 IPv6 的部署。很多公司都没有对自己所部署的 IPv6 大张旗鼓地宣传，从而导致人们误以为 IPv6 尚未得到部署。这些公司行事低调的原因主要有两点：经济因素和安全性考虑（既未掌握所有基于 IPv6 的攻击矢量，严密的

① 译者在通过 Google 也没搜到该计划，因此只能保留原文不翻。

安全措施也未落实到位)。本书其余章节将会讨论IPv6的部署事宜及相关要素。

1.6 总结

IPv6是下一代Internet协议，既可以解决IPv4地址短缺问题，也可以消除或减少一直沿用至今的NAT/PAT部署。取之不尽、用之不竭的IP地址是IPv6最主要的市场驱动力。这既保证了企业业务的连续性，也为Internet上的新型应用找到了新的出路。

IETF和其他标准化组织仍在评估新的解决方案，并同时制定着新的草案和RFC，以确保IPv6主机间的互操作性。

为了让新型应用程序迎接IPv6的大潮，大多数服务提供商、内容提供商以及诸多企业都计划或正在自己的网络基础设施中部署IPv6（有些甚至已经完成了IPv6的部署)。

本书的主要目标是，为服务提供商或企业提供IPv6设计框架，以帮助这些单位利用现有的过渡技术平稳过渡到IPv6；此外，本书还介绍了将IPv6集成进上述单位现有网络基础设施的方法。

1.7 其他参考资料

本章中的“注意”和“声明”[①]列出了完全掌握IPv6技术和协议方面的一切所需。实施IPv6时，在设计方面，有许多要素都需要仔细斟酌，其中包括安全、QoS、高可用性、管理、IT培训以及应用程序支持。

以下列出的参考资料提供了与IPv6、Cisco设计建议、产品与解决方案以及业界动向有关的详细信息，与上述主题相关的技术文档可谓是浩如烟海。

Aoun, C. and E. Davies. RFC 4966, “Reasons to Move the Network Address Translator - Protocol Translator (NAT-PT) to Historic Status.” http://www.rfc-editor.org/rfc/rfc4966.txt.

Cerf, Vinton G. “A Decade of Internet Evolution.” http://bit.ly/cNzjga.

Curran, J. RFC 5211, “An Internet Transition Plan.” http://www.rfc-editor.org/rfc/rfc5211.txt.

① 原文是“disclaimers”，据译者推测，这玩意儿可能是作者在正文中所给的链接。

IANA: http://www.iana.org.

ICANN: http://www.icann.org.

IETF: http://www.ietf.org.

IETF Behavior Engineering for Hindrance Avoidance (behave) drafts: https://datatracker.ietf.org/wg/behave.

IPv6 address report: http://www.potaroo.net/tools/ipv4.

Jeong, J., S. Park, L. Beloeil, and S. Madanapalli. RFC 5006, "IPv6 Router Advertisement Option for DNS Configuration." http://www.rfc-editor.org/rfc/rfc5006.txt.

Rekhter, Y., B. Moskowitz, D. Karrenberg, J. de Groot, and E. Lear. RFC 1918, "Address Allocation for Private Internets." http://www.rfc-editor.org/rfc/rfc1918.txt.

Van de Velde, Hain, Droms, Carpenter, and Klein. RFC 4864, "Local Network Protection for IPv6." http://www.rfc-editor.org/rfc/rfc4864.txt

CHAPTER 2

第 2 章　层次化网络设计

本章涵盖以下主题。

- **网络设计原则**：本节回顾了设计企业网络所必须遵循的三个重要原则：模块化、层次化以及高弹性。其后各节将会通过把企业网络划分为多个区块（模块）的方式，展开对以上三个概念的讨论。此外，还会对构成网络的各个模块做深入探讨。
- **企业核心网络区块设计**：本节回顾了企业核心层网络区块的设计需求和设计准则。
- **企业园区网络区块设计**：本节回顾了设计企业园区网络区块（行使接入-分布层功能）的各种选项。
- **企业网络服务设计**：本节简要介绍了企业从纯 IPv4 网络升级到 IPv4/IPv6 双协议栈网络时，所必须面对的网络设计需求。
- **企业数据中心网络区块设计**：本节回顾了数据中心网络所采用的设计方案，以及在数据中心网络中的每一层所要配置的特性。本节还会讨论如何制定存储区域网络设计方案。
- **企业边缘网络区块设计**：本节将讨论企业边缘网络的设计方案，其中包括：总部 WAN 设计、Internet 接入设计，以及分支网络设计等。

早期的计算机网络都是平面型拓扑，可根据需求，随时随地地在网络中添加设备。这样的平面型网络拓扑无论是设计、实施还是维护都很容易，但前提是网络的规模不能太大（主机数不能太多）。随着网络中的主机逐渐增多，而网络本身又缺乏故障隔离机制，因此一旦网络发生故障，不但影响面大，排除起来还非常麻烦。此外，在连接了大量主机时，这种平面型网络又会带来设计方面的难题。

鉴于平面型网络的种种缺陷，对企业来说，通过反复改进自己的网络设计，不但能够让网络跟得上企业发展的步伐，而且还能将其划分为一个个故障隔离域。因此，网络设计也就朝着模块化、层次化和高弹性的方向发展了。而这三个概念也最终变成了评判网络设计是否优秀的根本因素。

本章会简要介绍优秀网络设计的三要素——模块化、层次化以及高弹性——然后将通过把企业网络划分为多个区块（包括企业核心区块、企业园区、网络服务、数据中心区块以及企业边缘区块（Internet 子区、远程接入子区、WAN 子区以及分支网络））的方式，来展开对以上三个概念的讨论。此外，本章还会分别针对以上各个企业网区块，介绍相关网络设计注意事项，以使得网络设计者和架构师能够掌握不同的设计选项，从而能够在设计自己的企业网络时应用相关设计原则。本章同样提供了第 4 章中所讨论的各种网络服务的基础知识。

2.1 网络设计原则

商业网络应用程序历经多年的发展，已经从最简单的客户端/服务器模式发展到了多用户交互模式。如今，有些商业网络应用程序已充分利用上了诸如视频、语音以及无线方面的技术。随着各种技术的不断融合，对企业网的认知及需求都因交互式应用的出现而产生了巨大的转变。现在，不断变化的商业前景也迫切要求企业网络能够具备如下功能。

- **协作式应用程序的用户体验**：随着使用协作式、实时通信类、一次登录资源尽享型（single-sign-on）以及移动性应用程序的用户日渐增多，效果良好而又引人入胜的用户体验便成为了重中之重。
- **支持五花八门的末端用户设备类型**：除了个人电脑（PC）和 IP 电话之外，企业也已注意到，使用无线设备（包括支持 WiFi 的笔记本电脑和智能手机）和瘦客户终端的用户也越来越多。
- **网络的高弹性和更短的收敛时间**：商业运营要想适应全球化的发展，其网络需要每年 365 天，每天 24 小时持续运转；为了实现这一目标，则少不了高弹性的网络基础设施，来确保网络故障或升级期间不会对商业应用产生任何影响。
- **无处不在的安全性**：一直以来，无论是在数量上还是危害程度上，安全威胁都保持着增长态势，这也要求企业相应地提高其网络的安全性。此外，在安全性得到保障的同时，企业网络环境还应具备支持分布式和动

态应用程序的能力。随着企业业务伙伴的增多，就企业网络而言，也需满足随之渐长的业务伙伴或客户灵活访问的需求。

企业网是指互连末端用户和设备的基础设施[①]。就范围而言，企业网络既可以覆盖一座建筑物内的一层或多层，也可以“绵延”多个地理位置，横跨多座建筑物。此类网络不但要以高速链路来保障基本连通性，还要能具备运行关键商业应用所需的高弹性、安全性以及易于管理的网络服务架构。设计企业网时，请对以下三个原则铭记于心。

- 模块化。
- 层次化。
- 高弹性。

以下内容会详细介绍以上三个原则。

2.1.1 模块化

模块化是搭建网络的基本原则之一，可将按模块化设计的企业网视为由多个网络区块组成的集合，而企业网中的各个区块则分别依据系统化的方式来设计，在适当的情况下，还会引入层次化和冗余的设计理念。

对于网络设计来说，模块化的比重占得越多，将网络划分为各种功能齐备的网络区块（模块）也就愈发得简单，此后，便可让每个网络区块去支撑一个或一组特定的功能。如此一来，单个区块内的网络故障、网络升级以及对网络所做出的任何调整都不会影响到其他区块。只要遵循模块化的网络设计原则，就能够以每一网络区块为基础，来选择性地部署网络服务，但这一切仍需恪守网络的一般性设计原则。如图 2-1 所示，一个企业网可包括以下几个网络区块。

- **企业核心网络区块**：企业核心网络区块是企业网络的骨干，用来连接企业园区、数据中心、企业边缘以及网络服务区块。没有核心网络区块，上述各区块都将“分崩离析”，因此该区块需要提供不间断的 24 × 7 × 365 服务。应按一定的冗余和容错级别来设计核心网络层（区块），以确保在网络组件发生故障的情况下，能够即时恢复数据的流动。一旦在设计中引入了容错机制，网络便应能实现快速收敛和负载均衡，从而能够将核心网络区块中所有网元的作用发挥到极致。

① 原文就是这么蹩脚。

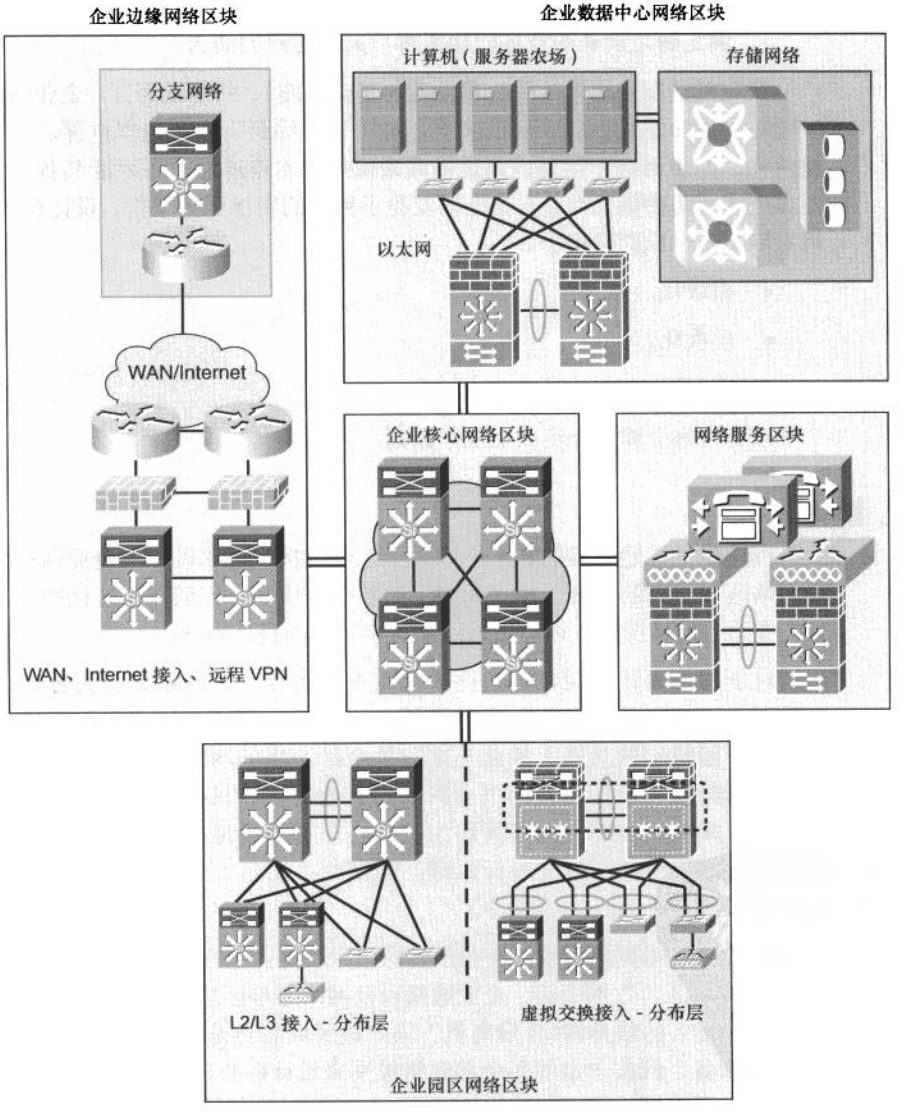


图 2-1 模块化网络设计示例

- **企业园区网络区块**：该网络区块包含所有独立运行的网元，用来提供终端用户、设备以及企业核心网络区块之间的连通性。可根据功能和服务

的需求，将该区块进一步划分为若干层次。在同一地理位置，取决于用户的数量，一个企业可拥有一个以上的园区网络区块。

- **企业数据中心网络区块**：该网络区块包含了三种重要网元：计算机（服务器）、网络设备（以太网交换机）以及存储（光纤通道）设备。在该区块内，还会部署用来构建“服务器农场”的大型计算机设备，以提供商业应用。这些服务器农场通常会既连接以太网交换基础设施，也连接光纤通道存储设备，当然，如以太网上的光纤通道（FCoE）之类的统一存储连接技术也在迅猛的普及之中。数据中心网络区块的作用是，能够让用户访问到由服务器农场来提供服务的商业应用程序。隶属于该区块的存储网络包括互连光纤通道主机的光纤通道交换机、连接服务器农场的 iSCSI 设备，以及支持 Cisco 统一 fabric 解决方案的产品（使用同一基础设施来连接光纤通道和以太网）。
- **企业边缘网络区块**：该区块包括三个子模块：Internet、远程访问 VPN 以及 WAN 模块，其作用是将企业网连接到服务提供商网络。该区块包含的所有网元的用途是：高效而又安全地保障企业园区网络模块与远程站点、商业伙伴、移动用户以及 Internet 之间的通信。企业边缘网络区块把来自各远程站点的连接“汇聚”在一起，对流量执行过滤，然后再将流量路由进企业园区网络区块。
- **网络服务区块**：对企业网来说，该区块还算是新生事物。该区块包括有统一通信、移动性以及用户认证所需的 IPv4 和 IPv6 服务。大多数企业往往都不会专门搭建网络服务区块，而是将其集成进数据中心区块。但网络服务区块能够为正部署 IPv6，但在部署当日，还不能完全部署好 IPv4/IPv6 双栈网络的企业提供很多好处。在这种情况下，企业需拿出一份临时性过渡解决方案：在不支持 IPv6 的网络上，“覆盖”上一层网络（架设隧道），来互联 IPv6 主机。要是把上述服务归并入一个单一的区块，在网络向双协议栈迁移的过渡时期，网络架构师们便无须架设遍及整个网络的多条隧道了（所有的 IPv6-in-IPv4 隧道，都由网络服务区块统一汇接）。

利用模块化网络设计，把网络划分为一个个小的区块（模块），除了易于管理以外，还具有以下优点。

- **便于管理**：采用了模块化的设计原则，每个网络区块不但可以单独管理，而且还能在每个区块内利用专用工具来管理所有的网络实体。让每个网

络区块“自立门户”——行使自己的一套功能，网络管理功能自然也不例外。

- **故障的隔离与排除**：各功能区块都提供与功能性和组织性支持架构（functional or organizational support structures）[①]相一致的边界。
- **改善的灵活性**：当网络设备需要升级时，模块化设计有利于对网络实施调整。可以精确地控制所要调整的设备范围，比方说，因需要增强功能而对少数或某个用户网段的设备升级。
- **降低运营开支（OpEx）**：模块化网络设计能够将网络分解为一个个较小的区块，每个小区块既简单而又易于维护。网络越简单，设计方案实施起来也就越快，对网络运维培训的时间要求也会短。由于按模块化设计的网络层次清晰，功能独立，因此，完工后，按照网络设计对网络工程实施验收也会非常容易。
- **简化对产品的选择**：采购网络设备时，模块化网络设计可将欲采购的网络设备与相应的网络层次绑定在一起，这就可以避免在不必要的功能上浪费采购资金。

由于每个网络区块都是独立的，故而一个模块内的故障、升级以及变更都会被限制在自己的边界以内。

2.1.2 层次化

层次化也是优秀的网络设计所必须遵循的重要原则之一。在为上一节所介绍的每一个网络区块制定设计方案时，都必须对高弹性和层次化的原则加以关注。随着企业业务的不断发展，企业网内的内部通信流量也随之增多，网络设计不但需要适应不断上线运行的新设备和新应用，而且还要能够满足渐增的容量需求，此外，还应尽量避免实施重大的系统升级。就灵活性而言，网络设计已经从传统的平面型网络，发展成为层次化网络拓扑，网络的各层之间既保持独立，又行使专用的功能，这使得网络架构师能够根据网络的相关层次，选择相应的硬件平台，安排所需的功能。划分出来的这些层次既具有各自的功能特征，也同时设有边界，以隔离网络故障。网络的每一层都具有独特的功能和独立的模块，用来提供网络服务。

企业园区网络区块一般分为三层（如图 2-2 所示），作用分别如下。

---

① 作者行文时根本不顾及读者的感受，非常喜欢用抽象形容词和抽象名词。

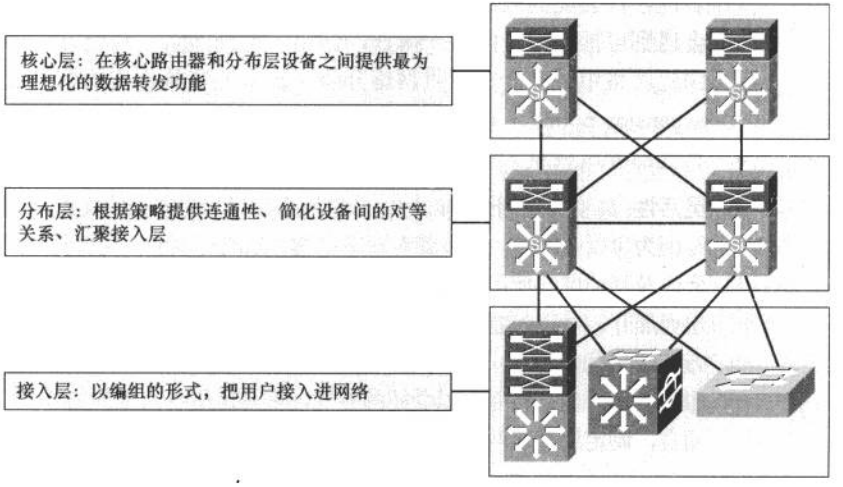


图 2-2 层次化网络设计

- **核心层**：在站点间提供最为理想化的数据转发功能、互连各分布层资源、提供访问数据中心网络区块的连通性，以及通过边缘网络区块提供对 Internet 的访问。
- **分布层**：让接入层设备（及用户）与提供网络服务的设备（比如防火墙、负载均衡器、IPS 等）之间建立起连通性，并实施与安全、流量负载，以及路由选择有关的策略[①]。
- **接入层**：提供终端用户和设备间的连通性[②]。

与平面型网络拓扑相比，层次化网络设计虽然代价不菲，但却优点多多。这些优点既可以增强网络的功能，又为网络设计引入了适当的模型，从而能够更好地满足商业需求。层次化网络设计的主要优点如下所示。

- **可扩展性**：引入了层次化的网络设计之后，如果网络规模发生了增长，网络设计师在网络中添加各网络区块时，只需照搬原有的设计。对于前面提到的各网络区块来说，要是在网络设计中不采用三级模型，其可扩

① 原文是“connects network services to the access layer and implements policies regarding security, traffic loading, and routing”，译文酌情改动。

② 原文是“provides connectivity to the end users and devices”，译者直译。

③ 原文是“Access layer：provides common group access to the network.”作者写作的目的只有一个，那就是只让自己能看懂！译文不是翻出来的，是猜出来的！

展性往往会受到限制。当然，在遵循层次化原则设计的网络中，可能仍会遇到可扩展性受限这一问题，但由于整个网络在功能方面彼此隔离，因此网络中势必会有可供网络升级（扩容）实施操作的施工点（natural expansion points），在这些地方实施与扩展性有关的操作，并不会对网络其他部分造成重大影响。

- **灵活性**：高度的灵活性是结构化分层网络设计与生俱来的优点，称其为优点，是因为可以分阶段、分步骤对网络中的各层进行调整，在调整的同时，并不会波及其他层的使用。比方说，对核心层数据传输方面的调整会独立于分布层而展开。对分布层性能和设计方面的调整，也可以通过分阶段或渐进式的方法来执行。此外，作为层次化设计整体的一部分，在设计架构中引入网络服务区块，也正是专为以受控的方式，来满足服务实施的需求[①]。就灵活性而言，制定层次化网络设计方案时，需要对以下关键领域仔细斟酌。
  - **控制平面的灵活性**：从第二层协议（STP）向第三层协议（路由协议）迁移。
  - **转发平面的灵活性**：将 IPv4 和 IPv6 一起考虑，为日后引入并使用 IPv6 埋下伏笔。
  - **编组用户的灵活性**：为了满足与公司间的合作、业务外包以及收购等商业需求有关的网络管理工作，在网络内，需考虑部署特定的网络访问方法，并做好激活相关服务的准备[②]。
  - **流量管理和控制的灵活性**：随着协作式应用程序的不断发展，要求园区网络区块的设计方案能够针对此类应用程序的流量，专门制定出一套简化的流量监控和故障排除机制。
  - **灵活的安全架构**：网络的安全架构应该能够适应受部署新应用程序的影响，而不断改变的流量模式。

---

① 原文是“Additionally, as a part of the overall hierarchical design, the introduction of the services block module into the architecture is specifically intended to address the need to implement services in a controlled fashion”。首先，原文按字面意思直译。其次，再来说一下作者，作者差就差在根本不替读者着想，总认为他说的话读者就一定会懂。果真如此的话，这书买来何用？最后，再替作者解释一下这句话。作者在抛出“services block module9(服务区块)”这一名词时，就没有解释清楚服务区块的作用。搭建该服务区块绝不是为了满足“the need to implement services（服务实施的需求）”，本书第 6 章对此有相关描述。

② 原文是“Enable network access and associated services within the network to support administration involving acquisition, partnering, or outsourcing of business functions”，译文按字面意思翻译不妥，所以译者认为，作者真正想表达的可能是：在网络中，将获准访问某些服务的特定用户编组，然后以用户组的方式来应用网络访问控制策略。

- **便于实施**：层次化的网络模型将网络划分为若干逻辑或物理区块，因此，对网络架构师来说，可在不影响现有基础设施正常运转的情况下，在逻辑或物理区块实施相关操作。举例来说，可在园区网络区块中部署新的接入层交换机，并将其与分布层交换机相连，而不会对现有接入层交换机，以及连接到这些交换机的用户有任何影响。
- **便于排除故障**：在按层次化原则设计的网络中，故障都被隔离在单独的区块中，故而可以很容易地定位并排除网络故障。比方说，某网络区块的分布层交换机发生了路由环路，这一故障只会影响到与其直连的接入层交换机，而不会对其他网络区块的分布层交换机有任何影响。
- **便于管理和容量规划**：就层次化网络模型而言，可把核心网络区块流量的递增视为需要对网络扩容的信号，因此在按层次化原则设计出的网络中，进行容量规划一般都非常简单[①]。按层次化原则设计出的网络往往都非常容易管理，原因是这样的网络具备以下优点——数据流流动方向的可预见性、可扩展性、独立实施性，且易于故障排除——所有这一切都能够使网络管理更为简单。

在本章的第 3 节和第 5 节会深入讨论层次化网络模型的三个层次。

2.1.3 高弹性

除了要制定出模块化和层次化的网络设计方案以外，对网络架构师来说，在网络设计方案的每一步中考虑高弹性也同样重要。让网络具备高弹性，以避免单点故障是确保高可用性和业务连续性的关键所在。在之前讨论过的各个不同网络区块及层级中，需要将交换机、链路以及网络设计中的弹性功能协调使用，相互贯穿。比方说，要是在接入层交换机配备了互为冗余的交换机 supervisor，即便主 supervisor 故障，业务连续性也可以得到保障。这还有助于确保不会对分布层（针对第二和第三层的部署）的网络（路由）收敛产生任何影响。

在网络设计方案中考虑高弹性，可能需要用到某些新的特性，但这往往只牵涉到如何实施层次化的网络，以及如何配置基本的 L2/L3 网络拓扑。

以下各节将会以模块化、层次化以及高弹性这三项原则为纲领，逐一介绍

① 原文是“Capacity planning is generally easier in the hierarchical model because the need for capacity usually increases as data moves toward the core”.从字面意思看，这段文字有太多的不确定性，作者这种写法，译者可以翻出 5 种以上不同的意思。译文根据译者自己的经验“杜撰”而成。

企业网各网络区块（企业核心网络区块、企业园区网络区块、企业网络服务区块、企业数据中心网络区块以及企业边缘网络区块（总部 WAN、Internet 接入以及分支机构网络））的网络设计方案。

2.2 企业核心网络区块设计

网络的核心层最简单，也最关键。该层构成了企业网络的骨干。网络的核心层既应具备高度的稳定性，还应具备高速交换流量的能力。该层虽然提供的服务（作者是指防火墙、负载均衡、入侵检测等服务）有限，但仍需冗余的设备和配置，以实现高可用性，并以此来保证执行软件升级或硬件更换时，不会造成流量的中断。企业网络核心区块设备会以第三层的方式“路由”进出企业网络的所有流量[①]。路由选择对数据中心网络区块的核心层（设备）来说十分关键，在配置时，需要启用路由协议内置的最高级别的安全特性，以避免与非法邻居路由设备建立路由邻接关系，并以此来防止非法路由的注入，避免路由环路的生成。为了避免上述问题，在网络的核心层需启用以下安全特性。

- 路由邻接关系认证。
- 路由过滤。
- 针对动态路由协议，记录邻接关系的变化情况。
- 防 IP 地址欺骗：启用单播 RPF(uRPF)以及速率限制（rate limiting）特性。

表 2-1 所示为设计网络核心层时的有关注意事项。

表 2-1 设计网络核心层的相关注意事项

推荐的做法	不推荐的做法
设计时要考虑到高可用性。应考虑采用 10GE 和 GE 以太网技术外加 port-channel 配置，来提高网络的带宽并实现链路冗余	采用（网络设备）基于软件的特性，此类特性会降低网络设备处理流量的速度
设计核心层设备的部署时应考虑到低延迟	在网络的核心层支持工作组的访问
采用收敛时间快的路由协议	在添加新的核心层节点时，为改善性能，才考虑对软硬件进行升级

① 原文是“The core provides a Layer 3 routing module for all traffic in and out of the enterprise network”，直译为“核心为进出企业网的流量提供第三层路由模块”。译者酌情更改译文。

2.3 企业园区网络区块设计

园区网络区块围绕着对两个基本子模块或子区块的使用来架构，这两个区块或模块都要连接到核心网络区块，它们是[①]：

- 分布层；
- 接入层。

以下内容会对以上两者做详细讨论。

2.3.1 分布层

分布层（交换机）上接网络的核心层（交换机），下连接入层（交换机）。在大型企业园区网络区块中，分布层交换机可能会不止一台，这要取决于与其相连的接入层交换机的数量。Cisco 推荐的最佳做法是，不要让超过 20 台交换机连接到单台分布层交换机。Cisco 之所以这样建议，是因为对于这种情况，无论采用哪种接入层设计方式（无论是第二层接入设计，还是路由式接入设计），在处理控制平面协议的能力方面，单台分布层交换机是万万满足不了这一设计需求的[②]。

当前，有下面三种基本的设计选项可用来设计接入层-分布层网络。

- 第二层（交换式）接入设计。
- 路由式接入设计。
- 虚拟交换机式设计。

第二层接入设计

传统的园区网接入-分布层设计都是采用第二层接入的方法——将所有接入层交换机全都配置为 L2 转发模式，分布层交换机则作为二、三层之间的“分界线”。对于这一特殊的设计方法，分布层交换机会担当终端主机的默认网关。

---

① 原文是“The campus network architecture is based on the use of two basic blocks or modules connected through the core of the network”，译者也不明白，为什么分布层和接入层都变成“模块”和“区块”了，既然不明白，译文只能按照原文字面意思翻译。

② 原文是“This is mostly limited by the control plane handling of the distribution layer, whether it is a Layer 2 or a routed access design”，这种文字直译出来不易读懂，译者只好自己杜撰译文。

基于 VLAN 的 Trunk 技术将一个个子网从分布层交换机“下放”到了接入层交换机。在分布层交换机上，不但会运行着第一跳冗余协议（比如，HSRP[热备份路由协议]或 GLBP[网关负载均衡协议]），还会运行路由协议，以提供与园区网络区块核心层之间的数据包转发功能。在适当的情况下，可能还会在交换机的接入端口或交换机间的链路上，激活某种版本的生成树协议，并启用相应的生成树加固特性（比如，Loopguard、Rootguard 以及 BPDUGuard 特性等）。尽管上述技术和特性对园区网络的部署非常重要，但与 IPv6 无关，因此本书不会在这些方面做过多纠缠。

第二层接入设计还有两个“变种”，定义 VLAN 的方式是两者之间最主要的差别。

- **环路式设计**：跨多台接入层交换机配置 1 个或多个 VLAN。如此一来，跨交换机配置的每一个 VLAN 都会形成生成树或第二层环路拓扑。

- **V 式（无环式）设计**：该设计方法遵循了当前多层网络设计的最佳做法指南——在每台接入层交换机上，只有一个 VLAN。这是一种在网络拓扑中消除环路的设计，其主要优点包括：利用 GLBP 特性来实现的每设备上行链路负载均衡、降低了网络恢复对生成树协议的依赖性、减轻了广播风暴的危害性、将单播泛洪发生的可能性（以及与非对称 L2/L3 网络拓扑有关的设计方面的难点）降至了最低。

本书第 6 章涵盖了上述两种设计方法及相关配置的详细信息。

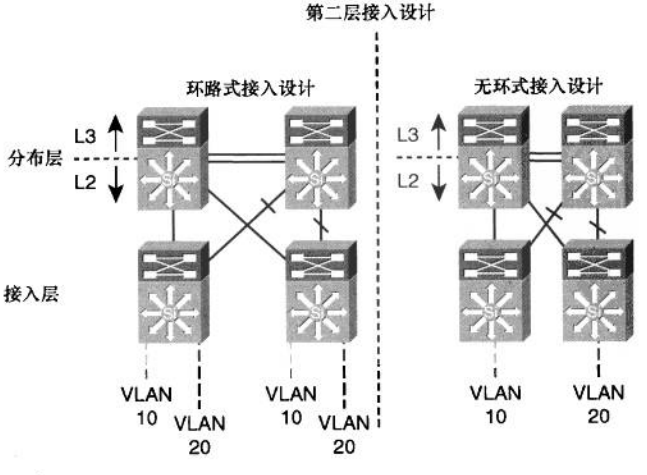


图 2-3 第二层接入设计

路由式接入设计

还有一种可替代传统分布层网络区块模型的设计是，让接入层交换机担当功能齐备的第三层路由节点（具备 L2/L3 交换机功能），并以第三层路由式点到点链路取代接入分布层交换机的第二层上行 Trunk 链路。对该设计来说，二/三层的分界线从分布层交换机变为了接入层交换机（如图 2-4 所示），乍一看，这似乎是设计方面的重大改变，但实际上只不过是略微扩展了多层网络设计的最佳做法。本书第 6 章涵盖了路由式接入设计及其相关配置的详细信息。

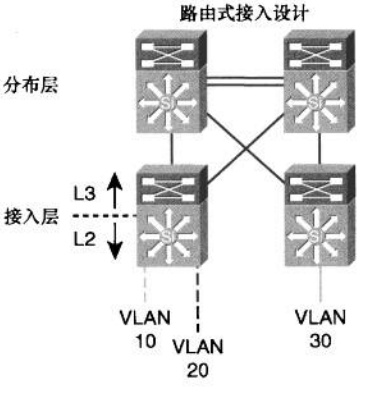


图 2-4 路由式接入设计

与利用第二层上行链路接入分布层的多层网络设计相比，路由式接入设计的优点如下所示。

- 可使用常见的故障排除工具（比如 ping 和 traceroute），端到端地去定位网络故障[①]；使用单一的控制平面协议（EIGRP 或 OSPF）[②]；无须使用如 HSRP 之类的特性。
- 该设计适用于绝大多数网络环境，但不支持让同一 Vlan 跨多台交换机运行的场合。

路由式接入设计虽然提供了额外的好处，但也给网络设计提出了如下挑战。

- 实施路由式接入设计时，需精心规划，以避免路由环路；需执行适当的

① 第二层（交换式）接入设计无法使用 IP 故障排除工具，比如 ping 和 traceroute 等，去端到端地定位网络故障，这是因为接入层交换机不运行 IP 协议。

② 不用运行 STP。

路由汇总，以确保网络的可扩展性（为新用户和新接入层交换机接入网络做好充分的预留）。

- 采用第二层接入设计时，可在连接到同一分布层的多台接入层交换机上，轻而易举地扩展同一子网。实施路由式接入设计时，要是跨两台交换机扩展同一子网，便会导致 IP 地址冲突，这对网络设计者来说可算是一道难题，因此除了在划分子网时要“精打细算”以外，在分布层还需执行路由汇总。
- 路由式接入实施起来代价不菲，因为每台接入层交换机的软硬件都需要支持第三层的功能。

虚拟交换系统分布层设计

虚拟交换系统（VSS）分布层设计（如图 2-5 所示）是对典型的交换/路由式接入层设计根本性的改变。过去，多台接入层交换机一般都上连到两台互为冗余的分布层交换机，网络控制平面协议的配置（比如，HSRP 和 802.1D STP）决定了交换机在每条上行链路上转发流量的方式，以及交换机本身或链路故障时网络的恢复方式。随着虚拟交换机概念的提出，如今，可配置一对分布层交换机，令两者以单台逻辑交换机的方式运行。

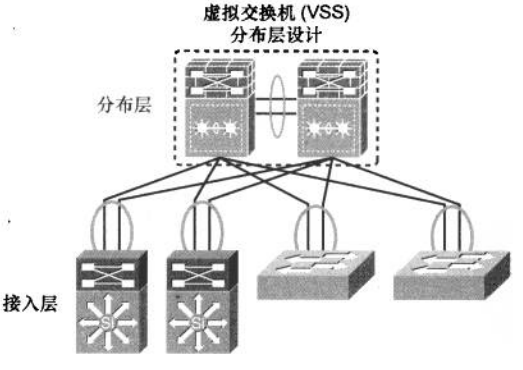


图 2-5 虚拟交换系统分布层设计

两台分布层物理交换机摇身一变为单台逻辑交换机之后，网络拓扑也势必会发生重大改变。在先前的接入层-分布层网络设计中，单台接入层交换机会设有两条上行链路连接两台分布层交换机，使用哪条链路传递流量则需控制平面协议来决定。部署了 VSS，接入层交换机会采用一条逻辑 MEC（多机箱以太网通道）上行链路连接到单台（虚拟出的）分布层交换机。VSS 架构将在第 6 章做深入探讨。

分布层设计比较

虽然以上三种接入层-分布层设计方法各有千秋，但虚拟交换机式和路由式接入设计方法要胜过传统的多层网络设计方法。以上两种新设计选项和传统的多层网络设计方法之间存在的差异包括网络配置和运行的简易程度、逐流（per-flow）的上下游负载均衡方式，以及收敛速度的快慢等。表 2-2 比较了本节所介绍的三种设计选项。

表 2-2 分布层设计方式的比较

特性	第二层接入设计	路由式接入设计	虚拟交换系统设计
接入分布层控制平面协议	STP（PVST+、快速 PVST+ 或 MST）	EIGRP 或 OSPF	PagP、LACP
同一 VLAN 是否跨越多台接入层交换机	是（形成 L2 生成树环路）	否	是
第三层边界	分布层	接入层	分布层
第一跳冗余协议	需 HSRP、GLBP、VRRP	无	无
接入分布层访问的逐流负载均衡	不支持	支持-等开销多路径（ECMP）	支持-MEC
收敛时间	900 毫秒～50 秒之间（取决于 STP 拓扑以及对 HSRP 调优）	50～600 毫秒	50～600 毫秒

2.3.2 接入层

接入层是企业网的最外围，或用户与企业网的第一“触碰”点。该层是终端设备接入网络的访问点。接入层也是可以发起对网络服务访问的最外围之处。可连接到接入层的设备五花八门，其中包括 PC、服务器、IP 电话、无线接入点、摄像头以及其他的PoE/PoE+设备。表2-3列出了接入层交换机所提供的某些服务。

表 2-3 接入层交换机所提供的服务和特性

服务要求	特性
协作式服务（collaboration service）	激活语言/视频应用：PoE 和 QoS 标记行为、监管、排队 应用程序可视化服务：灵活的 NetFlow 移动服务：统一有线/无线定位服务 虚拟化服务：VLAN、VRF-lite
自动化服务	自动化智能端口（Auto Smartports）、Smart CallHome
安全服务	访问控制：802.1x 和端口安全 控制平面监管（CoPP）、DHCPv6 中继、IPv6 路由器防护、IPv6 端口控制列表（PACL）

续表

服务要求	特性
高弹性	状态化切换（SSO）、不停止转发（NSF）、不中断业务的软件升级（ISSU）
智能网络控制服务	PVST+、快速 PVST+、EIGRP、OSPF、DTP、PAgP/LACP、UDLD、FlexLink、Portfast、UplinkFast、BackboneFast、LoopGuard、BPDUGuard、RootGuard

2.4 企业网络服务区块设计

相对于企业园区网网络区块，企业网络服务区块可算是相对新鲜的事物。当企业网络的规划者开始考虑将网络迁移到 IPv4/IPv6 双栈网络环境，并继续部署更为先进的统一通信服务时，势必会遭遇诸多挑战。至关重要的是，在将上述服务平滑地整合进企业园区网络区块的同时，要能够适度保持变更管理和故障隔离的可操作性。此外，在服务整合的过程中，园区网络区块还需继续保持灵活而又可扩展的设计风格。比方说，可通过搭建临时性的覆盖型隧道来部署 IPv6 服务，以使得 IPv6 设备发出的 IPv6 数据包以隧道封装的方式，通过园区网内尚未启用 IPv6 的区域。这一临时性措施可更为快捷地引入新型服务，而无须在全网范围内实施相关割接操作[①]。建议将以下设备（服务）部署在企业网络服务区块。

- **集中式无线控制器：**由控制器来调配并控制遍及整个企业园区网内的无线接入点。
- **集中终结企业园区网络区块至网络服务区块之间的站点内 IPv6 ISATAP（自动隧道寻址协议）隧道（Centralized IPv6 Intra-Site Automatic Tunnel Addressing Protocol (ISATAP) tunnel termination from the enterprise campus to the network services module）：**这可以对覆盖在现有网络之上的隧道网络形成紧密控制。类似于所有隧道技术，在网络中架设多条通往不同网段的 ISATAP 隧道，除了会增加网络管理的复杂性以外，还会增大故障排除的难度。
- **统一通信服务（Cisco 统一通信管理器、网关）：**为了方便统一通信服务的集中化管理，企业一般都会将呼叫管理器和其他的语言网关设备部署

① 原是“Such an interim approach enables a faster introduction of new services without requiring a networkwide, hot cutover.”译文为直译。在译者看来，作者只会用 enable 和 provide 这两个实意动词。到目前为止，enable 这个动词至少出现了不下 50 次。

在企业网络服务区块。

- **策略网关**：策略网关可用来行使用户认证、授权以及网络访问控制（NAC）等功能。典型的策略网关包括 AAA（认证、授权、记账）服务器、ACS（访问控制服务器）和 NAC 分析器（profiler）等。

2.5 企业数据中心网络区块设计

与企业园区网络区块相比，数据中心网络区块设计起来差别不大，只是在特性、产品以及性能方面有细微差别。除此之外，路由和交换方面的设计全都相同。数据中心网络区块也采用多层架构模型，所包含的层次如下。

- **汇聚层（分布层）**：接入层的终结点，用来将数据中心网络区块连接到核心网络区块。在某些设计中，许多企业都会考虑在数据中心区块中保留核心层，对于这一保留核心层的数据中心设计方案，仍需恪守前面提到的核心层网络设计原则。在该区块中，汇聚层也同样作为服务层，提供 4～7 层的服务，包括安全（防火墙）、服务器负载均衡（SLB）和监控服务等。
- **接入层**：接入层既可以使用 Catalyst 和 Nexus 交换机提供物理方式接入，也可以利用基于 hypervisor 的软件交换机（比如 Cisco Nexus 1000v）提供虚拟方式接入。接入层用来将虚拟机和物理服务器（bare-metal server）连接到网络。一般而言，服务器会以 10/100/1000M 的速率连接到接入层，而接入层交换机会以 10G 的上行链路连接到汇聚层。

采用上述的多层架构模型，使得数据中心网络区块的规模能够随着负载的增长和需求的变化而“与时俱进”[①]。

2.5.1 汇聚层

汇聚层是数据中心网络基础设施二、三层的分界线。通常，汇聚层交换机会以 10G 链路分别上连和下接企业核心网络区块和每台接入层交换机。汇聚层是非常理想的流量过滤点，可作为保护数据中心网络区块的第一道屏障。可在该层划分子区块来部署防火墙服务，过滤进进出出的流量。此外，还要在汇聚层提供第二层和第三层的流量对称模式，以支持状态化的包过滤机制。

在某些设计中，汇聚层也会作为服务层，来提供 4～7 层的服务，包括安全

① 原文是“This architecture enables data center modules to be added as the demand and load increase”，动词又是 enable，译者无语。

（防火墙、Web 应用防火墙[WAF]），服务器负载均衡（SLB）以及监控服务（网络分析模块[NAM]）等，如表 2-4 所示。

表 2-4 部署在汇聚层的 4～7 层服务

4～7 层服务	功能
安全	在许多网络设计中，并没有考虑到安全性（总是“亡羊补牢”）。其实，只要把安全性当做网络设计的基本要求（而不是作为附加选项）加以考虑，实现起来并不麻烦。对数据中心网络区块来说，对安全性的要求会随时间而变——新应用程序的推出、合规性要求的调整、商业的兼并以及安全性的违背等诸多因素，都会促使其发生改变。在数据中心网络中，需重点考虑的安全性包括三个方面：隔离（VLAN、VRF）、安全策略的执行，以及物理服务器/虚拟机的“能见度”
服务器负载均衡（SLB）	SLB 技术可将用来支撑服务的资源，归并为单个虚拟的访问点，以确保高性能和高弹性。归并资源的 SLB 设备“掩盖”住了多台服务器的真实 IP 地址，并以单 IP 地址取而代之，为客户端提供一种或多种服务，比如 HTTP、HTTPS 以及 FTP 等[①]
Web 应用防火墙（WAF）	WAF 为基于 Web 的应用提供防火墙服务。WAF 可加固并保护 Web 应用，令其免遭诸多常规攻击（比如身份/数据窃取、应用中断、欺骗以及其他针对性攻击等）之苦。常规攻击的具体形式包括跨站脚本攻击（XSS）、SQL 和命令注入、权限提升、跨站请求伪造（CSRF）、缓存溢出、cookie 篡改以及 DoS 攻击
监控	一般情况下，监控服务并不会部署在参与流量转发的设备上，而会集成进某台被动（采集）设备中，以监控连绵不断的数据流。除了具有调试和故障排除的用途之外，还可以利用监控服务来制定容量规划[②]

2.5.2 接入层

就传统意义而言，接入层交换机的作用是，连接服务器和应用程序农场（server farm and application farms）。发往/来自特定服务器或基于 Hypervisor 的主机，以及基于 Hypervisor 的主机之间的流量（Any communication to and from a particular server/hypervisor-based host or between hosts），都会途经接入层交换机或相关设备（比如，防火墙或负载均衡器等）。有以下两种部署模式适用于服务器接入，一般而言，这两种模式都会单独部署。

- **行尾（End of Row [EoR]）部署法（一排服务器机架的末尾安装一座网络机架，内置汇聚层交换机）**：EoR 模式是指利用一台汇聚层交换机提供多座服务器机架之间的连通性。一般情况下，在 EoR 设计中，都是单

① 整段原文是“The SLB technology abstracts the resources supporting the service to a single virtual point of contact to ensure performance and resiliency. Abstraction SLB devices mask the server’s real IP address and instead provide a single IP for clients to connect over a single or multiple protocols, including HTTP, HTTPS, and FTP”。这种文字简直是不忍卒读。

② 原文是“Monitoring services are typically not integrated into the actual data flow but are integrated as passive devices to help with ongoing monitoring of flows and for debugging and troubleshooting purposes. The monitoring services can also be leveraged for capacity planning”。作者的文字表达能力一般，译文酌情更改。

台交换机汇接多台服务器，因此可实施单点管理。

- **架顶（Top of Rack [ToR]）部署法：**ToR 模式是指在服务器机架内部署一台 1/2RU 的交换机，用来连接物理服务器（以 1G 或 10G 链路）。ToR 交换机会以 10G 上行链路汇接到汇聚层交换机。每座服务器机架的顶端都会安装一台独立的交换机（独立运行软件映像），因此需要单独配置和管理（每个机架内的设备）。

在接入层实施网络安全性的主要作用是保护第二层数据流。推荐的最佳做法是：利用 VLAN 来隔离流量，用访问控制列表（ACL）阻止任何不必要的流量。可在接入层部署的辅助安全特性包括私有 VLAN（PVLAN）和端口安全特性，其中端口安全特性包含了动态 ARP 检查、IPv6 路由器保护，以及 IPv6 端口访问列表（IPv6 ACL）等一系列子特性。此外，端口安全特性还可以将某些重要的服务器“锁定”于特定的端口。

虚拟接入层是指配置虚拟化时，驻留于物理服务器内的虚拟网络。服务器虚拟化给网络的安全性、能见度以及安全策略的执行情况提出了新的挑战，理由是虚拟机之间的通信流量可能既不会流出实际的物理服务器，也不会途径任何一台物理交换机。在此类网络环境中，要想贯彻并实施网络安全策略可不是一件容易的事。但在这一新型虚拟接入层中，网络设计目标追求的依然是：提供不逊于传统接入层所使用的诸多安全服务和安全特性。

虚拟接入层不但驻留于运行虚拟化软件的物理服务器之内，而且还有可能跨越多台运行虚拟化软件的物理服务器。要想建立虚拟机到物理服务器的连通性，还需在运行虚拟化软件的物理服务器内组件虚拟网络。

2.5.3 数据中心存储网络设计

存储区域网络（SAN）用来将服务器“接驳”至远程计算机存储设备（比如，磁盘阵列和磁带库等），其“接驳”方式酷似于将服务器的操作系统“挂接”在本地设备上。企业网中应用程序的增多，致使计算和存储空间的需求[①]保持同步增长。SAN 具备以下优点。

- **应用程序的高可用性：**应用程序独立于存储，可通过多种方式来访问。
- **更高的应用程序性能：**减轻了服务器处理数据存储的负担。
- **存储和可扩展性相结合：**让存储系统管理起来更为简单、更具灵活性、

[①] 原文是“demand of both compute as well as storage space”，“compute”可是动词！

更具可扩展性。

- **故障恢复**：可使用 FCIP（IP 上的光纤通道）等故障恢复特性远程复制数据。

自发明之日起，SAN 便一直将光纤通道用作为其底层传输技术。光纤通道技术与以太网技术截然不同，有时也称其为“服务器背后的网络”。SAN 已从传统的服务器发展到了包括交换机在内的设备间的存储连接，这样的交换机都专门内置了传输光纤通道协议（FCP）命令。在 SAN 领域，人们也把此类交换机称为“fabric”。

如图 2-6 所示，SAN 与以太 LAN 毫不相干。应用程序客户端通过园区以太网络访问服务器。服务器利用 SAN 执行数据 I/O 操作。

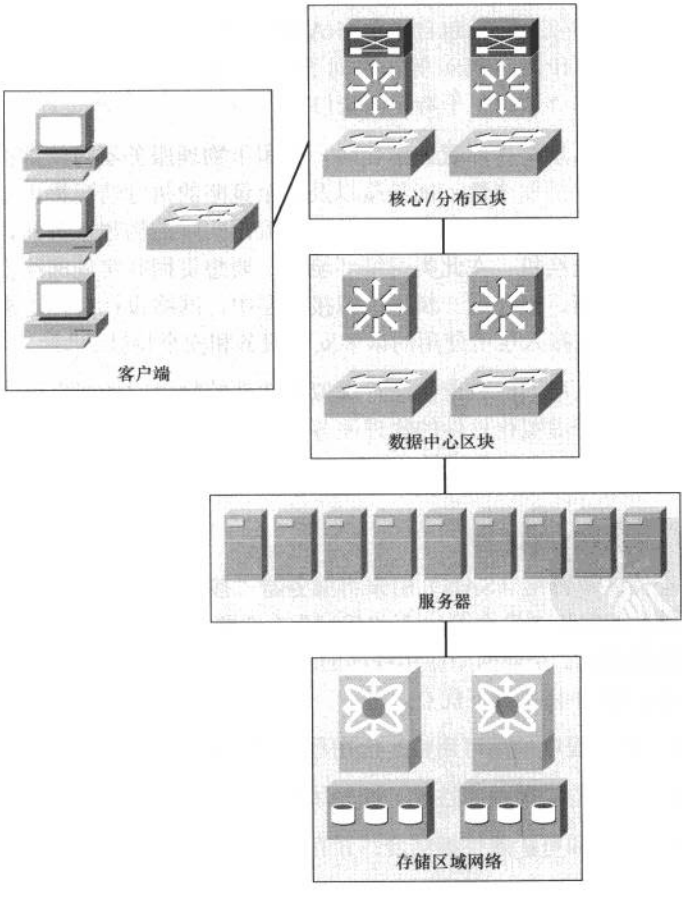


图 2-6 SAN 的设计

设计 SAN，不外以下两种拓扑。

- 紧缩式核心拓扑（Collapsed core topology）。
- 核心边缘拓扑（Core edge topology）。

以上两种拓扑将会在下面两小节介绍。

紧缩式核心拓扑

在该拓扑中，主机服务器和存储设备都连接在本地交换机上。连到 fabric 的交换机利用 ISL 作为交换机间的通信手段。

随着交换机上的端口逐渐被占满，人们会利用 ISL，将新交换机添加进 fabric。人们还会通过精心规划，尽量减少由 ISL 承载的交换机间的通信流量，理由是绝大多数流量应被限制在交换机本机。随着新交换机的不断添加，以及 fabric 规模的日益庞大，要想提前规划好存储的位置，保持原有的流量模式，势必难上艰难。在此情形，便需要重新评估交换机间 ISL 的带宽，其原因是由 ISL 来承载的用户到存储的访问流量所占链路带宽的比重会越来越高。就紧缩式核心拓扑而言，随着时间的流逝和 fabric 规模的“日益臃肿”，要想确定流量模式（ISL 链路所承载流量的比重）会变得愈发困难。

图 2-7 所示为紧缩式核心拓扑。

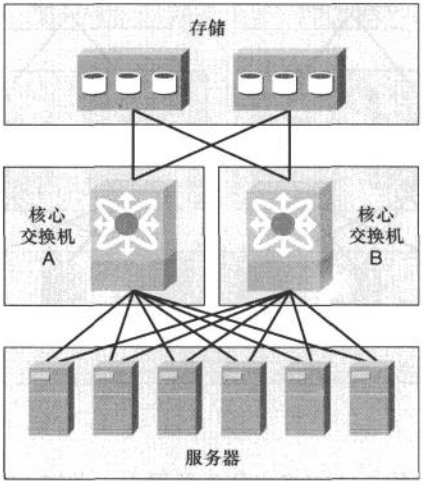


图 2-7 紧缩式核心拓扑

核心边缘拓扑

核心边缘拓扑由两个互为冗余的 fabric 构成。每个 fabric 都包括一台核心交换机和多台边缘交换机。对核心边缘拓扑架构来说，在每个 fabric 内，核心交换机支持所有存储或目标端口，并通过 ISL 与边缘交换机建立连通性。在该拓扑中，存储端口被统一汇接至核心层；主机则连接到边缘交换机，后者再通过 ISL trunk 链路连接至核心交换机。由于存储全都统一汇接至核心交换机，因此在核心交换机上便能开启某些 SAN 高级特性，以消除并降低管理 SAN 的复杂性。此外，采用该拓扑，还能够很容易地确定主机和存储之间的流量过预定（host-to-storage oversubscription）之比。

图 2-8 所示为核心边缘拓扑。

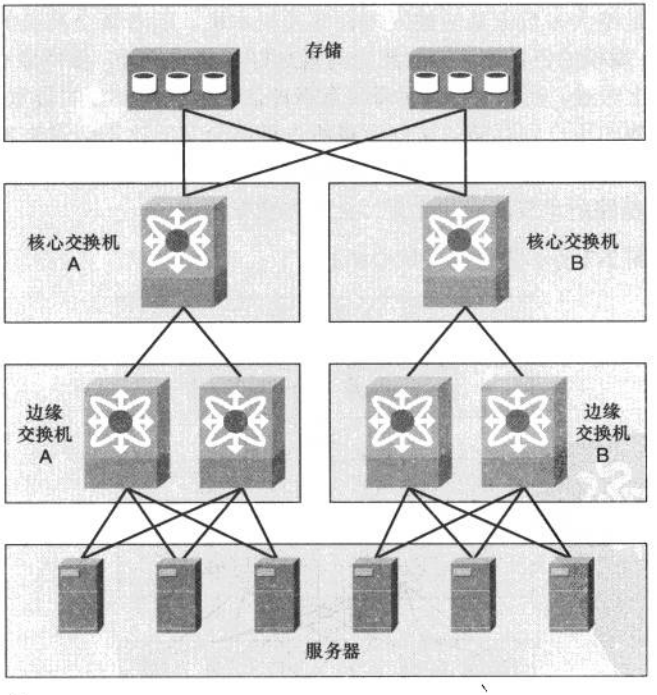


图 2-8 核心边缘拓扑

要想将孤立的 SAN（多半位于数据中心之间）连成一气，可采用诸如 FCIP（IP 上的光纤通道）之类的技术，以 IP 的形式来封装光纤通道，在 fabric 交换

机（比如核心 fabric 交换机，或其他支持 FCIP 的交换机产品）之间建立隧道，传递流量。

在一个公共的无损 10Gbit/s 以太网络上，同时传递以太网和存储流量被称为统一 fabric 或统一 IO。支撑同一 fabric 的底层技术被称为以太网上的光纤通道（FCoE）。

当前，对绝大多数企业网来说，都必须维护两套异地数据中心网络，这两套网络无论在管理上还是在实施上都截然不同。把存储流量也牵引至以太网，可让存储和用户应用程序数据在同一套物理网络上奔流不息。

图 2-9 所示为统一 fabric。

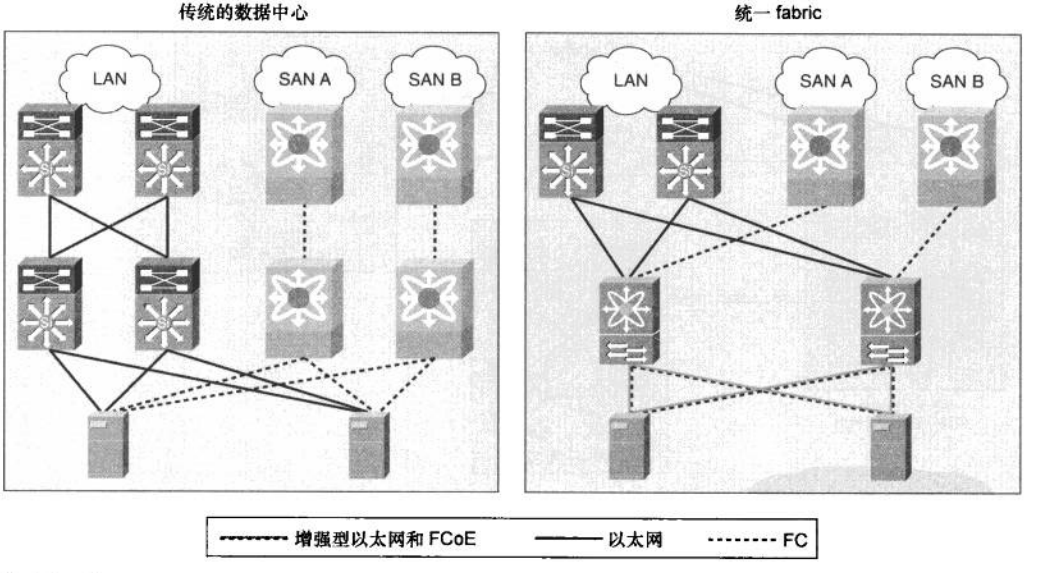


图 2-9 统一 fabric

2.6 企业边缘网络区块设计

企业边缘网络区块由 Internet、虚拟专用网（VPN）以及 WAN 模块组成，企业网正是通过这三个模块与服务提供商相连。企业边缘网络区块包含的所有网元的作用是：高效而又安全的保障企业园区网络区块和数据中心网络区块与远程站点、商业伙伴、移动用户以及 Internet 之间的通信。边缘网络区块汇集了

来自各远程站点的流量，对流量执行过滤，将流量路由进企业网内部。为满足不同规模的客户，以及各自的业务模式的需求，以模块化的方式设计边缘网络可使其兼具灵活性与专用性。对企业网来说，其边缘网络区块由以下三个关键网络区域组成。

- 前置设备区域，也常被称为 WAN 聚合网络。
- 连接到 WAN 聚合网络的分支机构网络。
- 面向 Internet 的网络区域，用来提供 VPN 和常规 Internet 接入服务。

图 2-10 所示为企业边缘网络区块所属各模块之间的关系。

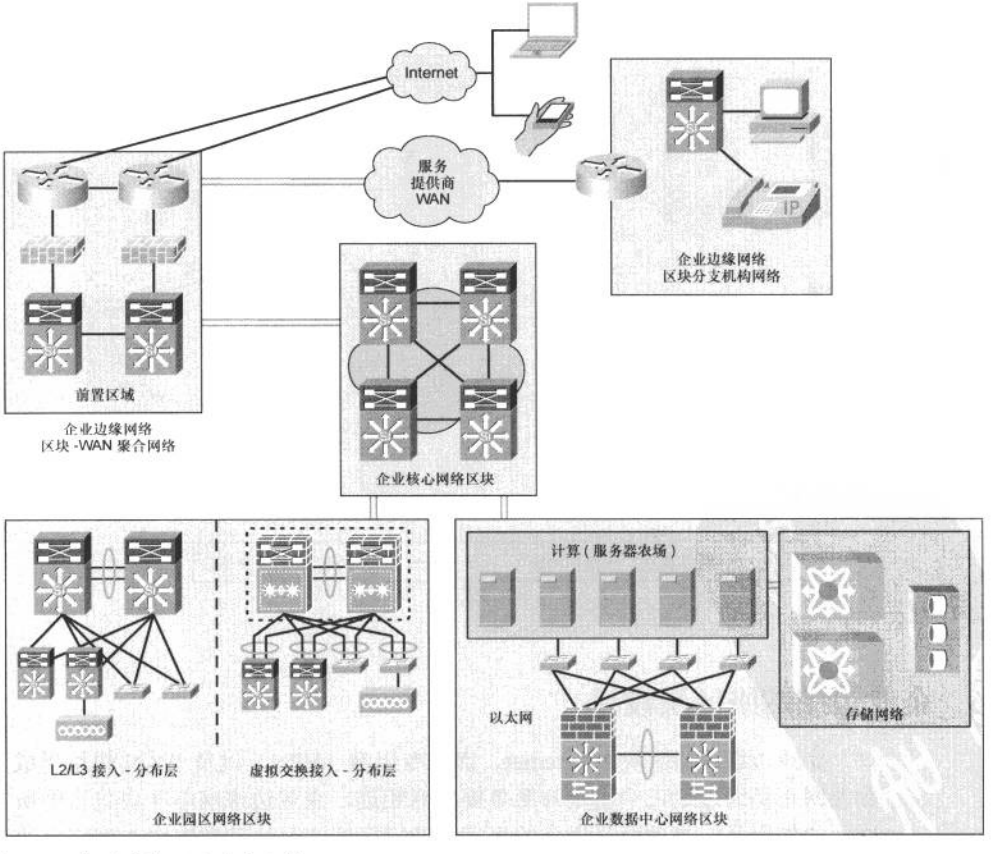


图 2-10 企业边缘网络区块设计

图 2-10 所示的每个网络区块（模块）都包含了特点鲜明的网络组件，并提供具备特定功能的服务。下一节将会介绍企业总部边缘网络区块中的各种常用组件。

2.6.1 企业总部边缘网络区块组件

在企业边缘网络区块内，部署了可提供各种功能的网络组件。取决于网络的规模和复杂程度，这些组件既可以集高性能、高可用性、高可扩展性、高安全性和各种丰富的特性于一身，也可以将多个组件合而为一，并令其只提供最基本的功能。表 2-5 列出了企业边缘网络区块常见的网络基础设施组件。

表 2-5 企业网边缘网络区块组件

网络基础设施	功能
边缘路由器	提供到服务提供商的 WAN/Internet 基本连通性 “奋战”在防欺骗和其他网络攻击的最前线 是网络流和事件信息的采集点，可对采集到的数据进行分析，或将其作为其他相关用途 部署两台或多台路由器并配置第一跳冗余协议（FHRP），以实现冗余
外部交换机	提供防火墙和边缘路由器间二、三层的连通性 部署两台交换机，以实现冗余 外部交换机还要负责提供到抗 DDoS（分布式拒绝服务）攻击设备的连通性（The outer switches provide connectivity to the distributed denial of service (DDoS) mitigation component）
防火墙	通过控制并检测进、出 DMZ 安全域和 WAN 聚合网络区域的流量，完成对以上两个模块的保护 以状态化的主/备(active/standby)故障切换模式部署防火墙，来实现冗余
附加的安全性	部署 IDS（入侵检测系统）的目的是，识别众所周知的攻击，对可疑行为发出警报 监控和分析系统会处理并分析由 IDS 和端点安全软件所生成的警报或告警信息[①] 识别 DDoS 和其他基于网络的攻击，并与部署在外部交换机上的抗 DDoS 系统协同工作

2.6.2 企业总部边缘网络区块设计

企业边缘网络区块提供了一套公共的网络基础设施，并在其上建立到分支机构、Internet、远程办公以及到商业伙伴（外联网）的连通性，以此来降

[①] 原文是“Alerts and alarms generated by the IDS and the endpoint security software are processed by a monitoring and analysis system for analysis and correlation purposes”，译者认为后面的“for analysis and correlation purposes”是废话，用不着翻译，译文做了相应的精简。

低基本建设费用和运营成本。表 2-6 罗列了这套网络区块边缘基础设施的设计需求。

表 2-6 企业总部边缘网络区块设计需求

设计需求	特性
网络的高弹性	在边缘网络区块内，需确保无单点故障，只有这样，才能最大限度地发挥这套网络基础设施的高可用性。该区块任意一次网络故障都会影响到对数据中心的网络资源（内联网或外联网资源），以及对商业伙伴的访问，从而造成业务损失。边缘网络区块是由多种平台、多个组件搭建而成，无论哪种平台、哪个组件都有可能会发生故障，并遭受攻击。因此，设计的关键是要通过部署多层次的冗余，来消除单点故障，这包括接口、设备以及网络拓扑层面的冗余
合规性	边缘网络区块的设计需满足以下标准的规定： 对于支付卡行业，需满足支付卡行业数据安全标准（PCI DSS） 对于医疗保健行业，需满足健康保险流通与责任法案（HIPAA） 任何违规行为都会导致吊销营业执照、严厉的处罚，甚至法律诉讼
模块化和灵活性	只要坚持模块化的网络设计原则，就能够在 Internet 和分支机构所用的网络和服务范围之间划清界限 只要坚持灵活性的设计原则，就能够在不影响现有设计的情况下，灵活地在网络中添加 IPSEC 和 WAN 优化等服务
无所不在的安全性	边缘网络区块需确保应用程序、端点以及网络的机密性、完整性和可用性

2.6.3 分支机构的网络架构

上一节罗列了位于企业总部的边缘网络区块架构和设计方面的注意事项。本节则会介绍位于企业的远程站点和分支机构的边缘网络设计组件。分支机构网络架构最主要的设计目标是要设计出具备灵活性、可扩展性、可靠性以及安全性的网络基础设施。

分支机构的网络架构包括以下关键组件。

- **边缘路由器**：边缘路由器（亦称为接入路由器或分支机构路由器）的作用是，在分支机构网络和总部企业网边缘网络区块之间（通过一条或多条 WAN 链路，或与一家或多家 Internet 服务提供商[ISP]）建立连通性。为获悉网络的行为和系统的状态，并能够“洞察”穿梭往来的流量，需在边缘路由器上激活 QoS、速率限制、安全（ACL 和 uRPF）、NetFlow 和 SNMP 等重要特性。此类路由器既可以利用内置的防火墙软件，也可以借助专门的硬件防火墙，来保护分支结构网络不受外部黑客的入侵，并能同时建立起一条分支机构网络和总部企业网边缘网络区块之间的安全通信信道。

- **内部交换机**：此类交换机用来建立边缘路由器与分支机构 LAN 之间的连通性。对于规模较小的分支机构，用来连接边缘路由器的交换机可能只有 1 至 2 台。而规模较大的分支机构可能会沿袭企业园区网区块的基本网络设计——采用传统的核心、分布以及接入三层设计模型。

2.6.4 分支机构边缘路由器的功能

边缘路由器的作用是，在分支机构网络和总部企业网边缘网络区块之间建立连通性。除了提供基本的网络连通性之外，边缘路由器还可以用来提供防火墙服务和应用程序识别（application intelligence）方面的服务，其目的都是为了在远程站点之间建立安全而又优化的通信信道。分支机构边缘路由器所提供的主要服务包括：

- 通过 QoS、集成进路由器软件的 NetFlow 以及应用加速（Cisco 广域应用服务 Cisco [WAAS]）等特性，来提供优化的应用程序交付服务；
- 安全性，包括硬件平台、WAN 及 VPN（涵盖了使用公钥基础设施[PKI]的安全 VPN 认证）。

整个分支机构网络都围绕边缘路由器基础设施来构建，许多高级特性也构建于其上。灵活性和可扩展性是边缘路由器基础设施的两个最重要的设计原则。只有本着灵活性的设计原创，才能将诸多高级网络服务（特性）无缝而又高效地集成进网络基础设施。只有本着可扩展性的设计原则，才能在不中断现有网络运行的情况下，轻而易举地提升网络基础设施的容量、添加分支机构站点。

在设计分支机构边缘网络解决方案时，应顾及各种各样的需求，比如，规模、业务的纵向发展（business vertical）、位置、成本等。针对分支机构的网络设计，人们推出了三套设计模板（模型），即：单层分支机构网络、双层分支机构网络以及多层分支机构网络模型。这三套设计模板将在第 8 章做深入讨论。表 2-7 则扼要罗列了这三套设计模板。

表 2-7 分支机构网络设计模板

分支机构网络设计模板	注意事项
单层分支机构网络	单层分支机构网络解决方案是指集所有功能与一体的单机箱解决方案。LAN 和 WAN 连通性都是由一台设备来提供。若企业分支机构不要求硬件平台方面的冗余，则可采用该设计模板。在按该模板设计的远程分支机构网络中，可容纳的用户数通常介于 20～30 之间

续表

分支机构网络设计模板	注意事项
双层分支机构网络	双层分支机构网络解决方案提供了两层架构[①]。第一层提供的是互为冗余的 WAN 功能，第二层则提供了 LAN 连通性。相较于单层分支机构网络，使用双层分支机构网络解决方案时，分支机构会申请两条 WAN 链路，以提供高可用性，此外，还会部署专用以太网交换机，为更多用户提供 LAN 接入
多层分支机构网络	该设计模板把整个网络的功能“分层治之”。每一层分别行使不同的功能，功能包括 WAN 终结功能、防火墙功能、服务终结以及 LAN 终结功能。该设计模板最主要的优点是：高冗余性、高可用性以及低路由器/交换机的 CPU 利用率[②]

2.6.5 典型的分支机构网络设计

每个分支机构的网络都需要与外部世界（总部或 Internet）连接，取决于分支机构的规模，连接的速率从 1.5Mbit/s（小型分支机构，最多 100 个用户）到 45Mbit/s（大型分支机构）不等。对于某些分支机构网络，可采用传统的私有 WAN 技术，比如 MPLS 或帧中继；当然，也可以采用 VPN 技术，对 Internet 连接充分加以利用，以连接远程站点。分支机构的路由器提供以下服务。

- **安全服务**：取决于分支机构所生成的流量，可选择将防火墙或 IPSEC 等安全服务集成进路由器，或利用单独的设备来提供此类服务——既可以使用安装在路由器中的模块，也可以采用 Cisco ASA 防火墙之类独立的专用设备。
- **统一通信服务**：服务包括本地呼叫控制、直连 PSTN（公众电话交换网）的 FXO/FXS 端口，以及备份连接。
- **应用智能服务**：一视同仁地为分支机构的网络用户提供公司总部的网络功能和服务级别，是分支机构网络解决方案的关键设计目标之一。因为 WAN 链路的低带宽及固有的延迟问题，这一目标实现起来并不容易。一般而言，对于 WAN 带宽有限的小型分支机构来说，可受益于应用智能服务，比如，使用 Cisco WAAS 来实现的 WAN 优化/应用加速。

以下列出了安装在分支机构的第二层接入交换机应具备的主要特性。

- 以太网上的供电（PoE）。

---

① 其实这是废话。

② 原文是“The significant benefits of this profile are redundancy, availability, and router/switchCPU utilization”，其中“router/switch CPU utilization”也能算是优点，译者真是无话可说，只能给出原文。译文微做调整。

- 生成树。
- 接入端口上的 CoS（服务类别），以及用来上连边缘路由器的端口上的 QoS 监管和整形特性[①]。

此外，某些分支机构还会部署某些本地化的网络服务，比如，本地的微软活动目录服务、本地的 DNS 和 DHCP 服务器等。由于这些服务都与 IPv6 相关，因此可配置分支机构边缘路由器，令其担当 Cisco IOS DHCPv6 服务器，为本地站点分配地址；也可以将分支机构边缘路由器配置为 DHCP 中继代理，以转发发往公司总部站点的本地 DHCP 请求。

就企业边缘网络区块的设计理念而言，大多数企业都已觉察到，自己目前所用的 IPv4 网络环境与日后将用的 IPv6 网络环境很可能别无二致。

2.7 总结

本章高屋建瓴地介绍了层次化企业网络设计。首先讨论了关键的设计三原则，即模块化、层次化和高弹性。然后，就需求方面，对平面化网络设计和分层的模块化网络设计进行了比较。因网络规模渐长，伴随着应用程序对网络带宽和网络服务的要求提高，故而也需要以新的方式来设计网络。

坚持网络设计三原则，有助于将企业网络划分为更为简单的多个区块，每个区块都行使自己的功能。本章还总结了设计上述区块时的注意事项，网络工程师/架构师需对不同的设计选项仔细研究，以便在自己的企业网设计中应用相关设计原则。

企业园区网络区块为末端用户和设备提供连通性，在介绍该区块时，还对隶属于其的各种分布层子区块设计（第二层接入设计[环路式和无环式]、路由式接入设计以及虚拟交换机式设计）进行了比较。

企业网络服务区块是一个全新概念。在 IPv4 网络向 IPv4/IPv6 双协议栈网络过渡期间，该区块可让网络架构师减少端点间隧道建立的范围。

本章 2.5 节涵盖了与服务器农场和存储网络 IP 连通性有关的网络设计，重点介绍了 SAN 设计。尽管 IPv6 能否与 SAN 和谐运作还尚未可知，但本节还是点出了 SAN 在企业网络中的重要性，并给出构建 SAN 的各种网络拓扑。

[①] 原文是“Class of service (CoS) on access ports and QoS policing and shaping on edge routers”。明明是在说分支机构的第二层接入交换机应具备的主要特性，却又扯到了边缘路由器上，不知道作者是在想什么。译者只好替作者打个圆场，修改了译文。

本章2.6节着重介绍了远程站点和分支机构连接到企业网络的各种方式，并讨论了与Internet连通性有关的高级概念[①]。此外，本节还详细介绍了WAN/分支机构网络所必备的各种特性和安全需求。

2.8 其他参考资料

Cisco Enterprise Campus Architecture: Overview and Framework:

http://www.cisco.com/en/US/docs/solutions/Enterprise/Campus/campover.html.

Cisco Data Center Infrastructure Design Guide:

http://www.cisco.com/en/US/docs/solutions/Enterprise/Data_Center/DC_Infra2_5/DCI_SRND_2_5_book.html.

Data Center Design—IP Network Infrastructure:

http://www.cisco.com/en/US/docs/solutions/Enterprise/Data_Center/DC_3_0/DC-3_0_IPInfra.html.

Data Center Service Integration: Service Chassis Design Guide:

http://www.cisco.com/en/US/docs/solutions/Enterprise/Data_Center/dc_servchas/service-chassis_design.html.

Deploying IPv6 Campus Design:

http://www.cisco.com/en/US/docs/solutions/Enterprise/Campus/CampIPv6.html.

[①] 作者似乎记性不好，2.6节应该没有“与Internet连通性有关”的内容。

CHAPTER 3

第 3 章　常用的 IPv4/IPv6 共存机制

本章涵盖以下内容。

- **纯 IPv6**：本节会介绍 IPv6 新建部署（greenfield deployment），并涵盖如何部署纯 IPv6 网络的内容。
- **过渡机制**：本节将讨论双协议栈（IPv4/IPv6）共存、IPv4 隧道上的 IPv6（IPv6 over IPv4）以及 MPLS 上的 IPv6（IPv6 over MPLS）[①]。
- **协议转换/代理机制**：本节将描述人们经常提到的“转换机制”，这样的机制操作于 TCP/IP 的各个层次（比如，从网络层到应用层）。NAT-PT（网络地址转换-端口地址转换）和 NAT64 机制是本节所要讨论的重点。

当前，基于 IPv4 的编址方案已不能满足企业网日渐增长的对 IP 地址量的需求。设计 IPv6 是为了解决 IPv4 地址短缺问题。可是，如何在企业网中高效部署 IPv6，而又不对生产造成任何影响，是摆在人们面前的一道难题。为此，人们提出的方法（机制）也五花八门。本章将会带读者简要回顾以下几种比较常用的共存机制。

纯 IPv6（只运行 IPv6 的网络）：纯 IPv6 机制是指在网络中所运行的传送协议只有 IPv6 这一种[②]。

双协议栈：双栈机制是指网络中的主机和网络设备既运行 IPv4，也运行 IPv6 协议。

---

① “IPv6 over IPv4”和“IPv6 over MPLS”精确的译法应该是“利用 IPv4 协议传递 IPv6 流量”和“利用 MPLS 传递 IPv6 流量”，但由于“X 上的 Y”这一称谓已成惯用语，比如，我们都把 PPPoE 称为“以太网上的 PPP（正确的译法是应该是“利用以太网传递 PPP 帧”）。因此，译者也遵循惯例，将“Y over X”翻译为“X 上的 Y”。

② 原文就是把 IPv6 称为“transport protocol（传输协议）”，译者翻为“传送协议”。

IPv4上的IPv6：这一过渡机制是指使用IPv4数据包来封装IPv6数据包。该方法一般会用在多个IPv6网络被IPv4网络隔离，而IPv4协议是唯一能够"打破"现有隔离的网络协议的情况下。对于这种情况，IPv6只会部署于边缘网络[①]。

MPLS上的IPv6：使用该过渡机制时，IPv6网域之间通常都被IPv4多协议标签交换（MPLS）核心网络隔开。利用MPLS来传递IPv6流量，既能够提高流量的传输性能，也能够实现以更为动态的方式来传递流量[②]。

转换机制：该机制利用了诸如sock网关、网络地址转换-端口地址转换（NAT-PT）、TCP UDP中继以及NAT64之类的技术，让只支持IPv6的设备和只支持IPv4的设备进行通信。

表3-1罗列并描述了上面提到的各种高层机制。

表3-1 共存的机制

机制	子机制	优点	困难
纯IPv6	——	可扩展性：IPv6提供的地址数更多，因此既可获得全球连通性，也可以对等到对等的方式来建立连接。改进了路由选择中的路由聚合功能（但仍需精心的规划和设计）	可能需要对网络基础设施、OS、应用程序、服务（DNS、DHCP）、网络管理，以及硬件（比如，NIC卡）实施升级
过渡机制	双协议栈	在保留现有IPv4网络的同时，可在网络中建立起纯IPv6连通性[③]	由于需要运维和管理IPv4和IPv6两种协议，因此同时支撑IPv4和IPv6代价不菲
	IPv4隧道上的IPv6	利用IPv4骨干网，来互连IPv6网络	对于大规模部署，尤其是网状拓扑结构的部署来说，不但十分复杂，而且代价高昂
	MPLS上的IPv6	在无须对整个MPLS核心网络升级的情况下，利用MPLS标签栈来封装IPv6流量，并令其穿越IP/MPLS网络	可能需要对互连IPv6网络和MPLS网络的设备（比如PE路由器）实施软硬件升级
转换机制	NAT-PT	在只支持IPv4和IPv6的主机间提供基本而又有限的转换功能	可扩展性和对应用程序的支持都受到了限制。整体性能受NAT-PT设备所限。还很有可能会导致其他问题，包括DNS转换问题、双栈主机会同时获得真实的地址和经过转换的地址[④]

① 原文是"In this scenario,the IPv6 is at the edge networks only"，这个"at"令译者无语。

② 原文是"In this transition mechanism, the IPv6 domains communicate with peer IPv6 domains over an IPv4 Multiprotocol Label Switching (MPLS) core, providing more dynamic and higher performance"。译者如果不"杜撰"译文，按原文字面翻译出来，应该也没有人能够看懂。译者还想问一下作者何为"providing more dynamic and higher performance（提供更动态和更高的性能）"？

③ 原文是"providing access to IPv6-enabled applications over IPv6"。作者除了"provide、enable"以外就不会用其他的实意动词了，而且作者的表达方式笨拙至极。

④ 后半句原文为"dual-stack hosts get both native and translated addresses"。又是用"get"这种很随意的词表达关键意思。译者选择直译。

续表

机制	子机制	优点	困难
	NAT64	在只支持 IPv4 和 IPv6 的主机间提供了两种转换模式：无状态转换和有状态转换模式。前者是 IPv4 和 IPv6 间一对一的映射；后者则是 IPv6 和 IPv4 间多对一的过载映射	会遭遇与任何 NAT 设备（包括 NAT-PT 设备在内）都相同的某些限制问题，比如，应用程序的互操作性问题、可扩展性问题，以及与性能有关的问题。当然，NAT64 在上述几方面要好过 NAT-PT

3.1 纯 IPv6

纯 IPv6 也称为“IPv6-only（网络中只运行 IPv6 协议）”。这意味着 IPv6 是网络中运行的唯一一种 IP 协议。

简而言之，可以把这种 IPv6 网络环境的运行方式想象为当今在用的 IPv4 网络环境，只是 IP 版本更高而已。可以预见，运行双栈网络所需的资金和运营成本势必会让公司不堪重负。在某些时点（随网络的不同，情况的差别可能会很大），完全禁用或至少在网络中的大多处禁用 IPv4 似乎也顺理成章。

已经有客户在新建网络或新建站点中部署了 IPv6，其中，无论对于网络基础设施还是应用程序，在功能方面不但最新而且也最强，此外，还几乎能够支持所有 IPv6 的丰富特性。毕竟，像这样的客户和网络只是少数，但随着时间的推移，会变得越来越多。

有了 IPv6，企业便能够部署满足特定商业需求的应用，然而，在短期内，部署 IPv6 所面临的最大挑战是，未必会得到所有人的支持。其他的挑战还有，在网络、安全性、管理、传输以及应用方面缺乏对 IPv6 端到端的强有力支持。许多大企业都自行开发自己专有的应用程序，在向 IPv6 过渡时，则又会带来运维方面的挑战。上述情形在业界普遍存在，要想实现对纯 IPv6 的全面支持，尚存不小差距。随着设备制造商和运营商不断填补这些差距，纯 IPv6 的部署已呈渐增之势。

图 3-1 所示为纯 IPv6 模型。模型中的主机和网络设备只有从头到尾全都启用 IPv6，才有可能实现互相通信[①]。

① 原文是“The hosts and network devices in this model need to operate in IPv6 end to end for communication between them to be possible”。译者的英文水平也十分有限，帮作者改写一下原文“The hosts and network devices in this model need to operate in IPv6 end to end in order to make it possible for communication each other”，译文按改写后的原文翻译。

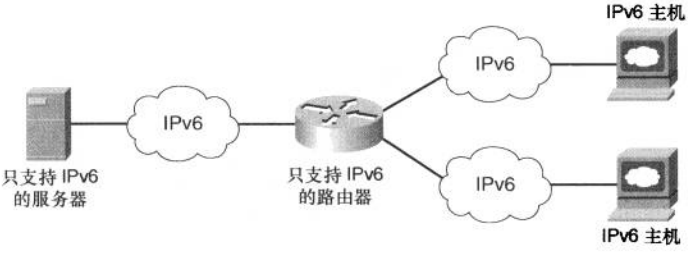


图 3-1 纯 IPv6 拓扑

3.2 过渡机制

顾名思义，过渡机制的宗旨是：帮助网络从一种协议过渡到另一种协议。对 IPv6 来说，“过渡”的基本含义是：把网络中运行的协议从 IPv4 逐渐割接至 IPv6，待到某日，时机成熟，再用 IPv6 完全取代之前所运行的 IPv4。可在短期内，许多过渡机制均要求网络同时开启 IPv4 和 IPv6 这两种协议。下面几节将会讨论某些广泛使用的过渡机制。

3.2.1 双栈机制

双协议栈机制既是首选的 IPv4 到 IPv6 的过渡机制，也是最为基本的过渡机制。由于该机制部署起来既不需要建立隧道，也不必执行（地址、端口方面的）转换，来建立端到端的连通性，因此对主机和网络来说，这也是一种最自然的部署 IPv6 的手段。在部署双栈的过程中，接入网络的所有部件（主机、服务器、路由器、交换机、防火墙等）需要同时运行 IPv6 和 IPv4。

双栈机制已有使用过的先例，其中包括 IPv4 与 IPX 和/或 AppleTalk 在同一节点上共存的例子。与使用其他双协议栈一样，IPv4 与 IPv6 并不彼此兼容。

双协议栈模型不但能实现从 IPv4 到 IPv6 网络环境的平稳过渡，而且还会将业务中断的可能性降至最低。该模型的运作方式是，在现有 IPv4 网络环境中，启用 IPv6 以及与其相关且必不可缺的各种特性，比如，实施 IPv6 路由选择、高可用性以及安全性等诸多特性。

双栈机制的最大优点是，无须在网络中“架设”隧道。该机制以“午夜航船（ships-in-the-night）”的方式来运行 IPv4 和 IPv6 两种协议，这意味着以上两种协议并列运行，除了共享相同的网络资源以外，在运作时，IPv4 和 IPv6 各不相干。无论是 IPv4 还是 IPv6，在路由选择、高可用性（HA）、服务质量（QoS）、

安全以及多播策略方面都“各自为政”。由于在转发数据包时，既不需要额外的封装，也没有任何查表方面的开销，因此与其他机制相比，双栈机制在转发性能方面还颇具优势。

支持双栈协议的主机，既可以配置 IPv4 地址也可以配置 IPv6 地址。在获取 IP 地址时，双协议栈节点会利用 DHCP（动态主机配置协议）这样的 IPv4 和 IPv6 地址分配机制，来获取与各自协议栈有关的配置信息。

图 3-2 所示为双协议栈模型。该模型中的节点和网络设备都要能够处理独立运行的 IPv4 和 IPv6 协议栈。

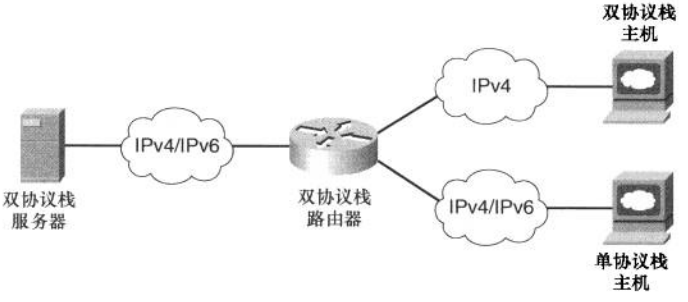


图 3-2 双栈 IPv6 拓扑

3.2.2 IPv6 上的 IPv4

IPv4 在企业网中的部署，可谓占统治性地位。随着 IPv6 的启用，及其在部署方面的增长，某些 IPv6 主机或整个网段可能需要通过 IPv6 来建立端到端的通信机制，但主机或网段之间却又有只运行 IPv4 的网络“加塞”。这在 WAN/分支网络的部署中非常常见，在此类场景中，分支机构和总部站点网络均支持 IPv6，但两者之间的 WAN 却只支持 IPv4。可利用建立在 IPv6 网络之间的 IPv4 隧道，将 IPv6 数据包封装在 IPv4 隧道之内传输，以此来建立端到端的 IPv6 通信机制。

一般而言，都会使用点对点隧道或点对多点隧道，在一种协议之上运载或传输另一种协议的数据包（对于本例，是指利用 IPv4 运载或传输 IPv6 数据包）。多年来，这样的隧道技术屡见不鲜，比如，利用 IP 隧道去传输 Novell IPX/SPX、AppleTalk、SNA 或其他协议的数据。

图 3-3 所示为在 IPv4 隧道报头内封装 IPv6 的基本框架。

隧道的类型不但多种多样，用途也各有不同。某些类型的隧道在使用方面要取决于隧道的终结点（主机、路由器）和终结点的数量，甚至还有可能要取

决于操作系统版本。

隧道类型包括：

- **路由器到路由器**：对于路由器到路由器的隧道，连接到IPv4基础设施的路由器，通过在IPv4报头内封装IPv6数据包的方式，来传输IPv6数据包。
- **主机到路由器**：对于主机到路由器的隧道，IPv4/IPv6主机可以以隧道封装的方式将IPv6数据包发送给IPv4/IPv6边界路由器。边界路由器终结此类隧道，然后再将“原汁原味”的IPv6数据包通过IPv6网络发送给末端主机。
- **主机到主机**：对于主机到主机的隧道，则会建立在两台或多台主机之间。IPv6/IPv4主机在IPv4报头内封装IPv6数据包，以建立隧道的方法来实现彼此通信。

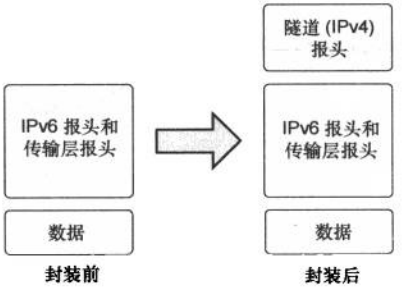


图3-3 IPv4上的IPv6隧道封装

可根据封装节点对隧道端点地址的解析方式，进一步将IPv4上的IPv6隧道机制划分为配置型隧道机制和动态隧道机制。对于手动配置的隧道，隧道端点地址由配置信息决定；而对于动态隧道，隧道端点地址则要以动态的方式来获取。

建立IPv4上的IPv6隧道的方法如下所示。

- 手动配置隧道（MCT）。
- 通过IPv4 GRE封装来传输IPv6数据包。
- 隧道代理（Tunnel broker）。
- 6to4隧道。
- 站点内的自动隧道地址协议（ISATAP）隧道。
- MPLS上的IPv6(6PE)。

本章会简要概括上述方法。具体的配置请参见后续章节。

注意

本章不会介绍像 6rd、双栈 Lite (DS-Lite)以及 Teredo 这样的隧道技术；前两种技术专用于服务提供商网络，而最后一种技术是微软专有且基于主机的技术，通常并不在企业网中使用。

表 3-2 罗列并描述了与上述隧道建立方法有关的缺点[①]。

从下一节开始，会深入讨论每一种 IPv4 上的 IPv6 隧道方法。

3.2.3 手工配置的隧道

手工配置的隧道（MCT）是指以静态方式来定义的隧道。大多数协议栈和路由器都支持 IPv6 MCT[②]，RFC 4213 则为在现有 IPv4 网络中，以手工方式配置 IPv4 上的 IPv6 隧道（传输 IPv6 数据包）提供了理论依据。MCT 也成为了首批为利用现有 IPv4 网络传输 IPv6 数据包而制定的过渡机制之一。

MTC 使用协议号 41 来封装流量[③]，隧道的封装方式则由隧道节点所呈现的静态配置信息来确定。隧道节点既可以是双栈路由器，也可以是双栈主机。对于 MTC 来说，需根据节点的配置信息和路由表来确定感兴趣数据包之类的额外信息。

图 3-4 所示为手工配置的隧道。模型中的末端主机运行 IPv6，而在主机间用来建立隧道的路由器运行双协议栈，隧道则建立于 IPv4 网络之上。

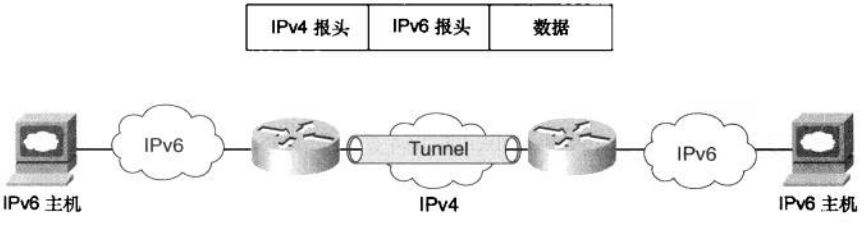


图 3-4 手工配置的隧道

例 3-1 所示为使用 MTC 时的数据包捕获。TCPDUMP 示出了 IPv6 报头

① 原文为表 3-1，估计作者写书时心不在焉。

② 原文是“MCT for IPv6 is supported by most of the stacks and routers”。首先，译者不知道这个 stack 是不是协议栈，如果是，译者想问一下作者“IPv4 上的 IPv6 隧道”究竟还能涉及到几种协议栈？

③ 原文是“MCTs use protocol 41 to encapsulate the traffic”。译者想问一下作者“protocol 41”怎么去封装“traffic”，译文为直译，这里也就不替作者遮丑了。至于这半句话的意思，读者应该“都懂的”。

之前的IPv4报头信息。由数据包的外部报头（IPv4报头）可以看出，IPv4隧道建立于10.1.21.3和10.1.21.1（已做高亮显示）之间，而内部IPv6报头的源和目的IP地址则分别为2000:cafe:2::1和2000:cafe:2::2（已做高亮显示）。

表3-2 IPv4上的IPv6隧道建立方法

方法	描述	缺点（若有）
手工配置的隧道	支持所有实现 基于标准（RFC 4213）	可扩展性不强，要是为了追求站点间的全互连，只要站点数增加，隧道的数量也会随之大幅增加，在运营方面所付出的努力也会呈指数级增长；随着站点数量的增长，管理起来便非常困难[①]
用来传递IPv6流量的IPv4 GRE隧道（IPv6-over-IPv4 GRE tunnel）	使用IPv4 GRE隧道 从隧道配置的角度来看，该方法与手工配置的隧道也没什么不同	与手工配置的隧道相同
隧道代理	基于标准（RFC 3053） 需要专门的隧道代理 动态隧道	隧道代理服务需接受远程配置改变，会带来安全隐患
6to4隧道	基于标准（RFC 3056） 隧道以自动的方式建立 利用IPv4基础设施，比如，虚拟广播链路 使用2002::/16作为IPv6地址前缀	底层的IPv4地址前缀决定了6to4 IPv6地址前缀，因此，要迁移到纯IPv6环境，需重新编址 需要使用公网IPv4地址 在隧道沿途的路径上，不支持NAT 不支持多播
ISATAP	自动覆盖机制（RFC 5214） 利用底层的IPv4作为非广播多路访问（NBMA）链路 易于配置，可对其扩展，以支持大量主机	不支持传输层NAT 出于安全方面的考虑，需对IPv4虚拟链路设定界线 不支持多播
MPLS上的IPv6	利用现有的MPLS基础设施，传输IPv6数据包	虽然比较灵活，取决于所采用的底层机制，配置起来可能会非常麻烦

① 原文是“Difficult to manage as the number of sites increases the operation effort increases exponentially as due to increase in number of sites the number of tunnels increases if fully mesh connections are desired (scalability)”，原文那么多连词（都是as），一个标点符号都没有。

例 3-1 与手工配置的隧道有关的 TCPDUMP 输出

```
root@0[~]# tcpdump -i br0 host 10.1.21.3 -v
tcpdump: listening on br0, link-type EN10MB (Ethernet), capture size
96 bytes
08:58:32.060893 IP (tos 0x0, ttl 255, id 30, offset 0, flags [none],
proto: IPv6 (41), length: 120) 10.1.21.3 > 10.1.21.1: IP6 (hlim 64,
next-header: ICMPv6 (58), length: 60) 2000:cafe:2::1 >
2000:cafe:2::2: ICMP6, echo request, length 60, seq 0
08:58:32.061648 IP (tos 0x0, ttl 255, id 102, offset 0, flags [none],
proto: IPv6 (41), length: 120) 10.1.21.1 > 10.1.21.3: IP6 (hlim 64,
next-header: ICMPv6 (58), length: 60) 2000:cafe:2::2 >
2000:cafe:2::1: ICMP6, echo reply, length 60, seq 0
08:58:32.062063 IP (tos 0x0, ttl 255, id 31, offset 0, flags [none],
proto: IPv6 (41), length: 120) 10.1.21.3 > 10.1.21.1: IP6 (hlim 64,
next-header: ICMPv6 (58), length: 60) 2000:cafe:2::1 >
2000:cafe:2::2: ICMP6, echo request, length 60, seq 1
08:58:32.062437 IP (tos 0x0, ttl 255, id 103, offset 0, flags [none],
proto: IPv6 (41), length: 120) 10.1.21.1 > 10.1.21.3: IP6 (hlim 64,
next-header: ICMPv6 (58), length: 60) 2000:cafe:2::2 >
2000:cafe:2::1: ICMP6, echo reply, length 60, seq 1
```

3.2.4 用来传递 IPv6 流量的 IPv4 GRE 隧道

一直以来，GRE 隧道都用在现有的 IPv4 网络上封装私有地址的 IPv4 数据包，或非 IP 流量（比如 AppleTalk）。对于传统的 GRE 隧道来说，数据包内部报头的 IPv4 地址在建立该 GRE 隧道的网络上是不可路由的。

用来传递 IPv6 流量的 IPv4 GRE 隧道是另一种手工配置隧道的机制，该机制利用点到点的封装方法来提供必要的服务。IPv6 数据包会作为此类 GRE 隧道的荷载。图 3-5 所示为使用 IPv4 GRE 隧道传递 IPv6 流量的网络示例。

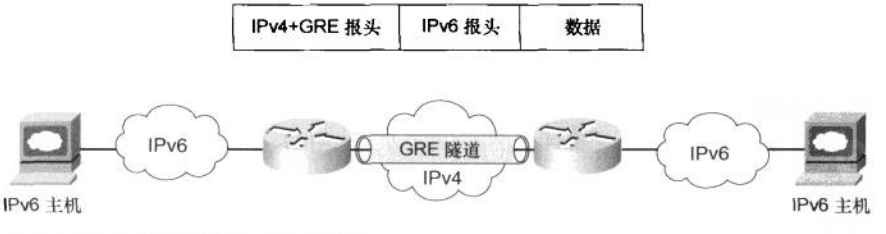


图 3-5 用来传递 IPv6 流量的 IPv4 GRE 隧道

此类隧道要求路由器支持双协议栈，以便在封装前、后能分别处理并路由IPv6、IPv4 数据包。从隧道配置的角度来看，这与手工配置的隧道非常相似，其原因是隧道同样是在一对路由器之间建立。

例3-2所示为与传递IPv6流量的IPv4 GRE隧道相关的数据包捕获。请注意输出中出现在IPv6报头前的IPv4 GRE报头。

例3-2 与传递IPv6流量的IPv4 GRE隧道有关的TCPDUMP的输出

```
root@0[~]# tcpdump -i br0 host 10.1.21.3 -v
tcpdump: listening on br0, link-type EN10MB (Ethernet), capture size
96 bytes
07:21:34.756456 IP (tos 0x0, ttl 255, id 72, offset 0, flags [none],
proto: GRE(47), length: 124) 10.1.21.3 > 10.1.21.1: GREv0, Flags
[none], length: 104) IP6 (hlim 64, next-header: ICMPv6 (58), length:
60) 2000:cafe:2::1 > 2000:cafe:2::2: ICMP6, echo request, length 60,
seq 0
07:21:34.756925 IP (tos 0x0, ttl 255, id 85, offset 0, flags [none],
proto: GRE(47), length: 124) 10.1.21.1 > 10.1.21.3: GREv0, Flags
[none], length: 104) IP6 (hlim 64, next-header: ICMPv6 (58), length:
60) 2000:cafe:2::2 > 2000:cafe:2::1: ICMP6, echo reply, length 60,
seq 0
07:21:34.757319 IP (tos 0x0, ttl 255, id 73, offset 0, flags [none],
proto: GRE(47), length: 124) 10.1.21.3 > 10.1.21.1: GREv0, Flags
[none], length: 104) IP6 (hlim 64, next-header: ICMPv6 (58), length:
60) 2000:cafe:2::1 > 2000:cafe:2::2: ICMP6, echo request, length 60,
seq 1
07:21:34.757581 IP (tos 0x0, ttl 255, id 86, offset 0, flags [none],
proto: GRE(47), length: 124) 10.1.21.1 > 10.1.21.3: GREv0, Flags
[none], length: 104) IP6 (hlim 64, next-header: ICMPv6 (58), length:
60) 2000:cafe:2::2 > 2000:cafe:2::1: ICMP6, echo reply, length 60,
seq 1
```

此类隧道的缺点是，随着参与建立隧道的路由器数据的增加，所要建立的隧道数便会呈几何级数增长，扩展该解决方案可谓是既费时又繁琐。随着站点数的增加，无论是对隧道的管理，还是排除与隧道有关的网络故障都会愈发的困难。

3.2.5 隧道代理

最初，IPv6网络通过手工配置隧道，借用IPv4所提供的传输机制[①]。人们

[①] 译文似乎不通，原文为“Initially IPv6 networks started using the transport facilities provided by IPv4 networks using the manual tunnel configurations”。如果译文不通，原文肯定不通。连续两个“using”，嗯，总算知道作者还会用其他的实意动词了。

也想出了各种各样的方法（比如，利用兼容 IPv4 地址的自动配置隧道和 6to4 隧道），试图以自动化的方式，来解决隧道的手工配置问题[①]。

隧道代理（定义于 RFC 3053 “IPv6 Tunnel Broker”）则是另外一种隧道建立机制，该机制会利用专门的服务器去简化隧道的建立过程，这样的服务器被称为隧道代理服务器。隧道代理服务器负责管理来自端点的隧道建立请求。该解决方案适用的场合为：小型而又孤立的 IPv6 站点（主机）。当今，那些主机彼此隔离的小型 IPv6 站点都使用现成的基于 IPv4 的基础设施来互相连接。启用了隧道代理服务，远程双栈系统上运行的应用程序便能够访问 IPv6 骨干网[②]。

图 3-6 所示为采用隧道代理时的自动隧道建立。

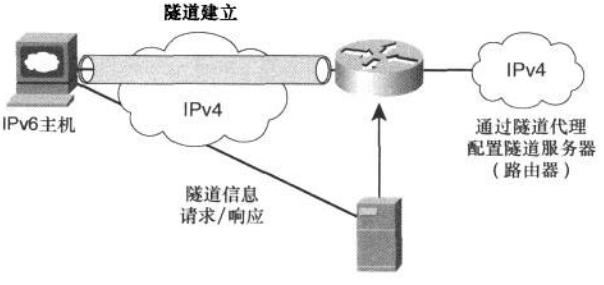


图 3-6 *启用隧道代理时的自动隧道建立*

隧道代理服务的一个最主要的短板是，路由器或隧道代理服务需要接受来自远程服务器的配置变更，这会招来严重的安全隐患。比方说，某企业可在一台专用 Web 站点上，向服务提供商注册一个 IPv4 地址，供自己的路由器使用[③]。服务提供商会交付一个脚本，用来构建通往 IPv6 网络的隧道、为终端系统分配 IPv6 地址，并为该路由器分配网络前缀，以建立与该站点其他网络的连通性。隧道代理服务器会管理通往隧道服务器的隧道的创建和删除[④]。采用该机制，任何配置错误或信息泄露都有可能会为站点带来严重安全隐患。

---

① 原文是 “Various ideas like automatically configured tunnels with IPv4-compatible addresses and 6to4 have tried to solve the manual configuration issues with automation”。没有办法，译者只能再次 “杜撰” 译文。

② 原文是 “Applications on the remote dual-stack systems can access the IPv6 backbone by enabling tunnel broker service”，“应用程序访问 IPv6 骨干网”，听起来怪怪的，但原文如此，译者选择直译。

③ 原文是 “For example, an enterprise could register the IPv4 address of the router with the service provider on a dedicated website”。原文不确定性因素太多，译者可以翻出 5 种意思不同的译文。

④ 原文是 “The tunnel broker manages the creation and deletion of the tunnel to the tunnel server”。原文蹩脚，译文自然也好不到哪儿去。

3.2.6 6to4隧道

6to4隧道（定义于RFC 3056“Connection of IPv6 Domains via IPv4 Clouds”）是一种自动隧道建立机制，通常都是由边界路由器来建立这样的隧道。6to4隧道机制既无须显式地配置，也用不上类似于虚拟广播链路那样的IPv4基础设施。隧道的目的地址是IPv6报头内的IPv6目的地址中内嵌的IPv4地址[①]。

以2002::/16前缀打头的IPv6地址被称为6to4地址，要想构造出48位的6to4前缀，需将“2002”安插在一个为主机或主机身后的网络所使用的IPv4地址之前。试举一例，用公网IP地址192.168.10.1[②]可构造出相应的6to4前缀2002:C0A8:0A01::/48。

该机制虽然也用IPv4数据包来封装IPv6数据包，但所用的IP协议号却是41。6to4隧道机制适用于以下两种部署场景。

- 互连6to4网域。
- 利用中继路由器互连6to4网域和纯IPv6网域。

图3-7所示为互连6to4网域的6to4部署，这也是用来互连多个IPv6站点的最为简单的部署场景。在该部署场景中，只会为路由器的外网接口分配一个6to4地址。

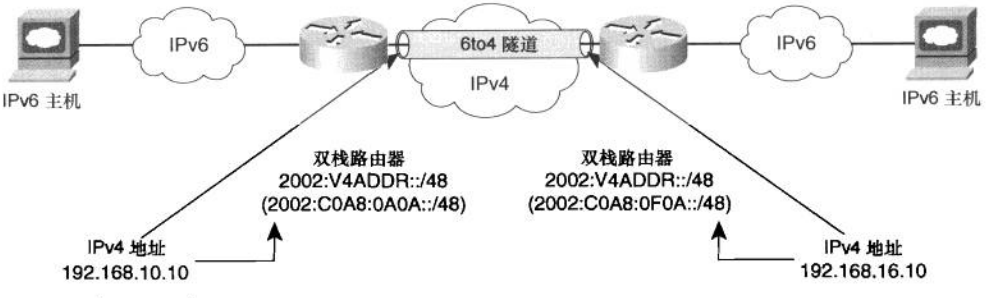


图3-7 互连6to4网域

只有当6to4网域和纯IPv6网域之间有通信需求时，这第二种部署场景才有可能会变得更加普及。这种通信方式会用到中继路由器（需使用同时配置了6to4地址和纯IPv6地址的标准双栈路由器）。图3-8所示的中继路由器将IPv4网络、

① 原文是：“The tunnel destination is the embedded IPv4 address from the IPv6 destination address in the IPv6 header”。不知道有没有译对，作者的文笔相当一般。

② 这个似乎不是公网地址。

纯 IPv6 网络以及 6to4 IP 站点网络连接在了一起。

图 3-8 所示为互连 6to4 网域和纯 IPv6 网域。

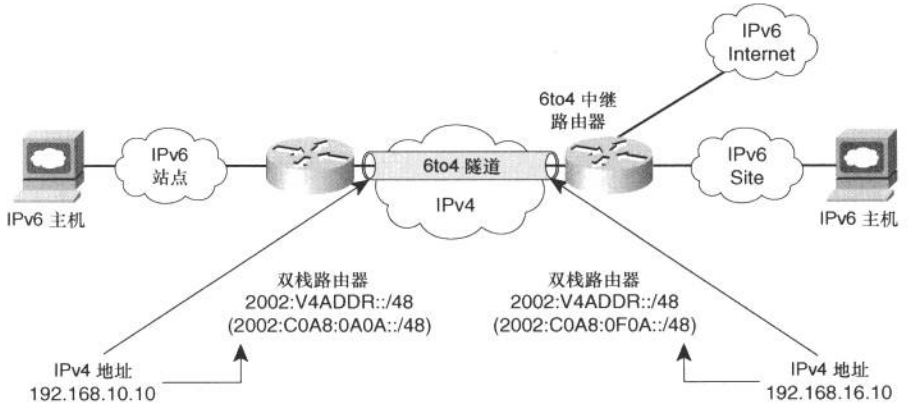


图 3-8 互连 6to4 网域和纯 IPv6 网域

互连 6to4 网域和纯 IPv6 网域时，可考虑采用诸如 EIGRP 之类的 IPv6 内部路由协议，来行使站点内的路由选择功能，以保证路由选择方面的正常运作。为了与纯 IPv6 网域内的设备通信，还需设置一条指向特定中继路由器的默认路由。由于内部路由协议无法跨 6to4 隧道运行，因此对于图 3-8 所示场景，只能采用静态路由或 BGP4+路由协议来完成路由选择。根据规范，6to4 中继路由器只能通告前缀 2002::/16，不能通告其所属的任何明细路由，这是为了避免路由表被明细路由“玷污”，从而能够更好地控制路由表的规模。

可使用对公众开放的 6to4 中继 2002:c058:6301:: (IPv4 地址- 192.88.99.1)，该 6to4 地址是一个离中继路由器最近的特殊的任意播地址。只要有 IPv6 Internet 访问方面的需求，就应该使用这一 6to4 地址。就本质而言，6to4 路由器和 6to4 中继路由器间的那条隧道并不具备任何安全性，故而会存在以下安全隐患。

- 6to4 中继路由器不会检查包含在数据包中的数据[①]。
- 地址欺骗是主要的安全隐患，可以很容易地对 IPv6 源地址进行欺骗。

3.2.7 站点间自动隧道地址协议(ISATAP)

定义于 RFC 5214 的 ISATAP 是一种自动化的覆盖型隧道机制（automatic overlay tunneling mechanism），站点内的 IPv6 主机可利用该机制来彼此通信。

① 原文是“The data contained in the packets is not checked by the 6to4 relay routers”，译文为直译。

ISATAP 将底层的 IPv4 基础设施视为非广播多路访问（NBMA）链路层。

ISATAP 隧道将网络设备（主机）接口的 IPv4 地址插入 IPv6 地址中，作为该地址的最后 32 位[①]。ISATAP 以隧道的方式传输 IPv6 数据报，适合部署于基础设施（底层）不支持 IPv6 协议，且只有为数不多的双栈主机需要建立连通性的站点。为了支持网络客户端的自动配置，ISATAP 路由器为 ISATAP 站点提供网络配置支持[②]。IPv6 客户端可借此自动完成自身的配置。

如图 3-9 所示，ISATAP 地址由三个部分组成。

前 64 位	32 位	32 位
全局 IPv6 或 本地链路前缀	0000:5EFE	接口的 IPv4 地址

图 3-9 ISATAP 地址格式

- 前 64 位为全局 IPv6 或本地链路前缀。
- 中间 32 位为 0000:5EFE（若使用公网单播地址，则为 0200:5EFE）。
- 最低 32 位为接口标识符 IPv4 地址。

图 3-10 所示为利用双栈 IPv6 主机和 ISATAP 路由器创建 ISATAP 隧道。

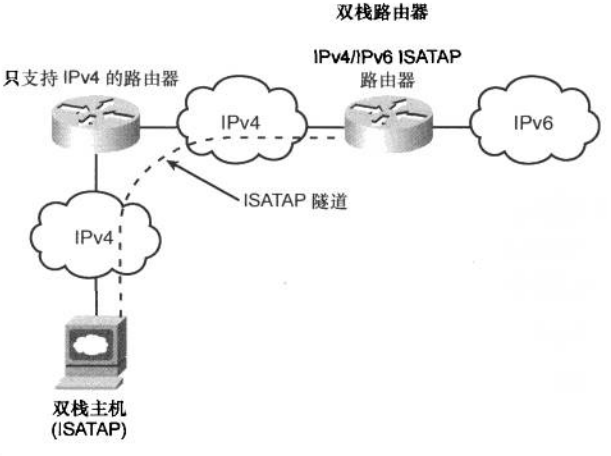


图 3-10 ISATAP 隧道

① 原文是“ISATAP tunnels embed the IPv4 address of the interface in the last 32 bits of the IPv6 address”。

② 原文是“To support automatic configuration to the network clients, the ISATAP routers provide network configuration support for the ISATAP site”。译者不明其意，按字面意思直译。

3.2.8 MPLS 上的 IPv6

服务提供商和大企业在“不急不躁”地改造自有网络基础设施以迎接 IPv6 的同时，还可利用现成的 IPv4 MPLS 网络基础设施来传输 IPv6 数据报。对于这种情况，PE（提供商边缘）路由器必须具备 IPv6 路由选择功能，而 P（服务器提供商）路由器却无须支持该功能。这使得服务提供商在无须升级自有骨干网的情况下，便能提供 IPv6 服务（在孤立的 IPv6 网域之间提供连通性服务）。

接下来，我们将会剖析以下 3 种利用 MPLS 传输 IPv6 数据报的概念。

- 利用建立在 MPLS 上的电路来传输 IPv6 数据报（IPv6 over circuit transport over MPLS）。
- 利用建立在 CE 路由器上的隧道来传输 IPv6 数据报（IPv6 using IPv4 tunnels over customer edge (CE) routers）。
- 利用基于 IPv4 的 MPLS 核心网络来传输 IPv6 数据报（IPv6 MPLS with IPv4-based core）(6PE/6VPE)。

表 3-3 对以上三种方法进行了比较。

表 3-3 MPLS 上的 IPv6

方法	描述	短板（如有）
利用建立在 MPLS 上的电路传输 IPv6 数据报	服务提供商(SP)向客户提供电路(例如，ATM、帧中继)	可扩展性
利用建立在 CE 路由器上的隧道传输 IPv6 数据报	这是一种建立 tunnel-in-tunnel(隧道中的隧道)的方法，要求 CE 路由器支持双协议栈 不波及 MPLS 基础设施 IPv6 数据报被封装 2 次：先封装进 IPv4 数据报；再封装进 MPLS 帧	隧道建立成本
利用基于 IPv4 MPLS 核心网络传输 IPv6 数据报	基于标准: RFC 4659, 6VPE RFC 4798, 6PE 利用现有的 IPv4 MPLS 基础设施提供服务 只会波及 PE 路由器	复杂性 MPLS 核心网络对 IPv6 一无所知 故障排除极为困难

利用建立在 MPLS 上的电路来传输 IPv6 数据报

由于该方法利用的是第二层电路仿真技术（MPLS 上的任意传输[AToM]），因此底层电路对 IPv6 的传输来说完全透明。使用该方法，无须在 PE 或 CE 路

由器上做任何配置变更。

图3-11所示为利用建立在MPLS上的电路传输IPv6数据报。

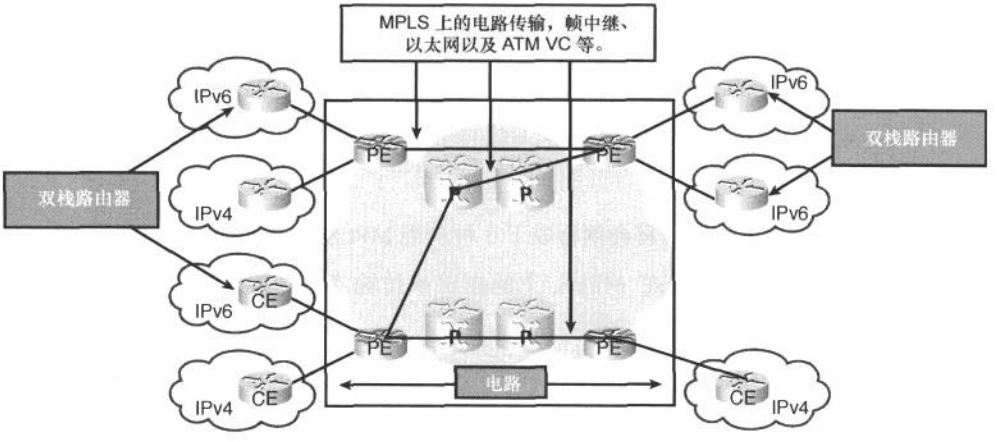


图3-11 利用建立在MPLS上的电路传输IPv6数据报

运营商可利用AToM为客户提供类似于ATM、帧中继以及PPP以太网之类的租用线路和第二层服务仿真。AToM唯一的短板便是可扩展性不强——CE路由器数量的增加，将会导致隧道数的激增。要是客户不担心次优路由选择的话，可采用中心到分支（部分互连）拓扑来规避可扩展性问题。

利用建立在CE路由器上的隧道来传输IPv6数据报

这是一种在支持MPLS的提供商核心网络中，建立tunnel-in-tunnel的方法，要求CE路由器支持双栈功能。站在转发平面的角度来看，要对IPv6流量封装2次：先以IPv4报头封装，然后再封装进MPLS帧。对服务提供商来说，该传输IPv6数据报的方法既不需要服务提供商从运维层面上干预，也不会波及网络基础设施，更无须更改P和PE路由器的配置变更，因此使用起来极为方便。然而，该方法也有其自身可扩展性方面的限制，这是因为一旦得到部署，不但要在CE路由器上手动更改配置，而且还需要在PE路由器之间建立全互连的隧道拓扑。

图3-12所示为在CE路由器上搭建隧道传输IPv6数据报。

基于MPLS的IPv6服务提供商边缘路由器（6PE）

6PE是基于MPLS的IPv6服务提供商边缘路由器的Cisco实现。6PE利用建

立在 IPv4 MPLS 核心网络之上的 MPLS LSP（标签交换路径），让多个 IPv6 站点相互通信。该特性需要在 PE 路由器上运行 MBGP（多协议 BGP），以交换 IPv6 前缀的 IPv6 可达性信息。支持双协议栈的 PE 路由器便可利用 IPv6 可达性信息，去应用相应的标签。由于 MPLS/VPN 架构中数据平面和转发平面的分离，因此便能“另辟蹊径”，在单一网络基础设施之上同时传输 IPv4 和 IPv6 数据报。

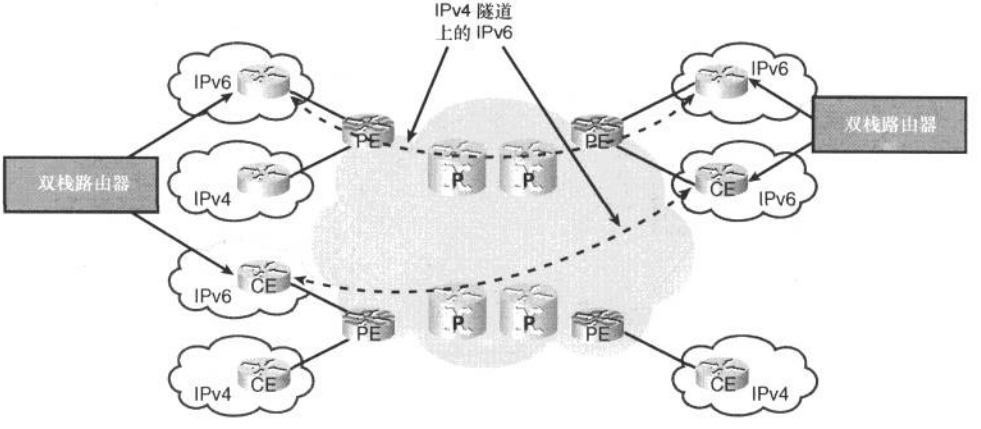


图 3-12 CE 路由器上搭建 IPv4 隧道传输 IPv6 数据报

在入站（ingress）PE 路由器上，流入的数据报附着了 2 个标签。内层标签是 MP-BGP 分配给目的 IPv6 前缀的标签，而入站 PE 路由器会基于外层标签将数据报送达出站 PE 路由器的 IPv4 地址，这台 PE 路由器能够将 IPv6 数据报转发到其最终目的地（提供了 IPv6 目的前缀的可达性）。

如图 3-13 所示，PE 路由器都是双栈路由器，且含有相应的 LDP（标签分发协议）和 MBGP 配置。若需要部署流量工程，还需配置 RSVP（资源预留协议）。所有的 PE 和 P 路由器都运行共用的 IGP（内部网关协议）。

IPv6 VPN 服务提供商边缘路由器（6VPE）

6VPE（定义于RFC 4659“BGP-MPLS IP Virtual Private Network (VPN) Extension for IPv6 VPN”）即 IPv6 VPN。IPv6 VPN 服务与 MPLS IPv4 VPN 服务几近相同。这一过渡机制使得服务提供商可以开展“IPv6 接入”和运营商支持运营商等业务。

由于没有 IPv6 编址方面的限制，Cisco 6VPE 实现具备良好的可扩展性。Cisco 6VPE 实现类似于 IPv4 MPLS VPN，因此该技术可让大企业或服务提供商只需将 PE 路由器的软件升级为支持双协议栈，便能在现有 IPv4 骨干网上开展

IPv6 MPLS VPN业务。

图3-14所示为6VPE架构。

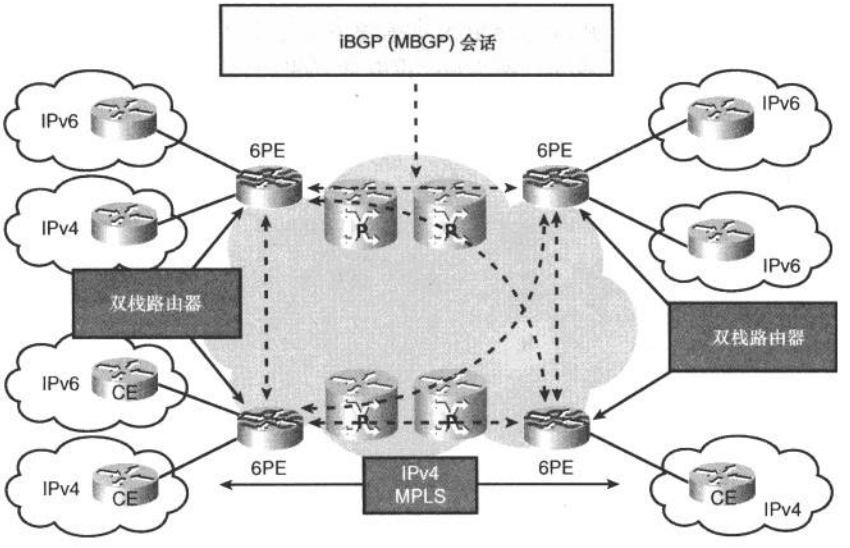


图3-13 6PE

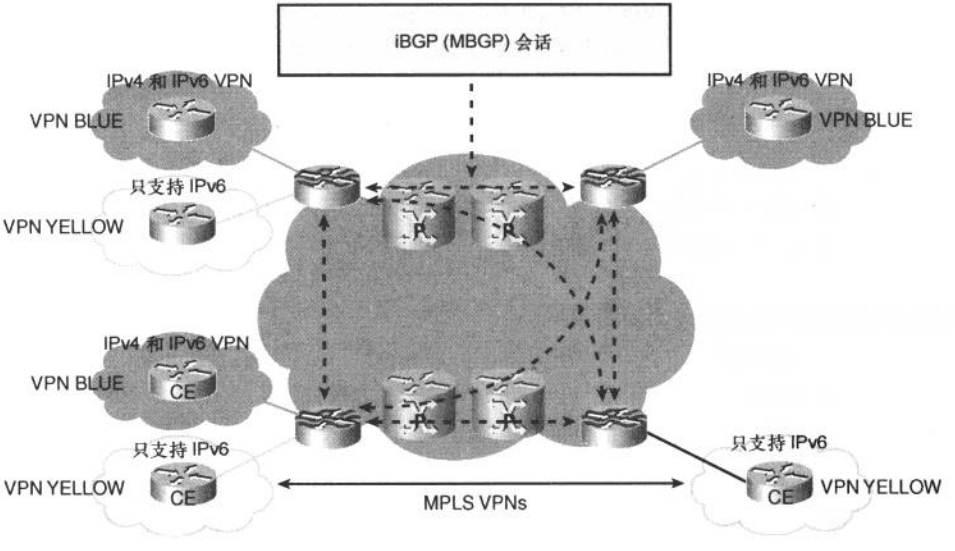


图3-14 6VPE架构

3.3 协议转换和代理机制

我们经常会碰到需要在 IPv4 和 IPv6 之间执行转换或代理的情况。比方说，为了能够让园区/分支机构网络中只支持 IPv6 的主机与数据中心接入层只支持 IPv4 的传统主机彼此通信，则须在数据中心内完成协议转换。

可部署一台中间设备或节点（比如防火墙、路由器或负载均衡器）来执行 IPv4 到 IPv6（或 IPv6 到 IPv4）转换功能。当然，运行在每个端点的操作系统也可以执行转换功能。

以下试举了一些执行 IPv4/IPv6 转换机制的示例。

- 网络地址转换-协议转换（NAT-PT）。
- NAT64。
- TCP-UDP 中继。
- Bump in the stack (BIS)。
- 基于 SOCKS 的 IPv6/IPv4 网关。

本节讨论 NAT-PT 和 NAT64 的原因是，这两种机制最为常用。而上面列出的其他几种机制（TCP-UDP 中继、BIS 以及基于 SOCKS 的 IPv6/IPv4 网关）则不在本章的讨论范围之内。

注意　IETF 已通过 RFC 4966 将 NAT-PT 正式置为“过时”状态，而大多数厂商也并不建议采用这种转换方式。

3.3.1 NAT-PT

NAT-PT 在网络层（第三层）执行 IPv4 和 IPv6 之间的转换。对于该机制来说，主要适用于 IPv6 网络中的端节点试图与 IPv4 网络中的节点通信的场景[①]。

NAT-PT 使用 IPv4 地址池，并会在 IPv4-IPv6 边界路由器上将地址池中的

---

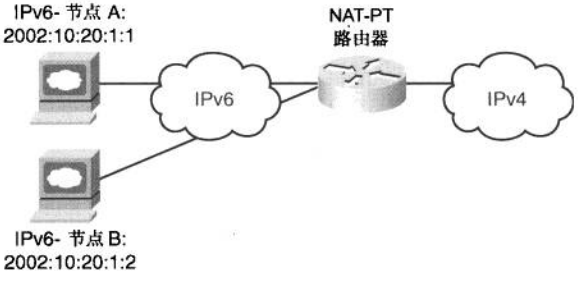
① 原文是“In this mechanism, end nodes in the IPv6 network are trying to communicate with the nodes in the IPv4 network. This method is primarily used for communication between the hosts that are IPv6-only to the ones that are IPv4-only”。译文两句变一句，作者的写法纯粹为了凑字数。

地址分配给IPv6端节点/主机。该机制与当今IPv4网络中所用的NAT机制非常相似。

NAT-PT基于的是无状态IP/ICMP转换(SIIT)算法，定义于RFC 2765（现已被RFC 6145取代）。该算法会在IPv4和IPv6报头之间执行转换，而无须知晓任意一条连接的状态[①]。

与IPv4所用的NAT相似，NAT-PT也支持静态转换和动态地址池两种形式。静态转换是指IPv4和IPv6地址之间进行一对一的映射。利用配置在NAT-PT路由器上IPv4地址和IPv6地址的映射关系，IPv6节点就能够与IPv4节点通信。动态NAT-PT则从地址池中分配多个地址，允许多对多的NAT-PT映射。

图3-15所示为NAT-PT路由器以及相应的地址转换表。



转换前		转换后	
源地址	目的地址	源地址	目的地址
2002:10:20:1:1	2002:CAFO:1:1	192.168.1.1	10.12.1.1
2002:10:20:1:2	2002:CAFO:1:2	192.168.1.2	10.12.1.2

图3-15 NAT-PT示例

该解决方案的短板与IPv4 NAT机制也几乎一模一样，如下所示。

- 不支持非对称路由，理由很简单，与会话相对应的回馈流量若不从同一台NAT-PT设备返回，便会被丢弃。

① 原文是“This algorithm translates between the IPv4 and IPv6 packet headers without requiring any per-connection state”。译文为直译。

- NAT-PT 设备在转换嵌入的地址时，需识别底层的应用或协议[①]。

3.3.2 NAT64

顾名思义，NAT64 转换机制是指将 IPv6 数据报转换为 IPv4 数据报。在 NAT64 网络环境中，数据报的发起者总是在 IPv6 端。虽然 NAT64 与其他 NAT 机制一样，也有着某些相同的短板，但却是最好的一种协议转换机制，这是因为 NAT64 是人们根据 IPv4 NAT 的多年使用经验雕琢而成，并修补了如 NAT-PT 等其他机制所具有的缺陷。NAT64 还提供了诸多附加特性，比如，NAT 映射、过滤以及 TCP 的同时打开（要求对等到对等的网络环境）等特性。

此外，NAT64 还提供了“发夹（hairpinning）”特性，该特性可让 NAT64 设备后的 IPv6 主机彼此通信。图 3-16 所示为包含 NAT64 设备、DNS64、IPv6 客户端以及 IPv4 服务器的网络设计。

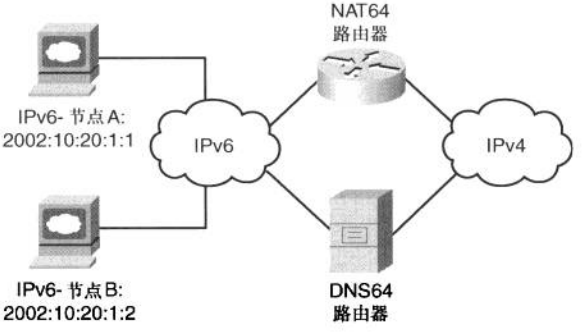


图 3-16 NAT64

3.4 总结

本章讨论了各种 IPv6 过渡机制，视如今底层网络基础设施的规模以及对 IPv4 的支持情况，可利用这些机制为客户网络从 IPv4 到 IPv6 的过渡做好准备。本章旨在简要介绍某些比较流行且已得到诸多厂家支持的过渡机制。后面的章节，比如第 6 章，将会详细讨论为何以及如何部署具体的过渡机制。

---

① 原文是“Any embedded address translation needs the knowledge of the underlying application/protocol”。译者也不知有没说透作者的意思，现给出原文。

3.5 参考资料

Carpenter, B. and K. Moore. RFC 3056, "Connection of IPv6 Domains via IPv4 Clouds." http://www.ietf.org/rfc/rfc3056.txt.

Cisco. Implementing IPv6 over MPLS: http://www.cisco.com/en/US/docs/ios/ipv6/configuration/guide/ip6-over_mpls.html.

Cisco. Implementing Tunnels for IPv6: http://www.cisco.com/en/US/docs/ios/ipv6/configuration/guide/ip6-tunnel.html#wp1055566.

Cisco. IPv6 over MPLS (Cisco 6PE): http://www.cisco.com/warp/public/cc/pd/iosw/prodlit/iosip_an.pdf.

Durand, A., P. Fasano, I. Guardini, and D. Lento. RFC 3053, "IPv6 Tunnel Broker." http://www.ietf.org/rfc/rfc3053.txt.

Carpenter, B. and K. Moore. RFC 3056, "Connection of IPv6 Domains via IPv4 Clouds." http://www.ietf.org/rfc/rfc3056.txt.

Nordmark, E. RFC 2765, "Stateless IP/ICMP Translation Algorithm (SIIT)." http://www.ietf.org/rfc/rfc2765.txt.

Gilligan, R. and E. Nordmark. RFC 2893, "Transition Mechanisms for IPv6 Hosts and Routers." http://www.ietf.org/rfc/rfc2893.

Tsirtsis, G. and P. Srisuresh. RFC 2766, "Network Address Translation - Protocol Translation (NAT-PT)." http://www.ietf.org/rfc/rfc2766.txt.

Nordmark, E. and R. Gilligan. RFC 4213, "Basic Transition Mechanisms for IPv6 Hosts and Routers." http://www.ietf.org/rfc/rfc4213.

De Clercq, J., D. Ooms, M. Carugi, and F. Le Faucheur. RFC 4659, "BGP-MPLS IP Virtual Private Network (VPN) Extension for IPv6 VPN." http://www.ietf.org/rfc/rfc4659.txt.

Templin, F., T. Gleeson, and D. Thaler. RFC 5214, "Intra-Site Automatic Tunnel Addressing Protocol (ISATAP)." http://www.ietf.org/rfc/rfc5214.txt.

Popoviciu, Ciprian P., Eric Levy-Abegnoli, and Patrick Grossetete. *Deploying IPv6 Networks*. Cisco Press. ISBN10: 1-58705-210-5; ISBN13: 978-1-58705-210-1.

CHAPTER 4

第 4 章　网络服务

本章涵盖以下主题。

- **多播**：本节会介绍如何在 IPv6 网络中配置多播。通过对多播寻址的讨论，来说明 IPv6 编址格式的高效性[①]。本节还会讨论各种 IPv6 多播路由协议。
- **服务质量（QoS）**：本节将侧重讨论 IPv4 与 IPv6 QoS 实现之间的主要差异，然后会介绍 IPv6 扩展报头[②]。
- **IPv6 路由选择**：本节将讨论几种使用最为广泛的路由协议，并会介绍这些协议如何在 IPv6 网络环境中运作。

如今，要把公司员工、客户和商业伙伴们紧密联系在一起，企业网络除了要完成数据传输功能之外，还要能提供其他各种不同服务[③]。对于当今的企业客户来说，要想在降低整体成本的同时，提高生产率，以下网络服务必不可缺[④]。

- 多播。
- 服务质量（QoS）。
- IPv6 路由选择。

本章会对以上三种网络服务做一般性介绍。Cisco Press 图书《Deploying IPv6 Networks》（Ciprian Popoviciu、Eric Levy-Abegnoli 和 Patrick Grossetete 著）则对

① 原文是“Multicast addressing introduces the effectiveness of the IPv6 addressing format”。但愿译文反映出了作者试图表达的意思。

② 原文是“then introduces extension headers”。这个“extension headers”到底是什么？

③ 原文是：“Enterprise networks today require various network services in addition to data transmission to bring employees, customers, and business partners together”。译文为译者猜测，译者已经不信任作者的文字了。

④ 难道开启多播、QoS 以及 IPv6 路由选择就能够提高生产率、降低成本？这种说法未免太过牵强。

以上三种网络服务做了深入探讨。

4.1 多播

就多播的传输机制而言，多播源会将每个多播数据包的一份拷贝发送到一个特殊的地址，对多播数据包感兴趣的多个接收者，会“借用”该地址来接收数据包。多播源和多播接收者都是指定多播组的成员，可遍布于网络的四面八方。支持多播转发的网络设备会将数据包的单份拷贝复制给多个接收者，而无须让每个接收者通过专门的单播连接从多播源接收数据。利用多播来传输视频流量，既可以降低整个网络的流量负载，也可以减小因复制不必要的通信流量而影响视频服务器的性能。现在，先举几个利用多播的应用程序示例，这包括视频会议、远程教学、软件发布、股票报价和新闻等。

图 4-1 所示为单条多播流，这条多播流由一 IP 摄像头生成，发往两台或多台媒体服务器，由服务器对视频信息存档。该 IP 摄像头便是多播源，媒体服务器则是多播接收者。参与多播传输的网络组件（比如，交换机和路由器等）会对多播数据包进行复制。多播数据包的复制只会发生在路由器或交换机上参与多播转发的多个出站接口处。要是使用单播来完成图 4-1 所示的应用，IP 摄像头则会分别为三台媒体服务器各生成一条单播数据流，这就造成了计算和网络资源的浪费。

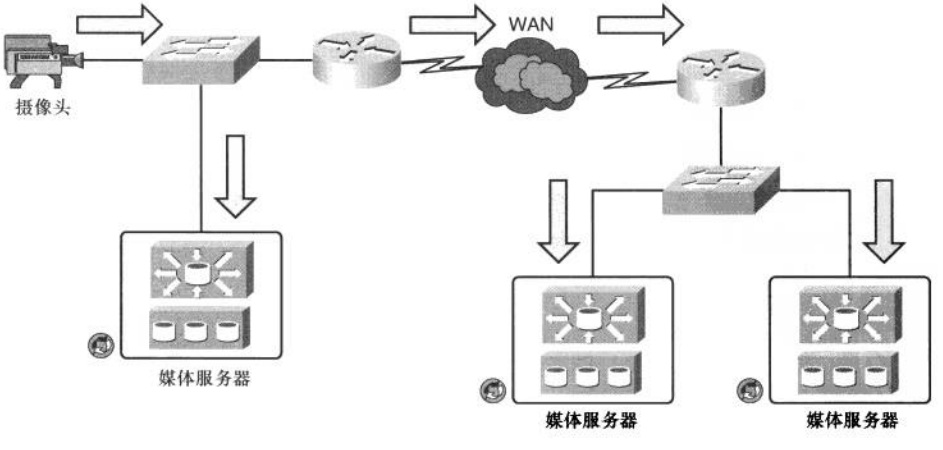


图 4-1 用作视频监控的多播

多播接收者可以是一个或多个多播组的成员，接收多播流量之前，需显式加入相应的多播组。由于多播流量基于UDP（用户数据报协议），而UDP与TCP不同，并没有内置诸如流控或错误恢复等可靠性机制，因此需利用QoS之类的工具来改善多播传输的可靠性[①]。

本章只是对多播技术做简要概括。欲了解IP多播技术的详细信息，请访问Cisco站点上的“Cisco IP Multicast Resources”，链接为http://www.cisco.com/en/US/products/ps6552/products_ios_technology_home.html。

4.1.1 IPv6多播编址

IPv6多播地址格式定义于RFC 4291。IPv6多播地址由8位地址、4位标志、4位范围以及112位组地址字段组成，如图4-2所示。

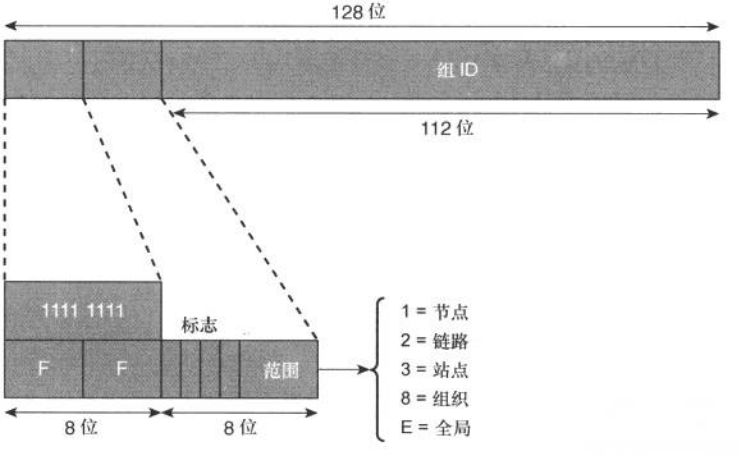


图4-2 IPv6多播地址格式

标志字段用来区分不同的地址类型。定于RFC 4291的T位用来标识多播地址是“永久”（置0）还是“临时”（置1）。定义于RFC 3306的P位提供了一种从IPv6单播地址“推算出”IPv6多播地址的手段。图4-3演示了RFC 3306

[①] 原文中，本段的末尾还有两句“Some edge devices can communicate with the media server using unicast or multicast communications. The use of multicast offers some benefits when a video stream is to be archived by several media servers because only a single stream is required from the IP camera or encoder”。译文为“图中的边缘设备既可以利用单播亦可以利用多播与媒体服务器通信。当多台服务器需要对视频信息存档时，采用多播传输机制是一种更好的选择，其原因是只需从IP摄像头或编码器发送单条数据流。”译者认为，这两句放在这里用处不大，可以删除。

地址的构成方式。修改后的标志字段又令 IPv6 多播地址新增了两个字段：前缀长度（Plen）和网络前缀字段，其中后者为全局分配的 IPv6 单播前缀。这既有助于人们识别出已分配的 IPv6 多播地址，又便于人们管理 IPv6 多播地址分配。

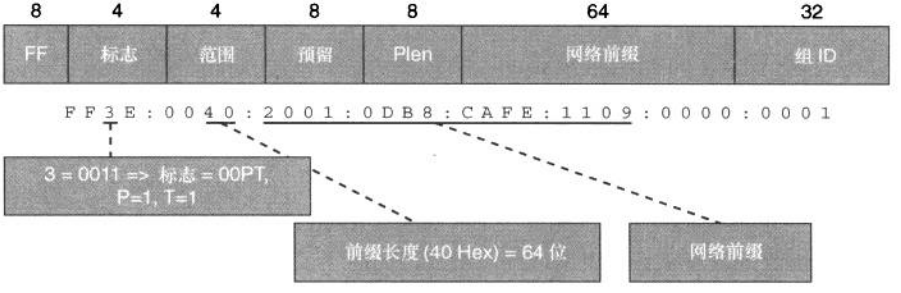


图 4-3 基于单播前缀的 IPv6 多播地址

为了更合理地去分配 IPv6 多播地址，RFC 3956 提出了一种将 RP（集合点）地址嵌入 IPv6 多播地址的方法。使用此类 IPv6 多播地址时，只有 RP 地址相同，才会有多播地址冲突的可能性。可通过将标志字段中的 R、P、T 位置 1 来解决这一问题。将标志字段配置为 0111 即意味着预留字段从 8 位骤减为 4 位——预留字段中的低 4 位被“征用”为 RPAdd 字段。可利用以下两个步骤从多播地址中获悉 RP 地址，如图 4-4 所示。

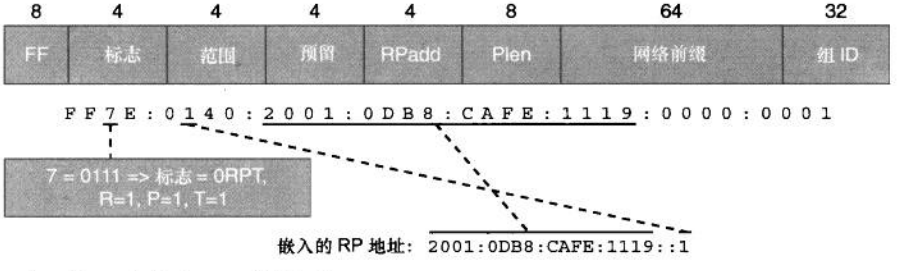


图 4-4 嵌入了 RP 地址的 IPv6 多播地址

步骤 1 将网络前缀字段的头若干位复制到一个新 128 位的 IPv6 地址（RP 地址）中，复制的具体位数由 Plen 字段定义[①]。

[①] 原文是“Copy the first number of bits as defined in the Plen field from the Network Prefix field to a new 128-bit IPv6 address”。请问读者，从字面上看，能看出这层意思吗？

步骤 2 以 RPAdd 字段值作为该 RP IPv6 地址的最后 4 位[①]。

某些 IPv6 多播地址被 Internet 编号分配机构（IANA）预留，可在 IANA 网站（http://www.iana.org）获取这些地址的最新列表。表 4-1 列出部分预留的 IPv6 多播地址。

表 4-1 IANA 登记在案的预留 IPv6 多播地址列表

地址	描述
ff02::1	本地网络中的所有节点
ff02::2	本地网络中的所有路由器
ff02::9	运行 RIP 的路由器
ff02::a	运行 EIGRP 的路由器
ff02::d	运行 PIM 的路由器
ff02::e	RSVP
ff02::1:2	本地网络中的 DHCP 服务器和中继代理
ff02::1:3	本地链路多播名字解析
ff05::1	本地网络站点中的所有节点
ff05::1:3	本地网络站点中的所有 DHCP 服务器
ff0x::fb	多播 DNS
ff0x::108	NIS
ff0x::114	实验

4.1.2 IPv6 多播侦听者发现（MLD）

在 IPv6 网络中，当某多播接收者告知路由器，自己欲从某特定多播组接收数据时，并不会使用 IPv4 网络中所用的 IGMP 协议，而是会用到 MLD 协议，这也是 ICMPv6 的附属协议。表 4-2 列出了 MLD 消息的类型。

[①] 译者再来解释一下 RP 地址从何而来，以图 4-4 为例。若多播地址为 FF7E:0140:2001:0DB8:CAFE:1119:0000:0001，其网络前缀字段值为 2001:0DB8:CAFE:1119，Plen 字段值为 0x40（10 进制值为 64），RPAdd 字段值为 1。按照作者所说的步骤 1，将“2001:0DB8:CAFE:1119”的前 64 位值复制进一个新 128 位的 IPv6 地址，得到“2001:0DB8:CAFE:1119::0”。再按步骤 2，将 RPAdd 字段值作为上面那个 IPv6 地址的最后 4 位，从而得到了 RP 的 IPv6 地址 2001:0DB8:CAFE:1119::1。

表 4-2 MLD 消息类型

消息类型	描述	ICMP 代码
查询	一般情况下，用于特定组或特定多播地址的查询[①]当路由器发出常规查询时，查询消息中的多播地址字段值被设置为 0。通过常规查询，路由器会获悉与己相连的链路上哪个多播地址拥有多播侦听者	130
报告	在报告消息中，发送此消息的多播侦听者会将多播地址字段值设置为自己所要侦听的特定 IPv6 多播地址	131
完成	在完成消息中，发送此消息的多播侦听者会将多播地址字段值设置为自己不再侦听的特定 IPv6 多播地址	132

MLD 利用 ICMP 来传递其消息。所有 MLD 消息都只能在本地链路范围内传递，IPv6 报头中的跳限制字段值总是为 1，而路由器告警选项也总被置位。路由器告警选项被置位，则意味着 MLD 消息附着有逐跳选项报头（hop-by-hop option header）。

MLD 报告消息必须封装在源地址为有效的本地链路地址或未指定地址的 IP 报头中发送，后一种情况是指发送 MLD 报告消息的设备接口尚未获取有效的 MLD 报告消息。允许 MLD 报告消息的源地址为未指定地址，其意在让邻居发现协议的消息传递机制支持 IPv6 多播。

常规查询消息（ICMPv6 类型 130）由 MLD 查询者（路由器）发往所有节点多播组地址（FF02::1）。主机/多播接收者随后会回复报告消息（ICMPv6 类型 131），告知自己感兴趣的多播组地址。上述两种消息 IP 报头中的跳限制字段值均为 1，故而路由器绝不会转发此类数据包。若多播接收者欲接收多播流量，应立刻发送报告消息，而不是坐等收到下一条周期性的查询消息之后，再做回应。

当一个多播接收者不再对某个多播组感兴趣时，则会向所有路由器多播地址（all-routers multicast group）（FF02::2）发送完成消息（CMPv6 类型 132）。路由器随后会发出特定组查询消息，检查该多播组是否还有其他的多播接收者。特定组查询消息 IP 报头的目的地址为上述多播组地址。如无主机响应特定组查询消息，路由器则会认为该组中不再有感兴趣的多播接收者。

定义于 RFC 3810 的 MLDv2 还具备多播源地址过滤功能。有了多播源地址

[①] 原文是“General, group-specific, and multicast-address-specific”。

过滤功能，多播接收者便可以告知路由器：自己要从特定的多播源（多播源地址已事先定义）接收多播流量了[①]。

此外，多播源地址过滤功能还可以让多播接收者选择从除特定多播源（多播源地址已事先定义）以外的所有多播源接收流量。MLDv2 被设计为能够与 MLDv1 互操作。

MLD 功能一般都会配置在网络（园区网、分支、数据中心网络区块）的接入层设备上。网络设计方案和二三层边界的布局则会决定在何处配置 MLD。

欲深入了解 MLD，请参阅本章“参考资料”一节中所列出的 RFC。与 MLD 配置有关的信息请见 Cisco Web 站点，链接为 http://www.cisco.com/go/ipv6。

4.1.3 多播路由：PIM

IP PIM 是业界事实上的标准，用来构建多播分发树。在大多数情况下，路由器会利用由 PIM 所获悉的信息，在多播路由表中安装共享树（*，G）和最短路径树（S，G）条目。

可使用以下命令来启用路由器的 IPv6 多播路由选择功能。

```
ipv6 multicast-routing
```

以下列出了 PIM 的几个主要版本。

- **PIM 稀疏模式（PIM-SM）**：PIM-SM 适用于一对多或多对多的应用，这样的应用是指一个或多个多播源向同一个多播组发送数据，典型的应用包括视频会议和对等到对等的游戏等。通过发送请求消息，路由器会明确告知 RP，自己要加入或离开相应的多播组。下游路由器除非向 RP 发出了加入消息，否则不会接收并转发相应的多播流量。RP 肩负着在不同的多播源和多播接收者之间转发多播数据的任务，并同时扮演着共享多播分发树树根的角色。
- **PIM 特定源多播（PIM-SSM）**：PIM-SSM 是 PIM-SM 的子集，采用这种多播流量转发方法时，交付给多播接收者的数据包仅来自于其所请求的特定源地址。典型的应用为内容交付，比如视频或音频节目（包括 IPTV）。

---

① 原文是“Source filtering provides the receiver with the option to report interest in listening to packets from a defined group of specific source addresses”。

- **双向 PIM（PIM-Bidir）**：开发 PIM-Bidir 是为了帮助部署新近崛起的与通信或与金融有关的应用程序，此类应用程序都依赖于多对多的应用模型。在部署了 hoot-n-holler 应用的网络环境中，PIM-Bidir 是推荐使用的多播路由协议，可允许任意一台主机向其所隶属的多播组发送消息。

注意 尚无 IPv6 的 PIM 密集模式（DM）实现。

以下各节将描述上述每一种控制平面协议。

PIM-SM

PIM-SM IPv6 实现的运作方式与其 IPv4 实现几近相同。PIM-SM 会构建共享树，并要求每个多播组都必须至少使用一个 RP。由于不同的 RP 可供不同的多播组使用，因此还需在 PIM 中引入一个补充机制，能够让路由器获悉：特定的多播组所使用的 RP。无论是在一个 PIM 域内，还是在不同的 PIM 域之间，这样的关联机制都必不可缺。此外，PIM 还会构造出最短路径树（SPT），以供默认情况下的流量转发使用[①]。

只有 IPv6 PIM 提供对嵌入 RP 的支持。这可让路由器利用多播数据包的目的地址，来获悉 RP 地址信息。路由器要想成为 RP，必须通过静态配置。其他路由器会从 MLD 报告消息/PIM 消息，以及多播数据包的目的地址中，获悉 RP 地址和与之相对应的多播组（地址）。然后，路由器便可以利用 RP，针对相关的多播组，执行全套 PIM 操作[②]。

启用 PIM-SM 时，可在路由器上使用以下命令静态配置 RP 的地址。

```
ipv6 pim rp-address 2001:DB8:CAFE:1002::1
```

RP 地址配毕之后，路由器会自动为其创建一个隧道接口，并会利用该虚拟单向隧道接口向 RP 发送注册消息。例 4-1 所示为 IPv6 PIM-SM 的配置和虚拟隧道的相关信息。

---

① 原文是“PIM also builds a Shortest Path Tree (SPT) that is used by default for traffic forwarding”，PIM-SM 默认情况下使用 SPT 转发流量吗，如果是这样的话，要 RP 何用？

② 以上两句的原文是“Routers learn the RP for the group from the group addresses in MLD reports or PIM messages and data packets. It can then use this RP for all PIM activity for the group”。后一句的主语“it”，译者实在不晓得指代什么，译文只能勉强杜撰。

例 4-1 IPv6 PIM-SM 配置和虚拟隧道相关信息

```
R1(config)# ipv6 multicast-routing
R1(config)# ipv6 pim rp-address 2001:db8:cafe:1002::1
R1(config)#


00:05:17: %LINEPROTO-5-UPDOWN: Line protocol on Interface Tunnel1, changed state
to up
R1# show ipv6 pim tunnel
Tunnel0*
  Type   : PIM Encap
  RP     : Embedded RP Tunnel
  Source: 2001:DB8:CAFE:1001::1
Tunnel1*
  Type   : PIM Encap
  RP     : 2001:DB8:CAFE:1002::1*
  Source: 2001:DB8:CAFE:1001::1
R1# show int Tunnel1
Tunnel1 is up, line protocol is up
  Hardware is Tunnel
  MTU 1466 bytes, BW 100 Kbit/sec, DLY 50000 usec,
     reliability 255/255, txload 1/255, rxload 1/255
  Encapsulation TUNNEL, loopback not set
  Keepalive not set
  Tunnel source 2001:DB8:CAFE:1001::1 (Ethernet0/0), destination
  2001:DB8:CAFE:1002::1
   Tunnel Subblocks:
      src-track:
         Tunnel1 source tracking subblock associated with Ethernet0/0
          Set of tunnels with source Ethernet0/0, 2 members (includes iterators),
          on interface <OK>
  Tunnel protocol/transport PIM/IPv6

  Tunnel TOS/Traffic Class 0xE0,  Tunnel TTL 65

<Output omitted for brevity>
```

当网络被划分为多个管理域时，实际则处于域间场景之下。基于 IPv4 的域间多播路由机制，使得每个 PIM 域都能管理自己的 RP。于是，人们开发出了多播源发现协议（MSDP），用在 PIM 域之间交换与多播源有关的信息。只要有组 G 的接收者，RP 便会向其他域中的多播源发送 PIM（S，G）注册消息。对许多应用来说，MSDP 的运作并不具备可扩展性，因此，也不为 IPv6 所支持。如今，

IPv6 的域间多播功能要靠 PIM-SSM 来完成了，如下一节所述。

PIM-SSM（PIM 特定源多播）

PIM-SSM 是 PIM-SM 的一个子集。对于 PIM-SSM 网络环境，多播侦听者会对自己欲加入的组和源（S，G）排定“座次”[①]。PIM-SSM 运作方式与 PIM-SM 几乎相同，但前者不会构建共享树，因此也就无需 RP。多播侦听者必须向 DR 表明，自己对哪个（S，G）感兴趣。正因如此，对 PIM-SSM 部署来说，则必须支持 MLDv2 多播侦听者或 SSM MLDv1 映射路由器[②]。

在 PIM-SSM 网络环境中，由于只会构造出（S，G）（源分发树），因此相较于 PIM-SM，PIM-SSM 在管理和部署方面要简单得多。而且，也不用其他协议来帮助管理 RP。但话又说回来，要想管理好多播源（地址）信息向多播侦听者的发布却并不容易。独立于 PIM 的应用层协议需帮助主机自动发现特定多播组的多播源（地址）。因此，除映射 MLD（将在后续章节中讨论）之外，无需针对 PIM-SSM 做额外配置。

SSM 服务模型要求主机同时指明自己所要加入的多播组，以及所要侦听的特定多播源。但对主机来说，只有 MLDv2 才能支持该功能。尽管 SSM 是一种非常受欢迎的部署模型，但在写作本书之际，在 IPv6 协议栈中实现 MLDv2 的主机并不多见，因此必须要拿出一种解决方案，能让 SSM 与 MLDv1 协同工作。该解决方案被称为 MLDv1 的 SSM 映射，并按如下两种模式运作。

- **静态映射**：在路由器上，以静态的方式为多播源（S）和特定的多播组建立映射关系。根据配置的映射信息，路由器会将任意的（*，G）MLDv1 报告消息与（S，G）关联起来。在路由器上，该映射特性为默认禁用。可使用全局配置命令 **ipv6 mld ssm-map enable**，来默认启用该特性。请使用以下命令去配置静态映射。

```
ipv6 mld ssm-map static ACL1 source 2001:DB8:CAFE:2001::10
ipv6 access-list ACL1
permit any ff08::1/64
```

访问控制列表用来标识映射至多播源（2001:DB8:CAFE:2001::10）的多播组地址。

---

① 原文是“In this case, the listener knows a priority of both the group and the source (S,G) it wants to join”。

② 原文是“For this reason,MLDv2 listener or the SSM MLDv1 mapping router support is required for a PIM-SSM deployment”。译文为直译。

- **动态映射**：在 DNS 服务器上，为 G（多播组）配置一条 AAAA 记录。当路由器收到与（*，G）有关的 MLDv1 报告消息时，会针对 G 记录做反向 DNS 查询。DNS 服务器会返回与 G 相对应的 S（单播地址）。如前所述，在路由器全局配置模式下激活了 SSM 映射功能之后，可使用以下命令来配置动态映射。

```
ipv6 multicast-routing
ipv6 mld ssm-map enable
ipv6 mld ssm-map static IPv6_SSM_MAP 2001:0DB8:1::1
ipv6 mld ssm-map query dns
```

  要想利用动态映射，在路由器上还须事先配置 DNS 服务器地址。在双栈网络中，路由器必须能够通过 IPv4 或 IPv6 访问到 DNS 服务器。

Cisco 路由器所支持的 SSM 映射特性，使得只支持 MLDv1 的 IPv6 主机能够接收到基于 SSM 的多播流量。

双向 PIM（PIM-Bidir）

在拥有大量多播源和多播接收者的网络中，双向 PIM 的效率更高。双向 PIM 与 PIM-SM 的差异在于：靠近多播源的路由器在转发多播数据包时，会顺着多播树先将数据包回发至 RP，再由 RP 沿着共享树转发。如此一来，便省却了创建 STP 和向多播源注册的过程，

对于 IPv6，Cisco 只支持双向 PIM 的静态配置，如下所示。

```
ipv6 pim rp-address 2001:DB8:CAFE:2001::20 bidir
```

本节简要概括了多播技术，并讨论了部署 IPv6 多播的各种方法。要深入了解 IPv6 多播路由技术，请访问 Cisco 站点，链接为 http://www.cisco.com/go/ipv6。

4.2 服务质量

如今的企业网络大都依靠全球性的基础设施来通信。这一全球性网络也是一个用来支持绝大多数应用和服务的平台。企业所提供的绝大多数应用和服务都以这样的网络为平台。在这样一个网络平台上，为不同应用数据流区分优先级，并控制诸如带宽、延迟、数据包之间的延迟变化（抖动）以及丢包之类的要素，便成为了 QoS 所兼具的任务。

以下几节将会讨论 IPv4 和 IPv6 QoS 之间的差异、IPv6 扩展报头，以及 IPv4 和 IPv6 的共存。这可以让读者更好地去理解两种协议的 QoS 实现，对规划 IPv6

迁移方案非常有用。

4.2.1 IPv4 和 IPv6 QoS 之间的差异

IPv4 和 IPv6 QoS 之间的差异大部分都出现在流量分类的过程方面，流量分类是指利用各种参数（比如 IP 源地址、IP 目的地址、DSCP 或 IP 优先级值，以及其他高层协议类型等）来区分出数据包或数据流。分类完毕之后，网络设备可根据 QoS 策略对数据包进行处理，而 QoS 策略可以反映出与数据包相对应业务的“贵贱程度”。表 4-3 总结了 IPv4 和 IPv6 QoS 机制方面所存在的差异。

表 4-3 IPv4 和 IPv6 QoS 机制方面所存在的差异

Qos 机制	实现	IPv4	IPv6
分类	优先级	Y	Y
	DSCP	Y	Y
标记 监管和整形	根据类别标记	Y	Y
	承诺访问速率	Y	Y
	策略路由	Y	Y
	速率限制	Y	Y
	根据类别监管	Y	Y
	通用流量整形	Y	N
	帧中继流量整形	Y	Y
拥塞避免	加权随机早期检测	Y	Y
拥塞管理	先进先出	Y	Y
	优先级队列	Y	Y（不支持传统方法）
	定制队列	Y	N
	低延迟队列	Y	Y

IPv4 报头中的服务类型（ToS）字段也被“照搬”进了 IPv6 报头中的流标签字段，而且用法也一模一样。然而，还需考虑额外的几个分类器，它们均与 IPv6 数据包的头格式相关。

- **协议类型或版本**：由于人们已经预计到了 IPv4 和 IPv6 协议的共存情况，因此还必须考虑为 IPv4 和 IPv6 流量分配不同的服务级别这种情形。可使用协议类型字段来区分两种协议[①]。此外，还可根据流量的各种协议

① 原文是“The Protocol Type field can distinguish between the two protocols”。在译者看来，协议类型字段是用来区分上层（传输层）协议的。据译者推测，应该是“可使用版本号字段，或数据帧中的以太类型字段来区分两种协议”。

类型来执行更多看似不相关的分类操作。

- **流标签**：流标签字段为 IPv6 报头所独有，人们最初是想要将其与基于资源预留的 QoS 体系结构结合使用。这是为了让路由器能够轻而易举地识别已做过资源预留的数据流。流标签的规范请见 RFC 3697。在 IPv6 报头中，流标签字段“坐落”于源地址和目的地址字段之前，因此有助于降低查找延迟，这也是其优势之一。

表 4-4 列出了 IPv4 和 IPv6 报头之间的差异。

表 4-4 IPv4 和 IPv4 报头之间的比较

IPv4 报头字段	IPv6 报头字段
版本	版本号不同，字段相同
报头长度	无此字段，IPv6 报头固定为 40 位。
总长度字段	净载长度字段
标识符、标记、分片偏移字段	无相关字段
生存时间（TTL）	跳限制
协议	下一个报头
头部校验和	无此字段
源地址	源地址（长度为 128 位）
目的地址	目的地址（长度为 128 位）
选项	无选项字段，改为扩展报头
服务类型	流标签

4.2.2 IPv6 扩展报头

在 IPv6 数据报中，被封装的数据净载之前可能会出现一个或多个扩展报头。扩展报头为 IPv6 数据报的创建提供了一种高效而又灵活的方法。必要时，扩展报头只会包含特殊用途的字段。IPv4 报头所提供的选项字段，本来也有可能会用在 IPv6 上。然而，开发 IPv6 时，人们对某些信息（比如 IP 数据报的分片信息和其他常用功能）做了优化设计。为了获得更大的灵活性，IPv6 仍需 IPv4 报头中选项字段所能提供的功能。于是，在 IPv6 数据报中添加了扩展报头，扩展报头一个接一个地排在主 IPv6 报头之后，如图 4-5 所示。

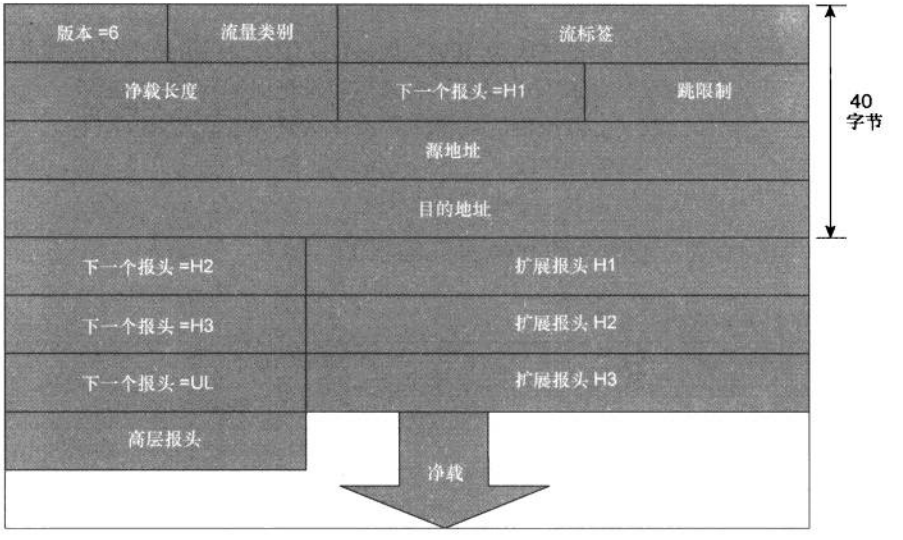


图 4-5 带扩展报头的 IPv6 数据包

有两个 IPv6 扩展报头可供 QoS 使用。

- **路由扩展报头：**可根据请求者对网络拓扑的认知程度和相关 QoS 敏感参数（比如，可能存在的吞吐量），来请求某条精确路由。
- **逐跳报头：**唯一需要被 QoS 敏感数据路径中所有设备完全处理的扩展报头。该报头会促使网络设备加快处理相关数据包，其原因是网络设备无需对高层协议做进一步分析。标准规定，若网络设备无法识别逐跳扩展报头，应将其忽略，并继续处理其他扩展报头。网络设备不允许修改在途数据包的逐跳扩展报头。

4.2.3 IPv4 和 IPv6 共存时的 QoS 机制

当网络中同时运行 IPv4 和 IPv6 两种协议时，存在以下两种可能的 QoS 机制。

- 为 IPv4 和 IPv6 流量分别施认不同的 QoS 策略。
- IPv4 和 IPv6 流量共用一套 QoS 策略，这套策略同时针对两种协议流量，执行分类和匹配操作。

两种协议流量的逐跳行为（PHB）则随以下情形而异。

- 与 IPv4 流量相对应的业务能够为企业带来收益，换句话说，较之于 IPv6 流量，至少在两种协议同时运行之初，IPv4 流量要更加重要。对于这种情况，网络工程师可能会让 IPv6 流量的优先级低于 IPv4 流量，并为前者提供较少的网络资源。
- 取决于不同应用程序所使用的流量模式，IPv4 和 IPv6 流量可能需要遵守不同的 PHB。

对于以上两种情形，应该为每种流量类型分类，并制定不同的策略。

就过渡机制而言，IPv6 流量可充分利用业已贯穿于 IPv4 基础设施而部署的 QoS 机制。在某些情况下，IPv6 流量在穿过 IPv4 网络之后，可能会丢失自己的标记。

随着应用程序对企业网提出了更高的性能方面的要求，在网络层协议（即 IPv4 和 IPv6）上斤斤计较也就没有必要了。运行特殊应用程序（比如视频）的末端用户不但不会关心所使用的网络层协议，而且无论使用哪种网络层协议，都不会对网络质量委曲求全。如此一来（IPv4 和 IPv6 流量使用同一套 QoS 策略），便降低了 QoS 部署的管理开销。如果要在第二层实施 QoS，推荐的做法是，不要对 IPv4 和 IPv6 的流量进行区分。

4.3 IPv6 路由选择

有多种 IPv4 路由协议（RP）可用在网络间发现路由，几乎每一种 IPv4 路由协议都有与之对应的 IPv6 路由协议或升级版本，包括下一代路由信息协议（RIPng）、最短路径开放优先协议版本 3（OSPFv3）、中间系统到中间系统（IS-IS）以及增强型内部网关协议（EIGRP）。尽管对路由协议的深入分析超出了本书的范围，但以下几节会简要介绍 OSPFv3、IS-IS、EIGRPv6 以及边界网关协议（BGP）。

4.3.1 OSPFv3

OSPF 属于链路状态路由协议。OSPFv2 是一种内部网关协议，用在同一自制系统内的路由器之间发布 IPv4 路由信息。为支持 IPv6，人们将 OSPF 升级到了版本 3（OSPFv3）。

运行 OSPF 的路由器会在链路状态通告（LSA）消息中通告链路状态、链路

前缀/掩码、链路权重，以及其他的本机连通性参数。为了确保每台 OSPF 路由器对网络拓扑有全面而又一致性的认识，LSA 会被可靠地泛洪给网络中的其他路由器。

在广播和非广播多路访问（NBMA）网络中，OSPF 邻居关系建立（利用 hello 协议）期间，会推举出一台指定路由器（DR）；DR 会在 OSPF 路由器之间担当通告 LSA 的中继路由器，这将有助于减少 OSPF 操作所产生的控制流量。此外，还会推举出一台备份指定路由器（BDR）。一旦 DR 发生故障，BDR 会挑起 DR 的重担，而无需发起新的 DR 推举过程。如图 4-6 所示。

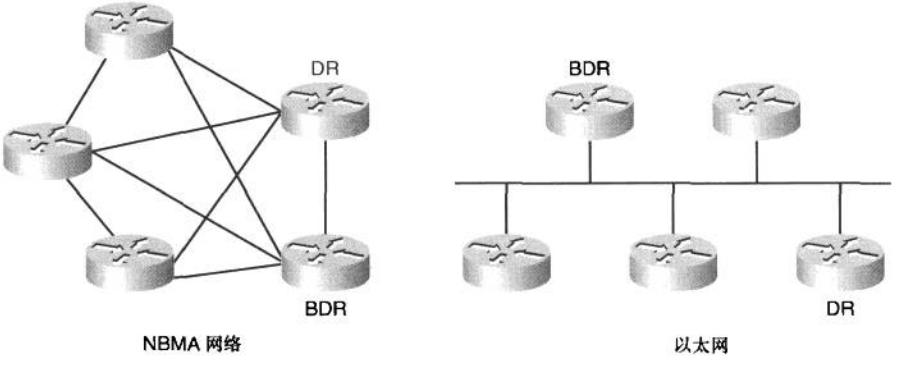


图 4-6 带 DR 和 BDR 的 OSPF 网络类型示例

OSPF 会把整个网络（自制系统）划分为一个个名为区域（area）的网络设备集合。路由器会维护自己所参与的各区域的拓扑数据库，自制系统内各区域之间的拓扑互不可见。OSPF 所提出的区域概念非常重要，有了这一概念，便可以实现两级路由选择，从而提高路由选择的可扩展性。由于路由器无需维护自己不参与区域的拓扑数据库，因此便能显著降低网络中穿梭往来的路由选择（控制）流量。降低路由选择流量的另一种方法是：在区域边界执行路由汇总。图 4-7 所示为 OSPF 区域的示例。

为保证路由信息分发的安全性，在 OSPF 报头中定义了 AuType 和 Authentication 字段（RFC 5709）。

最后，OSPF 在开发之初就支持无类别域间路由选择（CIDR）（由 OSPF 发布的每条路由都包含目的网络和掩码信息）。

定义于 RFC 5340 的 OSPFv3 是为支持 IPv6 路由选择而对 OSPF 所做的改进。OSPFv2 所用的基本机制，比如泛洪、DR 推举、区域支持、SPF 计算等，

对 OSPFv3 同样适用。在 OSPFv3 中，仍以 32 位的路由器 ID 来标识邻居路由器。然而，IPv4 和 IPv6 之间协议语义和地址格式方面的改变，也致使 OSPFv3 与 OSPFv2 之间产生了重大差异。

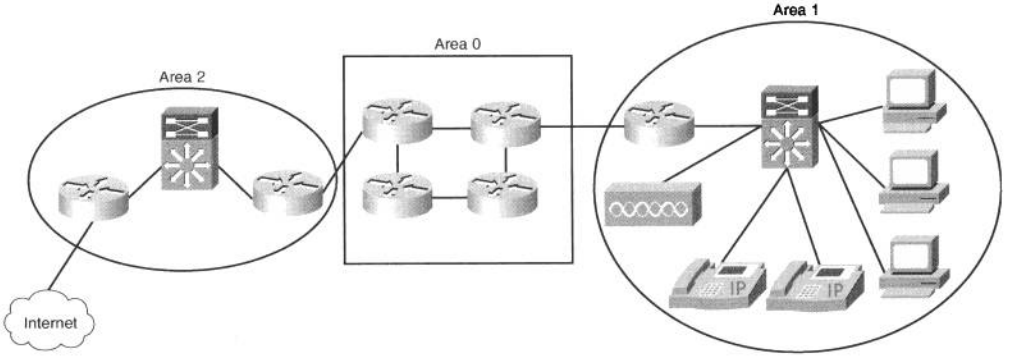


图 4-7 OSPF 区域

OSPF 协议的这两个版本彼此独立运行，并使用专有的数据库。OSPFv3 和 v2 并不相互兼容。例 4-2 所示为使用 NX-OS 时，OSPFv3 的配置。

例 4-2 NX-OS 中统一的 OSPF 配置示例

```
interface serial1
    ipv6 ospf 1 area 0
    ipv6 ospf cost 12
    ospfv3 cost 22
```

设计 OSPFv3 的目的是，创建一种不依赖于特定网络层的路由协议。为实现这一点，人们重新设计了 OSPFv3 路由器间所发送的消息，并从 OSPF 数据包和基本的 LSA 中删除了寻址语义（addressing semantics）。在 OSPFv3 中，诸如路由器或网络之类的 LSA 只携带拓扑信息。为携带 IPv6 地址和前缀，人们创建了以下两种 LSA。

- 链路 LSA：作用是向链路上的邻居路由器宣告路由器 IPv6 本地链路地址；向邻居路由器通告关联链路的 IPv6 地址列表；向邻居路由器宣告 LSA 消息的选项设置。
- 区域内前缀 LSA：用来携带通向区域内 OSPFv3 路由器的所有 IPv6 前缀信息（对于 IPv4，此类信息由路由器和网络 LSA 承载）。

人们还对以下几种 LSA 进行了修改。

- 路由器和网络 LSA[①]：不再携带前缀信息。对于 OSPFv3，以上两种 LSA 只携带网络拓扑信息，这也使得这两种 LSA 独立于网络协议。
- 区域间前缀 LSA：替代了网络汇总（3 类）LSA，其作用是向其他区域的路由器通告区域内部的网络。在 IPv6 中，此类 LSA 以<prefix, prefix length>而非<prefix, mask>的形式来表示。
- 区域间路由器 LSA：替代了自制系统边界路由器（ASBR）汇总（4 类）LSA，其作用是通告 ASBR 的位置。

OSPFv3 是以每条链路为基础，而不是像 OSPFv2 那样以每个子网为基础来运行。在物理链路上建立 OSPF 对等关系时，路由器会使用其接口的本地链路单播地址作为 OSPF 数据包的源地址。人们还针对数据包的泛洪范围进行了概括处理。OSPF 协议自身不再提供认证功能，而是借用 IPv6 协议内置的 AH（认证报头）和 ESP（安全封装净载）功能来执行认证。比之 IPv4，绝大多数的 IPv6 OSPF 数据包的格式都更为简洁，只是 OSPFv3 数据包中的 IPv6 地址字段要更长一些。

OSPFv3 具备在单条链路上运行多个 OSPF 协议实例的功能。OSPFv3 数据包报头包含了一个 8 位的实例 ID 字段，用来多路分解 OSPF 协议数据包。运行于路由器上的每个OSPFv3进程都会在自己所生成的OSPFv3数据包中自行设置实例值，并同时忽略具有其他实例值的 OSPF 数据包。

有了实例 ID，便能够对共享同一物理网络和 OSPF 区域的路由器之间的通信加以控制，而不用再依赖复杂的认证机制和访问控制列表了，这在以前都是必不可缺的。即便多台路由器共处于一个或多个物理网段上，服务提供商照样能够依靠实例 ID 去运行互相隔离的路由进程域。欲了解更多与 OSPFv3 有关的信息，请访问 Cisco 站点，链接为 http://www.cisco.com/go/ipv6。

4.3.2 EIGRPv6

Cisco 开发私有的增强型内部网关协议（EIGRP）的目的是，想要填补传统的距离矢量协议（IGRP, RIP）和先进的链路状态协议（OSPF、IS-IS）之间的空白。该协议吸收了链路状态协议某些行之有效的功能，并改进了距离矢量协议的可操作性和可扩展性。然而，其设计理念却是为了规避链路状态协议有时会遭遇到的拓扑方面的某些限制。Cisco 最终将其开发成了简单而又快速，且具备

[①] 原文是“Router link state advertisements and network LSAs”。

高弹性和高可扩展性的路由协议，并在许多公司的企业网络和某些服务提供商的边缘网络中得到了广泛部署。

EIGRPv6 是一种基于 IGRP 的路由协议，并针对 IGRP 做了如下改进。

- 扩散修正算法（DUAL）不但能够确定邻居所通告的路径是否无环，而且还能事先确立可替代的路径，无需等待其他路由器的通告。
- EIGRPv6 会让路由器存储所有学得的路由，而不是学自邻居的最佳路由。
- 当目的网络不可达时，EIGRPv6 会主动向邻居发起查询，从而在收敛时间方面具备更强的竞争力。
- 使用 hello 数据包维护邻接关系，收敛迅速。
- 利用可靠传输协议交换路由更新，从而消除了周期性通告完整路由更新的需求。
- 利用复杂的度量设置手段，来提供路由选择方面的灵活性。

为满足企业网路由选择的需求，EIGRP 协议本身采用了模块化的设计原则，其核心功能独立于网络层协议，但可配备与网络层协议相关的功能模块，因此，该协议可用于 IPv4、IPX 以及 AppleTalk 等网络层协议的路由选择。

鉴于 EIGRP 的广泛部署，也促使 Cisco 扩展其功能，去支持 IPv6。EIGRP 的模块化设计原则也简化了 IPv6 的实现——只需针对 IPv6，引入另一种协议相关的功能模块（协议标识符选定为 88，与 IPv4 相同），外加三种新型 TLV (IPv6_REQUEST_TYPE [0X0401]、IPv6_METRIC_TYPE [0X0402]和 IPv6_EXTERIOR_TYPE [0X0403])。

无论是 IPv4 还是 IPv6，在 EIGRP 的用法方面都非常相似，但在某些方面也存在差异。

- 为 EIGRP 进程所使用的 router ID 仍保持 32 位的长度，既可以手工定义，也可以取自配置在接口上的 IP 地址之一。
- 在只启用了 IPv6 功能的路由器上，如果不配置 router ID，就启动不了 EIGRP 进程。
- EIGRP hello 数据包的源地址为发送该包的接口本地链路地址；目的地址为 FF02::A（所有 EIGRP 路由器链路范围多播地址）。
- hello 数据包的格式可让两台邻居路由器不必在链路上共享同一子网

便能彼此“互见”（彼此互发 hello 数据包，因为 hello 包的目的地址为多播地址）。发送至特定对等体的 EIGRP 数据包都是以单播方式发送，对于这种情况，就要求两台路由器在链路上共享同一子网前缀了。

- 用于 IPv4 的 EIGRP 利用 MD5（消息摘要 5）来执行验证，而 IPv6 的 EIGRP 也支持同样的认证机制。写作本书之际，Cisco 正在开发利用 IPv6 协议内置的 IPSec 特性对 EIGRP 消息执行认证的功能，但 Cisco 路由器目前还不支持该功能。
- 用于 IPv4 时，EIGRP 的自动汇总功能默认开启，而对于 IPv6，则默认禁用，其原因是 IPv6 地址“生而无类”。
- 用于 IPv6 时，EIGRP 没有“水平分割”这一说，因为同一路由器接口上可能会出现多条前缀，这与 IPv4 不同。

可在 VPN（虚拟专用网）环境中，针对 IPv6 启用 EIGRP，其操作方式类似于针对 IPv4 启用 EIGRP[①]。例 4-3 所示为用于 IPv6 的 EIGRP 配置。

例 4-3 EIGRPv6 配置

```
interface GigabitEthernet7/1
  ipv6 enable
  ipv6 eigrp 100
ipv6 eigrp router 100
 no shutdown
 eigrp router-id 2.2.2.2
```

EIGRPv6 **show** 命令的输出也类似于 EIGRP，如例 4-4 所示。

例 4-4 EIGRPv6 show 命令的输出

```
R1# show ipv6 eigrp topology
EIGRP-IPv6 Topology Table for AS(100)/ID(2.2.2.2)
Codes: P - Passive, A - Active, U - Update, Q - Query, R - Reply,
       r - reply Status, s - sia Status

P 2001:DB8:CAFE:1004::/64, 1 successors, FD is 409600
```

（待续）

① 原文是“EIGRP for IPv6 will be enabled to operate within Virtual Private Networks (VPN) in a similar way as EIGRP for IPv4”。不知作者为何在此处莫名其妙地提到 VPN，致使这句话上下不接。

```
        via FE80::A8BB:CCFF:FE01:8700 (409600/128256), Ethernet0/0
P 2001:DB8:CAFE:1001::/64, 1 successors, FD is 281600
        via Connected, Ethernet0/0
P 2001:DB8:CAFE:1002::/64, 1 successors, FD is 128256
        via Connected, Loopback1

R1# show ipv6 eigrp neighbors
EIGRP-IPv6 Neighbors for AS(100)
H   Address                 Interface       Hold Uptime   SRTT   RTO  Q  Seq
                                            (sec)         (ms)       Cnt Num
0   Link-local address:     Et0/0            14 00:01:21   10   200  0  7
    FE80::A8BB:CCFF:FE01:8700
R1#
```

Cisco 对命令语法做了细微变动，用在 IPv4 中对 EIGRP 做调整和故障排除的命令同样适用于 IPv6 的 EIGRP。

4.3.3 IS-IS

中间系统到中间系统协议（IS-IS）定义于 ISO 标准 10589。起初，这一 ISO 的链路状态路由协议并非针对 IP 而开发，而是用在基于 CLNP（无连接网络协议）的网络中，在路由器之间行使路由功能。随着人们在该协议中增加了对 IP 的支持（RFC 1195）（有时，我们也将其称为集成的 IS-IS[I/IS-IS]），许多 ISP 和大企业也广泛将其选用为自有网络的 IGP（欲知更多与 IS-IS 有关的信息，请参考 Russ White 和 Alvaro Retana 所著的《IS-IS: Deployment in IP Networks》一书）。

IS-IS 的 IPv6 实现要求设置一个新的协议 ID 值（0X8E），IPv6 路由器会使用该值来表明自己具备 IS-ISv6 功能，并支持两种新型 TLV（IPv6_Reachability（0XEC）和 IPv6_Interface_Address (0XE8)）。扩展 IS-IS，令其支持 IPv6，与其当初支持 IPv4 一样，只需略加改动，这与 OSPF 的情况不同——为支持 IPv6 的路由选择，必须针对 OSPF 重新开发。正因如此，IS-ISv6 在运作方式上酷似 IS-ISv4，两者之间几乎没有差异。出现在 IS-IS 邻接关系列表中的邻居路由器地址都是路由器的本地链路地址。因为 Hello 数据包利用本地链路地址作为自己的源地址，所以即便邻居路由器之间不共享同一子网前缀，IS-IS 邻接关系照样能够得以建立。在用户看来，该路由协议只是针对 IPv6 添加了一个新地址家族。除了命令格式有细微变动以外，大多数可用于 IS-ISv4 的命令同样可用于 IS-ISv6。

IS-ISv6 利用单一拓扑结构，并针对自己所支持的所有协议（IPv4 和 v6）运行，执行同样的 SPF 计算。这一操作方式导致了某些部署方面的限制。以下两节分别介绍了单一拓扑结构和多拓扑结构。

单一拓扑结构

默认情况下，IS-ISv6 对所支持的所有协议运行单一拓扑结构，并在每个层次（level）上都运行单一实例的 SPF 计算（1 = area, 2 = domain）。由于路由器在执行上述操作时所占用的资源较少，因此这可算作一个优点。但换句话说，单一拓扑结构模式也具有某些短板。

- 区域（level1 或 level2）内的所有路由器必须在所有接口上支持同一套地址家族。这可以确保拓扑结构的一致性。这意味着，对于包含了多个孤立的 IPv6 网络的 IPv4 网络环境来说，并不适合部署单一拓扑结构模式的 IS-ISv6 协议。
- 配置在接口上的 metric 值对 IPv4 和 IPv6 同时生效。

要满足能力一致性的需求，还要解决如下问题：IS-ISv4 网络向 IS-ISv4+IS-ISv6 网络迁移时，会发生什么情况呢？由于要在路由器上配置额外的地址家族，一致性重新确立之前，邻接关系将会中断。为避免影响到运行中的 IPv4 服务，可在路由器上禁用 IS-IS 邻接关系检查特性。

注意 *在迁移过程中，为避免生产网络中 IS-IS 拓扑结构的不一致性，可在路由器上禁用 IS-IS 邻接关系检查特性。*

多拓扑结构

人们对 IS-IS 进行了改进，令其能够针对每种协议运行独立的拓扑结构和 SPF 计算。对于这种情况，不同的路由器可支持多套不同的地址家族。为了增加对 IS-ISv6 的多拓扑结构支持，Multi_Topology_Reachable_IPv6_Prefixes 这一新型 TLV 应运而生。可在 IPv6 地址家族下激活 IS-IS 的多拓扑结构操作模式。在该操作模式下，可针对 IPv6 路由设置独立于 IPv4 路由的 metric 值。

为了方便从单一拓扑结构到多拓扑结构的迁移，可采用某种过渡模式。在该过渡阶段，IS-IS 会在 LSP 中同时通告两种类型的 TLV。以通告较大的 LSP

（数据包）为代价，来换取平稳过渡[①]。

配置 IS-ISv6

在配置了 IPv6 地址的路由器接口上启用 IS-ISv6，是配置路由器运行该协议的最简单方法。无需改变 IPv4 IS-IS 进程的配置，如例 4-5 所示。

例 4-5 IS-IS 配置示例

```
isis example-area
net 49.0001.0000.0000.0001.00
!

interface FastEthernet0/1
 ip address 10.7.1.33 255.255.255.252
 ip router isis example-area
 ipv6 address 2001:FFFF:FFFF::2/64
 ipv6 enable
 ipv6 router isis example-area
```

4.3.4 BGP

边界网关协议版本 4（BGP4）是一种外部网关协议（EGP），用在 Internet 上的自治系统之间交换路由信息。BGP 根据人们对 EGP 的使用经验而设计，支持 CIDR 和路由聚合功能。BGP4 定义于 RFC 1771[②]；其他与 BGP 相关的文档包括：RFC 1772、RFC 1773 和 RFC 1774。欲深入了解与 BGP 有关的信息，请参考 Sam Halabi 所著的《Internet Routing Architectures，Second Edition》一书。

BGP 基本路由信息单元为 BGP 路径，即通往一组 CIDR 网络前缀的路由。BGP 路径会被附着上“五花八门”的路径属性，其中最重要的属性包括 AS_PATH 和 NEXT_HOP 属性。

AS_PATH 属性包含了一份路由传播过程中所途经的自制系统号的列表。通过将自己的 ASN 与接收自邻居自制系统路由的 AS_PATH 属性中的 ASN 进行比对，BGP 路由器便能检测并避免路由环路。

NEXT_HOP 属性是 BGP 路由通告机制的另一个重要环节。当 BGP 路由更新穿越自治系统边界进行通告时，其 NEXT_HOP 属性便会被改变为接收该 BGP

① 整段原文是“To facilitate the migration from single topology to multitopology, you can enable a transition mode. In this case, both types of TLVs are advertised in LSPs. Larger LSPs are thus traded off for a smooth transition”。译者未按原文字面翻译，如有不妥请指正。

② 最新的 BGP 标准是 RFC 4271。

路由更新的边界路由器的IP地址（请参见后文对eBGP的介绍）；但只要路由更新在AS内部传递，其下一跳属性绝不会做改动（请参见后文对iBGP的介绍）。这也确保了：在自制系统内，路由更新的下一跳总是宣告目的前缀的外部对等体的IP地址，而且，从数据包转发的角度来看，通告路由的内部BGP对等体，未必需要在通往由其所通告的目的前缀的数据包转发路径上担当转发路由器。

BGP的部署方式有两种：外部BGP（eBGP）和内部BGP（iBGP）。eBGP用在自治系统间建立对等关系，而iBGP则在同一自制系统内传递BGP路径信息。尽管由iBGP传递的某些路径信息（路由、metric）IGP（比如IS-IS或OSPF等）也能通告，这看起来似乎多余，但IGP却无法传递为BGP所独有的路径属性信息（比如，AS_PATH等）。因此，部署于同一自治系统的BGP边缘路由器，相互之间如欲传递接收自EBGP对等体的BGP路由属性，则必须采用IBGP。例4-6所示为通告IPv6网络2001:100::/24的BGP配置。

例4-6 IOS中的IPv6 BGP配置

```
router bgp 100
 no bgp default ipv4-unicast
 address-family ipv6
  network 2001:100::/24
```

如例4-7中**show bgp ipv6**命令的输出所示，BGP路由协议命令（包括IPv6协议族在内）与所有基于IPv4的BGP命令都以同样的方式配置。

例4-7 show bgp ipv6命令的输出

```
R1> show bgp ipv6
BGP table version is 8, local router ID is 200.10.10.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*>i2001:200:1::/48   2001:100:1:1::2          0    100      0 200 ?
*  2001:100:1:1::/64 2001:100:1:1::1          0             0 100 ?
*>                    ::                       0          32768 ?
*> 2001:100:2:1::/64 2001:100:1:1::1                       0 100 ?
*> 2001:100:3:1::/64 2001:100:1:1::1          0             0 100 ?
*> 2001:100:3:2::/64 2001:100:1:1::1          0             0 100 ?
*> 2001:100:3:3::/64 2001:100:1:1::1                       0 100 ?
*> 2001:100:3:4::/64 2001:100:1:1::1                       0 100 ?
```

BGP 运行于传输层协议 TCP 之上。连接建立时，BGP 对等体会交换各自所持完整路由表的拷贝。自那以后，BGP 对等体会维护各自的路由数据库，只会交换增量路由更新，由此可知，BGP 是一种高效的路由协议。

除了 BGP 属性之外，BGP 还可以利用 CIDR 来执行路由聚合，并以此来降低路由表的规模。假如一地址块为某 ISP 所有，该 ISP 将地址块中的一部分地址分配给自己的客户，那么 BGP 便可对接受自客户的路由做聚合处理，然后再向 EBGP 对等体宣告该完整的地址块，这样便能显著降低 BGP 路由表中路由的数量。

用于 IPv6 的多协议 BGP

RFC 4760 所载的多协议 BGP（MP-BGP）定义了 BGP4 的扩展功能，该功能可让 BGP 会话携带多种网络层协议的路由信息。

在最普遍的情况下，宣告 IPv6 路由的 BGP 对等会话将会建立在一条 IPv6 TCP 连接之上，并最终会与另一条用来宣告 IPv4 路由的 BGP 会话共存，如例 4-8 所示。

例 4-8 为 IPv4 和 IPv6 分别配置 BGP 会话

```
router bgp 10
 no bgp default ipv4-unicast
 neighbor 2001:db8:cafe:1019::1 remote-as 20
 neighbor 172.16.1.2 remote-as 30
!
 address-family ipv4
  neighbor 172.16.1.2 activate
  network 172.16.0.0
exit-address-family
!
 address-family ipv6
  neighbor 2001:db8:cafe:1019::1 activate
  network 2001:db8::/32
exit-address-family
```

多协议 BGP 是唯一能够同时携带 IPv4 和 IPv6 路由信息的路由协议。本节只是对多协议 BGP 做了简要介绍，详细的配置示例和配置原则请见 Cisco 官网，链接为 http://www.cisco.com/en/US/docs/ios/ipv6/configuration/guide/ip6-mptcl_bgp.html。

4.4 总结

本章在概述 IPv6 多播的同时，讨论了在 IPv6 网络环境中实施多播路由协议（PIM、SSM、PIM-Bidir 等）的不同方法。

服务质量（QoS）一节概述了与网络层协议（IP）版本无关的分类应用程序数据包的方法。此外，还对 IPv4 和 IPv6 两种协议做了简单比较，并简要介绍了 IPv6 数据包的扩展报头。

RIPng、OSPFv3、EIGRP、ISIS 和 BGP 之类的路由协议能够帮助路由器将数据包正确转发至目的网络。本章分别对每种路由协议用于 IPv4 和 IPv6 网络环境中的异同点进行了探讨。IPv6 与生俱来的可聚合特性能够协助网络设备有效应对快速增长的路由表规模，但与 IPv4 一样，IPv6 路由协议的收敛速度和稳定性不但仍是人们有待解决的难题，而且还很有可能会成为未来的创新领域。

4.5 参考资料

Bates, T., et. al. RFC 2858, "Multiprotocol Extensions for BGP-4." June 2000.

Deering, S., et. al. RFC 2710, "Multicast Listener Discovery (MLD) for IPv6." October 1999.

Halabi, Sam and Danny McPherson. *Internet Routing Architectures, Second Edition*, August 2000.

Popoviciu, Ciprian, Eric Levy-Abegnoli, and Patrick Grossetete. Deploying IPv6 Networks, Cisco Press, February 2006.

Rajahalme, J., et. al. RFC 3697, "IPv6 Flow Specification." March 2004.

CHAPTER 5

第 5 章　IPv6 部署规划

本章涵盖以下主题。

- **决定从何处着手**：除了要对部署 IPv6 的好处、成本和风险加以分析以外，企业还需将商业和技术因素结合起来进行考虑，所有这一切都将有助于确定从何处开始部署 IPv6。
- **试点规划**：IPv6 部署的试点阶段至关重要，这是因为只有经过该阶段的洗礼，才能暴露出培训、设备、设计以及技术等方面的缺陷。
- **规划编址方案**：在 IPv6 部署的试点阶段，出于实验的目的，可采用私网地址分配方案，但在不久以后，还需从区域性地址注册机构或服务提供商获取公网 IPv6 地址。

在向 IPv6 迁移的过程中，除了在设计和部署方面要精心规划以外，网络管理员还要参加相关培训。为应对 IPv6 部署方面的诸多挑战，需对现有网络进行评估，其中包括检查网络设备对 IPv6 的支持程度。

本章提供了一个框架，逐一罗列了部署 IPv6 前后的注意事项，可供读者在现有网络基础设施中集成 IPv6 时作为参考，以降低部署 IPv6 对业务造成的影响。

5.1　从何处着手

企业应明确采用 IPv6 的理由，例如，市场驱动或业务需求等。企业还应该将 IPv6 的互操作性与具体的业务目标联系起来[①]。

对公司来说，采用 IPv6 越早，在市场竞争力方面所占有的优势就越大。这不但能够巩固企业的市场地位，而且还能增强企业提供创新型服务的灵活性。

下面列出了企业希望部署 IPv6 的其他原因。

---

① 原文为“It should link IPv6 interoperability to specific business objectives”。译者不知何为“IPv6 interoperability”，只能按字面意思翻译。

- 需要更多的地址空间，以支持 IP 主机爆发性的增长，以及层出不穷的对等到对等和永远在线的应用程序。
- 应对 IPv4 地址耗尽的外部压力。
- 商业的连续性。
- 政府实施 IPv6 的政策。
- 客户需求。
- 全球通信。
- 供应链向 IPv6 的迁移。

在确定了部署 IPv6 的需求之后，企业应该对这项部署工作所能产生的效益加以分析。效益分析的结果可作为本次大规模部署 IPv6 的正当理由。

5.1.1 效益分析

企业应该对部署 IPv6 所能产生的效益进行分析，比方说，过渡到 IPv6 以后，涉足什么样的行业，经营什么样的项目，可从中获取多大好处。

企业需明确在部署 IPv6 之后，能否提高或能否保持自己的业务收入。一般而言，企业不外采取两种类型的 IPv6 部署方法：防御型和主动型。若企业认为部署 IPv6 并不能带来直接收益，但需要为即将涌现的新业务或新技术做好准备，便可采用防御型部署方法。试举一例，企业需确保自己面向公众的 Web 站点仍然可被持有支持（或只支持）IPv6 设备的潜在客户访问。只有这样，企业业务的连续性才能够得到保证。采用后一种部署方式向 IPv6 过渡，企业则可以通过赢得新的商机，或降低现有业务的维护成本，来增加收益。

在技术方面，部署 IPv6 可以消除配置隧道或 NAT 之类的临时性解决方案的烦恼。这不但会改善网络的性能，而且还会极大地简化运维工作。相较于 IPv4，IPv6 在技术特性的多样性方面要更胜一筹，这也令其在功能上更为强大，使用起来也更为灵活，从而能够让网络工程师更合理地去部署网络应用和服务。以下简要罗列了 IPv6 的主要特性。

- **IP 地址数**：IPv6 可供数十亿的新设备（比如移动设备、交通监控设备以及传感器等）接入 Internet。
- **自动配置**：IPv6 标准所定义的自动配置特性不但可让 IPv6 主机自行定义网络运行参数，而且还能自动建立与其他主机通信的信道。

- **安全性**：IPv6 标准要求所有 IPv6 主机都要支持 IP 层的安全性（IPSec），这可让 IPv6 主机之间能够安全的交换数字信息。
- **移动性**：就使用方面而言，移动 IPv6（MIPv6）模型要比其 IPv4 等价模型简单得多。
- **多播**：在 IPv6 网络中，对多播的运用要更为广泛，而且人们还对 IPv6 多播的通信机制进行了改进。此外，对 IPv6 多播地址空间的扩充和地址范围的明确，这也简化了对多播服务的调配工作。
- **扩展报头**：对扩展报头的支持，使得 IPv6 数据包除了基本报头以外，还能携带协议级信息。
- **降低了运营成本**：部署 IPv6 可降低日后的网络管理成本。

大块地址空间、NAT 网关的撤除、自动配置特性、集成了对 IPSEC 的支持以及移动性等诸多 IPv6 主要特性，使得企业能够降低网络的运营成本。

5.1.2 成本分析

企业也需要分析有关部署 IPv6 的成本，例如：

- 设计成本和工程（包括规划、系统集成、设计、测试、实施和部署等[①]）费用；
- 基础设施方面的升级和调整可能会包括：
  - 软、硬件（服务器、PC 以及其他所有 IT 装备）；
  - 应用程序；
  - 运营支撑系统。
- 与 IPv6 以及 IPv4 和 IPv6 互连有关日常技能培训费用；
- IPv6 部署启动之后的运营和维持成本。

简而言之，渐进式的 IPv6 部署方式需要企业投入资金[②]。可在正常的软硬件更换周期内，在网络中添加 IPv6 特性，以求将边际成本（incremental cost）降至最低。在能够支撑 IPv6 运行的基础设施就位之后，可利用双协议栈技术逐步从 IPv4 过渡到 IPv6。除了软、硬件成本之外，还应考虑：与部署 IPv6 有关

① “实施”和“部署”，“设计”和“规划”有何区别，两者择其一足矣。

② 这又是废话，请问哪种 IPv6 部署方式不需要花钱？

的成本（施工方面的成本）；经常性的开支和增量支出。

要是不能提前做好规划，那么企业对IPv6部署的前期投入（比如，基础性建设和技术方面的投资）将会十分之巨。同理，要是企业因市场状况而被迫在短期内部署IPv6的话，那么仍少不了巨额投入，需要巨额投入的原因是：现有设备的更新换代周期与基础设施的升级时机并不一致。对企业来说，可选择最佳时机去部署IPv6，以求开源节流（将建设成本降至最低）。在大多数情况下，可采用一种极具成本效益的方式部署IPv6——将部署IPv6纳入正常的设备更新换代采购周期，但若特定的IPv6需求并未包括在设备采购流程中，且随后需要快速迁移至IPv6，那么额外的开支便不可避免。

5.1.3 风险

应看清与部署IPv6有关的风险，这可能会涉及商业、法律以及技术层面。对于认识到的每一种风险，都应该采取一定的缓解手段，以求尽量减小或避免上述风险所带来的危害。

- **商业风险**：能否实现预期的收益和投资回报？公司的规模能否随IPv6的部署而发展壮大？看待商业风险的方法不止一种。可以从基础设施的维护方面来看待IPv6的部署，这与升级公共电话交换机（PBX）非常相似——尽管原有的基础设施已不堪重负，但仍能满足当前的商业需求。不堪重负的网络基础设施除了会增加商业风险以外，有时还会使运营成本上升。部署IPv6会降低商业风险，即便不能带来直接收益，但至少可在日后减少损失。
- **法律风险**：作为一个唯一的标识符，IPv6地址有利于对网络行为的跟踪。这可能会侵犯到某些人的隐私，因此网络管理员必须了解法律法规，以保护用户的隐私不受侵犯。
- **技术风险**：只要与IPv6有关的过渡机制实施不合理，管理不得当，那么安全风险便会逐渐显现。不同的IPv6过渡机制都会面临不同的技术风险。某些安全设备可能并不具备检测或过滤IPv6数据包的能力，从而会给攻击者提供可趁之机。需关注互操作方面的风险，以防止使用其他IPv6协议栈（比如，会话启动企业[SIP]和移动协议）时可能发生的互操作问题。

5.1.4 商务案例

新技术的部署应满足以下商业需求。

- 节省产出成本。
- 创造或增加收入来源。
- 形成竞争性或战略性优势。

在当今经济形式下，IT 决策都与其投资回报率（ROI）有关。虽然 ROI 可算是一种精确的测量手段，但却无法测算出以下因素。

- 风险。
- 复杂性。
- 无形资产。

在部署 IPv6 的商务案例中，需同时考虑上述因素和 ROI。此外，还需考虑某些财政措施，比如成本节约和收入来源等。部署 IPv6 的商务案例应着重突出项目重点、具体的解决方案，以及影响面。

5.1.5 过渡团队

成立名为过渡团队的项目组，去监督 IPv6 部署工作。团队成立之后，便可以在整个企业内规划、协调或传达与 IPv6 部署有关的事宜。此外，还需确保该团队享有充分的资源（例如，人员配制、培训和预算等），去完成计划内业已明确的任务。企业规模越大，越分散，便越需要成立这样的团队。

企业应对团队成员精挑细选，并明确其分工。团队成员可以来自不同的部门，比如，技术、业务或其他的商务部门。一般情况下，企业的网络/系统管理员、软件开发工程师以及技术支持工程师都会参与其中，并配合一名或一组项目经理工作。

过渡团队还需在企业内宣传 IPv6，让员工建立起对 IPv6 的认知。员工也应该大致知晓 IPv6 是怎么一回事儿，IPv6 对工作领域的影响，以及 IPv6 为什么对企业如此重要。

为了推动 IPv6 部署项目，制定相关规章制度，过渡团队在企业内还需拥有一定的行政权。在遇到资源分配问题时，握有一定的行政权力可以全面保证 IPv6 部署项目优先获得相关资源。

对于在企业内部署 IPv6 时可能会波及的领域，过渡团队要能有所预判。这可以通过项目管理手段，比如，定义里程碑来实现。

团队内的每个工作组要各尽其事，各司其职。应该为每个工作组配备一名组

长。团队可为每个工作组设立工作目标，并确保每个工作组人员充足，训练有素。

需定义指标，以追踪项目进展。通过汇报列入商务案例中的里程碑的完成情况，来增强管理层的信心。

过渡团队需明确制定并传达相应的规章制度，以确保企业所有相关部门在自己的未来规划中考虑IPv6。应及时发布并强调与IPv6相关的规章制度，选择并确定用来审查和执行规章制度的方法和工具。先举一个与IPv6相关的规章制度的例子：企业今后所采购的设备必须具备或支持IPv6功能。再举一个与执行上述规章制度有关的例子：只有兼容IPv6的商务项目才会获得批准。

在为整个企业制定全面的整体IPv6部署规划时，应确保步调一致，井然有序，并能够分清轻重缓急。

在执行完IPv6部署规划之后，过渡团队还需在企业范围内定期传达IPv6部署项目的进展情况。

5.1.6 培训

对于大企业，要想为IPv6部署项目的启动，以及日后的运维工作做好准备，开展与部署IPv6有关的技术或业务方面的培训势在必行。企业员工可根据自己所承担的工作，参加各种培训。下面列举了与IPv6有关的培训类型。

- **普及性培训（Awareness training）**：此类培训提供对IPv6技术的一般性介绍，侧重于让参训员工能够对IPv6有一个基本的认知，其中包括IPv6的商业驱动、部署问题、潜在的应用，以及对已支持IPv6的设备做简单介绍。
- **体系结构方面的培训（Architectural training）**：此类培训为专人而设，“专人”是指对架构、设计和部署IPv6起举足轻重作用的一干人等。
- **运维方面的培训（Operational training）**：负责已竣工IPv6网络运维支撑工作的工程师应参与此类培训。
- **专业培训（Specialized training）**：此类培训为专攻特殊技术领域（比如，移动、安全等）的业务专家（Subject Matter Experts [SME]）而设。

5.2 试点规划

企业在引入像IPv6这样负责提供应用程序底层连通性的新技术时，在实验

室中先行搭建一个试点/实验网络，不失为明智之举。在生产网络的实验室镜像网络中先行测试(Testing in a lab environment that mirrors the actual network)，既有助于网络工程师更好的掌握 IPv6 技术，又能够让企业员工对 IPv6 日后在企业网中的成功运行树立起信心。在成功完成实验环境的测试工作之后，应进入生产网络的试点实施阶段。这样的试点实施可以只在生产网络的某个小的区块展开。

以下各节将会简单概括 IPv6 网络试点工作所包括的某些行为。IPv6 网络试点工作完成之后，可根据 IPv6 大规模的部署/迁移方案，对上述行为进行改进。本书的第 12 章会对多个 IPv6 试点案例进行研究。

5.2.1 评估

在部署 IPv6 之前，企业应对目前在用的网络基础设施详细地的加以评估。评估的结果可为相关部件的升级或更换提供初步参考依据。以下举例说明了在部署 IPv6 之前，企业需要评估的内容。

- **现在及未来的地址分配规划**：这有助于在当前的地址分配方案和 IPv6 地址分配方案之间，建立起联系。
- **网络**：调查企业网中现有的网络类型。可能的网络类型包括 IP、无线、DSL、VoIP 以及 CPE。应整理出一份完整的网络设备库存清单。
- **网络服务**：确定企业网中目前运行着何种网络服务。例如 DNS、AAA、DHCP、NTP 等。
- **网络管理**：评估当网络部署了 IPv6 之后，将如何管理。可利用某些工具来管理网络，比如 NetFlow、MIBS、SNMP 等。
- **网络应用**：弄清诸如 VoIP、数据库等网络应用与 IPv6 的交互方式。应梳理出一份完整的在用软件及操作系统清单。
- **其他的基于 IP 或 IP 感知的服务**[①]：其他服务（诸如定位或移动）将使用 IPv6 作为通信机制。

5.2.2 设计

应针对受 IPv6 部署影响的各个领域，制定详尽的 IPv6 部署方案。IPv6 设计方案不但要符合标准，而且还应尽可能的提供与 IPv4 等价的特性，以满足平

① 原文是“Other IP-based/aware services:”，译文为直译。

稳过渡的需求。设计方案不应忽视上一节提及的任何一种网络或服务，还应重点考虑企业日后的流量增长。以下各节将会介绍制定IPv6设计方案时，所应考虑的各个方面。

IPv6编址方案

首先，需弄清企业网络的编址需求。如本节所述，编址方案应考虑此后数年企业网络的IP编址需求、地址分配、地址管理以及地址获取的流程，例如：

- 调查企业的编址情况，其中包括企业自有的基础设施、内联网和外联网、不受企业管理的站点，以及由企业提供或支撑的服务（比如，第三层VPN）；此外，还需对企业网络未来的IP地址使用情况做有效地预测。
- 制定IP地址分配方案。按此方案分配IPv6地址，除了能够满足网络设备和用户使用以外，还应能够为网络提供高效而又稳定的路由选择。是否在企业网内让主机使用无状态地址自动配置（SLAAC）特性，或状态化配置特性，也应作为IP地址分配方案的一部分来考虑。
- 确定是否采用IP地址隐私扩展功能。一旦采用，主机便会在不同的时间使用不同的IPv6地址作为源地址；例如，同一台主机的IPv6地址每日一换。
- 确定IPv6路由选择是否会影响到现有的IPv4路由选择，还需考虑IPv6路由选择如何与现有的IPv4路由选择结合使用，以及如何调整现有的IPv4路由选择[①]。
- 为通向Internet和其他站点的外部连接（链路）分配IPv6地址。这意味着需根据IPv6的对等情况，与ISP（Internet服务提供商）相互协调。同理，也要规划好使用IPv6的端到端站点链路的IPv6地址。

第12章会提供与IPv6编址规划有关的其他信息，还会给出IPv6编址规划的示例。

5.2.3 过渡机制

在向IPv6过渡的相当长一段时间内，IPv4和IPv6很可能会在网络中并肩

① 原文是“Determine the impact of IPv6 routing, its integration, and changes with the existing IPv4 routing”。译者不才，从字面实在看不出作者想表达什么，译文为译者杜撰。

运行。因此，企业需考虑不同的过渡机制，让 IPv6 和 IPv4 网络环境共存的同时，推动网络环境向 IPv6 的过渡。选择过渡机制时，企业需要考虑的问题包括：自己当前的网络环境、未来的 IPv6 流量、支持 IPv6 的设备或应用程序，以及 IPv6 部署方案。如本书的第 3 章所述，转换机制可分为三类，即双栈机制、隧道机制和转换机制。在 IPv6 部署的试点阶段需逐一测试以上三种过渡机制，以掌握每种机制之于网络的行为，此外，如此行事还可让网络工程师巩固自己的部署经验。

5.2.4 网络服务

部署 IPv6 时，势必会波及网络服务。领会其中的变化，将会是构造高效 IPv6 网络的关键。网络服务包括 DNS、DHCP、AAA 和 NTP 等。比方说，公司可能需要去决定是否让主机支持自动配置特性或 DHCPv6，或者同时支持两者。该公司还有可能要考虑实施双栈 DNS，以同时支持 IPv6 和现有的 IPv4 地址查询。作为试点的一部分，以上提及的网络服务都需要在网络中部署，对于在现有网络中为末端 IPv4 用户提供服务的等价 IPv6 网络服务来说，更应如此。

5.2.5 安全性

除了要防范类似于 IPv4 网络中存在的安全威胁以外，企业还需要提防向 IPv6 过渡期间，网络中新“冒”出的安全漏洞。比方说，原有的防火墙和入侵检测系统不提供 IPv6 数据包的检测和过滤功能，这些设备可能需要升级。恶意之徒可能会以隧道的方式封装 IPv6 流量，神不知鬼不觉地通过上述安全设备的检查。此外，IPv6 所具备的某些自动配置特性，虽然能够减低运维成本，但却又增加了安全隐患。例如，某些恶意之徒可能会对恳求、通告和绑定消息进行欺骗[①]。

对使用自动隧道的应用程序来说，其流量可以不受检测地穿越防火墙，并因此给外部世界攻击内部网络留有可乘之机。为了尽量避免上述问题，还需制定相关机制和安全策略来保护 IPv6 网络。

在 IPv6 部署的试点阶段，应考虑如何保护启用了 IPv6 的网络，比方说，可考虑部署基于 IPv6 的 IPsec。

5.2.6 IPv6 新特性

IPv6 引入了诸多新特性，包括状态化地址自动配置（SLAAC）、内置的 IPSec、

---

① 原文是“For example, malicious users can spoof solicitation, advertisement, and binding messages”。译者也不知道作者提及的“solicitation、advertisement 和 binding”具体是什么玩意儿。前两者很可能是 ICMPv6 里的“邻居恳求”和“邻居通告”消息，而“binding”消息很可能与移动 IPv6 有关。作者行文的最大特点是：句与句之间，段与段之间完全没有照应。

移动IPv6、流标签等。在IPv6部署的试点阶段，是评估上述特性的最佳时机。在此期间，每个企业都可以试用上面提到的全部或部分新特性。对上述特性的掌握将有助于企业确定IPv6的部署策略。

试举一例，开启了SLAAC特性的主机接入进一个路由式IPv6网络时，可自动获得IPv6地址。但这未必适用于所有应用程序，比如，必须使用DHCPv6状态化地址配置特性或静态配置IPv6地址的应用程序。

5.2.7 稳定性和可靠性

将IPv6引入现有网络时，企业必须能够确保IPv6设计方案本身的稳定性和可靠性。此外，还需确保IPv6解决方案不会给现有的网络环境造成任何负面影响。

双栈机制是指在设备上同时运行两种协议栈。需要紧密关注运行双协议栈给设备带来的开销。例如，运行双协议栈可能会致使某些设备CPU利用率的提升。

就可扩展性而言，则需关注运行路由协议的设备所能承载的路由条数，以及与其建立路由邻接关系的邻居数。为了确定设备的可扩展能力，企业可会同设备供应商来计算上述数字[①]。

5.2.8 服务等级协定

应制定能够反映出新IPv6策略，并包括过渡实施阶段的服务等级协定（SLA），这是因为过渡机制的实施往往会影响到SLA能否严格执行。一般而言，企业都会套用现有的SLA，然后再简单地添加有关IPv6的内容。但如此行事却未必合理，其原因是在一定的时间周期内，企业和服务提供商都会一直推广并使用过渡机制，这势必会给网络带来额外的延迟和抖动，以及更高的丢包率。

5.2.9 总结教训，开始实施

在试点阶段过后，需观其效果，并进行总结。在试点阶段所采集到的数据能够为企业顺利过渡到IPv6助一臂之力。此外，试点阶段还为企业员工学习IPv6这项技术提供了有利时机。取决于IPv6试运行（试点）的效果，已经证明了哪

① 原文是“In terms of scalability, the number of routes and neighbors needs to be addressed. This can be worked out with the vendor to determine IPv6 scalability on the device”。原文根本不能入眼，译文为译者杜撰。

些 IPv6 特性对企业有利，企业可针对此“去芜存菁”。

接下来，企业需要为现有的每台设备和每种应用制定企业内部的合规标准——这理所当然要根据行业内的最佳做法来制定。企业可以遵循的标准和认证有 USGv6 和 IPv6-ready logo（IPv6 就绪性徽标）等。对那些需要过渡到 IPv6 的设备和应用程序而言，还有两件事情有待完成。

- **对即将提供 IPv6 服务的设备和应用程序进行评估**：这可能需要联系设备及应用程序厂商，并要求厂商确认自己的 IPv6 产品详情及可用性[①]。
- **确定对所支撑的服务以及对客户的影响**：这可能需要在厂商和系统集成商之间建立起沟通渠道。厂商可能只会提供某一种支持 IPv6 的产品或服务。而系统集成商将会提供多种支持 IPv6 的产品或服务。因此，在 IPv4/IPv6 网络环境中，弄清应用程序和设备之间的相互影响和相互作用最为关键。

IPv6 的试运行大功告成之后，应立即制定针对整个企业网的 IPv6 过渡部署规划。在决定采用哪些 IPv6 特性，以及如何配置和管理这些特性时，IPv6 试运行（实验环境/试点阶段）的效果将会起到至关重要的作用。

可通过制定一份项目清单，来启动 IPv6 的实施工作。在这份清单中，应明确相关性，并正确地区分项目和任务的优先级[②]。尽量利用 IPv6 实验环境去验证 IPv6 的架构、设计以及业务规则等。

下一步，可在企业网网络基础设施（设备）的更新换代期间借机部署 IPv6，并应同时确保所有购得或改良的设备都能够支持 IPv6，更为重要的是，对这些设备来说，还需具备企业所需的各种 IPv6 特性，这一切都要视 IPv6 试点阶段所总结出的成果而定。

5.2.10 客户端/服务器的 IPv6 迁移方案

部署 IPv6 时，肯定不缺技术方案。选择正确的部署方案，不但取决于当前的网络环境，而且还要视末端用户设备及操作系统、路由器型号及版本、关键应用程序、预算和资源以及时间进度而定。以下各节将会介绍迁

① 原文为“the availability of their IPv6 product road map”，译者明知有点不对劲，但也只能这么译。

② 原文是“Within that list, identify dependencies and prioritize the projects and tasks appropriately”。译者也不晓得“dependencies”是什么，直译。

移阶段的准备工作。在介绍准备工作时，将会使用图 5-1 所示的网络作为基线，该网络代表所有只运行 IPv4 的网络。我们会逐一剖析三种最基本的迁移方案。读者可根据公司网络的复杂性、业务策略以及其他因素来套用这些方案。

图 5-1 所示为一台运行 IPv4 应用程序、套接字以及协议栈的客户端，外加一台具有相同配置的服务器；互联客户端和服务器的网络被划分为：可供服务器和客户端接入的接入层网络，和核心层骨干网络。该图显示的是最基本的客户端到服务器的连接。在特定情况下，该图还可以用来表示某内部客户端和某台内部服务器之间，通过所有内部网络的接入层和核心层网络，以及 Internet 建立连接。

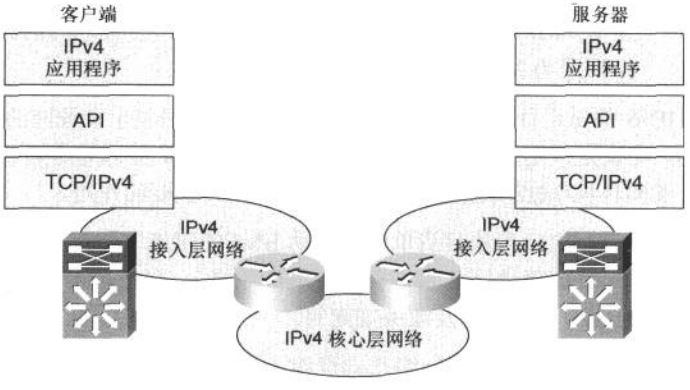


图 5-1 典型的使用 IPv4 的客户端/服务器模型

从网络的核心层开始部署 IPv6

本 IPv6 部署方案是指：先从网络的骨干或核心层开始部署 IPv6，然后向网络的边缘逐渐延伸，如图 5-2 所示。

采用本部署方案，要求网络的核心层路由器支持双协议栈（IPv4/IPv6）路由协议。通过在网络的核心层率先启用双协议栈，企业可以在不中断端点间生产网络流量的情况下，了解到 IPv4 与 IPv6 协议之间、路由协议之间以及其他诸多方面的差异。对于网络中的各端点来说，若没有明确的部署 IPv6 的时间限制，本部署方案可谓是上上之选。图中的核心网络既可以是企业内部的骨干网，也可以是 IPv6 ISP 网络，对于后一种情况，可能需要采用 6PE 或 6VPE 机制。

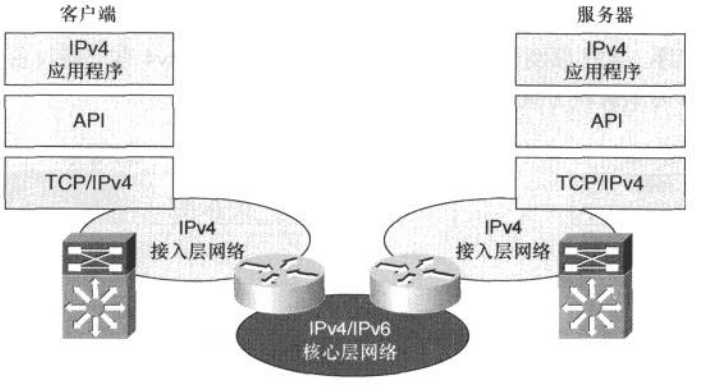


图 5-2 从网络的核心层开始部署 IPv6

从服务器端开始部署 IPv6

本方案是一种局部化 IPv6 部署方案，牵涉到将服务器和应用程序升级到支持双协议栈。服务器仍能提供 IPv4 的应用程序服务，末端用户也像以前一样通过 IPv4 访问服务器。但服务器也能提供 IPv6 的应用程序服务，可供支持双栈或纯 IPv6 的客户端访问。此时，除非客户端能够支持 IPv6，否则客户端与服务器要么建立的是 IPv4 连接，要么中间部署有地址转换或代理设备。当部署 IPv6 的工作重心落到数据中心网络时（让数据中心内的应用程序、操作系统以及网络设备支持 IPv6），本方案可谓是最佳选择。在客户端能够利用 IPv6 与服务器建立连通性之前，往往会暂时采用本方案。许多公司采用本方案的原因非常简单——服务器区域（数据中心）是部署 IPv6 难度最大的区域——与其他任何区域（比如，园区网络区域）相比，在设置服务器区域的操作系统、应用程序，以及支撑两者的网络基础设施的 IPv6 功能时，都需要做更为全面的考虑。在运用本方案时，有一点还需注意：某些网络服务，比如，DNS，只有当网络“从头到尾（end-to-end）”都支持 IPv6 时，才能有效运转。图 5-3 所示为本 IPv6 部署方案。

从客户端开始部署 IPv6

让客户端主机和接入层网络设备支持双协议栈，是另一种局部化 IPv6 部署方案。与上一种方案一样，本方案也需要充分的准备。由于本方案并不能建立端到端的 IPv6 连通性，因此若客户端不采用隧道或转换机制，便无法真正访问到启用了 IPv6 的服务。本方案可先在端点主机/网络设备上启用双栈功能，

然后，逐步在整个企业网内端到端地启用双栈功能（如图 5-5 所示）。IPv6 应用程序、API 以及 TCP/IP 协议栈会先在现有的 IPv4 客户端设备上安装。图 5-4 所示为本解决方案。

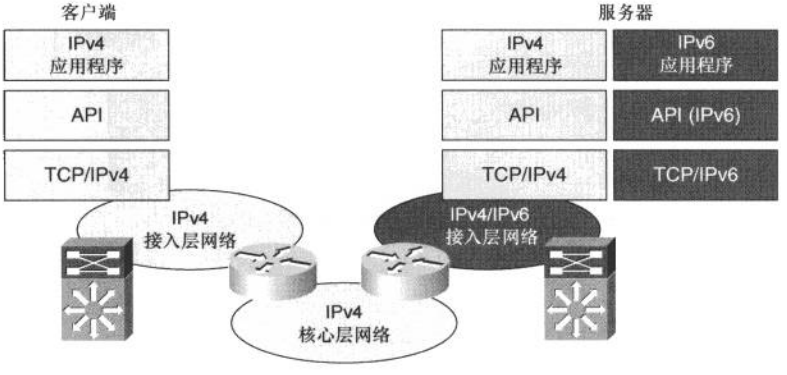


图 5-3 从服务器端开始部署 IPv6

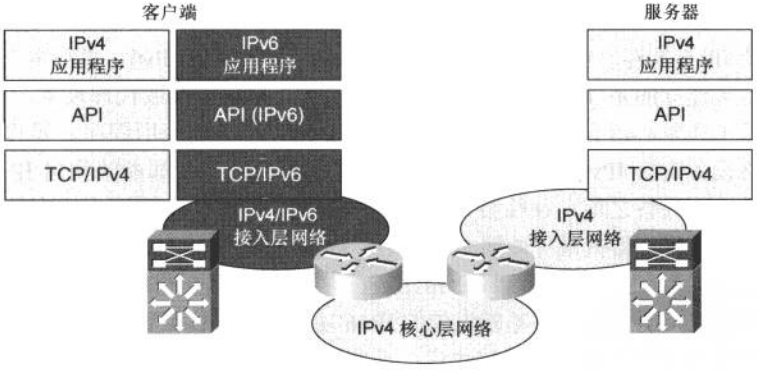


图 5-4 从客户端开始部署 IPv6

从客户端和服务器端双管齐下：部署双协议栈

同时升级客户端和服务器（适当时，应包括应用程序）也是一种在企业网内部署 IPv6 的方法。如图 5-5 所示为客户端和服务器同时部署双协议栈的方案[①]。

① 原文是“The use of dual-stack deployments, illustrated in Figure 5-5, facilitates the deployment to clients and servers over time”，原文让人费解，译文酌改。

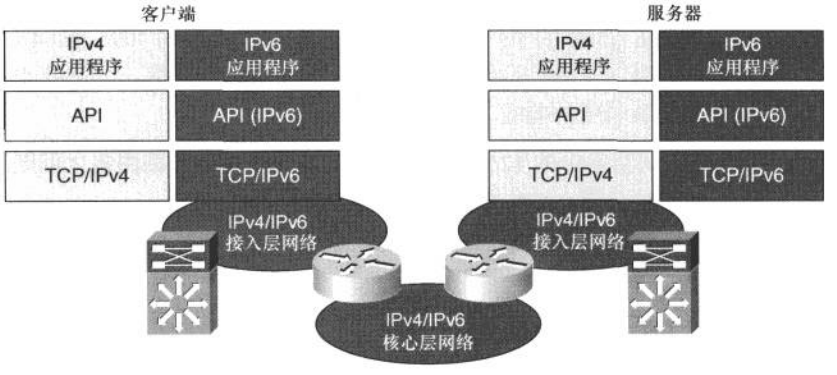


图 5-5 客户端和服务器端双管齐下：部署双协议栈

将应用程序迁移到 IPv6 时，需深思熟虑，在迁移被庞大用户群所访问的企业核心应用程序时，更需加倍小心。理想的做法是，在迁移过程中，让客户端同时支持 IPv4 和 IPv6，但这未必能够行得通。

5.3 规划编址方案

公网 IP 地址由 IANA 负责分配。IANA 会委派区域性的 Internet 注册机构（RIR）分配地址块。每个 RIR 都负责为某个地区分配地址。以下列举了一些 RIR。

- **ARIN**：美国 IPv6 注册服务机构（http://www.arin.net）。
- **APNIC**：亚太网络信息中心（http://www.apnic.net）。
- **RIPE**：欧洲地区 Internet 注册机构（http://www.ripe.net）。

RIR 会为本地 ISP 或末端用户分配 IPv6 地址。绝大多数的 ISP 都会遵循其 RIR 的地址分配指引，为客户分配单个/48 的子网。这一分配方式可让企业获得 65536 个/64 的网络。每个/64 的子网都拥有 2^{64} 个 IPv6 地址。

在拿到了一个 IPv6 子网之后，为执行高效路由选择和路由聚合，企业应对该子网做进一步的合理分配。目前还没有 IPv6 地址规划的最佳做法，但存在以下指导方针。

- **确保每个 LAN 子网都是/64 的网络**。在分配 IP 地址之前，先行统计出 VLAN 数，并能对未来所需的网络数做出合理的判断，可谓是良好的开端。

- **确定网络标识符所需的地址空间**。比方说，一个/56的地址空间将会得到256个/64的子网，而/52的地址空间将会得到4096个/64的子网。4位之隔（从/48开始划分），是为了“好看”。当然，可使用48/和64/之间任意长度的前缀。
- **0子网空间中的/126网络专用于链路（路由器和路由器之间链路）子网的地址分配**。这也是RFC 3627的推荐做法。

第12章提供具体的地址分配示例和相关建议。

5.4 总结

本章讨论了IPv6部署规划，并说明了如何启动IPv6部署过程，这包括规划、提出商务案例、组建监管IPv6部署的团队等。

为了摸索IPv6技术，并为IPv6在生产网络中的试点做好准备，应在实验室中先行搭建IPv6试验网络。

本章篇末则讨论了IPv6地址分配和如何获取IPv6地址，此外，还介绍了某些可用来完成IPv6编址方案的指导方针。

5.5 参考资料

Cisco. “IPv6 Deployment Strategies.”
http://www.cisco.com/en/US/docs/ios/solutions_docs/ipv6/IPv6dswp.html.

Savola, P. RFC 3627, “Use of /127 Prefix Length Between Routers Considered Harmful.” September 2003.

CHAPTER 6

第 6 章　园区网络中的 IPv6 部署

本章涵盖以下主题。

- **园区网络区块 IPv6 部署模型概述**：介绍企业网中最为常见的三种园区网络区块 IPv6 部署模型。
- **园区网络区块 IPv6 部署通则**：详细介绍适用于上述三种 IPv6 部署模型的一般性原则。
- **实施双栈模型**：详细列举了实施双栈模型的配置示例。
- **实施混合模型**：详细列举了实施混合模型的配置示例。
- **实施服务区块模型**：详细列举了实施服务区块模型的配置示例。

本书在介绍企业网的 IPv6 部署时，先将整个企业网划分为各个区块，然后再逐区块地剖析 IPv6 的部署要点。本章将讲解企业园区网络区块的 IPv6 部署。园区网络区块对 IPv6 的需求与 WAN/分支机构网络环境有所不同，这主要是因为在园区网络区块中，网络基础设施只有具备了线速转发 IPv6 数据包的能力，才能满足高性能需求，比方说，网络设备必须能够以 10/100/1000 Mbit/s，甚至是 10Gbit/s 速率来转发 IPv6 数据包。

若园区网络区块中的三层交换机不支持 IPv6 数据包的线速转发，则可以采用其他的设计选项，本章将会随文介绍这些设计选项。

6.1　园区网络区块 IPv6 部署模型概述

以下各节将会简要介绍三种园区网络区块 IPv6 部署模型，并会讨论每种部署模型的优点与适用场合。

- **双栈模型（DSM）**：在同一网络设备接口上同时启用 IPv4 和 IPv6 功能。
- **混合模型（HM）**：必要时，利用基于主机的隧道机制，将 IPv6 数据包

封装在IPv4数据包之内进行传输，而网络的其他地方则使用双栈机制。

- **服务区块模型（SM）**：类似于混合模型，由名为服务区块的专用网络区块终结IPv6-in-IPv4隧道。

6.1.1 双栈模型

双栈模型是一种完全依赖双协议栈的过渡机制。双栈模型是指在一个接口或一条链路上同时启用两种协议，并以双协议栈的模式运行。使用双栈模型的例子包括：在同一台设备上同时运行IPv4和IPX或IPv4和AppleTalk协议，这在前面的章节已经提到。

双栈模型是在现有IPv4网络环境中部署IPv6的首选而又通用的方法。采用此法，可在任何启用IPv4的网络中部署IPv6，并能够利用上与IPv6有关的任何特性，比如，路由选择、高可用性以及安全性等。

本章各节所采用的测试部件（试验用交换机平台）列表也简要罗列出了实施DSM的一般性需求。诸如交换机之类的园区网络组网设备能否以硬件方式（线速）支持IPv6转发，是部署IPv6时应首先考虑的因素。在园区网络区块内，链路的速率和容量往往要根据用户数、应用程序类型，以及用户所期望的延迟来决定。一般而言，由于园区网网络环境对数据传输的高速率要求，因此若部署于其中的网络设备只能以软件方式来转发IPv6单/多播数据包，Cisco就不建议在此网络环境中部署IPv6[①]。对读者来说，关键是要弄清自己的网络环境中是不是还遗留着能够以线速（硬件方式）转发IPv4数据包，但只能以软件方式转发IPv6数据包的网络设备。

对于前面提到的园区网网络设备，只适合在测试环境或小规模试点网络中启用其IPv6的转发功能，请千万不要在生产环境中部署此类设备。读者应务必对自己的园区网络区块中的所有交换机“明察秋毫”，以确保这些交换机都能以线速（以硬件方式）转发IPv6数据包。

以下各节将着重介绍DSM的优缺点，并会简要介绍其高层拓扑及实验用交换机平台。

6.1.2 DSM的优缺点

在园区网络区块中采用DSM部署IPv6，拥有诸多采用另外两种模型所不具备

[①] 原文是“Because of the typically high data rate requirements in this environment, Cisco does not recommend enabling IPv6 unicast or multicast layer switching on software forwarding-only platforms”。

的优点。采用 DSM 的最大优势是，无需在园区网络区块中“开凿”隧道。DSM 以“午夜航船”的方式运行两种协议，这意味着，除了共享相同的网络资源以外，IPv4 和 IPv6 独立运行，各行其是，不会互相干扰。无论是路由选择、高可用性（HA）、服务质量、安全性还是多播策略，IPv4 和 IPv6 都“各自为政”。此外，双栈机制还能使网络设备在数据包的处理性能方面更具优势，这是因为数据包都是被“原汁原味”地转发，既没有额外的封装，也不会造成网络设备查表方面的开销。

注意

准备或已经按 Cisco 路由式接入设计方案组网的客户将会发现，同样的设计方案也适用于 IPv6 的部署。

若现有的网络设备不支持 IPv6，便需要对设备进行升级，这也是 DSM 最致命的缺点。此外，同时运行两种协议还会增加运营成本，原因不言自明：无论从编址、路由协议、访问列表还是从管理等方面来看，IPv4 和 IPv6 都完全是两码事。

6.1.3 DSM 的拓扑结构

图 6-1 所示为在园区网络区块中采用 DSM 模型部署 IPv6 的简化拓扑结构图。该网络采用的是传统的三层设计模型，即整个网络区块由接入层、分布层和核心层构成。在所有网络设备/主机的相关接口上，同时启用了 IPv4 和 IPv6 功能，这也使得整个网络区块成为了纯双协议栈网络。本章随后将会围绕图 6-1 所示的网络拓扑，展开对 IPv6 DSM 部署模型配置方面的深入讨论。

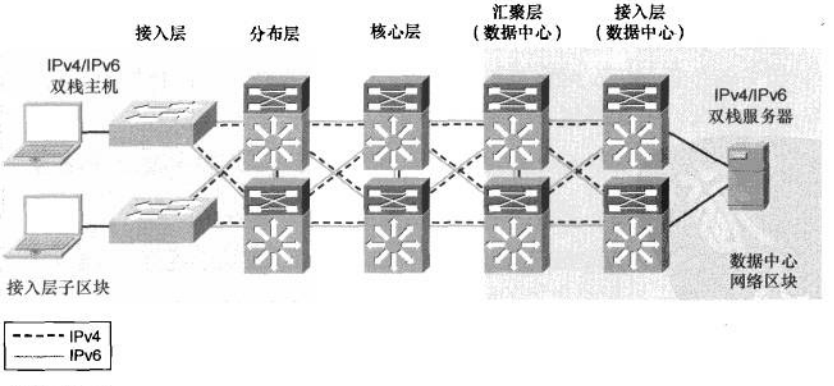


图 6-1 双栈模型示例

注意

图 6-1 示出数据中心网络区块只是为了起演示作用，本章不会对在该网络区块中部署 IPv6 进行讨论，本书第 9 章会对此做深入探讨。

6.1.4 DSM试验用网络部件

表6-1列出了搭建图6-1所示DSM网络拓扑，所使用的网络部件。

表6-1 DSM试验用网络部件

企业园区网络区块	硬件	软件
接入层	Cisco Catalyst 3750E/3560E Catalyst4500 Supervisor 6-E Catalyst 6500 Supervisor 32 或 720	12.2(46)SE 12.2(46)SG 12.2(33)SXI
主机设备	各种便携机：PC 或 Apple	Microsoft Windows XP、 Windows Vista、Windows 7、 Apple Mac OS X 以及 Red Hat Enterprise Linux WS
分布层	Catalyst 4500 Supervisor 6-E Catalyst 6500 Supervisor 32 或 720	12.2(46)SG 12.2(33)SXI
核心层	Catalyst 6500 Supervisor 720	12.2(33)SXI

6.1.5 混合模型

混合模型（HM）的部署策略是：采用两种或两种以上独立的过渡机制，以达成IPv6部署的设计目标。灵活性是混合模型的最大优点，采用混合模型部署IPv6时，可利用多种过渡机制相结合的方法，以最合理的方式搭建特定的网络环境。

采用混合模型部署IPv6，还能更好地利用上现有网络基础设施的功能。可根据多种标准（比如，网元的IPv6硬件转发能力、主机数、应用程序类型、IPv6服务的位置，以及网络基础设施和末端主机操作系统的特性对各种过渡机制的支持情况）来选择相应的过渡机制。

采用混合模型部署IPv6时，可充分利用以下三种主要的IPv6过渡机制。

- **双栈机制**：部署两种协议栈：IPv4和IPv6。
- **站点内自动隧道地址协议（ISATAP）**：依赖于现有IPv4网络基础设施的主机到路由器的隧道机制。
- **手动配置隧道**：依赖于现有IPv4网络基础设施的路由器到路由器的隧道机制。

有了HM，即便底层网络基础设施可能并不支持IPv6，但主机仍能访问到IPv6服务。

对于园区网络区块，在网络的分布层设备不支持或未启用IPv6的情况下，

让位于网络接入层的主机仍然可以访问到 IPv6 服务，是 HM 的价值所在。在这样的网络环境中，分布层交换机通常都担当接入层设备的第一跳三层网关。如果现有的分布层交换机并不支持 IPv6，那么接入层主机便获取不到 IPv6 地址（IPv6 状态化配置或 DHCP）和路由器地址的信息，从而无法访问园区网络区块中已启用了 IPv6 的其他子区块。

可在已启用 IPv6 功能的主机上创建隧道，来访问位于分布层之外的 IPv6 网络服务。在网络的接入层，可采用 HM 模型，充分利用主机上运行的 ISATAP 机制，来实现 IPv6 寻址和线外路由（off-link routing）。位于网络接入层且运行 Microsoft Windows XP、Vista 和 Windows 7 OS 的主机需启用 IPv6 功能，而且还要静态配置 ISATAP 路由器的地址或为其创建 DNS“A”记录。

注意 相关配置细节请见本章后面的“实施混合模式”一节。

图 6-2 所示为采用 HM 时的基本连通性流程图。

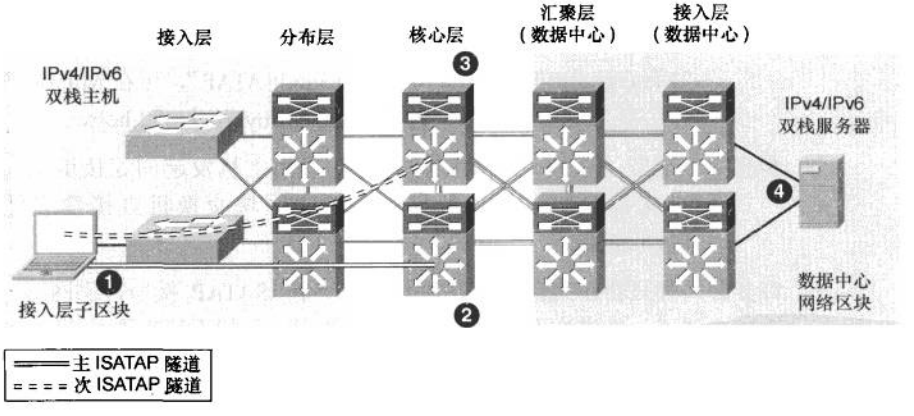


图 6-2 混合模式——连通性流程图

以下列出了图 6-2 所示的 HM 连通性流程图的 4 个步骤。

步骤 1 主机建立通往核心层的 ISATAP 隧道。

步骤 2 在核心层交换机上已配妥了 ISATAP 隧道接口。核心层交换机作为由主机建立的 ISATAP 隧道的终结点。如读者所见，这种隧道创建模式其实并不合理，原因是核心层网络设计目标追求的是功能上的简化，即让核心层设备去尽快地转发数据包；因此，在生产网络中，

核心层交换机并不适合作为ISATAP隧道的终结点。

步骤3 为使得ISATAP隧道具备高可用性，核心层交换机都是成对部署，并配置为冗余对的形式，去接受主机发起的ISATAP隧道连接。在两台核心层交换机上分别创建共享同一IPv4地址的loopback接口，来实现ISATAP隧道的冗余。两台交换机都将冗余（重复）的IPv4地址用作为ISATAP隧道的源地址。主机在连接IPv4 ISATAP隧道源地址时，会连接到两台核心层交换机中的一台（既可以用来实现隧道流量的负载均衡，也可以为此配置优先级，让主机优先连接到其中的一台交换机）。若其中的一台核心交换机故障，IPv4 IGP会迅速收敛，另一台具有相同IPv4 ISATAP地址的交换机会“变”为主用交换机。故障切换过程所花费的时间等于：IGP的收敛时间外加邻居不可达检测（NUD）的到期时间。更多与NUD有关的信息请见RFC 4861“Neighbor Discovery for IP version 6 （IPv6）”。Microsoft Windows Vista和Windows 7在配置中还实现了ISATAP隧道路由器（对于本例，即核心路由器）的流量负载均衡。更多与Microsoft Windows平台上ISATAP实现相关的信息请见白皮书“Manageable Transition to IPv6 using ISATAP”，可在Microsoft下载中心获取这份白皮书，链接为http://tinyurl.com/2jhdbw。

步骤4 配置了双协议栈的服务器会接受由主机发起的连接请求，并通过支持双协议栈的数据中心网络区块交换机直接建立外出的IPv6连接。

许多公司都只会在某台网络设备上配置一个ISATAP接口，供网络内的所有主机建立隧道。虽说这不会影响到正常的使用，但却无法实现基于IPv4源地址，对流量的访问控制。此外，如本书第3章所述，ISATAP隧道会创建一个巨大的平面型网络，这有可能并非人们所愿，尤其是对于用户数较多的企业。

为了控制ISATAP隧道终结之所在，以及主机通过ISATAP隧道能访问什么样的IPv6资源，可采用VLAN（或IPv4子网）到ISATAP隧道的“映射”技术。要是当前的网络设计方案满足以下条件：其一，接入层交换机上的端口与特定的VLAN相关联；其二，连接到接入层交换机上的用户根据自己所隶属的VLAN获取IPv4地址，那么便可以在VLAN与IPv6前缀和ISATAP隧道之间进行类似的映射（匹配）。

图6-3所示为特定VLAN中的用户和IPv4子网与特定ISATAP隧道的映射

（匹配）过程。

下面是对图 6-3 所示的 ISATAP 隧道映射和编号图标的解释。

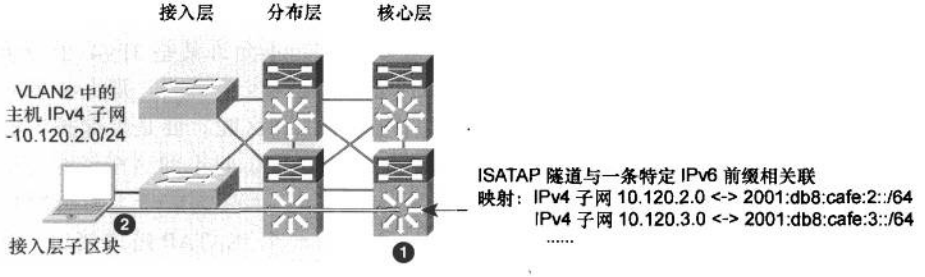


图 6-3 混合模式——ISATAP 隧道映射

步骤 1 在核心层交换机上创建 loopback 接口，并为其配置 IP 地址 10.122.10.2，该地址为 ISATAP 隧道源地址，且仅供位于 VLAN 2（即 10.120.2.0/24 子网）中的用户使用。

步骤 2 接入层的主机连接到隶属于特定 VLAN 的端口。对于本例，即为 VLAN 2。配置 DHCP 服务器，为 VLAN 2 中的主机设定地址范围（10.120.2.0/24）。

在主机上，也要执行与 ISATAP 隧道相关的配置，需静态设定 ISATAP 路由器的 IP 地址 10.122.10.2。静态配置 ISATAP 路由器 IP 地址的方法也有好几种。可在主机上通过命令行来设置 ISATAP 路由器的 IP 地址（**netsh interface ipv6 isatap set router 10.122.10.2**——这会在本章后面的“隧道配置”一节中细谈。），既可以手工输入命令，也可以通过脚本的方式（比如，可通过 Microsoft PowerShell、Microsoft SMS 服务器、组策略[Group Policy]以及若干种其他脚本方法）来配置。使用脚本配置方式时，脚本会通过检测主机现有的 IPv4 地址，来确定如何设置 ISATAP 路由器的 IP 地址。比方说，脚本会首先对主机现配的 IPv4 地址加以分析，并能够确定地址 10.120.2.x/24 中的“2”表示子网（VLAN）号。然后，脚本会在命令应用之前“确定”ISATAP 路由器的地址 10.122.10.2，其中“2”表示子网号或 VLAN 号。地址 10.122.10.2 实际上是核心层交换机上的一个 loopback 接口地址，交换机会将其用作为 ISATAP 隧道的端点。

注意 相关配置细节请见本章后面的“实施混合模式”一节。

出于以下原因，企业可能希望使用分离的 ISATAP 隧道 loopback 接口（即不同 VLAN 的用户，利用 ISATAP 路由器上的不同 loopback 接口建立隧道）。

- **控制和隔离**：若现有的安全策略规定，不允许某些 IPv4 子网访问特定的资源，并已经应用 ACL 执行了相关安全策略，那么在忽略该安全策略且实施了 HM 的情况下，会发生什么呢？要是受保护的资源可经由 IPv6 访问，那么先前通过 IPv4 协议访问不到这些资源的用户，现在可通过 IPv6 协议“成其美事”了。若要为成千上万个用户配置 ISATAP，且只在核心层设备上创建了一个 ISATAP 隧道接口，再要想通过 ACL 来控制源地址，则会既不好管理，又难于扩展。但要是像 VLAN 和 IP 子网那样，以逻辑方式用一条条 ISATAP 隧道来分隔用户的话，那么便可以部署 ACL，根据 IPv6 的源地址、源/目的地址，或甚至是第四层信息，去“轻松写意”地允许或拒绝用户对相关资源的访问。
- **可扩展性**：多年以来，对园区网络区块所属各 VLAN 内的主机数严加控制，一直是网络工程师们所奉行的铁律。其原因不外是要加强对广播域的控制。尽管 IPv6 和 ISATAP 隧道都不使用广播，但仍需考虑其可扩展性。这包括网络设备封装和解封装数据包对控制平面的影响，以及经过隧道封装的流量能否被网络设备线速转发等。利好消息是，配备了 Supervisor 720 和 Supervisor 32G 的 Cisco Catalyst 6500 交换机可以以硬件的方式来转发经过 ISATAP 隧道封装的流量。

以下列出了使用 HM 部署 IPv6 时的需求。

- 主机的操作系统必须支持 IPv6 和 ISATAP。
- 核心层交换机必须支持 IPv6/IPv4 dual-stack 和 ISATAP。

如前所述，在企业园区网络区块内，可结合使用多种过渡机制来建立 IPv6 连通性，比方说，在使用 HM 部署 IPv6 时，还可结合使用以下两种机制。

- 若接入层主机所使用的操作系统不止一种，比如，还有 Linux、FreeBSD、Sun Solaris 以及 MAC OS X 等，那么可采用 6to4 隧道来替代 ISATAP 隧道。
- 可选择在网络的其他层，比如，数据中心区块的汇聚层，而非核心层终结 6to4 隧道。

注意 本章并不会讨论 6to4 隧道和非核心层隧道终结机制，该机制只是使用 HM 部署 IPv6 时建议采用的后备选项。使用 6to4 隧道时，无论是设计方案，还是对安全性和设备扩展性的需求都会发生改变，因此所要考虑的因素也不相同。

6.1.6 混合模型的优缺点

HM 最主要的优势是，可充分利用现成的网络设备，无需升级，尤其是分布层交换机。要是当前在用的分布层交换机既能提供令人满意的 IPv4 服务，而又不存在性能方面的问题，且仍处于折旧期内，那么采用 HM 部署 IPv6 就比较“经济”。

弄清 HM 的缺点也非常重要。

- ISATAP 隧道不支持多播。
- 在网络核心层终结 ISATAP 隧道，使得核心层变为了 IPv6 流量的接入层。网络核心层的设计目标是稳定、简单、快速，而尽快地转发流量才是网络核心层最基本的功能。让网络核心层去干终结 ISATAP 隧道这样的“细活”（Adding a new level of intelligence to the core layer）可能不会被某些网络架构师接受。
- 要想更具粒度的在 VLAN 或 IP 子网与 ISATAP 隧道之间进行映射，配置起来非常繁琐，而且还不好维护。读者应该在整个企业网络只使用一个 ISATAP 隧道接口，与上面提到的“映射粒度”之间找到平衡。使用 HM 部署 IPv6 的目标是，在维护起来不太麻烦的情况下，部署足量的 ISATAP 隧道，实施流量隔离，以满足需求。为缓解采用 HM 所带来的网络运维方面的不便，应在网络中积极推行双栈机制。

任何网络设计方案，只要采用了隧道技术，就必须对性能、管理、安全性、可扩展性、路径 MTU 和高可用性等方面倍加关注。隧道技术只能作为 DSM 设计之后的备选方案。

6.1.7 HM 拓扑结构

图 6-4 所示为一个高度简化的园区网络区块 HM 拓扑结构。ISATAP 隧道由接入层交换机中的主机建立，终结于核心层交换机。图中的实线表示主用隧道。若终结主用隧道的核心层交换机故障，主机将会重新建立通往

另一台核心层交换机的隧道。本章稍后将会讨论ISATAP隧道的高可用性问题，还会围绕图6-4所示的网络拓扑，展开对IPv6 DM部署模型配置方面的深入讨论。

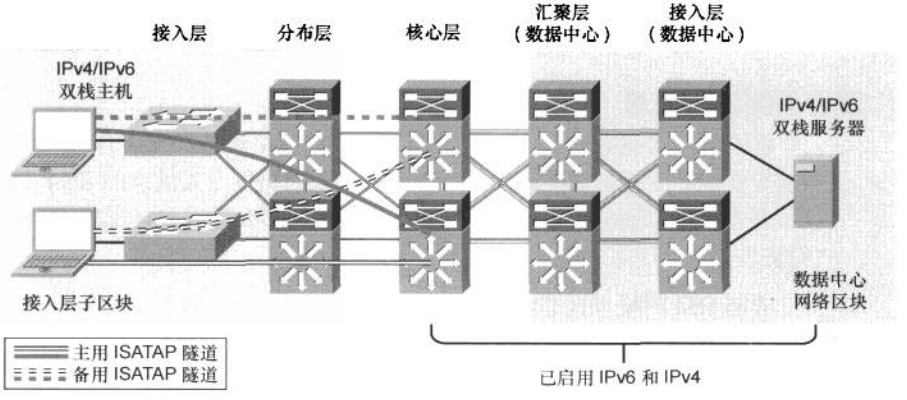


图6-4 混合模型拓扑结构

6.1.8 HM试验用网络部件

以下列出了搭建图6-4所示HM网络拓扑使用的网络部件。

- 园区网络区块：核心层。
- 硬件：Catalyst 6500 Supervisor 720。
- 软件：12.2(33)SXI。

请注意，只有终结ISATAP隧道的园区网络区块核心层Cisco Catalyst交换机，以及数据中心网络区块中运行双协议栈的Cisco Catalyst交换机需开启IPv6功能。因此，园区网络区块接入层和分布层Cisco Catalyst交换机的软件版本与本设计方案无关。

6.1.9 服务区块模型

服务区块模型（SBM）与本章讨论的其他两种模型截然不同。虽然服务区块的设计理念并不算新颖，但对于面临严峻挑战，需在短期内提供IPv6服务的企业来说，SBM却能以独到的方式为其解忧。服务区块的理念也曾为其他的网络设计领域所采用，比如，Cisco网络虚拟化，对于Cisco网络虚拟化解决技术来说，该理念被称为服务边缘。我们之所以说SBM与HM和DM不同，是因

为就部署方式而言，前者被归类于覆盖型网络部署，所以 SBM 一旦投入部署不但不会对现有 IPv4 网络产生影响，而且其实施部署的方式还属于完全集中化式的操作。这样，便可在保证 IPv6 的高可用性和 QoS 功能落实到位，并严格限制对 IPv6 资源访问的情况下，快速搭建起这一覆盖型网络，与此同时，只需对现有 IPv4 网络稍作（甚至不做）改动。

随着园区网络区块对 IPv6 的支持，SBM 也会慢慢“瓦解”。汇接进 SBM 的连接也会从隧道连接（ISATAP 和/或手工配置的隧道）逐渐“演变”为双栈连接。当园区网络区块的各个角落都支持 IPv6 时，便可以拆除 SMB 区块，让其设备另作它用。

需要两台配备了 Supervisor 32 或 Supervisor 720 的 Catalyst 6500 交换机来实施 SBM 的部署，这两台交换机应配置为冗余对。要是在 IPv6 部署的试点阶段，或在实验室/小型生产网络环境中部署 SBM，使用 Cisco ISR 或其他基于软件转发 IPv6 数据包的 Cisco 路由器，作为隧道的终结设备也无不可。

采用 SBM 部署 IPv6 时，保持其高度可扩展性和冗余性的关键是：需确保使用高性能交换机、Supervisor 及交换模块，去处理整个园区网区块中的 ISATAP 隧道流量、手工配置的隧道流量以及双栈连接流量的负载。随着隧道数量的增加，吞吐量势必也会增加，因此，有可能需要在 SBM 中部署额外的交换机（成对部署）去分担流量负载。

本章介绍的 SBM 示例与混合模型也有几分相似之处。底层的 IPv4 网络同样作为正待部署的覆盖型 IPv6 网络的根基；在园区网络区块接入层中的主机也要建立 ISATAP 隧道。以上两点都与混合模型相似。为了能够让园区网络区块接入层中的主机通过 IPv6，访问到位于数据中心网络区块接入层的应用程序和服务，还需在数据中心网络区块的汇聚层交换机和 SBM 交换机之间手动建立隧道。在园区网络区块核心层交换机与 SBM 交换机之间需配置 IPv4 IGP，以保证 IPv6-in-IPv4 隧道的底层 IPv4 连通性。在本章所讨论的示例中，我们分析的是一种极端的情况——在整个园区网络模块内（接入层、分布层和核心层）都未启用 IPv6。本章所举示例中的 SMB 交换机通过冗余的高速链路直连核心层交换机。

6.1.10 SBM 的优缺点

笼统说来，实施 SBM 的好处有：一，能够加快向主机交付 IPv6 服务的步伐；二，几乎不需要改变现有网络的配置（既不需要配置隧道的终结，也

不会增加太多的命令条数）；三，可以灵活地控制对支持 IPv6 的应用程序的访问。

就本质而言，只要能够有条不紊地做好以下工作，SMB 便能在 IPv6 服务新上线时，合理地控制其部署进度。

- 以每用户或每 VLAN 为基础，配置 ISATAP 隧道，控制数据流连接，并利用诸如 NetFlow 之类的工具监控网络设备的特定接口，以及特定源/目的地址对的流量，来起到测量 IPv6 流量使用情况的目的[①]。
- 通过在 SBM 交换机上配置 ACL 和/或路由策略，实施以每服务器或每应用程序为基础的访问控制。访问控制级别可以是开放对一种、若干或甚至是多种 IPv6 服务的访问，与此同时，网络中的其他服务则维持原样（只能通过 IPv4 访问），直到被升级或替换。这样的设置使得“每服务（per-service）”IPv6 的部署成为可能。
- 让 ISATAP 和手工配置的隧道，以及双栈连接都具备高可用性。
- 可通过两种方法让主机接入提供 IPv6 服务的 ISP：一，新建一条单独的 IPv6 连接，仅供主机传递 IPv6 Internet 流量；二，利用 Internet 边缘模块中现有链路同时传递 IPv4 和 IPv6 流量[②]。
- 实施 SBM 时既不会破坏现有的网络基础设施，也不会中断现有网络中的服务。

无论任何设计，只要依靠隧道机制来提供对网络服务的访问，都有其缺点，这在列举混合模式的缺点时已经提及。SBM 不但存在类似于 HM 的缺点（要建立大量的隧道），而且还需采购额外的设备，而 HM 却不必如此。需部署 SBM 交换机（包括线卡），用来与园区网络区块核心层交换机相连，此外，还要考虑软件和维护成本[③]。

由于 HM 和 SBM 的诸多弊端，故而 Cisco 始终建议采用 DSM 部署 IPv6。

---

① 原文是“Configuring per-user and/or per-VLAN tunnels through ISATAP to control the flow of connections and allow the measurement of IPv6 traffic use by allowing interface specific monitoring and specific source/destination pairing in network monitoring tools such as in NetFlow”。

② 原文是“Allowing hosts access to the IPv6-enabled Internet service provider (ISP) connections, either by allowing a segregated IPv6 connection used only for IPv6-based Internet traffic or by providing links to the existing Internet edge connections that have both IPv4 and IPv6 ISP connections”。

③ 原文是“More switches (the SBM switches) and line cards are needed to connect the SBM and core layer switches, and any maintenance or software required represents additional expenses”。

6.1.11 SBM 网络拓扑

本章会把 SBM 设计方案分为两部分来讨论。图 6-5 和图 6-6 分别示出了 SBM 设计方案中的 ISATAP 隧道部分和手工配置的隧道部分。在园区网络区块中，可结合多种技术来完成 IPv6 的部署，本节所介绍的 SBM 设计方案只是对其中两种技术的结合使用，需根据 IPv6 的设计目标和交换机/主机平台的软硬件功能对这些技术加以区分。

图 6-5 所示为冗余的 ISATAP 隧道，隧道由接入层主机发起，由 SBM 交换机终结。SBM 交换机通过 IPv4 链路直连核心层交换机，从而能够与园区网络模块的其他部分建立起连通性。出于 HA 以及同时执行 IPv4/IPv6 路由选择的考虑，两台 SBM 交换机通过双协议栈链路互连。

图 6-6 所示为冗余的手工配置的隧道，其作用是连接数据中心网络区块的汇聚层和服务区块。在数据中心网络区块中，服务器利用 ISATAP 隧道连接 HM 或 SBM 终结点的情况也非常常见。现在，位于园区网络区块接入层的主机可通过 IPv6 访问到部署于数据中心网络区块接入层的 IPv6 服务了。具体的配置信息请见本章稍后的“实施服务区块模型”一节。

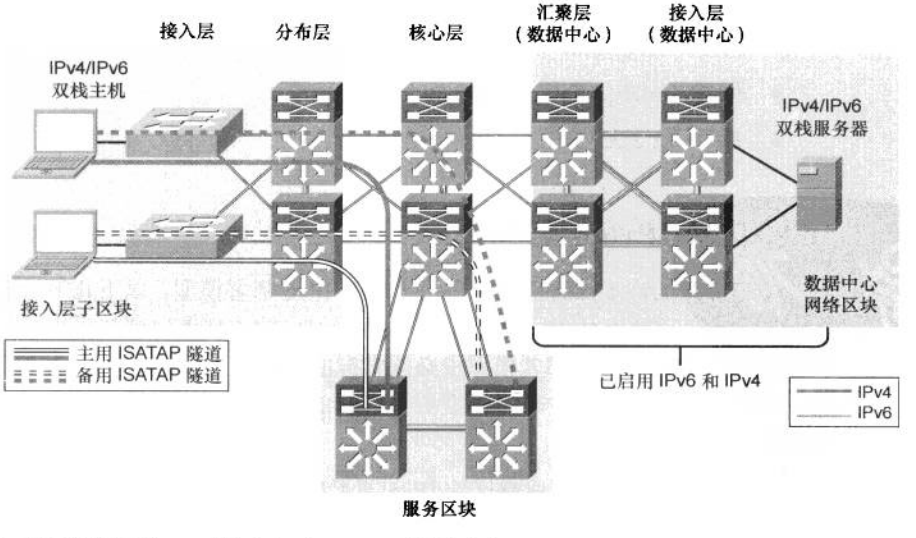


图 6-5 服务区块模型——连接主机（ISATAP 隧道部分）

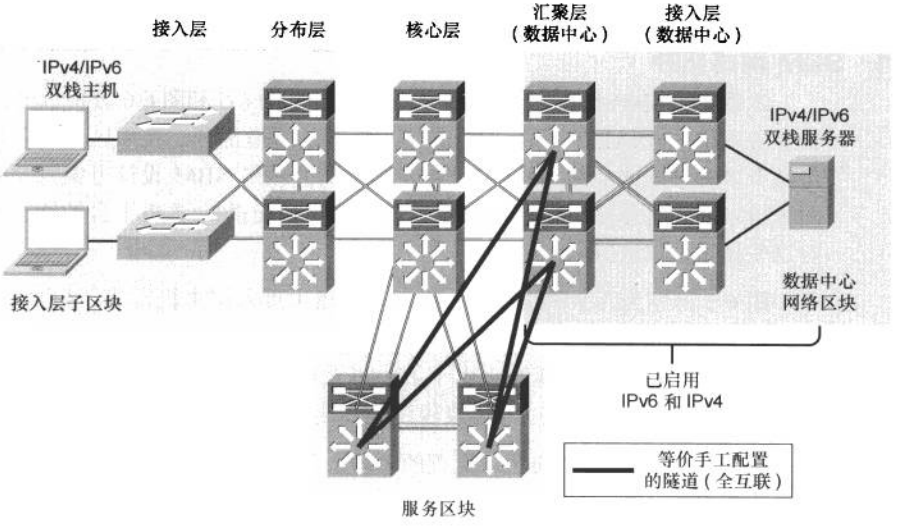


图 6-6 服务区块模型——连接数据中心（手工配置的隧道部分）

6.1.12 SBM 试验用网络部件

以下列出了搭建图 6-5 和图 6-6 所示 SBM 网络拓扑时，所使用的网络部件。

- 园区网络区块：核心层。
- 硬件：Catalyst 6500 Supervisor 720。
- 软件：12.2(33)SXI。

6.2 园区网络区块 IPv6 部署通则

本章介绍的许多部署原则都适用于所有 IPv6 部署模型。以下几节会着重介绍适用于园区网络区块 IPv6 部署的一般性原则，此类部署原则与采用哪种部署模型无关。对于在特定部署模型中必须掌握的特殊原则，我们也会随文一并指出。此外，对于任何一种部署模型所专有的部署原则，其配置可见于相应模型的实施一节。

本章所述的所有园区网络区块 IPv6 部署模型都会充分“尊重”现有的网络设计方案，并以此为基础，来为 IPv6 的部署提供物理接入、IPv4 路由选择（创建隧道时，用来支撑隧道源目之间的 IPv4 连通性）功能、QoS（用来保证隧道

流量的服务质量）、基础设施的安全性（保护隧道）以及高可用性（设备、链路、trunk 和路由选择的高可用性）。[①]采用双栈模型时，刊载于 Cisco 园区网络设计最佳做法文档中的绝大多数设计原则对 IPv4 和 IPv6 全都适用。

在按本章所述的园区网络区块部署模型部署 IPv6 之前，请务必对 Cisco 推荐的园区网络设计最佳做法烂熟于心。可去 Cisco 官网下载 Cisco 园区网络设计最佳做法文档，URL 为 http://www.cisco.com/en/US/netsol/ns815/networking_solutions_program_home.html。

6.2.1 编址

如前所述，本章并不会介绍 IPv6 编址的基本概念。但是，本章会重点讨论为网络设备（尤其是网络设备上的接口）分配 IPv6 地址的原则。

表 6-2 所列为针对网络设备上的链路分配 IPv6 地址时，如何选用不同的前缀长度。

表 6-2 前缀长度的指定原则

64 位	长于 64 位	短于 64 位
RFC 5375 和 IAB/IESG 推荐	可节约地址空间	可让每个广播域拥有更多的主机
既易于管理，又能保证地址分配的一致性	特殊情况： - /126 - p2p 链路的有效地址 - /127 – 只要不与任意播地址产生冲突，也可作为 p2p 链路有效地址 - /128 – Loopback 接口地址	不合理的做法
想启用 SLAAC、SEND 以及隐私扩展特性，就必须使用/64 位的前缀	增加了网络管理的复杂性	主机数太多，会让网络介质“难以招架”
地址空间损失严重	应竭力避免某些特殊地址所产生的地址冲突问题： - 路由器任意播地址 (RFC 3513) - 嵌入的 RP 地址 (RFC 3956)	没有使用该选项的任何理由

以下是对表 6-2 的说明。

- **/64**：在交换机下连用户 VLAN 的 SVI 接口上，推荐使用/64 位的前缀，这么做既易于管理，又能保证地址分配的一致性。要想启用无状态地址

[①] 原文是“All campus IPv6 models discussed in this chapter leverage the existing campus network design as the foundation for providing physical access, VLANs, IPv4 routing (for tunnels), QoS (for tunnels), infrastructure security (protecting the tunnels), and availability (device, link, trunk, and routing)”。

自动配置（SLAAC）、安全的邻居发现（SEND）以及隐私扩展特性，就必须使用/64位的前缀。

- **长于64位**：为P2P网络分配/64的前缀是对IPv6地址资源的浪费，这也是网络社团中诸多人士所持有的观点。关于应该为P2P链路分配何种前缀的争论仍在激烈地进行，但读者在分配IPv6地址时，既需尊重RFC制定的条条框框，也需考虑实际部署中的具体情况。在许多IPv6部署中，比较常见的做法是为VLAN（或主机所驻留的链路）分配/64的前缀；为P2P链路分配/126的前缀；要是能确保不与任意播地址冲突，也可以为P2P链路分配/127的前缀（写作本书之际，已有安全使用/127前缀的IETF草案“Using 127-bit IPv6 Prefixes on Inter-Router Links”问世）；为loopback接口分配/128的前缀。
- **短于64位**：一般而言，在任意一条链路上，其主机数都不可能多于/64的前缀所能提供的地址数，因此为一条链路分配短于64位的前缀是一种很烂的做法。

RFC 3627“Use of /127 Prefix Length Between Routers Considered Harmful”论述了使用/127的前缀有可能与某些专用地址产生冲突的原因所在。

6.2.2 物理连通性

除了以下几点之外，IPv6和IPv4物理连通性的部署原则均可以通用。

- **确保充足的带宽可供IPv4和IPv6流量同时使用**：这对任何新协议、技术或应用程序的部署来说都是重要因素。
- **弄清IPv6处理链路MTU（最大传输单元）的方式**：本书并不会介绍IPv6协议的基本操作和规范。读者可通过阅读RFC 2460“Internet Protocol, Version 6 (IPv6) Specification”和RFC 1981“Path MTU Discovery for IP version 6”，来了解IPv6 MTU和路径MTU发现（PMTUD）机制。
- **WLAN（无线LAN）上的IPv6操作**：IPv6在L2交换机上的操作方式应该与其在WLAN访问点上大致相同。不过，读者仍需对WLAN环境中的IPv6操作细节加以关注，主要包括通过IPv6管理WLAN设备（访问点[AP]和控制器），以及通过AP或基于控制器的QoS、VLAN和ACL对IPv6流量的控制。要想充分利用上WLAN设备上的上述智能服务，AP和/或控制器设备必须支持IPv6功能。

除了上面提到的几点以外，Cisco 还建议对主机和网络设备的内存、CPU 利用率以及现有流量的概况进行全面分析。此外，在按本章所讨论的 IPv6 部署模型实施 IPv6 部署之前，需完成 SLA（服务等级协定）的签订。

6.2.3 VLAN

IPv6 的 VLAN 划分原则与 IPv4 相同。采用双栈模型时，IPv6 和 IPv4 流量会穿越同一 VLAN。采用隧道模型时，IPv4 流量以及被封装进 IPv4 数据包的 IPv6 流量则会穿越同一 VLAN。

在通过 Trunk 与语音 VLAN 一并透传的数据 VLAN（IP 电话后）上，可完全支持 IPv6 功能。取决于数据和语音 VLAN 的配置，网络工程师有可能会遇到这样的问题：交换机第三层数据 VLAN 接口发出的第二层多播路由器通告可能会被附接的主机（连接到 IP 电话或启用了语音 VLAN 的交换机端口的主机）接收。大致说来，一台连接到 IP 电话的主机可能会从数据和语言 VLAN 收到 RA 消息。显而易见，这是一个与厂商以及交换机/电话版本无关的问题。[1]至于如何妥善处理这一第三层 VLAN 配置方面的问题，或通过 Cisco IP 电话中可能的设置来防止在数据 VLAN 上泛洪 IPv6 路由器通告，请对最新的 Cisco 最佳做法保持关注。

欲了解目前 Cisco 对 VLAN 设计给出的建议，请参考本章最后一节“参考资料”中所列出的 Cisco 园区网设计最佳做法文档。

6.2.4 路由选择

在园区网络区块中选择 IGP，受多种因素的影响，其中包括软硬件平台、IT 人员的专业技能、网络拓扑及规模等。对于本章所举的网络环境示例，IPv4 的路由选择协议将会采用 EIGRP，但也会用到 OSPFv2。本章会给出 EIGRP 或 OSPFv3 的 IPv6 配置。在某些内容中会对 IGP 进行互换，以向读者显示任意一种 IGP 的基本配置特征。

就上述 IGP 的运作方式而言，IPv4 和 IPv6 之间有诸多相似之处，最终的行

[1] 该段文字的原文是“IPv6 on data VLANs that are trunked along with voice VLANs (behind IP phones) is fully supported. Depending on the configuration of the data and voice VLANs, you might experience an issue where the Layer 2 multicast router advertisements from the Layer 3 data VLAN interface might bleed over to the attached host (connected to the IP phone or a voice VLAN–enabled switch port). Basically, a PC connected to an IP phone can receive RAs for both the data and voice VLANs. This is an issue that can be manifest regardless of vendor and switch/phone version”。译者不知作者所云，作者也未解释 IPv6 RA 消息在数据 VLAN 上泛洪有何坏处，因此只能按字面意思直译。

为与结果也大体相同。经验表明，在大多数 IPv6 的部署中，都会采用与 IPv4 相同的 IGP 作为 IPv6 的 IGP。这是因为网管人员已有过相应 IGP 的使用经验了，因此学习难度也会降低不少。然而，在某些情况下，对于某些 IPv6 的部署来说，也可以适时采用与 IPv4 不同的 IGP。由于这是关乎 IPv6 的白手起家式的部署（Greenfield deployment），因此网管人员需要弄清是不是真的有必要针对 IPv4 和 IPv6 分别运行不同的 IGP。或许读者会偏执于让 IPv4 和 IPv6 使用同一种 IGP，但只有更合理地汇总了路由，或真正利用上了路由协议内置的某项特定功能（比如，分发列表）时，对网络的高效运转才会起到推动作用。[①]底线是所选择的技术（IGP）要管用，网管人员还得熟悉；如果业务或技术部门的头头脑脑决意改变，就听命行事好了。

如前所述，本书已尽其所能地介绍了如何根据最佳做法，去实施 Cisco 园区网络设计。应尽可能地根据当前的最佳做法，去调优 IPv4 和 IPv6 所采用的 IGP。调优 IGP，令其能够保证网络的稳定性和可扩展性，并能够达成快速收敛，应成为任何网络设计的首要任务之一。

6.2.5 高可用性

本章并不会介绍高可用性（HA）的方方面面。对最新 Cisco 园区网络设计最佳做法的充分理解，是满足 HA 设计需求和建议的前提。以下列出了本章所涉 HA 的主要内容。

- **冗余的路由和转发路径**：在需要搭建隧道的情况下，为构建冗余的隧道路径，要靠 IPv4 的 EIGRP 或 OSPFv2，外加 Cisco 特快转发功能来实现；在启用双栈机制的情况下，则要靠 IPv6 的 EIGRP 或 OSPFv3，外加 Cisco 特快转发功能来实现。
- **用来终结 ISATAP 和手工配置隧道的冗余的第三层交换机**：在 HM 和 SBM 设计中，会用到此类交换机。除了硬件方面的冗余之外，实现隧道（ISATAP 和手工配置的隧道）的冗余也同样重要。对于 HM 和 SBM 设计的配置，以及采用这两种模式部署 IPv4 时使用冗余隧道的效果，都将会在本章的后面几节加以说明。
- **第一跳网关的高可用性**：对于 DSM 设计，分布层交换机会成为接入层

[①] 原文是“Perhaps you stick with the same IGP but do a better job of summarization or use different functions within the IGP such as a distribution list—or whatever makes sense in your network.”译文与原文不同，给出原文，以免误导读者。

主机的第一跳三层设备。传统的园区网络模块设计均普遍采用 HSRP 或 GLBP 作为第一跳冗余协议。本章会展示用于 IPv6 的 HSRP 的配置。第一跳冗余协议（比如，HSRP）的配置将会在本章稍后的“实施双栈模型”一节中进行讨论[①]。

6.2.6 QoS

采用 DSM 部署方案，能够充分利用现成的 IPv4 QoS 策略，并可很容易地对其进行扩展，以令其作用于穿越园区网络区块的新的 IPv6 流量。将 QoS 策略与应用程序或服务，而不是与网络协议（IPv6 或 IPv4）紧密结合，是实施 QoS 的通则，这也是 Cisco 推荐的做法。只要现有的 QoS 策略是针对应用程序来进行精细化的分类、监管以及排队，那么相应的 QoS 策略就应该能够对相关应用程序的 IPv4 和 IPv6 流量“一视同仁”。

应专门制定 QoS 策略，用来处理被隧道封装的 IPv6 流量。针对经过 ISATAP 隧道封装的 IPv6 流量执行 QoS 策略时，还应考虑到某些方面的限制。在网络中启用 ISATAP 隧道时，不应在接入层入站方向对 IPv6 数据包执行分类，接入层是信任或分类入站流量的理想所在[②]。而对于按 HM 和 SBM 设计的园区网网络区块，接入层根本不会有 IPv6 流量（因为不支持 IPv6）。隧道是建立在接入层主机和核心层交换机（HM）或 SBM 交换机之间，因此无法在接入层执行入站流量的分类操作。

对隧道流量解封装之后，可针对 IPv6 流量施以 QoS 策略，但这又给网络工程师提出了挑战。在抵达隧道端点之前，经过隧道封装的 IPv6 流量不可能被分类，这是因为只有对隧道流量解封装之后，才能对入站流量执行标记操作（入站流量的分类和标记操作应在物理接口而不是在隧道接口完成）。在终结隧道的交换机上，可针对任何已被解封装且由其转发的 IPv6 流量执行出站方向的 QoS 分类策略。于是，下游交换机便可以执行信任、监管和排队策略，以对 IPv6 流量做相应处理。

---

① 原文是“In this chapter, configurations are shown with HSRP for IPv6. The configuration for first-hop routing protocols such as HSRP are discussed in the section “Implementing the Dual-Stack Model,” later in this chapter”。首先，关于展示 HSRP 配置的内容，原文重复了两遍；其次，原文居然把 HSRP 当做第一跳路由协议。译者选择直译原文。

② 原文是“When ISATAP tunnels are used, the ingress classification of IPv6 packets cannot be made at the access layer, which is the recommended location for trusting or classifying ingress traffic”。前半句才说“不要在接入层分类入站 IPv6 数据包”，后半句却说“推荐在接入层对入站流量执行信任和分类操作”，难道 IPv6 数据包不算“流量”？其实，译到现在，译者已经发现原书有许多地方都难以自圆其说，受篇幅所限，译者只能指出其中的一小部分。

图 6-7 所示为在 HM 设计中采用 ISATAP 隧道时，可对 IPv6 流量实施 QoS 策略的位置。在图 6-7 中运行双协议栈的链路上，同时针对 IPv4 和 IPv6 流量应用了 QoS 策略。欲知更多有关在园区网络中实施 QoS 的信息，请参阅本章最后一节“参考资料”中所列出的 Cisco 园区网络 QoS 文档。

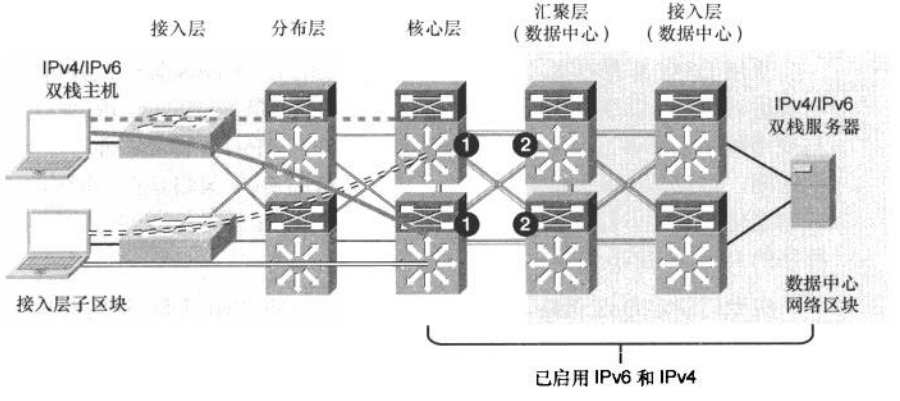


图 6-7 实施 QoS-混合模型

下面是对图 6-7 所示两个步骤的解释。

步骤 1 对于 HM，核心层交换机 IPv6 流量的外出接口是实施 QoS 分类和标记操作的首要位置。如前所述，IPv6 流量是接入层主机以隧道封装的方式发送到核心层交换机，在核心层交换机对隧道流量解封装之前，IPv6 数据包根本不可见。因为没有办法针对被隧道封装的流量执行入站方向的 IPv6 QoS 策略，因此核心层交换机 IPv6 流量的外出接口便成为了实施 QoS 策略的首要位置[①]。

步骤 2 现在，下游交换机（例如，数据中心区块汇聚层交换机）可对经过分类并被标记的 IPv6 数据包进行检查了，然后，在 IPv6 数据包的入站方向（接口）执行相应的 QoS 策略。QoS 策略可包括信任（入站接口）、监管（入站接口）以及排队（出站接口）。

图 6-8 所示为在 SBM 设计中采用 ISATAP 和手工配置的隧道时，针对 IPv6 流量，应用 QoS 策略的位置。

① 原文是“Because QoS policies for classification and marking cannot be applied to the ISATAP tunnels on ingress, the first place to apply the policy is on egress”。请问作者 ingress 和 egress 具体所指何处？

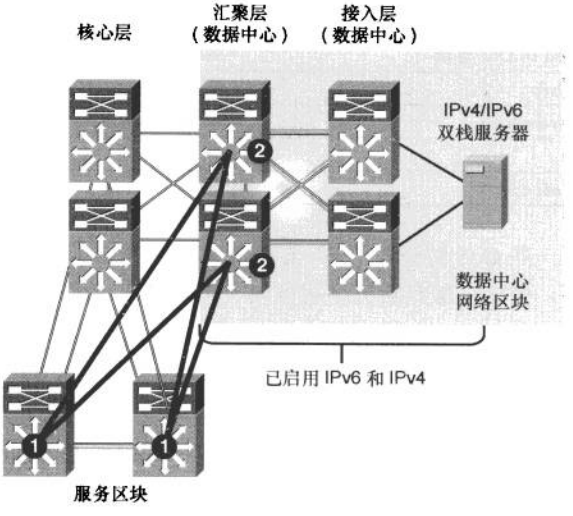


图 6-8 实施 QoS 策略-SBM（ISATAP 和手工配置的隧道）

下面是对图 6-8 所示的两个步骤的解释。

步骤 1 SBM 交换机从 ISATAP 隧道接口收到刚被解封装的 IPv6 数据包，然后再在手工配置隧道接口的出站方向上，针对 IPv6 数据包应用分类和标记策略。

步骤 2 于是，在数据中心网络区块的汇聚层，当 IPv6 数据包离开手工配置的隧道接口之后，下游交换机（数据中心网络区块汇聚和接入层交换机）可针对 IPv6 流量应用信任、监控以及排队策略。

注意 写作本书之际，Catalyst 6500 Supervisor 32/720 还不支持对 IPv6 多播数据包的微流监管。对于 SBM 设计，到本章截稿时为止，只能在交换机接口的入站方向（隧道接口除外）对 IPv6 数据包实施监管操作。欲知更多与 Catalyst 6500 平台 QoS 有关的信息，请参阅 PFC3 QoS 文档，URL 为 http://tinyurl.com/2ewv7mp。

为支持 IPv6，Cisco 在模块化 QoS CLI（MQC）命令中的 QoS **match** 和 **set** 语句中移除了 **ip** 关键字。经过修改的 QoS 命令语法可同时支持 IPv4 和 IPv6 数据包的匹配和设置操作，如表 6-3 所示。

表 6-3 QoS 配置命令方面的改变-IPv4 和 IPv6

只支持 IPv4 的 QoS 语法	同时支持 IPv4/IPv6 的 QoS 语法
match ip dscp	match dscp
match ip precedence	match precedence
set ip dscp	set dscp
set ip precedence	set precedence

还有一些 QoS 特性对 IPv4 和 IPv6 同时适用，但无需对命令行接口（CLI）进行修改，例如，加权随机早期检测（WRED）、监管、加权轮询（WRR）等。

在本章讲述实施三种 IPv6 部署模型（DSM、DM 以及 SBM）的各节中，并未对分类应用程序、关联映射 DSCP（区分服务代码点）值，以及带宽和排队方面的建议等与 QoS 有关的配置，展开深入讨论。Cisco 提供了一整套部署 QoS 的建议，请见 Cisco 官网以及 Cisco Press 图书《End-to-End QoS Network Design》。

欲了解更多与 Cisco 园区网络 QoS 设计方面的信息，请查阅本章最后一节“参考资料”中所列出的 Cisco 相关文档和 Cisco Press 书籍。

6.2.7 安全性

现有 IPv4 园区网络中常见的威胁和攻击，在 IPv6 网络中也同样存在。未经授权的访问、欺骗攻击、路由攻击、病毒/蠕虫、拒绝服务（DoS）攻击以及中间人攻击，只不过是能对 IPv4 和 IPv6 形成威胁的少数几种攻击而已。

许多对 IPv4 生效的潜在的攻击手段根本威胁不到，或至少不会以相同的方式威胁到 IPv6[①]。IPv6 有自己的一套邻居发现和路由器通告机制，此外，IPv6 的报头结构甚至对分片数据包的处理方式都与 IPv4 不同[②]（为了识别、弄清并缓解针对 IPv6 的安全威胁，无论是 Cisco 公司还是整个业界，都付出了许多努力。本章会指出园区网络区块中可能存在并有待解决的安全威胁，并会提供保护 IPv6 双栈和隧道流量[③]的基本示例。本章稍后的“实施双栈模型”一节中会讨论实施双栈部署时的第一跳安全性。

---

① 原文是“With IPv6, many threat possibilities do not apply or at least do not apply in the same way as with IPv4”。

② 译者注：原文是“There are inherent differences in how IPv6 handles neighbor and router advertisement and discovery, headers, and even fragmentation. ”原文真是差到了极点，译者改写了原文。

③ 原文是“protection for IPv6 dual-stack and tunneled traffic”。作者的写作方式就是一味图省事。精确的说法应该是“当网络运行双协议栈时，保护 IPv6 流量；否则，保护被 IPv4 隧道封装的 IPv6 流量。”写作技术书籍和创作文学作品或绘画不同，不能讲究“留白”。这就是说，对于技术方面的描述要不厌其烦，一定要把意思说透，但作者离这方面差的太远。

注意

本节所举示例绝非实战部署中的金科玉律，其目的是要让读者明白：在将安全策略应用于自己所维护的网络时，必须权谋再三，谋而后定。欲了解关乎 IPv6 安全的详细信息请阅读 Cisco Press 图书《IPv6 Security》。

以下各节描述了与网络设备防护有关的一般性安全准则，这些准则适用于所有园区网络设计模型。

通过复杂的地址分配，让“踩点”更加困难

在对园区网络区块中的网络设备编址时，需精心规划。安全专家们普遍认为，要想猜到以 ULA（全局唯一的单播局部地址）方式编址的交换机 IPv6 地址值非常困难。试举一个交换机接口 ID 的编址示例：假如 VLAN2 和 VLAN3 的 SVI 接口 IP 地址分别为 2001:db8:cafe:2::1/64 和 2001:db8:cafe:3::1/64，那么：：1 就是交换机的接口 ID。这样的编址示例不但世人皆知，而且很容易猜测，攻击者可借此迅速了解园区网络区块中网络设备的通用编址方案。另一种选择是，以随机化的方式为园区网络中网络设备的接口 ID 分配地址。还是以前例中的 VLAN2 和 VLAN3 为例，可使用诸如 2001:db8:cafe:2::a010:f1a1 和 2001:db8:cafe:3::c801:167a 之类的地址为 VLAN2 和 VLAN3 的 SVI 接口构造地址，其中“a010:f1a1”为交换机 VLAN2 SVI 接口的接口 ID。

本节所描述的编址方案给运维管理提出了严峻的挑战。为了简化对网络设备和编址的运维管理工作，网络工程师应通过半随机化的地址分配方案（semirandomized addresses），在网络设备接口 ID 的安全性和管理网络设备的方便性之间做出折中性的选择。

控制对园区网络交换机的管理访问

为严格限制通过 IPv6 对特定交换机的访问，可针对 line vty 设置 ACL，来放行特定源地址对交换机管理接口（loopback 接口）的访问[①]。应该把允许访问交换机的源地址设置为分配给企业网的 IPv6 前缀。对于全网范围的网络设备，为了使得包含在 ACL 中的 ACE 的创建更具可扩展性，定义 ACL 时，可在交换机的特定接口上以“permit”源为整个企业网网络前缀的方法为主，来控制对网络设备的访问，尽量不要使用先“deny”特定地址，再“permit any any”的方

① 原文是“To more tightly restrict access to a particular switch through IPv6, an ACL is used to permit access to the management interface (line vty) by way of the loopback interface”。译文并未按照原文翻译，请读者留意。

法。本例中的企业网站点所使用的IPv6前缀是2001:db8:cafe::/48。

例6-1所示为如何构造基本的ACL来严格限制对交换机VTY的访问。

例6-1 line VTY访问控制列表

```
ipv6 access-list MGMT-IN

 remark Permit MGMT only to Loopback0
 permit tcp 2001:DB8:CAFE::/48 host 2001:DB8:CAFE:6507::A111:1010

 deny ipv6 any any log-input

!
line vty 0 4

 session-timeout 3
 access-class MGMT-IN-v4 in
 password 7 08334D400E1C17
 ipv6 access-class MGMT-IN in

 logging synchronous
 login local
 exec prompt timestamp
 transport input ssh
```

对IPv6来说，运行SNMP（简单网络管理协议）时的安全需求等同于IPv4。如需使用SNMP，应先确定其版本，然后再考虑访问控制及选择认证/加密的方法。在本章所举的园区网络设计模型示例中，将使用NMPv3 (AuthNoPriv)为数据中心网络区块中的Cisco NMS服务器，提供轮询功能。

例6-2所示为本章中园区网络区块交换机所使用的配置示例。

例6-2 SNMP配置

```
snmp-server contact John Doe - ipv6rocks@example.com
snmp-server group IPv6-ADMIN v3 auth write v1default
snmp-server user jdoe IPv6-ADMIN v3 auth md5 cisco1234
snmp-server host 2001:DB8:CAFE:100::60 version 3 auth jdoe
```

写作本书之际，Cisco Catalyst交换机还不支持使用IPv6 HTTP ACL去控制对交换机自身的访问。记住这一点非常重要，因为目前支持**ip http access-class** ACL命令的交换机只能对使用IPv4对自身的HTTP访问进行控制，但对IPv6无能为力。这一IPv4和IPv6功能上的差异即意味着：此前通过IPv4

HTTP/HTTPS 访问不到交换机的用户，现在可通过 IPv6 来访问交换机了。除非必须通过 HTTP/HTTPS 对交换机直接访问，否则推荐的做法是在交换机上禁用 HTTP/ HTTPS 服务。

IPv6 流量监管

可将流量监管视为安全和/或 QoS 功能。企业园区网络中，很有可能会存在以聚合流或以每用户微流（aggregate or per-user microflow）为基础（即通过源/目的地址对和/或源/目的端口对），来实施流量监管的需求。对于本章所讨论的园区网络区块设计模型，有某些地方适合 IPv6 流量监管，尤其是实施每用户微流监管。

- **DSM**：在数据中心网络区块分布层 Catalyst 6500 交换机上，针对入站流量执行 IPv6 流量的每用户微流监管（理想位置）。
- **HM**：在数据中心网络区块汇聚层 Catalyst 6500 交换机上，针对入站流量（来自于园区网络区块接入层中的主机）执行 IPv6 流量的每用户微流监管。这一监控位置并不理想。首选位置应该是在数据中心网络区块核心层交换机上，针对入站流量执行微流监管，但对于 HM 模型，入站流量监管无法应用于隧道接口，故而只能在放在数据中心汇聚层交换机上执行。
- **SBM**：在本章所讨论的 SBM 特例中，执行 IPv6 流量的每用户微流监管可算是一种挑战。对于 SBM 模型，服务区块交换机为配备了 PFC3 卡的 Catalyst 6500。此类交换机支持入站方向的每用户微流监管，但目前还不支持出站方向的 IPv6 每用户微流监管。对于本章的 SBM 示例，IPv6 流量会在服务区块交换机上的 ISATAP 隧道和手工配置的隧道之间穿梭往来。由于入站流量监管不能应用于 ISATAP 隧道接口和手工配置的隧道接口之上，因此在服务区块中，无论什么位置都不适合执行流量监管。

例 6-3 所示为一个实施 IPv6 每用户微流监管的示例。对于本例，已在下游交换机上配置了 QoS 策略，令其匹配 IPv6 流量，并根据 Cisco 推荐的 QoS 策略配置之一为 IPv6 流量设置 DSCP 值。此外，还对该交换机进行了配置（如下所示），令其以每用户流（per-user flow）为基础（根据 IPv6 源地址），去执行流量监管。受到监管的流量被限速为 5Mbit/s，超出设定模板（profile）的流量将会被丢弃。

例 6-3 微流监管配置示例

```
mls qos
!
class-map match-all POLICE-MARK
  match access-group name V6-POLICE-MARK
!
policy-map IPv6-ACCESS
  class POLICE-MARK
     police flow mask src-only 5000000 8000 conform-action transmit exceed-
     action drop
  class class-default
   set dscp default
!
ipv6 access-list V6-POLICE-MARK
 permit ipv6 any any
!
interface GigabitEthernet3/1
 mls qos trust dscp
 service-policy input IPv6-ACCESS
```

写作本书之际，对 Catalyst 6500 Supervisor 32 和 720 来说，还不支持以硬件方式同时启用 IPv6 每用户微流监管和 IPv6 多播路由功能。但以上两种功能单独启用时，Supervisor 则能够以硬件方式来支持了。若在交换机上已经配置了 **ipv6 multicast-routing** 命令，然后再应用 IPv6 每用户微流监管策略，交换机就会返回系统消息，表明其正以软件方式交换 IPv6 数据包。反之，要是在某接口上应用了 IPv6 每用户微流监管策略，再激活 **ipv6 multicast-routing** 命令，那么同样的告警信息也会出现，如下所示。

```
006256: *Aug 31 08:23:22.426 mst: %FM_EARL7-2-
IPV6_PORT_QOS_MCAST_FLOWMASK_CONFLICT: IPv6 QoS Micro-flow policing configuration on port
GigabitEthernet3/1 conflicts for flowmask with IPv6 multicast hardware forwarding on SVI
interface Vlan2, IPv6 traffic on the SVI interface may be switched in software
006257: *Aug 31 08:23:22.430 mst: %FM_EARL7-4-FEAT_QOS_FLOWMASK_CONFLICT: Features
configured on interface Vlan2 conflict for flowmask with QoS configuration
on switch port GigabitEthernet3/1, traffic may be switched in software
```

在本章的最后一节“参考资料”中可找到更多与微流监管有关的信息。

利用控制平面监管（CoPP）

在本章所讨论的园区网络区块中，CoPP 只对 Catalyst 6500 Supervisor 32/720 生效。CoPP 可保护安装在交换机上的多功能交换特性卡（MSFC），令其免受 DoS 攻击和恶意流量对资源的占用。在网络中，重要的控制/管理平面流量都被

赋予了最高优先级。配备了 PFC3 的 Catalyst 6500 支持 IPv6 流量的 CoPP。CoPP 的配置取决于各种各样的因素，在部署建议方面可谓百花齐放，因为具体的 CoPP 策略都要以个案为基础来量身定制。

在本章的最后一节“参考资料”中可找到更多与 CoPP 有关的信息。

控制来自接入层的入站流量

可针对 IPv6 流量的源前缀（地址）来执行过滤。通常，都会在分布层交换机（DSM）SVI 接口的入站方向执行此类过滤，但对于 HM 或 SBM，也可在 ISATAP 隧道接口的入站方向来完成过滤。根据源前缀来控制 IPv6 流量，可有助于防范最基本的网络欺骗攻击。

例 6-4 所示的 ACL 只放行源为隶属于某特定 VLAN 的 IPv6 前缀的流量。

例 6-4 基本的作用于入站方向的 IPv6 ACL

```
ipv6 access-list VLAN2-v6-INGRESS
remark PERMIT ICMPv6 PACKETS FROM HOSTS WITH PREFIX 2001:DB8:CAFE:2::/64
permit icmp 2001:DB8:CAFE:2::/64 any
remark PERMIT IPv6 PACKETS FROM HOSTS WITH PREFIX 2001:DB8:CAFE:2::64
permit ipv6 2001:DB8:CAFE:2::/64 any
remark PERMIT ALL ICMPv6 PACKETS SOURCED BY HOSTS USING THE LINK-LOCAL PREFIX
permit icmp FE80::/10 any
remark DENY ALL OTHER IPv6 PACKETS AND LOG
deny ipv6 any any log-input
!
interface Vlan2
 ipv6 traffic-filter VLAN2-v6-INGRESS in
```

注意 Cisco IOS IPv6 ACL “暗伏”有放行 IPv6 邻居发现流量的“permit”条目。要是在 ACL 中配置有 **deny ipv6 any any**，那么将会覆盖“暗伏”的放行邻居发现流量的条目。若出于日志记录的目的，手动配置了 **deny ipv6 any any log-input** 语句（记录被 deny 掉的流量日志信息），那么必须在该语句之上添加以下两条语句。

```
permit icmp any any nd-na
permit icmp any any nd-ns
```

例 6-4 中名为 VLAN2-v6-INGRESS 的 ACL 包含了一条更为宽泛的条目（**permit icmp FE80::/10 any**），以放行邻居发现以及与其他任何 ICMPv6 服务有关的流

量，VLAN2中的主机或路由器必须依靠此类流量来彼此发现。

讨论应该放行或阻挡何种ICMP流量的RFC、草案以及IPv6部署方面的数据可谓汗牛充栋。欲了解更多有关IPv6数据包过滤的信息，请参阅本章最后一节“参考资料”所列出的IETF官网链接和Cisco Press书籍。

第一跳安全特性

当前，为防范无赖DHCP服务器、中间人攻击，以及其他在接入层针对IPv4的攻击，Cisco研发出了多种安全功能来加以应对。上述攻击矢量同样会对IPv6构成威胁，但无论是业界还是设备厂商，都缺乏相应的应对措施。Cisco将上述安全功能统一编组，并命名为“第一跳安全特性”。RA（通告）保护功能是Cisco正在开发的若干“第一跳安全特性”之一。RA保护功能能够在VLAN/链路上有效防范无赖RA（无赖路由器或主机“号称”自己是网关）。在园区网络区块的接入层，Cisco的许多安全特性（比如，安全邻居发现（SEND）、RA保护以及其他正在开发的新安全特性）都对主机或网络设备的安全防护工作大有裨益。RFC 3791“Secure Neighbor Discovery”所载的安全解决方案以及RA保护功能都是在IETF的“培育”下所结出的硕果。IETF应该对RA保护功能以及其他与IPv6功能性有关的标准的状态进行记录。与RA保护、端口ACL以及其他安全特性有关的额外信息和配置案例仅见本章篇末的“参考资料”一节。

阻止Microsoft Teredo的使用

Teredo的作用是：提供对NAT（网络地址转换）网关身后主机的IPv6的支持。网管人员必须知晓的是，Teredo会给网络带来许多安全隐患。在园区网络区块中，除非出台了针对Teredo的最佳做法（具备高等级的安全性），否则读者应确保在Microsoft Windows SP2及更高版本、Windows Vista和Windows 7中禁用Teredo。作为一种后备预防手段，读者可能还需考虑配置ACL（ACL可配置在接入层或更靠近上游的网络设备上，比如，边界路由器上[①]来阻断目的UDP端口号为3544的流量，以防止内网主机建立通往园区网络外部的Teredo隧道。更多与Teredo有关的信息请见本章篇末的“参考资料”一节。该节还列出了许多与IPv6安全相关的资源。

---

① 原文是“which can be done at the access layer or further upstream, such as at the border routers”。在译者看来，接入层的上一层应该是分布层，请问分布层怎么可能会是边界路由器。退一步来说，即便是园区网络区块中的接入层，也不能连接边界路由器，边界路由器都被部署于边缘网络区块。

6.2.8 多播

对任何企业网络设计来说，IPv6 多播都是其中的一项重要服务[①]，IPv6 多播的部署需求可能会让读者重新审视本章所讨论的三种 IPv6 部署模型。在三种 IPv6 部署模型与 IPv6 多播结合使用的情况下，有一个重要问题需要读者关注——ISATAP 隧道不支持多播流量的传播。这并非是软硬件方面的限制，而是 ISATAP 隧道机制本身的缺陷（请参见 RFC 5214 “Intra-Site Automatic Tunnel Addressing Protocol [ISATAP]” 的 6.2 节。）。

在企业园区网络区块中部署 IPv6 多播时，如何有效控制接入层中的主机和多播组之间的对应关系，是网络工程师需要考虑的最重要因素之一。IPv6 中 MLD（多播侦听者发现）协议所起的作用，等价于 IPv4 中的 IGMP（Internet 组管理协议）协议。两者都被用来控制多播组成员的关系。MLD Snooping 是一种在第二层交换机上启用的特性，用来控制交换机，令其只将多播流量发布到连接了多播侦听者的端口。要是没有该特性，连接了一个多播接收者的接入层交换机会将多播流量泛洪至其所有端口。在接入层，交换机的 MLD Snooping 特性能否支持 MLD 版本 1 和/或版本 2，可谓至关重要。请注意，上述需求只适用于接入层/分布层运行双栈或纯 IPv6 协议的部署。要是主机以搭建隧道的方式，让被封装的 IPv6 流量穿越接入层和分布层，那么 IPv6 多播流量并不会以其“本来面目”被主机所接收，因此接入层交换机也就无需支持 MLD Snooping 特性了。

注意　诸多 Linux 和 BSD 实现以及 Microsoft Windows Vista 和 Windows 7 都支持 MLDv2，这对于 PIM-SSM 的多播部署非常重要。对许多 IPv6 多播部署来说，采用 PIM-SSM 的同时，在接入层启用 MLDv2，可算是一种极佳的设计选择。Cisco IOS 开发出了一项名为 SSM 映射的特性，可将 MLDv1 报告消息映射为 MLDv2 报告消息，以供 PIM-SSM 使用。欲了解更多与 SSM（特定源多播）有关的信息，请见如下 URL（也可参考 Cisco Press 图书《Deploying IPv6 Networks》）：

http://www.cisco.com/en/US/docs/ios/ipv6/configuration/guide/ip6multicast_ps10591_TSD_Products_Configuration_Guide_Chapter.html#wp1058805.

---

① 原文是“IPv6 multicast is an important service for any enterprise network design”。原文的语法虽然没有问题，但从常识方面来讲根本就讲不通，IPv6 多播怎么可能会是“网络设计”中的“服务”呢？只能说“对于企业网络来说，IPv6 多播都是在其内运行的一项重要服务”，但是译文还是按字面原文的字面意思翻译。

在本章所举的示例中，只有在DSM的IPv6部署中才能见到支持多播的应用程序，这是因为DSM无需使用（不支持多播的）ISATAP隧道。在DSM示例中，测试用的支持多播的应用程序是Windows Media Services和VideoLAN Media Client (VLC)，这两种多播应用都使用嵌入式聚合点（嵌入式RP）和协议无关模式-特定源（PIM-SSM）多播组。多播源是位于数据中心网络区块中的服务器，分别运行Microsoft Windows Server 2003、Windows Server 2008和Red Hat 5.0操作系统。

本书的第3章已经讨论过了基本的IPv6多播概念。此外，Cisco.com上以及业界也有若干文档深入细致的介绍了多播。在本章与IPv6多播有关的配置示例中，除了启用基本IPv6多播功能的一般性命令参考以外，不会再提供其他方面的配置说明。欲知更多与多播有关的信息，请参见Cisco IPv6多播网页，链接为：http://www.cisco.com/en/US/products/ps6594/products_ios_protocol_group_home.html。

6.2.9 网络管理

当今，许多在用的传统网络管理工具都支持IPv6。但大多数网络管理工具和设备对IPv6的支持还任重道远。本章只会介绍与园区网络区块网络管理有关的基本网络管理服务（Telnet、SSH以及SNMP）。最新版本的IOS软件大都支持通过IPv6来传输SNMP消息，不过这还要取决于Catalyst交换机平台。请参考交换机硬件文档，以了解特定的交换机平台是否支持通过IPv6传输SNMP消息。要是读者所用的Catalyst交换机平台还不支持IPv6上的SNMP消息传输机制，则可利用IPv4上的SNMP消息传输机制来获取交换机上与IPv6相关的MIB/中断（trap）消息/通知（inform）消息。本书的第11章将会深入讨论IPv6网络管理。

6.2.10 地址分配

地址分配管理是网络工程师必须掌握的网络管理技能。一般情况下，为网络设备配置长达128位，以16进制表示的IPv6地址时，应尽量遵循自动化和简单化的原则。

为遵循上述原则，可利用通用前缀特性（general prefix feature），在园区网络模块内的交换机上配置地址前缀。该特性可让用户在交换机的全局配置模式下定义一条或多条前缀，并为这些前缀指定具有实际意义的名称。然后，可在接口配置模式下，调用该名称以替换接口上常用的IPv6前缀。例6-5所示为通

用前缀特性的使用方法。

例 6-5 配置通用前缀特性

```
!Define the General Prefix
6k-agg-1(config)#ipv6 general-prefix DC-1 2001:DB8:CAFE::/48

!Configure the general prefix named "DC-1" on a per-interface basis:
6k-agg-1(config-if)#ipv6 address DC-1 ::10:0:0:F1A1:6500/64

!Verify that the general prefix was correctly assigned to the interface:
6k-agg-1#show ipv6 interface vlan 10
Vlan10 is up, line protocol is up
  IPv6 is enabled, link-local address is FE80::211:BCFF:FEC0:C800
  Description: VLAN-SERVERFARM-WEB
  Global unicast address(es):
    2001:DB8:CAFE:10::F1A1:6500, subnet is 2001:DB8:CAFE:10::/64
```

提示 通用前缀特性在需要对接口重新编址时非常有用，这是因为只需改变通用前缀值便能快速更改整台交换机或路由器（所有相关接口）的地址前缀值。

更多与通用前缀特性有关的信息请见 Cisco 官网上的 Cisco IOS IPv6 文档页面（链接请见本章篇末的“参考资料”一节）。

可使用 SLAAC、DHCPv6 或静态指定等地址分配方式，为园区网络区块中的主机分配 IPv6 地址。DHCPv6 不但是一种最主要的地址分配方法，而且还是绝大多数网管人员都希望采用的地址分配方法。在 Catalyst 交换机支持 DHCPv6 中继代理特性之前，对园区网络区块接入层的主机来说，除了采用 SLAAC 作为分配 IPv6 地址的主要手段之外，还没有其他选择。如今，随着对 DHCPv6 中继代理特性的支持，网管人员在管理和分配 IPv6 地址时，已可按照 IPv4 那样“照方抓药”。可根据设计需求，在整个网络中同时启用 SLAAC 和 DHCPv6。比方说，可利用 SLAAC 为隶属于语音 VLAN 的 IP 电话分配 IPv6 地址，而隶属于数据 vlan 的 PC 则利用 DHCPv6 获取地址。

图 6-9 所示为园区网络区块内 DHCPv6 中继设备（园区网络区块分布层交换机）的布放位置，IPv4 所使用的 IP helper 特性也在该设备上启用。

DHCPv6 中继特性配置起来相当简单。例 6-6 所示为在面向接入层主机的 SVI 接口上 DHCPv6 中继特性的配置。

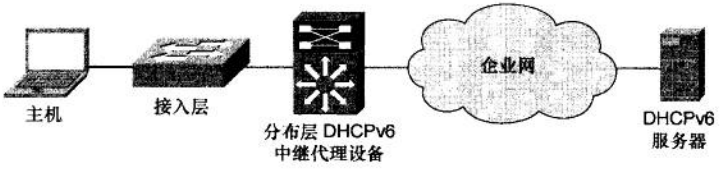


图 6-9 园区网络区块中 DHCPv6 中继特性的部署

例 6-6 DHCPv6 Relay Configuration

```
interface Vlan2
 description ACCESS-DATA-2
 ipv6 address 2001:DB8:CAFE:2::A111:1010/64
 ipv6 nd prefix 2001:DB8:CAFE:2::/64 no-advertise
 ipv6 nd managed-config-flag
 ipv6 dhcp relay destination 2001:DB8:CAFE:102::9
```

例 6-6 中的命令 **ipv6 dhcp relay destination** 指明了 DHCPv6 服务器的单播 IPv6 地址。命令 **ipv6 nd managed-config-flag** 用来在 RA 消息中设置“managed address configuration（管理地址配置）”标记，借此标记，主机便可获知：将要采用诸如 DHCPv6 等状态化地址配置机制。欲了解更多与 DHCPv6 中继代理有关的信息，请访问以下链接：

http://tinyurl.com/24kb5qy

在网络中，网络设备支持 DHCPv6 中继代理是一回事，主机能否支持 DHCPv6（比如，Microsoft Windows 7），服务器能否提供 DHCPv6 服务又是另外一回事。当前，有以下三种经过 Cisco 测试，可在园区网络中提供 DHCPv6 服务的 DHCPv6 服务器。

- **Cisco Network Registrar:** http://www.cisco.com/en/US/products/sw/netmgtsw/ps1982
- **运行于 IOS 的 Cisco DHCPv6 服务器:** http://tinyurl.com/24kb5qy
- **Microsoft Windows Server 2008:** http://technet.microsoft.com/enus/library/cc896553%28WS.10%29.aspx

Cisco 同样支持通过各种网络管理工具对只启用了 IPv6 的网络设备进行管理，管理的范围包括 DNS、DHCPv6、设备硬件管理以及对设备的监控等，此

外，还可利用这些网络管理工具对网络进行管理、故障排除和日志记录等[①]。欲了解各种 Cisco 网络管理解决方案的详细信息，请浏览 Cisco 官网“Network Management and Automation（自动化网络管理）”页面，连接为：

http://www.cisco.com/en/US/products/sw/netmgtsw/index.html.

6.2.11 性能和可扩展性

本章对性能和可扩展性的讨论主要侧重于在园区网络模块中规划和部署 IPv6 时，相对于特定交换机平台的通用部署原则。本章也并不准备针对各种试验用交换机平台，展开性能和可扩展性方面的分析。

一般而言，网管人员应该对现有网络中的链路、内存以及 CPU 的使用率了如指掌。如果其中的任何一项指标都居高不下，那么在现有网络中添加新技术、新特性或新协议（IPv6）便很有可能导致流量中断。然而，在实施 IPv6 的过程中，可以很容易地感受到园区网络区块中设备间链路带宽利用率方面的变化。随着 IPv6 的部署，用户会纷纷采用 IPv6 来访问他们之前用 IPv4 来访问的应用程序，因此链路上 IPv4 流量所占用带宽的比重通常会很快下降。但从整体来看，设备间链路带宽利用率却会呈上升趋势，这主要是拜 IPv6 路由选择所产生的控制流量所赐，要是在园区网中还搭建了 ISATAP 或手工配置的隧道，由于隧道封装而产生的字节开销也会增加链路的带宽利用率。

以下各节将会分别针对采用 DSM、HM 和 SBM 部署 IPv6 时的可扩展性和性能方面的考虑展开讨论。

采用 DSM 部署 IPv6 时的可扩展性和性能方面的考虑

本节将会讨论采用 DSM 部署 IPv6 时，对园区网络区块中的接入层、分布层以及核心层的可扩展性和性能方面的考虑。

在可扩展性方面，首先需要考虑的是在分布层和接入层交换机上同时运行的 IPv4 和 IPv6 这两种协议。在接入层（第二层接入或路由式接入）或分布层，交换机必须对 IPv4 和 IPv6 的邻居信息保持跟踪[②]。地址解析协议缓存之于 IPv4，

---

① 原文是“Cisco supports the management of IPv6-enabled network devices through a variety of network management products to include DNS, DHCPv6, device management, and monitoring, as well as network management, troubleshooting, and reporting. Chapter 11 discusses IPv6 management”。根据原文字面，很难弄清作者的意图，译者只能勉强翻译。本书的几位作者没有一位能把文字写通！

② 这里是指 IPv4 主机的 ARP 缓存信息和 IPv6 主机的邻居缓存信息。

如同邻居缓存之于IPv6。此处，读者需要知道，对于IPv4的ARP缓存，存储的是一对一的IPv4地址到MAC地址的映射信息。但对于IPv6，在交换机的邻居缓存表中，可能会存在多个IPv6地址（比如，本地链路地址、本地唯一地址以及多个公网地址）到单个MAC地址的映射。以下示例为分布层Catalyst 6500交换机上MAC地址为000d.6084.2c7a的主机的ARP和IPv6邻居缓存表项。该主机在分布层交换机的ARP表和IPv6邻居缓存表中的条目如下所示。

```
Internet 10.120.2.200          2          000d.6084.2c7a ARPA  Vlan2
The IPv6 neighbor cache entry is:
2001:DB8:CAFE:2:2891:1C0C:F52A:9DF1        4 000d.6084.2c7a STALE Vl2
2001:DB8:CAFE:2:7DE5:E2B0:D4DF:97EC        16 000d.6084.2c7a STALE Vl2
FE80::7DE5:E2B0:D4DF:97EC                  16 000d.6084.2c7a STALE Vl2
```

由IPv6邻居缓存表可知，MAC地址为000d.6084.2c7a的主机对应着3个IPv6地址（对于本例，该主机是一台在默认情况下不会使用EUI-64格式作为接口ID的Windows主机）。在IPv6邻居缓存表中，第一个地址是分配给该Windows主机的两个公网IPv6地址之一，这是由IPv6隐私扩展特性所生成的公网IPv6地址。第二个地址是主机通过IPv6无状态自动配置特性获取的另一个公网IPv6地址（该地址既可以手工指定，也可以通过DHCPv6获取），第三个地址是主机生成的本地链路地址。取决于主机所使用的地址类型，用于该主机的IPv6邻居缓存表的条目数既可降至最低一条（IPv6本地链路地址），也可攀升至多条。

在园区网络区块中部署IPv6时，需要考虑的另一个与性能和可扩展性有关的问题是，如何处理IPv6多播。如前所述，在启用IPv6多播时，为确保接入层交换机不在第二层将多播帧向所有端口泛洪，交换机能否支持MLD欺骗（snooping）功能（类似于IPv4的IGMP欺骗特性）可谓至关重要[①]。

除了前面提到的ARP/邻居缓存问题以外，采用DSM在园区网络区块中部署IPv6时，对于分布层交换机来说，还有另外两个性能和可扩展性方面的问题值得关注。

- 能否以硬件方式执行IPv6数据包的转发。
- 能否以硬件方式处理IPv6访问列表表项。在分布层交换机上配置ACL的主要作用包括：供QoS策略调用（针对来自接入层入站方向的数据包，执行分类和标记操作）、保障数据平面的安全性（阻挡来自接入层入站方向的DoS流量、欺骗攻击流量，以及未经授权的访问流量），外加保

[①] 原文是“As mentioned previously, it is important to ensure that MLD Snooping(similar to IGMP Snooping in IPv4) is supported at the access layer when IPv6 multicast is used to ensure that IPv6 multicast frames at Layer 2 are not flooded to all the ports”。翻译这样的文字，首要需要把原文转换为可以理解的语言，然后才轮到中英文之间的转换。

护交换机控制平面的安全性（与 QoS 策略结合使用）[1]。

就可扩展性和性能方面的考虑而言，园区网络区块中的核心层应该与分布层一视同仁。

采用 HM 部署 IPv6 时的可扩展性和性能方面的考虑

采用 HM 部署 IPv6 时，需按以下套路来考虑可扩展性和性能方面的问题。

- **接入层**：采用 HM 部署 IPv6 时，园区网络区块接入层的网络设备不支持 IPv6 功能。由于接入层链路上可能会收发额外的流量（被隧道封装的流量），因此网管人员可能需要考虑链路带宽利用率的问题。然而，如前所述，随着 IPv6 的部署，当用户纷纷采用 IPv6 来访问他们之前用 IPv4 来访问的应用程序时，IPv4 和 IPv6 所占链路带宽的比重势必会发生“逆转”。
- **分布层**：采用 HM 部署 IPv6 时，园区网络区块分布层网络设备在可扩展性和性能方面所受到的影响等同于接入层。
- **核心层**：采用 HM 部署 IPv6 时，园区网络区块核心层交换机可能会终结成百上千条 ISATAP 隧道。企业应该与集成商或 Cisco 相关团队协商，以确定现有的核心交换机能否处理设计所需的隧道数。若核心层交换机无力支撑由接入层主机所发起的隧道数，那么企业只有两种选择：其一，采用 DSM 部署 IPv6；其二，在无法采用 DSM 部署 IPv6 的情况下，弃用 HM，选择 SBM——部署专门的交换机来终结并管理隧道。对核心层网络设备来说，还需要考虑以下两个与可扩展性和性能有关的问题：
- **因 ISATAP 隧道接口的管理问题，而对交换机控制平面所产生的影响**：假设需要为接入层交换机上的每个 VLAN 都建立一条 ISATAP 隧道，那便很有可能会引发这一问题。在大型网络中，这一隧道建立模式会迫使交换机消耗其可观的 CPU 资源来维护大量的隧道。对一般的网络设备来说，虚拟接口控制平面的管理工作都是由 CPU 来执行的[2]。

---

[1] 原文是“IPv6 ACLs in the distribution layer are primarily used for QoS (classification and marking of ingress packets from the access layer), for security (controlling DoS, snooping, and unauthorized access for ingress traffic in the access layer), and for a combination of QoS and security to protect the control plane of the switch from attack”。译者没有按原文翻译，因为原文有违写作常识。

[2] 整段的原文是“Control plane impact for the management of ISATAP tunnel interfaces: This can be an issue if there is a one-to-one mapping between the number of VLANs and the number of ISATAP tunnels. In large networks, this mapping results in a substantial number of tunnels that the CPU must track. The control plane management of virtual interfaces is done by the CPU”。读者需要注意，译者没有按原文的字面意思翻译。肯定会有读者看不懂这段文字，但译者已经尽力了。

- **链路带宽利用率**：在核心层交换机与分布层交换机互连的链路上，带宽利用率会有所增长（拜经由IPv4隧道封装的IPv6流量所赐）；在核心层交换机与数据中心区块汇聚层交换机互连的链路上（已启用了双协议栈），带宽利用率可能会因IPv6流量的增加而有所增长[①]。

采用SBM部署IPv6时的可扩展性和性能方面的考虑

采用SBM部署IPv6时，需按以下套路来考虑可扩展性和性能方面的问题。

- **接入层**：采用SBM部署IPv6时，园区网络区块接入层的网络设备只支持IPv4功能，因此不需要在可扩展性和性能方面做特殊考虑。
- **分布层**：采用SBM部署IPv6时，园区网络区块分布层的网络设备只支持IPv4功能，因此不需要在可扩展性和性能方面做特殊考虑。
- **核心层**：采用SBM部署IPv6时，园区网络区块核心层的网络设备只支持IPv4功能，因此不需要在可扩展性和性能方面做特殊考虑。
- **服务区块**：在HM部署模式中，适用于园区网络区块核心层交换机的部署原则，大多适用于SBM部署中放置在服务区块中的交换机。两者之间唯一的差异是，对于SBM模式，放置在服务区块中的那对交换机既要终结ISATAP隧道，也要终结手工配置的隧道。采用SBM部署IPv6的好处有二：其一，专门部署了一对交换机用来负责隧道的建立和终结；其二，还可以在服务区块中部署更多的交换机来建立和终结更多的隧道。因此，该部署方式非常适合大规模的隧道部署。而对于HM部署模式，拜赐于园区网络区块核心层交换机在连接企业网各网络区块（数据中心区块、WAN区块等）时所扮演的重要角色，因而很难对园区网核心层进行扩展（添加额外的核心层交换机）。

6.3 实施双栈模型

以下各节将重点介绍DSM的配置。我们会将DSM的配置划分为几个方面（比如VLAN功能、路由选择功能和HA功能方面的配置等）来分别加以讲解。

---

① 原文是“There is an increase in link utilization coming from the distribution layer (tunneled traffic) and a possible increase in link utilization by adding IPv6 (now dual-stack) to the links from the core layer to the data center aggregation layers”。

其中某些方面的配置，比如 VLAN 功能和物理接口相关配置，并非专门针对 IPv6。对于 DSM 的 VLAN 配置来说，无论是 IPv4 还是 IPv6 都差异不大，但出于完整性的考虑，还是会分别示出。

注意

配置示例的示出原则：只会示出部署在同层或彼此互连的一对交换机的配置示例；在以下各节中，只会示出与当节所述内容相关的配置示例，比如，与路由选择或 HA 相关的配置示例。

6.3.1 网络拓扑

出现在本节的网络拓扑图用作为所有 DSM 配置示例的参照。图 6-10 所示为 DSM 配置示例所使用的交换机物理端口布局图[①]。

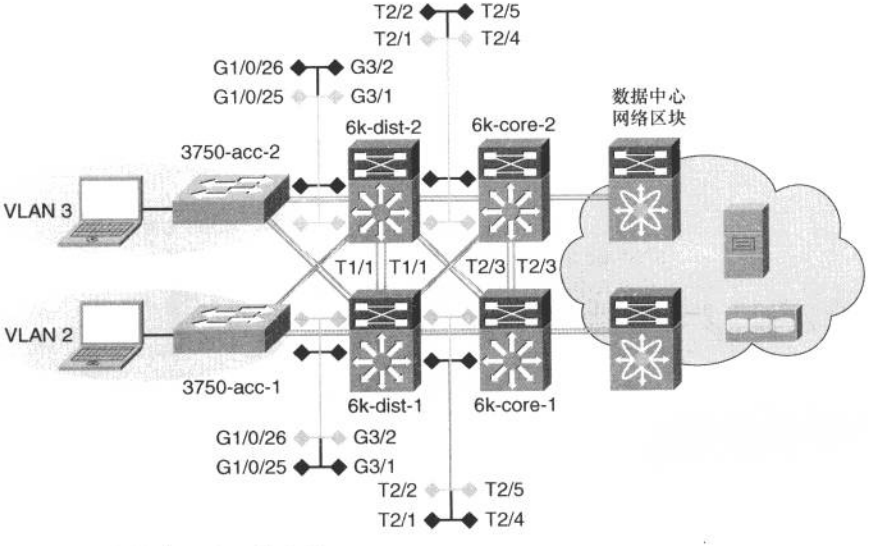


图 6-10 DSM 网络拓扑-物理端口布局图

注意

图 6-10 示出了园区网络区块的详细拓扑，但只给出了数据中心网络区块的示意图，欲了解数据中心网络区块的详细拓扑，请阅读本书的第 9 章。

① 原文是“The diagrams in this section are used as a reference for all DSM configuration examples. Figure 6-10 shows the physical port layout that is used for the DSM”。这种文字没法翻译，好歹和技术无关，译文也只有请读者将就一下了。

图 6-11 所示为 DSM 环境中的 IPv6 编址规划。为使该图更为清晰，所有 IPv6 地址均省略了/48 的网络前缀部分。这一省略的 IPv6/48 网络前缀为 2001:db8:cafe::/48，在本书所介绍的三种 IPv6 部署模型配置示例中，均选择使用这一网络前缀。

注意

请不要把本书用来举例的 IPv6 编址方案视为最佳做法。为方便读者阅读，本书用来举例的编址方案注重的是简单性。读者在生产网络中全面实施 IPv6 编址方案之前，应对 IPv6 地址分配规划细加斟酌。现有的 IPv4 编址规划已经给了人们太多深刻的教训，我们应该从中吸取教训，也许白手起家，重新开始着手 IPv6 的地址规划工作，能给我们以全新的机会合理地分配 IPv6 地址。

除了物理接口以外，还要给交换机的 loopback 接口和 SVI 接口分配 IPv6 地址。表 6-4 所示为各交换机 SVI 和 loopback 接口的 IPv6 地址。

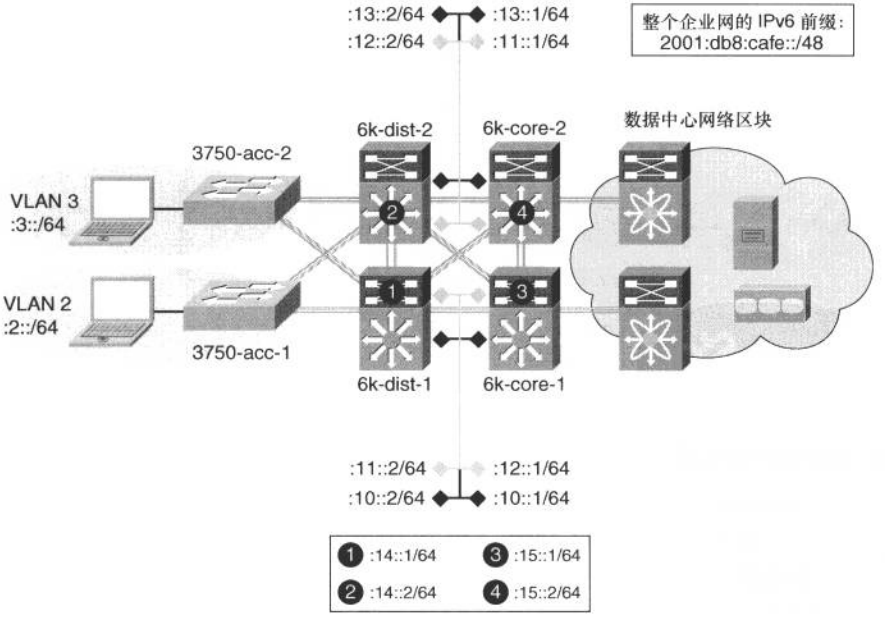


图 6-11 DSM 网络拓扑图-编址规划

表 6-4 交换机 SVI 和 loopback 接口 IPv6 地址

交换机	接口	IPv6 地址
3750-acc-1	loopback	2001:db8:cafe: lf3::5/128
	VLAN2	2001:db8:cafe:2::4/64
3750-acc-1	loopback	2001:db8:cafe: lf3::6/128
	VLAN3	2001:db8:cafe:3::4/64
6k-dist-1	loopback	2001:db8:cafe:lf3::3/128
	VLAN2	2001:db8:cafe:2::2/64
	VLAN3	2001:db8:cafe:3::2/64
6k-dist-2	loopback	2001:db8:cafe:lf3::4/128
	VLAN2	2001:db8:cafe:2::3/64
	VLAN3	2001:db8:cafe:3::3/64
6k-core-1	loopback	2001:db8:cafe:lf3::1/128
6k-core-2	loopback	2001:db8:cafe:lf3::2/128

6.3.2 物理接口和 SVI 接口的配置[①]

就交换机物理接口（点到点链路）方面的 IPv6 配置而言，与 IPv4 的配置大同小异。例 6-7 所示为 6k-dist-1 和 6k-core-1 之间 p2p 链路的物理接口配置。

例 6-7 6k-dist-1 P2P 链路的物理接口配置：

```
ipv6 unicast-routing                    !Globally enable IPv6 unicast routing

ip cef distributed                      !Ensure IP CEF is enabled (req. for IPv6 CEF)
ipv6 cef distributed                    !Globally enable IPv6 CEF.

!
interface TenGigabitEthernet2/1
 description to 6k-core-1
 dampening
 load-interval 30
 carrier-delay msec 0
 ipv6 address 2001:DB8:CAFE:10::2/64        !Assign IPv6 address

 no ipv6 redirects                           !Disable IPv6 redirects
```

配置中包含了 **no ipv6 redirects** 命令。该命令的作用是：禁用已在接口上默

① 本节的原文标题是“Physical/VLAN Configuration”，作者在写作方面只会“偷工减料”！

认启用的发送 ICMP 重定向消息的功能（意在告知主机：另有一台离目的网络更近的路由器可为你转发数据包）。在这条 p2p 链路上，无需启用 IPv6 重定向功能。在网络设备上，禁用不必要的服务，是一种最佳做法，特别是不要忘记禁用那些可能会成为攻击目标的服务。

例 6-8 所示为核心层交换机-分布层交换机之间的链路上与 IPv6 有关的配置。

例 6-8 6k-core-1 P2P 链路的物理接口配置

```
ipv6 unicast-routing
ip cef distributed
ipv6 cef distributed
!
interface TenGigabitEthernet2/4
 description to 6k-dist-1
 dampening
 load-interval 30
 carrier-delay msec 0
 ipv6 address 2001:DB8:CAFE:10::1/64

 no ipv6 redirects
```

例 6-9 所示为 6k-dist-1 SVI 接口（VLAN2）的配置。由配置可知，该交换机有一条 Trunk 链路下连接入层交换机，并透传数据 VLAN（VLAN2）的流量[①]。在本配置示例中，该交换机针对隶属于 VLAN 2 的主机开启了 DHCP 中继特性。带有 **no-advertise** 参数的行命令的作用是：防止链路上不支持 DHCPv6 功能的客户端使用 SLAAC 来获取地址。若网络环境中的接入层主机不全都支持 DHCPv6 功能，但欲让那些不支持 DHCPv6 功能的主机，也能通过其他方法获取 IPv6 地址，那就不要配置这条命令。命令 **managed-configflag** 则用来让交换机告知主机：请使用状态化的地址配置特性（DHCPv6）。

例 6-9 6k-dist-1 SVI（VLAN2）的配置

```
vtp domain ese-dc
vtp mode transparent
!
spanning-tree mode rapid-pvst
```

（待续）

① 原文是“The configuration shows a trunk link to the access layer and a data VLAN (VLAN2)”。译者对作者的这种写法提出强烈抗议！

```
spanning-tree loopguard default
spanning-tree vlan 2-3 priority 24576
!
vlan 2

 name ACCESS-DATA-2

!

interface GigabitEthernet3/1
 description to 3750-acc-1
 switchport
 switchport trunk encapsulation dot1q
 switchport trunk allowed vlan 2

 switchport mode trunk
 switchport nonegotiate
 no ip address
 spanning-tree guard root
!
interface Vlan2

 description ACCESS-DATA-2

 ipv6 address 2001:DB8:CAFE:2::2/64

 ipv6 nd prefix 2001:DB8:CAFE:2::/64 no-advertise !Don't send RA for this prefix

 ipv6 nd managed-config-flag          !Enabled managed address configuration
                                      flag
 ipv6 dhcp relay destination 2001:DB8:CAFE:102::9  !Define DHCPv6 server address

 no ipv6 redirects
```

在 Catalyst 3750 和 3560 交换机上，必须先启用正确的交换机数据库管理（SDM）模板，让交换机的 TCAM（三重可寻址内存）去完成相应的转发看表（forwarding lookups）功能[①]。在 3750-acc-1 和 3750-acc-2 上，已先期执行了 **sdm prefer dual-ipv4-and-ipv6 default** 命令，并完成了“dual-ipv4-and-ipv6”SDM 模板（需要重启设备才能生效）的配置。欲知与 **sdm prefer** 命令，以及与该命令

① 原文是“On the Catalyst 3750 and 3560 switches, you must enable the correct Switch Database Management (SDM) template to allow the ternary content addressable memory (TCAM) to be used for different purposes”。如果读者能看懂这段文字，还需要购买本书吗？译者没有直译原文。

相关联的模板有关的详细信息，请参考以下 URL：

http://tinyurl.com/28qj5lk.

在接入层，每台交换机上只运行 1 个 VLAN，接入层交换机上的管理 VLAN 和语音 VLAN 的配置方法并不在本节的讨论范围之内。接入层交换机之间不会互相透传 VLAN，分布层交换机会终结 VLAN 信息。例 6-10 所示为 3750-acc-1 的配置。

例 6-10 3750-acc-1 VLAN 配置

```
vtp domain ese-dc
vtp mode transparent
!
spanning-tree mode rapid-pvst
spanning-tree loopguard default
spanning-tree portfast bpduguard default
!
vlan 2                                  !VLAN2 - Data VLAN for 3750-acc-1

 name ACCESS-DATA-2

!
interface GigabitEthernet1/0/25
 description TRUNK TO 6k-dist-1
 switchport trunk encapsulation dot1q
 switchport trunk allowed vlan 2

 switchport mode trunk
 switchport nonegotiate
!
interface Vlan2                         !VLAN2 with IPv6 address

 ipv6 address 2001:DB8:CAFE:2::4/64

 no ipv6 redirects
!
interface GigabitEthernet1/0/10
 description TO PC
 switchport access vlan 2

 switchport mode access
 switchport port-security maximum 2
 switchport port-security aging time 20
 spanning-tree portfast
```

虽然在本章所讨论的各种 IPv6 部署模型并不会用到堆叠技术，但在许多其

他网络中的接入层 Catalyst 3750 和 3560 交换机上，却经常用到这项技术。使用堆叠技术时，IPv6 与 IPv4 并无不同。更多与 IPv6 环境中使用交换机堆叠技术有关的信息，请参考 URL: http://tinyurl.com/32324vo。

6.3.3 路由选择的配置

如前所述，本章将会以 EIGRP 为例（同时用于 IPv4 和 IPv6）来讲解采用 DSM 部署 IPv6 时，如何配置路由选择。本节所示的 EIGRP 配置会尽量遵循 Cisco 推荐的园区网络路由选择配置指南。

例 6-11 所示为交换机 6k-dist-1 与 IPv6 有关的 EIGRP 配置。

例 6-11 6k-dist-1 路由选择配置

```
key chain eigrp
 key 100
   key-string 7 1111
!
interface Loopback0
 ip address 10.122.10.9 255.255.255.255          !Address used for RID on EIGRP

 ipv6 address 2001:DB8:CAFE:1F3::3/128

 ipv6 eigrp 10

!
interface TenGigabitEthernet1/1
 description to 6k-dist-2
 ipv6 address 2001:DB8:CAFE:14::1/64

 ipv6 eigrp 10

 ipv6 hello-interval eigrp 10 1

 ipv6 hold-time eigrp 10 3

 ipv6 authentication mode eigrp 10 md5

 ipv6 authentication key-chain eigrp 10 eigrp

!
interface TenGigabitEthernet2/1
 description to 6k-core-1
```

（待续）

```
 ipv6 address 2001:DB8:CAFE:10::2/64
 ipv6 eigrp 10
 ipv6 hello-interval eigrp 10 1
 ipv6 hold-time eigrp 10 3
 ipv6 authentication mode eigrp 10 md5
 ipv6 authentication key-chain eigrp 10 eigrp
!
interface Vlan2
 description ACCESS-DATA-2
 ipv6 address 2001:DB8:CAFE:2::2/64
 ipv6 eigrp 10
!
ipv6 router eigrp 10
 router-id 10.122.10.9              !RID using Loopback0
 no shutdown                         !IPv6 EIGRP process is by default shut down
 passive-interface Vlan2
 passive-interface Vlan3
 passive-interface Loopback0
```

由配置可知，IPv6 EIGRP是以每接口为基础来配置的。在本例中，已遵照Cisco园区网络设计指南，调整了EIGRP 的HELLO和holdtimer值，以追求更快的收敛速度，并同时开启了EIGRP认证功能。EIGRP进程的router ID仍维持为32位，并取自配置在接口上，或手工定义的IPv4地址。此外，在实战中，我们强烈建议采用**ipv6 summary-address eigrp**命令，激活EIGRP IPv6路由汇总功能，向核心层交换机通告汇总路由。若交换机（路由器）只运行IPv6协议，则需要手工配置EIGRP router ID。

例6-12所示为交换机6k-core-1上与IPv6有关的EIGRP配置。由于在核心层交换机上，开通了两条冗余链路与接入层交换机互连，在这两条链路上，路由选择方面的配置便完全相同，因此例中省略了其中一条链路（一个接口）的

EIGRP 配置。

例 6-12 6k-core-1 路由选择配置

```
key chain eigrp
 key 100
   key-string 7 1111
!
interface Loopback0
 ip address 10.122.10.3 255.255.255.255
 ipv6 address 2001:DB8:CAFE:1F3::1/128
 ipv6 eigrp 10
!
interface TenGigabitEthernet2/4
 description to 6k-dist-1
 ipv6 address 2001:DB8:CAFE:10::1/64
 ipv6 eigrp 10
 ipv6 hello-interval eigrp 10 1
 ipv6 hold-time eigrp 10 3
 ipv6 authentication mode eigrp 10 md5
 ipv6 authentication key-chain eigrp 10 eigrp
!
ipv6 router eigrp 10
 router-id 10.122.10.3
 no shutdown
 passive-interface Loopback0
```

能否理解并熟知各种 IGP 计时器的含义和相应计时器的修改效果，是考察网络工程师是否称职的标杆之一。快速收敛是园区网络区块 IGP 设计所追求的主要目标。与分支机构和 WAN 网络环境相比，应该更为严格地去控制园区网络区块所运行的 IGP 计时器。在之前的配置示例中，与 IGP 计时器有关的配置全都遵照 Cisco 园区网络 IGP 配置最佳做法中的建议。在生产网络中部署 IPv6 之前，读者既要弄清每条 IGP 相关命令的含义，也要理解命令中计时器的取值。本章篇末“参考资料”一节所给出的链接列出了 Cisco 园区网络设计最佳做法文档。

6.3.4 第一跳冗余的配置

采用 DSM 部署 IPv6 时，与 HA 有关的设计方案除了涉及网络中各层（园区网络区块的核心层和分布层，数据中心网络区块的核心层和汇聚层）所

部署的两台交换机以外，还与IPv4和IPv6路由选择以及全面的故障容错机制有关[①]。在本章所要展示的园区网络区块分布层交换机的参考配置中，采用HSRP同时作为IPv4和IPv6的第一跳冗余机制。当然，采用GLBP也无不可[②]。

作用于IPv6的HSRP也是以每接口为基础来进行配置，这非常类似于IPv4的HSRP配置。例6-13所示为用于IPv4和IPv6的HSRP配置，请读者注意比较。由作用于IPv6的HSRP配置可知，HSRP版本2已经激活，并利用了该版本的某些新特性，比如，基于毫秒计时器值的通告和学习机制、扩展的组范围，以及对IPv6的支持等。

配置standby虚拟IPv6地址时，有手动定义本地链路地址（FE80::/10前缀）或自动配置两个选项。请注意，取决于硬件平台和HSRP代码版本，也可将公网IPv6地址定义为standby地址。**ipv6 autoconfig**命令可用来生成一个本地链路地址，该地址由前缀FE80::/10和HSRP IPv6虚拟MAC地址构造而成。HSRP IPv6虚拟MAC地址的范围为0005.73A0.000–0005.73A0.0FFF。

如例6-13的配置所示，为了缩短故障切换时间、及时调整HSRP优先级，以重新确立6k-dist-1和6kdist-2交换机间的active/standby角色，我们同时针对HSRP IPv4和IPv6配置了较低的hello计时器值。在6k-dist-1的HSRP配置中同样包括了占先（Preemption）的配置，只有这样，才能确保该交换机从具有较低HSRP优先级的6k-dist-2交换机手里重新夺回active（活跃路由器）的角色。由于该交换机是一台安装了多块线卡的Catalyst 6500，故而将HSRP的占先延迟时间值配置为了180秒。这是一种配置HSRP占先延迟时间值的最佳做法，其意在确保交换机有足够的时间让安装于自身的线卡加电并激活，只有在这之后，交换机才适合担当活跃路由器的角色。

最后，为保证安全性，在两台分布层交换机之间还开启了HSRP认证机制。欲知更多与IPv6 HSRP有关的信息，请见http://www.cisco.com/en/US/docs/ios/ipv6/configuration/guide/ip6-fhrp.html#wp1055254。

例6-13所示为6k-dist-1交换机与IPv4和IPv6 HSRP有关的配置。

---

① 原文是“The HA design in the DSM consists of running two of each switches (applicable in the distribution, core, and data center aggregation layers) and ensuring that the IPv4 and IPv6 routing configurations are tuned and completely fault-tolerant”，译文没有按照原文字面翻译。

② 原文是“All distribution pairs in the reference campus configuration are running HSRP for both IPv4 and IPv6.Optionally, GLBP can be used”。这种文字实在没法翻译，译文为译者杜撰。

例 6-13 6k-dist-1HSRP 配置

```
interface Vlan2
description ACCESS-DATA-2
standby version 2                                    !Required

standby 1 ip 10.120.2.1
standby 1 timers msec 250 msec 750
standby 1 priority 110
standby 1 preempt delay minimum 180
standby 1 authentication ese
standby 2 ipv6 autoconfig

standby 2 timers msec 250 msec 750

standby 2 priority 110

standby 2 preempt delay minimum 180

standby 2 authentication ese
```

6.3.5 QoS 配置

本章所展示的用于 DSM 的 QoS 配置代码均基于 Cisco 园区网 QoS 解决方案网络设计参考（QoS Solutions Reference Network Design[SRND]），这份网络设计参考的链接为 http://www.cisco.com/en/US/docs/solutions/Enterprise/WAN_and_MAN/QoS_SRND_40/QoSCampus_40.html。如本章 6.2 节所述，更改或创建同时作用于 IPv4 和 IPv6 流量的 QoS 策略时，在 QoS 的配置中，需要考虑的首要问题是，要确保将 IP 关键字从 **match** 和 **set** 语句中移除。只要不是因为作用于某应用程序 IPv6 流量的 QoS 策略与其 IPv4 流量完全不同，而需要分别“处理”的话，那么在制定 QoS 标准时，应对 IPv4 和 IPv6 流量一视同仁。用于执行分类、标记、排队以及监管操作的 QoS 策略会随客户的需求而产生很大的变化。而交换机所支持的队列类型和数量也会因交换机平台和所安装的线卡而异。由于 QoS 涉及的技术面太广，其配置亦与平台相关，故而在针对 IPv4 和 IPv6 部署 QoS 之前，读者应确保对 QoS 技术有透彻的理解。

本章并不包括 QoS 的入门内容，所举 QoS 的配置示例只是整体 QoS 解决方案中的“冰山一角”。例 6-14 所示为 6500-E 系列交换机上与 QoS 相关的总结性配置，该配置仅作参考之用。出于简化，配置中并未显示所有接口的配置。

例 6-14 6k-dist-1HSRP 配置

```
mls ipv6 acl compress address unicast     !Enable HW compression of address/ports

mls qos                                     !Enable QoS

!
class-map match-all BULK-DATA              !Associate IPv6 ACL with class-map

  match access-group name BULK-DATA
class-map match-all TRANSACTIONAL-DATA
  match access-group name TRANSACTIONAL-DATA
!
policy-map PER-PORT-MARKING                !Policy used for setting DSCP value

  class TRANSACTIONAL-DATA
   set dscp af21
  class BULK-DATA
   set dscp af11
  class class-default
   set dscp default
!
ipv6 access-list TRANSACTIONAL-DATA        !IPv6 ACLs used for classifcation
 remark HTTPS
 permit tcp any any eq 443
!
ipv6 access-list BULK-DATA
 remark FTP
 permit tcp any any eq ftp
 permit tcp any any eq ftp-data
!
interface GigabitEthernet3/1
 description to Access
 service-policy input PER-PORT-MARKING          #Apply marking policy
```

针对 IPv6 流量施以某些 QoS 特性时，无需在原有 IPv4 QoS 的配置基础上做任何改动。与排队处理有关的 QoS 配置正是其中的一例，该配置以每接口为基础，针对 IPv6 流量执行排队处理操作，无需添加任何特殊配置。例 6-15 所示为一份以每接口为基础来应用的 QoS 排队配置示例，该配置同时生效于 IPv4 和 IPv6 流量。

例 6-15 10G 以太网接口上的 QoS 排队配置示例

```
interface TenGigabitEthernet2/1
 description Uplink
 wrr-queue bandwidth 5 25 15 15 5 5 15     !Per-interface queueing

 wrr-queue queue-limit 5 25 15 15 5 5 15
 wrr-queue random-detect min-threshold 3 80 100 100 100 100 100 100 100
 wrr-queue random-detect min-threshold 4 80 100 100 100 100 100 100 100
 wrr-queue random-detect min-threshold 5 80 100 100 100 100 100 100 100
 wrr-queue random-detect min-threshold 6 80 100 100 100 100 100 100 100
 wrr-queue random-detect max-threshold 1 100 100 100 100 100 100 100 100
 wrr-queue random-detect max-threshold 2 100 100 100 100 100 100 100 100
 wrr-queue random-detect 4
 wrr-queue random-detect 5
 wrr-queue random-detect 6
 wrr-queue random-detect 7
 wrr-queue cos-map 1 1 1
 wrr-queue cos-map 2 1 0
 wrr-queue cos-map 3 1 2
 wrr-queue cos-map 4 1 3
 wrr-queue cos-map 5 1 6
 wrr-queue cos-map 6 1 7
 wrr-queue cos-map 7 1 4
 mls qos trust dscp                        !Trust previously marked DSCP values
```

6.3.6 多播配置

只要采用 DSM 部署 IPv6，IPv6 多播功能就可以全部用上。使用 PIM-SSM 或嵌入式 RP 时，IOS CLI 中并无相关的命令去激活设备的 IPv6 多播功能，这一点请读者务必理解。若单用 PIM-SSM 这一项特性，则只需在 IOS 的全局配置模式下执行 **ipv6 multicast-routing** 命令，该命令会在所有运行了 IPv6 的接口上，自动激活 PIM 功能。此外，该命令还会激活 PIM-SSM 特性以及与其相关联的多播组（地址）范围。在激活 IPv4 PIM-SSM 特性的过程中，所有的多播特性都需要手工配置，这与激活 IPv6-SSM 特性的过程差别很大。

以 3750-acc-1 交换机为例，为了控制 IPv6 多播流量的发布（只将多播流量发布给连接了活跃多播接收者的端口），该交换机需具备 IPv6 多播感知能力。可在该交换机上启用 MLD Snooping 特性来实现对 IPv6 多播的识别（感知）。在每台三层设备上，全局启用 IPv6 多播路由功能（在全局配置模式下执行 **ipv6 multicast- routing 命**

令），这将使得通向多播源（位于数据中心网络区块）的接口自始至终参与PIM进程。

虽然对IPv6多播设计的介绍超出了本章的范围，但本章会提供3750-acc-1、6k-dist-1以及6k-core-1交换机与多播有关的配置，网络中的多播应用程序也会使用PIM-SSM特性[①]。在执行过**ipv6 multicast-routing**命令，激活了PIM之后，交换机不但会自动创建用来完成多播源注册进程的PIM隧道接口，而且还会自动配置PIM-SSM的多播组（地址）范围。不过，上面提到的一切还要建立在以下基础之上：必须要让支持MLDv2功能的主机去访问安装在多播源主机上的多播应用程序，从而去触发播源主机发送目的地址为相应多播组地址的多播流量。

与IPv6多播相关的关键性配置其实很少，出于一致性的原则，以下所示为从接入层交换机到核心层交换机的IPv6多播配置。

- 3750-acc-1 在全局配置模式下启用MLD Snooping功能。

```
ipv6 mld snooping
```

- 6k-dist-1 在全局配置模式下启用IPv6多播路由功能。在采用IPv6多播加PIM的情况下，只要在全局配置模式下启用了IPv6多播功能，那么在任何已启用IPv6功能的接口（通过静态或动态方式分配了IPv6地址的接口）上，都会自动激活IPv6 PIM功能。

```
ipv6 multicast-routing
```

- 6k-core-1 在全局配置模式下启用IPv6多播路由功能。

```
ipv6 multicast-routing
```

在全局配置模式下启用了IPv6多播路由功能之后，可执行**show ipv6 pim range-list**命令去验证SSM的自动组范围特性是否已被激活，如例6-16所示。

例6-16 6k-core-1 - PIM-SSM多播组地址范围列表

```
6k-core-1# show ipv6 pim range-list
Static SSM Exp: never Learnt from : ::
  FF33::/32 Up: 00:00:05

  FF34::/32 Up: 00:00:05
```

（待续）

① 原文是“Although IPv6 multicast design is outside the scope of this chapter, configurations are shown for IPv6 multicast on the 3750-acc-1, 6k-dist-1, and 6k-core-1 switches, and the application is leveraging PIM-SSM”。译文为直译，因为替作者遮丑的成本实在太高。

```
FF35::/32 Up: 00:00:05
FF36::/32 Up: 00:00:05
FF37::/32 Up: 00:00:05
FF38::/32 Up: 00:00:05
FF39::/32 Up: 00:00:05
FF3A::/32 Up: 00:00:05
FF3B::/32 Up: 00:00:05
FF3C::/32 Up: 00:00:05
FF3D::/32 Up: 00:00:05
FF3E::/32 Up: 00:00:05
FF3F::/32 Up: 00:00:05
```

如例 6-17 的输出所示，在 3750-acc-1 交换机上可“看见”两台附接于本地的多播路由器，这两台多播路由器正是部署于分布层的两台交换机（与“ports”一栏下的端口相连）。

例 6-17　3750-acc-1 IPv6 多播 PIM 路由器状态

```
3750-acc-1# show ipv6 mld snooping mrouter
Vlan    ports
----    -----
   2    Gi1/0/25(dynamic), Gi1/0/26(dynamic)
```

当接入层交换机上的主机欲从某多播组接收流量时，可显示出与该多播组有关的信息，如例 6-18 所示。

例 6-18　3750-acc-1 IPv6 多播组的输出

```
3750-acc-1# show ipv6 mld snooping address
Vlan      Group           Type        Version     Port List
------------------------------------------------------------
2         FF35::1111      mld         v2          Gi1/0/25, Gi1/0/26
```

注意

读者很有可能遇到过这样一种情况：主机的操作系统不支持 MLDv2，导致其不能参与 PIM-SSM 网络环境。Cisco 开发出了一种名为 PIM-SSM 映射的特性，可利用该特性让只支持 MLDv1 的主机参与 PIM-SSM 网络环境。更多与该特性有关的信息请见 URL：http://www.cisco.com/en/US/docs/ios/ipv6/configura tion/guide/ip6-multicast_ps10591_TSD_Products_Configuration_Guide_Chapter. html#wp1058805。

在6k-dist-1上，可通过与IPv4相同的方式查看与PIM、多播路由、逆向路径转发以及多播组有关的信息。例6-19所示为6k-dist-1利用PIM-SSM转发活跃多播组（多播组地址为FF35::1111）流量的输出。由输出可知，多播流从与6k-core-1互连的接口流入，从interface **vlan2**（接入层主机所使用的网关接口）流出，该接口与3750-acc-1互连。

例6-19 显示6k-dist-1上的IPv6多播路由表项

```
6k-dist-1# show ipv6 mroute
Multicast Routing Table
Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group,
       C - Connected, L - Local, I - Received Source Specific Host Report,
       P - Pruned, R - RP-bit set, F - Register flag, T - SPT-bit set,
       J - Join SPT
Timers: Uptime/Expires
Interface state: Interface, State

(2001:DB8:CAFE:105:2E0:81FF:FE2C:9332, FF35::1111), 19:58:58/never, flags: sTI
  Incoming interface: TenGigabitEthernet2/1
  RPF nbr: FE80::215:C7FF:FE24:7440
  Immediate Outgoing interface list:
    Vlan2, Forward, 19:58:58/never
```

6.3.7 路由式接入的配置

比之传统的第二层接入设计方案，采用路由式接入设计方案时，在接入层和分布层交换机的配置上都会发生同样的变化。采用路由式接入方案时，由接入层交换机执行路由选择功能，而对于传统的接入层设计方案，则只让接入层交换机行使第二层功能，第一跳三层设备则被部署在分布层。本章虽不会讨论路由式接入设计方案的优势与劣势，但无论是对故障切换能力所实施的改进，还是将生成树这一复杂的控制平面协议“束之高阁”，都使得路由式接入设计能够吸引更多网络工程师的眼球。由于路由式设计方案在客户需求、性能以及运维方面的优势，本章将讨论采用路由式接入设计方案时IPv6的部署。

对DSM进行扩展，令其能够适用于路由式接入环境中IPv6的部署，其实颇为简单。对接入层主机而言，这消除了对第一跳冗余协议的依赖性。简而言之，就是在接入层交换机上开启IPv6路由功能，将接入层交换机和分布层交

换机之间的链路从 Trunk 链路改为路由式链路，令两者之间不再透传 VLAN 信息。

图 6-12 所示为经过更新的 DSM 网络拓扑，该网络拓扑采用了路由式接入设计方案。由于分布层以上未做任何改变，因此本图只示出了发生改变的接入层和分布层。此外，请读者注意，与图 6-11 相比，出于简洁，本图在显示 IPv6 地址时，省略了地址的前缀部分“2001:DB8:CAFE”，只示出了子网标识符（A、B、C 或 D）以及接口 ID。

如图 6-12 所示，现已将接入层交换机和分布层交换机之间的链路从 Trunk 改为了路由式链路。IPv6 地址和相应的路由选择功能亦需在链路两端的交换机上配置，VLAN 中的主机使用接入层交换机相关 SVI 接口的 IP 地址作为默认网关。

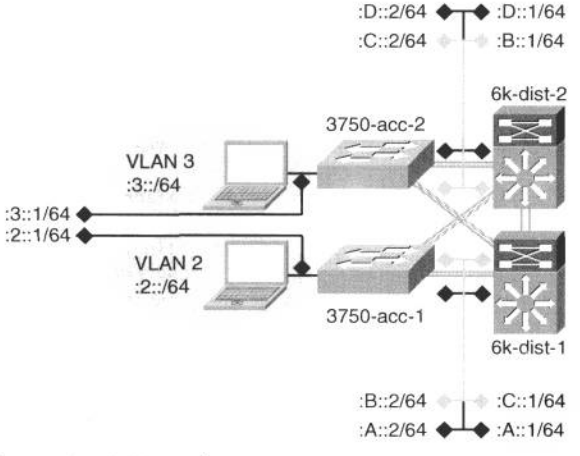


图 6-12 DSM 网络拓扑——路由式接入设计

注意

请那些在网络中使用 OSPF 的读者们注意，以下所示的 IGP 配置采用的正是 OSPFv3。可将其视为在园区网络区块中实施路由式接入设计方案时的 OSPFv3 配置范例。我们的目的是要帮助读者掌握在园区网络区块内配置 EIGRPv6 和 OSPFv3 的方法。

例 6-20 所示为 3750-acc-1 交换机与路由式接入有关的配置。

例6-20 路由式接入层交换机3750-acc-1

```
ipv6 unicast-routing                          !Globally enable IPv6 unicast routing
!
interface GigabitEthernet1/0/25
 description To 6k-dist-1
 ipv6 address 2001:DB8:CAFE:A::2/64           !Link is now a routed link
 ipv6 ospf network point-to-point
 ipv6 ospf hello-interval 1
 ipv6 ospf dead-interval 3
 ipv6 ospf 1 area 2                            !Link is in area 2
 no ipv6 redirects
 mls qos trust dscp
!
interface Vlan2
 load-interval 30
 ipv6 address 2001:DB8:CAFE:2::1/64            !VLAN2 on this switch becomes the
                                               !first layer 3 point for the
                                               hosts
                                               !in VLAN2 , the link-local
                                               address
                                               !on VLAN 2 will be the
                                               default
                                               !gateway for the hosts
 ipv6 ospf 1 area 2                            !VLAN2 is in area 2
 ipv6 nd managed-config-flag
 ipv6 dhcp relay destination 2001:DB8:CAFE:102::9
 no ipv6 redirects
!
ipv6 router ospf 1
 router-id 10.120.2.1
 log-adjacency-changes
 auto-cost reference-bandwidth 10000
```

（待续）

```
 area 2 stub no-summary              !Per the Routed Access Design guide - the
                                      !area (area 2) for the access layer
                                      !prefix is a totally stubby area
 passive-interface Vlan2
 timers spf 1 5
```

例 6-21 所示为 6k-dist-1 交换机与路由式接入有关的配置。

例 6-21 路由式接入层交换机 6k-dist-1

```
interface GigabitEthernet3/1
 description to 3750-acc-1
 ipv6 address 2001:DB8:CAFE:A::1/64          !Link is now a routed link

 no ipv6 redirects
 ipv6 ospf network point-to-point

 ipv6 ospf hello-interval 1

 ipv6 ospf dead-interval 3

 ipv6 ospf 1 area 2

!
ipv6 router ospf 1

 router-id 10.122.10.9
 log-adjacency-changes
 auto-cost reference-bandwidth 10000
 area 2 stub no-summary

 area 2 range 2001:DB8:CAFE:2::/64 cost 10       !Send a summary into area 0 for
                                                    !prefix "2" in area 2

 passive-interface Loopback0
 timers spf 1 5
```

如例 6-22 中 3750-acc-1 的 **show ipv6 route** 命令的汇总输出所示，该交换机已从两台分布层交换机学得默认路由（默认路由是由与核心层交换机相连的上游交换机注入，此类交换机下挂了连接到 Internet 的网络设备）[①]。

① 原文是“The default is injected by the upstream switches where the Internet edge connects to the core layer”。作者不解释默认路由的来历还好，一解释，反而让人摸不着头脑了。

例6-22 3750-acc-1 IPv6单播路由表的输出

```
3750-acc-1# show ipv6 route
IPv6 Routing Table - 13 entries
Codes: C - Connected, L - Local, S - Static, R - RIP, B - BGP
       U - Per-user Static route
       I1 - ISIS L1, I2 - ISIS L2, IA - ISIS interarea, IS - ISIS summary
       O - OSPF intra, OI - OSPF inter, OE1 - OSPF ext 1, OE2 - OSPF ext 2
       ON1 - OSPF NSSA ext 1, ON2 - OSPF NSSA ext 2
OI  ::/0 [110/11]

     via FE80::213:5FFF:FE1F:F840, GigabitEthernet1/0/26      !6k-dist-2

     via FE80::215:C7FF:FE25:9580, GigabitEthernet1/0/25      !6k-dist-1

C   2001:DB8:CAFE:2::/64 [0/0]
     via ::, Vlan2
L   2001:DB8:CAFE:2::1/128 [0/0]
     via ::, Vlan2
```

采用DSM并配搭路由式接入设计方案部署IPv6时，与6.3.6节所介绍的多播配置相比，也发生了一些变化。现在，由接入层交换机来执行路由选择功能，故而需在接入层交换机上启用PIM，PIM类型则要与网络中其他层设备上所运行的保持一致，即需将6.3.6节中所示的6k-dist-1交换机的多播配置“迁移”到接入层交换机上。请注意，在“迁移”配置之前，需确认接入层交换机平台是否支持IPv6多播，及其所运行的代码版本。

6.3.8 Cisco虚拟交换系统与IPv6

本节将介绍在园区网络区块中部署IPv6时对虚拟交换系统（VSS）的运用，并给出相关示例。

为防止单点故障，绝大多数企业都在分布层和核心层以故障切换对的方式部署交换机。这一部署方案给网络管理增加了难度——既增加了所需管理的节点数量，又需要运行诸如STP或HSRP之类的协议。此外，为实现无环的网络拓扑，STP会阻断通往接入层的一半链路，这也使得对链路带宽的利用率降低。

VSS通过把一对Catalyst 6500系列交换机归并为单一的网络/管理单元，以此来降低网络管理方面的难度。可将参与VSS的两台对等交换机视为一台逻辑

交换机（如图 6-13 所示）。图中的接入层交换机同时连接到两台分布层交换机物理机箱，这两台交换机物理机箱使用单条名为 MEC（多机箱 Ether-channel）的逻辑通道互联。利用 MEC，网络工程师便可在弃用 STP 或 HSRP 等协议的情况下，部署冗余而又无第二层环路的网络。

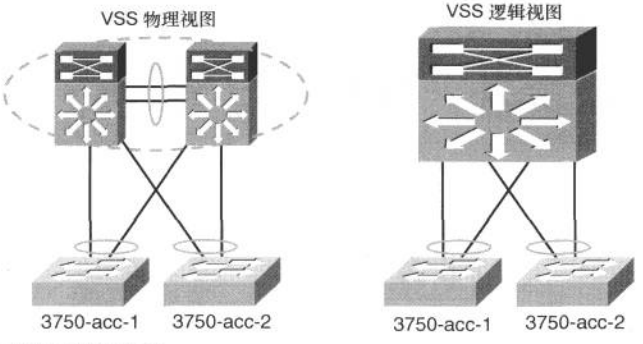


图 6-13 VSS 的物理和逻辑视图

站在数据平面的立场（从数据转发的角度来看），两台参与 VSS 的交换机处于 active-active 模式（同时参与数据的转发）；站在控制平面的立场（从路由信息收集的角度来看），两台参与 VSS 的交换机则处于 active-standby 模式。假如其中的一台交换故障，在控制平面收敛的同时，这种数据转发模式（active-active 模式）可保证流量不间断的转发。

图 6-14 所示为在先前所讨论的 DSM 拓扑中采用 VSS 时的部署；图中的核心层和分布层交换机均被配置为参与 VSS。

从网络的逻辑结构上来看，图 6-14 所示的网络拓扑看上去就像是由一台台交换机像链条那样串接起来，但实际上，交换机都是成对部署。比之图 6-10 所示的网络拓扑，在本图所示的 DSM 部署中，除了第二层无环特性以及独特的逻辑配置方式以外，分布层和核心层的其他所有一切均无任何变化。

从冗余性的角度来看，VSS 可支持跨交换机机箱的状态化故障切换（SSO）特性，因此假如 VSS 对中的一台交换机机箱发生故障，连接到该 VSS 的交换机只会将这一故障视为 MEC 中的一条链路故障，而不会认为发生了网络拓扑变更。最终的结果是，另外一台“健在”的交换机机箱仍会继续转发流量。此外，因跨交换机机箱所采用的故障切换机制是 SSO，上述故障发生时，路由选择拓扑不会发生改变，路由协议无需重新收敛。

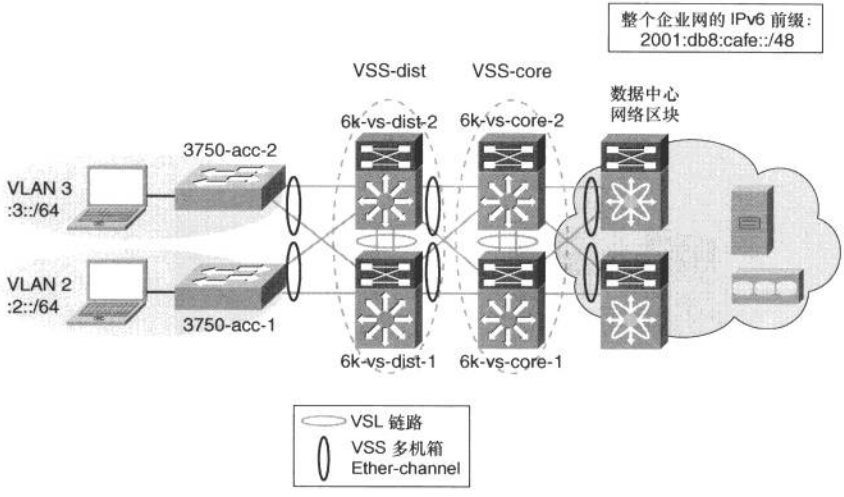


图 6-14 在IPv6部署中使用VSS

图6-11所示的编址方案对VSS部署同样适用。两者之间唯一的差异是，对于后者，每个VSS对需要一个IPv4和IPv6地址，而对于前者，则是每台交换机机箱需要一个IPv4和IPv6地址。表6-5所示为VSS配置示例中所使用的IPv6地址。

表 6-5 VSS示例中交换机Loopback/VLAN地址

交换机	接口	IPv6地址
VSS-dist	Loopback	2001:db8:cafe:1f3::3/128
	VLAN2	2001:db8:cafe:1f2::2/64
	VLAN3	2001:db8:cafe:1f3::2/64
VSS-core	Loopback	2001:db8:cafe:1f3::1/128

VSS配置

VSL（虚拟交换机链路）是一项用来支撑VSS的技术。参与VSS对的交换机通过VSL来交换配置和状态信息。备用（standby）交换机机箱通过VSL监控主用（active）交换机机箱。

为了完成VSS配对，两台独立的交换机需要经历从独立到VSS（standal one-to-VSS）的转换过程。下面是对三步转换过程的总结。

步骤1 选定一个虚拟交换机域编号（需确保其在同一个第二层网络中的唯一性）。

步骤2 为每台参与VSS对的交换机选定一个交换机编号（对于本例，

6k-vs-dist-1 为 1#交换机，6k-vs-dist-2 为 2#交换机）。

步骤 3 选定一条 VSL 链路（用来互连两台交换机的 etherchannel 端口，每台交换机上的 etherchannel 端口号需具备唯一性）。

在 VSL 链路上传输的帧都被封以一个特殊的帧头，长度为 32 字节，其作用是，为 VSS 提供在每台交换机机箱上所转发的数据包的信息。VSL 链路既会承载控制消息也会承载实际的用户数据[①]；在以下配置示例中，两台交换机分别使用 port-channel 10 和 20 作为 VSL 链路的端口。VSL 相关配置会在 VSS 转换配置之后示出（见例 6-23）。

例 6-23 从独立的交换机转换到 VSS（6k-vs-dist-1 和 6k-vs-dist-2）

```
Chassis Example: 6k-vs-dist-1
6k-vs-dist-1(config)# hostname VSS-dist
VSS-dist(config)# switch virtual domain 100
VSS-dist(config-vs-domain)# switch 1                !switch number 1

VSS-dist(config-vs-domain)# exit

VSS-dist(config)# interface Port-channel10
VSS-dist(config-if)# no switch virtual link

VSS-dist(config-if)# interface TenGigabitEthernet3/5
VSS-dist(config-if)# channel-group 10 mode on
VSS-dist(config-if)# interface TenGigabitEthernet3/6
VSS-dist(config-if)# channel-group 10 mode on

VSS-dist# switch convert mode virtual              !Enter yes when prompted

!Output removed for brevity

Chassis Example: 6k-vs-dist-2

6k-vs-dist-2(config)# hostname VSS-dist
VSS-dist(config)# switch virtual domain 100
VSS-dist(config-vs-domain)# switch 2                !switch number 2

VSS-dist(config-vs-domain)# exit

VSS-dist(config)# interface Port-channel20
```

（待续）

① 原文是“The VSL link carries the control and data”。能看懂这句话的人还用的着阅读本书吗?在译者看来，Cisco 文档都比本书容易阅读，而且 Cisco 文档还是免费的。

```
VSS-dist(config-if)# no switch virtual link

VSS-dist(config-if)# interface TenGigabitEthernet4/5
VSS-dist(config-if)# channel-group 20 mode on

VSS-dist(config-if)# Interface TenGigabitEthernet4/6
VSS-dist(config-if)# channel-group 20 mode on

VSS-dist# switch convert mode virtual          !Enter yes when prompted

!Output removed for brevity
```

注意

欲知更多与从独立的交换机转换到VSS有关的详细信息，请参阅从独立的Cisco Catalyst 6500交换机向Cisco Catalyst 6500虚拟交换系统转换的文档，链接为：http://www.cisco.com/en/US/ products/ps9336/products_tech_note09186a0080a7c74c.shtml。

例6-24和例6-25所示为从独立的交换机转换到VSS之后，各分布层交换机上的配置。

例6-24 布层交换机VSL的配置（6k-vs-dist-1）

```
Chassis: 6k-vs-dist-1                          !The VSL config before merge
switch virtual domain 100

 switch mode virtual
!
interface Port-channel10
 no switchport
 no ip address
 switch virtual link 1

 mls qos trust cos
 no mls qos channel-consistency
!
interface TenGigabitEthernet1/3/5
 no switchport
 no ip address
 mls qos trust cos
 channel-group 10 mode on
```

（待续）

```
!
interface TenGigabitEthernet1/3/6
 no switchport
 no ip address
 mls qos trust cos
 channel-group 10 mode on
```

VSS 的转换完成之后，两台交换机形成 VSS 时，例 6-25 所示的 VSL 专用配置将会被自动合并，从而致使两者的配置相同。[①]

例 6-25 分布层交换机 VSL 的配置（6k-vs-dist-2）

```
Chassis: 6k-vs-dist-2                          !The VSL config before merge

switch virtual domain 100

 switch mode virtual

!
interface Port-channel20
 no switchport
 no ip address
 switch virtual link 2

 mls qos trust cos
 no mls qos channel-consistency
!
interface TenGigabitEthernet2/4/5
 no switchport
 no ip address
 mls qos trust cos
 channel-group 20 mode on

!
interface TenGigabitEthernet2/4/5
```

（待续）

① 原文是“After the conversion to VSS, when both switches come up, the VSL-specific configuration shown in Example6-25will be dynamically merged, resulting in the same configuration on both switches”译文为直译。在译者看来，作者根本就没有弄懂如何在生产环境中让独立的交换机转换到 VSS，请读者千万不要在生产环境中按作者的配置行事，但“注意”中作者所给链接（Cisco 文档《Migrate Standalone Cisco Catalyst 6500 Switch to Cisco Catalyst 6500 Virtual Switching System》却很好地描述了如何将生产网络中独立的交换机迁移到 VSS。很遗憾，作者明显没有读懂这份文档。译者再次重申，阅读本书不如阅读 Cisco 文档）。

```
no switchport
no ip address
mls qos trust cos
channel-group 20 mode on
```

注意 在VSS中，接口的命名惯例会有所改变——取决于物理交换机机箱的编号，在接口编号前，会被冠之以额外的前缀1或2。例如，interface TenGigabitE thernet 2/4/5意指接口所在的物理交换机机箱号为2，槽位号为4，端口号为5。

VSS物理接口的IPv6配置

对于核心层和分布层交换机之间的点到点链路，VSS的配置方式与独立交换机的配置方式几乎相同，但要记住，现在，分布层和核心层的两台交换机都是按单台虚拟交换机的方式运作。因此，只需给用于点到点链路的port channel接口配置一个IP地址（而对于非VSS的部署，则需给每个物理接口分配一个IP地址）[①]。例6-26所示为分布层VSS交换机用来上连核心层交换机点对点链路的接口配置。

例6-26 分布层VSS交换机上连核心层交换机的Port-Channel接口的配置

```
ipv6 unicast-routing
ip cef distributed
ipv6 cef distributed
!
interface Port-channel30
 description to VSS-core
 ipv6 address 2001:DB8:CAFE:10::2/64
!
interface GigabitEthernet1/7/1
 no switchport
 no ip address
```

（待续）

① 原文是“Therefore, only one IPv6 address is used on the port channel interfaces (instead of one for each port channel to each switch in a non-VSS model)”。作者在行文时又没有考虑清楚。在不采用VSS的情况下，两台分布层交换机会发起两条物理链路，分别连接两台核心层交换机。在分布层交换机上，由于这两条链路并非由一台核心交换机终结，故而不可能形成Ether-channel；采用了VSS之后，由于分布层和核心层交换机都“缩水”为一台虚拟交换机，因此原本由两台交换机终结的物理链路现在由一台交换机终结，从而可以将这两条链路捆绑为Ether-channel。其实，再仔细考虑一下，原本分布层和核心层交换机之间的4条链路，现在可以“缩水”为2条。综上所述，译文并没有按照原文翻译。

```
 channel-group 30 mode desirable
!
interface GigabitEthernet2/7/1
 no switchport
 no ip address
 channel-group 30 mode desirable
```

例 6-27 所示为核心层 VSS 交换机下连分布层交换机点对点链路的配置。

例 6-27 核心层 VSS 交换机下连分布层交换机的 Port-Channel 接口的配置

```
ipv6 unicast-routing
ip cef distributed
ipv6 cef distributed
!
interface Port-channel30
 description to VSS-dist
 ipv6 address 2001:DB8:CAFE:10::1/64
!
interface GigabitEthernet1/7/1
 no switchport
 no ip address
 channel-group 30 mode desirable
!
interface GigabitEthernet2/7/1
 no switchport
 no ip address
 channel-group 30 mode desirable
```

当分布层设备选择使用 VSS 部署方式时，接入层仍采用每台交换机对应一个 VLAN 的 VLAN 分配方式，这与在分布层部署独立的交换机没有任何区别。接入层交换机之间不会透传 VLAN 信息，VLAN 信息由分布层 VSS 交换机终结。

至于所有其他的配置，比如，路由选择和 VLAN 的配置（包括接入层交换机在内），与采用非 VSS 部署的场景也一模一样。

6.4 实施混合模型

以混合模型（HM）部署 IPv6，是为了让企业园区网络区块中的某些用户能够访问到基于 IPv6 的应用程序，即便这些用户身处的园区网络区块还不能完全支持 IPv6。由于用户身处的园区网络区块并非完全支持 IPv6，因此大多数采用

HM部署IPv6的园区网络只支持IPv4。当企业网络向IPv6过渡时，对于其园区网络区块来说，应最先让其中的核心层部分支持IPv6。以下各节将会给出采用HM部署IPv6时，企业园区网络区块核心层交换机的配置，以及主机的ISATAP隧道的配置。如前所述，采用HM部署IPv6时，其前提条件是从园区网络区块核心层到数据中心网络区块全都支持IPv6/IPv4双协议栈。与双协议栈有关的配置均与实施HM无关，这是因为此类配置与前面提到的DSM配置相关[①]。那些与实施DSM有关的配置请见6.3节。

6.4.1 网络拓扑

如HM网络拓扑所示，在园区网络区块的分布层内，部署的是一对Catalyst 3750，而上一节DSM网络拓扑中的园区网络区块分布层内，则部署的是一对Catalyst 6500，这也是两种模型拓扑结构的唯一差异。产生这一变化仅仅是测试环境中硬件配备方面的原因，到不是因为有任何特殊的问题或建议。尽管Catalyst 3750不支持线速IPv6单播转发（以硬件的方式，转发IPv6单播包），但为了演示实施HM的配置，只需要让其模拟只支持IPv4的交换机平台，故而也无需在其上启用IPv6功能。图6-15所示为实施HM时的网络拓扑。

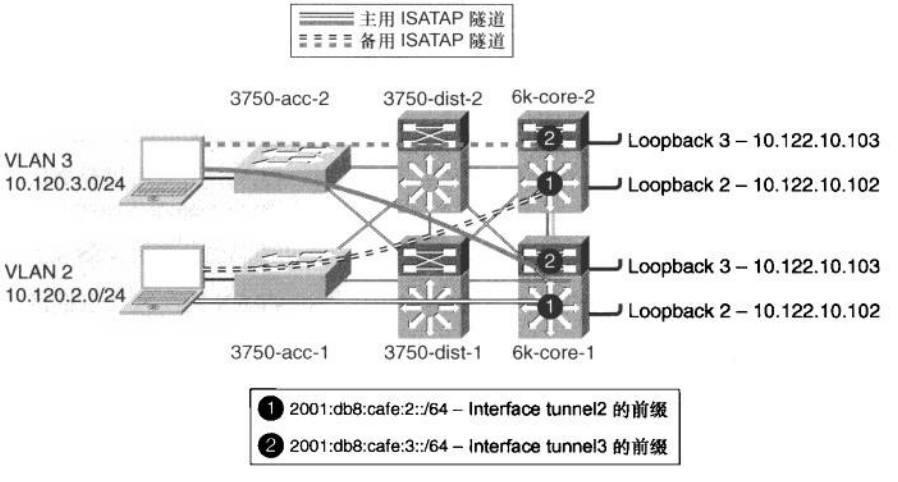


图6-15 混合模型网络拓扑

① 这句又是废话！

请读者关注图 6-15 中接入层和核心层的编址方案（即接入层主机用来建立 ISATAP 隧道的地址和接入层交换机用来终结 ISATAP 隧道的地址），此外，还应关注 ISATAP 隧道所使用的 IPv6 地址前缀。实施 HM 时，核心层交换机之间的互连链路以及核心层交换机与数据中心区块之间的链路全都支持双协议栈。图 6-15 并未显示这部分的 IPv6 编址，其原因是相关的编址情况与实施 DSM 完全相同。由上图可知，在两台核心层交换机上，分别配置了两个 IPv4 地址完全相同的 loopback 接口，不同交换机上 IPv4 地址完全相同的 loopback 接口形成了冗余配置，正是以此来实现 ISATAP 隧道的高可用性。为了让接入层主机在建立 ISATAP 隧道时保持一致性，主备隧道都要使用相同的前缀。

6.4.2 物理接口的配置

本节所示的物理接口配置只包括两台核心层交换机与分布层交换机互连接口的配置。请读者注意配置中的接口地址，因为这可以帮助理解后文将要给出的验证 ISATAP 隧道高可用性的输出。请不要忘记，从接入层到核心层只不过是 IPv4 网络，并不运行 IPv6，因此除了需要考虑 IPv4 网络的冗余和路由协议的快速收敛之外，也再无其他需要考虑了。

例 6-28 所示为 6k-core-1 交换机上与接入层交换机互连接口的配置。

例 6-28 6k-core-1 接口配置

```
interface GigabitEthernet1/1
 description to 3750-dist-1

 ip address 10.122.0.41 255.255.255.252

 ip hello-interval eigrp 10 1
 ip hold-time eigrp 10 3
!
interface GigabitEthernet1/2
 description to 3750-dist-2

 ip address 10.122.0.45 255.255.255.252

 ip hello-interval eigrp 10 1
 ip hold-time eigrp 10 3
```

例 6-29 所示为 6k-core-2 交换机上与接入层交换机互连接口的配置。

例6-29 6k-core-2接口配置

```
interface GigabitEthernet1/1
 description to 3750-dist-1

 ip address 10.122.0.49 255.255.255.252

 ip hello-interval eigrp 10 1
 ip hold-time eigrp 10 3
!
interface GigabitEthernet1/2
 description to 3750-dist-2

 ip address 10.122.0.53 255.255.255.252

 ip hello-interval eigrp 10 1
 ip hold-time eigrp 10 3
```

6.4.3 隧道的配置

如果只是配置ISATAP隧道本身的话，其实并不复杂，但只要ISATAP隧道与高可用性“搅合”在一起，情况比较棘手了。在接入层主机上，配置ISATAP隧道的基本步骤不外是：首先启用IPv6，然后指明ISATAP路由器的名称或IPv4地址。默认情况下，Microsoft Windows XP、Vista以及Windows 7都会针对“isatap.domain.com”进行DNS查询，其中“domain.com”是分配给企业园区网络区块的域名。假如已经针对路由器名“isatap”配置了DNS“A”记录，那么主机便会开始建立ISATAP隧道，隧道端点的IPv4地址与路由器名“isatap”相对应（即通过DNS解析得出的ISATAP路由器的IPv4地址）。一般情况下，这一默认工作方式肯定能够运转正常，直到ISATAP路由器或通向该路由器的链路发生故障，导致接入层主机不得不再建立一条通往备用ISATAP路由器的隧道为止。本节将要介绍的所有配置都虑及了IPv6服务的高可用性，并对此做了尽可能的优化。

实施HM时，确保ISATAP隧道的高可用性是其中的关键。有若干种方法能够实现ISATAP路由器的冗余。利用部署在园区网络区块核心层中的两台交换机，来实现ISATAP隧道的快速故障切换，是本节所要讨论的方法。其他常用的方法则大都依赖于DNS。尽管DNS方法实施起来要更快捷一些，但其同样对IPv6园区网络区块的整体性网络设计限制颇多，此外，该方法在故障切换时所消耗的时间也要更长一些。

要想实现 ISATAP 隧道的高可用性，不但要确保核心层交换机上隧道端点（从接入层主机的视角来看）的冗余，还要保证核心层交换机之间的 IPv4 和 IPv6 路由选择，以及核心层和接入层之间的 IPv4 路由选择配置得当。

请读者考虑以下两个常见的问题：其一，就 IPv4 路由选择而言，分布层交换机在将 IPv4 数据包转发给其中的一台 ISATAP 隧道路由器（核心层交换机）时，应该明确数据包转发的具体路径吗？其二，分布层交换机和 ISATAP 隧道路由器（核心层交换机）之间交换被隧道封装的流量时，有必要负载均衡吗？此外，读者还应关注以下两点。

- 操作系统中只有 Microsoft Windows Vista 和 Windows 7 支持出站方向上的 ISATAP 隧道流量的负载均衡。
- 在 ISATAP 主机和 ISATAP 路由器之间实施负载均衡并不会带来任何好处，这是许多部署和测试案例所得出的论证。测试表明，当使用具有冗余 IPv6 前缀的 ISATAP 隧道来实施负载均衡时，从发送流量的 ISATAP 主机的角度来看，会出现非对称路由问题，这意味着 ISATAP 主机发出的流量及相应的回馈流量不会走同一条路线（路由）。在本例中，会配置两台核心层交换机，令其中的一台去承载 ISATAP 隧道流量——若主交换机故障，备用交换机则会承载所有的 ISATAP 隧道流量——这是一种比较理想的隧道冗余部署方案。而让两台核心层交换机对 ISATAP 隧道流量实施负载均衡不但不会提高性能、降低负载、增强高可用性，而且还会让网络更难管理。此外，实施负载均衡还会提高排除网络故障的难度，其原因是很难摸清 ISATAP 隧道流量的流动路线。而确立一台核心层交换机，令其承载所有 ISATAP 隧道流量，除了能够降低流量管理和故障排除的难度以外，还能消除前面提到的非对称路由问题。

为了在核心层交换机出现故障时降低 ISATAP 隧道的故障切换时间[①]，必须在两台核心层交换机上配置重复的隧道端点地址。重复的隧道端点地址配置完成之后，在接入层主机上只需指明一个 ISATAP 路由器的名称或一个 IP 地址，无需利用 DNS 轮询技术让一个 ISATAP 路由器名称对应多个 IP 地址了。以下步骤是对上述过程的描述。

步骤 1 在两台核心层交换机上创建 loopback 接口（loopback2），并配置相同的 IPv4 地址（对于本例，IPv4 地址为 10.122.10.102）。由于 loop-

① 原文是“To maintain low convergence times for ISATAP tunnels when a core layer switch fails”。作者没有常识。只有路由协议才有“convergence time（收敛时间）”，ISATAP 隧道怎么能和收敛时间扯到一起快去呢，译文酌改。

back接口具有与生俱来的稳定性，因此首选其用来终结ISATAP隧道。请务必确保该loopback接口地址不为网络设备的Router-ID（RID）所用，否则将会在园区网络区块中导致重复的RID问题，这会影响到核心层交换机的CEF功能。

步骤2 在两台核心层交换机上创建ISATAP隧道，并以步骤1所创的loopback接口地址作为隧道的源地址（Loopback2—10.122.10.102）。为所创建的ISATAP隧道接口分配相同的IPv6前缀，如此一来，无论哪台交换机去终结由主机发起的ISATAP隧道都无关紧要了，主机可使用相同的前缀来建立连通性。快速故障切换的关键要素之一，便是不用保存状态。假如每次故障切换，主机都不得不获得一条新IPv6前缀的话，那么故障恢复时间将会受到这一新地址获取过程的影响。此外，在地址（租用）的生命周期内，主机还会缓存所有过去（老的ISATAP前缀）和新的IPv6地址。让所有由主机发起的ISATAP隧道连接都使用单条IPv6前缀，不但可显著缩短主机的故障切换时间，而且无论发生任何故障，ISATAP隧道都只会与一条IPv6路由表项相对应。

步骤3 配置两台核心层路由器，令两者通过IPv4 IGP通告步骤1中创建的loopback接口地址。主交换机（6k-core-1）使用默认的IGP度量值通告loopback接口路由。备用交换机（6k-core-2）以较高的IGP度量值通告loopback接口路由（本例采用EIGRP作为IPv4 IGP时，因此增大的是相应路由的延迟度量值），以使得分布层路由器优选6k-core-1所通告的loopback接口路由。当接入层交换机建立用来封装IPv6流量的ISATAP隧道时，Cisco的建议是，预先确立一个隧道端点地址，用来建立ISATAP隧道，因为利用两条等价的隧道端点路由来建立ISATAP隧道，以期执行被隧道封装流量的负载均衡并不可取[①]。

步骤4 配置两台核心层路由器，令两者通过IPv6 IGP通告ISATAP隧道的IPv6前缀。主交换机（6k-core-1）使用默认的IGP度量值通告ISATAP隧道的IPv6路由。备用交换机（6k-core-2）以较高的IGP度量值通告ISATAP隧道的IPv6路由（本例采用OSPFv3作为IPv6

① 原文是“Cisco recommends having a deterministic flow for the tunnels because load balancing between the tunnels using the same prefix is not desirable”，译者对原文进行了改写。

IGP，因此增大的是相应路由的 cost 值），以使得自己所通告的这条 IPv6 路由为次优路由。并不是非得调整 SATAP 隧道的 IPv6 路由的 IGP 度量值，本例如此设置是为了要预先确立 IPv4（请见步骤 3）和 IPv6 流量的转发路径。

步骤 5 在接入层主机上，手工指定 ISATAP 路由器的 IP 地址或名称（与 DNS "A" 记录相对应）。

注意

本节所举示例采用 EIGRP 作为 IPv4 IGP，采用 OSPFv3 作为 IPv6 IGP。当然，采用 EIGRP 或 OSPFv3 同时作为 IPv4 和 IPv6 的 IGP 也无不可。由于 6.3 节 DSM 的配置正是以 EIGRP 作为 IPv6 IGP，因此本节采用 OSPFv3 作为 IPv6 IGP 的目的是，为读者提供 IPv6 OSPFv3 的参考配置。

为了让 ISATAP 隧道、HA 以及路由选择的配置简单易懂，本节在向读者展示配置示例时会将以上三者一并示出。出于简化，配置实例中只会示出接入层 VLAN2 中的主机所使用的隧道配置。VLAN3 中的主机所使用的隧道配置只是在 IP 地址上有所不同。

以下所示为按前述 5 个步骤对核心层交换机的配置。配置中包括了 6kcore-1 用作为 ISATAP 隧道源地址的 loopback 接口的配置。隧道接口配置则由 IP 配置和相关 OSPFv3 配置组成。配置隧道接口，令其参与 IGP，其目的是要通告隧道接口的 IPv6 前缀，而不是要利用隧道接口去建立 IGP 邻接关系。默认情况下，隧道接口并不会通过隧道发送路由器通告消息，原因很简单——在点对点隧道上没有理由发送 RA 消息。然而，ISATAP 隧道是一种半自动多点隧道，这也表明隧道端点的信息并不确定。由于 ISATAP 隧道有那么点动态的味道，因此必须在隧道上禁用 RA 抑制（启用 RA）。此外，如前所述，本节所举示例采用 EIGRP 作为 IPv4 IGP，采用 OSPFv3 作为 IPv6 IGP。故而需通过 EIGRP 通告 loopback 接口的 IPv4 地址，让客户端主机能够和隧道端点建立 IPv4 连通性。同理，还需通过 OSPFv3，向企业网中已启用了 IPv6 的网络通告 ISATAP 隧道接口的 IPv6 前缀[①]。

例 6-30 所示为 6k-core-1 交换机 ISATAP 隧道的配置。配置 loopback 接口是为了让主机借此来建立 ISATAP 隧道。隧道接口的配置包含了 IPv6 地址配置、禁用 RA 抑制特性（该特性在隧道接口上默认开启，本例则须在隧道接口上发送 RA 消息）配置、IGP 配置（目的是通过 IGP，向企业网中已启用了 IPv6 的

① 该段大多文字都是废话，有凑字数之嫌，但译者还是全文翻译了。

网络通告ISATAP隧道接口的IPv6前缀），以及隧道源接口和隧道模式的配置[①]。

例6-30 6k-core-1交换机ISATAP隧道的配置

```
interface Loopback2
 description Tunnel source for ISATAP-VLAN2
 ip address 10.122.10.102 255.255.255.255  !Address that will be used as the
                                             !ISATAP tunnel2 source
!
interface Tunnel2
 description ISATAP VLAN2
 no ip address
 no ip redirects
 ipv6 address 2001:DB8:CAFE:2::/64 eui-64   !Tunnel prefix used for ISATAP
                                             !hosts connecting to this tunnel.
                                             !Interface-ID address for this
                                             !switch will be generated using
                                             !EUI-64
 no ipv6 nd suppress-ra                     !Tunnel interfaces disable the
                                             !sending of RA's. This command
                                            !re-enables RA's on this interface.
 ipv6 ospf 1 area 2                         !Just like the VLAN in the DSM,
                                             !this interface is part of area 2
 tunnel source Loopback2                    !Tunnel2 uses loopback2 as the
                                            !source
 tunnel mode ipv6ip isatap                  !Define the tunnel as ISATAP
!
router eigrp 10
 passive-interface Loopback2
 network 10.0.0.0                      !Ensure that the 10.122.10.102 address is
                                        !advertised to the rest of the network
 no auto-summary
```

（待续）

① 整段原文是“Example 6-30 shows the 6k-core-1 ISATAP configuration. Again, a loopback is configured and is used as the anchor for the ISATAP tunnel. The tunnel definition needs to have an IPv6 address configured, default suppression of RAs on tunnel interfaces is disabled (we want RAs sent), an IGP is enabled to advertise the IPv6 prefix to the rest of the network, and the tunnel source (loopback) and mode are defined”。整段原文不但不通，而且似无必要。

```
 eigrp router-id 10.122.10.9
!
ipv6 router ospf 1
 router-id 10.122.10.9
 area 2 range 2001:DB8:CAFE:2::/64 cost 10  !Advertise summary for the prefix on

                                             !Tunnel2 - just like a VLAN prefix
                                             !would be sent in the DSM
 passive-interface Loopback2
 passive-interface Tunnel2
 timers spf 1 5
```

例 6-31 的配置与例 6-30 基本相同，只是 IP 地址有所不同。此外，还对 6k-core-2 交换机的 IGP 配置进行了微调，因为该交换机为备用 ISATAP 路由器（其 loopback 地址和 ISATAP 隧道接口的路由都应作为次优路由通告（与 6k-core-1 上的 loopback 接口地址路由相比））。

例 6-31　6k-core-2 交换机 ISATAP 隧道的配置

```
 interface Loopback2
 description Tunnel source for ISATAP-VLAN2
 ip address 10.122.10.102 255.255.255.255

 delay 1000                          !Delay adjusted for EIGRP (IPv4)

                                     !in order to adjust preference
                                     !for the 10.122.10.102 host
                                     !route. This ensures that
                                     !6k-core-2 is SECONDARY to 6k-core-1
!
interface Tunnel2
 description ISATAP VLAN2
 no ip address
 no ip redirects
 ipv6 address 2001:DB8:CAFE:2::/64 eui-64

 no ipv6 nd suppress-ra
 ipv6 ospf 1 area 2
 tunnel source Loopback2
 tunnel mode ipv6ip isatap
!
router eigrp 10
 passive-interface Loopback2
```

（待续）

```
 network 10.0.0.0
 no auto-summary
 eigrp router-id 10.122.10.10
!
ipv6 router ospf 1
 router-id 10.122.10.10
 area 2 range 2001:DB8:CAFE:2::/64 cost 20    !Cost for prefix adjusted so that
                                                    !the route from 6k-core-2 is not
                                                    !preferred or equal to 6k-core-1
 passive-interface Loopback2
 passive-interface Tunnel2
 timers spf 1 5
```

图6-16所示为从分布层交换机向核心层交换机转发ISATAP隧道流量（即接入层主机将IPv4数据包转发至核心层交换机上的loopback接口）时的数据包行进路线图。6k-core-1上的loopback2接口地址被设置为接入层主机的主ISATAP路由器地址。如先前的IPv4 IGP所示，在6k-core-2上，调高了其loopback2接口地址10.122.10.102的主机路由的EIGRP度量值（延迟值），因此这条主机路由会以次优路由的形式通告（与6k-core-1上的loopback接口地址路由相比）。收到接入层VLAN2中的主机发往ISATAP路由器（10.122.10.102）的数据包时，分布层交换机会查找自己的路由表，并会查出相应路由的下一跳为6k-core-1。

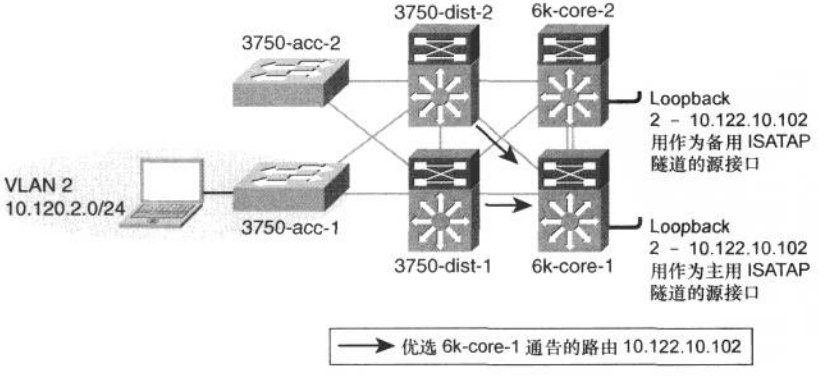


图6-16 混合模型——6k-core-1所通告的隧道路由器地址路由为最优路由

例6-32所示为3750-dist-1交换机上目的网络为10.122.10.102/32的路由表

项。有路由表可知，3750-dist-1 交换机通过 10.122.0.41（6k-core-1）学到了一条 10.122.10.102/32 的路由。

例 6-32 3750-dist-1 交换机通往核心交换机的路由表项的输出

```
3750-dist-1# show ip route ¦ b 10.122.10.102/32
D        10.122.10.102/32
[90/130816] via 10.122.0.41, 00:09:23, GigabitEthernet1/0/27
```

由 3750-dist-2 交换机的路由表项可知，路由 10.122.10.102/32 的下一跳也为 6k-core-1（10.122.0.45），如例 6-33 所示。

例 6-33 3750-dist-2 交换机通往核心交换机的路由表项的输出

```
3750-dist-2# show ip route ¦ b 10.122.10.102/32
D        10.122.10.102/32
[90/130816] via 10.122.0.45, 00:10:03, GigabitEthernet1/0/27
```

图 6-17 所示为 6k-core-1 出现故障，通向其 Loopback2 接口（10.122.10.102）的路由失效的情况。当 6k-core-1 通告的路由 10.122.10.102 失效之后，6k-core-2 所通告的路由 10.122.10.102 会变为优选路由，分布层交换机会将接入层主机发出的 ISATAP 隧道流量转发至 6k-core-2。

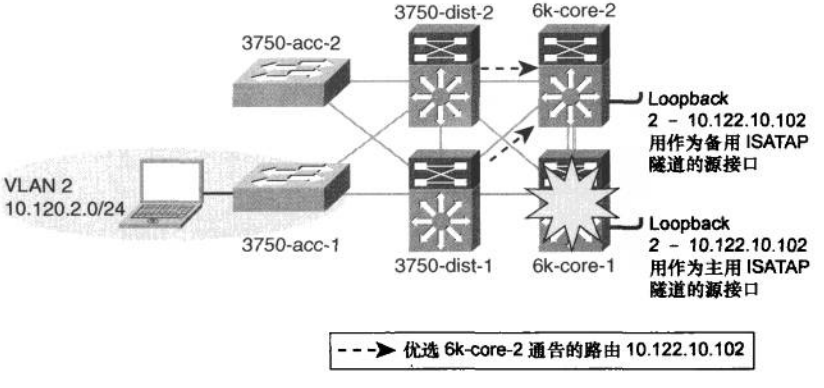


图 6-17 混合模型——6k-core-1 出现故障之后，6k-core-2 通告的隧道路由器地址路由变为最优路由

如例 6-34 分布层交换机上最新的路由表输出所示，现在，路由 10.122.10.102/32 的下一跳为 6k-core-2 交换机。

例6-34 6k-core-1故障之后，分布层交换机上的路由表输出

```
!3750-dist-1
!output omitted for brevity
3750-dist-1# show ip route | b 10.122.10.102/32
D        10.122.10.102/32
      [90/258816] via 10.122.0.49, 00:00:08, GigabitEthernet1/0/28

!3750-dist-2
3750-dist-1# show ip route | b 10.122.10.102/32
D        10.122.10.102/32
[90/258816] via 10.122.0.53, 00:00:08, GigabitEthernet1/0/28
```

在HM网络环境中，接入层主机可采用以下两种方法来建立ISATAP隧道。

- 手工指明ISATAP路由器的IPv4地址。
- 手工设定ISATAP路由器的DNS名称（需要做DNS解析）。

在不辅以某种主机管理工具或脚本的情况下，在主机上使用第一种方法，虽简单但却不具备可扩展性。如前所述，可利用Microsoft Group Policy、Windows PowerShell以及Microsoft SMS Server等若干工具，于登录或其他预定时间，在接入层主机上运行本地命令。

对于本例，在接入层VLAN2中的Windows主机上，激活了ISATAP隧道功能，并指明了ISATAP路由器的IPv4地址（Windows Vista和Windows 7的IPv6功能为默认启用）。在MH设计中，还会将接入层某个VLAN/子网中的主机与某个特定的ISATAP路由器IPv4地址建立起关联，这在前文已有所提及。虽然未必非得如此行事，但只要现有的安全策略是根据某给定VLAN的源IP地址来执行，那么这就是推荐的做法。

对于本例，VLAN 2中的主机隶属于子网10.120.2.0/24，故而在这些主机上所指明的ISATAP路由器的IPv4地址为10.122.10.102，其中，“102”中的“2”是指VLAN 2或子网号2。同理，VLAN 3中的主机隶属于子网10.120.3.0/24，该VLAN主机使用的ISATAP路由器的IPv4地址为10.122.10.103。以下为Microsoft Windows主机上设置ISATAP路由器的**netsh**命令。

```
C:\> netsh interface ipv6 isatap set router 10.122.10.102 enabled
Ok.
```

可使用例6-35所示的命令去验证Windows主机是否接受了所配置的ISATAP路由器。

例 6-35 Microsoft Windows 主机上 ISATAP 路由器 IP 地址的 netsh 输出

```
C:\> netsh interface ipv6 isatap show router
Router Name                : 10.122.10.102
Use Relay                  : enabled
Resolution Interval        : default
```

由下例可知，主机已成功建立了一条通往主核心层交换机（6k-core-1）的 ISATAP 隧道连接，并获得了一个有效的 IPv6 地址 2001:db8:cafe:2:0: 5efe:10.120.2.101（通过 RA 消息获得，ISATAP 隧道未使用 DHCPv6）。该 IPv6 地址的 64 位接口 ID 来历如下：主机自身的 IPv4 地址占据了最右边的 32 位，最左边的 32 位则由 0000:5efe 或 0200:5EFE 填充。需要说明的是，若主机的 IPv4 地址为公网单播地址，最左边的 32 位便会以后者来填充。IPv4 地址 10.120.2.101 既可由主机通过 DHCP 获取，也可在主机上静态配置，该地址会被用作 ISATAP 隧道主机端的源地址，而核心层交换机上的 Loopback2 地址（10.122.10.102）则被用作隧道的目的地址（即例 6-35 中的 Router Name）。

例 6-36 所示为主机上的隧道适配器自动生成的伪隧道接口。

例 6-36 Microsoft Windows 主机的 IPv6 地址汇总

```
Connection-specific DNS Suffix  . : cisco.com
IP Address. . . . . . . . . . . . : 2001:db8:cafe:2:0:5efe:10.120.2.101
IP Address. . . . . . . . . . . . : fe80::5efe:10.120.2.101%2
Default Gateway . . . . . . . . . : fe80::5efe:10.122.10.102%2
```

在接入层主机上手工设定 ISATAP 路由器名称的方法同样颇为简单，但需要做 DNS 解析。假如想和前例一样，让特定的主机使用特定的隧道，在不辅以某种主机管理工具或脚本的情况下，即便启用了 DNS，该方法也不具备可扩展性。与前例一致，采用该方法时，也可利用 Microsoft Group Policy、Windows PowerShell 以及 Microsoft SMS Server 等若干工具，于登录或其他预定时间，在接入层主机上运行本地命令。此外，在 Windows 主机上，还可以使用组策略（Group Policy）来设定 ISATAP 路由器的名称。

现在，我们来看一下如何在 Windows 主机上设置 ISATAP 路由器名称（非 IPv4 地址）来建立 ISATAP 隧道。本例中，接入层主机建立隧道时，所要解析的 DNS 名称是“isatap”（外加域名后缀）。比方说，接入层主机隶属于“cisco.com”域，那么当其建立 ISATAP 隧道时，会尝试解析“isatap.cisco.com”。用户可以像更改 IP 地址那样去更改这一名称。例 6-37 所示为 Microsoft Windows 主机上的 ISATAP 路由器名称的配置。由配置可知，“vlan2-isatap”被用作为 DNS 名。

例6-37 Microsoft Windows主机上ISATAP路由器名称的netsh输出

```
C:\> netsh interface ipv6 isatap set router vlan2-isatap enabled
Ok.

C:\> netsh interface ipv6 isatap show router
Router Name              : vlan2-isatap
Use Relay                : enabled
Resolution Interval      : default
```

在DNS服务器上，还需为接入层VLAN2和VLAN3中的主机添加以下两条DNS记录。

- **vlan2-isatap**: Host (A) 10.122.10.102
- **vlan3-isatap**: Host (A) 10.122.10.103

6.4.4 QoS配置

用于HM的QoS策略应该沿袭现有的IPv4 QoS策略。在HM模型中实施QoS时，会遇到在何处分类和标记IPv6数据包的难题，这一点在前面已经提及。IPv6数据包从接入层主机向核心层交换机的转发过程中，自始至终都被封装在ISATAP隧道内，因此IPv6 QoS策略无法对封装在ISATAP隧道内的IPv6流量生效。核心层交换机的出站接口是可对IPv6数据包施以IPv6 QoS策略的首要之处。例6-38所示的配置只是作为一个简单的示例。在该配置所示的QoS策略中，class map用来匹配表6-6所列的IPv6访问列表和DSCP设置。

表6-6 IPv6 QoS-Class Map，匹配ACL和DSCP设置

应用程序	Access group名	DSCP设置
FTP	BULK-DATA	AF11
Telnet	TRANSACTIONAL-DATA	AF21
SSH	TRANSACTIONAL-DATA	AF21
所有其他应用	N/A	0（默认）

还要将IPv6 QoS策略（service policy）应用在核心层交换机的出站接口上（针对来自接入层的IPv6流量）。上游交换机可以信任表6-6所列的DSCP设置，必要时，还可以将其应用于排队和监管（请参见6.3节）。

例6-38所示为采用HM部署IPv6时的6k-core-1上的QoS配置示例。QoS策略非常简单，只包括了两个class-map，用来分类IPv6应用程序的流量，即由

10G 接口（流量外出接口）外发的 FTP、Telnet 和 SSH IPv6 流量。匹配 class-map 的流量都会被设置相应的 DSCP 值（见表 6-6）。上游交换机会信任 IPv6 流量的 DSCP 值，并会根据这些值来执行相应的监管和排队操作。

例 6-38 HM 模型——6k-core-1 QoS 配置示例

```
mls ipv6 acl compress address unicast
mls qos
!
class-map match-all CAMPUS-BULK-DATA
  match access-group name BULK-DATA

class-map match-all CAMPUS-TRANSACTIONAL-DATA
  match access-group name TRANSACTIONAL-DATA

!
policy-map IPv6-ISATAP-MARK

  class CAMPUS-BULK-DATA
   set dscp af11

  class CAMPUS-TRANSACTIONAL-DATA
   set dscp af21

  class class-default
   set dscp default

!
ipv6 access-list BULK-DATA

 permit tcp any any eq ftp

 permit tcp any any eq ftp-data

!

ipv6 access-list TRANSACTIONAL-DATA

 permit tcp any any eq telnet

 permit tcp any any eq 22

!
interface TenGigabitEthernet2/3
 description to 6k-core-1
```

（待续）

```
 mls qos trust dscp
 service-policy output IPv6-ISATAP-MARK

!
interface TenGigabitEthernet3/1
 description to 6k-agg-1
 mls qos trust dscp
 service-policy output IPv6-ISATAP-MARK

!
interface TenGigabitEthernet3/2 description to 6k-agg-2
 mls qos trust dscp
 service-policy output IPv6-ISATAP-MARK
```

6.4.5 基础设施安全性配置

在实施HM模型时，还应在接入层针对由ISATAP隧道封装的IPv6流量做更为严格的访问控制。可在接入层交换机下连主机或上联分布层交换机的端口上，应用访问列表来实现上述访问控制。就访问列表的配置方法而言，单就网络管理这一项来说，在接入层交换机的上行链路端口统一应用要比在每个下连主机的端口上分别应用，要省事的多。

可在接入层交换机上，针对主机所在的VLAN，配置ACL，去放行目的地址为ISATAP隧道路由器地址的ISATAP隧道流量。试举一例，下列ACL只会放行目的地址为10.122.10.102（先前配置的ISATAP隧道路由器地址）的ISATAP隧道流量（IP协议号为47）。ACL中最后一条语句的作用是放行所有其他的IPv4流量（我们控制的是被ISATAP隧道封装的IPv6流量，因此必须放行其他的IPv4流量）。此外，还需将该ACL应用在接入层交换机下连每台主机端口的入方向（**ip access-group 100 in**），或上连分布层交换机端口的出方向（**ip access-group 100 out**）。

```
access-list 100 remark Permit approved IPv6-Tunnels
access-list 100 permit 41 any host 10.122.10.102
access-list 100 deny 41 any any
access-list 100 permit ip any any
```

6.5 实施服务区块模型（SBM）

与实施SBM模型有关的ISATAP隧道的部署方式与HM模型几近相同。在这两种模型中，都会成对部署冗余的交换机来实现ISATAP隧道终结功能，来为建立ISATAP隧道的接入层主机提供高可用性。两种模型在ISATAP隧道部署方

式上的唯一差别是：实施 SBM 时，会新部署一组（对）交换机专门用来终结隧道（包括 ISATAP 隧道和手工配置的隧道），而在实施 HM 时，则会利用现有的核心层交换机来终结 ISATAP 隧道。

以下内容将会重点关注服务区块交换机的接口（逻辑和物理接口）配置，出于完整性的考虑，还会给出数据中心区块汇聚层交换机隧道接口的配置。而整个底层的 IPv4 网络配置与上一节所述的 HM 模型配置完全相同。

在本节示例中，配置在核心层交换机上的 ISATAP 路由器地址（loopback 接口地址）与上一节相同，因此接入层主机上与 ISATAP 路由器有关的配置也和 HM 模型相同。此外，与上一节相仿，本节也会呈现所涉网络设备的 loopback 接口、隧道接口、路由选择以及高可用性等相关配置。

6.5.1 网络拓扑

为了让网络拓扑图清晰易懂，现将其分两部分来绘制，即 ISATAP 隧道拓扑和手工配置隧道拓扑。

图 6-18 所示为实施 SBM 时的 ISATAP 隧道拓扑。该图重点显示了园区网络区块接入层主机的地址（主机用来建立 ISATAP 隧道的地址）、服务区块中的网络设备（ISATAP 隧道的终结点），以及服务区块网络设备所使用的 IPv6 地址，

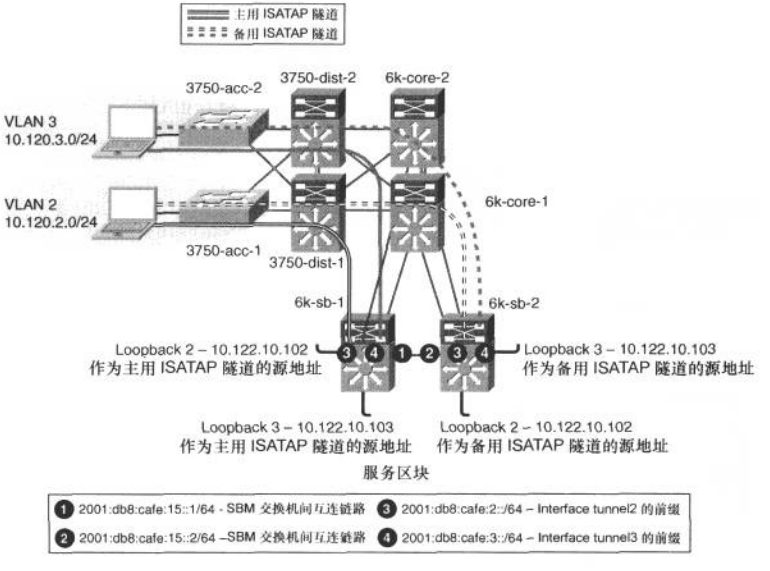


图 6-18 SBM ISATAP 网络拓扑

这些IPv6地址分别用于ISATAP隧道和设备互连。

图6-19所示为SBM交换机与数据中心网络区块汇聚层交换机之间点到点隧道的地址。考虑到所给配置的完整性，本节会给出点到点隧道的配置，该配置可作为互连SBM和数据中心服务的参考方案之一。当然，还使用其他方案（比如，ISATAP、6to4或手工配置的隧道），去连接数据中心内提供IPv6应用的每台服务器。此外，还可以在数据中心区块的汇聚层/接入层交换机与SBM交换机之间架设专用的IPv6链路，绕过核心层交换机与数据中心汇聚层交换机之间的非IPv6服务区域。一言以蔽之，在企业网内（园区网络区块接入层主机与数据中心区块接入层服务器之间）建立端到端IPv6连通性的方法可谓是多种多样。

图6-19所示的拓扑图示出了服务区块交换机上的loopback接口地址（用作为手工配置隧道的源地址），以及手工配置隧道所使用的IPv6地址。

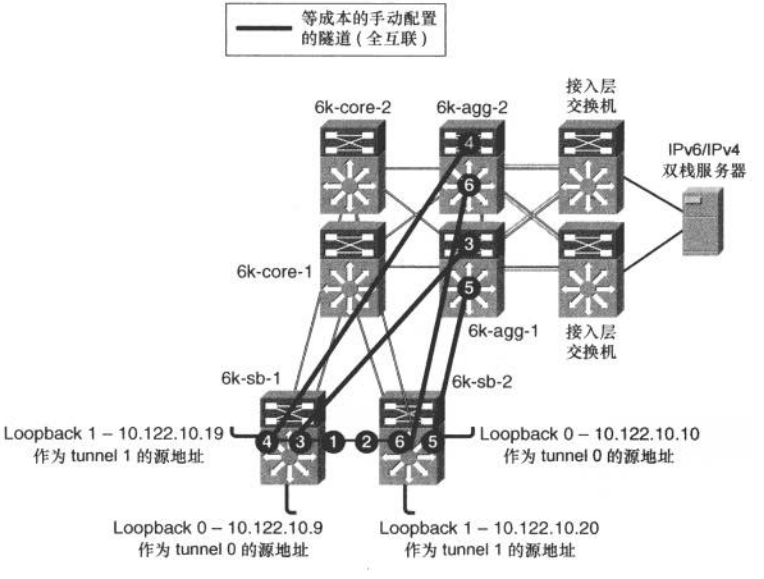


图6-19 SBM手工配置隧道拓扑结构图

6.5.2 物理接口的配置

本节会给出 2 台 SBM 交换机的配置，其中包括了与用来连接核心层交换机接口的配置。至于图 6-19 所示拓扑中与 IPv4 有关的配置，只会出现在 2 台 SBM 交换机的配置中。而所有其他与 IPv4 有关的配置全都谨遵 Cisco 园区网络设计最佳做法，这些内容已经超出了本节的范围。此外，如前所述，本场景完全支持 EIGRP 和 OSPFv3 等 IPv6 路由协议，来完成 IPv6 的路由选择。本节会给出 EIGRP 用于 IPv4 路由选择、OSPFv3 用于 IPv6 路由选择的配置示例。我们也针对 SBM 部署场景，做过了使用 EIGRP 完成 IPv4 和 IPv6 路由选择的实验，与本节所采用的路由选择方案相比，并无特别之处。

例 6-39 所示为 6k-sb-1 交换机互连核心层交换机接口的 IPv4 配置，以及互连 6k-sb-2 交换机接口的 IPv4 和 IPv6 的配置。

例 6-39 6k-sb-1 交换机接口配置

```
interface GigabitEthernet4/1
 description to 6k-core-1
 ip address 10.122.0.78 255.255.255.252       !IPv4-only connections to Core

 ip hello-interval eigrp 10 1
 ip hold-time eigrp 10 3
 ip authentication mode eigrp 10 md5
 ip authentication key-chain eigrp 10 eigrp
 mls qos trust dscp
!
interface GigabitEthernet4/2
 description to 6k-core-2
 ip address 10.122.0.86 255.255.255.252
 ip hello-interval eigrp 10 1
 ip hold-time eigrp 10 3
 ip authentication mode eigrp 10 md5
 ip authentication key-chain eigrp 10 eigrp
 mls qos trust dscp
!
interface GigabitEthernet4/3
 description to 6k-sb-2
 ip address 10.122.0.93 255.255.255.252
 ip hello-interval eigrp 10 1
 ip hold-time eigrp 10 3
 ip authentication mode eigrp 10 md5
```

（待续）

```
 ip authentication key-chain eigrp 10 eigrp
 mls qos trust dscp
!
interface GigabitEthernet4/3
 description to 6k-sb-2
 ip address 10.122.0.93 255.255.255.252
 ip hello-interval eigrp 10 1
 ip hold-time eigrp 10 3
 ip authentication mode eigrp 10 md5
 ip authentication key-chain eigrp 10 eigrp
 ipv6 address 2001:DB8:CAFE:15::1/64              ip2p link between SBM switches

 ipv6 ospf network point-to-point

 ipv6 ospf hello-interval 1

 ipv6 ospf dead-interval 3

 ipv6 ospf 1 area 0

 mls qos trust dscp
```

例 6-40 所示为 6k-sb-2 交换机的接口配置。除了编址以外，该配置与例 6-39 所示的配置相同。

例 6-40 6k-sb-2 交换机接口配置

```
interface GigabitEthernet4/1
 description to 6k-core-1
 ip address 10.122.0.82 255.255.255.252
 ip hello-interval eigrp 10 1
 ip hold-time eigrp 10 3
 ip authentication mode eigrp 10 md5
 ip authentication key-chain eigrp 10 eigrp
 mls qos trust dscp
!
interface GigabitEthernet4/2
 description to 6k-core-2
 ip address 10.122.0.90 255.255.255.252
 ip hello-interval eigrp 10 1
 ip hold-time eigrp 10 3
 ip authentication mode eigrp 10 md5
 ip authentication key-chain eigrp 10 eigrp
 mls qos trust dscp
```

（待续）

```
!
interface GigabitEthernet4/3
 description to 6k-sb-1
 ip address 10.122.0.94 255.255.255.252
 ip hello-interval eigrp 10 1
 ip hold-time eigrp 10 3
 ip authentication mode eigrp 10 md5
 ip authentication key-chain eigrp 10 eigrp
 ipv6 address 2001:DB8:CAFE:15::2/64              !p2p link between SBM switches
 ipv6 ospf network point-to-point
 ipv6 ospf hello-interval 1
 ipv6 ospf dead-interval 3
 ipv6 ospf 1 area 0
 mls qos trust dscp
```

6.5.3 隧道的配置

采用 SBM 部署 IPv6 时，SBM 交换机上的 ISATAP 隧道配置与实施 HM 时核心层交换机上的配置完全相同。为了避免重复，本节不再对与 ISATAP 隧道和路由选择有关的配置进行解释，但仍会给出相关配置（对配置的解释请见 6.4 节相关内容）。

例 6-41 所示为服务区块交换机 6k-sb-1 上与手工配置的隧道有关的配置。除了编址以外，该配置与数据中心区块汇聚层交换机（6k-agg-1/6k-agg-2）上相应的隧道配置完全相同。

例 6-41 SBM 6k-sb-1 上与手工配置的隧道有关的配置

```
interface Loopback0
 description Tunnel source for 6k-agg-1
 ip address 10.122.10.9 255.255.255.255
!
interface Loopback1
 description Tunnel source for 6k-agg-2
 ip address 10.122.10.19 255.255.255.255
!
interface Tunnel0
 description Manual Tunnel to 6k-agg-1
 no ip address
```

（待续）

```
 ipv6 address 2001:DB8:CAFE:110::1/64
 ipv6 ospf network point-to-point
 ipv6 ospf hello-interval 1
 ipv6 ospf dead-interval 3
 ipv6 ospf 1 area 0
 tunnel source Loopback0                    !Anchor p2p tunnel to Lo0

 tunnel destination 10.122.10.1             !10.122.10.1 is loopback0 on 6k-agg-1
 tunnel mode ipv6ip                         !IPv6-in-IPv4 tunnel

!
interface Tunnel1
 description Manual Tunnel to 6k-agg-2
 no ip address
 ipv6 address 2001:DB8:CAFE:111::1/64
 ipv6 ospf network point-to-point
 ipv6 ospf hello-interval 1
 ipv6 ospf dead-interval 3
 ipv6 ospf 1 area 0
 tunnel source Loopback1

 tunnel destination 10.122.10.2             !10.122.10.2 is loopback0 on 6k-agg-2
 tunnel mode ipv6ip
```

例6-42所示为6k-sb-1上与ISATAP隧道有关的配置。出于简化，不再显示6k-sb-2上相应的配置；除了编址以外，其他均无不同。

例6-42 SBM 6k-sb-1上与ISATAP隧道和路由选择有关的配置

```
interface Loopback2
 description Tunnel source for ISATAP-VLAN2
 ip address 10.122.10.102 255.255.255.255
!
interface Tunnel2
 description ISATAP VLAN2
 no ip address
 ipv6 address 2001:DB8:CAFE:2::/64 eui-64
 no ipv6 nd suppress-ra
 ipv6 ospf 1 area 2
 tunnel source Loopback2
 tunnel mode ipv6ip isatap
!
ipv6 router ospf 1
```

（待续）

```
router-id 10.122.10.9
log-adjacency-changes
auto-cost reference-bandwidth 10000
area 2 range 2001:DB8:CAFE:2::/64 cost 10
area 2 range 2001:DB8:CAFE:3::/64 cost 10
passive-interface Loopback0
passive-interface Loopback1
passive-interface Loopback2
passive-interface Loopback3
passive-interface Tunnel2
passive-interface Tunnel3
timers spf 1 5
```

6.5.4 QoS 配置

6.3 节所讨论的 QoS 配置原则同样适用于 SBM 模型。与实施 HM 模型时的 QoS 配置相比，只有应用 QoS 分类和标记策略的接口位置发生了变化。对于 SBM，service policy（QoS 分类和标记策略）要应用在 SBM 交换机通往 6k-agg-1 和 6k-agg-2 的手工配置的隧道接口上，其方向应该是出站方向。

例 6-43 所示为如何在 6k-sb-1 交换机的 Tunnel0 和 Tunnel1 接口上应用 service policy。

例 6-43 在 6k-sb-1 上隧道接口的出站方向应用 QoS 策略

```
interface Tunnel0
 description tunnel to 6k-agg-1
 service-policy output IPv6-ISATAP-MARK

!
interface Tunnel1
 description tunnel to 6k-agg-2
 service-policy output IPv6-ISATAP-MARK
```

6.3 节和 6.4 节所讨论的与安全性有关的配置和部署原则对 SBM 的部署全都适用。

6.6 总结

本章对企业园区网络区块中提供 IPv6 服务的几种部署模型逐一进行了分

析。由于可省去“开凿”隧道之苦，故而对于任何将要在园区网络中部署 IPv6 的企业来说，DSM 都应作为其首选目标。HM 部署模式也具有一定的实用价值，这是因为该模式可利用现有的网络基础设施，通过“开凿”ISATAP 隧道的形式，为园区网络区块接入层的末端主机提供 IPv6 服务。而 SBM 模式则是一种极为理想的临时性解决方案，在企业网内，可利用该解决方案在园区网络区块的接入层主机和数据中心网络区块内（或 Internet 上）提供 IPv6 服务的服务器之间，提供端到端的访问，而无需改动现有网络中的硬件布局。

本章所讨论的 IPv6 部署模型肯定不会是相应网络环境中硕果仅存的部署方法，但却为读者提供了部署 IPv6 的各种选项，读者可根据自己的网络环境、部署时间表，以及网络服务的具体情况，来充分利用这些选项。

表 6-7 总结了本章所讨论的三种 IPv6 部署模型的优势和短板。

表 6-7 种 IPv6 部署模型的优势与短板

模　型	优　势	短　板
双栈模型	无需“开凿”隧道 不依赖 IPv4 可为 IPv6 单多播提供卓越的性能和最高级别的高可用性 具备良好的可扩展性	要求园区网络区块交换设备具备硬件方式转发 IPv6 数据包的能力 在网络中运行双栈协议会面临运维和网络管理等方面的挑战
混合模型（HM）	可充分利用现有园区网络中只支持 IPv4 的网络设备（接入层和分布层设备） 可以每用户或每应用程序每基础来控制 IPv6 服务的交付 可部署冗余的 ISATAP 隧道，来实现 IPv6 访问的高可用性	需“开凿”隧道；会增加运维和管理成本 可扩展性因素（隧道数、每条隧道所承载的主机数） 不支持 IPv6 多播 隧道终结于核心层
服务区块模型（SBM）	限制缩短了交付 IPv6 服务的工期 无需对现有网络基础设施做任何改动 其他优势与 HM 类似	需新购支持 IPv6 功能的交换机 集成了 HM 所具备的所有缺点

6.7 参考资料

本章讨论了需要读者充分理解的 IPv6 技术及协议等相关内容，涵盖了许多与实施 IPv6 有关的设计考虑，其中包括安全性、QoS、高可用性、网络管理、IT 培训以及应用程序的支持等。

以下列出的参考资料包括了与 IPv6、Cisco 网络设计建议、产品和解决方案

以及业界活动有关的详细信息，但这些信息只不过是浩瀚如烟的技术文章中的冰山一角。

Arkko, J., ed., Kempf, J., Zill, B., and Nikander, P. RFC 3971, "SEcure Neighbor Discovery (SEND)."

Cisco. "Catalyst 3750-E and 3560-E IPv6 and Switch Stacks." http://www.cisco.com/en/US/docs/switches/lan/catalyst3750e_3560e/software/release/12.2_46_se/configuration/guide/swipv6.html#wp1091926.

Cisco. "Catalyst 6500 Software Configuration Guide - Configuring Denial of Service (DoS) Protection - CoPP." http://www.cisco.com/en/US/docs/switches/lan/catalyst6500/ios/12.2SXF/native/configuration/guide/dos.html.

Cisco. "Catalyst 6500 Software Configuration Guide - Configuring PFC QoS." http://www.cisco.com/en/US/docs/switches/lan/catalyst6500/ios/12.2SX/configuration/guide/qos.html.

Cisco. "Catalyst 6500 Software Configuration Guide - IPv6 ACL Compression." http://www.cisco.com/en/US/docs/switches/lan/catalyst6500/ios/12.2SXF/native/configuration/guide/acl.html#wp1090842.

Cisco. "Catalyst 6500 Virtual Switching System 1440." http://www.cisco.com/en/US/products/ps9336/index.html.

Cisco. "Cisco DHCPv6 Server in IOS." http://www.cisco.com/en/US/docs/ios/ipv6/configuration/guide/ip6-dhcp_ps6441_TSD_Products_Configuration_Guide_Chapter.html#wp1323295.

Cisco. "Cisco First Hop Security." http://www.cisco.com/en/US/docs/ios/ipv6/configuration/guide/ip6-first_hop_security.html.

Cisco. "Cisco IOS IPv6 Configuration Guide - HSRP for IPv6." http://www.cisco.com/en/US/docs/ios/ipv6/configuration/guide/ip6-fhrp.html#wp1055254.

Cisco. "Cisco IOS IPv6 Configuration Guide - Implementing IPv6 Multicast - SSM Mapping." http://www.cisco.com/en/US/docs/ios/ipv6/configuration/guide/ip6-multicast_ps10591_TSD_Products_Configuration_Guide_Chapter.html#wp1058805.

Cisco. "Cisco IPv6." http://www.cisco.com/web/solutions/netsys/ipv6/index.html.

Cisco. "Cisco IPv6 Multicast." http://www.cisco.com/en/US/products/ps6594/products_ios_protocol_group_home.html.

Cisco. "Cisco IPv6 Start Here Guide and Roadmap." http://www.cisco.com/en/US/docs/ios/ipv6/configuration/guide/ip6-roadmap.html.

Cisco. "Cisco Network Management and Automation." http://www.cisco.com/en/US/products/sw/netmgtsw/index.html.

Cisco. "Cisco Network Registrar." http://www.cisco.com/en/US/products/sw/netmgtsw/ps1982.

Cisco. "Design Zone for Campus - Routed Access Layer using EIGRP or OSPF." http://www.cisco.com/en/US/netsol/ns815/networking_solutions_program_home.html.

Cisco. "Enterprise QoS SRND." http://www.cisco.com/en/US/docs/solutions/Enterprise/WAN_and_MAN/QoS_SRND/QoS-SRND-Book.html.

Cisco. "Medianet Campus QoS Design 4.0." http://www.cisco.com/en/US/docs/solutions/Enterprise/WAN_and_MAN/QoS_SRND_40/QoSCampus_40.html.

Cisco. "Migrate Standalone Cisco Catalyst 6500 Switch to VSS." http://www.cisco.com/en/US/products/ps9336/products_tech_note09186a0080a7c74c.shtml.

Cisco. "Routed Access Q&A." http://www.cisco.com/en/US/netsol/ns340/ns394/ns147/ns17/netqa0900aecd8045965a.html.

Cisco. "Routing in the Wiring Closet White Paper." http://www.cisco.com/en/US/netsol/ns340/ns394/ns147/ns17/networking_solutions_white_paper0900aecd804c6e73.shtml.

Davies, Joseph. *Understanding IPv6*, second edition. (ISBN-10: 0-7356-2446-1. ISBN-13: 978-0-7356-2446-7).

Deering, S. and Hinden, R. RFC 2460, "Internet Protocol, Version 6 (IPv6) Specification."

Hinden, R. and Deering, S. RFC 3513, "Internet Protocol Version 6 (IPv6) Addressing Architecture."

Hogg, Scott and Vyncke, Eric. *IPv6 Security*. (ISBN-10: 1-58705-594-5. ISBN-13:978-1-58705-594-2).

McCann, J., Deering, S., and Mogul, J. RFC 1981, "Path MTU Discovery for IP version 6."
Microsoft. "Microsoft–Cisco ISATAP White Paper." http://www.microsoft.com/downloads/details.aspx?FamilyId=B8F50E07-17BF-4B5C-A1F9-5A09E2AF698B&displaylang=en.

Microsoft. "Microsoft IPv6 Home." http://technet.microsoft.com/en-us/network/bb530961.aspx.

Microsoft. "Microsoft Teredo Overview." http://technet.microsoft.com/en-us/library/bb457011.aspx.

Microsoft. "Microsoft Windows Server 2008 - DHCP Server." http://technet.microsoft.com/en-us/library/cc896553%28WS.10%29.aspx.

Narten, T., E. Nordmark, W. Simpson, and H. Soliman. RFC 4861, "Neighbor Discovery for IP version 6 (IPv6)." http://www.ietf.org/rfc/rfc4861.txt.

Popoviciu, Ciprian P.; Eric Levy-Abegnoli, and Patrick Grossetete. *Deploying IPv6 Networks.* ISBN-10: 1-58705-210-5. ISBN-13: 978-1-58705-210-1.

Savola, P. RFC 3627, "Use of /127 Prefix Length Between Routers Considered Harmful."

Savola, P. RFC 3956, "Embedding the Rendezvous Point (RP) Address in an IPv6 Multicast Address."

Szigeti, Tim and Christina Hattingh. *End-to-End QoS Network Design.* (ISBN-10: 1-58705-176-1. ISBN-13: 978-1-58705-176-0(.

Templin, F., T. Gleeson, and D. Thaler. RFC 5214, "Intra-Site Automatic Tunnel Addressing Protocol (ISATAP)."

Van de Velde, G., C. Popoviciu, T. Chown, O. Bonnes, and C. Hahn. RFC 5375, "IPv6 Unicast Address Assignment Considerations.

CHAPTER 7

第 7 章　部署虚拟化的 IPv6 网络

本章涵盖以下主题。

- **虚拟化概述**：本节会简要介绍虚拟化对设计和部署 IPv6 网络的好处和影响，以及虚拟化技术的分类。
- **网络虚拟化**：本节会围绕如何将 IPv6 与各种网络虚拟化技术（比如，交换机虚拟化和网段隔离[network segmentation]等）相结合展开讨论，并提供详尽的设计步骤和配置指导。网络架构师在将 IPv6 集成进现有网络之前，必须要对各种网络虚拟化设计方案及需求有着透彻的理解。
- **网络服务虚拟化**：对网络虚拟化技术的讨论在本节将会被扩大至 4～7 层。本节概述了启用 IPv6 时桌面虚拟化的设计原则。

本章涵盖了在虚拟化网络中部署 IPv6 的设计和实施技术。第一节会简单介绍什么是虚拟化，并连带对各种类型的虚拟化技术做简单概述。接下来，还会深入讨论网络服务和桌面 IPv6 虚拟化技术的设计、配置和验证。

“网络虚拟化”一节将会对如何利用 6PE 和 6vPE 技术，跨 MPLS 骨干网互连 IPv6 网络做深入探讨。此外，本节还会简单介绍服务虚拟化、服务虚拟化的优势，以及如何设计、配置和验证 Cisco ASA 5580 多上下文 IPv6 防火墙。本章的最后两节则侧重于介绍桌面和服务器虚拟化技术，但不包括对两者的部署，其原因是如今的绝大多数瘦客户端应用还不支持 IPv6。

7.1　虚拟化概述

现代化企业不但要求其 IT 基础设施能够灵活地去支持不断变化中的商业模式，而且还要在降低成本的同时，推动新型商业应用程序的上线运行。虚拟化技术可帮助企业充分利用现有的网络基础设施，来满足新型应用程序上线的商业需

求。此外，虚拟化技术的引入也从根本上颠覆了企业IT管理者对计算机、网络以及存储资源的认识。通过虚拟化技术，可利用抽象化的硬件资源来提供服务，这也使得IT的重点从技术转移到了由技术所提供的服务。

以下各节将会简单介绍启用虚拟化技术的优势，并会对网络、服务器（计算）以及存储基础设施的虚拟技术化加以分类。

7.1.1 虚拟化的优势

虚拟化并不是一个全新的概念，该技术已在大型机上使用了不下十年的时间。如今，虚拟化的理念仍未过时，并为现代化的个人计算机所用，来对其物理资源进行抽象。虚拟化技术具备以下优势。

- **通过改善IT基础设施的效率，来降低总拥有成本**：随着计算机计算能力的增长，不能“物尽其用”的可能性也随之增大。相较于为每一种应用程序配备专用的资源，采用虚拟化技术，可优化企业对硬件资源的使用。
- **节能降耗**：采用虚拟化技术，对现有基础设施的充分利用，可为企业实现节能降耗的目标。通过节能降耗，来创建一个符合环保要求的网络和基础设施环境。
- **通过资源共享来提高网络的灵活性**：虚拟化技术可确保各类应用程序共存于公共的物理硬件之上，而相安无事。
- **改进资源分配**：出于恪守标准、审计或性能方面的原因，可把独立的应用程序部署在虚拟化的基础设施之上，以逻辑的方式来隔离应用程序，以便对每种应用程序所需的特定资源进行控制。

7.1.2 虚拟化的分类

一般而言，当IT架构师言及术语虚拟化时，通常都意指服务器虚拟化，但有时，虚拟化也并不仅仅用来指代计算。如今，虚拟化的概念已从服务器延伸至了网络、存储以及桌面资源。表7-1列出了虚拟化的4个基本类别。

在将目光投向网络、服务器以及桌面的虚拟化之后，企业客户又开始将上述虚拟化技术扩展到了IPv4和IPv6网络。

表 7-1 虚拟化的分类

分 类	定 义	优 势
网络虚拟化	网络虚拟化包括交换机虚拟化： 类似于设备结对（VSS[虚拟交换系统]）和虚拟设备上下文（VDC）之类的技术 网络分割： 将可用带宽划分为可供不同主机使用的独立信道，从而整合网络资源 网络服务虚拟化： 将虚拟化扩展至网络设备（路由器、交换机）之外的四至七层节点	安全性的提升： 数据包转发路径的分离使得 A 网段的流量对 B 网段不可见。将广播限制在了本地网络。对外界屏蔽了内部网络的结构 避免网络拥塞： 对于被分割的网络，每个子网只有少量主机，可最大限度的降低了本地流量，故而提升了网络性能 遏制网络故障： 限制本地网络故障对网络其他部分的影响 提高运维效率： 提高服务节点的利用率 资源分配： 视应用程序的需求，为每个服务设备的实例分配资源[①]
服务器虚拟化	服务器虚拟化技术能够让多个虚拟服务器实例在单台物理服务器上运行。服务器虚拟化是由 VMware ESX、Xen、Linux KVM 之类的虚拟机管理程序来驱动	将服务器整合： • 改进了灾难恢复机制 • 利用虚拟机的灵活性来缩短宕机时间 • 缩短了部署新服务和新应用程序的时间
桌面虚拟化	桌面虚拟化是指在数据中心内基于虚拟机管理程序的系统上，架设虚拟机部署客户端操作系统	以比较简单的方式提供新的桌面系统： • 在服务器或客户端故障时，缩短宕机时间 • 以较低的成本来部署新的应用 • 桌面映像管理（Desktop image-management）功能 • 有效地保护知识产权
存储虚拟化	存储虚拟化技术可对分布在多台网络存储设备的物理存储加以整合，令其作为单台存储设备来提供服务	通过优化存储的使用降低存储成本： • 以分布式的存储来提供更合理的灾难恢复能力，并降低了企业用户的存储成本

本章只对存储虚拟化技术做一般性的介绍，并不会深入讨论该技术。

7.2 网络虚拟化

以下各节将会介绍如何在网络虚拟化解决方案中启用 IPv6。网络虚拟化包

① 原文是“Allocates resources to each instance of the services device depending on application needs”，“instance of the services device”译者也不知道是什么玩意，按字面直译为“服务设备的实例”。

括以下几个方面。

- 交换机虚拟化。
- 网络隔离。
- 网络服务虚拟化。

以下各节将会深入讨论网络虚拟化的这三个方面。

7.2.1 交换机虚拟化

交换机虚拟化是指让多台物理交换机像单台物理交换机那样运作。以下列出了两种主要的交换机虚拟化技术。

- **设备共用（Device pooling）（VSS）**：VSS之类的设备共用技术可让网管人员以逻辑的方式将两台物理交换机组合为单一管理域，从而简化自己所管理的网络。除了能够让网络管理变得更为简单以外，VSS还能支持active-active的数据转发路径，通过部署MEC（多机箱EtherChannel）来消除生成树之需[①]。本书的第6章涵盖了如何在IPv6网络中配置和部署VSS的详细内容。
- **Nexus 7000系列交换机上的虚拟设备上下文（VDC）**[②]：网管人员可利用VDC将单台交换机划分为多个虚拟上下文，其中每个上下文都拥有单独的管理平面，以及可针对不同网段/应用程序来实施的各种策略。本书的第9章将会介绍Nexus 7000系列交换机的IPv6配置。

7.2.2 网络隔离

本节主要介绍各种网络隔离技术，其中包括虚拟路由转发实例（VRF）、IPv6 L2/L3虚拟专用网（6PE/6VPE、VPLS）等。此外，本节还会讨论“纵贯”现有IPv4/MPLS网络，部署IPv6网络的设计和配置细节。以下列出了本节所要讨论的网络隔离技术。

- **虚拟路由转发实例（VRF）**：VRF可用在同一物理设备上隔离不同网段（比如，语音VLAN和guest VLAN）。
- **跨MPLS骨干网传输IPv6数据包（6PE/6VPE）**：在MPLS网络中，可

① 原文是“In addition to simplified management plane, VSS also enables active-active data paths, thereby eliminating the need for Spanning Tree by deploying Multi-chassis EtherChannels (MEC)”。

② 原文是“Virtual Device Contexts (VDC) on Nexus 7K”。

利用 6PE 来传输 IPv6 数据包，6VPE 则用在 MPLS VPN 网络环境中传输 IPv6 数据包。

- **虚拟专用 LAN 服务（VPLS）**：可使用 VPLS 技术，来互连分居异地的 IPv6 网络（L2 VPN）。

虚拟路由和转发（VRF-Lite）

在大型企业网络和服务提供商网络中，VRF 一般都会与 MPLS 结合使用。不过，虽然以上两种技术配合起来也算天衣无缝，但在路由器上，并不见非得配置 MPLS，然后才能激活 VRF。Cisco 为只激活了 VRF 特性，但未启用 MPLS 功能的部署方式起名为——VRF-Lite 或 Multi-VRF。本节将详述如何利用 VRF-Lite 技术去隔离 IPv4 和 IPv6 流量，以及如何分别针对 IPv4 和 IPv6 网络维护独立的 IP 路由表。

VRF-Lite 特性运作于第三层，该特性可允许同一台路由器拥有多个路由表实例，不同的路由表实例可包含相同或重叠的 IP 地址范围。在启用了 VRF-Lite 功能的情况下，路由器会根据分配给各 VRF 的接口来创建多张路由表和与之相对应的多张转发表。这是每个 VRF 保持都“与世隔绝”的前提条件，读者应对这一简单概念牢牢掌握。当然，要是想把流量从一个 VRF 路由至另一个 VRF，情况会复杂的多。

图 7-1 所示为一个拥有多个 IP 子网的企业网络，各 IP 子网之间被完全隔离，这就是说，这些 IP 子网被分别置入了隔离的 VRF。当网络架构师寻求在企业现有网络中集成一个 IPv6 主机网络，且该网络的流量和路由要与企业中的其他网络完全隔离时，那么便可利用 VRF-Lite 特性，来提出一个既简单而又强大的解决方案。

该企业网园区网络被划分为三个 VLAN。

- **VLAN 10（VRF-RED）**：只为 VRF Red 中的主机分配 IPv6 地址。将这些主机全部置入一个隔离的 VLAN——VLAN 10；同时利用 802.1Q Trunk，将 VLAN 10 透传至上游的分布层交换机。VRF Red 中的用户连接到 3 号楼中的接入层交换机。为 RED 网络中的主机分配的 IPv6 前缀是 2001:DB8:CAFE:2::/64。
- **VLAN 20（VRF-BLUE）**：为 VRF Red 中的主机同时分配 IPv4 和 IPv6 地址。VRF BLUE 中的用户连接到 2 号楼中的接入层交换机。为 Blue 网络分配的 IPv4 前缀为 92.168.1.0/24，IPv6 前缀为 2001:DB8:CAFE:1::/64。

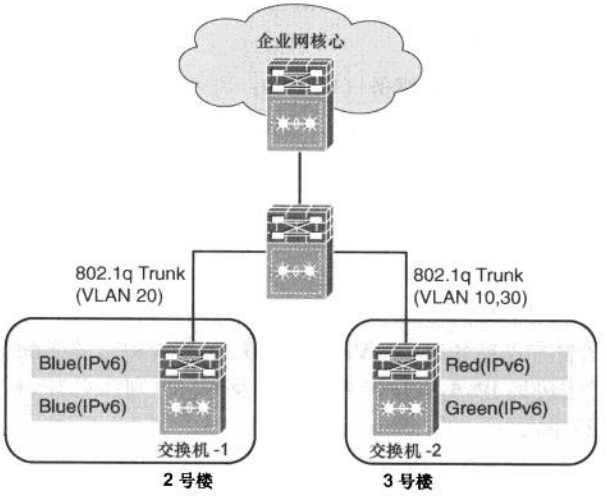


图 7-1 启用 VRF-Lite 功能的企业园区网络

- **VLAN 30（VRF-GREEN）**：只为 VRF Green 中的主机分配 IPv4 地址，但该 VRF 中的主机有可能会迁移到 IPv6。RF Green 中的用户连接到 3 号楼中的接入层交换机。为 Green 网络分配的 IPv4 前缀为 202.10.12.0/24。

2 号楼和 3 号楼内的交换机都通过 802.1Q Trunk 链路连接到分布层交换机（在本例中，分布层交换机为虚拟交换系统）。3 个 VRF 中的用户都依靠同一网络基础设施来建立连通性，但出于安全性方面的考虑，各 VRF 之间彼此隔离。例 7-1 所示为交换机-1 和交换机-2 的配置。

例 7-1 交换机上的 VRF 配置

```
switch-1#
!
vrf definition BLUE
 rd 100:3
 route-target export 100:3
 route-target import 100:3
!
 address-family ipv4
 exit-address-family
!
```

（待续）

```
 address-family ipv6
 exit-address-family
! /* VRF BLUE with IPv4 and IPv6 Addresses */
interface VLAN 20
 vrf forwarding BLUE
 ip address 192.168.1.1 255.255.255.0
 ipv6 address 2001:DB8:CAFE:1::1/64
!
interface GigabitEthernet1/3
 switchport
 switchport mode access
 switchport access vlan 20
!
!/* For the trunk on switch -1 */
interface GigabitEthernet1/3
 switchport
 switchport mode trunk
 switchport trunk allowed 20
!
switch-2#
vrf definition GREEN
 rd 100:2
 route-target export 100:2
 route-target import 100:2
 !
 address-family ipv4
 exit-address-family
!
vrf definition RED
 rd 100:1
 route-target export 100:1
 route-target import 100:1
 !
 address-family ipv6
 exit-address-family
!/* VRF RED with IPv6 address */
interface VLAN 10
 vrf forwarding RED
 ipv6 address 2001:DB8:CAFE:2::1/64
!
interface GigabitEthernet1/1
 switchport
 switchport mode access
 switchport access vlan 10
```

（待续）

```
!
interface GigabitEthernet1/2
 switchport
 switchport mode access
 switchport access vlan 10
! /* VRF GREEN with IPv4 Addresses */
interface VLAN 30
 vrf forwarding GREEN
 ip address 202.10.12.1 255.255.255.0
!
interface GigabitEthernet1/5
 switchport
 switchport mode access
 switchport access vlan 30
!
```

注意

将某接口划入VRF时，IOS会自动删除该接口上任何预配置的IP地址，只有这样，才能将与该IP地址相对应的主机路由从全局路由表中删除。此后，再在该接口上设置IP地址时，与其IP地址相对应的明细路由便会在VRF路由表中现身。

配毕VRF以及与VRF相关联的接口之后，网管人员可执行**show vrf**命令对配置进行验证（见例7-2）。

例7-2 显示show vrf命令的信息

```
switch-2# show vrf
  Name                                     Default RD          Protocols   Interfaces
  GREEN                                    100:2               ipv4        Gi1/5
  RED                                      100:1               ipv6        Gi1/1
                                                                           Gi1/2
```

注意

VRF与VLAN一样，只对交换机（路由器）本机生效。在分布层交换机上，也需要针对这3个VRF完成类似的配置。

因为Multi-VRF特性为每个VRF域都分别提供了一组接口、一张路由表以及一张转发表，而VRF-Lite信息对与交换机（路由器）来说只具有本机意义，故而相关信息不能在整个网络中传播。为此，人们设法将标签交换路径从PE路由器延伸到了CE（客户边缘）路由器。如此一来，单台CE路由器便可以像PE

路由器那样，作为多台虚拟路由器来运行，如图 7-2 所示。只有这样，网管人员才能充分利用 MPLS/VPN 技术，把 VRF-Lite 映射为 VPN-ID，“纵贯”核心网络，传播 VRF-Lite 信息[①]。

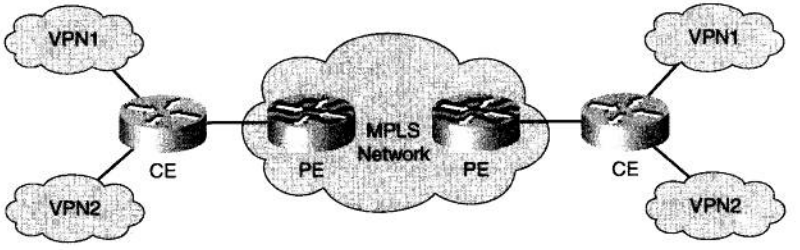


图 7-2 CE 路由器作为多台虚拟路由器运行

跨 MPLS 骨干网传输 IPv6 数据包

随着某些大企业的 IT 基础设施规模的不断扩张，其角色也已转变为企业内的“服务提供商”骨干网，并为异地的不同业务部门提供服务。此类大企业可利用 MPLS 作为底层传输技术，为分处异地的站点交付 L3 VPN 解决方案。除了 L3 VPN 功能以外，还可以利用 MPLS 来实施流量工程，这便能有助于 IT 部门在不同的链路之间，控制和安排不同的 VPN 流量，并可据此对各业务部门征收或要求分摊长途链路的费用。当此类企业考虑在自己的网络中部署 IPv6 时，还需“纵贯”MPLS 网络互连各个孤立的 IPv6 网域。以下各节将会重点介绍如何在整个企业网中扩展（只对路由器本机有意义的）VRF-Lite 特性。网络架构师可利用 Cisco 6PE 和 6VPE 技术，在现有的 IPv4 MPLS 基础设施中部署 IPv6 服务。

6PE

BGP 对等体之间建立会话时，会发送包含可选参数的 BGP OPEN 消息。BGP 能力是可选参数之一。BGP 能力包括多协议扩展、路由重刷新、出站路由过滤（ORF）等。交换多协议扩展能力时，BGP 对等体会顺带交换地址家族标识符（AFI）和子地址家族标识符（SAFI），并以此来确定与己对等的 BGP 路由器的能力。BGP AFI 和 SAFI 值请见表 7-2。

---

① 译者迫不得已改写了整段原文，现给出整段原文“Being locally significant to the router, VRF-Lite information cannot be carried across the entire network, because the Multi-VRF provides each VRF domain with its own set of interfaces, routing table, and forwarding tables. Thus, it enables the extension of the label-switched paths to the Customer Edge (CE) routers. As shown in Figure 7-2, this allows each CE router to act as several different CEs. Thus, the network architects can leverage MPLS VPN technology to map VRF-Lite to a VPN-ID and carry it across the core of the network”。VRF-Lite 和 Multi-VRF 特性如此复杂，以作者写作能力要是还能说清的话，那真是咄咄怪事了。

表 7-2 BGP AFI 和 SAFI 值

BGP 属性	详 细 信 息
地址家族标识符（AFI）	1：IPv4 2：IPv6
子地址家族标识符（SAFI）	1：单播 2：多播 3：单播和多播 4：MPLS 标签 128：带 MPLS 标签的 VPN（MPLS-labeled VPN）

与 MPLS/VPN 一样，BGP 的 IPv6 也是通过多协议 BGP（MP-BGP）（RFC 4760）来实现，其路由则由两个新的属性来承载，这两个新属性分别是 MP_UNRE ACH_NLRI 和 MP_REACH_NLRI[①]。MP_UNREACH_NLRI 携带不可达目的网络信息，而 MP_REACH_NLRI 则携带可达目的网络信息。由于这两个属性都是可选属性，因此路由通告路由器便可与不支持此类能力的 BGP 路由器交互，以此来保持后向兼容能力，亦即：无论 BGP 路由器支持该属性与否，彼此间均可相互通信[②]。

如果 BGP 携带的是 IPv6 路由信息，对于单/多播路由信息来说，路由信息中所包含的 AFI 值总为 1，SAFI 值则分别为 1 和 2。这样一来，6PE 路由器在通告 IPv6 前缀时，会将自己的 IPv4 地址作为相应路由的 BGP 下一跳。在 IPv6 BGP 路由的下一跳字段中，出站 6PE 路由器（即通告 IPv6 路由的 BGP 路由器）的 IPv4 地址会被编码为由 IPv4 地址映射而成的 IPv6 地址（IPv4-mapped IPv6 address），其格式为“::FFFF:<IPv6 BGP 路由的 IPv4 下一跳地址>”。此外，入站 6PE 路由器（接收 IPv6 路由通告的 BGP 路由器）会给 IPv6 前缀绑定一个标签。MP-BGP 所使用的 SAFI 为“标签”（其值为 4）。标签绑定信息也会被编码进 MP_REACH_NLRI 属性，随前缀信息一并通告。

因为 IPv6 路由的下一跳是由 IPv4 地址映射而成的 IPv6 地址，那么需要转发 IPv6 数据包的 6PE 路由器便可通过检查 MP-BGP 路由信息，来“发现”一条基于 IPv4 的（IPv4-enabled）标签交换路径（LSP），并使用这条 LSP 去转发特定 IPv6 目的地址的数据包[③]。可使用标签分发协议（LDP）或资源预留协议-流

① 原文是“IPv6 in BGP is implemented through Multi-Protocol BGP (MP-BGP) (RFC 4760), as are MPLS and VPNs through two new attributes: MP_UNREACH_NLRI and MP_REACH_ NLRI”。译者实在不晓得作者想表达什么，只能勉强翻译。

② 原文是“As these attributes are optional so the allow the speaker to talk to the other BGP routers which don't support these capabilities, this allows backward compatibility as the BGP speakers without these attributes can speak to the BGP speakers with these attributes”。这种文字实在没法翻译，只能勉强翻译加杜撰。

③ 原文是“The IPv4-mapped IPv6 addresses allow a 6PE router that has to forward an IPv6 packet to automatically determine the IPv4-enabled Label Switched Path (LSP) to use for a particular IPv6 destination by looking at the MP-BGP routing information”。

量工程（RSVP-TE）去建立那条基于 IPv4 的 LSP。

当 IPv6 数据包进入 IPv4 核心网络时，入站 6PE 路由器会对 IPv6 数据包执行标签压入操作。该路由器会首先压入由出站 6PE 路由器利用 MP-BGP 通告给自己的内层标签，出站 6PE 路由器收到附着了此类标签的 MPLS 帧时，便知其封装的是 IPv6 数据包。入站 6PE 路由器还会在流入的 IPv6 数据包报头前压入一个外层标签，该标签与通过 IPv4 建立起来的 LSP 相对应，这条 LSP“纵贯”出/入站 6PE 路由器。图 7-3 所示为 6PE 网络。

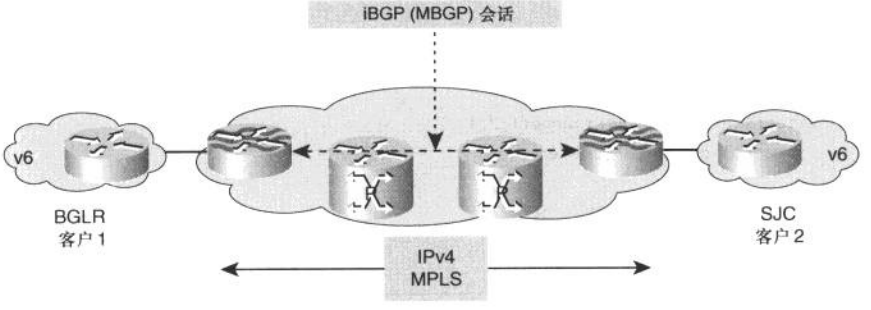


图 7-3 6PE 网络

图 7-3 中的 PE 和 CE 路由器之间运行的动态路由协议是下一代 RIP（RIPng）。PE 和 CE 设备都是虚拟交换系统（VSS）。以下列出了 6PE 的配置步骤。

步骤 1 配置 CE 路由器，如例 7-3 所示。

例 7-3 CE 路由器的接口配置

```
BGLR-CE:                                  # Bangalore CE Router

ipv6 unicast-routing
!
interface Loopback0
 ipv6 address 2001:DB8:CAFE:1:1::1/128
!
interface Port-channel30
 description to PE-VSS-core-BGLR
 ipv6 address 2001:DB8:CAFE:2::1/64      #Assign IPv6 address
  !
interface TenGigabitEthernet1/7/1
  channel-group 30 mode on
!
```

（待续）

```
interface TenGigabitEthernet1/7/1
  channel-group 30 mode on
!
SJC-CE:                                  # SanJose CE Router

ipv6 unicast-routing
!
interface Loopback0
 ipv6 address 2001:DB8:CAFE:1:2::1/128
!
interface Port-channel30
 description to PE-VSS-core-BGLR
 ipv6 address 2001:DB8:CAFE:3::1/64      #Assign IPv6 address
!
interface TenGigabitEthernet1/7/1
  channel-group 30 mode on
!
interface TenGigabitEthernet1/7/1
  channel-group 30 mode on
    !
```

步骤2 配毕CE路由器接口之后，还需配置PE路由器接口，以完成上连P路由器，下连CE路由器的需求，如例7-4所示。

例7-4 PE路由器的接口配置

```
BGLR-CE:                                 # Bangalore CE Router

ipv6 unicast-routing

!
interface Loopback0
 ipv6 address 2001:DB8:CAFE:1:1::1/128
!
interface Port-channel30
 description to PE-VSS-core-BGLR
 ipv6 address 2001:DB8:CAFE:2::1/64      #Assign IPv6 address
  !
interface TenGigabitEthernet1/7/1
  channel-group 30 mode on
!
interface TenGigabitEthernet1/7/1
  channel-group 30 mode on
!
```

（待续）

```
SJC-CE:                                    # SanJose CE Router

ipv6 unicast-routing
!
interface Loopback0
 ipv6 address 2001:DB8:CAFE:1:2::1/128
!
interface Port-channel30
 description to PE-VSS-core-BGLR
 ipv6 address 2001:DB8:CAFE:3::1/64      #Assign IPv6 address
!
interface TenGigabitEthernet1/7/1
  channel-group 30 mode on
!
interface TenGigabitEthernet1/7/1
  channel-group 30 mode on
    !
```

步骤 3 在完成 PE 和 CE 路由器的基本配置之后，还需完成路由协议的配置。例 7-5 所示为 RIPng 的配置。

例 7-5 CE 路由器路由协议的配置

```
BGLR-PE:                          # Bangalore PE Router

ipv6 unicast-routing
!
int loopback0
 ip address 10.10.10.1 255.255.255.255
!
interface Portchannel10
 description interface to P network
 ip address 10.1.2.1 255.255.255.0
 mpls ip
 !
interface PortChannel30
 description interface to VSS-CE BGLR
 ipv6 address 2001:DB8:CAFE:2::2/64
 ipv6 enable
!
interface TenGigabitEthernet1/7/1
  channel-group 30 mode on
!
interface TenGigabitEthernet2/7/1
  channel-group 30 mode on
```

（待续）

```
!
SJC-PE:                          # SanJose PE Router

ipv6 unicast-routing
!
int loopback0
 ip address 10.10.10.4 255.255.255.255
!
interface Portchannel10
 description interface to P network
 ip address 10.1.5.1 255.255.255.0
 mpls ip
!
interface PortChannel30
 description interface to VSS-CE SJC
 no ip address
 ipv6 address 2001:DB8:CAFE:3::2/64
 ipv6 enable
 !
!Interface assignment to port-channels removed for brevity
BGLR-CE:                                  # Bangalore CE Router

ipv6 unicast-routing
!
interface Loopback0
  ipv6 rip customer enable
!
interface Port-channel30
  ipv6 rip customer  enable
ipv6 router rip customer
!
SJC-CE:                                   # SanJose CE Router

ipv6 unicast-routing
!
interface Loopback0
!
 ipv6 rip customer enable
!
interface Port-channel30
 ipv6 rip customer  enable
!
ipv6 router rip customer
    !
```

步骤 4 在 PE 路由器上，同样也要完成路由协议的配置。例 7-6 所示 PE 路由器的配置建立在已完成 P 路由器配置的基础之上。

例 7-6 PE 路由器的配置

```
BGLR-PE:                          # Bangalore PE Routing Configuration

interface PortChannel30
 ipv6 rip customer enable
!
router bgp 100
 neighbor 10.10.10.4 remote-as 100
 neighbor 10.10.10.4 update-source Loopback0
 no auto-summary
!
 address-family ipv6
  neighbor 10.10.10.4 activate
  neighbor 10.10.10.4 send-community both
  neighbor 10.10.10.4 send-label
  redistribute connected
  redistribute rip customer
 exit-address-family
!
ipv6 router rip customer
 redistribute bgp 100
!
SJC-PE:                           # SanJose PE Routing Configuration

interface PortChannel30
 ipv6 rip customer enable
!
router bgp 100
 neighbor 10.10.10.1 remote-as 100
 neighbor 10.10.10.1 update-source Loopback0
!
 address-family ipv6
  neighbor 10.10.10.1 activate
  neighbor 10.10.10.1 send-community both
  neighbor 10.10.10.1 send-label
  redistribute connected
  redistribute rip customer
  no synchronization
```

（待续）

```
 exit-address-family
!
ipv6 router rip customer
 redistribute bgp 100
    !
```

步骤 5 配置 IPv4 骨干网的 IGP，即 P 和 PE 路由器之间所运行的动态路由协议，本例采用的是 OSPF。例 7-7 所示为 BGLR_PE 和 P 路由器 OSPF 的配置。

例 7-7 BGLR-PE 和 P 路由器 OSPF 的配置

```
BGLR-PE# show running | beg router ospf
router ospf 100
 log-adjacency-changes
 network 10.1.2.0 0.0.0.255 area 0
 network 10.10.10.0 0.0.0.255 area 0
!
/* Bangalore P Router */
BGLR-P# show running ¦ begin ospf 100
router ospf 100
 log-adjacency-changes
 network 10.1.2.0 0.0.0.255 area 0
 network 10.1.4.0 0.0.0.255 area 0
 network 10.10.10.0 0.0.0.255 area 0
    !
```

例 7-8 所示 **debug** 命令的输出有助于验证 6PE 的配置。

例 7-8 利用 debug ip bgp 命令验证 BGP 路由更新的交换

```
Additionally you c BGLR-PE# debug ip bgp

! BPG negotiation, state moving from Idle to Active
*Jul 23 05:36:28.563: BGP: 10.10.10.4 went from Idle to Active
*Jul 23 05:36:28.563: BGP: 10.10.10.4 open active, local address 10.10.10.1
!Output omitted for brevity

! BGP sending OPEN ,capability exchange and other messages
*Jul 23 05:36:28.575: BGP: 10.10.10.4 OPEN has CAPABILITY code: 1, length 4
*Jul 23 05:36:28.575: BGP: 10.10.10.4 OPEN has MP_EXT CAP for afi/safi: 1/1
*Jul 23 05:36:28.575: BGP: 10.10.10.4 rcvd OPEN w/ optional parameter type 2
```

（待续）

```
(Capability) len 6
!Output omitted for brevity
*Jul 23 05:36:28.575: BGP: 10.10.10.4 OPEN has CAPABILITYcode: 128, length 0
*Jul 23 05:36:28.575: BGP: 10.10.10.4 OPEN has ROUTE-REFRESH capability(old) for
all address-families
*Jul 23 05:36:28.575: BGP: 10.10.10.4 rcvd OPEN w/ optional parameter type 2
(Capability) len 2
*Jul 23 05:36:28.575: BGP: 10.10.10.4 OPEN has CAPABILITY code: 2, length 0
*Jul 23 05:36:28.575: BGP: 10.10.10.4 OPEN has ROUTE-REFRESH capability(new) for
all address-families
BGP: 10.10.10.4 rcvd OPEN w/ remote AS 100

! BGP moved to Established
*Jul 23 05:36:28.575: BGP: 10.10.10.4 went from OpenSent to OpenConfirm
*Jul 23 05:36:28.579: BGP: 10.10.10.4 went from OpenConfirm to Established

*Jul 23 05:36:28.579: %BGP-5-ADJCHANGE: neighbor 10.10.10.4 Up
!Output omitted for brevity
```

此外，还可通过观察 **show ip bgp unicast neighbors** 命令的输出，来验证 MP-BGP 的会话建立情况。例 7-9 所示为这条命令的输出，以及与 IPv6 路由有关的信息。

例 7-9　show ip bgp unicast neighbors 命令的输出

```
BGLR-PE# debug ip bgp

! BPG negotiation, state moving from Idle to Active
*Jul 23 05:36:28.563: BGP: 10.10.10.4 went from Idle to Active
*Jul 23 05:36:28.563: BGP: 10.10.10.4 open active, local address 10.10.10.1
!Output omitted for brevity

! BGP sending OPEN ,capability exchange and other messages
*Jul 23 05:36:28.575: BGP: 10.10.10.4 OPEN has CAPABILITY code: 1, length 4
*Jul 23 05:36:28.575: BGP: 10.10.10.4 OPEN has MP_EXT CAP for afi/safi: 1/1
*Jul 23 05:36:28.575: BGP: 10.10.10.4 rcvd OPEN w/ optional parameter type 2
(Capability) len 6
!Output omitted for brevity
*Jul 23 05:36:28.575: BGP: 10.10.10.4 OPEN has CAPABILITYcode: 128, length 0
*Jul 23 05:36:28.575: BGP: 10.10.10.4 OPEN has ROUTE-REFRESH capability(old) for
all address-families
*Jul 23 05:36:28.575: BGP: 10.10.10.4 rcvd OPEN w/ optional parameter type 2
(Capability) len 2
*Jul 23 05:36:28.575: BGP: 10.10.10.4 OPEN has CAPABILITY code: 2, length 0
*Jul 23 05:36:28.575: BGP: 10.10.10.4 OPEN has ROUTE-REFRESH capability(new) for
```

（待续）

```
all address-families
BGP: 10.10.10.4 rcvd OPEN w/ remote AS 100

! BGP moved to Established
*Jul 23 05:36:28.575: BGP: 10.10.10.4 went from OpenSent to OpenConfirm
*Jul 23 05:36:28.579: BGP: 10.10.10.4 went from OpenConfirm to Established

*Jul 23 05:36:28.579: %BGP-5-ADJCHANGE: neighbor 10.10.10.4 Up
!Output omitted for brevity
BGLR-PE# show ip bgp ipv6 unicast neighbors          #BGLR-PE

BGP neighbor is 10.10.10.4,  remote AS 100, internal link
  BGP version 4, remote router ID 10.10.10.4
  BGP state = Established, up for 00:06:21
  Last read 00:00:15, last write 00:00:17, hold time is 180, keepalive interval
is 60 seconds
  Neighbor capabilities:
    Route refresh: advertised and received(new)
    Address family IPv4 Unicast: advertised and received
    Address family IPv6 Unicast: advertised and received
    ipv6 MPLS Label capability: advertised and received
!Output omitted for brevity
  For address family: IPv6 Unicast
  BGP table version 5, neighbor version 5/0
  Output queue size : 0
  Index 1, Offset 0, Mask 0x2
  1 update-group member
  Community attribute sent to this neighbor
  Sending Prefix & Label
                                 Sent        Rcvd
  Prefix activity:              ----        ----
    Prefixes Current:              1           1 (Consumes 76 bytes)
    Prefixes Total:                2           2
    Implicit Withdraw:             1           1
    Explicit Withdraw:             0           0
    Used as bestpath:            n/a           1
    Used as multipath:           n/a           0
!Output omitted for brevity
BGLR-PE# show ipv6 route   2001:DB8:CAFE:3::1/64  # Route to SJC

Routing entry for 2001:DB8:CAFE:3::/64
  Known via "bgp 100", distance 200, metric 0, type internal
  Redistributing via rip customer
```

（待续）

```
  Route count is 1/1, share count 0
  Routing paths:
    10.10.10.4%Default-IP-Routing-Table indirectly connected
      MPLS Required
      Last updated 00:07:51 ago
! show ipv6 cef , please note the imposition of two labels
BGLR-PE# show ipv6 cef 2000:CAFE:3::/64
2000:CAFE:3::/64
  nexthop 10.1.2.2 Ethernet0/0 label 18 17
```

6VPE

6VPE 可用在 MPLS 骨干网中传递 IPv6 VPN 流量。在 MPLS 网络中，IPv6 VPN 的运作方式与 IPv4 VPN 大致相同。对那些提供 IPv4 MPLS VPN 服务的服务提供商来说，只需升级 PE 路由器的 IOS 版本，便可以平稳过渡到提供 IPv6 VPN 服务，而无需对核心（P）路由器做任何升级。

图 7-4 所示为一个简化的 6VPE 网络，该网络只包括了 ABC 公司的两个站点 ABC-SanJose（SJC）和 Bangalore（BGLR）。本节将会以此为例来帮助读者理解 6VPE 网络[①]。

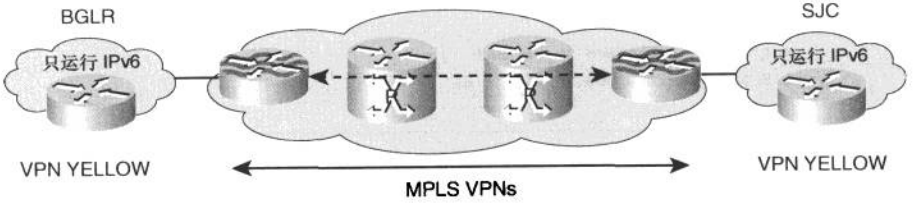


图 7-4 6VPE 网络示例

例 7-11 所示为图 7-4 中 PE 路由器上与 ABC 公司相关联的 VRF ABC-vrf 的配置。以下列出了 PE 和 CE 路由器所需的配置步骤。[②]

步骤 1 配置 CE 路由器，如例 7-10 所示。

---

① 原文是“A simplified network with just two sites of a company ABC—SanJose (SJC) and Bangalore (BGLR)—to explore the concept is used. Figure 7-3 shows a sample network to help you understand 6VPE”。作者太没有责任心了，译文为译者杜撰。

② 原文是“Figure 7-3 is a basic network that demonstrates 6VPE. In Example 7-10, customer ABC corresponds to VRF ABC-vrf on the PE routers. The following steps and configurations on the PE and CE routers are required”。原文完全语无伦次，作者在写作方面似乎找不到思路。译者对译文做了精简。

例 7-10 BGLR CE 路由器的配置

```
BGLR-CE:                          # Bangalore CE Router

!
ipv6 unicast-routing
!
interface Loopback0
 ip address 10.10.10.10 255.255.255.255
 ipv6 address 2001:DB8:CAFE:1:1::1/128
!
interface Portchannel30
 description interface to PE BGLR
 ipv6 address 2001:DB8:CAFE:2::1/64
!
/* Assignment of interfaces to channel-group same as 6 PE hence skipped */
SJC-CE:                           # SanJose CE Router

!
ipv6 unicast-routing
!
interface Loopback0
 ip address 10.10.10.11 255.255.255.25
 ipv6 address 2001:DB8:CAFE:1:2::1/128
!
interface Portchannel30
 description interface to PE BGLR
 ipv6 address 2000:DB8:CAFE:3::1/64
    !
```

步骤 2 配毕 CE 路由器的接口之后，在 PE 路由器上配置 6VPE 功能，并配置上连 P 路由器和下连 CE 路由的接口，以建立与 P 和 CE 路由器的连通性，如例 7-11 所示。

例 7-11 PE 路由器接口 IP 地址和 6VPE 的配置

```
BGLR-PE:                          # Bangalore PE Router 6VPE Config

ipv6 unicast-routing
!
vrf definition ABC-vrf
 rd 100:1
 !
 address-family ipv6
  route-target both 100:1
```

（待续）

```
 exit-address-family
!
mpls label protocol ldp
!
int loopback0
 ip address 10.10.10.1 255.255.255.255
!
interface Portchannel10
 description interface to P network
 ip address 10.1.2.1 255.255.255.0
 mpls ip
!
interface Portchannel30
 description interface to CE-BGLR
 vrf forwarding ABC-vrf
 ipv6 address 2001:DB8:CAFE:2::2/64
 ipv6 enable
!
SJC-PE:                          # SanJose PE Router 6VPE Config

ipv6 unicast-routing
!
vrf definition ABC-vrf
 rd 100:1
 !
 address-family ipv6
  route-target both 100:1
 exit-address-family
!
mpls label protocol ldp
!
int loopback0
 ip address 10.10.10.4 255.255.255.255
 !
interface Portchannel10
 description interface to P network
 ip address 10.1.5.1 255.255.255.0
 mpls ip

interface Portchannel30
 description interface to CE SJC
 vrf forwarding ABC-vrf
 ipv6 address 2001:DB8:CAFE:3::2/64
 ipv6 enable
    !
```

步骤3 配置CE路由器，以激活路由选择功能。如例7-12所示，CE路由器采用的路由协议是BGP。

例7-12 CE路由器路由协议的配置

```
BGLR-CE:                          # Bangalore CE Router

!/* Note: This bgp session will also negotiate exchange if IPv4 AF */
router bgp 65000
 no synchronization
 bgp log-neighbor-changes
 neighbor 2001:DB8:CAFE:2::2 remote-as 100
 no auto-summary
 !
 address-family ipv6
  neighbor 2001:DB8:CAFE:2::2 activate
  network 2001:DB8:CAFE:1:1::1/128
 exit-address-family
!
SJC-CE:                           # SanJose CE Router

router bgp 65000
 no synchronization
 bgp log-neighbor-changes
 neighbor 2001:DB8:CAFE:3::2 remote-as 100
 no auto-summary
 !
 address-family ipv6
  neighbor 2001:DB8:CAFE:3::2 activate
  network 2001:DB8:CAFE:1:2::1/128
  no synchronization
 exit-address-family
!
    !
```

步骤4 配置PE路由器，以激活路由选择功能。例7-13所示为两台PE路由器路由选择方面的配置。

例7-13 PE路由器路由协议的配置

```
BGLR-PE:                          #Bangalore PE Routing Configuration

router ospf 100
```

（待续）

```
 log-adjacency-changes
 network 10.1.0.0 0.0.255.255 area 0
 network 10.10.10.0 0.0.255.255 area 0
!
router bgp 100
 bgp log-neighbor-changes
 neighbor 10.10.10.4 remote-as 100
 neighbor 10.10.10.4 update-source Loopback0
 !
 address-family ipv4
  no neighbor 10.10.10.4 activate
  no auto-summary
  no synchronization
 exit-address-family
 !
 address-family vpnv6
  neighbor 10.10.10.4 activate
  neighbor 10.10.10.4 send-community both
 exit-address-family
 !
 address-family ipv6
  redistribute connected
  no synchronization
 exit-address-family
 !
 address-family ipv6 vrf ABC-vrf
  neighbor 2001:DB8:CAFE:2::1 remote-as 65000
  neighbor 2001:DB8:CAFE:2::1 activate
  override
  redistribute connected
  no synchronization
 exit-address-family
!
SJC-PE:                          #SanJose PE Routing Configuration
router ospf 100
 log-adjacency-changes
 network 10.1.0.0 0.0.255.255 area 0
 network 10.10.10.0 0.0.255.255 area 0
!
router bgp 100
```

（待续）

```
 bgp log-neighbor-changes
 neighbor 10.10.10.1 remote-as 100
 neighbor 10.10.10.1 update-source Loopback0
!
 address-family ipv4
  no auto-summary
  no synchronization
 exit-address-family
 !
 address-family vpnv6
  neighbor 10.10.10.1 activate
  neighbor 10.10.10.1 send-community both
 exit-address-family
 !
 address-family ipv6 vrf ABC-vrf
  neighbor 2001:DB8:CAFE:3::1 remote-as 65000
  neighbor 2001:DB8:CAFE:3::1 activate
  redistribute connected
  no synchronization
 exit-address-family
!
```

完成上述配置之后，还需验证6VPE网络能否正常运转。验证6VPE配置之前，请先在PE路由器上验证客户的VRF配置，如例7-14所示。

例7-14 验证客户VRF及其他基本参数的配置

```
BGLR-PE# show vrf ipv6
  Name                                     Default RD          Protocols   Interfaces
  ABC-vrf                                  100:1               ipv6        Po30
BGLR-PE# show vrf ipv6 interfaces
Interface              VRF                              Protocol    Address
Po30                   customer-vrf                       up
2001:DB8:CAFE:2::2
BGLR-PE# show ip bgp all neighbors
!Output omitted for brevity
For address family: VPNv4 Unicast
!Output omitted for brevity
For address family: VPNv6 Unicast

BGP neighbor is 10.10.10.4,  remote AS 100, internal link
  BGP version 4, remote router ID 10.10.10.4
  BGP state = Established, up for 00:03:35
  Last read 00:00:30, last write 00:00:44, hold time is 180, keepalive interval
```

（待续）

```
is 60 seconds
  Neighbor capabilities:
    Route refresh: advertised and received(new)
    Address family VPNv6 Unicast: advertised and received

!Output omitted for brevity
BGLR-PE# show bgp vpnv6 unicast vrf customer-vrf 2001:DB8:CAFE:1:2::1/128
BGP routing table entry for [100:1]2001:DB8:CAFE:1:2::1/128, version 7
Paths: (1 available, best #1, table customer-vrf)
  Advertised to update-groups:
        1
  65000
    ::FFFF:10.10.10.4 (metric 22) from 10.10.10.4 (10.10.10.4)
      Origin IGP, metric 0, localpref 100, valid, internal, best
      Extended Community: RT:100:1
      mpls labels in/out nolabel/19
! show ipv6 and the imposition of two labels
BGLR-PE#show ipv6 cef vrf customer-vrf 2000:CAFE:1:2::1/128
2000:CAFE:1:2::1/128
  nexthop 10.1.2.2 Portchannel10 label 18 19
BGLR-PE#
```

要想验证 CE 和 PE 路由器是否学到了正确的路由，可针对前者执行 **show ipv6 route** 命令，针对后者执行 **show ipv6 route vrf *vrf name*** 命令，如例 7-15 所示。

例 7-15 show ipv6 route vrf vrf name 命令的输出

```
BGLR-PE# show ipv6 route vrf customer-vrf
IPv6 Routing Table - customer-vrf - 6 entries
Codes: C - Connected, L - Local, S - Static, U - Per-user Static route
       B - BGP, M - MIPv6, R - RIP, I1 - ISIS L1
       I2 - ISIS L2, IA - ISIS interarea, IS - ISIS summary, D - EIGRP
       EX - EIGRP external
B   2001:DB8:CAFE:1:1::1/128 [20/0]
     via FE80::A8BB:CCFF:FE00:6500, Port-channel30
B   2001:DB8:CAFE:1:2::1/128 [200/0]
     via 10.10.10.4%Default-IP-Routing-Table, indirectly connected
C   2001:DB8:CAFE:2::/64 [0/0]
     via Port-channel10, directly connected
L   2001:DB8:CAFE:2::2/128 [0/0]
     via Port-channel10, receive
B   2001:DB8:CAFE:3::/64 [200/0]
```

（待续）

```
     via 10.10.10.4%Default-IP-Routing-Table, indirectly connected
L    FF00::/8 [0/0]
     via Null0, receive
BGLR-CE# show ipv6 route
IPv6 Routing Table - Default - 6 entries
Codes: C - Connected, L - Local, S - Static, U - Per-user Static route
       B - BGP, M - MIPv6, R - RIP, I1 - ISIS L1
       I2 - ISIS L2, IA - ISIS interarea, IS - ISIS summary, D - EIGRP
       EX - EIGRP external
LC   2001:DB8:CAFE:1:1::1/128 [0/0]
     via Loopback0, receive
B    2001:DB8:CAFE:1:2::1/128 [20/0]
     via FE80::A8BB:CCFF:FE00:6601, Port-channel30
C    2001:DB8:CAFE:2::/64 [0/0]
     via Ethernet0/0, directly connected
L    2001:DB8:CAFE:2::1/128 [0/0]
     via Ethernet0/0, receive
B    2001:DB8:CAFE:3::/64 [20/0]
     via FE80::A8BB:CCFF:FE00:6601, Port-channel30
L    FF00::/8 [0/0]
     via Null0, receive
```

要想得知6VPE路由信息是否被交换，请开启相应的**debug**命令，如例7-16所示。

例7-16 debug bgp all命令的输出

```
/* Negotiation of BGLR PE with the SJC PE router */

BGLR-PE# debug bgp all


! BPG negotiation, state moving from Idle to Active
*Jul 24 18:21:19.225: BGP: 10.10.10.4 went from Idle to Active
*Jul 24 18:21:19.229: BGP: 10.10.10.4 open active, local address 10.10.10.1
*Jul 24 18:21:19.229: BGP: 10.10.10.4 read request no-op
*Jul 24 18:21:19.229: BGP: 10.10.10.4 went from Active to OpenSent

! BGP sending OPEN ,version, Autonomous system number and other messages
*Jul 24 18:21:19.229: BGP: 10.10.10.4 sending OPEN, version 4, my as: 100, hold-
time 180 seconds
*Jul 24 18:21:19.237: BGP: 10.10.10.4 send message type 1, length (incl. header) 45

! BGP receiving OPEN with capabilities
```

（待续）

```
*Jul 24 18:21:19.241: BGP: 10.10.10.4 rcv message type 1, length (excl. header) 26
*Jul 24 18:21:19.241: BGP: 10.10.10.4 rcv OPEN, version 4, holdtime 180 seconds
*Jul 24 18:21:19.241: BGP: 10.10.10.4 rcv OPEN w/ OPTION parameter len: 16
*Jul 24 18:21:19.241: BGP: 10.10.10.4 rcvd OPEN w/ optional parameter type 2
(Capability) len 6
*Jul 24 18:21:19.241: BGP: 10.10.10.4 OPEN has CAPABILITY code: 1, length 4
*Jul 24 18:21:19.241: BGP: 10.10.10.4 OPEN has MP_EXT CAP for afi/safi: 2/128
*Jul 24 18:21:19.241: BGP: 10.10.10.4 rcvd OPEN w/ optional parameter type 2
(Capability) len 2
*Jul 24 18:21:19.241: BGP: 10.10.10.4 OPEN has CAPABILITY code: 128, length 0
*Jul 24 18:21:19.241: BGP: 10.10.10.4 OPEN has ROUTE-REFRESH capability(old) for
all address-families
*Jul 24 18:21:19.241: BGP: 10.10.10.4 rcvd OPEN w/ optional parameter type 2
(Capability) len 2
*Jul 24 18:21:19.241: BGP: 10.10.10.4 OPEN has CAPABILITY code: 2, length 0
*Jul 24 18:21:19.241: BGP: 10.10.10.4 OPEN has ROUTE-REFRESH capability(new) for
all address-families
BGP: 10.10.10.4 rcvd OPEN w/ remote AS 100
*Jul 24 18:21:19.241: BGP: 10.10.10.4 went from OpenSent to OpenConfirm

! BGP moved to Established
*Jul 24 18:21:19.241: BGP: 10.10.10.4 went from OpenConfirm to Established
*Jul 24 18:21:19.241: %BGP-5-ADJCHANGE: neighbor 10.10.10.4 Up
/* Negotiation with the BGLR CE router */

BGLR-PE#

! BPG negotiation, sending OPEN, version and other parameters
*Jul 24 18:22:00.321: BGP: 2001:DB8:CAFE:2::1 went from Active to OpenSent
*Jul 24 18:22:00.321: BGP: 2001:DB8:CAFE:2::1 sending OPEN, version 4, my as: 100,
holdtime 180 seconds
*Jul 24 18:22:00.329: BGP: 2001:DB8:CAFE:2::1 send message type 1, length (incl.
header) 45
! BPG receiving OPEN with other parameters
*Jul 24 18:22:00.337: BGP: 2001:DB8:CAFE:2::1 rcv message type 1, length (excl.
header) 34
*Jul 24 18:22:00.337: BGP: 2001:DB8:CAFE:2::1 rcv OPEN, version 4, holdtime 180
seconds
*Jul 24 18:22:00.337: BGP: 2001:DB8:CAFE:2::1 rcv OPEN w/ OPTION parameter len: 24
*Jul 24 18:22:00.337: BGP: 2001:DB8:CAFE:2::1 rcvd OPEN w/ optional parameter type
2 (Capability) len 6
*Jul 24 18:22:00.337: BGP: 2001:DB8:CAFE:2::1 OPEN has CAPABILITY code: 1, length 4
*Jul 24 18:22:00.337: BGP: 2001:DB8:CAFE:2::1 OPEN has MP_EXT CAP for afi/safi: 1/1
*Jul 24 18:22:00.337: BGP: 2001:DB8:CAFE:2::1 rcvd OPEN w/ optional parameter type
2 (Capability) len 6
```

（待续）

```
*Jul 24 18:22:00.337: BGP: 2001:DB8:CAFE:2::1 OPEN has CAPABILITY code: 1, length 4
*Jul 24 18:22:00.337: BGP: 2001:DB8:CAFE:2::1 OPEN has MP_EXT CAP for afi/safi: 2/1
*Jul 24 18:22:00.337: BGP: 2001:DB8:CAFE:2::1 rcvd OPEN w/ optional parameter type
2 (Capability) len 2
*Jul 24 18:22:00.337: BGP: 2001:DB8:CAFE:2::1 OPEN has CAPABILITY code: 128,
length 0
*Jul 24 18:22:00.337: BGP: 2001:DB8:CAFE:2::1 OPEN has ROUTE-REFRESH
capability(old) for all address-families
*Jul 24 18:22:00.337: BGP: 2001:DB8:CAFE:2::1 rcvd OPEN w/ optional parameter type
2 (Capability) len 2
*Jul 24 18:22:00.337: BGP: 2001:DB8:CAFE:2::1 OPEN has CAPABILITY code: 2, length 0
*Jul 24 18:22:00.337: BGP: 2001:DB8:CAFE:2::1 OPEN has ROUTE-REFRESH
capability(new) for all address-families
BGP: 2001:DB8:CAFE:2::1 rcvd OPEN w/ remote AS 65000
! BGP going to Established state with the peer and then declaring Adjacency to UP
*Jul 24 18:22:00.337: BGP: 2001:DB8:CAFE:2::1 went from OpenSent to OpenConfirm
*Jul 24 18:22:00.341: BGP: 2001:DB8:CAFE:2::1 went from OpenConfirm to Established
*Jul 24 18:22:00.341: %BGP-5-ADJCHANGE: neighbor 2001:DB8:CAFE:2::1 vpn vrf cus-
tomer Up
```

虚拟专用LAN服务

虚拟专用LAN服务（VPLS）提供了基于MPLS网络的点到多点通信机制。VPLS通过伪线（pseudo-wires）来互连分居各地的每个站点，并使得这些站点能够共享同一以太广播域。能够实现这种互连方案的技术包括MPLS上的以太网（EoMPLS）、L2TPv3以及GRE隧道。有两份记录VPLS的IETF草案，分别是RFC 4761“Virtual Private LAN Services（VPLS）Using BGP for Auto-Discovery and Signaling”和RFC 4762“Virtual Private LAN Services（VPLS）Using Label Distribution Protocol（LDP）Signaling”。

作为一项第二层的技术，VPLS在互连IPv6网络时，其运作方式与连接IPv4网域并无差别。本节只会试举一个利用VPLS建立站点间第二层连通性的简单示例，并不会探究VPLS的诸多细节。

图7-5所示为VPLS网络拓扑，其中，某客户有三个站点通过MPLS核心网络彼此互连。图中的三台PE路由器通过伪线形成全互连。

例7-17所示为上图中一台PE路由器（PE-6500-1）的配置示例。至于其他的PE路由器，除了IP地址以外，配置语法全都相同。

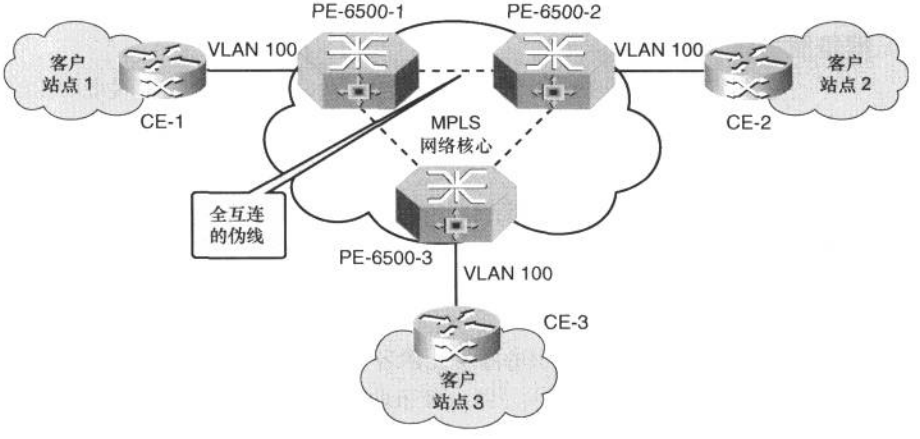


图 7-5 VPLS 网络拓扑示例

例 7-17 VFI 配置

```
Router#
!
 l2 vfi customer manual
! Create a vpn id and configure the vpn neighbors
vpn id 100
 neighbor 20.20.20.20 encapsulation mpls
 neighbor 30.30.30.30 encapsulation mpls
!
interface loopback 0
ip address 10.10.10.10 255.255.255.255
!
interface fastethernet1/0
 switchport
 switchport mode dot1qtunnel
 switchport access vlan 100
!
interface vlan 100
 no ip address
 xconnect vfi customer
```

VPLS 提供的是第二层的服务，由 6 个组件构成：附接电路、包交换网络（PSN）、伪线、自动发现、自动配置和虚拟交换实例（VSI）。对于 IPv6 来说，VPLS 所提供的服务是透明的。在互连分居各地的数据中心网络的场景中，VPLS 可在数据中心网络之间提供第二层路径，详情请见本书的第 9 章。

7.2.3 网络服务虚拟化

在本节，网络服务是指防火墙、负载均衡、监控以及其他网络所需的4～7层服务。写作本书之际，在某些运行上述服务的网络设备上，对IPv6的支持还受到一定的限制。现场实施时，我们建议读者核对以下设备的最新文档。

- Cisco防火墙（ASA家族，包括ASA 5500和防火墙服务模块[FWSM]）。
- Cisco ACE。
- Cisco入侵检测服务模块。

以下各节将关注网络服务虚拟化的各个方面，并以在多上下文中支持IPv6的Cisco ASA 5500解决方案为例，来帮助读者理解什么是网络服务虚拟化。

防火墙虚拟化

防火墙虚拟化是指把安全策略从物理硬件移植到抽象层来实现，旨在提高网络资源利用率和灵活性。可利用防火墙虚拟化技术，让多台设有不同安全策略的虚拟防火墙（上下文）在同一物理防火墙硬件模块上运行，并“各行其是”。每台虚拟防火墙都有属于自己的一组接口，安全策略可在这些接口上应用。站在安全上下文（虚拟防火墙）的角度来看，所看到的是一致而又标准的一套硬件设备，而与物理部件无关。每个安全上下文都是一台独立的防火墙，都有属于自己的配置、安全策略和管理者。

随着安全策略被分散配置在各逻辑虚拟防火墙上，安全虚拟上下文易于配置和便于管理的重要优势才能够得以显现。而单台大型防火墙，接口众多，安全策略配置复杂，将其拆分为多台小型防火墙后，既易于管理，策略也更为简单，从而避免了因配置错误所招致的系统隐患和安全漏洞。

Cisco ASA虚拟化架构

如今的Cisco防火墙都支持虚拟防火墙技术，由于一个虚拟防火墙可算作全功能防火墙的一个实例或一个分区，因此也将其称为上下文。图7-6所示为由单个防火墙CPU所虚拟出的多个不同的上下文。

就传统意义而言，防火墙都是以路由模式运行，可将其算做第三层设备[①]。运行于路由模式时，ASA作为受保护主机的默认网关。此外，Cisco ASA还支

① 原文是“Traditionally, firewalls provide a Layer 3 hop or the routed mode”。译者好想对作者说，咱实意动词不用“provide”或“enable”了，成不？

持透明模式，其运作方式类似于“bump in the wire（线路插件）”。有了防火墙虚拟化，便可将 ASA 划分为多个安全上下文，其详细内容将会在下一节进行讨论。图 7-7 所示为如何将单台 Cisco ASA 划分为多个安全上下文——既有运行路由模式的上下文又有运行透明模式的上下文。

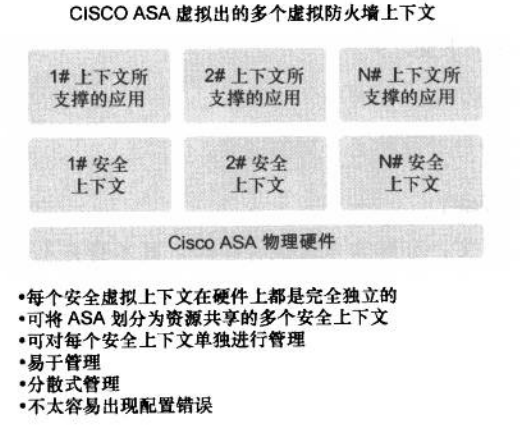


图 7-6 利用 Cisco ASA 来实现虚拟化

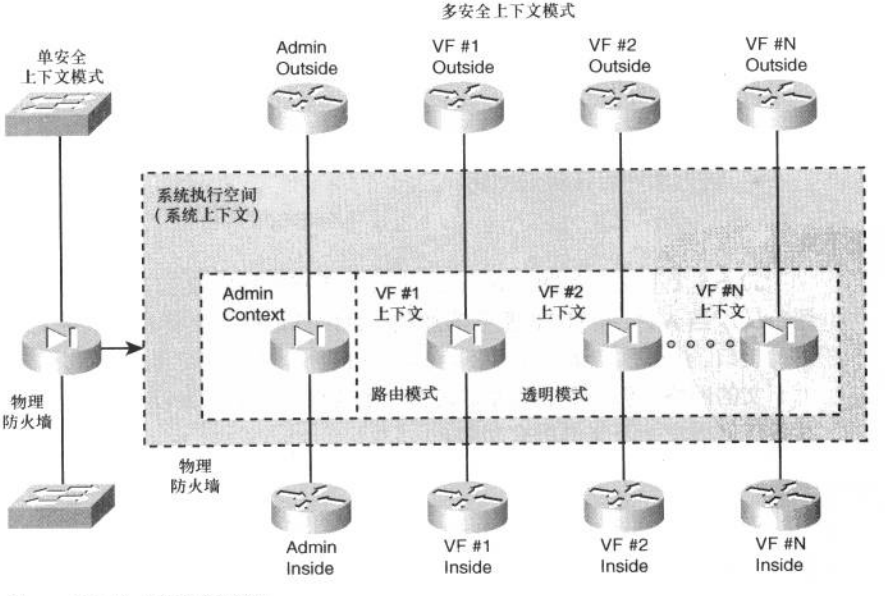


图 7-7 Cisco ASA 多上下文逻辑图

理解 Cisco ASA 中的虚拟上下文

以下各节将会介绍虚拟上下文，为每个虚拟上下文合理分配资源的指导方针。在 Cisco ASA 上，网管人员可对以下几种类型的上下文进行配置。

- 系统上下文。
- 管理上下文。
- 用户上下文。

以下各节是对每种上下文的定义。

系统上下文（系统执行空间）

系统上下文，亦称为系统执行空间，其本身并不是一个真正的上下文，但可用来定义和配置其他的用户上下文。网管人员可利用该上下文去创建上下文，并为所创建的每个上下文分配接口和系统资源。在系统执行空间中，网管人员可配置以下各种特性。各种类型的上下文以及系统上下文可见图 7-6。

- 防火墙启动配置文件。
- 防火墙模式（路由模式或透明模式）。
- 上下文定义（配置文件、接口分配或映射）。
- 保存 crash 信息。
- 防火墙系统时钟。
- 防火墙许可证激活密钥。

admin 上下文

admin 上下文（如图 7-6 所示），具备所有常规虚拟上下文的特性，但此外还定义了当 ASA 从 TFTP 服务器加载软件映像或配置文件时，系统上下文所使用的接口。由于系统上下文并不直接与任何物理接口打交道，故而会借用 admin 上下文的网络资源。除了配置方式类似于其他任何上下文以外，admin 上下文还具备任何用户上下文的所有功能。

用户上下文

每个用户上下文（如图 7-6 所示），都是一台独立的虚拟防火墙，并设有为自己所独有且独立于其他用户上下文的一组安全策略。可将每个用户上下文配置为“路由”或“透明”模式，无论哪种模式，都可拥有属于自己的一套安全

策略、网络地址以及访问控制方法等。所有可配置在单防火墙平台上的安全策略，都可以配置在用户上下文中。存在于这种安全上下文（亦称为虚拟上下文）中的每台虚拟防火墙都包括：

- 一组逻辑接口；
- 设在每个接口上的安全参数；
- 作用于虚拟防火墙的全局数据、状态信和统计数据；一个唯一的虚拟防火墙标识符，可供网关人员在访问上下文时引用；
- 可供配置并被强行分配的整个系统硬件资源的一部分。

在 Cisco ASA 上配置多上下文

本节将描述如何在 Cisco ASA 虚拟防火墙上配置各种上下文和基本的 IPv6 功能。图 7-8 所示为包含了一个 admin 上下文（名为 Admin）和两个用户上下文（名为 Red 和 Blue）的 ASA 防火墙网络。

如图 7-9 所示，admin 上下文仍使用 IPv4 地址，而其他的上下文均使用 IPv6 地址。紧随其后的配置示例则说明了如何在多上下文模式（虚拟防火墙）中转换和配置 Cisco ASA 设备。

要想在 Cisco ASA 上激活多上下文模式，请执行例 7-18 中的命令。只有在激活了多上下文模式之后，才能在 Cisco ASA 上创建各种上下文。要想在 Cisco ASA 上禁用多上下文模式，可执行该命令的 **no** 形式。更改过上下文模式之后，Cisco ASA 设备需要重启。

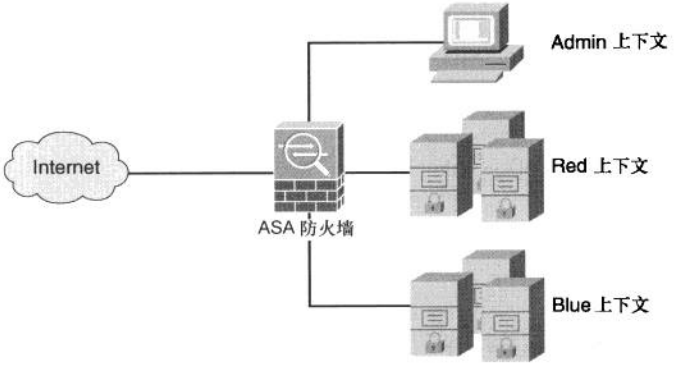


图 7-8 ASA 防火墙网络

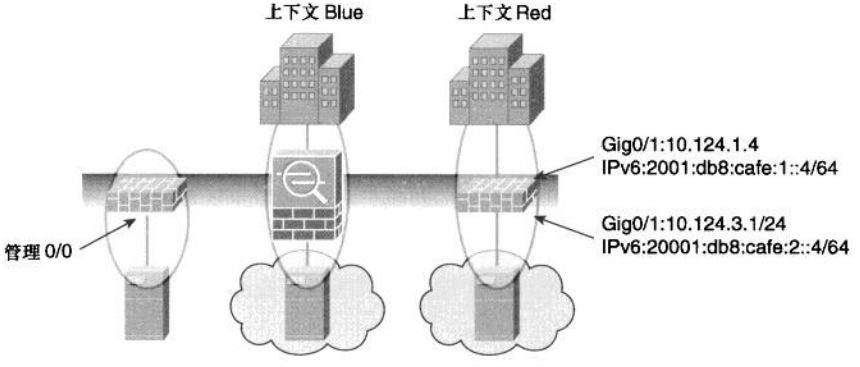


图 7-9 Cisco ASA 上下文示例

例 7-18 激活多上下文模式

```
ciscoasa(config)# mode multiple
WARNING: This command will change the behavior of the device
WARNING: This command will initiate a Reboot
Proceed with change mode? [confirm]

Convert the system configuration? [confirm]

!
The old running configuration file will be written to flash

The admin context configuration will be written to flash

The new running configuration file was written to flash
Security context mode: multiple
***
*** --- SHUTDOWN NOW ---
***
*** Message to all terminals:
***
***   change mode
Process shutdown finished
Rebooting.....
Restarting system.
!Output omitted for brevity
```

（待续）

```
Creating context 'admin'... Done. (1)
*** Output from config line 34, "admin-context admin"

Cryptochecksum (changed): 85ac17c5 76153dab e815c777 22b18a8d

*** Output from config line 36, "  config-url flash:/admi..."

Cryptochecksum (changed): 487a8b1e 3574c5d7 dc07697e 6349934f
Type help or '?' for a list of available commands.
ciscoasa>
```

在 Cisco ASA 上，将其配置为在多上下文模式运行之后，可继续配置用户上下文并为其分配地址。在进入上下文展开配置之前，网管人员应配置接口，并根据服务等级协定（SLA）为上下文分配资源；这包括接口的分配、URL 的配置、入侵防御系统（IPS）sensor 的分配以及故障切换组的加入。例 7-19 所示为具体上下文的配置。

例 7-19 配置多上下文防火墙

```
ciscoasa# conf t
ciscoasa(config)# context Red
Creating context 'Red'... Done. (2)
ciscoasa(config-ctx)# config-url disk0:/Red.cfg

WARNING: Could not fetch the URL disk0:/Red.cfg
INFO: Creating context with default config
ciscoasa(config-ctx)# allocate-interface GigabitEthernet0/0
ciscoasa(config-ctx)# allocate-interface GigabitEthernet0/1
ciscoasa(config-ctx)# context Blue

Creating context 'Blue'... Done. (3)
ciscoasa(config-ctx)# config-url disk0:/Blue.cfg
WARNING: Could not fetch the URL disk0:/Blue.cfg
INFO: Creating context with default config
ciscoasa(config-ctx)# context admin
ciscoasa(config-ctx)# allocate-interface GigabitEthernet0/3

/* Make sure that the contexts are created and interfaces assigned */
ciscoasa# sh context
Context Name      Class      Interfaces           URL
*admin             default    Management0/0        disk0:/admin.cfg
```

（待续）

```
 Red                  default     GigabitEthernet0/0,   disk0:/Red.cfg
                                   GigabitEthernet0/1
 Blue                 default     GigabitEthernet0/3    disk0:/Blue.cfg
Ciscoasa(config-ctx)# int GigabitEthernet0/0
ciscoasa(config-if)# no shutdown
ciscoasa(config-if)# int GigabitEthernet0/1
ciscoasa(config-if)# no shutdown
ciscoasa(config-if)# int GigabitEthernet0/2
ciscoasa(config-if)# no shutdown
ciscoasa(config-if)# int GigabitEthernet0/3
ciscoasa(config-if)# no shutdown
```

配置具体上下文的下一步是配置防火墙接口和各种特性（包括故障切换、路由选择、访问列表及其他ASA特性等）。例7-20所示为接口、静态路由以及故障切换的配置。

例7-20 配置防火墙接口和故障切换选项

```
/* Red Context Config */

ASA Version 8.2(2) <context>
!
hostname Red
enable password 8Ry2YjIyt7RRXU24 encrypted
passwd 2KFQnbNIdI.2KYOU encrypted
names
!
interface GigabitEthernet0/0
 nameif Outside
 security-level 0
 ip address 202.1.1.1 255.255.255.0 standby 202.1.1.2
 ipv6 address 2001:db8:cafe:1::1/64 standby 2001:db8:cafe:1::2
!
interface GigabitEthernet0/1
 nameif Inside
 security-level 100
 ip address 10.1.1.1 255.255.255.0 standby 10.1.1.2
 ipv6 address 2001:db8:cafe:2::1/64 standby 2001:db8:cafe:2::2
!
pager lines 24
mtu Outside 1500
mtu Inside 1500
```

（待续）

```
ipv6 route Outside ::/0 2001:db8:cafe:1::10
!Output omitted for brevity
System Context Failover Config:

failover
failover lan unit primary
failover lan interface FO-LINK GigabitEthernet0/2
failover link FO-LINK GigabitEthernet0/2
failover interface ip FO-LINK 2001:db8:cafe:5::1/64 standby 2001:db8:cafe:5::2
!
```

配置 IPv6 访问列表

在 Cisco ASA 上配置访问列表时，IPv4 和 IPv6 差别不大，不外是先创建访问控制条目（ACE），然后在接口上应用所创建的 ACE。以下列出了适用于 IPv6 访问列表的指导方针和局限性。

- 与用于 IPv4 的 **access-list** 命令相比，除了多出关键字“**ipv6**”以外，用于 IPv6 的 **access-list** 命令几乎相同。
- 默认情况下，对流量的放行方式随产品而异。比方说，对于 ASA，流量总能从安全级别高的区域流向安全级别低的区域，无需配置访问列表；而对于 FWSM（防火墙服务模块），除非设置了相应的访问列表去放行特定的流量，否则流量不能在任何安全区域之间流动[①]。
- 在 **access-list/ipv6 access-list** 命令中，可调用 **object-group** 命令所配置的对象组。

例 7-21 所示为先创建名为 ACL_IN 的访问列表，让 IPv6 地址为 2001:db8:cafe:2::/64 和 2001:db8:cafe:3::/64 的主机能够访问任意主机；然后，再将其应用于 Inside 接口。

例 7-21 配置 IPv6 访问列表/show access-list 命令

```
ciscoasa/Red# conf t
ciscoasa/Red(config)#ciscoasa/Red(config)# ipv6 access-list ACL_IN permit ip
2001:db8:cafe:2::/64 any
ciscoasa/Red(config)# ipv6 access-list ACL_IN permit ip 2001:db8:cafe:3::/64 any
```

（待续）

① 原文是“The traffic that is allowed by default differs from product to product. For example, for the access lists on the ASA, all packets from the outside to the interface are denied until specific access lists are configured. On FWSM (Firewall Services Modules), all traffic is denied from any to any interface until access lists are applied”。译文和原文明显不同，请读者注意。

```
ciscoasa/Red(config)# access-group ACL_IN in interface inside
ciscoasa/Red# show access-list ACL_IN
ipv6 access-list ACL_IN; 2 elements; name hash: 0xc0a1e5cd
ipv6 access-list ACL_IN line 1 permit ip 2001:db8:cafe:2::/64 any (hitcnt=0)
0xb72c7224
ipv6 access-list ACL_IN line 2 permit ip 2001:db8:cafe:3::/64 any (hitcnt=0)
0x83b1bec0
# Access list configuration with object-groups
ciscoasa/Red(config)# object-group network out_net
ciscoasa/Red(config-network)# network-object 2001:db8:cafe:10::/64
ciscoasa/Red(config-network)# exit
ciscoasa/Red(config)# object-group network in_net
ciscoasa/Red(config-network)# network-object 2001:db8:cafe:2::/64
ciscoasa/Red(config-network)# exit
ciscoasa/Red(config)# ipv6 access-list ACL_OUT permit ip object-group out_net
object-group in_net
ciscoasa/Red(config)# exit
ciscoasa/Red# show access-list ACL_OUT
ipv6 access-list ACL_OUT; 1 elements; name hash: 0x21ec8810
ipv6 access-list ACL_OUT line 1 permit ip object-group out_net object-group in_net
0x912c1636
ipv6 access-list ACL_OUT line 1 permit ip 2001:db8:cafe:10::/64
2001:db8:cafe:2::/64 (hitcnt=0) 0x984c7723
ciscoasa/Red#
```

警告 更为细致的Cisco ASA防火墙配置步骤，请参考以下配置指南：http://www.cisco.com/en/US/docs/security/asa/asa82/configuration/guide/config.html。

7.3 桌面虚拟化

随着对能够降低桌面管理成本的桌面虚拟化技术的运用，计算应用和生产力应用不断从本地化向以数据中心为中心的集中化演进。采用了桌面虚拟化技术，便能将末端主机桌面与客户机的底层硬件及操作系统相分离。取决于客户端桌面所运行的实际环境，可将虚拟桌面划分为两类。

- **基于服务器的计算模型**：使用该模型，服务器主机操作系统和所安装的应用程序要么运行在虚拟空间内（例如，Microsoft 终端服务），要么运行基于虚拟机管理程序的专用虚拟机之上（例如，集成了 VMware View 的 VMware vSphere）。远程用户可使用远程显示协议（remote display

protocol)，比如，微软远程桌面协议(RDP)和 Citrix 独立计算架构(ICI)，去访问服务器。

- **基于客户端的计算模型**：使用该模型，客户可以在本地的虚拟机上运行访客操作系统及应用程序，或者直接在本地运行应用程序，更有甚者，应用程序还能收听（看）来自数据中心的串流（stream）。VMware Fusion（用于 Apple Mac OS X 的客户端虚拟化管理工具）和 Microsoft App-V（Application virtualization/streaming[应用程序虚拟化/串流]）是两个基于客户端的计算模型的例子。

以下各节将讨论 IPv6 与桌面虚拟化设计之间的联系，以及在瘦客户端设备上启用 IPv6 时厂商所应恪守的设计原则。

7.3.1 IPv6 和桌面虚拟化

作为一项底层连通性技术，IPv6 通过支持瘦客户端的自动配置及提供超大型地址空间，来支持桌面虚拟化。随着瘦客户端的部署，企业需要两倍于前的 IP 地址数：一半用于瘦客户端，一半用于实际的计算系统（服务器上运行的虚拟机）。

市面上支持 IPv6 的瘦客户端软硬件和显示协议寥寥无几，众多厂商都在持观望态度。而对于那些在自己的产品中支持 IPv6 的厂商，通常需要在桌面虚拟化的三个领域，去考虑 IPv6 的部署。

- **无/瘦/胖客户端（Zero/thin/thick-client）配置**：瘦客户端既可以通过状态化地址自动配置（SLAAC）特性或 DHCPv6 来获取 IPv6 地址，也可以静态配置 IPv6 地址。DHCP 服务器需要提供一条通往 firmware TFTP 服务器的路径，否则便要利用 DHCP 选项引导 TFTP 服务器[①]。
- **连接代理/门户/网关（broker/portal/gateway）**：虚拟桌面接口环境（VDI）使用的是连接代理，而基于服务器的计算（SBC）环境则通常会利用门户或代理网关，去执行认证、授权并为客户提供桌面/会话。若将连接代理配置为“隧道”或“代理”模式，其中连接代理所起的是后台桌面的作用（虚拟机或 SBC 会话），那么连接代理以及支撑其运行的显

① 原文是“The DHCP server needs to provide a path to the firmware TFTP or boot the server using DHCP options”。作者这种语言，实在不是译者这种凡人所能理解，译文只能勉强翻译。

示协议都必须支持IPv6。

- **虚拟机/物理计算机/刀片PC：**在诸如VMware View和Citrix XenDesktop之类的托管虚拟桌面（HVD）环境中，会使用虚拟机作为客户的桌面。虚拟机必须支持IPv6，支持的方式包括虚拟机操作系统的TCP/IP协议栈（比如，Windows 7虚拟机）或虚拟机上运行的代理程序。远程客户会利用各种显示协议（包括RDP、ICA以及Teradici PCoIP等），连接虚拟机。为了能够让客户通过IPv6成功连接到虚拟机，显示协议必须支持IPv6，但目前支持IPv6的显示协议不多。此外，还可以利用显示协议去直接连接运行Oracle Solaris、Linux、Microsoft Windows以及Apple Mac OS的物理计算机或刀片PC。如今，Microsoft RDP和虚拟网络计算（VNC）都可以在IPv6上操作，并可以用来连接绝大多数远程操作系统。

下一节将介绍采用Oracle Sun Ray实施桌面虚拟化的场景。

7.3.2 桌面虚拟化示例：Oracle Sun Ray

本节将描述如何采用Oracle Sun Ray瘦客户端实施桌面虚拟化。图7-10所示为用于本场景的网络拓扑。

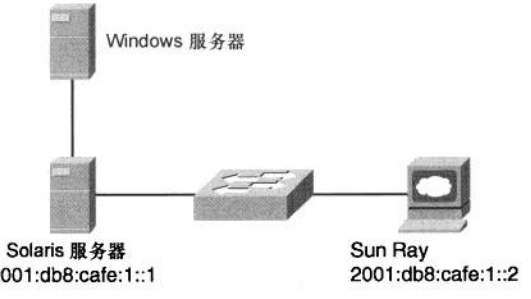


图7-10 Sun Ray瘦客户端解决方案网络拓扑

首先，要配置瘦客户端（对于本场景，即为Sun Ray Integrated Display[集成显示]），然后再执行相关IPv6的配置。Sun Ray客户端的详细配置步骤请参考Oracle Wikis主页，链接为：

http://wikis.sun.com/display/SRSS4dot2/Sun+Ray+Client+Boot+Process。

对于GUI的配置，请参考Oracle主页上的文档“How to Set DTU Configuration

Parameters”，链接为：

http://wikis.sun.com/display/SRSS4dot2/How+to+Set+DTU+Configuration+Parameters+%28Pop-up+GUI%29。

在图 7-10 所示的拓扑中，为 Sun Ray 配置的 IPv6 地址为 2001:db8:cafe:1::1[①]。配置 Solaris 服务器通过 IPv6 与客户端链接之后，服务器会下载自己的 firmware（固件）以及虚拟机的映射信息。服务器和客户机通过 IPv6 建立起连通性之后，网管人员可用 Sun Ray Administration GUI 去验证 IPv6 地址、firmware 版本以及连接情况。例 7-22 所示为 Sun Ray 连接会话的日志输出。

例 7-22 Sun Ray 连接的日志输出

```
Jul 18 17:15:56 GSBU-TME-SOLARIS utauthd: [ID 715959 user.info]
Worker0 NOTICE: CLAIMED by StartSession.m5 NAME: pseudo.00144fad725b
PARAMETERS: {stealProtected=true,
terminalIPA=2001:db8:cafe:1:0:0:0:2, type=pseudo,

 fw=GUI4.2_77_2009.10.19.17.01,Boot:2.0; 2007.08.17-17:32:09-PDT,
state=disconnected, cause=insert, doamgh=true, barrierLevel=420,
lockaction=disconnect, rawId=00144fad725b,
terminalCID=IEEE802.00144fad725b, MTU=1500, tokenSeq=1,
firstServer=20010db8cafe0001-01, namespace=IEEE802,

keyTypes=dsa-sha1-x1,dsa-sha1, ddcconfig=1,
clientRand=ZfbRy4BLiO8ZAE2SEu5gXV6okILcQvtx/pJxHPKREUi,
id=00144fad725b, realIP=20010db8cafe0001-02,
startRes=1280x1024:1280x1024, useReal=true, event=insert,
sn=00144fad725b, rawType=pseudo, hw=SunRayP8-270, initState=1,
usersession=false, _=1}
```

7.4 服务器虚拟化

服务器虚拟化是一种将裸机中安装的服务器操作系统移植到虚拟机的主流方法。VMware vSphere、Citrix XenServer 以及 Microsoft Hyper-V 都是市面上领先的服务器虚拟化基础设施产品。随着服务器虚拟化与 IPv6 的关系越来越紧密，服务器虚拟化平台所包括的两个重要领域也需要支持 IPv6，这两个重要领域是：

① 原文是“In this topology, Sun Ray is configured with an IPv6 address of 2001:db8:cafe:1::1”。可在图 7-10 中，这个地址是分配给 Solaris 服务器的。对于 Sun Ray 解决方案，译者是外行，因此只能按字面意思直译。

虚拟化管理程序（hypervisor）OS 以及管理工具。

写作本书之际，对 IPv6 支持最好的服务器虚拟化平台是带 ESX/ESXi hypervisor 的 VMware vSphere 4.1。大多数 VMware vSphere 4.1 组件，比如，hypervisor OS（ESX/ESXi）和 vCenter 都支持 IPv6。本书第 12 章会举一个如何在上述组件上启用 IPv6 的实例。

7.5 总结

本章涵盖了如何在虚拟化网络环境下设计和实施 IPv6。阅读本章之后，读者应该能够学会如何利用 6PE 和 6VPE 技术，跨 MPLS 骨干网，互连 IPv6 网络。读者还能通过本章了解到什么是网络服务虚拟化，以及部署网络服务虚拟化的优点，为此，本章展示了如何在 Cisco ASA 5580 上配置多上下文 IPv6 防火墙，以及验证配置的方法。此外，本章还简要介绍了桌面和服务器 IPv6 虚拟化技术。

7.6 参考资料

Cisco. Deploying Secure Multi-Tenancy into Virtualized Data Centers: http://www.cisco.com/en/US/partner/docs/solutions/Enterprise/Data_Center/Virtualization/securecldeployg.html.

Cisco. Implementing IPv6 over MPLS: http://www.ciscosystems.ch/en/US/docs/ios/ipv6/configuration/guide/ip6-over_mpls_ps6922_TSD_Products_Configuration_Guide_Chapter.html.

Cisco. Network Considerations to Optimize Virtual Desktop Deployment: http://www.cisco.com/en/US/prod/collateral/switches/ps5718/ps4324/white_paper_c11-531553_ns725_Networking_Solutions_White_Paper.html.

Durand, A., P. Fasano, and D. Lento. RFC 3053, "IPv6 Tunnel Broker." http://www.rfc-editor.org/rfc/rfc3053.txt.

Carpenter, B. and K. Moore. RFC 3056, "Connection of IPv6 Domains via IPv4 Clouds." http://www.rfc-editor.org/rfc/rfc3056.txt.

Vixie, P. and D. Wessels. RFC 2756, "Hyper Text Caching Protocol (HTCP/0.0)." http://www.rfc-editor.org/rfc/rfc2756.txt.

Gillgan, R. and E. Nordmark. RFC 2893, "Transition Mechanisms for IPv6 Hosts and

Routers." http://www.rfc-editor.org/rfc/rfc2893.txt.

Handley, M., D. Thaler, and R. Kermode. RFC 2776, "Multicast-Scope Zone Announcement Protocol (MZAP)." http://www.rfc-editor.org/rfc/rfc2776.txt.

Nordmack, E. and R. Gilligan. RFC 4213, "Basic Transition Mechanisms for IPv6 Hosts and Routers." http://www.rfc-editor.org/rfc/rfc4213.txt.

De Clercq, J., D. Ooms, M. Carugi, and F. Le Faucheur. RFC 4659, "BGP-MPLS IP Virtual Private Network (VPN) Extension for IPv6 VPN." http://www.rfc-editor.org/rfc/rfc4659.txt.

Templin, F., T. Gleeson, and D. Thaler. RFC 5214, "Intra-Site Automatic Tunnel Addressing Protocol (ISATAP)." http://www.rfc-editor.org/rfc/rfc5214.txt.

Durand, A., P. Fasano, I. Guardini, and D. Lento. RFC 3053, "IPv6 Tunnel Broker." http://www.rfc-editor.org/rfc/rfc3053.txt.

CHAPTER 8

第8章　在WAN/分支机构网络中部署IPv6

本章涵盖以下主题。

- **WAN/分支机构网络部署概述**：本节将会介绍分支机构网络的单层、双层和多层部署模型（模板）。
- **WAN/分支机构网络 IPv6 部署通则**：本节将会呈现适用于所有 WAN/分支机构网络的 IPv6 部署通则。
- **WAN/分支机构中 IPv6 部署示例**：结合 WAN/分支机构网络的单层、双层和多层部署模型的特征要素，举出与 IPv6 有关的网络设备配置示例。
- **基于纯 IPv6 部署 WAN/分支机构网络**：详尽展示了站点之间使用纯 IPv6 互连时，WAN/分支机构网络 IPv6 相关设备的配置示例，而在先前几节的配置示例中，站点之间都是使用 IPv4 上的 IPv6（IPv6-over-IPv4）互连。

本章将重点介绍如何利用 IPv6 外加各种不同的方法，来互连分支机构和总部/区域站点的网络。如本书其他各章所述，有时，为了让 IPv6 数据报得以在 WAN 上传输，需要将其封装进 IPv4 报头。对于此类情况的发生，原因是大都是 WAN 提供商不具备提供纯 IPv6 服务的能力，而并非企业用来联网的设备缺乏相关的 IPv6 特性或功能。对于 WAN/分支机构网络环境，有若干种部署选项可用来为其用户提供 IPv6 连通性，从而能够让用户访问到位于主站点及主站点一侧的 IPv6 应用程序或服务[①]。

本章还会讨论在 WAN/分支机构网络环境中部署纯 IPv6 的方法。只要 WAN

① 原文是“There are a wide variety of deployment options in WAN/branch scenarios that can provide you with a way to provide IPv6 connectivity to branch users and access applications and services located at the main site and beyond”。

服务提供商支持端到端的IPv6访问，便无需依赖IPv4 IPSec或SSL对IPv6数据包的封装，来保障数据传输的安全性。而如今，都是通过在运行Cisco IOS的分支机构路由器和总部WAN前置路由器之间，利用IPv4 IPSec来传输IPv6数据报，以保证数据传输的安全性[①]。

8.1 WAN/分支机构网络部署概述

以下各节将会简要介绍三种最为常见的Cisco分支机构网络架构，以及与其相关联的总部WAN前置区域部署模型[②]。本节内容将会说明如何在以下几种分支机构网络模型中集成IPv6的基本知识。

- 单层部署模型。
- 双层部署模型。
- 三层部署模型。

8.1.1 单层部署模型

单层分支机构网络模型是一种集多种功能于一台设备的设计方案，其WAN连通性基于Cisco动态多点虚拟专用网络（DMVPN）解决方案。在这样的网络环境中，会部署一台Cisco集成业务路由器(ISR)来满足LAN和WAN连通性和安全性的需求。更多与Cisco ISR平台有关的信息请见本章最后一节所列出的相关参考资料。图8-1所示为分支机构网络单层部署模型的简要拓扑结构图。

在图8-1所示的单层网络架构中，只部署了一台ISR，通过ISP所提供的T1/E1专线来建立WAN连通性。这条T1/E1专线作为与公司总部（HQ）站点互连的主用链路。为实现WAN连接的冗余，还可再申请一条ADSL线路作为T1/E1专线的备用链路。当然，使用其他类型的WAN介质或VPN技术来实现上述设计也无不可。

---

① 以上两句的原文是“When port-to-port IPv6 access is available by the WAN service provider, the dependency for encapsulating IPv6 into IPv4 IPsec or SSL is no longer present.IPv6 over IPsec can be deployed today between Cisco IOS branch routers and the WAN head-end routers”。译者认为，作者行文不成体系，想到哪儿写到哪儿，不具备任何一点写作常识，译文只能勉强翻译。

② 原文是“The following sections provide a high-level overview of the three most commonly deployed Cisco branch profiles and the associated WAN head-end”。

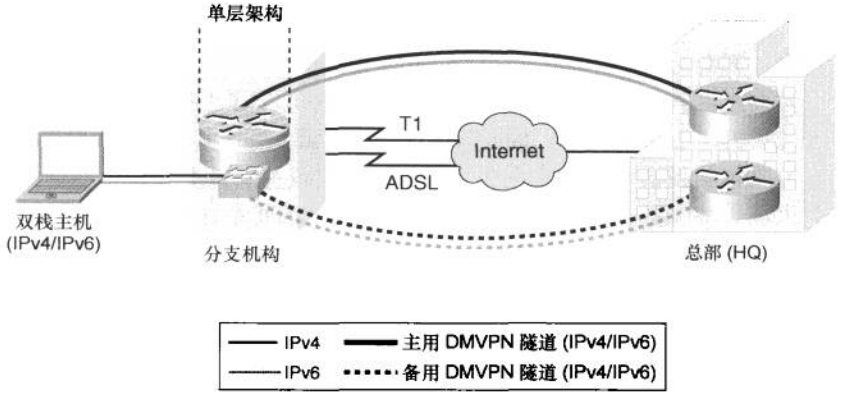


图 8-1 分支机构网络单层部署模型

分支机构网络与 HQ 站点之间的 IPv4 和 IPv6 连通性要靠 IPv4 IPSec（利用 DMVPN 技术）来建立（DMVPN 支持 IPv4 和 IPv6-over-IPv4 IPSec）。DMVPN 的运作方式是，将 IPv4 和 IPv6 流量封装进 GRE 隧道，通过 GRE 隧道在站点间递送流量，利用 IPv4 IPSec 对 GRE 隧道进行加密。DMVPN 隧道建立在 T1 链路之上，作为站点间数据传输的主用路径，一旦 T1 链路失效，则会通过 ADSL 链路建立 DMVPN 隧道。

所有离开分支机构网络的流量都会经由 VPN 连接流入 HQ，其中自然包括了访问 Internet 的流量。一般情况下，Cisco 不建议在分支机构站点中采用隧道分离技术。要是用户执意采用该技术（这就是说，分支机构访问 Internet 的流量直接发送给本地运营商，无需发送至 HQ，而访问公司的业务流量仍需经由 VPN 连接发送），Cisco 给出的建议是：应仔细审查并测试路由选择和安全性对该部署方式所造成的影响。

在单层部署模型中，LAN 连通性由安装在 ISR 上的集成交换模块（EtherSwitch 业务模块）来提供。具体做法是，配置 ISR，令下连分支机构用户的 SVI 接口运行双协议栈（既运行 IPv4 也运行 IPv6 协议栈）。

除了设置于 HQ 网络的所有安全策略以外，还会专门针对分支机构制定一套基础设施安全策略，其相关配置同时作用于分支机构本地的 IPv4 和 IPv6 流量。此外，还会利用上 ISR IOS 内置的软件防火墙特性（比如，Cisco IOS 防火墙或 Cisco IOS 基于区域的防火墙特性）。在 QoS 方面，一套 QoS 策略会同时对

IPv4 和 IPv6 流量生效。

单层部署模型有一个很明显的缺点，那就是无论是在交换还是在路由选择方面，都毫无冗余性可言。只有 Internet 线路和通往总部的 VPN 连接具备冗余性。然而，由于在分支机构网络中只部署了 ISR 这一根“独苗”（既是路由器，又做交换机，还当防火墙），因此只要这台 ISR 或安装于其上的模块发生故障，那么分支机构便会和总部完全失去联系。双层或多层部署模型可在网络部件（防火墙、路由器、交换机以及到总部的连接）的冗余性方面做出改进，从而能够满足客户提出的冗余性需求。

8.1.2 双层部署模型

双层部署模型则在分支机构网络中引入了物理交换机来行使交换功能，并且在保留那台原有 ISR 的基础上，再部署另外一台 ISR，如此一来，便可以实现路由设备和链路级别的冗余性。

图 8-2 所示为双层部署模型示意图。

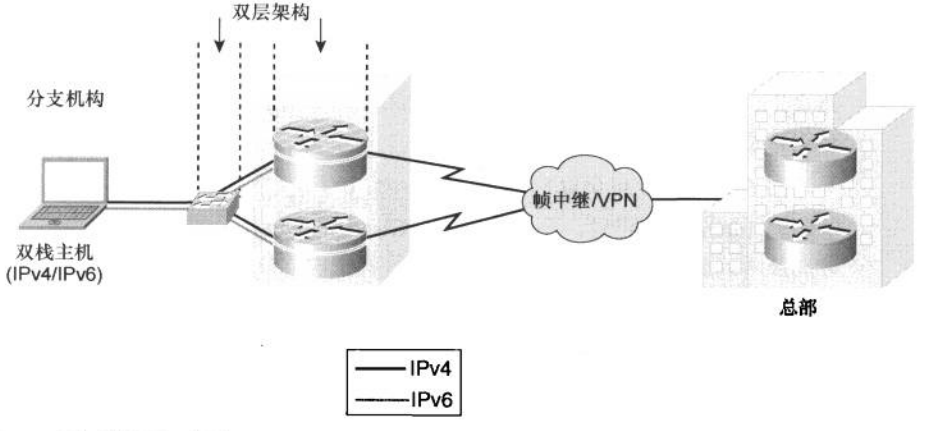


图 8-2 双层部署模型示意图

单层部署模型和双层部署模型之间主要有以下几个方面的差异。

- 冗余性。
- 可扩展性。
- WAN 连接方式。

冗余性

在双层分支机构部署模型中，为了增强高可用性，会部署不同的设备分别行使路由和交换功能。在分支机构中，可部署一台或多台交换机以行使 LAN 接入功能。此外，还会部署 2 台 WAN 路由器以冗余的方式上连帧中继云（或其他类型的 WAN/VPN 连接），下连 LAN 交换机。

可扩展性

与单层分支机构部署模型相比，双层模型要更易于扩展，理由是单层模型中的那台 ISR 集众多功能于一体，无法拆分。正因如此，单层分支机构部署模型不但具备巨大的成本优势，而且组建起来的网络还易于管理，但在高可用性和可扩展性方面却有严重的缺陷。采用单层分支机构部署模型时，分支机构规模越大，那台 ISR 所肩负的任务就越重，于是，因性能过载而造成停运的风险也将随之大增。部署更为高端的 ISR 可以缓解上述风险，但这却并不能满足网络的高可用性需求。采用双层分支机构部署模型，则可在必要时，在分支机构网络中添加多台 Catalyst 交换机，并采用 Cisco 智能堆叠技术来满足 LAN 接入的高可用性需求。

WAN 连接方式

采用双层分支机构部署模型时，可供采用的 WAN 连接类型包括帧中继、点对点 IPSec VPN、DMVPN、多协议标签交换（MPLS），以及由运营商提供的其他任何一种 WAN 连接方式。Cisco IOS 路由器完全支持帧中继上的 IPv6，因此只要采用帧中继作为 WAN 连接方式，那便无需在分支机构和总部网络之间“开凿”隧道，并令其封装 IPv6 流量。这能为 IPv6 的部署和网络管理提供极大的便利，原因很简单——无论是分支机构网络，还是 HQ 网络都可以同时运行双协议栈（IPv4 和 IPv6 并肩运行）。由于无需“开凿”隧道，因此与架设隧道有关的部署工作（比如，隧道的高可用性、安全性、QoS 以及多播流量的发送等）可以全部省略，从而能够大大降低运维成本。

就分支机构网络的安全性而言，单层和双层模型之间并无本质区别，只是在双层模型的部署中，会使用两台路由器来充当安全设备。

8.1.3 多层部署模型

多层分支机构网络部署模型的目标是，根据业务需求，将分支机构网络划

分为若干层次，每个层次都能够实现设备及链路冗余。就本质而言，多层模型结合了单层和双层模型的优点，但在网络的高可用性和可扩展性方面要更胜一筹，此外，还可以提供功能更为强大的防火墙服务。多层分支机构网络部署起来与（企业网中的）小型园区网络（区块）极为相似，两者都采用相同的产品、设计理念和配置。两者之间的差异一般会显现在多层分支机构网络中防火墙和 WAN 路由器的部署方面。

图 8-3 所示为多层分支机构网络部署模型示意图。

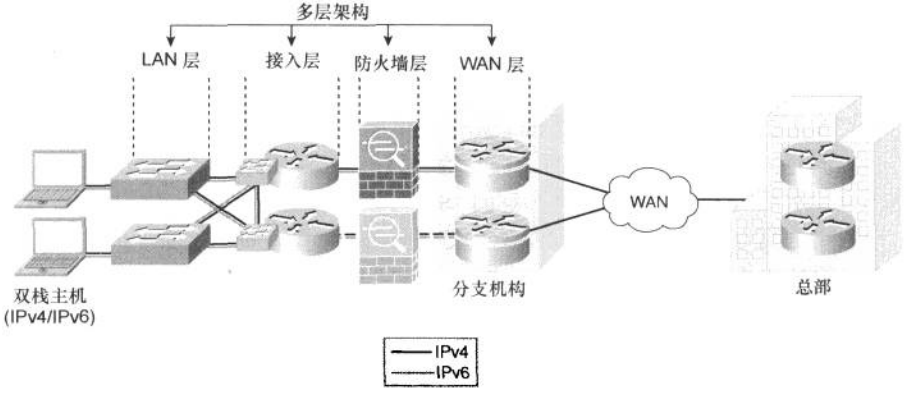


图 8-3 多层分支机构网络部署模型示意图

图 8-3 所示为各层的分布方式。与双层模型相比，多层模型发生了以下明显变化。

- **防火墙层**：不再让 WAN 路由器承担防火墙的功能。图中所示场景为部署了两台 ASA 5500 系列防火墙，针对 IPv4 和 IPv6 流量提供状态化的防火墙服务。图中底部的 ASA 运行于状态化故障切换模式。采用状态化故障切换配置时，ASA 既可运行于 active/standby（如图 8-3 所示）模式，也可以运行于 active/active 模式。
- **接入层**：用来提供分支机构内部的路由选择和 VLAN 终结功能。就功能而言，分支机构网络中的接入层在很多方面都类似于园区网络中的分布层。
- **LAN 层**：图 8-3 所示的 LAN 层承担双层部署模型中那台 LAN 交换机的功能。较之于更大的分支机构网络，在较大范围的需求调研里，这种

情况占比更高[①]。

8.2 WAN/分支机构网络 IPv6 部署通则

有许多 IPv6 部署通则都适用于本章所述的三种分支机构网络部署模型。以下各节将会介绍在分支机构网络中部署 IPv6 时所需考虑的通用原则，这些原则与采用哪种部署模型无关。如果有特定的部署原则需请读者关注，本节亦会随文指出适用于相应原则的部署模型。

本章所讨论的分支机构网络部署模型充分借鉴了最新的 Cisco 分支机构网络设计最佳做法，并以此作为本节内容的基础。就部署方式而言，分支机构网络模型中的 IPv6 部件应尽量与 IPv4 部件相同。

在实战中运用本章所述的分支机构 IPv6 部署模型之前，请读者务必理解 Cisco 推荐的分支机构网络设计最佳做法。Cisco 分支机构网络设计最佳做法文档请见以下链接：http://www.cisco.com/en/US/netsol/ns816/networking_solutions_program_home.html。

8.2.1 编址

大多数情况下，为点对点（P2P）链路分配/64 的前缀，已经是绰绰有余了。设计 IPv6 是为了解决地址空间短缺问题，在分支机构网络中，即便采用了很烂的地址管理方案，应该也不会出现地址不够用的情况。

有些读者可能会问，为 P2P 链路分配/64 的前缀会不会太过浪费。这种做法也一度在 IPv6 社团中引起了争论。对于那些立志要严格管控 IPv6 地址空间的网管人员来说，则可以很安全的在 P2P 链路上使用/126 的前缀，其做法类似于/30 位的 IPv4 前缀用于 P2P 链路。若能确保不与具有特殊用途的 IP 地址发生冲突，甚至还可以在 P2P 链路上使用/127 的前缀。IPv6 地址分配原则请见 RFC 5375，链接为 http://www.ietf.org/rfc/rfc5375.txt。

本章所举示例中的 P2P 链路均使用/64 的前缀。而末端主机则通过两种方式获取 IPv6 地址：其一，利用状态化地址自动分配特性(SLAAC)（详见 RFC 4862 “IPv6 Stateless Address Auto configuration”）获取地址，即由路由器的 VLAN（主机所驻留的 VLAN）子接口通过 RA 消息，通告 IPv6 前缀；其二，通过状态化

① 原文是 “There are just more of them to account for the larger-scale requirements that are most likely found in a larger branch”。这种文字要想翻译，只能靠猜了。

的DHCPv6获取地址。而DNS服务器的各种选项以及域名信息，则利用无状态或状态化的DHCPv6机制来分配。本章稍后将会给出SLAAC、状态化以及无状态DHCPv6的具体配置。

更多与IPv6编址服务有关的信息请见以下URL：

http://www.cisco.com/en/US/docs/ios/ipv6/configuration/guide/15_0/ipv6_15_0_book.html。

8.2.2 物理连接

部署IPv6对物理连接方面的要求与IPv4几乎相同，但还需考虑以下5项要素。

- **是否有足够的带宽**：无论部署哪种新技术、新协议或新应用，都应该首先考虑是否有足够的带宽可用来传递原有和新近生成的流量。对分支机构来说，之所以要更加关注带宽问题，是因为在许多情况下，分支机构网络的WAN连接都是低速链路，故而需依赖QoS去解决带宽问题，其做法也仅限于此。本章并不会罗列IPv6部署对带宽的需求，因为其中的变数太多，故而会以个案为基础来加以讨论。
- **最大传输单元（MTU）和分片问题**：IPv6要求的最小MTU字节数为1260字节。如果IPv6数据包的大小超过了链路层MTU，那就必须在链路层执行数据包的分片和重组，但这一切对IPv6都是透明的。读者可通过阅读RFC 2460 (http://www.ietf.org/rfc/rfc2460.txt) 和 RFC 1981(http://www.ietf.org/rfc/rfc1981.txt)来了解MUT和路径MTU发现（PMTUD）。
- **IPSec VPN**：当IPSec与GRE或手工创建的隧道结合使用时，由于隧道的封装和IPSec报头都会增加IP数据包的长度，因此需重点考虑如何调整路由器接口的MTU值，以确保路由器不强制执行对IPv4数据包的分片。手工配置路由器接口的MTU值，将该值调整为IPv6数据包被GRE和IPSec报头封装前的大小，便可以满足IPv6的MTU需求，而不会存在数据包分片问题[①]。更多与IPSec VPN有关的信息请见Cisco网站列出的

[①] 原文是“By manually configuring the MTU values prior to IPv6 encapsulation, the MTU requirements can be met for IPv6 without fragmentation concerns”。译文按字面意思翻译。但在译者看来，作者还是未把意思说透。译者来举个例子，要是不想让路由器对IPSec封装的数据包分片，需在加密路由器上，把路由器接口的IP MTU值改小。比如，路由器接口的IP MTU值为1500字节，若IPSec+GRE报头的长度为100字节，那就应该把路由器接口的IP MTU值改为1400字节。

IPSec 设计指南，链接为：http://www.cisco.com/en/US/tech/tk583/tk372/tech_design_guides_list.html。

- **无线 LAN（WLAN）上的 IPv6：**应保证 IPv6 也能在无线接入点上正常运转，这与保证其在第二层交换机上平稳运行也没什么不同。然而，对于 WLAN 环境中的 IPv6 部署，还有某些事宜有待考虑，比如，通过 IPv6 来管理 WLAN 设备（AP、控制器等）；通过基于 AP 或控制器的 QoS、VLAN 以及 ACL，去控制 IPv6 流量等。要想充分利用 WLAN 设备所提供的更多智能化特性，AP 和控制器都必须支持 IPv6。写作本章之际，Cisco 在自己的 WLAN 产品线上还不能提供对 IPv6 的强有力支持。
- **IPv6 电话端口：**需要指出的是，对于直连于 Cisco IP 电话端口的主机来说，Cisco 可保障其 IPv6 功能。Cisco IP 电话上用来连接主机的端口也是交换机端口，启用了 IPv6 功能的主机在此类端口上的运作方式与接入 Catalyst 第二层交换机端口无异。

在将本章所述的任意一种 IPv6 部署模型应用于实战之前，除需考虑以上 5 要素之外，我们给读者的建议是：不但要对网络中现有流量特征和网络设备/主机的 CPU/内存使用情况进行分析，而且还需对服务等级协定（SLA）中的条款做进一步的梳理。

8.2.3 VLAN

VLAN 之于 IPv6 如同其之于 IPv4。在网络中运行双协议栈时，IPv4 和 IPv6 流量会在同一 VLAN 内川流不息。欲知最新的分支机构网络 VLAN 设计建议，请参考 Cisco 分支机构 LAN 设计最佳做法文档，链接为：http://www.cisco.com/en/US/ docs/ solutions/Enterprise/Branch/Overview.html。

Cisco 完全支持在数据 VLAN 上使用 IPv6，数据 VLAN（IP 电话之后的 VLAN）会随语音 VLAN 通过 Trunk 一并透传。请务必确保 IP 电话灌有正确的 firmware，并对统一通信管理器做相应的设置，以保证 IPv6 路由器通告消息（基于多播）不在数据和语音 VLAN 间扩散[①]。

---

① 原文是："Care must be taken to ensure that the correct firmware and proper Cisco Unified Communications Manager configurations are made to ensure that the data and voice VLANs do not allow IPv6 router advertisements (multicast-based) to be bled between VLANs"。译者也不晓得这"firmware"是什么玩意，译文只能杜撰。作者有很强的公司自豪感，处处拿 Cisco 专有的东西说事。可作者的能力，尤其是写作能力却又配不上他那份自豪感。

更多与IPv6和Cisco IP电话有关的信息，以及如何让统一通信端点更好地去行使IPv6功能，请见http://www.cisco.com/en/US/docs/voice_ip_comm/cucm/srnd/ipv6/ipv6srnd.html中的“Unified Communications Endpoints”一节。更多与如何在统一通信管理器上部署IPv6有关的信息，请参见以下链接：http://www.cisco.com/ en/ US/docs/voice_ip_comm/cucm/srnd/ipv6/ipv6srnd.html。

8.2.4 路由选择

在企业分支机构网络中，可根据多种因素来选择所运行的IGP，这些因素包括硬件功能、IT人员技能以及网络规模等。本章示例大都采用EIGRP作为IPv4和IPv6的IGP。此外，也会顺带提及作为IPv4和IPv6 IGP的OSPFv2和OSPFv3。

在本章的示例网络中，已按最新的Cisco分支机构网络设计文档，将其所载的最佳做法落实到位。此外，还同时启用了IPv4和IPv6 EIGRP的邻接关系和路由更新通告的认证功能。已根据最新的分支机构网络实施最佳做法，对示例网络中运行的IGP IPv4和IPv6进行了调优。调优IGP，以确保网络的稳定性、可扩展性和路由协议的快速收敛，理应成为任何网络设计的重中之重。

8.2.5 高可用性

与网络高可用性有关的许多内容都超出了本章的范围。只要能够对最新的Cisco分支机构网络设计最佳做法有着透彻的理解，便足以满足诸多高可用性方面的需求了。本章将主要关注以下高可用性方面的内容。

- **冗余的WAN链路**：每个客户可能都会根据自己的实际情况来部署冗余的WAN链路。某些客户可能会选择部署T1（主用）+宽带DSL链路（备用）。帧中继+MPLS连接的组合也很常见。
- **冗余的路由和转发路径**：可利用EIGRPv4和v6来实现。在某些情况下，可以使用等开销多路径（Equal Cost Multi-Path [ECMP]），而在另一些情况下（IPSec GRE和手工配置的隧道），则两条路径分别用作主用和备用链路[①]。

---

① 原文是“This is accomplished by leveraging EIGRP for IPv4 and IPv6. In some cases, Equal Cost Multi-Path (ECMP) is used, and in other cases (IPsec GRE and manual tunnels), one path is preferred over another,but the secondary path is available for redundancy”。在译者看来，对某些初级水平的读者来说，作者不做以上解释可能还好懂一点。其实原文的标题“冗余的路由和转发路径”在起名上就有问题，因为许多新手根本弄不清“路由”和“转发”之间的关系，因此若把标题改为“冗余的数据包转发路径”似乎更好懂一点。

- **第一跳网关的高可用性**：该级别的 HA 适用于任何部署了多台（2 台或 2 台以上）WAN 路由器的分支机构网络环境[①]。本章示例将采用 HSRPv2 作为 IPv4 和 IPv6 的第一跳网关冗余协议。Cisco 亦开发出了可用于 IPv4 和 IPv6 的网关负载均衡协议（GLPB）。

8.2.6 QoS

实施 QoS 时，应根据应用或服务（即 TCP/UDP 端口号）而不是网络协议（IPv4 或 IPv6）来制定 QoS 策略。一般而言，若已经针对具体的应用制定出了现成的 QoS 分类、监管以及排队策略，那么这些 QoS 策略应对 IPv4 和 IPv6 流量一视同仁。

通过 MQC（模块化 QoS CLI）配置 IPv6 QoS 策略时，在 QoS **match** 和 **set** 语句中无需使用 **ip** 关键字，这一点请读者倍加关注。

表 8-1 所示为 Cisco 为同时支持 IPv6 和 IPv4，而对 IOS 所做的语法方面的改进。

表 8-1 IOS 中对 QoS 命令的语法修正

只用于 IPv4 的 QoS 命令语法	IPv4 和 IPv6 共用的 QoS 命令语法
match ip dscp	match dscp
match ip precedence	match precedence
set ip dscp	set dscp
set ip precedence	set precedence

当然，对于某些同时对 IPv6 和 IPv4 流量生效的 QoS 特性（比如，加权随机早期检测（WRED）、监管以及加权轮询（WRR）等），Cisco 并未修改与这些 QoS 特性相对应的 CLI QoS 配置命令。

Cisco 提出了一整套 WAN/分支机构网络 QoS 部署建议，详细信息请参见第 9 章最后一节所给的链接。

8.2.7 安全

在 IPv4 园区网络中，许多常见的威胁和攻击对 IPv6 网络也同样适用。未经

① 原文是“This level of HA applies to any branch and/or WAN head-end connection where there are two or more routers”。其中的“WAN head-end connection”略过未翻，“WAN 前置连接（路由器）”应该是指公司总部用来终结分支机构 WAN 链路的路由器，作者可能想要表达：HA 对部署了多台 WAN 前置路由器的公司总部网络环境同样适用。

授权的访问、欺骗攻击、路由攻击、病毒/蠕虫、拒绝服务（DoS）攻击以及中间人攻击，都只不过是能对IPv4和IPv6同时构成威胁的少数几种攻击手段而已。

对于诸多对IPv4构成潜在威胁的攻击手段来说，根本威胁不到或至少不会以相同的“作案手段”威胁到IPv6。IPv6有自己的一套邻居发现和路由器通告机制，此外，IPv6的报头结构甚至分片方式都与IPv4不同。基于此，本章只涉及IPv6安全的一般性主题，而不会详细讨论IPv6的安全建议和配置。为了识别、弄清并缓解针对IPv6的安全威胁，无论是Cisco公司还是整个业界，都付出了许多努力。我们为读者推荐一本专门介绍IPv6安全主题的Cisco Press图书：《IPv6 Security》，作者是Scott Hogg和Eric Vyncke。欲知更多与IPv6安全有关的信息，请查阅本章最后一节所列出的参考文献。

本节将会指出分支机构网络中有待安全加固的几个区域，并会给出具体示例，以演示如何保护IPv6双栈和隧道流量。

注意

本节所举示例绝非实战部署中的金科玉律，其目的是要让读者明白：在将安全策略应用于自己所维护的网络时，必须权谋再三，谋而后定。在分支机构/WAN网络中实施IPv6安全时，同样需要精心规划。

适用于分支机构网络模型的网络设备安全防护通则如下所示。

- **控制对分支机构路由器和交换机的管理访问**：在每种部署模型所使用的交换机和路由器上，都会配置相关命令，来控制登录自身的管理访问。在所有路由器上创建loopback接口，来作为网络管理和路由选择之用。

  为了严控通过IPv6对网络设备的访问，可在设备上配置ACL，以放行经loopback接口访问管理接口（line vty）的流量。获准访问网络设备的源地址为分配给企业网的IPv6前缀。为使得配置在各类网络设备上的ACL更具可扩展性，在控制对设备的管理访问时，应该以放行（permit）源地址为整个企业网网络前缀为主，而不是拒绝（deny）某些源地址对特定接口的访问[①]。本例中，分配给该企业网站点的IPv6前缀为2001:db8:cafe::/48，如例8-1所示。

① 原文是“To make ACL generation more scalable for a wide range of network devices, the ACL definition can permit the entire enterprise prefix as the primary method for controlling management access to the device instead of filtering to a specific interface on the device”。鉴于作者的写作水平，译者给出这句话的原文，译文仅供参考。

例 8-1 路由器 VTY 配置

```
interface Loopback0
 ipv6 address 2001:DB8:CAFE:1F3::9/128
!
ipv6 access-list MGMT-IN
remark Permit MGMT only to Loopback0
permit tcp 2001:DB8:CAFE::/48 host 2001:DB8:CAFE:1F3::9
deny ipv6 any any log-input
!
line vty 0 4
session-timeout 3
access-class MGMT-IN-v4 in
password 7 08334D400E1C17
ipv6 access-class MGMT-IN in        #Apply IPv6 ACL to restrict
                                    #access
logging synchronous
login local
exec prompt timestamp
transport input ssh
```

- **控制通过 HTTP 对网络设备的访问**：写作本章之际，Cisco IOS 还不支持使用 IPv6 HTTP ACL 去控制对网络设备的访问。记住这一点非常重要，因为目前，对于支持 **ip http access-class** ACL 命令的交换机和路由器来说，只能对使用 IPv4 HTTP 访问交换机的用户进行控制，但对使用 IPv6 HTTP 访问的用户则鞭长莫及。这意味着，此前通过 IPv4 HTTP/HTTPS 访问不到交换机的用户和子网，现在通过 IPv6 能够访问到交换机了。
- **控制平面监管（CoPP）**：CoPP 之所以能够对路由器起到防护作用，是因为该特性能够阻止 DoS 或其他非必要流量平白无故地消耗路由器的 CPU 资源。重要的控制平面和管理平面流量都被赋予了较高的优先级。CoPP 的配置取决于各种因素，并无圭臬可奉，这是因为具体的 CoPP 策略都要以个案为基础来量身定制。更多与 CoPP 有关的信息请见以下链接：http://www.cisco.com/en/US/docs/ios/12_3t/12_3t4/feature/guide/gtr-tlimt.html。
- **控制分支机构 LAN 发送的入站流量**：基于源地址来过滤流量。在分支机构路由器连接 LAN 交换机的接口（子接口）上，根据源地址过滤入

站方向的流量是一种很常见的做法。基于源地址来控制 IPv6 流量，可保护网络避免遭受地址欺骗攻击[①]。

例 8-2 所示为一个用来行使基本过滤功能的 ACL 示例：该 ACL 应用于分支路由器 LAN 接口的入站方向。

例 8-2 具备基本过滤功能的 ACL（应用于分支机构路由器的 LAN 接口）

```
ipv6 access-list DATA_LAN-v6
remark PERMIT ICMPv6 PACKETS FROM HOSTS WITH PREFIX 2001:DB8:CAFE:1004::/64
permit icmp 2001:DB8:CAFE:1004::/64 any
remark PERMIT IPv6 PACKETS FROM HOSTS WITH PREFIX 2001:DB8:CAFE:1004::64
permit ipv6 2001:DB8:CAFE:1004::/64 any
remark PERMIT ICMPv6 PACKETS SOURCED BY HOSTS USING LINK-LOCAL
permit icmp FE80::/10 any
remark PERMIT DHCPv6 ALL-DHCP-AGENTS REQUESTS FROM HOSTS
permit udp any eq 546 any eq 547
remark DENY ALL OTHER IPv6 PACKETS AND LOG
deny ipv6 any any log-input
!
interface GigabitEthernet0/0.104
 description VLAN-PC
 ipv6 traffic-filter DATA_LAN-v6 in
```

Cisco IOS IPv6 ACL 中总是“暗伏”了放行 IPv6 邻居发现消息流量的 ACE。若配置了 **deny ipv6 any any**，那么暗伏的放行 IPv6 邻居发现消息流量的 ACE 将会失效。要是出于日志记录的目的而在 ACL 中手动配置了 **deny ipv6 any any log-input** 语句，那么在此之前一定要配置 **permit icmp any any nd-na** 和 **permit icmp any any nd-ns** 这两条 **permit** 语句，记住这一点非常重要。

- **IPv6 状态化防火墙服务**：防火墙可按状态化的方式对进出分支机构网络的流量执行安全检测。写作本书之际，Cisco ASA 5500 系列产品、Cisco IOS 防火墙以及基于区域的 Cisco IOS 防火墙都在不同层次上支持对 IPv6 流量的检测。至于以上哪种防火墙解决方案适用于自己的网络环境，则需查阅 Cisco 文档、请教 Cisco 顾问团队或 Cisco 合作伙伴。
- **禁用尚未使用的服务**：许多服务（比如 HTTP）可同时使用 IPv4 和 IPv6。

① 原文是“Filter which prefixes are allowed to source traffic. This is most commonly done on ingress on the LAN or subinterface on the branch router. Controlling IPv6 traffic based on source prefix can help protect the network against basic spoofing”。

一般情况下，禁用或启用这些服务会对 IPv4 和 IPv6 同时生效。禁用任何未使用的服务是网管人员应长期奉行的准则。

8.2.8 多播

IPv6 多播对任何企业网络来说都是一项重要服务。在分支机构网络的 LAN 内，能否有效控制主机/多播组之间的对应关系，是在分支机构网络中部署 IPv6 多播所需考虑的重要因素之一。在功能上，IPv6 所使用的多播侦听者发现（MLD）协议等价于 IPv4 所使用的 Internet 组管理协议（IGMP）。在 IPv6 和 IPv4 中，MLD 和 IGMP 分别用来控制多播组成员关系。MLD 欺骗特性能够控制多播流量的分布——令交换机只把多播流量从连接了多播接收者的接口外发。要是没有该特性，即便只连接了一个（或几个）多播接收者，分支机构 LAN 交换机也会在隶属同一 VLAN 的所有端口上泛洪多播流量。因此，在分支机构 LAN 中，交换机能否支持 MLDv1 和 v2 的 MLD 欺骗特性就变得至关重要了。

如今，Cisco 支持以下 PIM 实现：PIM-SM、PIM-BSR、PIM-SSM、双向 PIM、嵌入 RP 以及用于 IPv6 多播地址家族的多协议 BGP。

业界和 Cisco.网络都有许多详细介绍 IPv6 多播的文档。本章除了会给出用来启用 IPv6 多播以及配置嵌入 RP 所需的命令以外，不会再有其他与 IPv6 多播相关的配置说明。更多与 IPv6 多播有关的信息，请参考以下 URL。

- **Cisco IPv6 多播：**

http://www.cisco.com/en/US/technologies/tk648/tk828/tk363/technologies_white_paper0900aecd8014d6dd.html

- **Cisco IOS IPv6 多播配置：**

http://www.cisco.com/en/US/tech/tk828/technologies_white_paper09186a00802 03f7a.shtml

8.2.9 管理

IPv6 网络管理尚处于发展之中，还有很长的一段路要走[①]。诸多可用于 IPv4 的传统型网络管理工具同样适用于 IPv6。本章所讨论的分支机构网络管理原则只涉及基本的网络管控服务（Telnet、SSH 和 SNMP）。可利用 Telnet 和 SSH，

① 原文是“Management for IPv6 is under development and has a long way to go”。“Management for IPv6”是指 IPv6 网络管理工具、方法还是技术，抑或是网络设备对 IPv6 管理的支持？作者既然没有提，译者也懒得费心机去调整译文了，译文为直译。

通过IPv6，管理到前述的两种分支机构部署模型中的所有启用了IPv6的网络设备。

IPv6的SNMP部署类似于IPv4。在本章所述的各种分支机构网络部署模型中，可采用SNMPv3 (AuthNoPriv)，为总部数据中心网络区块内的Cisco网管系统（NMS）提供轮询功能[①]。以下便是在本章所涉分支机构路由器上所采用的SNMPv3配置示例。

```
snmp-server contact John Doe - ipv6rocks@example.com
snmp-server group IPv6-ADMIN v3 auth write v1default
snmp-server user jdoe IPv6-ADMIN v3 auth md5 cisco1234
```

要想将SNMP信息发送给Cisco NMS服务器，便需要在分支机构网络设备上配置SNMP网管主机的IP地址。可对分支机构网络设备进行配置，令其通过IPv4或IPv6将SNMP信息发送给NMS服务器。

```
snmp-server host 2001:DB8:CAFE:100::60 version 3 auth jdoe
```

身为网管人员，还需参透网络管理的另一领域，即地址管理领域的方方面面。为众多网络设备分配超长的16进制IPv6地址时，即便不能采用自动化的分配手段，所采用的手段最起码也要比如今在用的手动分配法要更加人性化一点。

如今，Cisco开发出了一种在Cisco设备上分配IPv6前缀的方法，即通用前缀特性。使用该特性，网管人员可在Cisco设备的全局配置模式下定义一条或多条前缀，并为相关前缀赋予人性化的名称。随后，以每接口为基础配置IPv6地址时，可使用该人性化名称来替代惯用的IPv6前缀。通用前缀特性最适用于IPv6地址前缀频繁变动的场景，比如，IPv6试点阶段或早期的生产网络环境中，因为此时最后的IPv6编址策略尚未完全落实到位。以下所示为配置通用前缀特性的示例。

步骤1 定义通用前缀。

```
br1-1(config)# ipv6 general-prefix BRANCH-1 2001:DB8:CAFE::/48
```

步骤2 在接口配置模式下，配置名为BRANCH-1的通用前缀。

```
br1-1(config-if)# ipv6 address BRANCH-1 ::1005:0:0:0:1/64
```

① 作者的原意应该是，在支机构网络中的网络设备上开启SNMP功能，以便于总部数据中心网络区块内的Cisco网管系统能够通过SNMP管理到这些设备。作者本来文笔就不行，行文还喜欢七拐八绕，绕来绕去只会让人不知所云。

步骤 3 验证配置在接口上的通用前缀。

```
br1-1# show ipv6 interface g1/0.100
GigabitEthernet1/0.100 is up, line protocol is up
  IPv6 is enabled, link-local address is FE80::217:94FF:FE90:2829
  No Virtual link-local address(es):
  Description: DATA VLAN for Computers
  Global unicast address(es):
   2001:DB8:CAFE:1005::1, subnet is 2001:DB8:CAFE:1005::/64
```

与通用前缀特性有关的详细信息请见 Cisco 官网 Cisco IOS IPv6 文档页面，链接为：http://www.cisco.com/en/US/docs/ios/ipv6/configuration/guide/ip6-addrg_bsc_con_ps10591_TSD_ Products_Configuration_Guide_Chapter.html#wp1132473。

可通过各种网络管理工具对具备 IPv6 功能的 Cisco 网络设备进行管理[①]。欲知更多与 Cisco 网络管理解决方案有关的信息，请见 http://www.cisco.com/en/US/products/sw/netmgtsw/index.html。

本书第 11 章将会深入探讨 IPv6 网络管理。

8.2.10 可扩展性和性能

本章并不准备去分析与各种测试用硬件平台的扩展性和性能有关的细枝末节。本章将侧重于介绍在分支机构网络中规划和部署 IPv6 时，所应恪守的一般性原则，这些原则与特定的硬件平台无关。

在分支机构网络中规划和部署 IPv6 时，应考虑如下事宜。

- **链路带宽占用率**[②]：实施 IPv6 部署时，分支机构网络链路上的流量占用率（IPv4 和 IPv6 流量所占链路带宽的比重）发生改变的情况会非常常见[③]。随着 IPv6 部署如火如荼地进行，当用户使用 IPv6 来发送他们之前用 IPv4（也只能用 IPv4）来发送的应用程序数据时，IPv4 流量所占链路的带宽比重会呈递减之势。而整体链路带宽利用率则通常会略有增加，这通常是拜控制平面的路由选择流量和隧道封装的开销（如果网

① 原文是“Cisco supports the management of IPv6-enabled network devices through a variety of network management products to include DNS, DHCPv6, device management and monitoring, and network management, troubleshooting, and reporting”。译者不晓得作者到底想表达什么，只能对原文的后半句“忍痛割爱”。

② 原文是“Traffic utilization（流量利用率）”，译文按中国人的说话习惯，翻为“链路带宽占用率”。

③ 原文是“In IPv6 implementations, it is common to see a change in traffic utilization ratios on the branch network links”。译文没有按照作者字面意思翻译，请读者注意。

络中架设了各种隧道）所赐。

- **路由/转发**：身为网管人员，对分支机构网络路由器的路由和转发能力一定要心中有数。如果现有的分支机构路由器已经为处理 IPv4 路由选择，耗费了众多 CPU 和内存资源，就不应再在其上新增 IPv6 功能。要是分支机构路由器以硬件方式来执行路由选择功能，那么开启 IPv6 路由选择对其影响不大。
- **对 ACL 的处理**：应用 ACL 之前，一定要斟酌再三，读者应视其为一条铁律。在分支机构路由器上，IPv6 ACL 主要有三个作用：其一，用于 QoS（对来自接入层的入站数据包进行分类和标记）；其二，用于安全（用在接入层防范 DoS、欺骗攻击以及未经授权的访问）；其三，用于 QoS+安全，保护路由器的控制平面不受攻击。此外，还可以通过分支机构路由器为 IPv4 和新近部署的 IPv6，提供状态化防火墙服务、IDS/IPS 服务以及其他新服务。在分支机构路由器上开启新的高级服务时，应同时支持 IPv4 和 IPv6。在分支机构路由器上开启上述服务，且配置了 IPv6 功能之后，其性能必会受到影响[①]。

Cisco 出过一份文档，该文档记录了启用 IPv4 和 IPv6 时，Cisco 路由器平台性能方面的比较参数，链接为 http://www.cisco.com/web/strategy/docs/gov/IPv6perf_wp1f.pdf。

8.3 WAN/分支机构网络实施示例

就三种不同的 WAN/分支机构网络模型的配置和设计方案而言，其中亦有许多相似之处。一般而言，变数最多的可能要数分支机构网络中的设备数量了，出于高可用性的目的，外加对整体网络可扩展性的考虑，有一点变数也不奇怪。

本章所举的实施示例集三种 WAN/分支机构网络模型的特质于一身，以使读者能够更好的理解 WAN/分支机构网络的各个层次结构、网络功能以及具体的产品和特性，在实战场景中配置 IPv6 时，可对此加以应用。

在本章剩余的内容中，我们会把该示例网络拓扑称之为“混合型分支机构网络示例”，或简称为 HBE（hybrid branch example）。此外，本章以 HBE 为例，

① 从“此外……”开始，直至段末，和 ACL 可有一丝半点关系？作者安插这段文字除了误导读者以外，还有其他目的吗？

只是意在将三种 WAN/分支机构网络模型的要素结合在一起，而无意使其成为推荐的“样板（best practice）”设计方案。

图 8-4 所示为 HBE 网络环境的示意图。

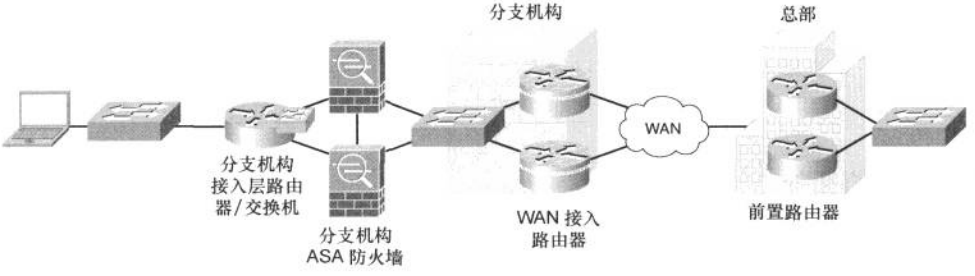


图 8-4 混合型分支机构网络示例概述

该示例网络（HBE）具备很强的灵活性，可采用几乎各种 WAN 连接方式，包括帧中继、MPLS、点到点 IPSec、DMVPN 等。在 HBE 的分支机构网络中，部署了两台冗余的 WAN 接入路由器，并利用冗余的链路，连接到总部的两台前置路由器。在 WAN 接入路由器身后，还部署了一台 Cisco ASA 5500 系列防火墙。当然，出于高可用性的目的，也可以部署一对互为冗余的 ASA。还可部署一台配备了 Cisco EtherSwitch 模块的 Cisco ISR 系列路由器（或不配 EtherSwitch 模块，单独配备一台 Catalyst 交换机），来连接分支机构本地的主机资源（包括 PC、打印机以及其他网络资源等）。

为满足各分支机构不同的业务需求，可能还需添加额外的设备，比如，可用来增强网络高可用性、安全性、分支机构网络服务健壮性的路由器、交换机和其他网络设备等。

注意 本章不会给出网络设备的完整配置，只会示出与部署 IPv6 有关的配置摘要。

8.3.1 试验用网络设备

表 8-2 列出了混合型分支机构网络示例中所采用的实验用网络设备。

表 8-2 HBE 实验用网络设备

功 能	硬 件	软 件
路由器	集成服务路由器：2800 和 3800 系列	Advanced Enterprise Services 15.0.1M1
交换机	Cisco Catalyst 3750E/3560E	12.2(46)SE
防火墙	Cisco ASA 5510	8.2(2)
主机设备	各种笔记本、PC	Microsoft Windows Vista、Windows 7

8.3.2 网络拓扑

图 8-5 所示为 HBE 的精确拓扑结构图。该图示出了分支机构网络连接，及分支机构和总部之间网络连接的 IPv6 编址规划。

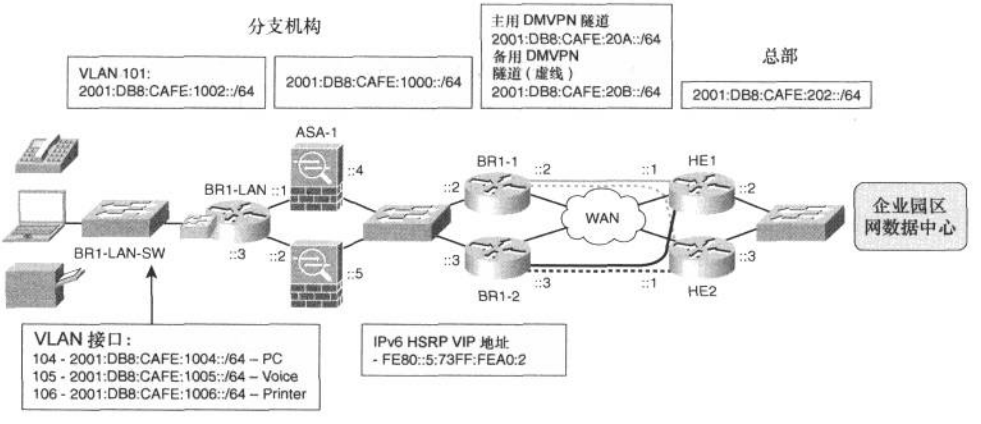


图 8-5 HBE IPv6 编址详情

以下各节将从物理和逻辑层面探讨与 WAN 接入、分支机构 LAN 以及防火墙有关的连通性问题。

WAN 连通性

HBE 采用支持 spoke-to-spoke 的双 DMVPN 云技术（Dual DMVPN Cloud Topology with spoke-to-spoke support）来建立 WAN 连通性，Cisco DMVPN 设计指南请见如下链接：

http://www.cisco.com/en/US/docs/solutions/Enterprise/WAN_and_MAN/DMVPDG.html

双 DMVPN 云技术是指分支机构与总部之间采用主/备 DMVPN 隧道的方式，去建立连通性（主/备 DMVPN 隧道在图中分别用实/虚线表示）。主/备 DMVPN 隧道各形成一个独立的 IPv4 和 IPv6 网络。还需对 IGP 加以调整，以使得分支机构路由器与总部前置路由器之间优选其中的一条隧道传递流量，若该主用隧道故障，IGP 应快速收敛，备用隧道会取而代之。

使用传统的帧中继、MPLS 或 IPSec VPN 在支机构和总部之间建立 WAN 连通性也无不可。本例选择 DMVPN 作为 WAN 连接手段的原因是，向读者展示一份实用的支持 IPv6 的 Cisco DMVPN 配置。

既然这只是一个示例，而且影响网络连接和配置的因素又有很多，故而在寻址和物理连接方面就采用了一种简单的方法。本例中很重要的一点是：场景与 IPv4 环境大体相同。本例的目的则是展示语法（配置 IPv6 和 IPv4 的语法）调整的细微之处。

分支机构 LAN 连通性

在 WAN 接入路由器和 Cisco ASA 之间通过一台 Catalyst 交换机来建立 LAN 连通性。将两台路由器配置为 HSRP 组成员，并同时作用于 IPv4 和 IPv6。在 Cisco ASA 上，设有一条默认路由，下一跳为 HSRP VIP 地址。

LAN 接入路由器通过内置其中的一块 EtherSwitch 模块与 ASA 互连。当然，也可以在两者之间部署一台专用的 Catalyst 交换机来代替这块 EtherSwitch 模块。

在该分支机构网络中，使用 Catalyst 交换机来提供主机、打印机以及 IP 电话的 LAN 接入。在 HBE 中启用了三个 VLAN 来用于主机的接入。

- **VLAN 104**：用于数据 VLAN，供 PC 接入。由配置在 LAN 接入路由器上的本地 DHCP 地址池为主机分配 IPv4 地址。IPv6 地址则由分支机构路由器利用 SLAAC 来提供，而 DNS/域名则由分支机构路由器上所配的本地 IPv6 DHCP 地址池提供。当然，也可以选择由总部站点通过 DHCP 为该 VLAN 提供“一条龙式”的 IPv4 和 IPv6 服务。
- **VLAN 105**：用于语音 VLAN。由配置在 LAN 接入路由器上的本地 DHCP 地址池来分配 IPv4 地址，并同时提供其他语音专用的 DHCP 选项（TFTP 服务器信息）。IPv6 地址则由状态化的 DHCPv6 机制来提供。当然，使用无状态的 IPv6 DHCP 也无不可。
- **VLAN 106**：用于打印机 VLAN。由配置在 LAN 接入路由器上的本地

DHCP 地址池来分配 IPv4 地址。分支机构中的打印机服务器网卡通过无状态的自动配置机制，从路由器接口获取 IPv6 地址。当然，也可以选择由总部站点通过 DHCP 为该 VLAN 提供“一条龙式”的 IPv4 和 IPv6 服务。

防火墙连通性

在分支机构网络中，可部署也可以不部署专用防火墙，这要视网络设计和具体的安全策略而定。分支机构网络防火墙的部署用途不外有二：其一，用于分支机构本地的 Internet 访问（隧道分离的场景）；其二，防火墙自身用作为分支机构的 VPN 设备。此外，还可在 WAN 接入路由器上激活 IOS 防火墙功能，提供边界防护功能，以替代专用的 ASA 防火墙。

在 HBE 中，Cisco ASA 防火墙的配法和用法无非都是一些最基本的功能。需配置“outside”和“inside”接口各一。Cisco ASA 既可按不带冗余功能的单台防火墙来部署，也可按具备状态化故障切换功能的故障切换对形式来部署，对于后一种情形，需部署第二台 ASA 作为备用单元（standby unit）来使用（如图 8-4 所示）。

可将 Cisco ASA 配置为路由模式或透明模式（有时也称为桥接模式）。本例选择使用路由模式，该模式也是最流行的部署方式。简单来说，运行于路由模式下的 ASA 拥有多个第三层接口，每个接口分别隶属于不同的 IPv4 和 IPv6 网络，并算作相应 IPv4 和 IPv6 网络的路由跳（运行于路由模式下的 ASA 既可运行静态路由协议，也可运行动态路由协议）。可把运行于透明模式的 ASA 视为一台第二层网桥，数据包以桥接的方式穿越 ASA，并被 ASA 检测[①]；一般来说，运行于该模式下的 ASA 等于是“bump-in-the-wire”。以上只是对防火墙两种模式的简单介绍，读者应充分理解两种模式之间的差别和优劣。更多与 ASA 透明/路由模式有关的信息请见：http://www.cisco.com/en/US/docs/ security/asa/asa83/configuration/guide/fwmode.html。

前置路由器的配置

HBE WAN 的配置同样要涉及总部站点，在总部站点内，亦部署有两台 Cisco 路由器作为双 DMVPN 云拓扑的终结路由器。

---

① 原文是“Transparent mode has the ASA in a Layer 2 configuration where packets are bridged across and inspected”。译文与原文字面意思有一定的出入，请读者注意。

注意

前置路由器的硬件平台和型号随网络的规模和可扩展性方面的需求而异。请咨询 Cisco 顾问团队和/或 Cisco 合作伙伴，以选择最适合的产品来行使前置路由器的功能。

这两台前置路由器通过 FE 链路接入 ISP，当然也可以通过 T1/E1、DS3 或任何其他连接方式接入 ISP。本章选择使用 FE 链路去生成相关配置。

DMVPN 是一种可承载 IPv4 和 IPv6 流量的 VPN 技术。本章所示的 DMVPN 配置采用了 Cisco IOS 对 DMVPN 第三等级的支持（Phase 3 of Cisco IOS support for DMVPN）。以下所列为针对 DMVPN 定义的三个等级。

- **第一等级**：只支持分支站点到中心站点的直接访问（Hub-and-spoke capability only）。
- **第二等级**：初步支持分支站点间的直接互访（Initial spoke-to-spoke capability）。
- **第三等级**：支持 IPv6 以及分支站点间直接互访的增强功能，从而能够更好的适用于大型 NBMA（非广播多路访问）网络（Support for IPv6 and enhancements for spoke-to-spoke to support larger-scale non-broadcast multiaccess (NBMA) networks）。

读者无论是想了解与 IPv6 DMVPN 的原理、操作以及配置有关的信息，还是想弄清 DMVPN 第三等级的增强功能与下一跳解析协议（NHRP）的操作方式，均可参阅以下 URL。

- **实施 IPv6 DMVPN：**

  http://www.cisco.com/en/US/docs/ios/ipv6/configuration/guide/ip6-dmvpn_ps10591_TSD_Products_Configuration_Guide_Chapter.html。

- **NHRP 的快速交换增强功能（Shortcut switching enhancements for NHRP）：**

  http://www.cisco.com/en/US/docs/ios/ipaddr/configuration/guide/iad_nhrp_dmvpn.html#wp1072593。

- **配置 NHRP：**

  http://cisco.com/en/US/docs/ios/ipaddr/configuration/guide/iad_cfg_nhrp.html#wp1078234。

需要为DMVPN配置不同的特性和参数，如密钥、保持时间等[①]。

前置路由器HE1和HE2上，各创建了一条隧道（接口）[②]。HE1为主用前置路由器，并因本例采用的是双DMVPN云技术，故而为HE1所用的隧道与HE2所使用的隧道分处不同的IPv4和IPv6子网。还有一件要务请读者注意，在路由器上激活IPv6多播时，PIM会使用Tunnel0和Tunnel1与RP和隧道源通信。因此，在创建隧道（接口）时建议从Tunnel2开始创建。

例8-3所示为HE1的配置。除了IPv4和IPv6地址和路由优先级以外，HE2的配置与HE1完全相同。HE2的配置便不再给出。

例8-3 HE1的配置

```
ipv6 unicast-routing
ipv6 cef
!
crypto isakmp policy 1       #Set ISAKMP Policy using pre-shared
                               #keys
 encr aes 256
 authentication pre-share
 group 2
!
crypto isakmp key CISCO address 0.0.0.0 0.0.0.0
crypto isakmp key CISCO address ipv6 ::/0  #Pre-share key for
                                             #any (::/0) peer
!
crypto ipsec transform-set HUB esp-aes 256 esp-sha-hmac
!
crypto ipsec profile HUB
 set transform-set HUB
!
interface Tunnel2                    #If deployed, PIMv6 uses
                                       #tunnel 0 and 1 by default
                                       #so it is recommended to start
                                       #at 2
 description DMVPN Tunnel 1
 ip address 10.126.1.1 255.255.255.0
 no ip redirects
```

（待续）

① 原文是“You need to configure different features and values for the DMVPN configuration such as keys, hold times, and so on”。原文如此，译者直译。

② 原文是“HE1 and HE2 have one tunnel configuration each”。

```
 no ip unreachables
 no ip proxy-arp
 ipv6 address 2001:DB8:CAFE:20A::1/64
 ipv6 mtu 1416                        #Set MTU to account for
                                       #Tunnel/IPSec overhead
 ipv6 eigrp 10                        #Enable IPv6 EIGRP
 ipv6 hold-time eigrp 10 35
 no ipv6 next-hop-self eigrp 10
 no ipv6 split-horizon eigrp 10
 ipv6 nhrp authentication CISCO      #Set authentication string
                                       #for NHRP
 ipv6 nhrp map multicast dynamic     #Automatically add routers to
                                       #NHRP mappings
 ipv6 nhrp network-id 10             #Enables NHRP on interface
 ipv6 nhrp holdtime 600
 ipv6 nhrp redirect                   #Phase 3 NHRP redirect for
                                       #spoke-to-spoke
 tunnel source Serial1/0
tunnel mode gre multipoint             #Multipoint GRE to support
                                         #multiple end-points
tunnel key 10
tunnel protection ipsec profile HUB  #Apply IPSec profile
!
interface GigabitEthernet2/0         #LAN interface to HQ network
 description to HQ
 ip address 10.123.1.2 255.255.255.0
 ipv6 address 2001:DB8:CAFE:202::2/64
 ipv6 eigrp 10
 standby version 2
 standby 2 ipv6 autoconfig
 standby 2 priority 120
 standby 2 preempt delay minimum 30
 standby 2 authentication CISCO
 standby 2 track 2 decrement 90
!
interface FastEthernet0/0
 description to ISP
 ip address 172.16.1.1 255.255.255.252
!
ip route 0.0.0.0 0.0.0.0 172.16.1.2
!
ipv6 router eigrp 10                          #Enable EIGRP for IPv6
 no shutdown
```

分支机构WAN接入路由器的配置

例中的分支机构路由器采用串行链路（T1/E1）接入ISP。当然，选择宽带（DSL/cable/无线）、以太网或DS3链路接入ISP也无不可。分支机构WAN接入路由器与ISP只建立了IPv4连通性，应在其上配置ACL，以放行进/出ISP的必要流量，即只放行用来与总部前置路由器建立DMVPN连接所涉及的相关流量（匹配具体的端口号/协议类型）（假定本例不允许隧道分离[split tunneling]）。分支机构WAN接入路由器IPv6部分的配置类似于总部前置路由器的配置，IPv6的配置主要涉及本地以太网接口和DMVPN隧道接口。

两台分支机构WAN接入路由器（BR1-1和BR1-2）的配置几乎相同，只是在IPv4和IPv6地址、路由优先级以及HSRP优先级的唯一性方面有所差异。例8-4所示为BR1-1的配置（只示出了两条DMVPN隧道中一条隧道的配置）。

例8-4 BR1-1的配置

```
ipv6 unicast-routing
ipv6 cef
!
crypto isakmp policy 1
 encr aes 256
 authentication pre-share
 group 2
!
crypto isakmp key CISCO address 0.0.0.0 0.0.0.0
crypto isakmp key CISCO address ipv6 ::/0
!
crypto ipsec transform-set SPOKE esp-aes 256 esp-sha-hmac
!
crypto ipsec profile SPOKE
 set transform-set SPOKE
!
interface Tunnel2
description to HUB
 ip address 10.126.1.2 255.255.255.0
 no ip redirects
 no ip unreachables
 no ip proxy-arp
 ipv6 address 2001:DB8:CAFE:20A::2/64
 ipv6 mtu 1416
 ipv6 eigrp 10
```

（待续）

```
 ipv6 hold-time eigrp 10 35
 no ipv6 next-hop-self eigrp 10
 no ipv6 split-horizon eigrp 10
 ipv6 nhrp authentication CISCO
 ipv6 nhrp map 2001:DB8:CAFE:20A::1/64 172.16.1.1
 ipv6 nhrp map multicast 172.16.1.1
 ipv6 nhrp network-id 10
 ipv6 nhrp holdtime 600
 ipv6 nhrp nhs 2001:DB8:CAFE:20A::1
 ipv6 nhrp shortcut
 tunnel source Serial1/0
 tunnel mode gre multipoint
 tunnel key 10
 tunnel protection ipsec profile SPOKE
interface Serial1/0
 description to ISP
 ip address 172.16.1.9 255.255.255.252
!
interface GigabitEthernet2/0
 description to BRANCH LAN
 ip address 10.124.1.2 255.255.255.0
 negotiation auto
 ipv6 address 2001:DB8:CAFE:1000::2/64
 ipv6 eigrp 10
 standby version 2
 standby 1 ip 10.124.1.1
 standby 1 priority 120
 standby 1 preempt delay minimum 30
 standby 1 authentication CISCO
 standby 1 track 1 decrement 90
 standby 2 ipv6 autoconfig
 standby 2 priority 120
 standby 2 preempt delay minimum 30
 standby 2 authentication CISCO
 standby 2 track 2 decrement 90
!
router eigrp 10
 network 10.0.0.0
!
ip route 0.0.0.0 0.0.0.0 172.16.1.10
!
ipv6 router eigrp 10
 no shutdown
```

分支机构防火墙的配置

如前所述，HBE 中 Cisco ASA 防火墙的部署方式非常简单，仅供读者参考。由于人们普遍认为分支机构网络是通过私有 WAN/VPN 链路接入总部网络的受信网络，因此总会竭力避免在分支机构网络中部署防火墙，并希望以此来降低采购和管理成本。正因如此，人们往往总会在 WAN 接入路由器上配些 ACL，来防范那些最原始的攻击。对于本例，由于分支机构中的用户无法直接访问 Internet，因此无需设置全面的防火墙防护策略，故而可避免部署专用防火墙（或防火墙的故障切换对形式）给网络管理增添的成本和复杂度，在这一点上，网络工程师们的想法势必总能保持一致。

此处，作者并不想去争论在分支机构网络中部署专用防火墙的价值所在，只是意在给出将专用 Cisco ASA 防火墙部署于分支机构网络的基本设计思路和配置示例。

以下给出的 Cisco ASA 防火墙的配置基于软件版本 8.2(2)，防火墙运行于路由模式，且配置为互为冗余。

由于每个客户所使用的应用程序类型和 ACL 配置选项都有所不同，因此本章只会给出基本的 ACL 配置示例以供读者参考。

注意

本章只会给出与 IPv6 有关的配置，这在前面已多次提到。本节所示 Cisco ASA 的配置并不完整，并未包括在分支机构内部署全功能企业级防火墙的所有必要配置。切勿将本节所示配置视为一般情况下的 Cisco ASA 或安全的最佳做法。

例 8-5 中的配置以 alias（别名）的配置打头，将一个用户自定义的名称与 IPv6 前缀相关联，以“BR1-LAN”指代前缀 2001:DB8:CAFE:1003::/64。此外，还创建了另一个 alias（别名），将一个完整的 IPv6 地址（部署在分支机构网络内的一台 IPv6 服务器）与一个用户自定义的名称相关联。

为 outside 和 inside 接口定义了安全级别、IPv4 和 IPv6 地址。命令中的 **standby** 关键字用来定义互为冗余的对等 ASA 防火墙的地址。

配置了一个用来举例的 object group（并不是非配不可），作用于使用 TCP 端口号 3389 的 RDP 协议流量。配置 ACL 时，会调用该 object group，以放行源为 2001:DB8:CAFE::/48，发往先前定义的分支机构服务器（Br1-v6-Server）的 RDP 协议流量。然后，将该 ACL 应用于 outside 接口。

写作本书之际，Cisco ASA 只支持 IPv4 动态 IGP。对于 IPv6，则支持静态路由。对于本例，需在 Cisco ASA 上，针对分支机构内部 LAN 所隶属的各 VLAN 子网，以及 Cisco ASA 与安装在 BR1-LAN 路由器上的 EtherSwitch 模块之间的网络，设置静态路由。其中的一条静态路由调用了先前定义的 alias。针对 outside 接口配置了一条静态默认路由，下一跳指向两台分支机构 WAN 接入路由器 LAN 接口的 HSRP VIP 地址。

Cisco ASA 的 GigabitEthernet0/3 用作为故障切换接口，其中 ASA-1 为主用单元。在故障切换接口上，需配置 IPv4 或 IPv6 地址，但不能同时配置。在本例中，采用 IPv6 地址作为故障切换接口 IP 地址。

最后，在 inside 接口上激活 SSH，以开放源地址落在所配前缀范围内的主机对 ASA 的远程管理权限。

注意

下面将要展示的配置全都通过 CLI 完成。当然，也可以通过 Cisco 自适应安全管理器（ASDM）CLI 来完成配置。

例 8-5 ASA-1 的配置

```
name 2001:db8:cafe:1003:: BR1-LAN description VLAN on EtherSwitch
name 2001:db8:cafe:1004:9db8:3df1:814c:d3bc Br1-v6-Server
!
interface GigabitEthernet0/0
 description TO WAN
 nameif outside
 security-level 0
 ip address 10.124.1.4 255.255.255.0 standby 10.124.1.5
 ipv6 address 2001:db8:cafe:1000::4/64 standby 2001:db8:cafe:1000::5
!
interface GigabitEthernet0/1
 description TO BRANCH LAN
 nameif inside
 security-level 100
 ip address 10.124.3.1 255.255.255.0 standby 10.124.3.2
 ipv6 address 2001:db8:cafe:1002::1/64 standby 2001:db8:cafe:1002::2
!
interface GigabitEthernet0/3
 description LAN Failover Interface
```

（待续）

```
!
object-group service RDP tcp
 description Microsoft RDP
 port-object eq 3389
!
ipv6 route inside BR1-LAN/64 2001:db8:cafe:1002::3
ipv6 route inside 2001:db8:cafe:1004::/64 2001:db8:cafe:1002::3
ipv6 route inside 2001:db8:cafe:1005::/64 2001:db8:cafe:1002::3
ipv6 route inside 2001:db8:cafe:1006::/64 2001:db8:cafe:1002::3

#Default route to HSRP address on WAN access routers
ipv6 route outside ::/0 fe80::5:73ff:fea0:2
ipv6 access-list v6-ALLOW permit icmp6 any any
ipv6 access-list v6-ALLOW permit tcp 2001:db8:cafe::/48 host Br1-v6-Server object-
group RDP
failover
failover lan unit primary
failover lan interface FO-LINK GigabitEthernet0/3
failover interface ip FO-LINK 2001:db8:cafe:1001::1/64 standby
2001:db8:cafe:1001::2
access-group v6-ALLOW in interface outside
ssh 2001:db8:cafe::/48 inside
```

例 8-6 所示为故障切换接口（G0/3）配置的总结性输出。

例 8-6 ASA-1 show failover interface 命令输出

```
asa-1# show failover interface
        interface FO-LINK GigabitEthernet0/3
                System IP Address: 2001:db8:cafe:1001::1/64
                My IP Address     : 2001:db8:cafe:1001::1
                Other IP Address : 2001:db8:cafe:1001::2
```

例 8-7 所示为 Cisco ASA 的正常故障切换状态和配置的总结性输出。由输出可知，这台 ASA 为主用单元，并处于 active 状态。输出中还显示了 ASA 的 outside 和 inside 接口信息，ASA 对故障切换进行跟踪时，会用到接口的 IPv4 和 IPv6 地址信息。

例 8-7 ASA-1 show failover 命令输出

```
asa-1# show failover
Failover On
Failover unit Primary
Failover LAN Interface: FO-LINK GigabitEthernet0/3 (up)
Unit Poll frequency 1 seconds, holdtime 15 seconds
```

（待续）

```
Interface Poll frequency 5 seconds, holdtime 25 seconds
Interface Policy 1
Monitored Interfaces 2 of 160 maximum
Version: Ours 8.2(2), Mate 8.2(2)
Last Failover at: 05:15:12 UTC Apr 12 2010
        This host: Primary - Active
                Active time: 48 (sec)
                slot 0: ASA5520 hw/sw rev (2.0/8.2(2)) status (Up Sys)
                  Interface outside (10.124.1.4/fe80::21e:7aff:fe81:8e2c): Normal
                  Interface inside (10.124.3.1/fe80::21e:7aff:fe81:8e2d): Normal
                slot 1: ASA-SSM-4GE hw/sw rev (1.0/1.0(0)10) status (Up)
        Other host: Secondary - Standby Ready
                Active time: 261 (sec)
                slot 0: ASA5520 hw/sw rev (2.0/8.2(2)) status (Up Sys)
                  Interface outside (10.124.1.5/fe80::21d:a2ff:fe59:5fe4): Normal
                  Interface inside (10.124.3.2/fe80::21d:a2ff:fe59:5fe5): Normal
                slot 1: ASA-SSM-4GE hw/sw rev (1.0/1.0(0)10)
status (Up)
```

例 8-8 所示为防火墙检测到的连接状态。防火墙检测到了分属 outside 和 inside 安全域两台主机间的 Telnet 会话连接（TCP 端口号 23）。

例 8-8　防火墙检测到的连接状态

```
asa-1# show conn
6 in use, 13 most used
TCP outside 2001:db8:cafe:1000::2:23 inside
2001:db8:cafe:1004:c53c:2d6a:ccef:f2c5:1044, idle 0:02:49, bytes 115, flags UIO
```

EtherSwitch 模块的配置

对 HBE 来说，EtherSwitch 模块属于可选部件，用来替代传统的 Catalyst 交换机。本章给出其配置的目的，是要让读者知道，该模块的配置方法与配置 Catalyst 3560/3750 交换机几乎完全相同。本例所使用的 EtherSwitch 模块的部件编号为 NME-16ES-1G。

在 HBE 中，EtherSwitch 模块用来连接分支机构 LAN 接入路由器和两台 ASA 防火墙。激活 EtherSwitch 模块的 IPv6 特性和功能之前，需配置交换机数据库管理模板（SDM），以令其同时支持 IPv4 和 IPv6。支持 IPv4 和 IPv6 的三个 SDM 模板是：

- Dual IPv4 and IPv6 default template（双 IPv4 和 IPv6 默认模板）；
- Dual IPv4 and IPv6 routing template（双 IPv4 和 IPv6 路由模板）；

- Dual IPv4 and IPv6 VLAN template（双 IPv4 和 IPv6 VLAN 模板）。

可在全局配置模式下，通过下面这条命令来定义 Dual IPv4 and IPv6 default template。

```
BR1-EtherSwitch(config)#sdm prefer dual-ipv4-and-ipv6 {default | routing | vlan}
```

配毕之后，需重启设备，命令才能生效。重启 EtherSwitch 模块之后，可执行 **show sdm prefer** 命令（见例 8-9），来验证当前生效的 SDM 模板正确与否。

例 8-9 EtherSwitch 模块 show sdm prefer 命令的输出

```
BR1-EtherSwitch# show sdm prefer
 The current template is "desktop IPv4 and IPv6 default" template.
 The selected template optimizes the resources in
 the switch to support this level of features for
 8 routed interfaces and 1024 VLANs.

  number of unicast mac addresses:                  2K
  number of IPv4 IGMP groups + multicast routes:    1K
  number of IPv4 unicast routes:                    3K
  number of directly-connected IPv4 hosts:          2K
  number of indirect IPv4 routes:                   1K
  number of IPv6 multicast groups:                  1.125k
  number of directly-connected IPv6 addresses:      2K
  number of indirect IPv6 unicast routes:           1K
  number of IPv4 policy based routing aces:         0
  number of IPv4/MAC qos aces:                      0.5K
  number of IPv4/MAC security aces:                 1K
  number of IPv6 policy based routing aces:         0
  number of IPv6 qos aces:                          0.625k
  number of IPv6 security aces:                     0.5K
```

更多与配置 SDM 有关的信息请见以下 URL：

http://www.cisco.com/en/US/docs/switches/lan/catalyst3560/software/release/12.2_25_see/configuration/guide/swsdm.html#wp1077854。

EtherSwitch 模块的 IPv6 配置部分相当简单。在 HBE 中，只会用到该模块上的三个接口。其中一个是 EtherSwitch-to-router（EtherSwitch 模块和路由器之间的内连接口）接口（GigabitEthernet 1/0/2），另外两个以太网接口用来连接两台 Cisco ASA 防火墙。

写作本书之际，Cisco ASA 还不支持用于 IPv6 的动态路由协议，故而只能

在 EtherSwitch 模块上设置默认路由，并将其下一跳指向 Cisco ASA 的故障切换 IP 地址。作为一种选择，可在该模块上启用 EIGRPv6，将默认路由通告给分支机构内部的 BR1-LAN 路由器，并同时使得该路由器将学得的所有内部路由通告给 EtherSwitch 模块。当然，在 BR1-LAN 路由器和 EtherSwitch 模块上分别设置相应的静态路由，亦可起到相同的效果。例 8-10 所示为 EtherSwitch 模块的配置。

例 8-10 EtherSwitch 模块的配置

```
ipv6 unicast-routing
!
interface FastEthernet1/0/1
 description TO ASA-1
 switchport access vlan 101
!
interface FastEthernet1/0/2
 description TO ASA-2
 switchport access vlan 101
!
interface GigabitEthernet1/0/2     #Interface connecting to
                                     #branch LAN access
                                     #router (EtherSwitch internal
                                     #interface)
 description to BR1-LAN
 no switchport
 ip address 10.124.4.2 255.255.255.0
 ipv6 address 2001:DB8:CAFE:1003::2/64
 ipv6 eigrp 10                      #Optional - dynamic routing
                                     #for IPv6 inside the branch
!
interface Vlan101
 ip address 10.124.3.3 255.255.255.0
 ipv6 address 2001:DB8:CAFE:1002::3/64    #VLAN for network
                                             #connecting ASA
 ipv6 eigrp 10
!
ipv6 route ::/0 2001:DB8:CAFE:1002::1     #Default route pointing
                                             #to ASA
ipv6 router eigrp 10                 #Enable EIGRP for IPv6
 redistribute static                 #Redistribute default route
                                       #to LAN router
 passive-interface Vlan101           #Do not attempt adjacency on
                                       #VLAN101
```

分支机构 LAN 路由器的配置

BR1-LAN 分支机构 LAN 接入路由器（配置请见例 8-11）担当分支机构网络的 L3 分布层设备。该路由器终结了配置在 L2 接入层交换机（BR1-LAN-SW）上的 VLAN trunk，BR1-LAN-SW 交换机用来连接分支机构网络中的个人主机。除了行使基本的 L3 终结和路由选择功能以外，BR1-LAN 路由器还能通过无状态的 DHCPv6 (RFC 3736)，提供地址分配服务，此外，该设备还担当了状态化的 DHCPv6 中继代理路由器。利用无状态的 DHCPv6，BR1-LAN 路由器便能通过 SLAAC (RFC 4862)，来行使地址分配服务，但其他与地址有关的信息，比如，DNS 名和 DNS 服务器信息，则通过有状态的 DHCPv6 地址池提供（例 8-11 中接口 G0/0.104 的配置）。启用了状态化 DHCPv6 中继特性之后，该路由器便能够将 DHCP 请求消息转发给预先定义的 DHCPv6 服务器（例 8-11 中接口 G0/0.105 的配置）。

例 8-11 BR1-LAN 路由器配置示例

```
ipv6 unicast-routing
ipv6 cef
!
ipv6 dhcp pool DATA_W7                    #DHCPv6 pool name
 dns-server 2001:DB8:CAFE:102::8         #Primary IPv6 DNS server
 domain-name cisco.com                    #DNS domain name passed
                                            #to client
!
interface GigabitEthernet0/0
 description to BR1-LAN-SW
 no ip address
 duplex auto
 speed auto
!
interface GigabitEthernet0/0.104
 description VLAN-PC
 encapsulation dot1Q 104
 ip address 10.124.104.1 255.255.255.0
 ipv6 address 2001:DB8:CAFE:1004::1/64       #Client uses SLAAC
                                                #with this prefix
 ipv6 nd other-config-flag              #Set flag in RA to instruct
                                          #host how to obtain "other"
                                          #information such as domain
 ipv6 dhcp server DATA_W7               #Use DHCP pool above for
                                          #options
```

（待续）

```
 ipv6 eigrp 10
!
interface GigabitEthernet0/0.105
 description VLAN-PHONE
 encapsulation dot1Q 105
 ip address 10.124.105.1 255.255.255.0
 ipv6 address 2001:DB8:CAFE:1005::1/64
 ipv6 nd prefix 2001:DB8:CAFE:1005::/64 0 0 no-autoconfig #Do
                                                  #not use prefix for
                                                  #autoconfiguration
ipv6 nd managed-config-flag       #Set flag in RA to instruct
                                    #host to use DHCPv6
 ipv6 dhcp relay destination 2001:DB8:CAFE:102::9     #Relay for
                                                      #DHCPv6 server
 ipv6 eigrp 10
 interface GigabitEthernet0/0.106
 description VLAN-PRINTER
 encapsulation dot1Q 106
 ip address 10.124.106.1 255.255.255.0
 ipv6 address 2001:DB8:CAFE:1006::1/64
 ipv6 eigrp 10
!
interface GigabitEthernet1/0
 description TO ETHERSWITCH MODULE
 ip address 10.124.4.1 255.255.255.0
 ipv6 address 2001:DB8:CAFE:1003::1/64
 ipv6 eigrp 10
!
ipv6 router eigrp 10
 no shutdown
```

BR1-LAN-SW Catalyst 交换机的配置包括将连接到 BR1-LAN 路由器的接口设置为 IEEE 802.1Q trunk，并令该 Trunk 接口放行 VLAN 104～106 的流量通过。除了配置管理接口的 IPv4 和 IPv6 地址，以令网管人员能够通过 IPv4/IPv6 管理到该交换机以外，再无与 IPv6 有关的配置了。因此，此处不再给出 BR1-LAN-SW 交换机的配置。

8.4 基于纯 IPv6 部署 WAN/分支机构网络

写作本书之际，还几乎没有任何企业能将自己的网络建设成为端到端，

完全通过 IPv6 来建立连通性的网络，即从分支机构站点到总部 WAN 前缀区域也使用纯 IPv6 网络建立连通性。随着越来越多的 SP 为其客户提供 IPv6 服务，企业也可以利用 IPv6 作为站点间传输加密 IPv6 流量的一种手段，并能够弃 IPv6-in-IPv4 加密隧道部署而不用，这些内容在本章的前几节中已做过讨论。

IOS 路由器能够支持 IPv6 上的 IPSEC 部署。接下来将给出如何在两台 Cisco IOS 路由器之间，部署 IPv6 上的 IPSEC 的基本配置。

图 8-6 所示为两台路由器通过 IPv6 Internet 互连的网络拓扑。在本例中，尽管两台路由器可通过开启双栈（IPv4 和 IPv6）机制接入 Internet，但实际上两者只开启了 IPv6 功能，并通过 SP 的 IPv6 设备接入 Internet。

本例的配置不但非常简单，而且与 IPv4 上的点到点 IPSEC 配置极为相近，配置间的差异大多与路由器接口所配的 IP 地址有关。

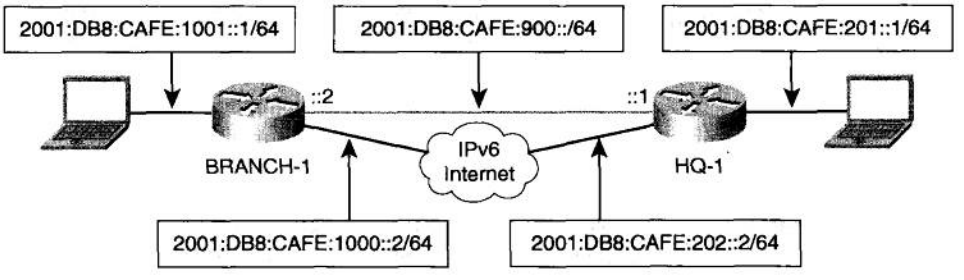


图 8-6 基于 IPv6 Internet 的 IPSEC VPN

例 8-12 所示为 HQ-1 路由器的基本配置。其中的 ISAKMP 及 IPSEC 策略信息的配置均与先前 HBE 中所采用的配置相同，只是与隧道有关的配置有所不同。隧道的源和目的地址如今为 IPv6，而非先前的 IPv4 地址了。此外，隧道模式也变为了 IPv6 上的 IPSEC 传输。最后，还要在路由器的串行接口上配置 IPv6 地址，与 ISP 的 IPv6 设备建立连通性。在 HQ-1 路由器上，激活了单播逆向路径转发（uRPF）特性，以遏制地址欺骗攻击。如果用在生产网络，还需在串行接口上设置 ACL，只放行分支机构和总部间特定协议及源/目的地址的流量。

例 8-12 HQ-1 路由器的配置

```
ipv6 unicast-routing
ipv6 cef
!
crypto isakmp policy 1
 encr aes 256
 authentication pre-share
 group 2
crypto isakmp key CISCO address ipv6 ::/0
!
crypto ipsec transform-set HUB esp-aes 256 esp-sha-hmac
!
crypto ipsec profile HUB
 set transform-set HUB
!
interface Tunnel2
 no ip address
 ipv6 address 2001:DB8:CAFE:900::1/64
 ipv6 eigrp 10
 tunnel source 2001:DB8:CAFE:202::2    #Source is now using IPv6
 tunnel mode ipsec ipv6              #IPSec over IPv6 tunnel mode
 tunnel destination 2001:DB8:CAFE:1000::2  #Dest. now using IPv6
 tunnel protection ipsec profile HUB
 !
interface GigabitEthernet1/0
 description LAN
 no ip address
 ipv6 address 2001:DB8:CAFE:201::1/64
 ipv6 eigrp 10
!
interface Serial2/0
 description to ISP
 no ip address
 ipv6 address 2001:DB8:CAFE:202::2/64     #v6 connection to ISP
 ipv6 verify unicast reverse-path              #uRPF for IPv6
!
ipv6 route ::/0 2001:DB8:CAFE:202::1           #Default to ISP
ipv6 router eigrp 10
 eigrp router-id 1.1.1.2
```

例 8-13 所示为 BRANCH-1 路由器的配置。除了 IP 地址和 IPsec profile 名

以外，该配置与 HQ-1 路由器的配置几乎相同。

例 8-13 BRANCH-1 路由器的配置

```
ipv6 unicast-routing
ipv6 cef
!
crypto isakmp policy 1
 encr aes 256
 authentication pre-share
 group 2
crypto isakmp key CISCO address ipv6 ::/0
!
crypto ipsec transform-set SPOKE esp-aes 256 esp-sha-hmac
!
crypto ipsec profile SPOKE
 set transform-set SPOKE
!
interface Tunnel2
 no ip address
 ipv6 address 2001:DB8:CAFE:900::2/64
 ipv6 eigrp 10
 tunnel source 2001:DB8:CAFE:1000::2
 tunnel mode ipsec ipv6
 tunnel destination 2001:DB8:CAFE:202::2
 tunnel protection ipsec profile SPOKE
!
interface GigabitEthernet1/0
 description LAN
 no ip address
 ipv6 address 2001:DB8:CAFE:1001::1/64
 ipv6 eigrp 10
!
interface Serial2/0
 description to ISP
 no ip address
 ipv6 address 2001:DB8:CAFE:1000::2/64
 ipv6 verify unicast reverse-path
!
ipv6 route ::/0 2001:DB8:CAFE:1000::1
ipv6 router eigrp 10
 eigrp router-id 1.1.1.3
```

例 8-14 所示为 ISAKMP 对等体和安全关联（SA）的状态。

例 8-14 HQ-1 路由器上 ISAKMP 对等体和 SA 的输出

```
HQ-1# show crypto isakmp peers
Peer: 2001:DB8:CAFE:1000::2 Port: 500 Local: 2001:DB8:CAFE:202::2
 Phase1 id: 2001:DB8:CAFE:1000::2

HQ-1# show crypto isakmp sa
IPv4 Crypto ISAKMP SA
dst              src              state          conn-id status

IPv6 Crypto ISAKMP SA

 dst: 2001:DB8:CAFE:1000::2
 src: 2001:DB8:CAFE:202::2
 state: QM_IDLE          conn-id:   1002 status: ACTIVE

 dst: 2001:DB8:CAFE:202::2
 src: 2001:DB8:CAFE:1000::2
 state: QM_IDLE          conn-id:   1003 status: ACTIVE
```

8.5 总结

本章描述了在分支机构网络中部署 IPv6 的方法。可将分支机构网络部署模型归纳为：单层、双层和多层三种。本章所举的配置示例基于混合型分支机构网络部署示例，该示例集三种分支机构网络部署模型的要素于一身。除了介绍 IPv4 上的 IPv6 IPSec VPN 解决方案以外，本章还专门针对具备端到端 IPv6 传输能力的企业客户（需 SP 支持），给出了纯 IPv6 IPSec 配置示例。在分支机构网络环境中部署 IPv6 的方法多种多样，并不局限于本章所介绍的那三种套路（模型），读者可根据自己的实际网络环境来选择合适的部署方式。

8.6 参考资料

对 IPv6 技术及协议的充分理解，是读懂本章内容的前提条件。实施 IPv6 时，需要对诸多设计要素加以考虑，其中包括安全、QoS、高可用性、管理、IT 培训和应用程序的支持等。

以下列出了众多介绍 IPv6 详细信息的参考资料中的部分，其中包括 Cisco 推荐的设计方案、产品和解决方案以及业界活动等。

Popoviciu, Ciprian P., Eric Levy-Abegnoli, and Patrick Grossetete. *Deploying IPv6 Networks*. Cisco Press. (ISBN10: 1-58705-210-5; ISBN13: 978-1-58705-210-1).

Hogg, Scott and Eric Vyncke. *IPv6 Security*. Cisco Press. (ISBN10: 1-58705-594-5; ISBN13: 978-1-58705-594-2).

Szigeti, Tim and Christina Hattingh. *End-to-END QoS Network Design*. Cisco Press. (ISBN10: 1-58705-176-1; ISBN13: 978-1-58705-176-0).

Cisco. Cisco IOS IPv6 Configuration Guide:
http://www.cisco.com/en/US/docs/ios/ipv6/configuration/guide/15_0/ipv6_15_0_book.html.

Cisco. Design Zone for Branch:
http://www.cisco.com/en/US/netsol/ns816/networking_solutions_program_home.html.

Cisco. Deploying IPv6 in Unified Communications Networks with Cisco Unified Communications Manager:
http://www.cisco.com/en/US/docs/voice_ip_comm/cucm/srnd/ipv6/ipv6srnd.html.

Cisco. IPsec Design Guides:
http://www.cisco.com/en/US/tech/tk583/tk372/tech_design_guides_list.html.

Cisco. Cisco IOS Control Plane Policing:
http://www.cisco.com/en/US/docs/ios/12_3t/12_3t4/feature/guide/gtrtlimt.html.

Cisco. Enterprise QoS Solution Reference Network Design Guide:
http://www.cisco.com/en/US/docs/solutions/Enterprise/WAN_and_MAN/QoS_SRND/QoS-SRND-Book.html.

Cisco. Cisco IOS IPv6 Multicast Technologies:
http://www.cisco.com/en/US/technologies/tk648/tk828/tk363/technologies_white_paper0900aecd8014d6dd.html.

Cisco. Cisco IOS IPv6 Multicast Configuration:
http://www.cisco.com/en/US/tech/tk828/technologies_white_paper09186a0080203f7a.shtml.

Cisco. Cisco Implementing IPv6 Multicast:
http://www.cisco.com/en/US/docs/ios/ipv6/configuration/guide/ip6-multicast_ps10591_TSD_Products_Configuration_Guide_Chapter.html.

Cisco. Defining and Using IPv6 General Prefixes:
http://www.cisco.com/en/US/docs/ios/ipv6/configuration/guide/ip6-addrg_bsc_con_ps10591_TSD_Products_Configuration_Guide_Chapter

.html#wp1132473.

Cisco. Network Management and Automation:
http://www.cisco.com/en/US/products/sw/netmgtsw/index.html.

Cisco. Dynamic Multipoint VPN (DMVPN) Design Guide:
http://www.cisco.com/en/US/docs/solutions/Enterprise/WAN_and_MAN/DMVPDG.html.

Cisco. Cisco IOS Release 15.0 - Implementing Dynamic Multipoint VPN for IPv6:
http://www.cisco.com/en/US/docs/ios/ipv6/configuration/guide/
ip6-dmvpn_ps10591_TSD_Products_Configuration_Guide_Chapter.html.

Cisco. Shortcut Switching Enhancements for NHRP in DMVPN Networks:
http://www.cisco.com/en/US/docs/ios/ipaddr/configuration/guide/iad_nhrp_dmvpn.html
#wp1072593.

Cisco. Configuring NHRP:
http://cisco.com/en/US/docs/ios/ipaddr/configuration/guide/iad_cfg_nhrp.html#wp1078
234.

Cisco. Catalyst 3560 Switch Configuration Guide: Configuring SDM Templates:
http://www.cisco.com/en/US/docs/switches/lan/catalyst3560/software/release/12.2_25_
see/configuration/guide/swsdm.html#wp1077854.

Cisco. Branch Routers (including ISR):
http://www.cisco.com/en/US/products/ps10906/Products_Sub_Category_Home.html.

Savola, P. RFC 3627, "Use of /127 Prefix Length Between Routers Considered Harmful."
http://www.ietf.org/rfc/rfc3627.txt.

Hinden, R. and S. Deering. RFC 3513, "Internet Protocol Version 6 (IPv6) Addressing
Architecture." http://www.ietf.org/rfc/rfc3513.txt.

Savola, P. and B. Haberman. RFC 3956, "Embedding the Rendezvous Point (RP) Address
in an IPv6 Multicast Address." http://www.ietf.org/rfc/rfc3956.txt.

Deering, S. and R. Hinden. RFC 2460, "Internet Protocol, Version 6 (IPv6) Specification."
http://www.ietf.org/rfc/rfc2460.txt.

Thomson, S., T. Narten, and T. Jinmei. RFC 4862, "IPv6 Stateless Address
Autoconfiguration." http://www.ietf.org/rfc/rfc4862.txt.

Droms, R. RFC 3736, "Stateless Dynamic Host Configuration Protocol (DHCP) Service
for IPv6." http://www.ietf.org/rfc/rfc3736.txt.

McCann, J., S. Deering, and J. Mogul. RFC 1981, "Path MTU Discovery for IP version 6."
http://www.ietf.org/rfc/rfc1981.txt.

Arkko, J. (Ed.), J. Kempf, and P. Nikander. RFC 3971, "SEcure Neighbor Discovery
(SEND)." http://www.ietf.org/rfc/rfc3971.txt.

Templin, F., T. Gleeson, M. Talwar, and D. Thaler. RFC 4214, "Intra-Site Automatic Tunnel Addressing Protocol (ISATAP)." http://www.ietf.org/rfc/rfc4214.txt.

Van de Velde, G., T. Chown, O. Bonness, and C. Hahn. RFC 5375, "IPv6 Unicast Address Assignment Considerations." http://www.ietf.org/rfc/rfc5375.txt.

CHAPTER 9

第 9 章　数据中心网络中的 IPv6 部署

本章涵盖以下主题。

- **设计和实施运行双协议栈的数据中心网络**：本节会介绍在数据中心网络的接入层、汇聚层以及核心层内部署 IPv6 的原则。
- **在采用了虚拟化技术的数据中心网络中实施 IPv6**：本节将描述在采用了虚拟化技术的数据中心网络中实施 IPv6 的原则。
- **在 SAN 网络中实施 IPv6**：本节将介绍如何在存储区域网络中实施 IPv6。
- **数据中心互连场景中的 IPv6 设计**：本节讲解多数据中心互连场景中的 IPv6 设计方案。

对企业网来说，数据中心网络环境中的 IPv6 部署工作既不受人重视，也常被人低估。与企业园区网络相比，在某些方面，数据中心网络与其有相似之处，例如，两者都涉及交换和路由，但相同点也就这么多了。对于数据中心网络中的 IPv6 部署，许多企业既未下大力气研究，也未花时间评估，一旦部署开来，就会经常被受 IPv6 部署波及的产品和技术弄得手足无措。

除了传统的路由和交换功能以外，数据中心网络还需具备以下功能。

- 可利用（安装在交换机中的）服务模块或独立的硬件，来提供联网、安全以及性能方面的服务[①]。
- 存储区域网络（SAN），比如，附接到存储系统和交换机的 Fibre Channel、Ethernet 以及 FcoE（以太网上的 Fibre Channe）。
- 数据中心虚拟化，包括服务器、网络以及存储虚拟化等。
- 设备、操作系统以及应用管理，包括：

① 原文是 "Network, security, and performanceservices in the form of service modules and/or appliances"。作者居然写出这种文字，无语。

- 服务器带外管理（也称为lights-out管理[LoM]），比如，Hewlett-Packard Integrated Lights Out (iLO)或Dell Remote Assistance Card (DRAC)；
- CiscoWorks LAN管理解决方案、Microsoft系统中心，以及用来管理设备、OS以及应用程序的其他管理系统。

- 设备和站点的灾难恢复和高可用性，包括：
  - 数据中心互连（DCI），即利用各种传输技术互连两个或多个数据中心网络，其目的是要为诸如集群和实时的虚拟机迁移（比如，VMware vMotion和Microsoft Hyper-V LiveMigration等）之类的网络服务和应用建立起2、3层的邻接关系；
  - 全局和站点的负载均衡；
  - 网络和安全设备（比如，防火墙等）的无状态/状态化故障切换。
- 服务器农场，用来为以下各种各样的应用程序提供服务：
  - 关键性应用（比如，SAP，Oracle，其他在线的商业应用）；
  - 虚拟桌面基础设施（VDI）；
  - 消息传递及协作（比如，Microsoft Exchange）；
  - 文件和打印机；
  - DNS和DHCP；
  - 专业性的网格计算应用。

本章只会对上面提及的与IPv6部署有关的某些要点展开深入讨论。在路由和交换领域，各厂商对IPv6的支持已日臻完善，但在其他领域（比如，网络服务、关键性应用，以及任意一种操作在第三层以上的产品和技术）则反差很大，各厂商对IPv6的支持也参差不齐。

就企业网内部所属的各个区块而言，数据中心的情况最为复杂，技术和产品的差异性也最大，故而建立和维护一张差异清单是重中之重，要清楚哪些支持IPv6，哪些不支持，并要和厂商协作来缩小这些差距。

9.1 设计和实施双栈数据中心网络

本章将重点讨论数据中心网络区块的三个层次（接入层、汇聚层以及核心层），

并会探究数据中心网络的设计原则，亦会给出相关实施案例（在网络的三个层次全都启用双协议栈）。本节所提及的众多概念、产品、特性和配置亦适用于企业内/外部数据中心的网络建设，其中，外部数据中心是指面向 Internet 的数据中心网络。但有关与服务提供商对等、与 ISP 互连链路的配置、多宿主以及 Internet 数据中心高可用性等方面的概念，却不在本章的探讨范围之内。Cisco、其他设备厂商以及服务提供商客户也在不断地钻研、测试和验证上面提到的多项设计要素[①]。

注意

阅读本章，读者可见识到来自 Cisco 和其他厂商的种类繁多的技术和产品。本章的目的不在于介绍技术和产品本身，以及它们的运作方式，而是意在说明它们在启用了 IPv6 网络中的部署方式。其中的某些产品和技术不但功能强大，而且使用灵活。因此，所涉及的设计和实施方案也多种多样，这意味着本章所介绍的设计和实施方案并非一成不变，在实战中，读者可根据自身的网络环境，自行定制方案。

数据中心网络区块分为三层，即接入层、汇聚层和核心层。在数据中心内部，还存在与存储和应用相关的其他“层次”，但本节所讨论的层次只关乎网络。

图 9-1 所示为前面提到的数据中心网络分层示意图。图中的每一层都包括了某些常见的 Cisco 产品。当然，也可以选用其他的产品和技术来搭建这样的网络，本图只是用作演示。

部署在图 9-1 中接入层的设备类型可谓多种多样，这种情况常见诸于许多数据中心，其原因是这些数据中心的交换设备型号（种类）会随用途而异。某些服务器可能会通过 10/100/1000M（甚至是 10G 以太网）链路，连接到接入层 Catalyst 交换机（比如，Catalyst 6500、4900M/4948E）和/或 Nexus 5000/2000 交换机。在数据中心的其他区域，可能还引入了 Cisco 统一计算系统（UCS）（其中包括 Fabric 互连[FI]和刀片机箱[blade chassis]）用来运行服务器虚拟化解决方案（比如，VMware ESX）。对于这样的网络环境，则可以部署含虚拟以太网模块（VEM）（运行在每台 VMware ESX 主机上）和 Virtual Supervisor 模块（VSM）（用来管理 VEM）的 Nexus 1000 系列交换机。一旦将虚拟化服务

① 原文是“Many of those design elements are still being worked out, tested, and validated by Cisco and enterprise and service provider customers”。作者的企业荣誉感还挺强，难道“Cisco”不算“enterprise”吗？如果作者真的那么“爱惜羽毛”的话，为什么不在写作上下点工夫呢？

器和网络部件合二为一，通常便很难界定“真正的”接入层。很多时候，人们都会把Nexus 1000 VEM或UCS Fabric互连（FI）视为接入层设备。不管部署在哪一层，网络设备都具有共同的特征，即都能把主机连接到网络，这也是接入层设备的基本功能。

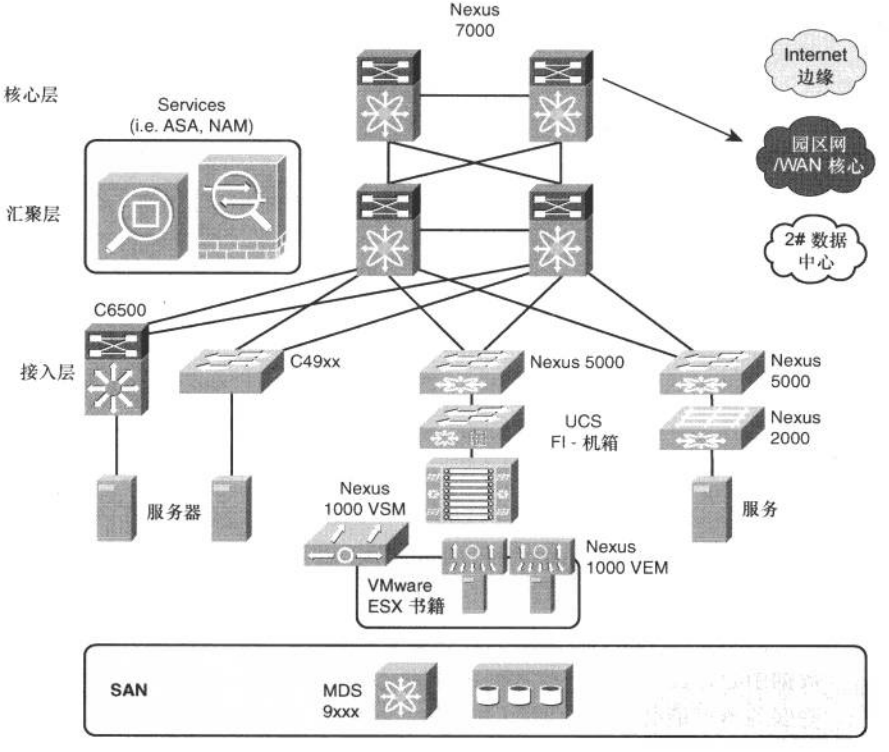


图 9-1 数据中心网络拓扑示意图

在数据中心网络中，汇聚层交换机是终结或“汇接”所有接入层交换机上行链路的设备。除了终结接入层交换机上行链路以外，网络和应用服务往往也“坐落”于汇聚层交换机，网络和应用服务是指安全服务（比如，Cisco ASA或IDS/IPS所提供的服务）、网络监控（比如，Cisco NAM）服务、服务器负载均衡服务、网管服务（连接网管系统）等。汇聚/服务层设备既可以选用物理交换机，也可以选用安装在交换机机箱内的服务模块，但最为常见的汇聚/服务层设备不是“初出茅庐”的Nexus 7000交换机，就是“历史悠久”的Catalyst 6500

交换机，这当然是为了充分利用这两种产品的稳定性和虚拟化功能[①]。

核心层交换机的作用是将数据中心网络区块连接到企业网络的其他区块（即 Internet 边缘区块、园区网络区块以及 WAN 核心区块等；当然，也可以通过各种数据中心互连技术，来连接多个数据中心网络。）

以下各节将讨论如何在数据中心网络的每一层部署 IPv6，意在使用简单的网络拓扑介绍最基本的概念，而非特定的产品组合。本章所涉及的 Cisco 数据中心交换产品配置起来大多（即便不是全部）相同，因此只会给出与产品无关但涉及 IPv6 的配置。

9.1.1 数据中心接入层

如前所述，以物理或虚拟的方式将主机接入网络，是接入层设备的主要功能之一。多种类型的主机和许多 Cisco 产品都能够以物理或虚拟的方式将主机连接到网络[②]。在数据中心网络接入层部署 IPv6 时，需考虑诸多事宜，其中的绝大多数都与前面提到的园区网络区块接入层相同。在数据中心网络接入层部署 IPv6 时应考虑如下事宜。

- **IPv6 多播**：若连接到接入层交换机的主机需接收 IPv6 多播，接入层交换机便需具备 MLDV1 或 V2 欺骗功能。如此一来，只有那些需要接收多播数据包的交换机端口才需要具备以硬件的方式转发第二层多播数据包的功能[③]。
- **IPv6 QoS**：在数据中心网络的接入层，可对（可不对）IPv6 数据包执行 QoS 分类/标记处理。某些企业可能会选择在接入层、汇聚层或甚至在网络内部，去执行 QoS 的分类/标记操作。若确需在接入层执行 QoS 操作，则应对 IPv6 数据包执行分类和标记处理，即让接入层交换机信任或改写由主机标记的数据包报头中的相关 QoS 字段[④]。

---

① 原文是“This aggregation/services layer can use appliances or service modules and most often leverages the large-scale and virtualization capabilities of the Nexus 7000 and the long-standing Catalyst 6500”。原文和译文的字面意思差别很大，请读者注意。

② 原文是“Many types of hosts and many Cisco products can connect these hosts to the network physically and virtually”。原文如此，译文也只好这么译。

③ 原文是“This enables the Layer 2 hardware-supported constraint of multicast to only those ports that need to receive the multicast packets。”实在看不透作者到底想表达什么，译文为译者杜撰。

④ 译文是“If it is required in the access layer, IPv6 packets should be classified and marked, allow re-marking of the packets, and trust markings that took place on the host”。这段文字放在这里不但讲不通，而且也无必要，译文酌改。

- **IPv6 的安全性**：可在接入层应用多种与安全相关的特性和技术，包括：ACL、控制平面监管、第一跳安全特性（比如，无赖 RA 防护特性等），以及其他与网络层协议无关的特性（比如，私有 VLAN 特性）等。
- **IPv6 管理**：通过 IPv6 来执行设备和系统级别的网络管理十分重要。比较常见的网络管理协议包括：SNMP、SSH/Telnet、HTTP，以及其他运行在设备管理级别的访问协议。对于利用上述网管协议执行网络/系统管理的程序或工具（“manager of manager” or systemwide management tools）来说，能够在启用了 IPv6 功能的端点上运行，并具备 IPv6 所独有的功能也同样重要。
- **性能**：以前，许多厂商（包括 Cisco 在内）都有这么一条不成文的规矩：所交付的某些网络设备在处理 IPv6 数据包时，其性能要打对折（即在性能方面，只能达到处理 IPv4 数据包时的一半）。在过去，这条规矩似乎还能讲得过去，但随着网络中 IPv6 数据流的激增，以及越来越多的用户开始使用支持 IPv6 的主机和应用程序，显而易见，对于下一代网络设备来说，在处理 IPv6 数据包时所表现出来的性能至少应不逊于其对 IPv4 数据包的处理。

诚然，取决于数据中心接入层的设计方案，还有许多其他因素也需考虑，但以上罗列的诸项则被认为是设计方案中的重点，在部署 IPv6 时，读者应对这些要素做仔细评估。

接入层设备的 IPv6 配置

一般而言，数据中心接入层交换机的 IPv6 配置简单而又琐碎。若未采用路由式的接入层设计，那便无需针对 IPv6 实施专门的配置。诸多适用于 IPv4 的配置可原封不动地照搬至 IPv6。

例 9-1 所示为数据中心接入层交换机的基本配置示例，其中的某些配置语句也适用于生产环境中的数据中心接入层设备。这份 Catalyst 6500 交换机的配置示例包含了通往汇聚层 Nexus 7000 交换机的 10G 上行链路的配置（这台 Catalyst 6500 交换机双上联至两台 Nexus 7000，但配置中只显示了一条上行链路的配置），以及下连主机端口的配置（这台 Catalyst 6500 交换机下连多台主机，配置中只显示了一条下连主机的链路配置）。由图 9-1 可知，在数据中心接入层，可部署各种型号的接入层交换机，具体的配置语法随交换机的型号而异，但基

本原理都是相同的。

例 9-1 数据中心接入层交换机的配置

```
vlan 14
 name WebSVR
!
interface TenGigabitEthernet2/1
 description to N7k-agg-1
 switchport
 switchport trunk encapsulation dot1q
 switchport trunk allowed vlan 10-14,30,31
 switchport mode trunk
 spanning-tree guard loop
 mls qos trust dscp
!
interface GigabitEthernet1/31
 description W2K8 Server
 switchport
 switchport access vlan 14
 switchport mode access
 spanning-tree portfast edge
 spanning-tree bpduguard enable
 mls qos trust dscp
 !
interface GigabitEthernet1/48
 description to OOB-NETWORK
 ipv6 address 2001:DB8:CAFE:100::14/64

!
ipv6 route ::/0 GigabitEthernet1/48 FE80::211:BCFF:FEC0:C800
```

在数据中心接入层，除了物理交换机以外，在利用 Hypervisor（[服务器虚拟化管理程序]，比如，VMware ESX、Microsoft Hyper-V、Xen 等）的服务器虚拟化的部署中，还会用上虚拟交换机。通常，人们会将此类基于软件的虚拟交换机视为“虚拟的接入层设备”，其中，虚拟机（VM）连接到主机内部的虚拟交换机。先进如 Cisco Nexus 1000v 这样的基于软件的虚拟交换机，其运作方式也酷似传统的物理交换机，转发 IPv6 帧自然也不在话下。此外，亦能通过 IPv6 对 Cisco Nexus 1000v 交换机进行管理，其配置如例 9-2 所示。

例 9-2 Cisco Nexus 1000v IPv6 管理配置

```
ip host Nexus1000v-1 172.16.100.18 2001:db8:cafe:100::18


vrf context management
  ip route 0.0.0.0/0 172.16.100.1

  ipv6 route 0::/0 fe80::0211:bcff:fec0:c800 mgmt0

interface mgmt0
  ip address 172.16.100.18/24
  ipv6 address 2001:0db8:cafe:0100::0018/64
```

网卡结对（NIC-Teaming）

在数据中心接入层部署 IPv6 时，还需考虑到使用网卡结对技术连接至接入层交换机的主机。网卡结对是指将一台服务器上的两块或多块网卡捆绑在一起，作为单块网卡来使用的技术。这不但能够提升服务器的网络吞吐量，而且还能增强物理网卡的高可用性。

如果每个提供网卡结对技术的厂商都有自己的一套对 IPv6 的支持方式，那么情况就会变得很糟。以下便是一个经常发生在启用了网卡结对功能的主机上的问题：网卡结对软件本身不支持 IPv6，但网管人员在创建网卡结对之前，便启用了 IPv6，或在创建了结对接口之后，手动为该接口分配了 IPv6 地址。现试举一个手工为 NIC 结对接口添加 IPv6 地址的示例，如例 9-3 所示。

例 9-3 在 Microsoft Windows 主机上，为 NIC 结对接口静态分配 IPv6 地址——网卡结对软件不支持 IPv6

```
netsh interface ipv6> add address "Local Area Connection" 2001:db8:cafe:10::7
Ok.
netsh interface ipv6>sh add
Querying active state...
Interface 10: Local Area Connection
Addr Type  DAD State  Valid Life   Pref. Life   Address
---------  ---------  -----------  -----------  ---------------------
Manual     Duplicate      infinite     infinite 2001:db8:cafe:10::7
```

例 9-3 所示为一台 Microsoft Windows Server 2008 主机的命令输出，在创建

了 NIC 结对接口之后，为该结对接口（Interface 10: “Local Area Connection”）手工配置静态 IPv6 地址。由于例中主机所运行的网卡结对软件不能识别 IPv6，因此输出并没有报错，但这会引发一个与 IPv6 有关的问题。当 Microsoft Windows TCP/IP 协议栈接受该 IPv6 地址，但却无法控制这一新创建的虚拟适配器（归网卡结对软件控制）时，问题便得以显现（其他操作系统可能也会有类似问题）——IPv6 地址被分配给了构成结对接口的物理网卡。由于构成结对接口的两块物理网卡同时获取了同一 IPv6 地址，故而会触发重复地址检测（DAD）过程。读者可在例 9-3 的输出中看见 DAD 的状态为“Duplicate（重复）”。这样一来，便破坏了 IPv6 的连通性。

综上所述，读者在启用了网卡结对功能的主机上配置 IPv6 地址时，应弄清网卡结对软件是否支持 IPv6，若否，则需更换新的网卡，或执行软件升级。例 9-4 所示为网卡结对软件支持 IPv6 时，Microsoft Windows 主机上与 NIC 适配器有关的输出。也可以通过图形界面的方式（即相关接口[网卡]的“属性”对话框），来获取相关信息。

例 9-4 在 Microsoft Windows 主机上，为 NIC 结对接口静态分配 IPv6 地址——网卡结对软件支持 IPv6

```
Ethernet adapter DC-ACCESS-1:                    #Physical Interfaces- Pre Team

   Connection-specific DNS Suffix  . :
   Autoconfiguration IP Address. . . : 169.254.25.192
   Subnet Mask . . . . . . . . . . . : 255.255.0.0
   IP Address. . . . . . . . . . . . : fe80::204:23ff:fec7:b0d7%12
   Default Gateway . . . . . . . . . : fe80::212:d9ff:fe92:de76%12

Ethernet adapter DC-ACCESS-2:
   Connection-specific DNS Suffix  . :
   IP Address. . . . . . . . . . . . : 172.16.10.20
   Subnet Mask . . . . . . . . . . . : 255.255.255.0
   IP Address. . . . . . . . . . . . : 2001:db8:cafe:10::20

   IP Address. . . . . . . . . . . . : fe80::204:23ff:fec7:b0d6%11

   Default Gateway . . . . . . . . . : fe80::212:d9ff:fe92:de76%11


Ethernet adapter TEAM-1:                        #Team Interface - Post Team
```

（待续）

```
   Connection-specific DNS Suffix  . :
   IP Address. . . . . . . . . . . . : 172.16.10.20
   Subnet Mask . . . . . . . . . . . : 255.255.255.0
   IP Address. . . . . . . . . . . . : 2001:db8:cafe:10::20
   IP Address. . . . . . . . . . . . : fe80::204:23ff:fec7:b0d6%13

   Default Gateway . . . . . . . . . : fe80::212:d9ff:fe92:de76%13


Interface 13: TEAM-1
Addr Type  DAD State  Valid Life   Pref. Life   Address
---------  ---------  -----------  -----------  ---------------------
Public     Preferred  4m11s        4m11s        2001:db8:cafe:10::20

Link       Preferred  infinite     infinite     fe80::204:23ff:fec7:b0d6
```

例9-4首先显示了创建NIC结对接口之前，两块物理网卡（DC-ACCESS-1和DC-ACCESS-2）的配置信息。接下来，又显示了担当NIC结对接口的TEAM-1接口的配置，该接口有一个IPv4地址和一个IPv6地址。若安装了完全支持IPv6的网卡结对软件，为NIC结对接口配置的IPv6地址会通过DAD检查，其状态也会被置为在用状态（“preferred”状态）。

9.1.2 数据中心汇聚层

数据中心汇聚层，顾名思义，是指各接入层交换机“汇合”之处，多台接入层交换机（的上行链路）在此处汇接于几台功能强劲的汇聚层交换机。汇聚层位居核心层和接入层之间，通过核心层，就能为驻留于接入层的主机提供外部（比如，Internet边缘区块、WAN以及园区网络区块等）连通性。除了汇接接入层交换机的物理上行链路，并以逻辑的方式终结VLAN以外，在数据中心汇聚层，还要执行并提供偏向于应用的负载均衡、安全性以及offload服务。这些服务包括：

- 防火墙服务；
- 深度包检测/入侵检测/入侵防护服务；
- 服务器负载均衡服务；
- SSL offload服务；
- 网络监控和分析服务。

既可使用独立的硬件也可使用（安装在交换机机箱内的）服务模块来提供上述服务。由于这些专业化的产品全都操作于网络层之上，因此相关产品必须全面支持 IPv6——不但要具备基本的 IPv6 寻址、路由和转发以及管理功能，而且还需在自己的“主项（主要功能）”上，从网络层到应用层，全方位地支持 IPv6。可是，包括 Cisco 在内的许多厂商并不能为自己的服务产品[①]及相关技术提供全方位的 IPv6 支持，外加不逊于 IPv4 的转发性能。读者应联系与自己有业务往来的各厂商，以获取其支持 IPv6 的产品列表以及对 IPv6 支持的详细信息。

以下各节将探讨在数据中心汇聚层部署 IPv6 及相关应用服务[②]时，诸多注意事项中的几则。

在汇聚层另辟蹊径传递 IPv6 流量[③]

在数据中心接入层，部署纯 IPv4 服务产品（即只能为 IPv4 流量提供防火墙、负载均衡等服务）的情况可谓是司空见惯。在此情形，即便部署于汇聚层的服务产品不具备 IPv6 功能，但有时，仍需让 IPv6 数据流通过汇聚层。

大多数情况下，在汇聚层，服务模块或独立的硬件[④]都可采用“单臂”模式来部署，以提供防火墙、负载均衡等服务。当然，上述设备以透明（桥接）模式或路由模式来部署也无不可。

图 9-2 逐一列出了上述三种部署模式的示意图，图中显示了 IPv6 和 IPv4 流量在各部署模式中的穿行方式。

图 9-2 所示为在数据中心汇网络聚层部署服务产品的三种模式：透明模式、单臂模式和路由模式，其中，所部署的服务产品既可以是集成进交换机机箱内的服务模块，也可是独立的硬件设备，但这些设备不是尚未启用就是不具备 IPv6 功能。图中，有一台双栈服务器连接到接入层交换机（图中并未显示），并与汇聚层交换机建立了连通性，而服务产品的部署方式不外有三：一，作为服务模块直接安装在汇聚层交换机内（比如，使用 Catalyst 6500 作为汇聚层交换机，并配备相应的服务模块）；二，专门配备一台用来提供各种服务的交换机机箱（该机箱用来安装所有的服务模块，然后再与汇聚层交换机相连）；三，配备用来提供各种服务的独立硬件设备（比如，ASA、PIX、IDS、负载均衡器等），然后再

① 即 FWSM、ASA、CSM 之类提供防火墙、负载均衡等服务的独立设备或交换机设备。

② 即上面提到的防火墙、LB、IPS/IDS 等服务。

③ 标题原文是“Bypassing IPv4-Only Services at the Aggregation Layer”，译者认为文不对题，酌改。

④ ASA、FWSM、CSM 甚至是 F5 负载均衡器之类的设备。

将其与汇聚层交换机相连。[①]

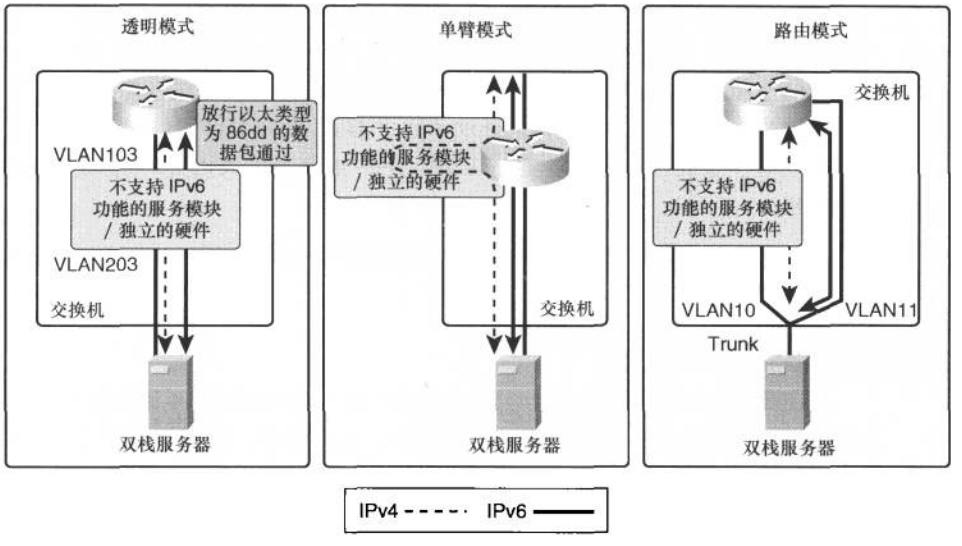


图9-2 不支持IPv6功能的服务产品——三种部署模式[②]

无论采用那种模式，IPv6流量都可获准穿越汇聚层交换机以及上面所提及的服务产品，只是无法“享受到”相关服务产品（集成进交换机机箱内的服务模块或独立的硬件设备）所提供的流量检测、负载均衡以及其他任何4～7层的服务（即不对IPv6流量执行流量检测、负载均衡等动作）。以下列出了在每种部署模式中所要关注的细节。

- **透明模式**：服务产品以桥接的方式转发IPv4流量，并为其提供相关服务（即针对IPv4流量执行流量检测、负载均衡以及流量过滤等动作）；但只会以桥接的方式转发IPv6流量，不会为其提供服务（只是以桥接的方式放行IPv6流量通过，流量检测、负载均衡以及流量过滤功能对

① 译者对整段原文做了大幅度修改，为了不影响读者理解，现给出整段原文“Figure 9-2 shows transparent, one-arm, and routed mode examples where an integrated services module or, alternatively, an appliance is deployed, but either does not have IPv6 enabled or cannot support it yet. A dual-stack server is connected to the access layer (not shown) and has connectivity to the aggregation layer switch, where a service module is deployed directly on the aggregation switch (through a Catalyst 6500 with service module), a services chassis design (all modules are located in external services switches that are physically attached to the aggregation layer switch), or an appliance directly attached to the aggregation layer switch”。

② 原文是“Non-IPv6-Capable Service Products - Options”。译者认为文不对图，作者的表达能力差到了极点，译文为直译。

IPv6 流量不生效）。一般情况下，会在服务产品上配置 Ethertype ACL，去放行以太类型为 0x86DD（IPv6 数据包的以太类型）的以太网帧通过。如此一来，所有匹配 Ethertype ACL 的 IPv6 流量都会被服务产品以桥接的方式放行（转发），而不会针对这些流量做任何检测和处理。

- **单臂模式**：要让 IPv6 流量顺利穿越服务层设备（即上文提及的那些不支持 IPv6 功能的服务产品），单臂模式部署起来最为简单。在单臂部署模式中，会通过 PBR（策略路由）或其他基于策略的机制，将流量转发至特定的服务产品，这样的服务产品通常以虚拟 IP 地址（VIP）的“面目”示人。若未针对 IPv6 流量配置 BPR 或其他路由策略，那么设有 VIP 的服务产品便接收不到 IPv6 流量，于是 IPv6 流量会被正常转发至目的网络，而不“享受”任何其他服务。说透了，采用单臂模式，要想让 IPv6 流量能够正常抵达目的网络，无需针对其配置 BPR 或其他路由策略[①]。
- **路由模式**：这是三种服务产品部署模式中最为复杂的一种，实施起来也非常困难，因为以该模式部署的服务产品可能会在 IPv6 流量的转发途中“横插一杠”。以路由模式部署的服务产品都具备三层或三层以上的功能，这意味着若服务产品未启用或不具备 IPv6 数据包的转发功能，便需按以下方法之一让 IPv6 流量绕过按路由模式部署的服务产品。
  - 将 IPv6 流量路由至连接了下游服务器的专用 VLAN 或第三层物理接口。
  - 将主机的网卡配置为 Trunk。这可让带 VLAN 标记的数据帧在汇聚层交换机、接入层交换机以及下游主机之间穿梭往来。如图 9-2 所示，带 VLAN 10 标记的帧用来承载 IPv4 流量，带 VLAN 11 标记的帧则用来承载 IPv6 流量。在数据中心网络中如此行事，将会使得 IPv6 的管理和部署极其复杂。

部署纯 IPv6 服务器农场

部署纯 IPv6 服务器农场（即部署纯 IPv6 网络设备、主机、操作系统以及服

---

① 译者认为原文不通，译文做了重大调整，为避免误导读者，现给出整段原文“This mode is the easiest mode to deploy when it comes to allowing IPv6 past the services layer. In one-arm mode, traffic is specifically forwarded to the module or appliance based on the destination IP address of a configured Virtual IP (VIP) address, through Policy-Based Routing (PBR), or through another policybased mechanism. If there are no IPv6-enabled VIPs, PBR, or other routing policies configured, the IPv6 traffic continues on to the destination with no services applied. Basically, there is nothing to configure to allow IPv6 to reach its destination in this mode”。

务）是另外一种让IPv6流量通过汇聚层的手段。在企业网内部署IPv6之初，该方法可让IPv6流量、启用了IPv6的端点以及应用程序与现有的IPv4网络相隔离，这也是其大行其道的原因所在。可随着时间的推移，这种运作模式的成本（无论是投资成本还是业务成本）可谓非常之高，原因不言自明：在两种网络环境（IPv4和IPv6）中，需对网络设备、操作系统许可证/服务器硬件，以及运行维护做重复投资。在中心网络中构建双协议栈服务器农场，应当成为每个企业的终极目标。

在双栈网络中支持纯IPv4服务

企业网已实现双栈运行，但在其数据中心网络区块中，依旧部署有纯IPv4服务器的例子也屡见不鲜。或许这是拜服务器的OS不能升级为支持IPv6协议栈所赐。也有可能是因为运行在服务器上的OS能够支持IPv6，但运行在OS上的应用程序与网络层协议密不可分——即应用程序的代码只支持IPv4而不支持IPv6。无论怎样，在少数情况下，却偏偏需要让运行IPv6的主机连接纯IPv4服务器。

有几种方法可供人们在一定时期内去解决这一问题。这些方法包括IPv4/IPv6代理、NAT64（写作本书之际，NAT64还属于草案形式，由IETF的BEHAVE工作组牵头制定）以及SLB64。

某些用户会使用Apache的HTTP代理功能，在IPv6和IPv4主机之间提供基本级别的基于HTTP的数据流服务。另一些用户则会采用无状态的NAT64机制，去执行一对一的IPv6到IPv4地址转换，或采用有状态的NAT64机制，去执行多对一/多对几的IPv6到IPv4地址转换（有时也称其为地址/端口过载）。最后，若既需在两个地址家族间执行转换，又需执行服务器负载均衡服务，还可利用一下名为SLB（服务器负载均衡）64的机制，在支持该机制的设备上，会同时行使IPv6到IPv4转换和服务器负载均衡功能

在SLB64机制为设备厂商的主流产品所支持，且能在其上高效运转之前，在一定时期内，某些用户还得继续指望NAT64+SLB44（即常见的服务器负载均衡技术，但兼具IPv4到IPv4转换功能）。

无论采用以上哪种方法，都应将其视为过渡性解决方案。我们的最终目标是要将数据中心网络（包括应用程序、与应用相关的服务以及操作系统在内）建设成为真正的双协议栈网络。

在汇聚层部署支持IPv6的服务产品

在数据中心汇聚层部署支持IPv6的服务和设备时，适用于IPv4的设计

和部署原则也大半适用于 IPv6。对某特定的平台来说，等价特性可有效防止 IPv4 和 IPv6 之间完全的一对一映射，但在大多数情况下，搭建一个兼具高可用性和安全性，且便于管理的服务层所需的各项要素，对于两种协议全都相同[①]。

以下内容将会举两个在数据中心汇聚层部署 IPv6 服务的示例。第一个示例所用的设备是 Cisco 网络分析模块（NAM）。第二个示例将描述如何在数据中心网络汇聚层部署 Cisco ASA。

配置部署在汇聚层的 Cisco NAM

将 NAM 部署于汇聚层之后，便可利用其去监控进出汇聚层或来自其他数据源（比如，启用了 NetFlow 特性的设备所驻留的网段）的 IPv4/IPv6 流量，并能生成相应的报表（或执行日志汇报功能），供排除网络故障时使用。配置 Cisco NAM，令其监控 IPv6 流量并形成相关报表（或生成相关日志），与网络层协议本身无关[②]。NAM 可作为模块安装在 Catalyst 6500 交换机或 Cisco ISR 路由器内，配置时，只需将正确的 VLAN 或接口定义为数据源。配置 NAM 实际上与 IPv6 协议本身无关，如例 9-5 所示。

例 9-5 配备了 Cisco NAM 模块的 Catalyst 6500 交换机

```
6k-agg-1# show module
Mod Ports Card Type                                  Model              Serial No.
--- ----- -------------------------------------- ---------- ----------

. . . # Output summarized
  9     8  Network Analysis Module                 WS-SVC-NAM-2       SAD074900GH

!
analysis module 9 management-port access-vlan 16
!
monitor session 5 source vlan 10
monitor session 5 destination analysis-module 9 data-port 1
```

① 原文是“Feature parity for a given platform can prevent a perfect 1:1 mapping between IPv4 and IPv6, but for the most part, the elements needed for a highly available, secure, and well-managed services layer are the same between to the two protocols”。译者不明其意，完全按照字面意思翻译。不知“Feature parity”是不是指同时适用于 IPv4 和 IPv6 的特性，作者没有交代，译者也不敢借题发挥。

② 原文是“The configuration of the Cisco NAM to support the monitoring and reporting of IPv6 traffic is protocol agnostic in nature”。

例9-5所示为一台在数据中心接入层作为服务交换机的Catalyst 6500[①]，其9#槽位已安装了Cisco NAM-2模块。例中的命令analysis module 9 management-port access-vlan 16，明确了NAM模块所安装的槽位和管理端口所在的VLAN。例中还定义了与监控数据流相关的配置。对于本例，所要监控的数据源为VLAN 10中的数据，即要将VLAN 10中的数据发送至交换机第9槽Cisco NAM的数据端口（data-port）1以供分析。图9-3所示为NAM图形界面的配置，该配置与例9-5中命令行的配置所起的作用相同。

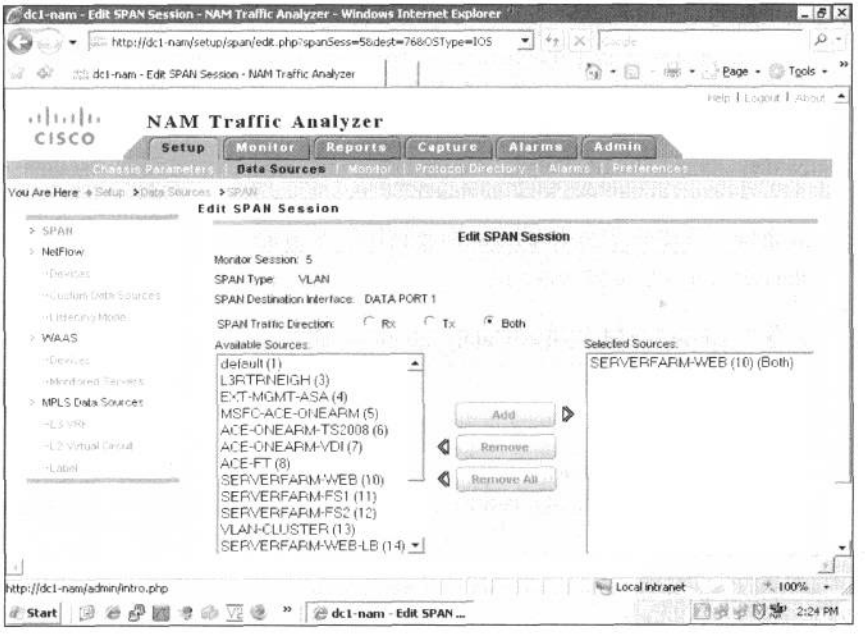


图9-3 Cisco NAM监控数据源的配置

配毕所要监控的数据源之后，NAM便能够对数据源（VLAN）中过往的数据进行监控和捕获，并能生成相关报表了。NAM的图形配置界面会随其版本不同，而发生细微的变化。图9-4所示为主机2001:db8:cafe:10:da61与其他主机之间应用程序会话的详细视图。

要是将Cisco NAM模块（设备）直接安装在汇聚层交换机内或网络中的其

① 该交换机与汇聚层交换机互连。

他位置，所涉及到的配置应该也与本节示例基本相同。

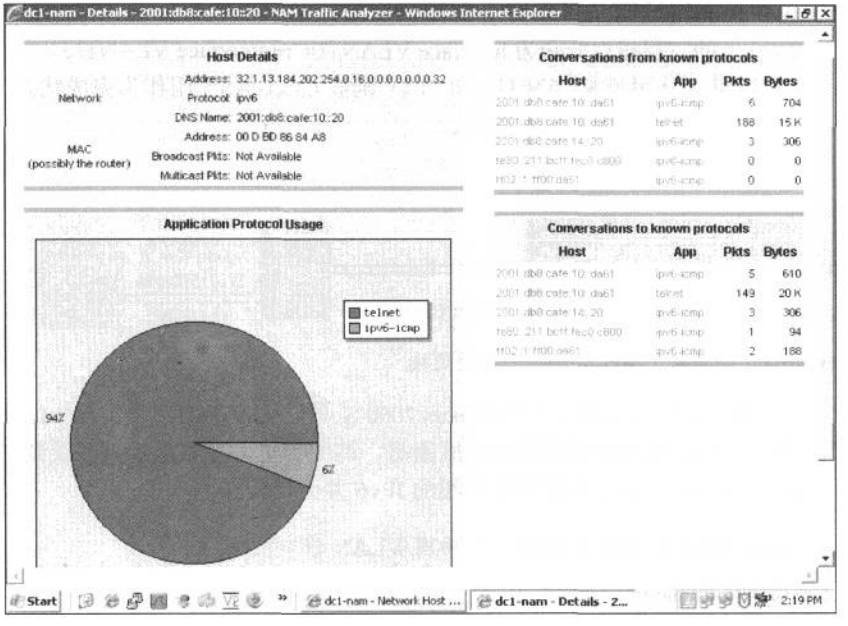


图 9-4 Cisco NAM 生成的网络中主机应用程序会话图

更多与 Cisco NAM 产品及配置有关的信息请见：http://www.cisco.com/en/US/products/ ps5740/Products_Sub_Category_Home.html。

配置部署在汇聚层的 Cisco ASA

在数据中心汇聚层部署 Cisco ASA 是我们所举的第二个示例。一般而言，在数据中心汇聚层中，Cisco ASA 的物理布局方式与一般的纯 IPv4、纯 IPv6 或双栈配置场景并无任何区别。Cisco ASA 的虚拟或逻辑设置方式应该也与此类似，即便不同，也要取决于部署时所要用到的特性。

物理链路的连接方式，以及运行于这些物理链路之上的 VLAN 配置方式可谓多种多样。图 9-5 中的网络拓扑图所示为成对部署于汇聚层的 Nexus 7000 数据中心交换机，以及分别连接到这两台交换机的 ASA 5580 防火墙（运行于透明模式）。

由图 9-5 可知，每台 Cisco ASA 和 Nexus 7000 之间有 4 条物理链路互连。在 Nexus 7000 上，有两条 10G 以太网链路分别连接 Cisco ASA 的 outside 和 inside

接口，已将这两条链路配成了Trunk，以期透传多个VLAN的流量；Cisco ASA也正是利用这两条链路对进出汇聚层的流量加以检测，这两条链路在Nexus 7000上的三层接口分别为interface VLAN114和interface VLAN115。而另外两条GE链路（分属于VLAN 116和117）则被Cisco ASA用作为完成状态化故障切换的故障切换链路和状态链路。

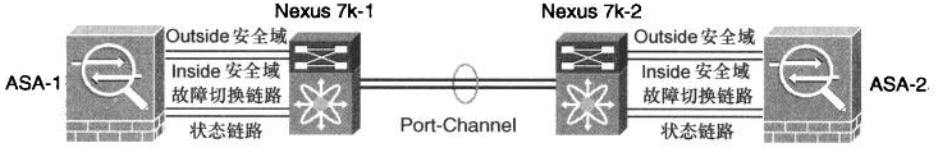


图9-5 Nexus 7000 汇聚层交换机与ASA之间的连接

例9-6所示为其中一台Nexus 7000汇聚层交换机的部分物理/VLAN接口配置。之所以未给出完整配置的原因是，其中的绝大多数配置只涉及到Trunk和port-channel，而与本章所要介绍的IPv6并无瓜葛。

例9-6 Nexus 7000汇聚层交换机——物理/VLAN接口配置

```
interface Vlan114

  no shutdown
  description Outside FW VLAN
  ipv6 address 2001:0db8:cafe:0114::0002/64

  hsrp version 2

  hsrp 114 ipv6

    preempt delay minimum 180

    timers 1 3

    ip autoconfig


interface Vlan115

  no shutdown
  description Inside FW VLAN
  ipv6 address 2001:0db8:cafe:0115::0002/64
```

（待续）

```
  hsrp version 2

  hsrp 115 ipv6

    preempt delay minimum 180

    timers 1 3

    ip autoconfig


interface Ethernet1/2
  description ASA-1-Outside
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 114

  logging event port link-status
  logging event port trunk-status
  no shutdown

interface Ethernet1/3
  description ASA-1-Inside
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 115

  logging event port link-status
  logging event port trunk-status
  no shutdown
interface Ethernet1/4
  description ASA Failover Link
  switchport
  switchport access vlan 116

  no shutdown

interface Ethernet1/5
  description ASA State Link
  switchport
  switchport access vlan 117

  no shutdown
```

对部署于数据中心的 Cisco ASA 而言，不但在配置方面所涉甚多，而且部

署起来也相对复杂，其中势必涉及到启用不同安全特性和安全策略的多上下文配置。例9-7所示的配置摘录自适用于图9-5所示场景的Cisco ASA防火墙的基本配置。请注意，admin上下文的配置并未示出。

例9-7 Cisco ASA系统配置示例

```
firewall transparent
!
interface GigabitEthernet3/0
 description LAN Failover Interface
!
interface GigabitEthernet3/1
 description STATE Failover Interface

interface TenGigabitEthernet5/0
 description N7k1 Outside
!
interface TenGigabitEthernet5/0.114
 vlan 114
!
interface TenGigabitEthernet5/1
 description N7k1 Inside
!
interface TenGigabitEthernet5/1.115
 vlan 115
!
failover
failover lan unit primary
failover lan interface FO GigabitEthernet3/0
failover replication http
failover link STATE GigabitEthernet3/1
failover interface ip FO 2001:db8:cafe:116::1/64 standby 2001:db8:cafe:116::2

failover interface ip STATE 2001:db8:cafe:117::1/64 standby 2001:db8:cafe:117::2

!
context sf-1
  allocate-interface TenGigabitEthernet5/0.114 outside
  allocate-interface TenGigabitEthernet5/1.115 inside
  config-url disk0:/sf-1.cfg
```

由例9-7可知，Cisco ASA的物理接口分别与VLAN 114和115相关联。故障切换的配置则包括了故障切换链路、状态链路以及IPv6主备地址的配置。请

注意，在配置 Cisco ASA 的故障切换功能时，IPv4 和 IPv6 地址只能两者择其一，不能“得陇望蜀”。最后，在上下文 sf-1 中，定义并分配用于 outside 和 iniside 安全域的接口。

例 9-8 所示的配置摘录自 Cisco ASA 的 sf-1 上下文的配置。

例 9-8 Cisco ASA 上下文配置示例

```
firewall transparent
hostname sf-1
!
interface outside
 nameif outside
 security-level 0
!
interface inside
 nameif inside
 security-level 100
!
access-list bpdu ethertype permit bpdu
ipv6 address 2001:db8:cafe:114::5/64 standby 2001:db8:cafe:114::6

ipv6 access-list v6-ALLOW permit ip any any

ipv6 access-list v6-ALLOW permit icmp any any

ipv6 access-list v6-ALLOW permit eigrp any any

access-group bpdu in interface outside
access-group v6-ALLOW in interface outside

access-group bpdu in interface inside
access-group v6-ALLOW in interface inside
```

sf-1 上下文的配置相对简单。该上下文（防火墙）运行于透明模式，并拥有 2 个接口（由例 9-7 中的配置可知，已将接口 t5/0.114 和 t5/1.115 分配给了该上下文）。由于 ASA 运行于透明模式，因此需放行 BPDU 数据包通过。此外，还为该上下文分配了一个 IPv6 地址，在其上设置的 ACL 也颇为宽松（permit any any）。在透明模式中，名为 ethertype 的过滤器作用于第二层数据帧，而 IPv6 访问列表则作用于第三层数据包。最后，别忘了在接口上应用访问列表。

其他种类的网络服务产品（比如，Cisco IPS/IDS）同样适用于 IPv6 网络环境，也常被部署在数据中心汇聚层。对于设备或服务来说，无论是令两者去支

持 IPv6，还是去支持 IPv4，其中的大部分配置工作都十分类似。欲知如何配置 Cisco IPS 以支持 IPv6 的详细信息请参见 http://www.cisco.com/en/US/docs/security/ips/7.0/configuration/guide/idm/idmguide7.html。

在汇聚层，除了要部署上述应用服务（application service）以外，还会存在出于复制/连接存储和网络服务（比如，用于集群和虚拟机移动性[例如，VMware vMotion]，第二层网络延伸）的目的，而互连多个数据中心的情况。与该需求有关的设计和实施方案既要支持 IPv6，也要支持 IPv4。本章稍后将会讨论这方面内容。

9.1.3 数据中心核心层

有诸多要素可以决定，一个企业的数据中心网络是否有必要包含专门的核心层，或者说，是否只需要让数据中心网络汇聚层与园区网络区块的核心层直接相连（以避免构建数据中心网络的核心层）。数据中心的规模是决定因素之一，只要规模够大，便会迫使企业去建设专门的数据中心网络核心层，以连接所有数据中心汇聚层交换机。无论什么样的理由，也不管数据中心网络核心层在整体上是如何设计，在该层，IPv6 的部署通常都不值一提[①]。因为保证网络的核心层快速、稳定、可控且兼具可扩展性，是网络设计的基本目标之一，故而很少有人会在核心层交换机上激活众多特性并部署大量服务。

由于在数据中心网络核心层交换机上所需激活的特性不会太多，因此，核心层交换机上与 IPv6 有关的配置仅限于：一，配置与其他网络区块（比如，WAN、全区网络区块以及 Internet 边缘区块）互连所使用的相关第三层接口的地址；二，配置路由协议，例如，EIGRPv6、OSPFv3 或 IS-ISv6。至于核心层交换机上的其他 IPv6 配置（与 QoS 和设备/网络的安全性有关的配置），则往往套用已就位的 IPv4 配置。

读至此处，读者想必已掌握了如何配置（核心层、）网络设备的 IPv6 地址、路由协议以及对三层设备的管理访问；因此，本节也就不再给出任何配置示例了。

9.2 在采用虚拟化技术的数据中心内实施 IPv6

如果业务连续性是企业追求的主要目标，那便意味着网络应 7 × 24 小时可

① 原文是“Regardless of the reasoning or even the overall design of the data center core, the IPv6 deployment at this layer is often trivial in design and deployment”。

用。为了达成这一目标，某些企业可能会实施异地冗余，在异地建设多个数据中心，每个数据中心都会执行数据和应用程序的复制功能。

对虚拟化的应用也早已遍地开花，在网络、服务器、桌面、安全、存储以及应用程序领域都能够见到虚拟化的身影。将虚拟化应用于以上所有技术，再对这些技术加以组合，便能构建出一个在功能上更为灵活的数据中心环境。这不但更具成本效益，而且还搭建起了一个具备高可用性且易于管理的数据中心体系架构。由于数据中心是以虚拟化的方式来构建，因此 IP 地址的规模和使用量势必也会水涨船高。在此类网络环境中，继续使用 IPv4 便会暴露出某些局限性，比如，IPv4 固有的安全性问题，以及因 IPv4 地址空间缺乏，而很难满足迅猛增长的末端设备上线的需求。

如今，诸如 VMware vSphere 之类的服务器虚拟化软件已能支持 IPv6。其解决方案则由若干零散的物理和虚拟要素构成，如图 9-6 所示[①]。

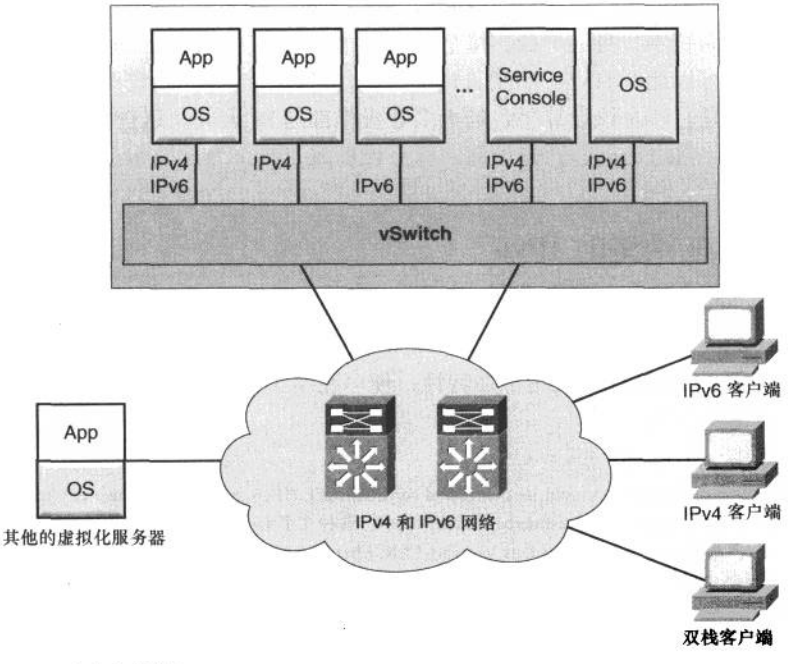


图 9-6 使用 IPv6 的服务器虚拟化

① 原文是“The solution comprises a number of individual physical and virtual elements, as shown in Figure 9-6”。译者不明作者原意，译文为直译。

如今，绝大多数的guest操作系统（驻留于物理服务器上的虚拟机的操作系统）都已能支持IPv6的运行。与物理服务器一样，取决于安装在虚拟机上的个别操作系统的能力，亦可将虚拟机配置为只运行IPv6、只运IPv4或同时运行IPv4和IPv6。应将底层网络配置为支持双协议栈（IPv4和IPv6），以确保执行以下操作时能够建立起IP连通性：一，对虚拟机的管理访问；二，虚拟机对基于IP的存储访问；三，迁移/备份虚拟机上所运行服务[①]。

VMware ESX、Microsoft Hyper V和Xen之类的服务器虚拟化软件不但支持IPv6，而且还支持运行于Hypervisor（虚拟化管理软件）之上的VM运行双协议栈[②]。除了能够让Hypervisor上的虚拟交换机以桥接的形式转发IPv6数据包以外，虚拟系统的管理组件还应该支持IPv6[③]。至于如何配置VMware vSphere和Hypervisor ESXi，以令网管人员能够通过IPv6管理到这两种虚拟系统，请见第12章的配置示例。

Cisco统一计算系统（UCS）是一个集计算能力、网络、存储以及服务器虚拟化系统于一身的平台。每个UCS系统都能支持成千上万个虚拟机，每个虚拟机都可以运行IPv6。出于管理的目的，可针对该计算系统开启IPv6功能，但相反的是，涉及该计算平台的大多数操作都与IP无关，这些操作与服务器硬件、供电、网络访问、存储访问，以及裸机或Hypervisor操作系统管理有关[④]。

9.3 实施SAN网络的IPv6

Cisco多层导向器级交换机（Cisco Multilayer Director Switch）[⑤]MDS 9000可提供SAN网络接入特性（比如，IP上的光纤通道[FCIP]和iSCSI），并同时支持IPv4和IPv6。利用上述特性，便可在基于IP的现有网络上提供基于光纤通道的服务[⑥]。

---

① 原文是“The underlying network should be configured for dual-stack (IPv6 and IPv4) connectivity to ensure access for management, IP-based storage, and migration/backup services”。这种文字实在无法翻译，译文只能拼凑而成。

② 原文是“Server virtualization software such as VMware ESX, Microsoft Hyper V, and Xen support IPv6 and dual-stack for the VMs that run on the Hypervisor”。对于这种文字，译者不敢保证译文能否正确传达作者的原意，特给出原文。

③ 作者的原意应该是要表达：能够通过IPv6管理到虚拟化系统。

④ 作者的原意应该是，对IPv6的支持是UCS全部功能的一小块，只有当通过IPv6管理UCS时，才涉及IPv6，UCS的许多功能都与IPv6（甚至是IP）无关。译者只是不明白作者为何要在此处插入这一段。

⑤ 这种型号的交换机译者还未曾见过真身，中文译名来自Cisco官网。

⑥ 原文是“These services leverage theexisting IP-based network to provide Fibre Channel–based services.”，直译为“这些服务可充分利用现有的基于IP的网络，以提供基于光纤通道的服务。”译者觉得“服务‘提供’服务”的句式似乎不通，因此在译文中调整了主语。

9.3.1 FCIP

有了 FCIP 技术，网管人员便可利用基于 IP 的现有网络，在企业网内延伸 SAN（存储区域网络）所能企及的范围。由于使用专门的光纤通道去延伸（扩展）SAN，不但实施起来代价高昂，而且还增加了管理成本，因此 FCIP 技术被证明为性价比极高。NX OS 可在单个物理接口上支持双协议和多条 FCIP 隧道（每条隧道可同时支持 IPv4 和 IPv6），如图 9-7 所示。

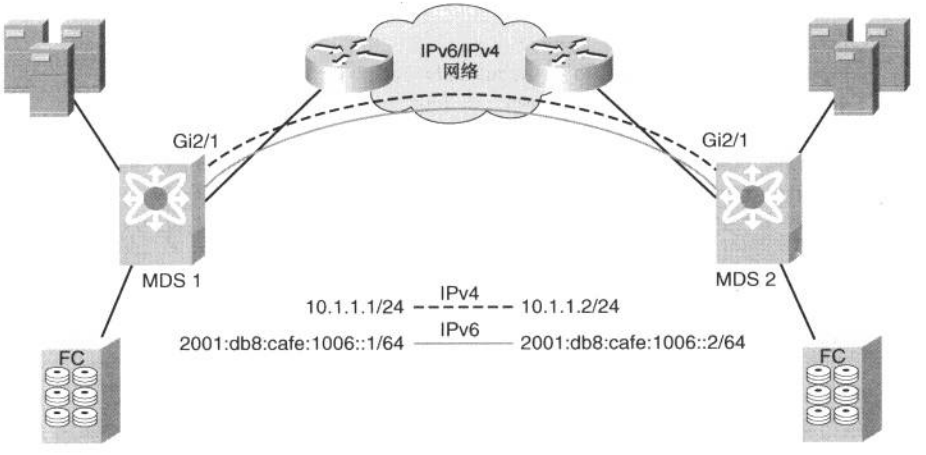


图 9-7 MDS 上使用 IPv4 和 IPv6 的 FCIP

FCIP IPv6 配置的第一步是要在 MDS 交换机上启用 IPv6 路由选择和 FCIP 特性，如例 9-9 所示。

例 9-9 在 MDS 交换机上启用 IPv6 和 FCIP

```
ipv6 routing
feature fcip
```

紧接着，在 MDS 交换机上配置 GE 接口，如例 9-10 和例 9-11 所示。

例 9-10 MDS1 上 GE 接口的配置

```
interface GigabitEthernet2/1
  ip address 10.1.1.1 255.255.255.0
  ipv6 address 2001:db8:cafe:1006::1/64
  no shutdown
```

例 9-11 MDS1 上 GE 接口的配置

```
interface GigabitEthernet2/1
  ip address 10.1.1.2 255.255.255.0
  ipv6 address 2001:db8:cafe:1006::2/64
  no shutdown
```

可使用 **ping** 命令来验证连通性，如例 9-12 所示。

例 9-12 MDS 上 ping 命令的输出

```
ping ipv6 2001:db8:cafe:1006::1
PING 2001:db8:cafe:1006::1(2001:db8:cafe:1006::1) 56 data bytes
64 bytes from 2001:db8:cafe:1006::1: icmp_seq=1 ttl=64 time=0.636 ms
64 bytes from 2001:db8:cafe:1006::1: icmp_seq=2 ttl=64 time=0.597 ms
64 bytes from 2001:db8:cafe:1006::1: icmp_seq=3 ttl=64 time=0.578 ms
64 bytes from 2001:db8:cafe:1006::1: icmp_seq=4 ttl=64 time=0.590 ms
64 bytes from 2001:db8:cafe:1006::1: icmp_seq=5 ttl=64 time=0.571 ms

--- 2001:db8:cafe:1006::1 ping statistics ---
5 packets transmitted, 5 received, 0% packet loss, time 3996ms
rtt min/avg/max/mdev = 0.571/0.594/0.636/0.031 ms
```

接下来，需配置 FCIP profile 和 FCIP 隧道。对于本例，会利用一个 GE 接口来建立两条（虚拟）FCIP 隧道。其中的一条隧道为现有的 IPv4 隧道，另一条为 IPv6 隧道。如需配置多条 FCIP 隧道，可分别为它们配置不同的 TCP 端口。在本例中，我们用 TCP 端口 3326 来配置这条新的 IPv6 FCIP 隧道。可将流量慢慢地“割接”进这条新的 IPv6 FCIP 隧道。例 9-13 和例 9-14 所示为 FCIP profile 和 FCIP 隧道的配置。本例虽然创建了两条隧道（一条用于 IPv4，另一条用于 IPv6），但在单条隧道上同时传递 IPv4 和 IPv6 流量也无不可，记住这一点非常重要。

例 9-13 MDS1 交换机上 FCIP 的配置

```
fcip profile 1
  ip address 10.1.1.1
!
fcip profile 2
  port 3226
  ip address 2001:db8:cafe:1006::1
```

（待续）

```
!
interface fcip1
  use-profile 1
  peer-info ipaddr 10.1.1.2
  write-accelerator
  ip-compression auto
  no shutdown
!
interface fcip2
  use-profile 2
  peer-info ipaddr 2001:db8:cafe:1006::2 port 3226
  write-accelerator
  ip-compression auto
  no shutdown
```

例 9-14 MDS2 交换机上 FCIP 的配置

```
fcip profile 1
  ip address 10.1.1.2
!
fcip profile 2
  port 3226
  ip address 2001:db8:cafe:1006::2
!
interface fcip1
  use-profile 1
  peer-info ipaddr 10.1.1.1
  write-accelerator
  ip-compression auto
  no shutdown
!
interface fcip2
  use-profile 2
  peer-info ipaddr 2001:db8:cafe:1006::1 port 3226
  write-accelerator
  ip-compression auto
  no shutdown
```

可执行 **show interface** 或 **show fcip summary** 命令，来验证 FCIP 的配置，如例 9-15 和例 9-16 所示。

例 9-15 MDS1 上 show fcip summary 命令的输出

```
mds-1# show fcip summary
--------------------------------------------------------------------------------
Tun prof    Eth-if    peer-ip          Status T W T Enc Comp  Bandwidth    rtt
                                              E A A                max/min    (us)
--------------------------------------------------------------------------------
1   1    GE2/1      10.1.1.2          TRNK  Y Y N  N    A    1000M/500M  1000
2   2    GE2/1      2001:db8:cafe:    TRNK  Y Y N  N    A    1000M/500M  1000
                     1006::2
```

例 9-16 MDS2 上 show fcip summary 命令的输出

```
mds-2# show fcip summary
--------------------------------------------------------------------------------
Tun prof    Eth-if    peer-ip          Status T W T Enc Comp  Bandwidth    rtt
                                              E A A                max/min    (us)
--------------------------------------------------------------------------------
1   1    GE2/1      10.1.1.1          TRNK  Y Y N  N    A    1000M/500M  1000
2   2    GE2/1      2001:db8:cafe:    TRNK  Y Y N  N    A    1000M/500M  1000
                     1006::1
```

9.3.2 iSCSI

Internet SCSI（iSCSI）广泛应用于 TCP/IP 上的 SCSI 协议传输。iSCSI 是一种基于标准的协议，可用来传递 SCSI 命令及响应。有了该技术，主机便可利用网卡，基于 TCP/IP，来完成对存储阵列的数据块级访问（block-level access）。

主机需要安装 iSCSI 驱动程序，iSCSI 驱动程序为 SCSI 和 TCP/IP 协议之间架起了一座沟通的桥梁。驱动程序会把主机和存储之间所使用的 SCSI 命令，转换为 TCP/IP 网络上传播的 iSCSI 荷载。此外，驱动程序还会将接收到的（从存储发往服务器的）iSCSI 荷载，反转为 iSCSI 命令。

在服务器上，由于吉比特网卡可在主机和存贮之间提供吉比特的速率[①]，因此广泛用作为执行 iSCSI 功能的标配网卡。

MDS 9000 系列交换机支持 IPv6 上的 iSCSI 配置，如图 9-8 所示。MDS 9000 交换机的 GE 接口和 iSCSI Initiator 都可以配置 IPv6 地址，如例 9-17 所示。

① 这明摆着是废话。

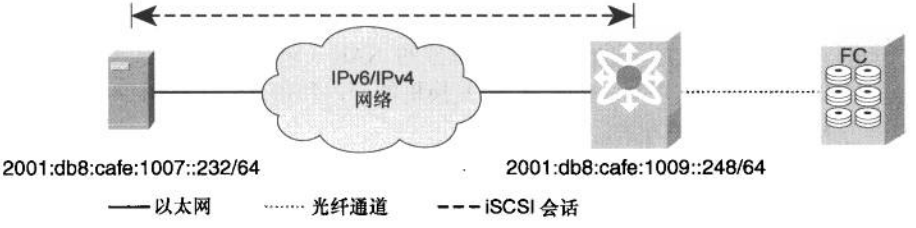


图 9-8 Cisco MDS iSCSI 配置

例 9-17 Cisco MDS iSCSI 的 IPv6 配置

```
feature iscsi
ipv6 routing
!
interface GigabitEthernet2/2
  ipv6 address 2001:db8:cafe:1009::248/64
  no shutdown
!
iscsi initiator ip-address 2001:db8:cafe:1007::232
```

图 9-9 所示为一台 Microsoft Windows Server 2008 R2 主机上的 iSCSI Initiator 客户端，该主机使用 MDS 9000 交换机作为 iSCSI 会话的 target portal。

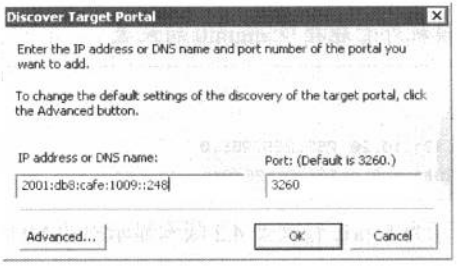


图 9-9 iSCSI Initiator 客户端

9.3.3 管理 Cisco MDS

Cisco Fabric 管理器是一款网络管理工具，可利用 SNMPv3，以图形界面的方式来展示网络中设备的实时视图。Cisco Fabric 管理器依赖于底层操作系统的

IPv6配置。Cisco Fabric管理器和MDS交换机之间建立IPv6连通性之后，SNMP的get和set操作与IPv4完全相同。

可分别通过mgmt0端口和虚拟的SAN（VSAN）接口（亦称为FC上的IP[IPFC]），实现对MDS 9000交换机的带内和带外管理。图9-10所示为管理MDS交换机的方法。

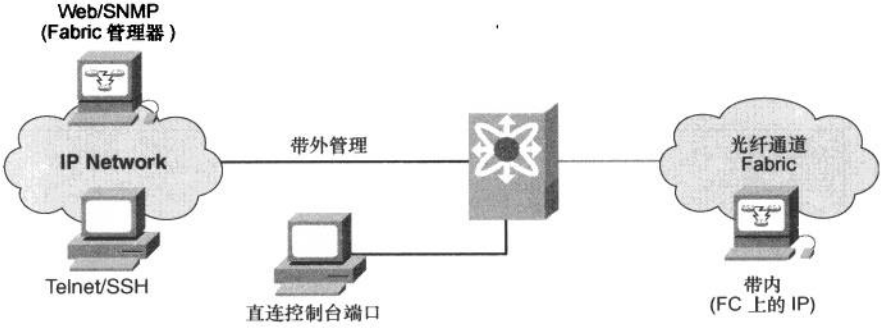


图9-10 管理MDS交换机

图9-10所示的接口，只要能配置IP地址，则既可以配置IPv4地址，也可以配置IPv6地址。例9-18所示为带外管理接口mgmt0的配置示例和使用VSAN10的带内接口的配置示例[①]。

例9-18 Cisco MDS交换机外管理接口mgmt0的配置

```
interface mgmt0
  ipv6 enable
  ip address 10.121.10.20 255.255.255.0
  ipv6 address 2001:db8:cafe:10::20/64
```

图9-11所示为Fabric管理器4.2版本显示出的MDS交换机管理接口的配置。可使用Fabric管理器针对MDS交换机分别进行IPv4和IPv6的配置，包括网络管理、FCIP以及iSCSI等。

① 例9-18只给出了mgmt0的配置。

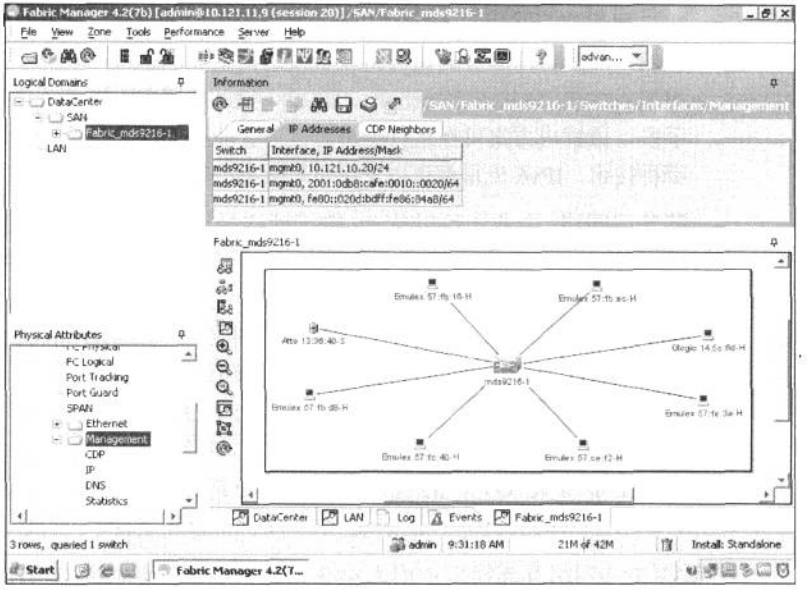


图 9-11 Cisco MDS Fabric 管理器：管理接口的 IPv6 配置

9.4 数据中心 IPv6 互连设计

本节将介绍 LAN 延伸技术（LAN extensions）（亦称为数据中心互连[DCI]），并会描述该技术与路由选择和存储空间之间的关系。本章之前的“FCIP”一节已经论述了如何利用第三层网络，在数据中心网络环境内延伸存储区域。同理，DCI 技术及设计也常用在多个数据中心之间延伸子网。

可利用 DCI 将第二层子网延伸至传统的数据中心第三层网络边界之外。用 DCL 互连多个数据中心，其目的是要显示出虚拟化数据中心模型在应用程序以及服务器的移动性方面的优势。利用应用程序以及服务器的移动性，便能够在灾难恢复、数据中心的迁移/合并,或计划内的维护时提供底层连通性。

9.4.1 设计要领：裸纤（Dark Fiber）、MPLS 和 IP

可采用以下不同的传输技术，以各种各样的方式来互连多个数据中心。

- **裸光纤**：这是典型的第一层服务。这种类型的服务虽然相对昂贵，但却大受欢迎，理由是可利用其来传递各种类型的流量，比如，以太网或 SAN 流量。

- **第二层网络**：采用这一传输方式，即等于先把纯以太网流量发送给服务提供商，然后再由服务提供商将流量发送给企业的远程站点。当然，采用覆盖型 L2VPN 解决方案（比如，虚拟专用 LAN 服务[VPLS]）也是一种替代手段，该解决方案可在运行维护方面为企业带来更多的灵活性。无论采用哪种技术，IPv6 流量都可以“纵贯”数据中心网络间的第二层网络。
- **第三层网络**：企业还可以从 SP 购买或选择使用依托于 SP 所提供的第三层服务，来互连自己的多个数据中心[①]。对于这种情况，企业网边缘设备既可与 SP 设备建立对等关系，也可依托于 SP 网络，与远程站点边缘设备建立对等关系。可在企业的各数据中心站点之间，搭建一个覆盖型网络拓扑，来延伸 LAN。OTV（覆盖传输虚拟化）技术是 Cisco 为互连多个数据中心而研发出的新成果。说穿了，OTV 即为 MAC 地址路由选择，其中 MAC 地址为目的地址（网络），但使用 IP 地址作为其下一跳。欲知更多有关 OTV 的信息，请见以下链接：http://www.cisco.com/en/US/prod/switches/ps9441/nexus7000_promo.html。

在数据中心网络互连模型中引入第 3 层概念时，无论将第 3 层视为协议的传输手段，还是封装手段，其对 ipv6 的支持方式要与对 ipv4 的支持方式保持一致，这一点至关紧要[②]。

图 9-12 所示为多数据中心互连示意图。

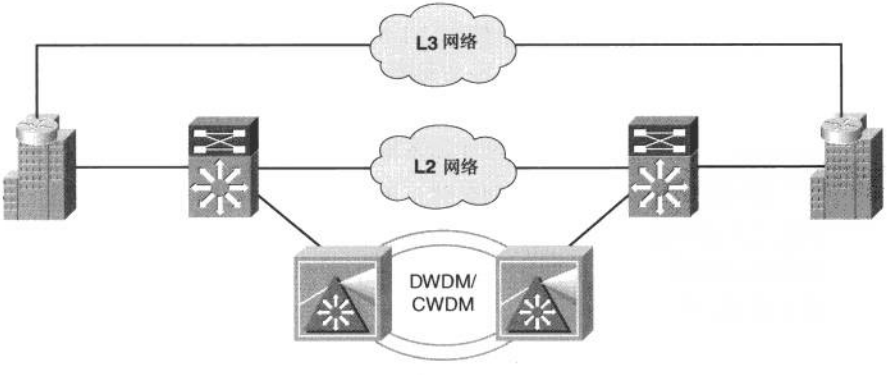


图 9-12 互连数据中心

① 原文是“Enterprises can also use Layer 3 connectivity from the SP or over the SP”。

② 原文是“When Layer 3 awareness comes into the picture, it is critical that IPv6 be supported in the same way IPv4 is, either as a means of transporting the protocol or using the protocol as a means of encapsulation”。即便按照字面意思翻出了译文，其实又有何用？给出这段文字只是不想提作者背黑锅。

9.4.2 DCI 互连和解决方案

设计 DCI 互连方案最重要的一个环节是，通过 WAN 来延伸 LAN，并需确保其无环性。表 9-1 罗列了可在 1～3 层使用的 DCI 互连解决方案。

表 9-1 DCI 解决方案

需求	解决方案
裸纤	虚拟交换系统（VSS） 虚拟 PortChannel（vPC）
L2	MPLS 上的以太网（EoMPLS） 虚拟专用 LAN 服务（VPLS） 高级 VPLS（A-VPLS）
L3	GRE 上的 EoMPLS（EoMPLSoGRE）
加密	IEEE 802.1ae IP 安全（IPSec）
QoS	分层 QoS（HQoS）

设计 DCI 时，取决于所采用的设计方案和技术，具体配置与上层的 IP 协议版本可能有关也可能无关。在规划或部署诸如 DCI 这样的新服务、技术或设计时，请确认是否将 IPv6 也纳入其中一并考虑。

与表 9-1 所列解决方案有关的配置，只要牵涉 IPv6，本书先前各章均已示出，比如，第 8 章给出了 IPSec 之于 IPv6 的相关配置，第 6 章给出了有关 VSS 的配置。更多与 DCI 有关的详细信息，请见 http://www.cisco.com/go/dci。

9.5 总结

本章涵盖数据中心网络的 IPv6 部署，并简要概述了数据中心网络双栈部署和设计原则。本章亦提供了数据中心接入层设备（Nexus 7000 和 1000v）的配置示例，重点介绍了如何在数据中心接入层配置 NAM 服务模块、Nexus 7000 交换机以及 ASA，令上述设备支持 IPv6 功能；而对于数据中心核心层，则应尽量令其保持简单，以快速执行 IPv6 数据包的转发（Layer 3 IPv6 addressing）。

在数据中心网络中，虚拟化在性能和高可用性的提升方面一直扮演着至关重要的角色。而存储亦作为数据中心网不可分割的重要组成部分——能够充分借助现有的第三层网络资源，利用IPv6上的FCIP技术，将SAN延伸至异地数据中心网络。对于本章所示出的FCIP、iSCSI以及SAN管理的配置示例，可作为企业数据中心网络环境中配置IPv6时的参考配置。

本章第四节简要介绍了在多个数据中心之间，用来延伸第二层网络的各种技术手段。这些技术及相关设计都已日渐成熟，身为网管人员不但要跟踪这些技术的发展趋势，而且还应重点关注这些技术对IPv6的支持方式，以及采用这些技术的前提要素。

9.6 参考资料

Popoviciu, Ciprian P., Eric Levy-Abegnoli, and Patrick Grossetete. *Deploying IPv6 Networks*. Cisco Press, 2006 (ISBN-10: 1-58705-210-5; ISBN-13: 978-1-58705-210-1).

Cisco Data Center Switches:
http://www.cisco.com/en/US/products/ps9441/Products_Sub_Category_Home.html.

Cisco Catalyst 6500 Series Switches:
http://www.cisco.com/en/US/products/hw/switches/ps708/index.html.

Cisco Catalyst 4900 Series Switches:
http://www.cisco.com/en/US/products/ps6021/index.html.

Cisco Network Analysis Module Products:
http://www.cisco.com/en/US/products/ps5740/Products_Sub_Category_Home.html.

Cisco ASA 5550 Series Products:
http://www.cisco.com/en/US/products/ps6120/index.html.

Cisco ASA 5500 Series Configuration Guide Using the CLI, 8.3:
http://www.cisco.com/en/US/docs/security/asa/asa83/configuration/guide/config.html.

Cisco Data Center Interconnect: http://www.cisco.com/en/US/netsol/ns975/index.html.

Cisco MDS 9000 Family NX-OS IP Services Configuration Guide:
http://www.cisco.com/en/US/docs/switches/datacenter/mds9000/sw/5_0/configuration/guides/ipsvc/nxos/ipsvc.html.

Data Center Interconnect (DCI): Layer 2 Extension Between Remote Data Center:
http://www.cisco.com/en/US/prod/collateral/switches/ps5718/ps708/white_paper_c11_493718.html.

Cisco Nexus 7000: Overlay Transport Virtualization (OTV):
http://www.cisco.com/en/US/prod/switches/ps9441/nexus7000_promo.html.

Cisco Nexus 1000V Series Switches:
http://www.cisco.com/en/US/products/ps9902/index.html.

CHAPTER 10

第 10 章　IPv6 远程访问 VPN 的部署

本章涵盖以下主题。

- **基于 Cisco AnyConnect 的 IPv6 远程访问**：本节将介绍如何利用 Cisco AnyConnect SSL VPN 客户端，以建立双栈 SSL VPN 会话的方式，提供对企业网内网网络服务的 IPv6 远程访问。
- **基于 Cisco VPN 客户端的 IPv6 远程访问**：本节将介绍如何利用 Cisco VPN 客户端和基于主机的 IPv6 隧道，以建立 IPSec 会话的方式，提供对企业网内网网络服务的 IPv6 访问。

对许多 IT 团队而言，在传统的企业网边界范围以内，实现 IPv6 访问都费了老鼻子劲，提供对远程用户访问的支持往往还要更费劲一些。可利用传统的加密型基于 VPN 客户端的解决方案，去提供 IPv6 远程访问服务，但前提是 VPN 解决方案至少具备以下三种功能之一。

- 既能够通过 IPv4 SSL（安全套接字层）VPN 传递 IPv6 流量，也要求 VPN 终结设备支持双协议栈。
- 能够借助 IPv4 IPSec VPN 会话，建立一条通往企业网内网隧道终结设备的基于 IPv6 的隧道[①]。
- 能够在远程客户端和企业站点之间，通过一条受保护的连接（比如，通过 IPSec 或 SSL），来传输纯 IPv6 流量。

第一种解决方案是指利用 Cisco AnyConnect SSL VPN 客户端（SVC），通过 IPv4 建立一条通往 Cisco ASA 的 SSL（SSL-over-IPv4）连接。随后，借助已建立的 IPv4/SSL 连接，在客户端和 ASA 之间传输 IPv6 流量，最后，在终结 IPv4/SSL

① 原文是 "Enable IPv6-based tunnels through an established IPv4 IPSec VPN session to an IPv6 tunnel termination point inside the enterprise"。作者的文字可谓处处生歧义，特给出原文。

连接的 Cisco ASA 上再将 IPv6 流量作为纯 IPv6 数据包路由至最终目的地。

第二种解决方案是利用 Cisco VPN 客户端，去建立一条基于 IPv4 的 IPSec（IPSec-over-IPv4）会话，这条 IPSec 会话通往总部站点内的某台前置设备（比如，Cisco ASA、Cisco IOS 路由器或 Cisco VPN 3000 集中器）。可采用某种隧道机制（比如，ISATAP、6to4 或手工配置的隧道）在 IPv4 报头内封装 IPv6 数据包，然后，再将经过封装的 IPv6 流量注入 IPSec VPN 连接。总部站点内的 Cisco VPN 前置设备会终结那条 IPSec VPN 会话，但仍保持流量的 IPv6-in-IPv4 的隧道封装方式，并会将经过封装的 IPv6 流量路由至企业网内部的隧道终结设备。隧道终结设备在对 IPv6-in-IPv4 隧道流量解封装之后，再对解封装的流量（纯 IPv6 数据包）执行路由选择。

如今，又有了第三种解决方案，即借助于 Microsoft DirectAccess（DA）特性。利用该特性，可在 Microsoft Windows 7 和 Windows Server 2008 R2 主机间提供纯 IPv6 的远程访问功能。Microsoft DA 要求在受保护的端点之间建立纯 IPv6 连通性。若端点间不能建立 IPv6 连通性，Microsoft DA 会尝试利用若干种隧道机制（比如，6to4、Teredo、ISATAP 以及 IP-HTTPS）来封装 IPv6 流量。使用该技术，需要充分了解 Microsoft DA 的运作机制，但对 Microsoft DA 原理、设计以及部署的介绍超出了本书的范围。要想了解 Microsoft DA 是否适用于自己的网络环境，请访问以下 Microsoft 站点链接，并参考相应的指南：

http://technet.microsoft.com/en-us/network/dd420463.aspx

本章只会着重介绍前两种解决方案，即利用 Cisco AnyConnect 和 Cisco VPN 客户端的 IPv6 远程访问。写作本书之际，Cisco 尚不支持纯 IPv6 网络上的远程访问，但 IPv6 AnyConnect 解决方案已在开发之中。欲了解 IPv6 AnyConnect 解决方案开发的最新进展，请咨询 Cisco 顾问团队，或访问 Cisco WEB 站点上的相关产品页面。

10.1 利用 Cisco AnyConnect 的 IPv6 远程访问

利用 Cisco ASA 的 Cisco AnyConnect 解决方案，远程用户可通过以下两种方式安全地连接到企业站点。

- 无客户端 SSL VPN。
- Cisco AnyConnect VPN 客户端。

无客户端 SSL VPN 解决方案（亦称 WebVPN）是指，用户打开 Web 浏览

器，连接 Cisco ASA portal，通过 IPv4/TCP 443 端口，建立起 TLS 连接。此后，客户端便可以访问驻留于企业网内部的服务了。只要对 Cisco ASA 和通过 portal 来访问的后台应用程序进行配置，令它们运作于 IPv6 之上，那么客户端便能够通过 IPv6 访问到相关应用程序。本章并不会介绍无客户端 SSL VPN 的配置，请参阅下列 Cisco 文档来获悉相关信息：

http://www.cisco.com/en/US/docs/security/asa/asa80/asdm60/ssl_vpn_deployment_guide/deploy.html#wp1016526

Cisco AnyConnect VPN 客户端是一款安装在用户主机上的应用程序。用户可以执行 Cisco AnyConnect VPN 客户端程序，并通过其建立通往 Cisco ASA 的数据报传输层安全（Datagram Transport Layer Security，DTLS）连接（通过 IPv4/UDP 443 端口）。尽管该客户端程序同样支持使用 TCP 443 端口去建立 TLS 连接，但 DTLS 通过 UDP 建立起的是低延迟路径，可规避使用纯 SSL 连接时所经常遇到的延迟和带宽问题，因此，DTLS 对延迟敏感型的应用（比如，语言）有非常大的帮助。针对 Cisco AnyConnect 环境启用 DTLS 时[①]，会同时建立起两条并行的隧道：一条用于 TLS，另一条用于 DTLS。要是 UDP 隧道因故未能建立，或建立后中断，流量还可以“走”基于 TLS 的隧道。

图 10-1 所示为运行双协议栈的主机利用 SVC 访问公司企业网的示意图。主机利用 SVC，在 IPv4 Internet 上，建立了一条通向 Cisco ASA 的 DTLS 会话。在 Cisco ASA 上，自然也要启用双协议栈功能。Cisco ASA 收到主机通过 DTLS 连接发送的 IPv6 数据包之后，会将数据包路由至公司内网的目的网络。

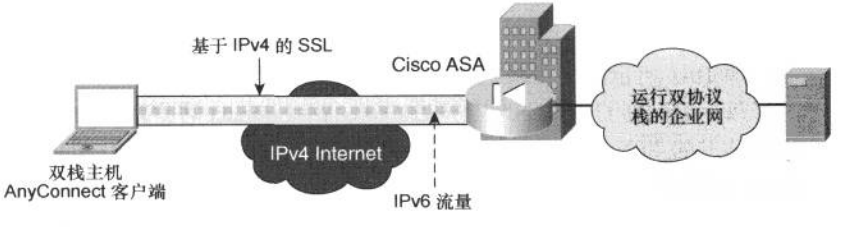


图 10-1 Cisco AnyConnect VPN 客户端连接

图 10-2 所示为利用 Cisco AnyConnect VPN 客户端实施远程访问的拓扑示例，本节所展示的配置均出自该拓扑。客户端的 IPv4 地址为 172.16.1.2，通过 IPv4 连接到了 Internet。Cisco ASA 也通过一台 Cisco 路由器（图中并未显示）

① 原文是“When DTLS is enabled for the Cisco AnyConnect environment”，译文似乎不通，但是译者还是选择直译。

连接到 IPv4 Internet，这台 Cisco 路由器提供 Internet 接入、基本过滤以及 IPv4 地址转换（NAT）功能。Cisco ASA 有 inside 和 outside 接口各一，IP 地址分别为 10.124.3.1 和 10.124.1.4。Cisco ASA 在其 inside 安全域（接口）开启了双栈功能，在其 inside 接口还配置了 IPv6 地址：2001:DB8:CAFE:1002::1。

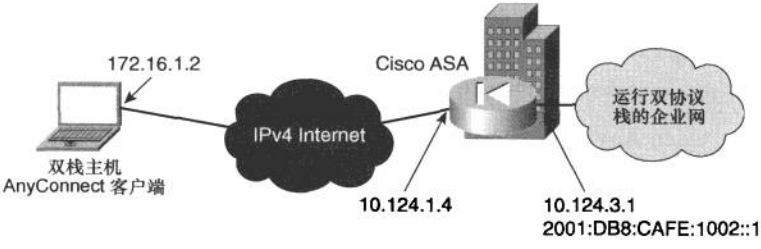


图 10-2 Cisco AnyConnect VPN 客户端拓扑示例

在 Cisco ASA 上，针对由客户端发起的 AnyConnect 会话配置了 2 个地址池。一个是 IPv4 地址池，地址范围为 10.124.3.30～10.124.3.80。另一个是 IPv6 地址池，可提供 50 个 IPv6 地址（出自前缀 2001:DB8:CAFE:1002::/64），从 2001:DB8:CAFE:1002::100/64 开始分配。Cisco AnyConnect 客户端与 Cisco ASA 之间的连接建妥之后，客户端将会从上述地址池获取到 IPv4 和 IPv6 地址。此外，其他的网络服务（比如，DNS、用户的认证和授权、Microsoft 活动目录集成等）既可以通过 IPv4 也可以通过 IPv6 来提供。与纯 IPv4 远程访问模式相比，本模式所使用的安全过滤和检测机制并无差异。至于为流入企业内网的 IPv6 流量所制定的安全策略，可与 IPv4 流量的安全策略一起，应用于网络中的同一位置。

例 10-1 所示的配置只是 Cisco ASA 完整配置中的一小部分，这些配置只是为了演示如何在 Cisco ASA 上开启 Cisco AnyConnect 客户端对 IPv6 的支持，请切勿将其视为配置中的最佳做法。请注意，例中所示的所有配置也可以使用 Cisco 自适应安全设备管理器（ASDM）GUI 来完成。

例 10-1 包含了 Cisco ASA Inside 和 outside 接口的配置。虽然 outside 接口不会去处理 VPN 会话中的 IPv6 数据包，但还是在该接口的配置中激活了 IPv6 功能（使用 **ipv6 enable** 命令），这只是 Cisco ASA 软件的硬性规定。自然也无需针对 outside 接口设置任何特别的安全策略，去防范来自 Internet 的 IPv6 攻击，这是因为 ipv6 enable 命令只是本地有效——虽会为该接口创建一个本地链路地址，但通过 IPv6 根本无法从 Internet 上访问到 outside 接口（Internet 连接为 IPv4），这就是说，攻击者需与该接口/链路处于同一第二层域内，才能有机会攻

击到 outside 接口的 IPv6 本地链路地址。

例 10-1 Cisco ASA AnyConnect 的配置

```
interface GigabitEthernet0/0
 nameif outside
 security-level 0
 ip address 10.124.1.4 255.255.255.0
 ipv6 enable                        #Software requirement to enable IPv6 on the

                                    #outside interface

!
interface GigabitEthernet0/1
 nameif inside
 security-level 100
 ip address 10.124.3.1 255.255.255.0
 ipv6 address 2001:db8:cafe:1002::1/64

!
ip local pool v4Pool 10.124.3.30-10.124.3.80 mask 255.255.255.0
ipv6 local pool v6Pool 2001:db8:cafe:1002::100/64 50    #v6 pool (50 addresses)

ipv6 route inside ::/0 2001:db8:cafe:1002::3            #default route pointing to

                                                        #next-hop inside enterprise


!
route outside 0.0.0.0 0.0.0.0 10.124.1.1 1
!
webvpn
 enable outside
 svc image disk0:/anyconnect-win-2.4.1012-k9.pkg 1
 svc enable
group-policy ANYCONNECTGRP internal
group-policy ANYCONNECTGRP attributes
 vpn-tunnel-protocol svc webvpn             #Enable SSL VPN Client and Clientless

 split-tunnel-policy tunnelall              #Prohibit split-tunneling

 webvpn
  svc dtls enable                           #Enable DTLS (TLS over UDP)

  svc keep-installer installed
```

（待续）

```
  svc ask enable default svc timeout 15
group-policy DfltGrpPolicy attributes
 vpn-tunnel-protocol IPSec l2tp-ipsec svc webvpn
 address-pools value v4Pool
 ipv6-address-pools value v6Pool

webvpn
  svc ask enable default svc timeout 15
username sslvpn1 password bzN3HgmMqoLp3Liy encrypted

username sslvpn1 attributes                #Associate test user with Group Policy

 vpn-group-policy ANYCONNECTGRP

tunnel-group DefaultRAGroup general-attributes
 address-pool v4Pool
tunnel-group DefaultWEBVPNGroup general-attributes
 address-pool v4Pool
tunnel-group ANY-TG type remote-access
tunnel-group ANY-TG general-attributes   #Assign pool/group policy to tunnel grp

 address-pool v4Pool

 ipv6-address-pool v6Pool

 default-group-policy ANYCONNECTGRP
```

客户端建妥了通往 Cisco ASA 的有效 SVC 连接之后，可执行若干条命令来查看连接会话的状态及相应的统计信息。例 10-2 所示为两条不同命令的输出。

第一条命令的输出显示出了 IPv6 地址池的名称、地址范围、大小以及在用和可用地址的数量。这份输出还显示出了“In Use addresses（在用地址）”和“Available Addresses（可用地址）”列表（由于输出过长，未完全显示）。

第二条命令的输出显示的是 **vpn-sessiondb** 的汇总信息，其中，DTLS 隧道的状态信息尤为重要。输出中包含了已分配出去的 IPv4 和 IPv6 地址，以及客户端用来建立连接的公网地址。

例 10-2 IPv6 Pool 和 vpn-sessiondb 命令的输出

```
asa-1# show ipv6 local pool v6Pool
IPv6 Pool v6Pool
Begin Address: 2001:db8:cafe:1002::100

End Address: 2001:db8:cafe:1002::131

Prefix Length: 64
Pool Size: 50

Number of used addresses: 2
Number of available addresses: 48

In Use Addresses:
2001:db8:cafe:1002::100
2001:db8:cafe:1002::101
Available Addresses:
2001:db8:cafe:1002::102
2001:db8:cafe:1002::103
!OUTPUT OMITTED

asa-1# show vpn-sessiondb detail svc
!OUTPUT SUMMARIZED...
DTLS-Tunnel:
  Tunnel ID     : 6.3
  Assigned IP   : 10.124.3.30            Public IP     : 172.16.1.2

  Assigned IPv6: 2001:db8:cafe:1002::100

  Encryption    : AES128                  Hashing       : SHA1
  Encapsulation: DTLSv1.0                 UDP Src Port : 4430

  UDP Dst Port : 443                      Auth Mode     : userPassword
  Idle Time Out: 30 Minutes               Idle TO Left : 30 Minutes
  Client Type  : DTLS VPN Client

  Client Ver    : AnyConnect Windows 2.4.1012
  Bytes Tx      : 8720                    Bytes Rx      : 26074
  Pkts Tx       : 109                     Pkts Rx       : 303
  Pkts Tx Drop : 0                        Pkts Rx Drop : 0
```

图 10-3 所示为用来建立上述连接的 Cisco AnyConnect VPN 客户端的统计信息。

图 10-3 Cisco AnyConnect VPN 客户端的统计信息

现在，用户可以通过一条 AnyConnect SSL 会话，访问到企业网内网的 IPv4 和 IPv6 应用程序和服务了。从远程用户到企业网内部前置设备（Cisco ASA）的纯 IPv6 远程访问（纯 IPv6 上的 SSL VPN）解决方案尚处于开发之中，预计很快将会投放市场。

10.2 利用 Cisco VPN 客户端的 IPv6 远程访问

与 Cisco AnyConnect 解决方案不同，Cisco VPN 客户端本身并不支持 IPv6。然而，利用 Cisco VPN 客户端，再结合使用基于主机的隧道机制（动态或静态），便能够实现 IPv6 的远程访问。在远程客户端上同时使用 ISATAP 隧道（RFC 5214），并建立 Cisco VPN 连接来传递 IPv6 数据报，正是该应用的一个实例。请牢记，ISATAP 是基于主机的隧道，可用在主机与路由器/交换机/服务器之间建立IPv6连通性。本解决方案的工作原理是：在远程主机上，先行建立Cisco VPN 客户端连接，以求在远程主机和企业网内部的隧道端点设备之间建起连通性。连通性一旦建立，便能够通过 IPv4 IPSec 连接（由 Cisco VPN 客户端建立），来发送经过隧道封装的 IPv6 流量。

在图 10-4 所示的网络拓扑场景中，某远程用户通过一条 IPSec 会话，连接到 Cisco IPsec VPN 终结设备（可以是 Cisco IOS 设备、ASA，或 VPN 3000 集中器）。IPsec 连接建妥之后，还要在远程主机和企业网内部的 IPv6 隧道终结设备之间，建立 ISATAP 隧道（或其他类型的基于主机的隧道）。IPv6 隧道终结设备终结了 ISATAP 隧道之后，会通过双栈或纯 IPv6 连接，将远程主机发来的 IPv6

流量路由至企业内部的目的网络。对于本部署方式，在企业网内部（即企业内部的隧道终结点），将隧道流量解封装之后，需对解封装的 IPv6 流量施以 IPv6 安全 ACL 以及流量检测策略。

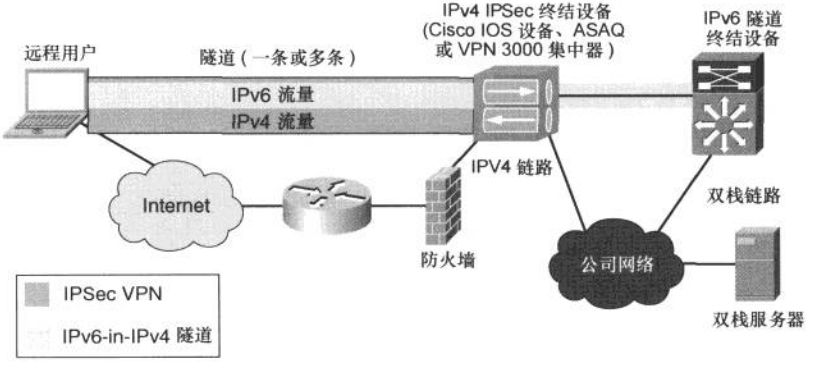


图 10-4 Cisco VPN 客户端与基于主机的隧道结合使用

注意 本书的第 6 章已经对 ISATAP 及其在企业网中的高可用性的部署方式展开过了深入讨论。

例 10-3 所示为企业网内部 ISATAP 隧道终结设备的基本配置。在该设备（部署在企业网内部的 Cisco IOS 路由器或配备了 Supervisor 720 的 Catalyst 6500 交换机）上，创建了一个隧道接口，在接口上定义了一条 IPv6 前缀，其接口 ID 采用经过修改的源于 EUI-64 的地址（定义于 RFC 4291）。隧道连接建立之后，ISATAP 客户端会使用该 IPv6 前缀（生成 EUI-64 地址）。在隧道接口上屏蔽路由器通告（RA）消息，是 Cisco IOS 的默认行为。由于将会使用这条 ISATAP 隧道连接多个客户端，并且还需要将那条 IPv6 前缀通告给多个客户端主机，因此必须要让该隧道接口发送 RA 消息。可在 ISATAP 隧道终结设备上，禁用默认的 RA 消息抑制行为，令隧道接口发送 RA 通告消息。对 ISATAP 隧道终结设备来说，隧道的源地址为 loopback 接口地址，该隧道并无目的地址，这是因为隧道模式既然为“isatap”，即意味着该隧道的运作方式为点到多点。

注意 鉴于实际的 IPv6 流量与 VPN 终结设备无关，以下并未给出其完整配置。

例 10-3 VPN 终结设备上的 ISATAP 隧道配置（用于 Cisco VPN 客户端）

```
interface Loopback0
 ip address 10.124.109.1 255.255.255.255
!
interface Tunnel4
 no ip address
 no ip redirects
 ipv6 address 2001:DB8:CAFE:1009::/64 eui-64
 no ipv6 nd ra suppress
 tunnel source Loopback0
 tunnel mode ipv6ip isatap
```

如第 6 章所述，ISATAP 主机主要通过两种方法来获悉 ISATAP 路由器的地址，这两种方法是指：静态配置法和动态 DNS 域名解析法。使用第一种方法可有助于网管人员就何时、何人以及如何在企业中推行 ISATAP 做更为细致地控制。在运行 Microsoft Windows 操作系统的主机上，可使用以下命令来配置 ISATAP 路由器的地址（路由器 loopback 接口地址）。在该命令中既可以包括 IPv4 地址，也可以使用主机名（使用 DNS 解析主机名）。对于本例，**set router 10.124.109.1** 语句中的 IPv4 地址正是 ISATAP 隧道终结设备的 loopback0 接口地址——即 ISATAP 隧道的源地址。

```
C:\>netsh interface ipv6 isatap set router 10.124.109.1
Ok.
```

图 10-5 所示为 Cisco VPN 客户端已经建立起了 VPN 会话，客户端的 IP 地址为 10.124.3.30。

例 10-4 所示为 Microsoft Windows 7 主机（Cisco VPN 客户端和 ISATAP 主机）上 **ipconfig** 命令的汇总输出。由输出可知，该主机的 IPv6 地址为 2001:db8:cafe:1009:0:5efe:10.124.3.30。其中，前缀 2001:db8:cafe:1009 正是例 10-4 中为 tunnel 4 接口所定义的前缀；0:5efe 为 ISATAP 隧道接口标识符，定义于 RFC5241；10.124.3.30 则来源于主机的 IPv4 地址。

此外，例 10-4 还展示了该主机可以成功地 ping 通位于企业网内部的服务器。

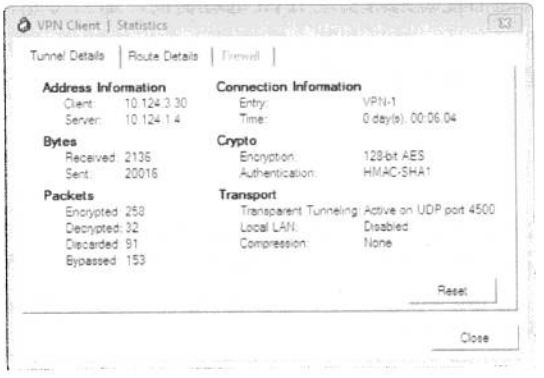


图 10-5 Cisco VPN 客户端会话

例 10-4 Microsoft Windows 主机上的 ping 和 ipconfig 输出

```
Tunnel adapter isatap.cisco.com:

   Connection-specific DNS Suffix  . : cisco.com
   IPv6 Address. . . . . . . . . . . : 2001:db8:cafe:1009:0:5efe:10.124.3.30

   Link-local IPv6 Address . . . . . : fe80::5efe:10.124.3.30%13
   Default Gateway . . . . . . . . . : fe80::5efe:10.124.109.1%13

C:\> ping 2001:db8:cafe:1005::1

Pinging 2001:db8:cafe:1005::1 with 32 bytes of data:
Reply from 2001:db8:cafe:1005::1: time=2ms
Reply from 2001:db8:cafe:1005::1: time<1ms
Reply from 2001:db8:cafe:1005::1: time=2ms
Reply from 2001:db8:cafe:1005::1: time=3ms
```

10.3 总结

网管人员需要能够解决用户对 IPv6 应用程序和服务的访问问题，即要让用户能够从企业内外的任何地方访问到部署在企业网内部的 IPv6 应用程序和服务。远程 VPN 上的 IPv6 访问解决方案正日臻完善。Cisco 提供了基于 SSL VPN 的 Cisco AnyConnect 上的 IPv6 解决方案，该解决方案支持远程用户对企业网内部的双栈访问。如果有人目前还在使用 IPSec 上的 Cisco VPN 客户端，那么我们建议你使用 Cisco AnyConnect 解决方案。要是不能采用 Cisco AnyConnect 解

决方案，且还需在短期内实现 IPv6 远程访问，那么可先行通过 Cisco VPN 客户端建立 IPSec 会话，然后在其上搭建基于主机的隧道（ISATAP、6to4 以及手工配置的隧道等）来满足上述需求。

10.4 参考资料

Popoviciu, Ciprian P., Eric Levy-Abegnoli, and Patrick Grossetete. *Deploying IPv6 Networks*. Cisco Press. (ISBN-10: 1-58705-210-5; ISBN-13: 978-1-58705-210-1).

Hogg, Scott and Eric Vyncke. *IPv6 Security*. Cisco Press. (ISBN-10: 1-58705-594-5; ISBN-13: 978-1-58705-594-2).

Microsoft. Microsoft DirectAccess: http://technet.microsoft.com/en-us/network/dd420463.aspx.

Cisco. Cisco ASA 5500 SSL VPN Deployment Guide, Version 8.x: http://www.cisco.com/en/US/docs/security/asa/asa80/asdm60/ssl_vpn_deployment_guide/deploy.html.

Cisco. Cisco ASA 5500 SSL VPN Deployment Guide - Clientless SSL Documentation: http://www.cisco.com/en/US/docs/security/asa/asa80/asdm60/ssl_vpn_deployment_guide/deploy.html#wp1016526.

Rescorla, E. and N. Modadugu. RFC 4347, "Datagram Transport Layer Security."

Templin, F., T. Gleeson, and D. Thaler. RFC 5214, "Intra-Site Automatic Tunnel Addressing Protocol (ISATAP)."

Hinden, R. and S. Deering. RFC 4291, "IP Version 6 Addressing Architecture."

CHAPTER 11

第 11 章　管理 IPv6 网络

本章涵盖以下主题。

- **网络管理框架**：本章第一节将会介绍独立于网络协议（IPv4 和 IPv6）的网络管理框架。这一框架是管理任何网络基础设施的根基。
- **IPv6 网络管理应用程序**：本节描述的内容与 IPv6 协议（IPv6 网络）和现有 IPv6 网管应用程序的结合使用有关。
- **IPv6 网络管理工具**：随着网络中两种协议（IPv6 和 IPv4）的并肩运行，网络基础设施设备也扩大了对现有内置工具（诸如 MIB、NetFlow、IP SLA 以及支持 IPv6 部署的 EEM）的支持范围。
- **IPv6 网络管理**：在网络中部署完 IPv6，并在末端主机间建立起了 IPv6 连通性之后，对企业来说，便会侧重于通过 IPv6，来扩展自己的网络基础设施设备的管理范围[①]。本节将会详尽介绍用于 IPv6 网络管理的各种应用程序。
- **IPv6 流量监控工具**：本节将重点介绍 IPv6 流量监控、捕获以及分析工具，无论是故障排除，还是网络异常行为检测都离不开这样的工具。

网络管理和网络监控不但是企业级网络运维工作的根基，而且是其重要的组成部分。网络管理的主要目标是，为网络基础设施的运维流程（例如调配、监控、日志汇报或故障排除）提供支撑。当企业在现有网络基础设施内集成了 IPv6 功能之后，则会增加网络管理的难度。如本书第一章所述，从根本上来讲，IPv6 是 IP 网络层模型的全新实现，而非其中的一项新特性。在不久的将来，IPv6 将会与 IPv4 共存于网络基础设施之上，因此需改进现有的网络管理工具、网管流程和网络管理方法，并以此来支持异构的 IPv4 和 IPv6 网络管理。

---

① 作者的原意应该是："企业在成功部署了 IPv6 之后，重点工作便落在了如何通过 IPv6 来管理网络设备的身上，而在此之前，都是通过 IPv4 来管理网络设备的。"在译者看来，作者的英文表达能力时常导致原文的字面意思与作者的原意不符。

如今，可用来管理现有 IPv4 网络的工具可谓多种多样，其范围从简单的脚本，到复杂如需下血本购买许可证，并要求后台数据库支持的网络管理应用程序。尽管，大多数网络管理工具已历经多年的发展，对于其所管理的 IPv4 网络基础设施来说已算是成熟产品，但其所拥有的特性并不见得就适用于 IPv6 网络。因此，当读者扩大 IPv4 网络管理工具的使用范围，以期管理运行双协议栈——即 IPv4 和 IPv6 两种协议混合运行的网络时，应首先关注以下问题。

- 当前在用的网络管理应用程序（工具）能否同时通过 IPv4 和 IPv6 管理到网络基础设施？哪些在用的应用程序只能支持 IPv4 网络的管理，哪些能够同时支持 IPv4 和 IPv6 两种协议？
- 当前有哪些设备处于网络管理工具的管控之下？能否通过 IPv6，用带 IPv6 扩展功能的 SNMP 或其他各种网管协议，管理到这些网络设备。
- 当前使用的是哪种 MIB（管理信息库）？是分别针对 IPv4 和 IPv6 使用一套 MIB，还是使用一套统一的与协议无关的新型 MIB？现有的网络基础设施支持 IPv6 MIB 吗？
- 如今在用的网络管理工具适用于 IPv6 过渡部署规划和长期的纯 IPv6 部署规划吗[①]？

本章会简述关键概念，并会甄别管理 IPv6 部署所需的管理工具[②]。对于那些需要在现有的 IPv4 网络管理基础设施之上，管理由自己设计并部署的 IPv6 网络基础设施的网络工程师们来说，阅读本章会对他们有所帮助[③]。表 11-1 列出了常用的 Cisco 网络设备的能力[④]。贯穿本章，我们会深入探讨这些特性，即便不讨论这些特性，表 11-1 也是对支持此类特性的设备很好的总结[⑤]。

---

① 原文是“Will the tools used today fit into the transition plan and a long-term IPv6 “native or“IPv6-only” plan?”译者只能把“a long-term IPv6 “native or“IPv6-only” plan?”翻译为“长期的纯 IPv6 部署规划”。

② 原文是“This chapter outlines key concepts and identifies tools needed for managing IPv6 deployments”。译文按字面意思直译，译者实在不晓得到底作者想表达什么。

③ 原文是“The chapter assists network architects in managing their IPv6 designs and deployments as part of their existing IPv4 network management infrastructure”。译文只能做较大改动，否则翻出来的一定让人头晕。

④ 原文是“Table 11-1 lists the capabilities of popular Cisco devices”。具体是什么能力，作者却又只字未提，译者选择直译，也就不费脑筋去替作者遮丑了。

⑤ 原文是“Throughout this chapter, these features will be discussed in detail; however, the following table provides a good summary of which devices offer these features”。译者这也知道译文很“傻”，但原文中的“however”比译文还“傻”。给出译文，坚决不替作者背黑锅。

表 11-1 常用的 Cisco 交换机 IPv6 能力

管理	协议	Nexus 7000	Catalyst 6500	Catalyst 4900	Catalyst 4500
监控和报告	SNMP	Y	Y	Y	Y
	Syslog	Y	Y	Y	Y
网络服务	NTP	Y	Y	Y	Y
	TFTP	Y	Y	Y	Y
控制和操作	Telnet	Y	Y	Y	Y
	SSH	Y	Y	Y	Y
	HTTP	Y	Y	Y	Y
	Netconf	Y	N	Y	N
MIB		Y	Y	Y	Y
接口信息		Y	Y	Y	Y
统计		Y	Y	Y	Y
ICMP		Y	Y	Y	Y
NetFlow		Y	Y	N	Y(Sup 7E)
IPSLA		N	Y	Y	Y

11.1 网络管理框架：FCAPS

为实现网络管理的简单化和标准化，网络管理者们制定出了一整套网络管理框架，旨在实现网络运维所需的一系列网络管理功能。这套通用的网络管理框架与 IPv4 或 IPv6 本身无关，但却成为了任何网络基础设施管理的基础。该框架从功能上将网络管理工作划分为 5 块（5 个功能领域）：故障管理、配置管理、记账管理、性能管理和安全管理（FCAPS）。这一整套网络管理框架实现了网络运维所需的一系列网络管理功能。

FCAPS 最先由 ISO 以工作草案的形式提出（N1719）。其中的每个功能领域都自成体系。比方说，排除网络故障时的故障诊断行为与配置网络设备就区别很大。以下各节将会简要介绍 FCAPS 的各个功能域。

注意 更多与 FCAPS 框架有关的细节请见：http://www.techfaq.com/fcaps.html。

11.1.1 故障管理

故障管理是指通过主动监控或人工响应（比如，用户申告）的方法，对故

障进行检测、分类、记录以及报告的过程。此外，故障管理还包括故障起因（root cause）和事件的关联性分析。

11.1.2 配置管理

配置管理既是控制机制，也是规章制度，在生产网络环境中恪守配置管理制度，就能够提高变更实施的成功率。配置管理还可以包括网络设备清单的自动化管理、软件序列号管理以及系统认证管理。

11.1.3 记账管理

记账管理是指在一个计费周期内，针对最终用户计量所管理的资源和服务，并将结果呈现给最终用户的方法。这包括资产跟踪、服务等级报告以及供应商管理等。

11.1.4 性能管理

性能管理包括满足目标服务等级，以及建立详细的服务等级基线的能力[①]。该功能域还包括容量规划以及网络性能分析与报告等。

11.1.5 安全管理

安全管理包括网络加密；灾难恢复和应急规划；安全告警监控、报告和策略管理等。

表11-2总结了以上5大功能域的特征和产品示例。

表11-2 对FCAPS的总结

FCAPS功能域	特征	产品举例
故障	故障检测及修复 故障分离（隔离）及网络恢复 告警信息的处理、过滤及生成 故障检测及记录 故障诊断 测试及验收 网络恢复 故障汇报	Cisco Works Tivoli Netview Cisco Fabric Manager for managing storage networks

① 原文是“Performance management includes the capability to meet target service levels and to establish a detailed service-level baseline”。译文为直译。

续表

FCAPS 功能域	特征	产品举例
配置	资源预制 网络调配 自动发现 备份及恢复 数据库的操作 变更/资产管理 认证	Cisco Works Cisco Fabric Manager Cisco Data Center Network Management (DCNM) Network Registrar Network Compliance Manager
记账	使用情况跟踪 计费 资产跟踪 服务等级管理 供应商管理	Cisco Works NetFlow Collectors: NetQoS
性能	基线定义 容量规划 性能分析 监控 报告	Cisco Works NetFlow
安全	访问控制 安全管理 安全审计 告警监控 加密 策略管理	Cisco ACS (TACACS+/RADIUS)

11.2 IPv6 网络管理应用程序

网管系统（NMS）是一种应用程序，可通过网络设备所提供的管理接口，来管理网络设备并与网络设备进行通信[①]。NMS 也被称为网络管理器。

网络管理应用程序可帮助在网络中集成了 IPv6 的企业客户完成以下任务。

- 部署 IPv4/IPv6 双栈或纯 IPv6 网络。
- 将网络管理的范围从现有的纯 IPv4 网络基础设施设备，扩大到 IPv4/IPv6 双栈或纯 IPv6 网络基础设施设备。

① 原文是“A network management system (NMS) is an application that manages the agents and communicates with the agent through a management interface provided by the agent”。译文未完全按照原文字面意思翻译，请读者留意。

Cisco 推出了一大票网络管理应用程序套件，来帮助人们管理网络基础设施。除了 Cisco 自行研发的网络管理应用程序以外，其他厂商的网络管理解决方案（比如，HP OpenView 和 IBM Tivoli）也能通过 Cisco 网络设备提供的管理接口，管理到 Cisco 网络设备[①]。就 Cisco 的产品来说，Cisco LAN 管理解决方案（Cisco LAN Management Station[②]）是业界领先的 IPv6 网管解决方案，由 CiscoView、Campus Manager 以及 Resource Manager（RME）这三个主要部件构成。

表 11-3 对 Cisco LAN 管理解决方案的三个重要组件的特性进行了总结。

表 11-3 总结 Cisco LAN 管理解决方案的重要组件的特性

重要组件	特征
CiscoView	配置，包括配置 IPv6 地址（单/多播） IPv6 邻居发现（ND） 路由器通告（RA） 多播侦听着发现（MLD） 接口信息 IPv6 邻居（RD 和 RA） 流量统计信息（IPv6 VS. IPv4 流量）
CiscoWorks Campus Manager	IPv6 拓扑服务 IPv6 ACL 用户跟踪
CiscoWorks RME	以报表形式呈现的 IPv6 地址 NetConfig 模板对 IPv6 的支持 IPv6 的 **netshow** 命令 专门用来解析和比较涉及 IPv6 的配置文件的工具 configlet

11.3 IPv6 网络工具（Instrumentation）

一般而言，借助于开放的应用编程接口（API），Cisco 或其他厂商的网络管理应用程序便能够对网络基础设施进行管理。如今，那些 API 的功能又得到了进一步的扩展，可为以下网络故障隔离、故障调试和故障排除工具所使用。

- 利用 SNMP 管理信息库（MIB）的网络设备管理工具。

① 译者注：原文是“In addition to the Cisco offerings of network management applications, third-party network management stations such as HP OpenView and IBM Tivoli can manage Cisco network devices through the built-in instrumentation on routers and switches”。译文与原文字面上有很大出入，请读者留意。此外，译者认为，原文中的“station”应该为“Solution”。

② 译者认为应该是“Solution”。

- IPv6 应用程序可视性及监控[①]，包括灵活的 NetFlow、采样流（sFLOW）、IPFIX 以及 IPv6 IPSLA。
- 调用 EEM（嵌入式事件管理器）可灵活定制网络管理任务的自动化网络管理工具（Automation using flexible programming with Embedded Event Manager [EEM]）。

接下来，会对上述工具展开详细介绍。

11.3.1 利用 SNMP MIB 的网络设备管理工具

管理信息库（MIB）以“代理”的形式提供了对网络设备的抽象，网管程序（比如，CiscoWorks）可从该代理获取到网络设备的信息[②]。（网管程序）对网络设备的管理请求可以是先前罗列的 FCAPS 功能中的任意一项，包括硬件配置信息[③]、历史上的（网络）性能趋势以及当前的设备资源状态等。每个受（网管程序）管理的对象都会有一个唯一的名为 OID（对象标识符）的东西来表示。

图 11-1 所示为 MIB 和网络设备的逻辑视图。该图以逻辑的方式，把网络设备划分为实际的资源平面和管理平面。实际的资源平面展示了网络设备完成网络通信功能所用到的真实实体。管理平面则展示了受（网管程序）管理的对象，网管程序会分别利用这些对象来管理实际的资源。将受管理对象组合在一起则表示 MIB，而 MIB 则用来表示设备本身（The managed objects combined represent the MIB, which in turn represents the device）。

（网管程序）可通过 SNMP 来访问 MIB，从而执行以下网络管理操作。

- 读取 MIB 变量：网管工作站（NMS）可通过指明 OID 的方法，从网络设备的管理代理请求 MIB 信息。管理代理会获取 OID 值，并将 MIB 信息发送给 NMS。
- 设置 MIB 变量：NMS 向网络设备的管理代理发送一个 OID 值。管理代理会将相应 OID 值更改为 NMS 要求设置的值。

最初，IETF 针对 IPv6 创建了该协议专有的 MIB，其在功能上，与现有的

---

① 原文是“IPv6 application visibility and monitoring”，译者认为原文不通，但译者还是选择按字面意思翻译。

② 原文是“A Management Information Base (MIB) provides an abstraction of the network device in he form of an “agent” that can be retrieved by a network management application (for example, CiscoWorks)”。译文基本为直译，杜撰了少许。反正译者是没弄懂作者想表达什么。

③ 原文是“physical configuration information”，译者认为“物理的配置信息”不成体统，译文酌改。

IPv4 MIB 等价。随着企业们在自己的网络中引入 IPv6，为了方便网络管理，让现有 IPv4 网络的 SNMP 功能与 IPv6 结合使用，这些企业纷纷要求 IETF 定义协议无关的 MIB。表 11-4 所列为 IPv4 和 IPv6 协议专用的 MIB RFC。

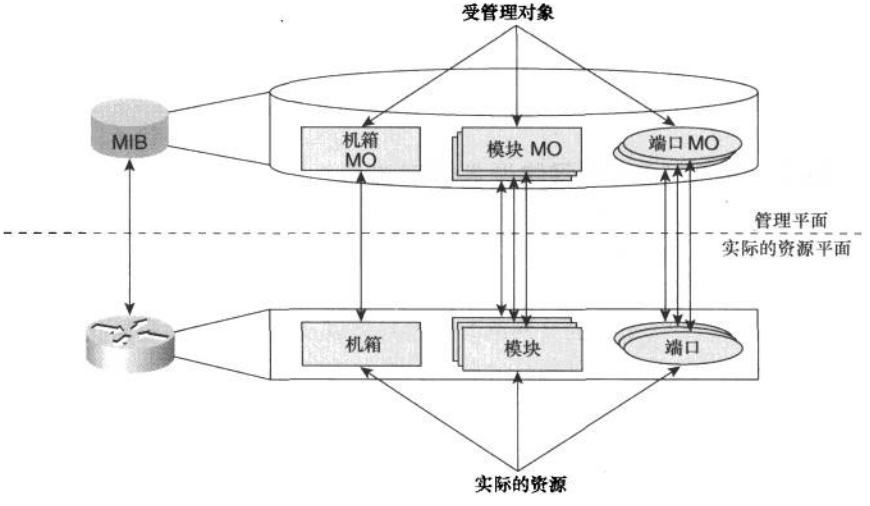


图 11-1 管理代理和 MO

表 11-4 协议专用的 MIB

IPv4 专用的 RFC	IPv6 专用的 RFC	RFC MIB 的定义
RFC 2011	RFC 2465	IP/IP 转发 MIB
RFC 2012	RFC 2452	TCP MIB
RFC 2013	RFC 2454	UDP MIB
RFC 2096（取代了 RFC 1354）	——	IP 转发表 MIB

随着采用 IPv4-IPv6 双协议栈部署的企业日渐增多，IETF 也觉得有必要去定义与协议无关（独立于 IPv4 或 IPv6 协议），并能归纳出（网络设备）管理信息的 MIB。图 11-2 所示为 IETF MIB RFC 从协议专用到协议无关的演变过程。以下 RFC 定义了协议无关的 MIB。

- RFC 4292 “IP Forwarding Table MIB”，其内提供的建议与转发 MIB（forwarding MIB）的实现有关，这一实现与 IP 版本无关[①]。该 RFC 为

[①] 原文是 “provides recommendations to support IP version-independent implementations of the forwarding MIB”。原文很“绕”，为保证译文的精确性，译文也只能做如此变通了。

IPv4 和 IPv6 提供了一个通用的转发 MIB 框架。

- RFC 4293 “Management Information Base for the Internet Protocol (IP)”，是对 IP MIB 的修订，意在创建一套描述和管理 IP 模块的对象，该对象与 IP 版本无关。这份 RFC 合并了 RFC 2465、2466 和 2011，以提高 IPv6 设备的可管理性。
- RFC 4022 “Management Information Base for the Transmission Control Protocol (TCP)”，定义了协议无关的 TCP MIB。
- RFC 4113 “Management Information Base for the User Datagram Protocol (UDP)”，定义了协议无关的 UDP MIB。

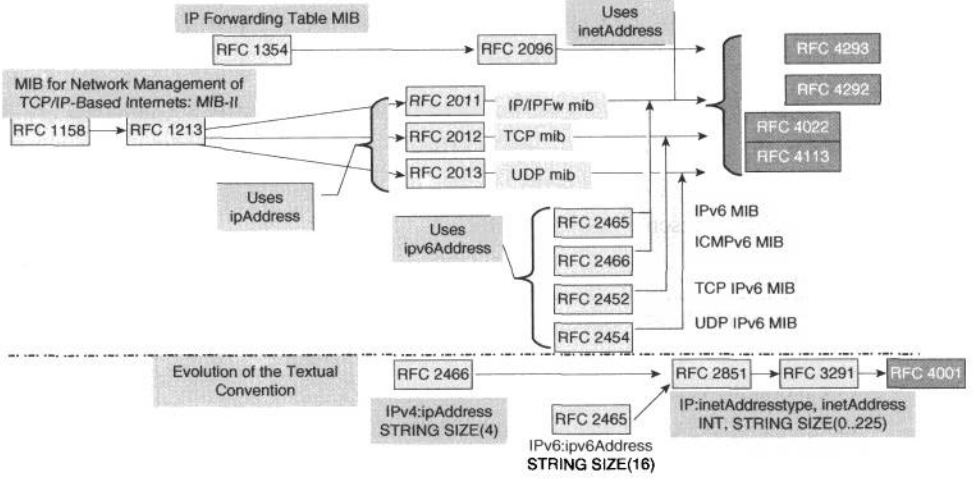


图 11-2 MIB 的演进过程

注意 有了对独立于协议的 MIB 的定义之后，所有的 Cisco 设备都可以支持协议无关的 MIB 了。

有关 IPv6 MIB

MIB 以 SNMP MIB 语言的形式来定义，并作为 RFC 文档提交 IETF 审核通过。除非另有说明，否则 Cisco enterprise MIB 一定会恪守相关 RFC 所定义的准

则。RFC 总是不断被更新，MIB 自然也不甘落后。一般而言，NMS 应使用最新的 MIB，以确保对自己所管理的网络中的“风吹草动”了如指掌（to ensure maximum network visibility）。

若 NMS 无法从一台受其管理的设备（比如，一台 Cisco 路由器）获取所请求的 SNMP 信息，其原因可能是缺失相关的 MIB，该缺失的 MIB 允许 NMS（针对受管理设备）执行具体的数据采集任务。一般而言，若 NMS 获取不到某特定的 MIB 变量，要么是拜 NMS 无法识别该 MIB 变量所赐，要么归咎于设备的管理代理不支持该 MIB 变量。对于第一种情况（NMS 不识别指定的 MIB 变量），可能需要将相关 MIB 加载进 NMS（借助于 MIB 编译器）。说明白一点，为针对 Cisco 设备执行特定的数据采集任务，NMS 管理员可能需将 Cisco MIB 或兼容的 RFC MIB 加载进 NMS。对于第二种情况（设备的管理代理不支持 MIB 变量），则需检查 Cisco IOS 版本或系统软件。不同的 MIB 所支持的 Cisco IOS 版本也不一样。具体信息可见 Cisco IOS 版本的相关说明。

5000 号以内的 RFC 文档中，大约有 300 份涉及 MIB（Approximately 300 RFCs in the first 5000 RFCs contain MIBs）。虽说讨论所有的 MIB 超出了本章的范围，但本章会重点介绍与协议无关，但与管理网络设备有关的 MIB。表 11-5 所列为集成进 Cisco 产品中的 MIB。

表 11-5 IPv6 MIB 描述

与 IPv6 有关的 MIB	描述	RFC
CISCO-CONFIG-COPY-MIB	用来复制 Cisco 路由器配置的 MIB	Cisco 私有 MIB
CISCO-CONFIG-MAN-MIB	配置管理 MIB	Cisco 私有 MIB
CISCO-DATA-COLLECTIONMIB	用来周期性采集数据的 MIB	Cisco 私有 MIB
CISCO-FLASH-MIB	用来管理 flash 设备	Cisco 私有 MIB
CISCO-IETF-IP-FORWARDING-MIB	用来管理 IP 路由器。这是 IP-FORWARD-MIB 的 Cisco 版本	Cisco 私有 MIB
CISCO-IETF-IP-MIB	用来管理 IP 和 ICMP 操作。这是 IP-MIB 的 Cisco 版本	Cisco 私有 MIB
IP-FORWARD-MIB	用来管理 IP 路由器	RFC 4292
IP-MIB	用来管理 IP 和 ICMP	RFC 4293
ENTITY-MIB	通过单一 SNMP 代理，管理多个逻辑实体	RFC 4133
NOTIFICATION-LOG-MIB	用来记录 trap 和 inform 这样的 SNMP 事件	RFC 3014
SNMP-TARGET-MIB	用来远程配置 SNMP 参数	RFC 2273
CISCO-SNMP-TARGETEXT-MIB	用来远程配置 SNMP 参数。这是 SNMP-TARGET-MIB 的 Cisco 版本	Cisco 私有 MIB

续表

与 IPv6 有关的 MIB	描述	RFC
INET-ADDRESS-MIB	针对所定义的 IP 地址，定义文本约定	RFC 4001
TCP-MIB	用来管理 TCP 连接	RFC 4022
UDP-MIB	用来管理 UDP 连接	RFC 4113

11.3.2 IPv6 应用程序可视性及监控

本小节将简要介绍 Cisco 路由器所具备的 IPv6 应用程序识别特性和监控工具[①]。本小节将讨论以下特性。

- 灵活的 NetFlow。
- IPFIX。
- IPv6 SLA。

灵活的 NetFlow

从 1996 年开始，Cisco 路由器和交换机自出厂就具备 NetFlow 特性[②]。NetFlow 是获取 IP 业务数据的事实上的标准，可帮助网管人员实施网络及安全监控、规划、流量分析和记账等。此外，NetFlow 还能让网管人员更好地理解：

- 如何通过监控网元的使用情况，来建立网络基线；
- 网络中流淌着什么样的流量，从哪儿来到哪儿去，何时流过，方向如何；
- 如何检测和分类安全事故。

为了充分利用 NetFlow 这一特性，需务必弄清该特性在企业网中所部属的层次，如图 11-3 所示。

NetFlow 版本

NetFlow 历经多年的发展，已能够支持许多新协议并引入了许多新的增强功

① 原文是“This section provides an overview of features that enable application visibility as well as monitoring tools that can be used with IPv6”。译文未按原文译出，作者的文字实在太过一般。

② 译者实在忍不住想说两句，作者也太高看自己效力的公司了。1996 年 Cisco 生不生产交换机暂且不提，即便时至今日，绝大多数新型 Cisco 交换机也绝不是出厂就自带了 NetFlow 特性。

能，如表 11-6 所列。NetFlow 的最新版本 V9 提供了对 IPv6 的支持，并新引入了许多其他额外的增强功能。本书将重点介绍 NetFlow V9（亦称为灵活的 NetFlow）。

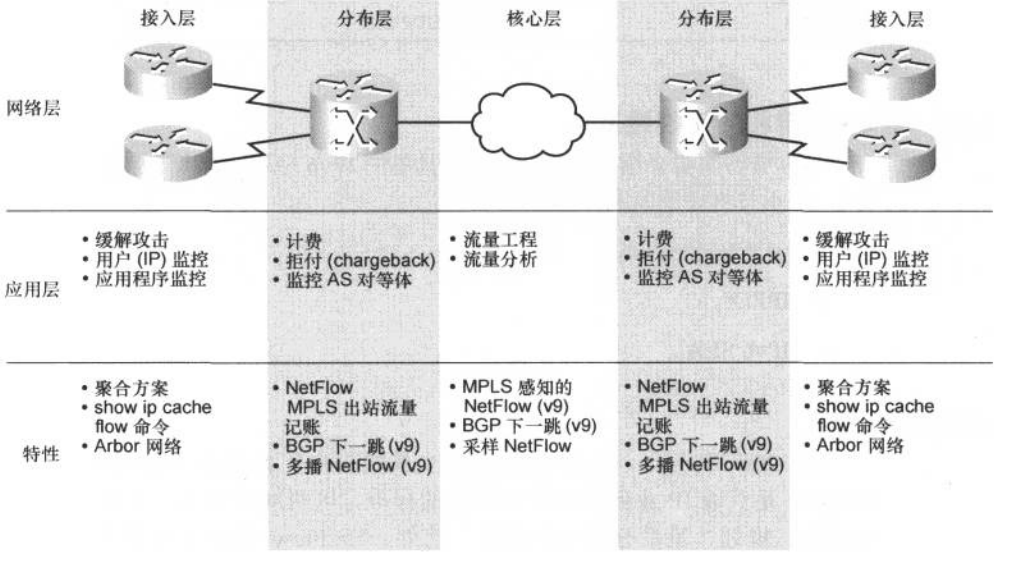


图 11-3 NetFlow 的部署

表 11-6 NetFlow 输出版本特性

输出格式	所支持的特性
版本 9	支持 IPv6。本章将重点介绍 NetFlow 版本 9。该版本支持多播、DoS 检测、BGP 下一跳等[①]。定义于 RFC 3954
版本 8	支持聚合方案，从而降低了对资源的占用情况
版本 5	最为常见，也是被广泛使用的版本
版本 1	原始版本，一般只有老式采集系统才需要使用该版本

注意 欲了解其他版本的 NetFlow 信息，请浏览 http://www.cisco.com/en/US/prod/collateral/iosswrel/ps6537/ps6555/ps6601/prod_white_paper0900aecd80406232.html。

① 原文是“This version also supports Multicast, DoS, BGP next hop, and so on”。原文真是吓死人不赔命，NetFlow 版本 9 居然支持 DoS，请问是能够利用其发动 DoS 攻击吗？

NetFlow 版本 9（灵活的 NetFlow[FnF]）

就传统意义而言，网管人员使用 NetFlow 不外是执行网络/安全监控、网络规划、流量分析以及 IP 记账等任务。灵活的 NetFlow 是对原始的 NetFlow 的改进，具备定制流量分析参数的能力。NetFlow 可按两种模式运行，即完整模式（full mode）和采样模式（sampled mode）。在完整模式中，只要 NetFlow 表的容量允许，NetFlow 会对流入接口的每个数据包“明察秋毫”。而在采样模式中，NetFlow 只会随机选择部分数据包（数据包的采样率通常为 1/32），并利用其来总结网络流量的特征。用 NetFlow 采样模式，既可以降低所采集的 NetFlow 输出数据总量，亦能够减少发送至外部采集设备（以供进一步分析）的 NetFlow 输出数据流量[①]。NetFlow 的采样模式常用于以下两种情形。

- 容量规划：使用采样模式，虽然只能了解网络流量的概况，但也能为网络容量规划提供足量的信息。
- 高速接口（比如，10GE 接口）上的网络流监控：为每个数据包建流（creating flows），会加重 NetFlow 缓存和采集设备的负担。

与传统的 NetFlow V5 相比，灵活的 NetFlow 又添加了几项重要的增强功能，如表 11-7 所列。

表 11-7 NetFlow 和 NFN 的比较

传统的 NetFlow V5	灵活的 NetFlow（FnF）	优点
所有应用共用一组流信息和单一缓存	分别对不同的 NetFlow 应用进行跟踪	对安全性数据和流量分析数据的跟踪分开进行 可将不同的流量监控结果输出到不同的目的地 可针对每种应用，展开详尽的分析
所有应用共用单一环境	可创建虚拟的 NetFlow 缓存，从而能够以更为灵活的方式“观察到”网络中的应用	可创建虚拟的 NetFlow 缓存。可对网络中的安全或流量事故进行隔离 能够以自定义的方式来标识流量 在判定和隔离安全事故的准确性方面几乎可达 100%
使用固定的流字段，可实现有限的数据聚合	用户可对流信息进行配置，并且还可以利用 NetFlow 版本 9 输出中的新数据类型	具备选择相关信息的能力 来自第二层及以上诸层的新信息，包括数据包部分[②]

① 原文是“Sampled NetFlow enables network architects to reduce the volume of NetFlow export that is collected and sent to the external NetFlow collector for further analytics”。能把这句话用中文讲通并不容易。

② 原文是“Ability to select relevant information. New information from Layer 2 and above, including packet sections”。译者实在不知作者的本意，因此按字面意思直译。

以下几节将深入探讨 NetFlow 的完整和采样模式。

NetFlow 的完整模式

NetFlow 的完整模式支持对构成特定流的每个数据包进行采集更新[①]。这也是 NetFlow 的默认行为，可帮助网管人员“查明”穿梭于特定交换机的每一个数据包。由于在高端路由和交换设备上，总会以硬件来“接管”NetFlow 特性，因此为每个流创建流记录，并不会增加设备自身的 CPU 利用率。然而，若要将 NetFlow 流记录输出至外部 NetFlow 采集设备，那就需要 CPU 出面干预了，在流量居高不下的情况下，对 NetFlow 流记录的输出处理势必会导致设备的 CPU 利用率上升。取决于配置在流量接收接口上的流掩码，每个流在 NetFlow 硬件表中都有相对应的表项。

流掩码

流掩码用来确定所要采集的统计数据的粒度。举例来说，若将流掩码设置为 source-only，那么 NetFlow 表中的每一条记录只会包含源 IP 地址。NetFlow 在 NetFlow 表中存储流记录，只要流掩码配置不当，NetFlow 表就会变得异常庞大。可使用流掩码来指明流入（网络设备接口的）数据包中的以下字段，NetFlow 会利用这些字段来标识数据包流。

- 源和目的 IP 地址。
- 源和目的 TCP 端口号。
- IP 协议类型、数据包流入接口的 VLAN 号（input VLAN）以及数据包报头中的 ToS 位。

NetFlow 的运作方式

NetFlow 在运作时，会用到 7 种参数，如图 11-4 所示。NetFlow 会针对这 7 种键值（key），对数据包进行检测，并在 NetFlow 缓存中创建数据包流。随后，还可将数据包流输出至外部 NetFlow 采集设备（比如 CiscoWorks 和 NetQoS）。

---

① 原文是“Full NetFlow allows collection updates for every packet that constitutes the identified flow”。译文为直译，反正译者是没弄懂作者的原意。据译者推测，如果待采集的特定流中的每个数据包发生了变化，Netflow 特性对此“了然于胸”，这才是所谓的采集更新。作者并没有这么说，这只是译者推测，仅供读者参考。

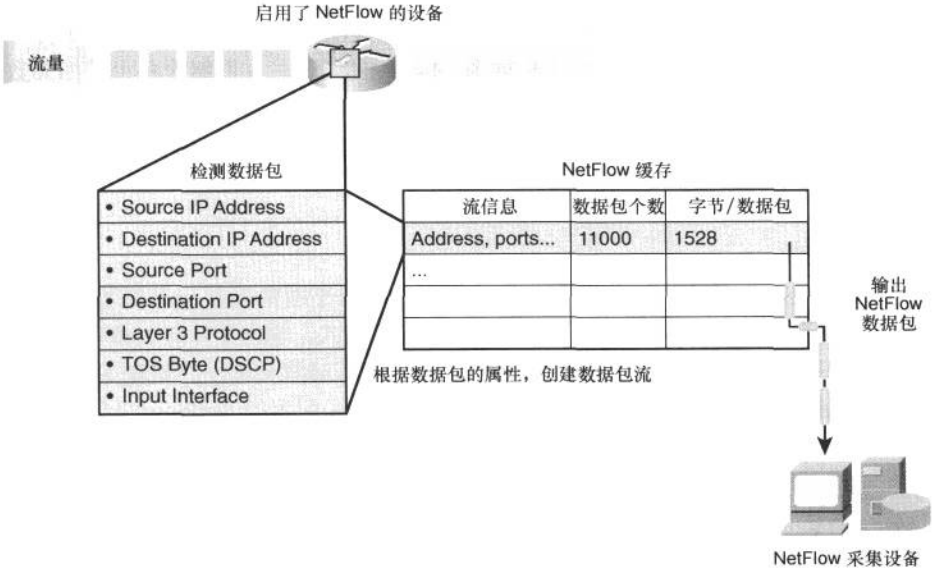


图 11-4 NetFlow 的操作方式

注意

欲了解与商业及免费的 NetFlow 采集程序有关的信息，请浏览以下 URL: http://www.cisco.com/en/US/prod/collateral/iosswrel/ps6537/ps6555/ps6601/prod_white_paper0900aecd80406232.html。

NetFlow 的采样模式

设计 NetFlow 的采样模式，是为了让网管人员根据自己所定义的基于时间的采样率或数据包采样率，来获取采样数据。若网管人员选择根据时间进行采样，并将采样数据输出，那么还需定义采样率（sampling rate）。采样率定义了给定时间间隔内的一段时长，NetFlow 会在每个时间间隔中的这段时长内采集流信息。基于数据包的采样率也使用类似的套路，由流中数据包的个数除以采样率，来得出采样次数。

NetFlow 采样模式的最明显的缺点是，不能准确给出流的组成情况。正因如此，通常，NetFlow 采样模式只适用于无需“览网络之全豹”（a full view of the data is not required）的场景。在做网络及容量规划时，正是 NetFlow 采样模式一显身

手的好时机。在采样模式的操作中，只要 NetFlow 表的容量允许，NetFlow 引擎就会按 M/N 的数据包采样率，对接口上的数据包进行采样（采样率介于 1:2 到 1:8000 之间）。现在进一步拓展图 11-4 中的示例，在网络设备上，如果数据包的接收接口（egress interface）为高带宽接口，其带宽在 10GE 或 10GE 以上（比如，40G 或 100G），那么采样模式正是推荐使用的 NetFlow 操作模式，因为这可以节省路由器资源的消耗。

NetFlow 数据输出

NetFlow 数据输出（NDE）是一种数据平面的操作，其过程是首先提取 NetFlow 记录，然后再将其按所指定的记录格式，输出至外部 NetFlow 采集设备（采集程序）。运行于交换机/路由器上的 NDE 进程会输出存储在 TCAM 表中的 NetFlow 数据报，并利用 UDP 将其发送到预先配置的 NetFlow 采集设备。由于使用 UDP 来发送 NetFlow 数据包，网络流数据输出操作缺乏相应的确认机制，故而会出现数据输出记录丢失的情况。不过，NetFlow 采集设备可依赖 NetFlow 数据报报头中包含的序列号字段（类似于 TCP 头部中的序列号字段）来查明数据输出记录是否丢失。NetFlow 数据记录中的序列号等于最后收到的数据报的序列号加上前一个数据报中流的个数[①]。在序列号的帮助下，NetFlow 采集设备再减去期待的序列号，以确定输出过程中是否发生了数据报丢失的情况[②]。

网管人员可基于以下任意一种情形，去触发 NetFlow 流记录的输出。

- 接收流的网络设备端口宕。
- 发生了路由翻动。
- 管理员手工清除了 NetFlow 表项。
- NetFlow 表项因 NetFlow 老化计时器之一过期而超时。
- 一个老化计时器（aging timer）会告知 NED 进程，何时可以输出流记

① 原文是“The sequence number in the NetFlow data record is equal to the sequence number of the last datagram plus the number of flows in the previous datagram”。译文按原文字面意思直译。但在译者看来，这句话应该这么写“网络采集设备收到的 NetFlow 数据流记录数应该等于：最后收到的数据报中所含流的记录数加上前一个数据报（NetFlow 报头）中的序列号。”

② 原文是“With the help of the sequence number, the NetFlow collector can then subtract the expected sequence number to determine whether any datagrams were lost in the export process”。译文为直译。但在译者看来，这句话应该这么写“在收到了新的数据报之后，NetFlow 采集设备会从 NetFlow 报头中提取预期的序列号，然后再用该序列号值减去自己已接收的 NetFlow 数据流记录数，并以此来断定是否发生了 NetFlow 数据流记录的丢失。”译者申明，译者的看法纯属个人意见，仅供读者参考。译者再唠叨一句，按作者的写法，不但文字不通，而且还不合逻辑。

录。流记录“老化”意谓删除过期的 NetFlow 表项的过程。只有当流记录到期时，才会从 NetFlow 表中删除，而且还会将其随同统计信息一并输出。

网管人员需对现有的 NetFlow 采集应用程序进行适当的调整，以支持不同的 NetFlow V9 格式。

在 Nexus OS 中配置灵活的 NetFlow

在 Nexus OS 中配置灵活的 NetFlow 包括以下 4 个步骤。

步骤 1 配置 exporter（NetFlow 采集设备参数）。

步骤 2 配置流记录。

步骤 3 配置流监控。

创建一个新的 NetFlow 缓存；

挂接流记录；

将先前配置的 exporter 挂接到新创建的 NetFlow 缓存。

步骤 4 在接口上配置 NetFlow。

以下内容将对以上 4 个配置步骤展开深入讨论。

配置 exporter

可通过 CLI 或其他应用程序来访问（查看）NetFlow 数据。配置 exporter 的目的有三：一，指明将 NetFlow 数据输出发送至何处；二，定义 NetFlow 数据输出的传输类型；三，定义 NetFlow 数据输出的属性。

注意 写作本书之际，只能将 IPv4 地址用作为 NetFlow 数据输出的目的地址。

配置流记录

流记录用来表征 NetFlow 缓存中的流。NetFlow 流记录包含一组键值字段和非键值字段。网络管理应用程序将会根据所监控到的流量类型，来支持用户定义或预定义的流记录[①]。

① 原文是“Network management applications will support user-defined and predefined flow records based on the type of traffic being monitored”。译文为直译。

配置流监控

流监控器（flow monitor）定义了存贮在 NetFlow 缓存中的信息。流监控包含流记录。[①]

配置接口

最后一步，还要将流监控器应用于需做流量分析的接口。例 11-1 所示为一份简单的 NetFlow 配置示例。

例 11-1 Cisco Nexus OS 灵活的 NetFlow 配置

```
! Create Exporter
flow exporter my-exporter
destination 1.1.1.1
! Create Flow Record
flow record my-record
match ipv6 destination address
match ipv6 source address
collect counter bytes
! Create Flow Monitor
flow monitor my-monitor
exporter my-exporter
record my-record
! Apply Flow Monitor on interface
int gig3/1
ip flow monitor my-monitor input
```

在 Cisco IOS 中配置灵活的 NetFlow

在 IOS、IOS-XE 以及 IOS-XR 中，IPv6 的 NetFlow 配法与 IPv4 类似，但需关注下列事宜。

- 必须使用 NetFlow 输出版本 9。
- NetFlow 同时作用于出/入站流量。
- NetFlow 支持 IPv6 的 L2 和安全监控。
- 需要基于 IPv4 的采集设备来完成 NetFlow 数据输出功能。

---

① 原文是“The flow monitor defines the information stored in the NetFlow cache. The flow monitor contains the flow records”。译文为直译。

- 必须配置 IPv6 Cisco 特快转发（CEF）。
- 在接口上必须配置 IPv6 flow 命令。

例 11-2 所示为 Cisco IOS 中的 IPv6 NetFlow 配置。

例 11-2 Cisco IOS 中的 IPv6 NetFlow 配置

```
ipv6 unicast-routing
ipv6 cef
ip flow-export destination 172.28.103.122 99
!
interface FastEthernet0/0
 no ip address
 duplex auto
 speed auto
 ipv6 address 2001:DB8:CAFE:155::1/64
 ipv6 enable
 ipv6 flow ingress
 ipv6 flow egress
```

例 11-3 所示为 IPv6 NetFlow 的输出命令，其语法与 IPv4 大同小异。

例 11-3 show ipv6 flow cache 命令输出

```
Router-1# show ipv6 flow cache
IP packet size distribution (49990 total packets):
   1-32   64   96  128  160  192  224  256  288  320  352  384  416  448  480
   .000 .000 .000 .999 .000 .000 .000 .000 .000 .000 .000 .000 .000 .000 .000

    512  544  576 1024 1536 2048 2560 3072 3584 4096 4608
   .000 .000 .000 .000 .000 .000 .000 .000 .000 .000 .000

IP Flow Switching Cache, 475168 bytes
  8 active, 4088 inactive, 8 added

  100 ager polls, 0 flow alloc failures
  Active flows timeout in 30 minutes
  Inactive flows timeout in 15 seconds
IP Sub Flow Cache, 33928 bytes
  12 active, 1012 inactive, 12 added, 8 added to flow
  0 alloc failures, 0 force free
  1 chunk, 1 chunk added
SrcAddress         InpIf    DstAddress       OutIf    Prot SrcPrt DstPrt Packets
```

（待续）

```
2001:DB...:155::2 Fa0/0     FE80::2...E:9478 Local     0x3A 0x0000 0x8800 1
FE80::2...49:1540 Fa0/0     FE80::2...E:9478 Local     0x3A 0x0000 0x8800 1
FE80::2...6E:9478 Local     FE80::2...9:1540 Fa0/0     0x3A 0x0000 0x8800 1
2001:DB...:155::1 Local     2001:DB...155::2 Fa0/0     0x3A 0x0000 0x8100 16K
2001:DB...:155::2 Fa0/0     2001:DB...155::1 Local     0x3A 0x0000 0x8000 35K
FE80::2...49:1540 Fa0/0     FE80::2...E:9478 Local     0x3A 0x0000 0x8700 1
FE80::2...6E:9478 Local     FE80::2...9:1540 Fa0/0     0x3A 0x0000 0x8700 1
```

IPFIX

IP 协议流信息输出（IPFIX）定义于 RFC 3917，是一种尚在完善中的 IETF 标准，基于 NetFlow 版本 9。比之 NetFlow V9，IPFIX 在提法和用法上都有所不同，但工作原理几近相同。IPFIX 基本上可算作“推（push）”协议。这意味着每个发送者只会周期性地将 IPFIX 消息发送给接收者，而接收者不会对此做任何干预。发送者可在自己所发的消息中自行定义数据类型，这就是说，IPFIX 不但可自由扩展，而且还适用于不同的场景。图 11-5 所示为 IPFIX 流。

与灵活的 NetFlow 类似，IPFIX 需要几个步骤才能采集到流量流。这些步骤包括：观测点观测、计量过程、输出过程以及最后的采集过程[①]。可将一 IP 流量流视为（量化为）途经某观测点的一组数据包，这组数据包具有某些共同的属性，比如具有共同的：

- 报头字段（IP 地址和端口号等）；
- 数据包字段（多协议标签交换[MPLS]标签等）；
- 数据包处理字段（下一跳 IP 地址和接口等）。

以下内容描述了图 11-5 所示的 IPFIX 流的每一步骤[②]。

观测点

观测点是指可观测到网络中过往数据包之所在，比如，网络设备上的某个接口[③]。一个观测点可能是其他多个观测点的超集。比方说，可将整块线卡（line card）视为一个观测点，这一观测点便是线卡上单个接口观测点的超集。

---

① 以上两句的原文是“Similar to Flexible NetFlow, IPFIX requires the following steps for a traffic flow. These steps include the observation, metering process, exporting process, and finally the collecting process”。

② 原文是“The following sections describe each step of the IPFIX flow shown in Figure 11-5”，译文为直译。

③ 原文是“An observation point is a location in the network point where packets can be observed,such as an interface”。

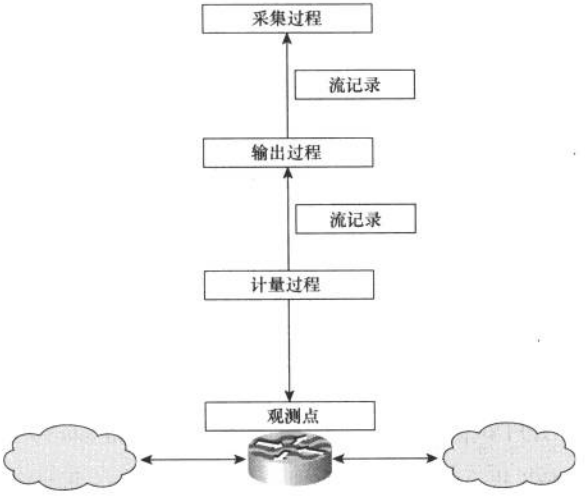


图 11-5 IPFIX 流

计量过程

在观测点，会对感兴趣流量进行分类，并对流量信息做计量处理，以生成相应的流记录。流记录包含的是与特定流（在观测点计量得出）有关的属性和特征，属性和特征均通过测量得出[①]。

在计量过程中，观测点（网络设备的接口）会检查包头及网络设备对数据包的处理方式。在计量过程中，观测点（网络设备的接口）的任务如下：

- 创建新的流记录；
- 更新现有的流记录；
- 计算流的统计信息；
- 将流记录交给输出过程处理；
- 删除流记录。

输出过程

在计量过程中，网络设备会创建一个或多个进程来生成流记录。在输出过程中，网络设备也会创建相关进程去生成 IPFIX 消息，该消息用来携带计

① 原文是“A flow record contains measured properties and characteristics about a specific flow that was metered at an observation point”。

量过程中所创建的流记录。IPFIX 消息由传输层协议封装，然后再交由采集过程处理。

采集过程

在采集过程中，网络设备会创建相应的进程去接收来自一个或多个输出进程的流记录。基于用户的配置，在采集过程中，可存储或进一步处理流记录。

以下增强功能使得 IPFIX 比 NetFlow v9 更为稳定、更加高效。图 11-6 所示为 IPFIX 和 NetFlow v9 数据包头部格式之间的差异。

- 与 NetFlow 数据包相比，由于（报头中）UNIX secs 字段的移除，IPFIX 数据包要更小一些。
- 在 IPFIX 报头中，以的长度字段替换了 NetFlow v9 报头中的 Count 字段。
- IPFIX 报头中的版本号字段为 10，NetFlow v9 报头中的版本号字段值为 9。

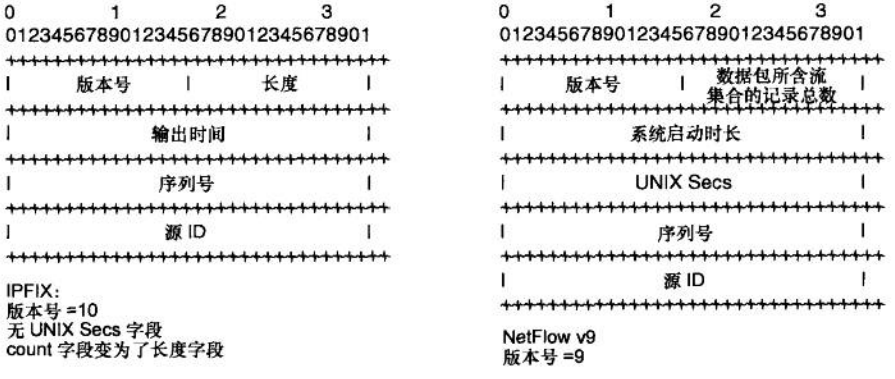


图 11-6 IPFIX 和 NetFlow v9 数据包（报头）格式之间的差异

IP SLA 之于 IPv6

一直以来，网管工作站（NMS）都被部署在网络的中央，从此处针对其他远程站点（或区块）执行连通性测试，并获取与网络管理有关的数据信息。随着网络连接方式的日渐转变，和网络功能的日益增多，无论是网络管理还是网络性能数据的采集都变得愈加复杂。解决该问题的方法之一是，在网络的不同区域部署 NMS 设备或软件，但这却既增加了成本，甚至还提高了网络管理工具（或软件）自身的管理难度。

Cisco IOS SLA（服务等级协定）特性，可令 Cisco 设备去采集并向 NMS

发送网络性能和管理数据。Cisco SLA 特性与第二层无关，可用来提供端到端的有关网络性能的数据（指标），这些数据事关末端用户的体验[①]。

IPSLA 特性通过 SNMP 来发送各项网络性能指标，借此，诸如 CiscoWorks 系能监控器（IPM）之类的性能监控应用程序或其他厂商的性能管理产品，便能获取到这些网络性能指标数据。当网络性能下降到特定等级或性能问题被纠正时，在操作上，Cisco IP SLA 特性可让路由器接收相应的告警信息[②]。外部 NMS 应用程序与 Cisco 设备之间，通过 Cisco RTTMON MIB 来进行有关 IPSLA 的交互。对 Cisco RTTMON MIB 对象变量的完整描述，请参阅 CISCO- RTTMON MIB.my 文件的描述，可从 Cisco MIB 站点获取，链接为：

http://www.cisco.com/public/sw-center/netmgmt/cmtk/mibs.shtml。

虽然对 IPSLA 操作的完整介绍超出了本书的范围，但本章会讨论某些常见场景中与 IPv6 有关的示例。

- ICMP 回显（echo）操作：测量 Cisco 路由器和任何 IP 设备之间的逐跳响应时间。
- TCP 连接操作：测量 Cisco 路由器和任何 IP 设备之间的 TCP 操作响应时间。
- UDP 回显操作：测量 Cisco 设备之间的 UDP 响应时间。
- UDP 抖动操作：利用 UDP 流量，针对 VOIP 的语音质量给出近似的评分，并以此来计算 UDP 数据包到达时间的变化（抖动）[③]。

以下内容是对上述 IPSLA 操作的概述。

ICMP 回显操作

ICMP 回显操作可用来监控 Cisco 路由器和 IP 设备之间端到端的响应时间。排除网络连通性故障时，IPSLA ICMP 回显操作会非常有用。图 11-7 所示为路由器计算数据包往返时间的方法（T2 减去 T1）[④]。按照该方法所计算出的数据

① 原文是“Cisco IOS IP SLAs are Layer 2 independent, which provides end-to-end metrics that an end user is likely to experience”。译文未完全按字面意思翻译，请读者注意。

② 原文是“Cisco IOS IP SLA operations allow the router to receive alerts when performance drops below a specified level and when problems are corrected”。译者认为原文不通，只能勉强翻译。

③ 原文是“Uses UDP traffic to generate an approximate Voice over IP score as well as to calculate jitter variance for the data”。译者改写了原文，为避免误导读者，特给出原文。

④ 原文是“Figure 11-7 shows that the router calculates the round-trip time by calculating T2-T1”。

包往返时间，并不包含目标主机对数据包的处理时间。

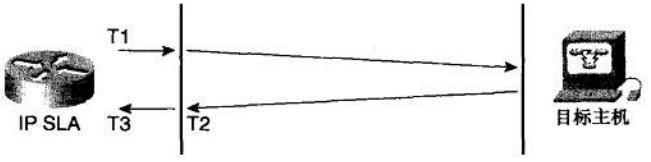


图 11-7 IPSLA ICMP 回显操作

例 11-4 所示为使用 IPv6 时，IP SLA ICMP 回显操作的配置。

例 11-4 Cisco IOS IPSLA ICMP 回显操作配置示例

```
ip sla 10
 icmp-echo 2001:db8:cafe:151::10 source-ip 2001:db8:cafe:121::1
 frequency 300
 request-data-size 28
 tag Gateway_WebServer
!
ip sla schedule 10 life forever start-time now
```

由例 11-4 可知，IPSLA 操作的编号为 10，并且是立即启动，无限期运行。这一 IPSLA 操作会令路由器发送 28 字节长的 IPv6 ICMP echo request 数据包，IP 报头的源和目的地址分别为 2001:db8:cafe:121::1 和 2001:db8:cafe:151::10，每隔 300 秒发送一次。配置 IPSLA 操作标记（编号）的目的是对操作进行识别。

例 11-5 所示为 IPSLA 所捕获到的统计信息。

例 11-5 Cisco IP SLA ICMP 回显统计信息

```
show ip sla statistics 10 details

Round Trip Time (RTT) for        Index 10
Type of operation: icmp-echo
        Latest RTT: NoConnection/Busy/Timeout
Latest operation start time: *01:40:13.920 UTC Tue Jan 18 2011
Latest operation return code: Timeout
Over thresholds occurred: FALSE
Number of successes: 0
Number of failures: 1
Operation time to live: Forever
Operational state of entry: Active
Last time this entry was reset: Never
```

TCP 连接操作

Cisco IPSLA TCP 连接操作的作用是：在传递 IPv6 流量的 Cisco 路由器和设备之间，对执行一次 TCP 连接所花费的时间进行测量。TCP 属于传输层协议，是一种可靠的全双工数据传输机制。目的端设备可以是任何 IPv6 设备或 IP SLA 响应设备。

可利用 IPSLA TCP 连接操作来测试虚电路或应用程序的可用性。如图 11-8 所示，测得的 TCP 连接建立时间为：SYN/ACK 数据包的接收时间减去 SYN 数据包的发出时间，即 T2−T1。

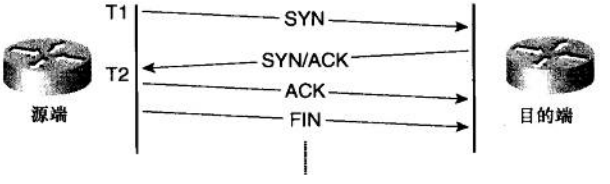


图 11-8 IP SLA TCP 连接操作

在例 11-6 中，目的端是一台进行 Telnet 应用的路由器上（2001:db8:cafe:101::51）。[①]

例 11-6 Cisco IOS IP SLA TCP 连接操作目的端配置示例

```
ip sla responder
```

例 11-7 所示为执行 IP SLA TCP 连接操作的源端路由器配置示例。该操作被设置为立即启动。

例 11-7 Cisco IOS IP SLA TCP 连接操作源端配置示例

```
ip sla 10
 tcp-connect 2001:DB8:CAFE:101::51 23
 frequency 30
ip sla schedule 10 life forever start-time now
```

UDP 回显操作

UDP 回显操作可用在 Cisco 路由器和 IP 网络设备之间，监控端到端的 UDP

[①] 原文是“In Example 11-6, we use a host (2001:db8:cafe:101::51) running a Telnet application at the destination router.”这完全是不知所云，译文酌改。

响应时间，这与 ICMP 回显操作非常相似。可使用 UDP 回显操作去测量 UDP 响应时间和端到端的 UDP 连通性。许多应用程序都利用 UDP 来提供 IP 服务。在支持 RFC 862（echo 协议）的任何一台网络设备之间，都可以使用 UDP 回显操作，但我们强烈推荐采用 Cisco 网络设备作为该操作的目的端设备。

图 11-9 所示为在源路由器和响应路由器之间测量 UDP 回显延迟（时间）。T5 减 T4 可算作源端路由器的处理延迟，T3−T2 可算作目的端路由器的处理延迟。UDP 回显的总延迟时间为 T2+T4−T1−T3。在负载较重的接口上执行 UDP 回显操作可能会有些误差，但可以忽略不计[①]。

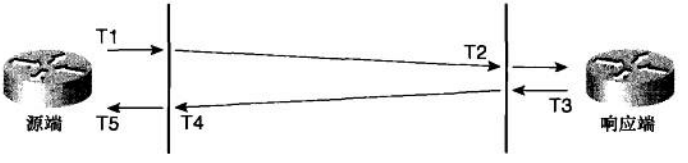


图 11-9 IPSLA UDP 回显操作

在响应端路由器上，可执行 **ip sla responder** 命令来完成 IPSLA UDP 回显操作（如图 11-9 所示）的配置。例 11-8 所示为源端路由器的配置。

例 11-8 Cisco IOS IP SLA UDP 回显操作源端路由器配置示例

```
ip sla 10
 udp-echo 2001:DB8:CAFE:101::51 5000
 frequency 5
ip sla schedule 10 life forever start-time now
```

UDP 抖动操作

可使用 UDP 抖动操作来监控实时性应用程序，比如，IP 上的语言（VoIP）、IP 上的视频，或实时性的视频会议等，如图 11-10 所示。单这一项操作便能同时捕获到双向传输数据包的延迟、抖动以及丢失等数据[②]。在操作的输出中会包含一份全面的信息报告。此外，还可利用 UDP 抖动操作，来模拟即将提供的 IPv6 服务数据。

抖动测量是指当数据包从某源发往某目的地时，对其延迟的变化进行测量。在连续发送多个数据包（比如，VoIP 流）时，对一个稳定的网络来说，所接收

① 原文是“On loaded interfaces, queuing delay might become a problem; otherwise, it is negligible”。

② 原文是“This single operation can capture at the same time the delay, jitter, and loss in both directions”。

的数据包在“节奏”上应保持不变，如图 11-10 中端到端的 IP 电话流量。若网络中存在延迟，在接收端，数据包到达的间隔时间可能会随之增加。这也被称为正向抖动（positive jitter），即接收端所收数据包的时间间隔会增大。反之，则被称为逆向抖动（negative jitter）。那些“天生”对延迟敏感的应用程序数据流，如 VoIP 流量，正向抖动是绝对不可接受的，理想情况应该是“零抖动”[①]。

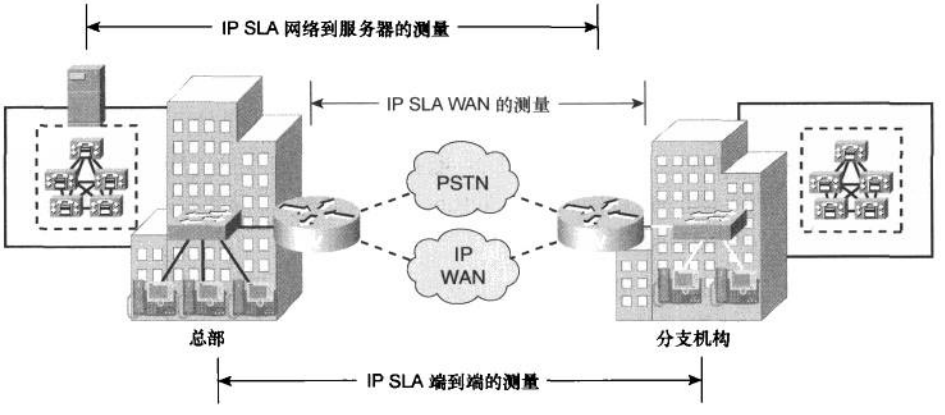


图 11-10 SLA UDP 抖动操作的运用

UDP 抖动操作会模拟 UDP 流量。默认情况下，路由器会每隔 10ms 生成 10 个 UDP 数据包（数据净载为 10 字节），这一操作会每分钟重复一次。用户可通过配置来修改上述参数。在源端路由器上配置 UDP 抖动操作之前，必须先在目的端设备上激活其 IPSLA 响应端设备的功能。在此之后，Cisco IOS 才能支持该功能。例 11-9 所示为 IP SLA UDP 抖动操作的配置。

例 11-9 Cisco IOS IP SLA UDP 抖动操作源路由器（router B）配置示例

```
ip sla 10
udp-jitter 2001:DB8:CAFE:101::51 65051
ip sla schedule 10 life forever start-time now
```

① 译者认为作者对抖动的解释并不全面，译文并未完全按字面意思翻译，为避免译文误导读者，现给出整段原文“Jitter is the measure of variance in delay of a packet sent from the source to the destination. When multiple packets—for example, a VoIP stream—are sent consecutively, the timing in receiving the packets should be constant in a stable network, as seen in Figure 11-10 between end-to-end IP phones. If there are delays in the network, the arrival delay between the packets might increase. This is known as a positive jitter, which indicates that the time between receiving the packets is increasing. The reverse is known as negative jitter. In a VoIP stream, which by nature is a delay-sensitive network application, a positive jitter is undesirable and a value of 0 is ideal”。

例 11-10 所示为 **show ip sla detail** 命令的输出。该输出提供了一份与 UDP 抖动指标有关的综合性报告，这些抖动指标如下所示。

- Delay（延迟）：UDP 抖动操作执行完毕之后，路由器会通过 delay 值来汇报数据包往返时间延迟，可以微秒或毫秒为单位（视配置而定）。
- Jitter（抖动）：**show ip sla detail** 命令的输出会给出事关正/负向抖动的大量输出，这些输出非常容易理解。
- Packet（数据包）：使用 IPSLA 可测量出 4 类数据包的丢失或同化事件（assimilated event）。
- Packet loss（数据包丢失）：数据包丢失可能发生在从源发往目的地的途中（在 MIB 中作为 packetLossSD 来通报），或反方向（在 MIB 中作为 packetLossDS 来通报）。
- Tail drop（尾丢弃）：知道发生了丢包现象，但不知发生在哪个方向上。当测试流的最后一个（或多个）数据包被丢弃时，便会发生这种情况，其原因是 UDP 抖动操作并不涉及序列号。在旧版本中，这被称为数据包失踪现象，简称 Packet MIA。在 MIB 中，PacketMIA 这一称谓仍在使用。
- Packet late arrival（数据包延迟到达）：数据包虽能到达，但“姗姗来迟”，以至于高层应用程序认为其已被丢弃（或认为其收到时已为时已晚）。请考虑一下 VoIP 应用。若接收端迟于预期才收到某语音包，对于连续进行的语音通话来说，未免也太过晚矣。该语音包将被视作丢弃。
- Packet misordering（数据包失序）：数据包未按发送顺序抵达。应用程序可能将这样的数据包视为丢弃，也可能不这么认为（packetOutOfOrder）。

例 11-10 Cisco IOS IP SLA UDP 抖动统计信息

```
c38d14-1# show ip sla sta 1 details
IPSLAs Latest Operation Statistics
IPSLA operation id: 10
Type of operation: udp-jitter
        Latest RTT: 1 milliseconds
Latest operation start time: 15:16:59.005 UTC Thu Jan 5 2010
Latest operation return code: OK
```

（待续）

```
RTT Values:
        Number Of RTT: 1000              RTT Min/Avg/Max: 1/1/2 milliseconds
Latency one-way time:
        Number of Latency one-way Samples: 0
        Source to Destination Latency one way Min/Avg/Max: 0/0/0 milliseconds
        Destination to Source Latency one way Min/Avg/Max: 0/0/0 milliseconds
        Source to Destination Latency one way Sum/Sum2: 0/0
        Destination to Source Latency one way Sum/Sum2: 0/0
Jitter Time:
        Number of SD Jitter Samples: 999
        Number of DS Jitter Samples: 999
        Source to Destination Jitter Min/Avg/Max: 0/1/1 milliseconds
        Destination to Source Jitter Min/Avg/Max: 0/1/1 milliseconds
        Source to destination positive jitter Min/Avg/Max: 1/1/1 milliseconds
        Source to destination positive jitter Number/Sum/Sum2: 29/29/29
        Source to destination negative jitter Min/Avg/Max: 1/1/1 milliseconds
        Source to destination negative jitter Number/Sum/Sum2: 29/29/29
        Destination to Source positive jitter Min/Avg/Max: 1/1/1 milliseconds
        Destination to Source positive jitter Number/Sum/Sum2: 22/22/22
        Destination to Source negative jitter Min/Avg/Max: 1/1/1 milliseconds
        Destination to Source negative jitter Number/Sum/Sum2: 22/22/22
        Interarrival jitterout: 0        Interarrival jitterin: 0
        Over thresholds occurred: FALSE
Packet Loss Values:
        Loss Source to Destination: 0            Loss Destination to Source: 0
        Out Of Sequence: 0       Tail Drop: 0    Packet Late Arrival: 0
Packet Skipped: 0
Voice Score Values:
        Calculated Planning Impairment Factor (ICPIF): 0
        Mean Opinion Score (MOS): 0
Number of successes: 120
Number of failures: 0
Operation time to live: Forever
Operational state of entry: Active
Last time this entry was reset: Never
```

利用嵌入式事件管理器灵活编程功能的自动化网络管理

EEM（嵌入式事件管理器）是一种灵活的基于策略的网络管理框架，可通过编程来实现。有了 EEM，网管人员便可根据一组特定事件的发生情况，来定制脚本，执行相应的网络管理动作。图 11-11 所示为 Cisco Catalyst 和 Cisco Nexus 交换机中的 EEM 框架流程图。

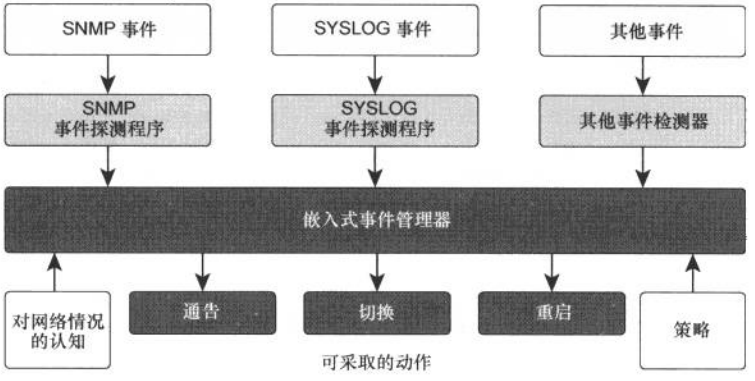


图 11-11 Cisco Catalyst 6500 中的 EMM 流程图

以下对 EMM 的基本运作方式进行了总结。

步骤 1 某系统事件发生，被事件探测程序捕获。该系统事件可以是由系统生成的一条 syslog 消息——通过 CLI 执行某条命令、某给定的计数器超出了预设的阈值，或有新的线卡插入设备时，都会致使系统生成 syslog 消息。EEM 集众多事件探测程序为一体，这些事件探测程序可算作其子系统进程，分别用来监控重要事件的发生。

步骤 2 给定的事件探测程序将会向相关 EEM 子系统（进程）发出警报，并向其传递与该事件有关的信息。在网络设备上，会启动一个由网管人员事先定义，且已向 EEM 注册过的脚本，这一脚本会利用上述事件信息，去执行预先安排好的动作。

EEM 提供了以下两种类型的脚本，可利用这两种脚本，并根据预计的事件发生情况，来作为执行动作的标准。

- **Applet**：此类脚本通过 CLI 运行。
- **TCL 脚本**：Catalyst 6500 EEM 子系统对 TCL 的支持基于的是 TCL 版本 8.3.4。这与 Cisco IOS 中 TCL shell 所使用的 TCL 版本一致。

虽然基于 Applet 的脚本提供了一个简单的选项，可利用该选项将脚本加载进交换机，但借助于 TCL，仍可编写出更加灵活、功能更为强大的脚本。本节会对这两种脚本展开深入讨论。

EEM 架构由两个层次组成，TCL 脚本正是运行在这两个层次之上。用户编写的脚本可能会无意间访问到系统资源，从而破坏系统的完整性，借助于这两个层次，便能够保障交换机系统的安全性。

Cisco 硬性脚本（Cisco mandatory script）运行于 TCL 完全模式。该模式提供了对所有交换机资源的完全访问。

而用户编写的脚本（User-built script）则运行于 TCL 安全模式。在 TCL 安全模式的脚本操作中，脚本全都在安全解释器（safe interpreter）内运行，与其他应用保持隔离。在 TCL 安全模式中执行的脚本受控于主解释器（master interpreter），由主解释器对运行中的脚本所发出的服务请求进行控制。TCL 安全模式可根据需求，让 Cisco 设备去逐条屏蔽或定制 TCL 命令，从而提供了一种手段，来防止系统受有瑕疵脚本（runaway script）的危害。此外，该模式还可让网管人员，从 CLI 执行单条命令，去禁用基于用户的脚本。修改 Cisco 硬性策略（Cisco mandatory policy）也不是不可以，但用户需要将经过修改的策略转移到用户目录，并令其在安全模式运行。Cisco 还对那些调用 CLI 命令的脚本实施了额外的安全措施。EEM 甚至还提供了一条指明 IOS user-id 的命令，可让 TCl 脚本去使用 TACACS+命令授权服务[①]。

对于安全模式下的 TCL 脚本编程来说，使用环境变量是常有的事儿。这种环境变量是指定义于 TCL 脚本之外的全局变量，但可在脚本内引用。环境变量可作为一个有用的工具，来帮助脚本去了解其所运行的系统，以及触发事件的环境。有以下几种类型的环境变量。

- **用户定义的环境变量**：由用户定义。
- **Cisco 定义的环境变量**：由 Cisco 定义，或针对某条特定的策略而创建。

在交换机上，环境变量由事件管理器的“environment（环境）”命令来定义。所有由 Cisco 定义的环境变量都以“_”开头。“_”为预留字符，用户在定义变量时不能使用。可直接使用若干 Cisco 定义的环境变量，具体用法请见 IOS 手册。欲了解与 EEM 脚本有关的信息，请访问 Cisco 官网，连接为：http://tinyurl. com/2dvjvnu。

11.4 IPv6 网络管理

表 11-8 所列为 IPv6 网络管理的三个主要环节（监控和报告、网络服务、访问控制和运维）所涉及的协议[②]。为了协助用户从纯 IPv4 管理网络，迁移至混

---

① 原文是“EEM provides a command that allows the specification of the IOS user-id, allowing a TACACS+ command authorization service to be used”。鉴于作者的笔法，此句译文仅供参考，译者不能保证其正确性。

② 原文是“Table 11-8 outlines the three main categories of IPv6 network management transport: monitoring and reporting, network services, and access control and operations”。原文真是烂得够呛，译文酌改。

合的双栈 IPv4/IPv6 管理管理，Cisco 为旗下的 IOS、Nexus OS、IOS-XE 以及 Cisco IOS-XR，新增了大量特性和功能。

表 11-8 对 IPv6 网络管理的总结

IPv6 网络管理	所用协议[①]
监控和报告	SNMP
	SYSLOG
	ICMP
网络服务	TFTP
	NTP
访问控制和运维	Telnet
	SSH
	HTTP

对 Cisco IOS 软件来说，IPv6 路由功能是默认禁用的。要想利用 IPv6 来管理网络，为获取网络管理信息，在网元之间必须先行建立 IPv6 连通性[②]。

11.4.1 监控和报告

以下内容将讨论各种网络监控和报告提供的手段。在本章先前的内容中，已详细介绍了 MIB，本节将会讲述如何配置 SNMP 去访问 MIB。SYSLOG 为网管人员提供了识别网络事件的能力，本节亦会对此做相应的介绍。最后，还会讨论如何利用 ICMP 来验证网络的连通性。

IPv6 上的 SNMP

IPv6 主机可针对 Cisco 设备执行 SNMP 查询，并从其接收 SNMP 通告[③]。

如本章对 MIB 的介绍一节所述，为了支持 IPv6，Cisco 已对集成进路由器/交换机内的 SNMP 代理和 MIB 功能，进行了改进。可启用更为安全的 SNMP 功能，甚至还可利用 3DES 和 AES 来加密 SNMP 数据包。

SNMP 团体字串对于基本的 SNMP 配置来说必不可缺。出于安全方面的考虑，可配置 ACL，对使用团体字串访问 Cisco 设备的主机的 IP 地址加以限制，

① 原文居然是“transport（传输）”，真太长见识了。

② 原文是“To use the IPv6 management features, an IPv6 communication path must be available between the source and destination for the network management information”。原文惨不忍睹，译文酌改。

③ 原文是“An IPv6 host can perform queries and receive SNMP notifications from a Cisco IOS device”。原文不通，译文酌改。

此外，还可配置 community，对访问 MIB 的读写权限进行限制。

例 11-11 所示为 IP 地址为 2001:db8:cafe:151::51 的主机将使用团体字串“public”，从路由器接收 OSPF SNMP 通告。

例 11-11 IPv6 上的 SNMP 配置

```
snmp-server community public RW
snmp-server enable traps ospf
snmp-server host 2001:DB8:CAFE:151::51 public
```

IPv6 上的 Syslog

运行于 Cisco 交换机/路由器上的 Syslog 进程，可让用户将 Syslog 信息记录进设有 IPv6 地址的服务器或主机。自 Cisco IOS Release 12.4(4)T 和 12.2(33)SRC/12.2(33)SXH 版本起，在 Cisco 路由器/交换机上，可使用 IPv6 地址去配置 Syslog 服务器。IOS 上的 Syslog 配置如下所示。

```
logging host ipv6 2001:DB8:CAFE:151::62
```

Cisco NX OS 上的 Syslog 配置如下所示。

```
logging server 2001:DB8:CAFE:151::62
```

ICMPv6

就运作方式而言，ICMPv6 和 IPv4 ICMP 基本相同。ICMP 可生成信息性消息和错误性消息，在 IPv6 中，还会利用 ICMPv6 数据包来完成 IPv6 邻居发现过程、执行 MTU 发现以及多播侦听者发现功能等。图 11-12 所示为 ICMPv6 消息的格式。在 IPv6 报头的“下一个报头”字段中，以值 58 来标识 ICMPv6 数据包。

例 11-12 所示为 ping 命令的示例。

例 11-12 IPv6 ping 命令的输出

```
Router# ping ipv6 2001:DB8:CAFE:155::6509

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 2001:DB8:CAFE:155::6509, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 0/0/4 ms
Router#
```

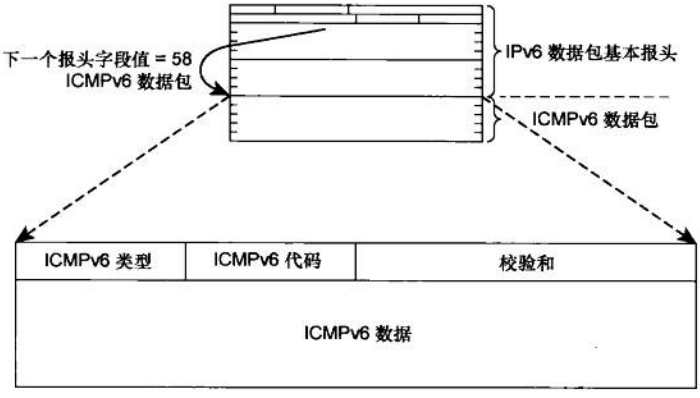


图 11-12 ICMPv6 数据包格式

执行故障排除时，还可以观察 ICMPv6 的 debug 信息，如例 11-13 所示。

例 11-13 IPv6 ICMP debug

```
Router# debug ipv6 icmp
6w3d: ICMPv6: Sending echo request to 2001:DB8:CAFE:155::6509
6w3d: ICMPv6: Received echo request from 2001:DB8:CAFE:152::3550
6w3d: ICMPv6: Sending echo reply to 2001:DB8:CAFE:152::3550
6w3d: ICMPv6: Received echo reply from 2001:DB8:CAFE:151::6513
6w3d: ICMPv6: Received echo reply from 2001:DB8:CAFE:155::6509
6w3d: ICMPv6: Sending echo request to 2001:DB8:CAFE:155::6509
```

11.4.2 网络服务

本节将介绍几种与 IPv6 有关的网络服务，Cisco 网络设备可利用这些服务去行使某些功能，比如，上传/下载软件映像文件和配置文件，以及时钟同步等。

TFTP

Cisco 设备支持通过 IPv6 来上传/下载 IOS 映像文件，其运作方式与 IPv4 类似。比如，在 Cisco 设备上，可使用以下命令来备份运行配置。

```
copy running-config tftp://2001:DB8:CAFE:151::51/running-config
```

在基于 NX-OS 的网络设备（比如，MDS 9500 导向器级交换机）上，对 TFTP 和其他应用的 IPv6 支持无异于 Cisco IOS 设备。

NTP

网络时间协议（NTP）是一种用来同步各种网络设备时钟的协议。NTP 可提供精确而又可靠的参考时间，这对金融和法律事务来说都是必不可缺的。此外，排除网络故障时，还可使用 NTP 来确定故障发生时间。所有基于 Cisco IOS 和 NX OS 的设备都支持 NTP。

NTP 运行于 UDP 之上，而 UDP 则运行于 IP 之上。在网络中，部署 NTP 方法通常是，让一台（多台）NTP 服务器从某个 NTP 权威时间源获取时间，然后由其在整个网络内发布时间。

NTP 版本 4（NTPv4）是 Cisco IOS 的特性之一。NTPv4 除了能够支持 IPv6 以外，还可提供老版本 NTP 所没有的其他增强功能。NTPv4 最重要的增强功能之一便是对特定多播组的使用，即利用 IPv6 本地站点多播地址（去发布时间信息）。这使得 NTP 服务器具备自动配置功能，从而能够以极低的带宽成本，提供最精确的时间。

在 Cisco NX-OS 设备上，有关 IPv6 NTP 服务器和对等体的配置与 IPv4 几乎相同。

```
ntp server 2001:0DB8:151:0:0:0:0:3555 version 4
ntp peer 2001:0DB8:151:0:0:0:0:6509 version 4
```

以下给出了利用 IPv6 本地站点多播地址时的 NTPv4 配置。

```
interface fastethernet 0/1
ntp multicast FF02::1:FF0E:8C4E
```

11.4.3 访问控制及运维

本节将讨论如何利用 IPv6 去访问并操作网络设备，将会介绍 Telnet、SSH 以及 HTTP。

Telnet

要想能够 Telnet 到 Cisco IOS 设备，须先创建 VTY 接口并设置密码。Cisco IOS 设备既可以发起 IPv6 Telnet 会话，也可以作为 Telnet 服务器被 IPv6 设备访问。在 Cisco IOS 设备上，IPv6 Telnet 特性的运作方式与 IPv4 几乎相同。例 11-14 所示为利用 IPv6 进行 Telnet 访问的配置示例。

例 11-14 IPv4 Telnet 配置

```
ipv6 unicast-routing
!
interface gigabitethernet1/1
 ipv6 address 2001:DB8:CAFE:155::6509/64
 ipv6 enable
!
line vty 0 4
 login
 password cisco
```

要想验证配置是否生效，请在 Telnet 客户端（主机或另一台路由器）上针对该设备发起 Telnet 连接，如例 11-15 所示。

例 11-15 使用 IPv6 Telnet 路由器

```
Router_A# telnet 2001:db8:cafe:155::6509
Trying 2001:DB8:zCAFE:155::6509 ... Open


User Access Verification

Password:
Router_B>en
Password:
Router_B#
```

show sessions 命令的输出将会给出与 IPv6 Telnet 连接有关的信息，如例 11-16 所示。

例 11-16 显示 IPv6 Telnet 会话的 show sessions 命令

```
show sessions

Conn Host                    Address                 Byte   Idle Conn Name
*   1 2001:0db8:20:1::12 2001:0db8:20:1::12         0      0 2001:0db8:20:1::12
```

SSH

Cisco 路由器和交换机集 SSH 客户端和服务器特性于一身，可借此来创建安全的加密管理连接。此外，Cisco 路由器和交换机还支持 SSHv2，并支持 SSH debug 增强功能和 VRF 感知的 SSH 特性。与 IPv4 一样，在 Cisco 路由器/交换

机上配置 IPv6 SSH 时，也需满足以下先决条件。

- 要么配置本地用户名/密码，要么配置 AAA（认证、授权、记账）。
- 必须配置域名。
- 需生成 SSH 密钥。

配毕 SSH 之后，还可使用 SCP 来拷贝映像和配置文件。例 11-17 所示为 IOS 中的 IPv6 SSH 配置。

例 11-17 IOS IPv6 SSH 配置

```
ipv6 unicast-routing
aaa new-model
user cisco password cisco
interface gigabitethernet1/1
 ipv6 address 2001:DB8:CAFE:155::6509/64
 ipv6 enable
 exit
ip domain-name cisco.com
crypto key generate
line vty 0 4
 transport input ssh
```

例 11-18 所示为 NX OS 中的 SSH 配置。

例 11-18 NX OS IPv6 SSH 配置

```
ipv6 routing
ssh key rsa
feature ssh
```

HTTP

运行于 Cisco IOS 中的 HTTP 服务器可同时为 IPv4 和 IPv6 HTTP 客户端提供服务，反之，运行于 Cisco IOS 中的 HTTP 客户端也可同时向 IPv4 和 Ipv6 HTTP 服务器发送服务请求。使用 HTTP 客户端时，必须使用 RFC 2732 所列的规则来书写包含 IPv6 地址的 URL，即以“[]”来括起 URL 中的 IPv6 地址部分。比方说，对于包含 IPv6 地址 2001:db8:cafe:1001::4507 的 URL，应该以 http://[2001:db8:cafe:1001::4507]的方式来书写。

图 11-13 所示为使用 IPv6 地址发送 HTTP 请求。

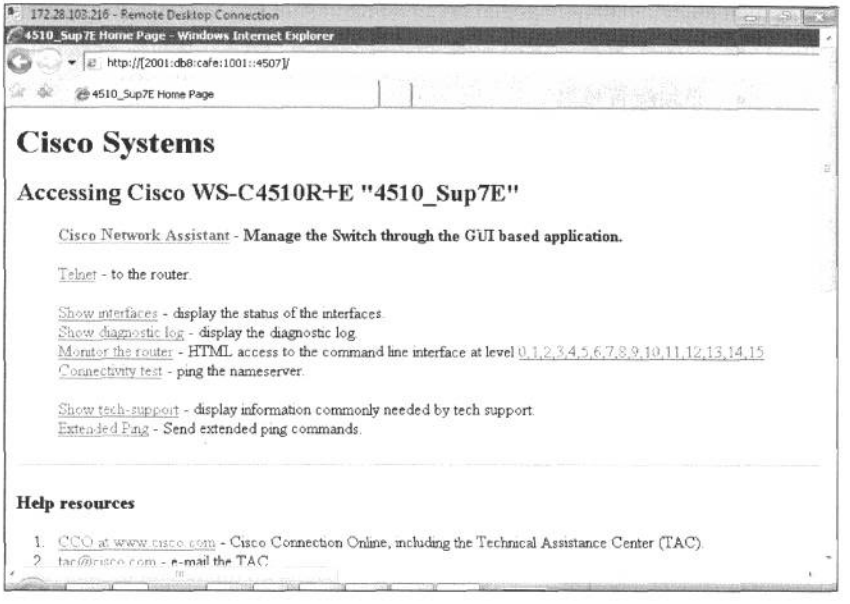


图 11-13 使用 IPv6 HTTP 访问 Cisco 设备

11.5 IPv6 流量监控工具

以下各节将介绍几种 IPv6 流量监控工具。本节将首先介绍流量捕获工具，其捕获到的流量则可作流量分析之用。接下来，将介绍 Catalyst 6500 交换机自身携带的 mini 协议分析器，可利用其分析由 6500 交换机所捕获到的流量。最后，本章将讨论如何在 Nexus 7000 VLAN 层面进行流量分析。

11.5.1 SPAN、RSPAN 和 ERSPAN

交换机端口分析仪（SPAN）是一种高效、高性能的流量监控系统。当流量穿越交换机时，SPAN 会将其复制到一个或多个监控接口。该特性在排除网络连通性故障及计算网络带宽利用率等方面得到了广泛运用。Cisco 产品（包括 Cisco Catalyst 6500、4500E、3750 以及 Cisco Nexus 交换机）支持三种类型的 SPAN，如图 11-14 所示。

- **本地（机）SPAN：**在同一台交换机上的接口之间（可以是一对一、一对多、多对一或多对多）镜像流量。

- **远程 SPAN（RSPAN）**：在交换机上的一个或多个接口和某个特殊的 RSPAN VLAN 之间镜像流量，被镜像的流量将会跨越第二层网络，送达另外一台（或多台）交换机。收到镜像流量的交换机会将其从 RSPAN VLAN 复制到自身的一个或多个接口。
- **封装式远程 SPAN（ERSPAN）**：在交换机上的一个或多个接口和 IP GRE 隧道接口之间镜像流量。被镜像的流量可跨任意类型的第三层网络发送。只有 Catalyst 6500 和 Nexus 7000 系列交换机才支持该特性。

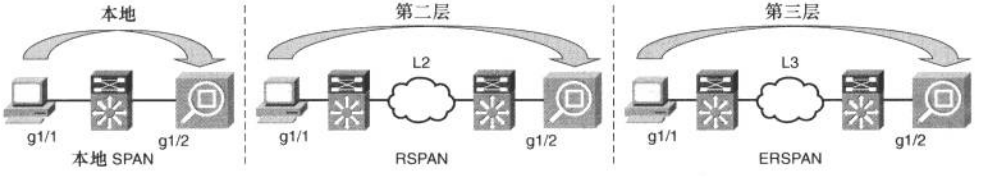


图 11-14 SPAN 的类型

以下各节将会给出创建 SPAN、RSPAN 和 ERSPAN 的配置示例。此外，读者还会学习到如何利用 mini 协议分析器，排除 Catalyst 6500 交换机故障。

配置各类 SPAN

本地（机）SPAN 的配置由两条 monitor session 命令组成，分别指明 SPAN 源端口和目的端口，如例 11-19 所示。该例所示为在 Cisco IOS Release 12.2(18)SXF 及以前版本中，配置 SPAN 的命令。

例 11-19 Cisco IOS Release 12.2(18)SXF 及以前版本中本地 SPAN 的配置

```
monitor session 1 source int fa 4/1
monitor session 1 destination int fa 2/2
```

在 Cisco IOS 12.2(33)SXH 版本中，配置 SPAN 的命令有所改变，如例 11-20 所示。

例 11-20 Cisco IOS 12.2(33)SXH 及后续版本中，本地 SPAN 的配置

```
monitor session 1 type local
 source int fa 4/1
 destination int fa 2/2
```

配置 RSPAN 时，请牢记以下注意事项。

- 通过 RSPAN 会话传递受监控流量时，会丢弃流量原来的 VLAN ID 值。
- 使用 RSPAN 捕获不到 BPDU 级别的数据帧。

例 11-21 和例 11-22 所示为在两台交换机上，利用一个专用的 RSPAN VLAN，配置 RSPAN 的方法。

例 11-21 RSPAN 示例——交换机 1

```
vlan 120
 remote-span
 exit
monitor session 1 source int fa1/1
monitor session 1 destination remote vlan 120
```

例 11-22 RSPAN 示例——交换机 2

```
vlan 120
 remote-span
 exit
monitor session 1 source remote vlan 120
monitor session 1 destination interface int fa 3/12
```

配置 ERSPAN 时，建议在源会话端和目的会话端 ERSPAN 设备上，使用专用的 loopback 接口来建立会话。ERSPAN 在 SPAN 子配置模式中进行配置。在 ERSPAN 源会话端设备上，需定义 ERSPAN 源端会话。只有在定义了 SPAN 源、目的端 ERSPAN 会话 IP 地址、ERSPAN ID 以及一个“origin ip address”之后，系统才会激活 ERSPAN 源端会话。默认情况下，ERSPAN 隧道为“shutdown”模式，在退出 SPAN 配置模式之前，需执行 **no shutdown** 命令激活 ERSPAN 隧道。以下为 ERSPAN 会话源端的配置示例[①]。

ERSPAN 会话源端配置示例。

```
monitor session 1 type erspan-source
source interface GigabitEthernet1/0/1 rx
source interface GigabitEthernet1/0/4 - 8 tx
source interface GigabitEthernet1/0/3
destination
erspan-id 100
ip address 10.10.0.1
ip prec 5
ip ttl 32
origin ip address 10.1.0.1
```

① 原文其实并未给出任何 ERSPAN 配置示例，作者未完成的事业，就由译者来完成吧，配置示例摘自 Cisco 官网。

ERSPAN 会话目的端配置示例。

```
monitor session 2 type erspan-destination
destination interface GigabitEthernet1/3/2
destination interface GigabitEthernet2/2/0
source
erspan-id 100
ip address 10.10.0.1
```

注意 写作本书之际，配置 ERSPAN 时，还只能使用 IPv4 地址。

Mini 协议分析器

可在 Catalyst 6500 系列交换机上，利用 mini 协议分析器，来排除网络故障。Mini 协议分析器能够先从 SPAN 会话捕获流量，然后，再在交换机的本机内存缓冲区（local memory buffer）内存储捕获到的数据包。最后，网管人员既可在交换机的控制台上“观看”捕获到的数据包，也可以将其存储进交换机的本机文件系统，甚至还可以将数据包输出至外部服务器。

例 11-23 所示为在 Catalyst 6500 交换机上与 mini 协议分析器有关的配置。

例 11-23 Catalyst 6500 交换机 mini 协议分析器的配置

```
monitor session 2 type capture
buffer-size 65535
 source interface Fa4/2
```

试举一例，首先，配置数据包生成器，令其生成源和目的地址分别为 2001:db8:cafe:155::4001 和 2001:db8:cafe:155::6509 的 IPv6 数据包。然后，让数据包生成器将此类数据包发送给一台 Catalyst 6500 交换机的 FastEthernet 4/2 接口[①]。

对于本例[②]，先配置交换机，启动监控捕获会话（执行启动监控捕获会话的命令时，带捕获 100 个数据包的配置选项，在捕获到 100 个数据包后，监控捕获会话将自动停止，交换机会在控制台上生成提示信息）。紧接着，让数据包生成器生成 100 个上面提到的那类数据包。如此一来，在控制台上，执行 **show monitor capture buffer detail dump** 命令，便可列出捕获到的数据包。在 Catalyst

① 数据包生成器就连接在这台交换机的 FastEthernet 4/2 接口上。

② 本段开头原文还有一句“To turn on the capture of packets on the switch, the monitor capture session has to be started and stopped”。翻译过来为“为了在交换机上捕获到数据包，必须先启动、再停止监控捕获会话。”译者认为，此处横插一句只会误导读者，故而自作主张将其从正文中剔除。

交换机上所捕获到的数据包的内容都以十六进制的形式来显示。例 11-24 中的 **show** 命令输出高亮显示了捕获到的数据包的内容（以十六进制值显示），其中包括了数据包的源和目的 IPv6 地址，以及数据包的其他内容。

例 11-24 Cisco IOS Mini 协议分析器输出示例

```
6509_A# monitor capture start for 100 packets
6509_A#

*Mar  8 01:39:27.012: %SPAN-5-PKTCAP_START: Packet capture session 2 started

*Mar  8 01:39:55.913: %SPAN-5-PKTCAP_STOP: Packet capture session 2 ended as the
specified number of packets are captured, 100 packets captured
6509_A# show monitor capture buffer detail dump nowrap 256

 1       Arrival time : 01:39:42.997 UTC Mon Mar 8 2010
         Packet Length : 60 , Capture Length : 60
         len 60   ,   0000.0300.0100   0000.0300.0000   86DD
6030000000063BFF20010DB8CAFE01550000000000004001200010DB8CAFE01550000000000006509000
102030405
 2       Arrival time : 01:39:42.997 UTC Mon Mar 8 2010
         Packet Length : 60 , Capture Length : 60
         len 60   ,   0000.0300.0100   0000.0300.0000   86DD
6030000000063BFF20010DB8CAFE01550000000000004001200010DB8CAFE01550000000000006509000
102030405
 3       Arrival time : 01:39:42.997 UTC Mon Mar 8 2010
         Packet Length : 60 , Capture Length : 60
         len 60   ,   0000.0300.0100   0000.0300.0000   86DD
6030000000063BFF20010DB8CAFE01550000000000004001200010DB8CAFE01550000000000006509000
102030405
 4       Arrival time : 01:39:42.997 UTC Mon Mar 8 2010
         Packet Length : 60 , Capture Length : 60
         len 60   ,   0000.0300.0100   0000.0300.0000   86DD
6030000000063BFF20010DB8CAFE01550000000000004001200010DB8CAFE01550000000000006509000
102030405
 5       Arrival time : 01:39:42.997 UTC Mon Mar 8 2010
         Packet Length : 60 , Capture Length : 60
         len 60   ,   0000.0300.0100   0000.0300.0000   86DD
6030000000063BFF20010DB8CAFE01550000000000004001200010DB8CAFE01550000000000006509000
102030405
 6       Arrival time : 01:39:42.997 UTC Mon Mar 8 2010
         Packet Length : 60 , Capture Length : 60
         len 60   ,   0000.0300.0100   0000.0300.0000   86DD
6030000000063BFF20010DB8CAFE01550000000000004001200010DB8CAFE01550000000000006509000
102030405
```

11.5.2 利用 VACL（VLAN 访问列表）捕获数据包

VACL 数据包捕获特性提供可了一种更具粒度的网络流量分析方法。利用该特性，可过滤并转移 VLAN 中的流量。与 SPAN 端口不同，该特性并不具备选择出站或入站流量的能力，而是作用于所有匹配过滤标准（过滤标准由访问列表中的 ACE 指明）的流量。目前，Catalyst 6500 和 Nexus 7000 平台，都支持 VACL 数据包捕获功能。表 11-9 所列为 VACL 与 SPAN 之间的某些差异。

表 11-9 VACL 数据包捕获特性与 SPAN 之间的差异

VACL	SPAN
根据 ACE 来捕获流量，能够有选择地捕获流量	捕获单个或多个源接口的所有流量
只有 Nexus OS 才支持利用 VACL 捕获 IPv6 流量	Cisco 所有交换机平台均支持 SPAN
以硬件方式来执行	SPAN/RSPAN 需借用交换机的 ASIC 硬件资源来监控数据包的数据信息
只能配置一个会话	SPAN 要启用多个会话
匹配 VACL 捕获的流量会被发送到所有配置了流量捕获的端口	从一个接口捕获流量要占用一个 SPAN 会话，将捕获到的流量发送到另一个接口还得占用另一个会话[①]

配置 VACL 数据包捕获特性的步骤如下所示。

- 配置一个 IP 访问列表。
- 将该 IP 访问列表与 VACL access-map 相关联。
- 将 VACL access-map 应用于相应的 VLAN。
- 配置捕获端口。

例 11-25 所示为利用 VACL 捕获 IPv6 数据包的配置。

例 11-25 NXOS 上利用 VACL 捕获 IPv6 数据包的配置[②]

```
ipv6 access-list vacl
10 permit tcp any any eq www
vlan access-map vacl 10
match ipv6 address vacl
action forward
vlan filter vacl vlan-list 200
int fa3/30
switchport capture allowed vlan 200
switchport capture allowed vlan 200
switchport capture
```

① 原文是“Different SPAN sessions enable capturing traffic from one interface and sending it to another interface”。作者的文字太过深奥，译文仅供参考。

② 原文未给“捕获端口”的配置，译者只好替作者给出。

11.6 总结

IPv6网络管理是本章的重点内容[①]。本章简要介绍了FCAPS管理框架，并总结了该管理框架与各种网络管理应用程序之间的对应关系。网元[②]大都内置有MIB这样的工具（instrumentation），Cisco及其他厂商的网络管理应用程序能藉此管理到这些网元。Cisco交换机和路由器普遍支持NetFlow、IP SLA以及EEM特性，利用这些特性，不但能轻易地隔离网络故障，而且还能更好地窥“网络之全豹”。

对网关人员来说，NetFlow的作用有三：其一，可通览网络中过往的流量；其二，可确定安全事件或网络变更对网络使用方面的影响；其三，可藉NetFlow所收集到的信息，改进网络的性能。只有NetFlow版本9支持IPv6，IETF也将其发展成了IETF标准——IPFIX。

IPSLA是一种能模拟出不同流量模式的探测机制，该机制可帮助网管人员在全网范围内隔离网络故障。主要的IPv6 SLA探测操作包括IP回显操作、UDP回显操作、TCP连接操作和UDP抖动操作，这几种SLA探测操作能够在网络性能和网络管理方面，为网管人员提供有用的数据。

本章第四节介绍了几种访问和管理网络（设备）的方法和工具，比如，SNMP、Telnet、SSH和syslog等，上述工具在IPv4网络管理中应用广泛，但仍可用于IPv6网络管理，在配置方面也几乎不需要做任何改动。

本章篇末则重点介绍了几种流量监控工具，比如，SPAN、Mini协议分析器，以及VACL捕获等，这些工具有利于网管人员排除网络故障。

本章讨论的所有内容不但可帮助网管人员去完成从IPv4到IPv6网络管理的过渡工作，而且在必要时，还能在排除网络故障方面助人一臂之力。

11.7 参考资料

Cisco. Cisco IOS配置指南：

① 原文是“In this chapter, we took a closer look at managing different design IPv6 options”，在译者看来“managing different design IPv6 options（管理不同的设计IPv6选项）”不是人话，故而未按原文翻译。

② 原文是“网络基础设施”。

http://www.cisco.com/en/US/products/ps6350/tsd_products_support_configure.html.

Cisco. Cisco Nexus 7000 Series Switches Configuration Guides: http://www.cisco.com/en/US/products/ps9402/products_installation_and_configuration_guides_list.html.

Cisco. Cisco MDS 9000 NX-OS and SAN-OS Software Configuration Guides: http://www.cisco.com/en/US/products/ps5989/products_installation_and_configuration_guides_list.html.

Cisco. Catalyst 6500 Release 12.2SXH and Later Software Configuration Guide: http://www.cisco.com/en/US/docs/switches/lan/catalyst6500/ios/12.2SX/configuration/guide/book.html.

Haberman, B. RFC 4294, "IP Forwarding Table MIB." http://tools.ietf.org/html/rfc4292.

Hinken, R., et. al. RFC 2732, "Format for Literal IPv6 Addresses in URLs." http://tools.ietf.org/html/rfc2732.

S. Routhier, Ed. RFC 4293, "Management Information Base for the Internet Protocol." http://tools.ietf.org/html/rfc4293.

Bierman, A. RFC 4133, "SNMP Entity MIB". http://tools.ietf.org/html/rfc4133.

Kavasseri, R. RFC 3014, "SNMP Notification MIB." http://tools.ietf.org/html/rfc3014.

Levi, D., et. al. RFC 2273, "SNMPv3 Applications." http://tools.ietf.org/html/rfc2273.

Daniele M., et. al. RFC 4001, "Textual Conventions for Internet Network Addresses." http://tools.ietf.org/html/rfc4001.

Raghunarayan, R. RFC 4022, "Management Information Base for the Tranmission Control Protocol." http://tools.ietf.org/html/rfc4022.

Fenner B., et al. RFC 4113, "Management Information Base for the User Datagram Protocol." http://tools.ietf.org/html/rfc4113.

CHAPTER 12

第 12 章　按部就班：搭建 IPv6 实验网络，启动生产网络的试点工作

本章涵盖以下主题。

- **IPv6 实验室环境的网络布局**：一般而言，不太可能在实验室中再现与生产环境相仿的 IPv6 网络，于是我们搭建出了一个可将就使用的示例网络（该网络也被称为实验室网络环境，简称实验环境）。
- **IPv6 实验环境的编址情况**：制定出了一份简单的编址规划方案，其目的是要向读者演示层次化编址方案的决策过程，读者不应将其视为编址规划的基准方案（baseline）。
- **配置网络设备**：在实验环境中以及生产网络的 IPv6 试点阶段内，配置网络设备时，会以本书先前各章所示出的配置作为基准。
- **操作系统、应用程序以及管理工具的部署**：网络建成之后，如不用来“跑”应用程序，那就毫无存在的必要。本节所举的几个示例，与在 IPv6 网络中安置操作系统和应用程序，以及部署用来管理操作系统、应用程序以及 Hypervisor（虚拟机管理程序）的管理工具有关。
- **开展生产网络的 IPv6 试点工作**：部署 IPv6 并非只涉及技术，同样要牵扯到行政及业务流程，本节将会对此加以总结。

IT 项目要想取得成功，除了要在项目开展之前未雨绸缪，在项目开展之后同舟共济之外，还需有宽裕的时间让 IT 人员掌握新技术、新协议、新设备以及新的流程。在企业中推行 IPv6，与任何其他重大技术推广并没有任何差别；除了需要精心规划，还需丰富的实施经验，只有这样，才能够避免在试点阶段和生产网络中“捅篓子”。

IPv6 部署的实验和试点阶段之所以常为人们所忽视，是因为人们总抱有这样一种观念——“IPv4 和 IPv6 能有多大差别呢？”——该想法确实也不能算错，IPv4 与 IPv6 之间的确有不少相同之处。从网络设计的角度来看，两种协议的部署方式，尤其是双协议栈的部署，几乎完全相同。然而，这两个版本的协议之间也有诸多不同之处，有必要对 IPv6 的配置过程及其与主机、操作系统和应用的协作过程形成审慎而又务实的观点。

本章将讨论搭建 IPv6 实验环境所要关注的事宜，该实验环境可满足以下测试需求。

- 第二层需求（比如，第二层交换机具备 MLD 欺骗功能）。
- 第三层需求（比如，良好的编址规划和路由选择方案）。
- WAN 分支机构的连通性需求（比如，能够通过各种 WAN 传输选项，建立 WAN 连通性）。
- Internet 连通性和安全性需求（比如，建立 IPv6 Internet 访问，并提供防火墙及远程访问 VPN 服务）。
- OS 的配置需求（比如，服务器和客户端 OS 的配置）。
- 活动目录、DNS 和 DHCP。
- 应用程序测试（比如，文件服务器、WEB 服务器以及服务器虚拟化的访问和管理）。

根据上述需求，以及本书其他章节所介绍的设计方案/配置举例，我们将搭建一个具备 IPv6 基本功能的实验或生产网络的测试环境。这一实验或生产网络的测试环境不但能够帮助读者加深对 IPv6 的理解，而且还能让读者见识一下如何在网络中部署 IPv6。

12.1 实验环境拓扑示例

搭建 IPv6 实验环境的方式多种多样。本节会示出一个 IPv6 实验环境的拓扑示例，该拓扑示例是本章所有内容（论述、配置、测试）的基础。

如图 12-1 所示为该 IPv6 实验环境拓扑的逻辑视图。将其称为逻辑视图，是因为图中所示的某些组件其实是利用虚拟化技术针对单台物理设备的“分割”，这些虚拟出的逻辑组件被部署在网络的不同位置，行使不同的功能（比如，将物理交换机

划分为多个 VLAN，可把每个 VLAN 视为一台逻辑交换机）。此外，还在一台物理主机（物理服务器）上安装了虚拟化软件，虚拟出了多台虚拟机，并遍布网络各处。

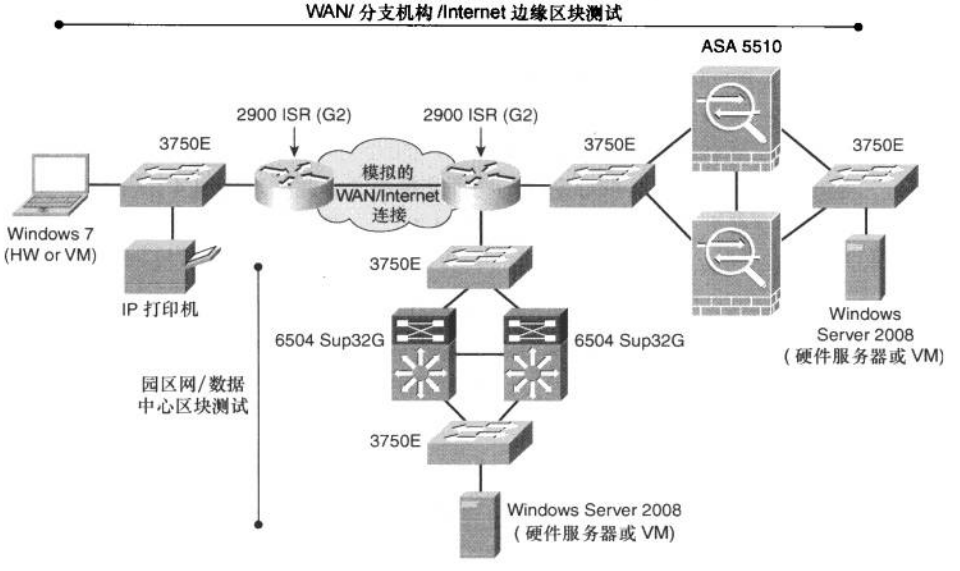


图 12-1 IPv6 实验环境拓扑示意图

该 IPv6 实验环境拓扑所包括的设备（主机）清单如下所示。

- 1 台 Cisco Catalyst 3750E-48TD 交换机。
- 2 台 Cisco 2911 ISR G2 路由器。
- 2 台配备了 Supervisor 32G 的 Cisco Catalyst 6504 交换机。
- 2 台 Cisco ASA 5510 防火墙。
- 一台安装了 VMware vSphere 4.1（带 ESX i 功能）的硬件服务器。
- 上述服务器配备 4 块 10/100/1000M 网卡（NIC）。
- vSphere 4.1“托管”着以下 VM。
  - Microsoft Windows 2008 R2 64-bit 上安装了 VMware vCenter。
  - Microsoft Windows Server 2008 R2 64-bit (AD/DNS/DHCP)。
  - Microsoft Windows Server 2008 R2 64-bit (Internet 模拟机)。
  - Microsoft Windows 7 Enterprise 32-bit。

- 各种 Linux VM。
- 一台支持 IPv6 打印的网络打印机（型号为 Brother MFC 7840W）。

乍一看，本 IPv6 实验环境的物理拓扑与逻辑拓扑并不相同，其原因是，我们在那台 Cisco Catalyst 3750E-48TD 交换机上划分了多个 VLAN（VLAN 信息为本机有效，不为外部设备所知），从而将其划分为多台虚拟交换机。这样的虚拟交换机可在网络各区域顶替物理交换机，如此一来，本实验环境中只需一台物理交换机，而无需为各区域分别订购物理交换机了。此外，还用 4 根网线将那台安装在硬件服务器上的 4 块 10/100/1000M 网卡分别与 Cisco Catalyst 3750E-48TD 交换机的不同端口相连，以便运行于其上的各 VM 连接到相应的端口和 VLAN。

理想情况下，在实验环境中所使用的基础设施部件应该与用于生产网络的部件保持一致，以便于网管员人员提前感受在生产网络中部署 IPv6 的体验。而对网络拓扑来说，则更应如此。本章所举示例中的装备和拓扑只是搭建 IPv6 实验环境的无数种方法之一。

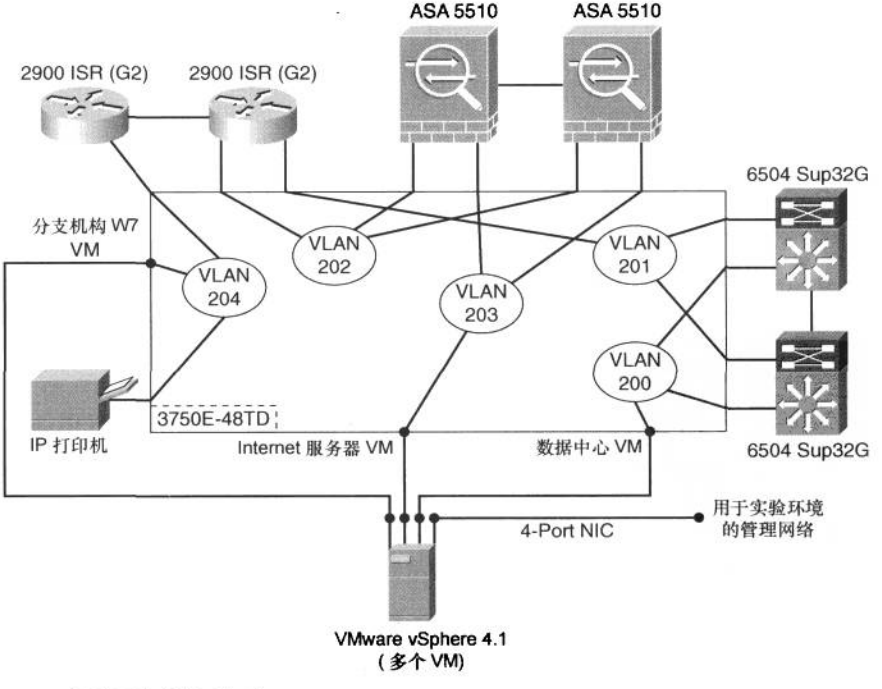


图 12-2 IPv6 实验环境的物理拓扑

图 12-2 所示的实验环境麻雀虽小，但五脏俱全（With this simplistic but functional lab layout），接下来的工作包括制定编址方案、配置设备、安装操作系统、设置应用程序，开展测试。

12.2 实验环境的编址方案

如果实验环境的规模够大，那么 IPv6 编址方案可能会稍微麻烦一点，否则，编址方案的制定则会简单得多。大多数情况下，企业都不会将实验环境所采用的 IPv6 编址方案应用于生产网络，这自然是归因于两种网络环境的规模一小一大。但在实验环境中，可验证生产网络能否在点到点链路上使用/64 的前缀，而非/126、/127 或其他类型的前缀。此外，在为实验环境的网络设备分配 IPv6 地址时，也可以使用非常复杂的 64 位接口 ID 进行实验，以取代简单的接口 ID（比如，2001:DB8:CAFE:100:6500:ABFD:1928:3746 vs. 2001:DB8:CAFE:100::1）。大多数实验表明，使用复杂的 IP 地址分配方案，不但不会提高网络安全性，反而会给网络运维制造障碍。请读者牢记，可利用实验环境去尝试与 IPv6 编址有关的各种可能性，但不足以据此来确定大型生产网络的 IPv6 编址规划方案。

为实验环境的规模扩充做好准备，也是一个不坏的主意。在许多情况下，其他 IT 团队都会逐渐意识到实验环境的价值，并会觉察到也需依靠该实验环境来测试自用的 IPv6 部件。正因如此，先制定出一份可满足实验环境规模增长的网络编址环境，并同时针对企业员工开展与生产网络 IPv6 编址规划有关的培训工作，可算作一条良策。以下所示为一份简单的编址方案，该方案并未基于任何最佳做法，读者可将其视为如何细分 IPv6 前缀以供实验环境（或生产网络的启动阶段）使用的参考指南。

2001:db8:cafe::/48 (company prefix, using RFC 3849 prefix)
/50 - Theatre (US, EMEA, APAC, spare)
/51 - Region (West/East)
/55 - City/site (Denver, SFO, NYC, and so on), 16 cities per region
/64 - Link prefix at each site (512 /64s used)

An example of this addressing hierarchy is as follows (note: nonzero prefixes are shown):

Company - 2001:db8:cafe::/48
USA - 2001:db8:cafe::/50
West - 2001:db8:cafe::/51
San Jose, CA (HQ) - 2001:db8:cafe::/55

Campus - 2001:db8:cafe::/64 - cafe:ff::/64
Data center - 2001:db8:cafe:100::/64 - cafe:1f2::/64
Loopback range - 2001:db8:cafe:1f3::/128 (out of /64)
WAN core/DMZ - 2001:db8:cafe:200::/55
(Sites are branches)
Los Angeles, CA - 2001:db8:cafe:400::/55
San Diego, CA - 2001:db8:cafe:600::/55
Seattle, WA - 2001:db8:cafe:800::/55
Las Vegas, NV - 2001:db8:cafe:a00::/55
Phoenix, AZ - 2001:db8:cafe:c00::/55
Salt Lake City, UT - 2001:db8:cafe:e00::/55
Denver, CO - 2001:db8:cafe:1000::/55
and so on

可供实验环境使用的地址分配方案多种多样，以上只举出了其中的一种。其他与编址规划有关的注意事项还包括：一，如何建立并测试与 SP 对接的多宿主连接场景；二，如何针对不同的 SP，过滤不同的前缀；三，如何以最佳方式聚合路由等。

12.3 网络设备配置

网络设备具体如何配置要取决于特定的网络场景，以及实验环境中所要完成的具体测试内容。配置实验环境网络设备时，读者完全可以借鉴本书先前各章所提供的完整网络设备配置清单，归因于此，本章也就不再呈现每台设备的配置了。表 12-1 列出了搭建实验环境时，书中所讨论过的相关配置出处。

表 12-1 Cisco 网络设备配置出处

实验环境各网络区块	配置出处
园区网/数据中心区块	第 6 章、第 9 章
WAN/分支机构网络区块 Internet 边缘网络区块	第 8 章、第 6 章

12.4 操作系统/应用程序的安装以及网络管理

网络设备连接完毕、配置完 IPv6 功能，并验证过端到端的连通性之后，是时候在实验环境中安装终端、安装应用程序/操作系统，并设置必要的系统或网络管理工具了。由图 12-1 的实验环境概况图可知，需要安装的操作系统包括 Microsoft Windows 7、Microsoft Windows Server 2008 R2 以及带 Hypervisor ESXi 功能的 VMware vSphere 4.1。对于大多数企业来说，在 IPv6 实验环境中，安装其他种类的操作系统（比如，Linux 以及 Microsoft Hyper-V 之类的 Hypervisor 解决

方案）也属常事。出于演示的目的，本实验环境中只需安装 windows 操作系统。

首先，在物理服务器上安装 VMware vSphere 4.1 和 Hypervisor ESXi 4.1。VMware vSphere 和 ESXi 的安装非常简单，VMware 环境的规划和部署说明请见 VMware vSphere 支持中心网页：http://www.vmware.com/support/product-support/vsphere。

本章只关注 ESXi 的 IPv6 配置，这一步完成之后，实验环境中的 VMware vCenter 和 VMware 基础设施客户端（VI 客户端）才能通过 IPv6，去连接并管理主机。

图 12-3 所示为 ESXi 控制台，在实验环境中正是利用该控制台来安装 ESXi。

在主控制屏（如图 12-3 所示）的“Configure Management Network”下选择“IPv6 Configuration”。

默认情况下，VMware ESXi 上的 IPv6 功能为禁用状态，因此在配置 IPv6 之前，需先行激活其 IPv6 功能。图 12-4 所示的截屏表明，IPv6 功能已被激活。

在控制台上激活了 IPv6 之后，需重启主机。只有在主机重启之后，其他的 IPv6 配置参数才能生效。图 12-5 所示为为 ESXi 主机静态分配 IPv6 地址。在实验环境中，第一跳路由器会发送路由器通告消息，接收该消息的主机，会把第一跳路由器通告的地址作为自己的默认网关地址（最好是 HSRP 或 GLPB IPv6 虚拟地址）。当然，也可以在主机上，静态配置默认网关的地址（即 HSRP 或 GLPB IPv6 虚拟地址）。

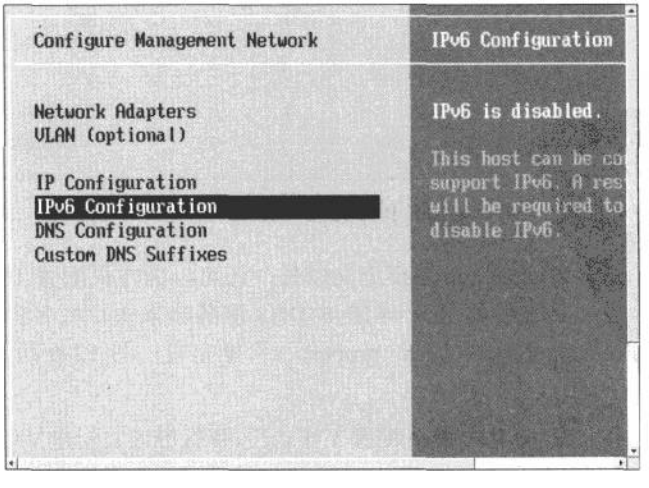


图 12-3 VMware ESXi 控制台——配置管理网络截屏

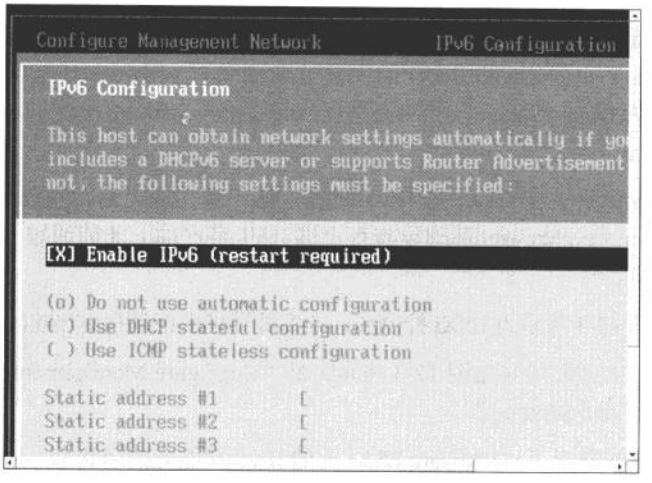


图 12-4 VMware ESXi 控制台——激活 IPv6

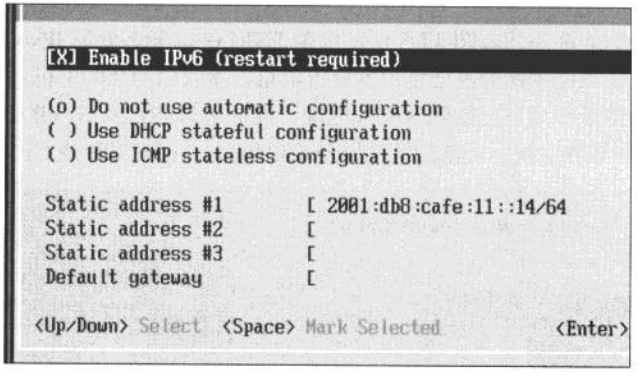


图 12-5 VMware ESXi 控制台——静态配置 IPv6 地址

除了上述静态地址需要定义以外，还可以为主机配置 IPv6 DNS 服务器地址，以使其能够利用 IPv4/IPv6 DNS 服务器来执行域名解析。这一步配置工作可在 vCenter 的“host properties”里完成。图 12-6 所示为主机的 DNS 配置。

主机上线，并完成 IPv6 配置工作之后，那就应该去创建 VMware vCenter VM 了（也可以在别的主机上安装 vCenter，然后远程管理这台主机，并以远程管理

的方式来创建 VM)。图 12-7 所示为来自 VMware vCenter VM 的“AddHost Wizard”截屏，创建其是为了用来管理 ESXi 主机。

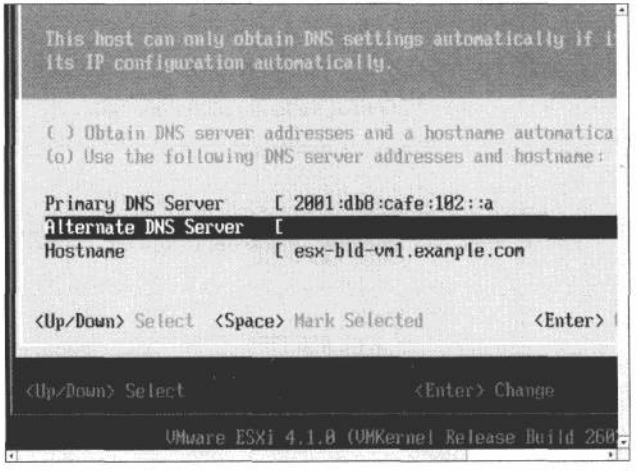


图 12-6 VMware ESXi 控制台——配置 DNS

如图 12-7 所示，在 vCenter 的“Add Host Wizard”的“host”字段中，填入了一个 IPv6 地址，该地址为之前所定义的 ESXi 主机地址。如果已事先部署了 DNS 服务器，或编辑了本机的 hosts 文件并添加了相关记录，以行使域名解析功能，那么也可以在“host”字段中填入 ESXi 主机名来代替 IPv6 地址。在提供了用户名/密码之后，vCenter 便会开始连接 ESXi 主机，并会将其添加进自己的数据库。

图 12-8 所示为已将一台新 ESXi 主机（IPv6 地址为 2001:db8:cafe:11::14）添加进了 vCenter 数据库，读者还可以看见另外两台 ESXi 主机（这两台主机与本章内容无关)。

截至目前，我们已经搭建出了一个功能齐备的 IPv6 实验环境，并安装好了一台 VMware vCenter 和 Hypervisor ESXi 主机，从这台主机可构建出所有其他的 VM。

我们构建出了一台 OS 为 Microsoft Windows Server 2008 R2 的 VM，用作为实验环境中的活动目录（AD)，DNS 以及 DHCP 服务器。为了省事，这台 VM 还将作为实验环境中的 Web 和 FTP 服务器，为 Windows 7 客户端提供 IPv6 上的 www 和 FTP 服务。

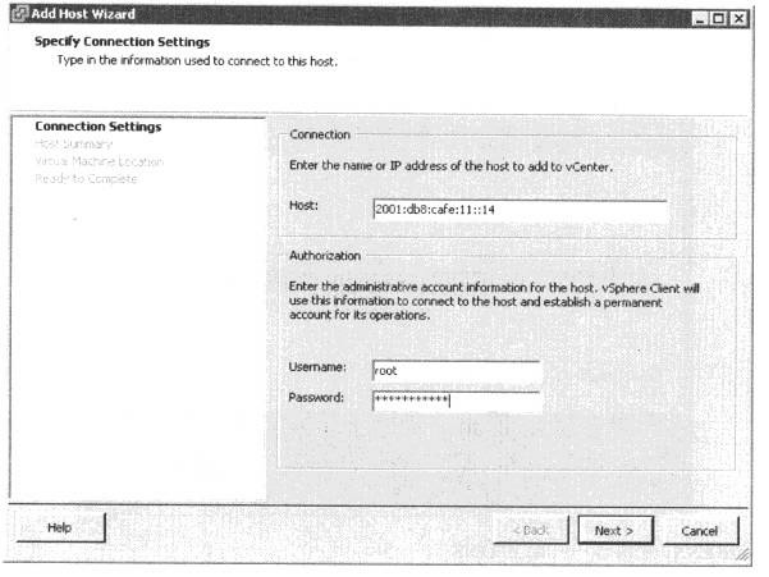


图 12-7 VMware vCenter 控制台 - Add Host Wizard

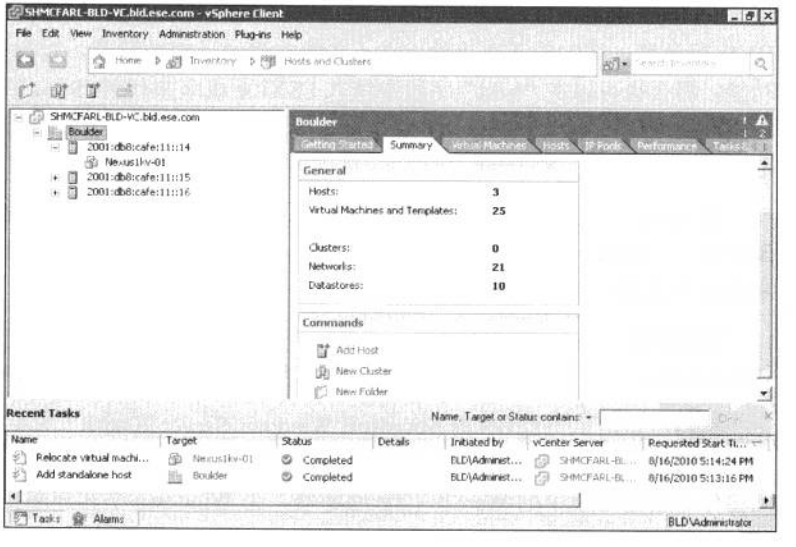


图 12-8 VMware vCenter 控制台-支持 IPv6 的 ESXi 主机

在 Microsoft Windows Vista、Windows 7、Windows Server 2008 以及 2008 R2 上，都默认支持 IPv6。对于这台实验环境中的 Windows 服务器来说，其 IPv6 地址和 DNS 服务器地址全都采取静态配置，DNS 服务器的地址配置为自身的 loopback 地址::1，如图 12-9 所示。

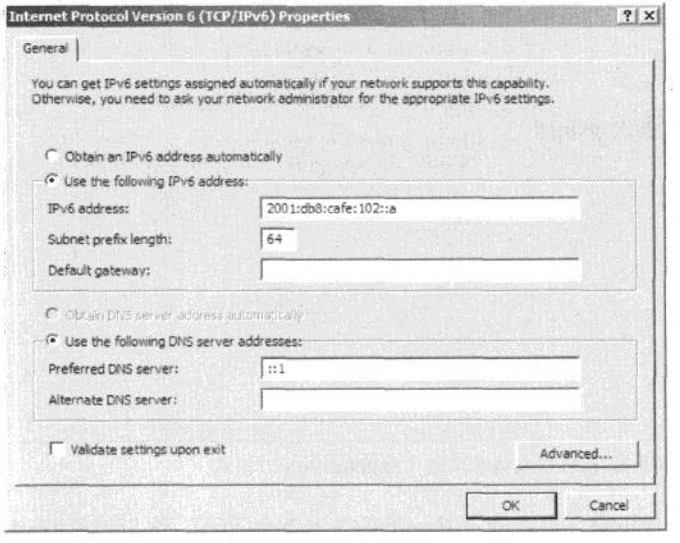


图 12-9 Microsoft Windows Server 2008 R2-静态配置 IPv6 地址

构建完 Windows Server 2008 R2 VM、在其上配置完 AD 和 DNS 服务，并安装好 DHCP 和 WEB 服务器之后，该到了配置 IPv6 DHCP 服务，以便 Windows 7 客户端能够获取到 IPv6 地址的时候了。

Microsoft Windows Server 2008 DHCP 服务器支持两种自动获取地址的模式：无状态模式和有状态模式。无状态模式是指：DHCP 客户端通过其他方式（比如，利用自动配置特性，从本地路由器接受 IPv6 前缀）接收 IPv6 地址，但从运行 DHCP 服务的 Microsoft Windows 服务器接收其他 DHCP 选项信息。只有有状态模式才与 IPv4 DHCP 机制类似——DHCP 客户端通过 DHCP 服务器同时接收地址和 DHCP 选项信息。可根据实际需求来启用或禁用 DHCP 的无状态模式。在本实验环境中，为日后测试方便，在 Microsoft Windows Server 2008 DHCP 服务器上开启了 DHCP 无状态模式。

图 12-10 所示为选择开启 DHCP 模式的截屏图。

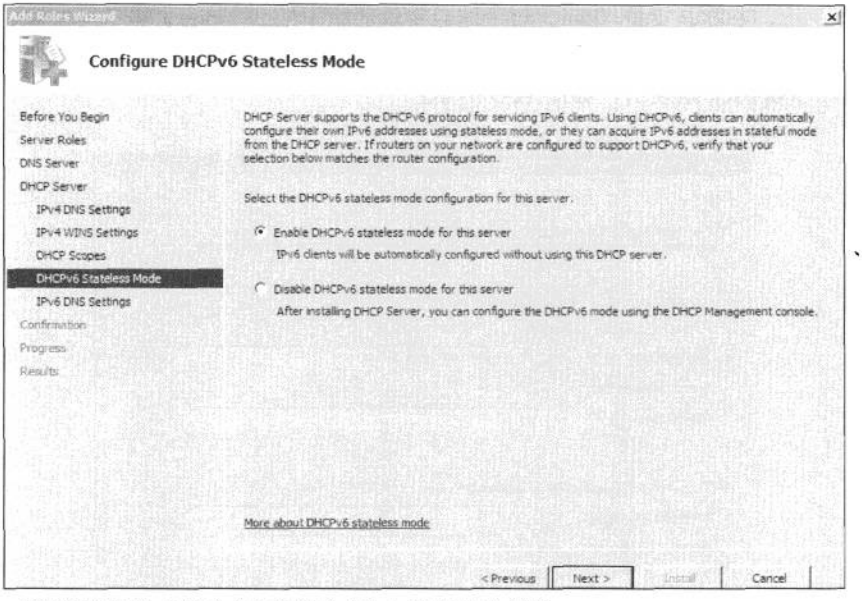


图 12-10 DHCP 服务器-无状态或纯有状态（Stateful-Only）模式

接下来，需要确定 IPv6 DHCP 地址池范围，这与创建 IPv4 DHCP 地址池范围几乎完全一样。需提供“Scope name”、“address range”、“exclusions”等选项。

图 12-11 所示为 DHCP New Scope Wizard 的第一屏截屏图。

在实验环境中，与 DHCP 有关的第一步测试工作是，需要确保 Windows 7 客户端能通过 DHCP，接收到相应 VLAN 的 IPv6 地址。

在 DHCP 服务器上，我们定义的第一个地址池范围即为 VLAN 102 使用的前缀 2001:db8:cafe:102::/64，DHCP 服务器自身也隶属于 VLAN 102。先将 Windows 7 客户端划到 VLAN 102，以测试其能否收到 IPv6 地址，是否已成功地加入域，并对其访问实验环境网络的基本情况进行相关测试。图 12-12 所示为 VLAN 102 所使用的 IPv6 前缀范围：2001:db8:cafe:102::/64。

下一步（单击图 12-12 中的 Next 按钮），便来到了“排除”地址范围界面，DHCP 服务器不会把落在这一地址范围内的地址分配出去，对与本实验环境来说，如需配置此选项。

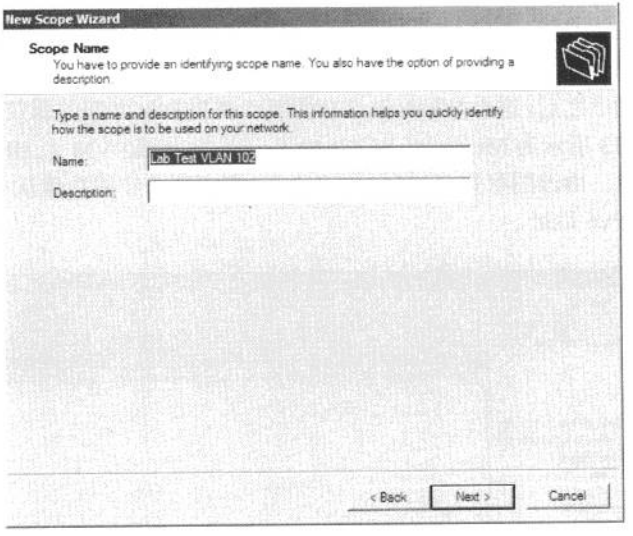


图 12-11 IPv6 DHCP 服务器 New Scope Wizard-scope name

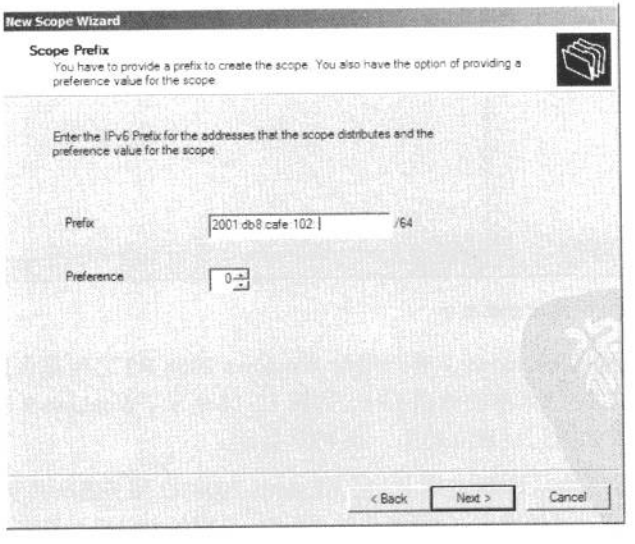


图 12-12 IPv6 DHCP 服务器 New Scope Wizard-Scope Prefix

DHCP 地址范围定义完毕之后，还需为 DHCP 客户端配置额外的 DHCP 服务器选项（比如，域名搜索列表等）和其他选项。

配毕 AD、DNS 以及 DHCP 服务的基本参数之后，就可以在 vCenter 内构建 Microsoft Windows 7 VM，并能够将其连接到 VLAN 102 了。

加电之后，这台 Windows 7 VM 便可以从 DHCP 服务器获取到 IPv6 地址了。图 12-13 所示为 Microsoft Windows Server 2008 R2 VM 上 DHCP 管理控制台的截屏图。由该图可知，这台 Windows 7 客户端已经成功地从 DHCP 服务器接收到了 IPv6 地址。

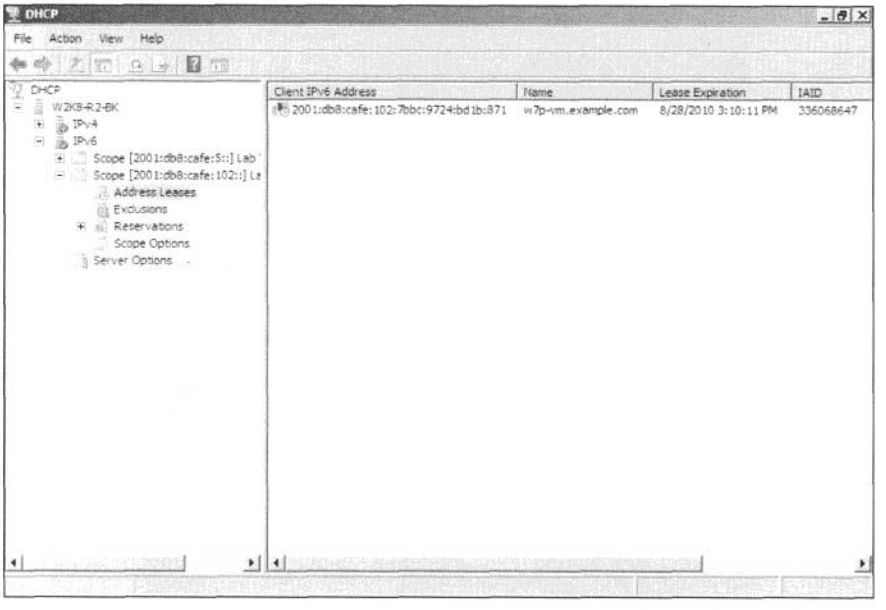


图 12-13 IPv6 DHCP——地址分配成功

为确定 Windows 7 客户端与 Windows 2008 R2 之间是否建立起了连通性，还需执行几步基本的测试工作。如图 12-14 所示，Windows 7 客户端与 Windows 2008 R2 之间已经建立起了一条 FTP 会话。

在 Windows 7 客户端上，对 Windows 2008 R2 服务器所共享的磁盘，执行了映射网络驱动器的操作（执行操作时，使用的是服务器名而不是服务器的 IPv6 地址）。此外，还在 DNS 服务器上创建了一条 DNS AAAA 记录，以测试 Windows 7 客户端能否正常解析 DNS 记录。如图 12-15 所示，在 DNS 服务器上添加了一条记录“test”，并将其与 Windows 2008 R2 服务器的 IPv6 地址相关联。

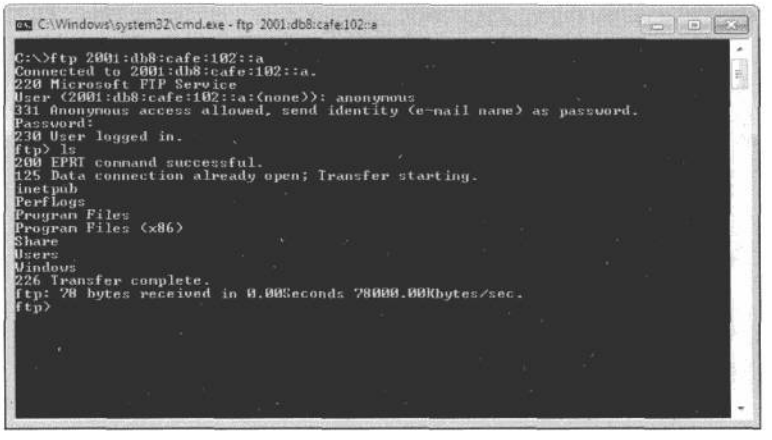


图 12-14 IPv6 上的 FTP 连接测试

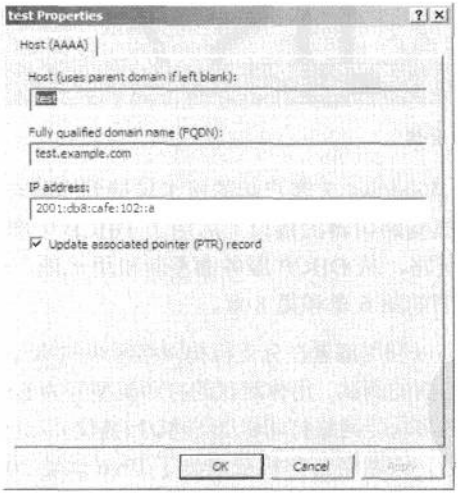


图 12-15 新的 DNS AAAA 记录

然后，在 Windows 7 客户端上开打浏览器，输入网址：http://test.example.com，便打开了 Windows 2008 R2 服务器（IPv6 地址为 2001:db8:cafe:102::a）上的默认 Web 站点（Web 服务器为 IIS）。现在，在 Windows 7 客户端上，已经建妥了基于 IPv6 的 FTP 和 HTTP 连接，并利用 IPv6 访问到了映射的网络驱动器。图 12-16 所示为 Windows 7 客户端上的活跃 TCP 连接。

图 12-16 IPv6 上的活跃 TCP 会话

我们再将 Windows 7 客户端的网卡转接到实验环境的 WAN/分支机构网络区块。在分支机构路由器的接口上启用了 DHCP 中继特性，以便于该客户端能够通过 WAN 链路，从 DHCP 服务器重新租用地址。有关 DHCP 中继特性的配置，请参考本书的第 6 章和第 8 章。

此外，还可以利用部署在分支机构网络区块的那台网络打印机（支持 IPv6），来做一些其他方面的测试。用作测试的打印机型号为 Brother MFC 7840W（这是一台小型办公室的多功能网络打印机/复印机/扫描仪）。这台 MFC 7840W 支持 IPv6，如图 12-17 所示，该网络打印机获取到了 IPv6 地址 2001:db8:cafe:1000::5，其前缀 2001:db8:cafe:1000::/55 正好是分配给该分支机构链路的前缀。

到目前为止，实验环境的网络不但可以运转正常，而且也部署了诸如 Windows 7 和 Windows Server 2008 R2 VM 这样的主机，各网络端点（比如，支持 IPv6 的网络打印机）也已连接上网。还需构建出一台 VM，作为部署在 Internet 边缘区块之外的主机或服务器。可利用该 VM 来模拟 Internet 服务器，以进行外部访问测试、Cisco Any Connect SSL VPN 客户端远程访问测试，以及 Cisco ASA 的入站端口流量过滤测试。

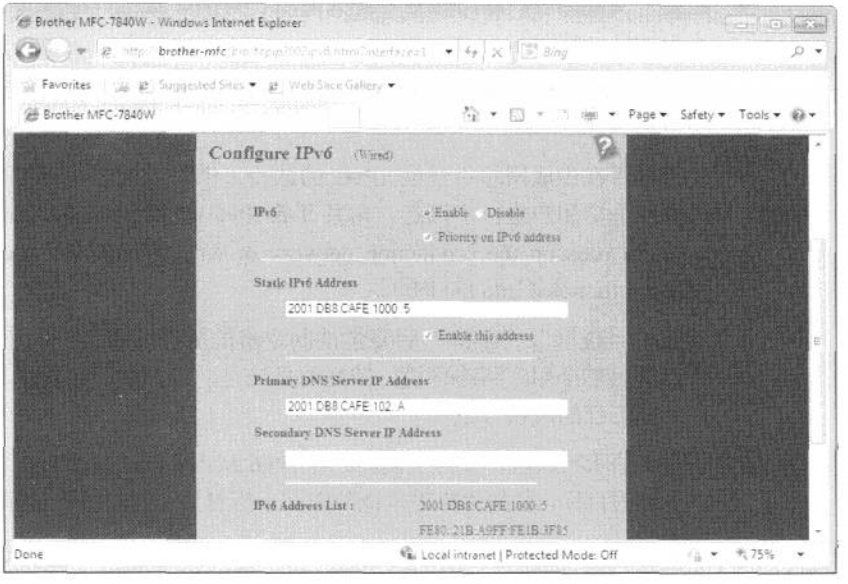


图 12-17 网络打印机的 IPv6 相关设置

针对本 IPv6 实验环境的基本测试完成之后，试验结果表明，本实验环境已经具备了进行全面测试和应用程序部署的条件。

12.5 开展生产网络的 IPv6 试点工作

开启生产网络的 IPv6 试点工作，总会涉及将 IPv6 引入生产网络，只是在程度上有所不同，或许，还要牵涉到让生产网络访问到搭建在实验室中的实验环境。此时，也正是检验并解决诸多与操作层面有关的 IT 问题的最佳时机。在此阶段，可能会暴露出来的问题包括：日常运维问题、对 IPv6 端点的技术支持问题、应用程序管理问题、补丁/升级问题，以及 help desk 问题。

在 IPv6 试点阶段，最关键的因素往往都与行政而不是和技术挂钩。当务之急是，应成立一支由各方 IT 代表组成的跨职能虚拟团队，来参与规划和部署生产网络的测试环境。如果网络团队一意孤行，强行在生产网络的 IPv6 先期试点阶段推广 IPv6，而不事先与其他 IT 团队（比如，桌面、安全、运维、应用/数据中心等其他 IT 团队）沟通的话，那么 IPv6 试点阶段必将以失败而告终。

要想成功的规划并执行 IPv6 试点阶段的部署，应采取以下行动。

- 组建一支囊括 IT 部门内各专业（网络、桌面、安全、运维等）成员的虚拟团队（VT）。
- 根据 VT 成员和高级管理人员的意见，设定 IPv6 试点阶段所能达到的期望值和目标。
- 决定是否在生成网络中开展 IPv6 的试点工作，或决定是否对试点阶段的某些选定用户进一步筛选，向其开放 IPv6 试验环境（Decide whether the pilot runs on the production network or whether selected users of the pilot are funneled into the lab）。
- 不要怕“碰壁”，“碰壁”后要尝试制定新的部署方案。相对于用 IPv6 直接覆盖整个生产网络环境，较好的做法是，边摸索、边发现，以期找到一套可行的解决方案。
- 不要害怕承担风险，要勇于创新。在 IPv6 试点阶段，只要参与各方设定了合理的目标，任何网络停运或故障都不应对生产网络的流量造成影响。
- 寻求帮助。Ineternet 上与 IPv6 有关的资源可谓无穷无尽，这些都可帮助我们敲定配置、制定测试方案、搞定网络故障。此外，还可向设备供应商伸出援助之手，若得不到帮助，那就换掉它。
- 一旦发现 bug，应立即提交给设备供应商，以便其在正式交付产品之前，能够尽快修复 bug。要获得 cisco 的支持，可访问 cisco.com 上的支持站点：http://www.cisco.com/cisco/web/support/index.html、直接致电 Cisco TAC，或与 Cisco 客户团队沟通。
- 学习。与管理 IPv6 生产网络环境相比，在试验环境以及生产网络的 IPv6 试点阶段中实施操作，压力要小很多。因此，要多花时间，学习新的知识，并制定出一份取材自 IPv6 试点阶段的完备方案，为在生产网络中正式部署 IPv6 做好准备。

12.6 总结

对于大多数 IT 项目来说，成败的关键要归结于：能否精心规划和充分测试，以及是否能够开展有针对性的试点运行。对任何 IPv6 部署来说，IPv6 实验室网络环境的搭建自然也决定其成败，搭建实验环境的工作重心包括：网络设备的配置、应用程序的部署、对操作系统间差别的比较以及对整个试验环境的管理。

积累自实验环境的经验可直接为 IPv6 试点阶段所用，在该阶段，对 IPv6 的部署应更为高效。来自企业 IT 部门各领域（专业）的代表应参与生产网络的每一次 IPv6 试点工作，并需制定出一整套明确的目标和期望值。只要 IT 部门的全体员工都把自己看作该项目的一份子，勇于承担责任，并能够对积累自实验室环境和试点阶段的经验加以总结，那么 IPv6 在生产网络中的正式部署就指日可待了。

12.7 参考资料

Cisco. Cisco Catalyst 3750E Series Switches: http://www.cisco.com/en/US/products/ps7077/index.html.

Cisco. Cisco Branch ISR Routers: http://www.cisco.com/en/US/products/ps10906/Products_Sub_Category_Home.html.

Cisco. Cisco Catalyst 6500 Series Switches: http://www.cisco.com/en/US/products/hw/switches/ps708/index.html.

Cisco. Cisco ASA 5500 Series Adaptive Security Appliances: http://www.cisco.com/en/US/products/ps6120/index.html.

Cisco. Cisco Support Tools: http://www.cisco.com/cisco/web/support/index.html.

Microsoft. Microsoft Windows Server 2008 R2: http://www.microsoft.com/windowsserver2008/en/us/default.aspx.

Microsoft. Microsoft Windows 7: http://www.microsoft.com/windows/windows-7/default.aspx.

Microsoft. Microsoft IPv6: http://technet.microsoft.com/en-us/network/bb530961.aspx.

Huston, G., A. Lord, and P. Smith. RFC 3849, "IPv6 Address Prefix Reserved for Documentation." http://www.ietf.org/rfc/rfc3849.txt.

VMware. VMware vSphere 4: http://www.vmware.com/products/vsphere/ and http://www.vmware.com/support/product-support/vsphere.

IPv6在企业网络中的部署

- 了解IPv6对企业的影响
- 理解IPv6服务以及支撑这些服务的IPv6特性
- 回顾最常使用的转换机制，其中包括双栈（IPv4/IPv6）网络、IPv4隧道上的IPv6、MPLS上的IPv6
- 设计具备模块化、层次化以及高弹性的IPv6网络
- 选择IPv6最佳实施方法
- 搭建IPv6实验网络环境
- 按部就班地配置园区网络、WAN/分支机构网络以及数据中心网络
- 将可满足生产需求的IPv6服务纳入IPv4网络
- 实施虚拟化IPv6网络
- 部署IPv6远程访问
- 以高效率、低成本的方式管理IPv6网络

本书是Cisco Press出版的网络技术系列丛书之一，该系列丛书可以为网络从业人员提供搭建高效网络、学习最新技术、打造辉煌职业生涯所需要的宝贵信息。

本书涵盖了在园区网络、WAN/分支机构网络、数据中心网络、虚拟化网络环境内成功部署IPv6必知必会的内容。在书中，Cisco公司的4位顶尖IPv6专家呈现了在大型网络中部署IPv6的切实可行的方法。他们向读者展示了IPv6对现有网络的影响，描述了IPv4/IPv6共存机制，提供了IPv6网络规划的指导意见，给出了用来构建IPv6实验环境、试点网络以及生产网络的配置示例。

作者首先回顾了IPv6技术的基本原理，探讨了企业加快部署IPv6的市场驱动力，然后介绍了与IPv6路由选择、QoS、多播、网络管理等技术有关的新特性，并拿上述特性与IPv4的类似特性做了一番比较。最后还列举了不少相关的配置示例。

对任何一名需要评估、规划、管理IPv6网络或向IPv6网络迁移的网络工程师、架构师、管理员、咨询师来说，本书是不可或缺的资料。

Shannon McFarland，CCIE #5245，Cisco公司企业咨询工程师，是企业网IPv6部署和数据中心设计方面的技术顾问，专注于应用程序部署和虚拟化桌面基础设施领域的研究。16年来，他从事过的工作包括大型企业园区网络、WAN/分支机构网络的设计；数据中心网络设计和优化；虚拟桌面基础设施的设计、部署和优化。最近10年，Shannon经常参加各种全球性的IPv6活动（其中包括Cisco Live），并在会议上踊跃发言。

Muninder Sambi，CCIE #13915，是Cisco Catalyst 4500/4900系列平台的产品营销经理，以及Cisco IPv6开发理事会的核心成员，同时还是IETF的IPv6领域的关键参与人员。

Nikhil Sharma，CCIE #21273，Cisco公司技术营销工程师，负责为Catalyst 4500产品线的软硬件雕琢新特性。

Sanjay Hooda，CCIE #11737，Cisco公司的一名技术领导，专注于嵌入式系统的研究，并协助定义新产品的体系结构。他当前关注的领域包括高可用性和大型分布式交换系统中的消息传递。

ISBN 978-7-115-26836-5

美术编辑：王建国

分类建议：计算机／网络技术／思科技术

人民邮电出版社网址：www.ptpress.com.cn

ISBN 978-7-115-26836-5

定价：69.00 元