

CCIE Professional Development: Advanced IP Network Design

网络核心技术内幕



专业 IP 网络规划与设计



本书配套光盘内容包括:

1. 与本书配套电子书
2. 送“计算机基础知识全面速成”多媒体学习软件

Alvaro Retana

[美] Don Slice 著

Russ White

希望图书创作室 译



北京希望电子出版社

Beijing Hope Electronic Press

www.bhp.com.cn

CCIE Professional Development: Advanced IP Network Design

网络核心技术内幕



专业 IP 网络规划与设计



本书配套光盘内容包括:

1. 与本书配套电子书
2. 送“计算机基础知识全面速成”多媒体学习软件

Alvaro Retana

〔美〕 Don Slice 著

Russ White

希望图书创作室 译



Z088937



北京希望电子出版社

Beijing Hope Electronic Press
www.bhp.com.cn

2000

内 容 简 介

本书是 21 世纪网络工程师设计宝典系列之一，是美国 Cisco 网络专家资格认证的权威培训教材。

随着网络的日益扩展，成功的网络规划对网络的稳定起着举足轻重的作用。通过本书，读者便能了解组建一个可扩展性网络所必须的基本概念，掌握网络规划的好方法，屏弃坏方法，并深入学习总的规划原则，设计出良好的网络规划。

本书共有四部分（分九章、五个附录）。第一部分讲述了网络稳定性的基础——网络的分层，讨论了分层规划的原则、地址分配和聚合、各层的冗余和网络规划原则的应用。第二部分讲述了各种先进的内部网关协议，包括 OSPF，IS-IS，EIGRP 网络规划。第三部分讲述了网络的扩展，讨论了 BGP 核心层和网络的可扩展性以及其它大规模核心层。第四部分作为本书的附录讲述了 OSPF，IS-IS，EIGRP，BGP 的基础。在讲述理论概念的同时，本书各章最后都附有实例学习和复习题，针对部分疑难问题提出相应的解决方案，并在附录 E 中有各复习题的答案。

本书结构清晰，内容丰富，技术新、实用性强，不但是想获取 Cisco 网络专家资格认证的广大科技人员必读的教科书，同时也是从事网络应用设计和开发的广大工程人员、开发人员、网络管理人员的重要参考书，高等院校相关专业师生重要的自学、教学参考用书和社会相关领域培训班教材。

本书配套光盘内容包括：1. 与本书配套电子书；2. 送“计算机基础知识全面速成”多媒体学习软件。

版 权 声 明

本书英文版名为“CCIE Professional Development: Advanced IP Network Design”，由 Cisco Press 出版，版权归 Cisco Press 所有。本书中文版由 Cisco Press 授权出版。未经出版者书面许可，本书的任何部分不得以任何形式或任何手段复制或传播。

版权登记 图字 01-1999-2738 号

系 列 书：21 世纪网络工程师设计宝典系列（1）

书 名：网络核心技术内幕——专业 IP 网络规划与设计

文 本 著 者：（美）Alvaro Retana Don Slice Russ White 著 希望图书创作室译

C D 制 作 者：希望多媒体开发中心

C D 测 试 者：希望多媒体测试部

责 任 编 辑：马宏华

出版、发 行 者：北京希望电脑公司 北京希望电子出版社

地 址：北京海淀路 82 号 100080

网 址：www.bhp.com.cn E-mail: lwm@.hope.com.cn

电 话：010-62562329,62541992,62637101,62637102, 62633308,62633309

（发行和技术支持）

010-62613322-215（门市） 010-62531267（编辑部）

经 销：各地新华书店、软件连锁店

排 版：希望图书输出中心

C D 生 产 者：文录激光科技有限公司

文 本 印 刷 者：北京双青印刷厂

开 本 / 规 格：787×1092 16 开本 15.625 印张 354 千字

版 次 / 印 次：2000 年 2 月第 1 版 2000 年 2 月第 1 次印刷

印 数：0001～5000 册

本 版 号：ISBN7-900031-63-4/TP·63

定 价：32.00 元（1CD，含配套书）

说明：凡我社光盘配套图书若有自然破损、缺页、倒页、脱页，本社负责调换。

译者的话

近几年来，随着 Internet/Intranet 为代表的计算机网络技术的日新月异，网络已成为当今世界上高科技领域最热门的技术之一。广大网络工作者急需了解最先进的网络技术知识。Cisco 作为世界上最大的网络专业公司，对网络技术的发展起着举足轻重的作用。本书主要针对网络的规划和稳定性提出相应的理念和解决方案。提供了组建一个可扩展性网络所必需的基本概念。只有有了好的、建立在被时间和经验验证过的规划，才能轻松的获得网络的稳定性。本书内容详尽、是极具参考价值的专业书籍，适于已经了解了路由及路由协议，并准备进一步深入学习的专业网络工作人员。

本书主要翻译人员：邓晓刚；参与本书翻译的其他人员：贾洪江、吴红梅、何琳琳、王晨、丁玉绢等。感谢北京希望电子出版社各位老师对本书译稿的精心审校。由于时间匆忙，译者水平有限，难免存在问题，恳请指正，不胜感谢。

JSN 66

关于作者

Alvaro Refana ,CCIE # 1609, 目前是 Large Scale Switching and Routing Team 的测试工程师。此前, Alvaro 是 Internet Service Provider Support Team 和 Routing Protocols Team 的技术领导。他在 BGP 和 Internet 结构方面是个知名的专家。

Don Slice, CCIE # 1929, 是一个 RTP 的升级工程师。他从前是 RTP TAC 的 Routing Protocols Team 的高级工程师。他在 EIGRP、OSPF 中一般的 IP 路由问题上是个知名的专家, 并且以他对 DECnet, CLNS/ISIS 的知识而闻名。Don 为世界范围的 Cisco 工程师提供升级支持。

Russ White, CCIE # 2635, 是一个致力于路由协议和路由结构的升级工程师。Russ 以他对 EIGRP, BGP 和其它 IP 路由方面的知识而闻名。

关于技术评审

William V.Chernock III, 是网络结构和设计方面的高级顾问。在过去的 8 年里, 他在经济和卫生行业排名的公司建造了大规模的战略网络。可以在 wchernock@aol.com 上访问到 William。

Vijay Bollapragada 是 Cisco 系统的 Internet Service Provider team 的高级工程师。他与在大规模网络设计和结构问题上的 Core Service Provider 合作。Vijay 可以在 vbollapr@cisco.com 上访问到。

前言

网络发展的必然规律：由小规模发展到大规模，再到超大规模，以至多个国家的网络。基于上述规律有如下必然的结果：一旦某个网络发展为多个国家的网络，相应的就会有人来决定由一个网络路由协议换成另一个网络路由协议的方案。他们会增加一个应用，或者一个核心链接会出现，然后再消失。

本书将提供组建一个可扩展性网络所必需的基本概念。因为我们工作在“网络崩溃了，快修”的情况下，这些基本概念将涵盖案例学习和理论讨论。本书包含了网络规划的好方法、坏方法及总的规划原则。在适当的时候，我们将会提供一些疑难问题的解决技巧。你会学到把网络扩展到它所需的程度（当然，越大越好了）。

本书内容

本书面向已经了解了路由及路由协议的基础，并准备进一步深入学习的专业网络工作人员。

下面列出的是本书未包含的内容：

- 非 Cisco 的路由器：读者总不会期望 Cisco 出版社发行一本包含其它厂商的路由器配置样例的书吧？
- 路由器配置：本书不会具体教读者如何配置 Cisco 路由器。它的主要目的在于介绍结构和原理，我们希望每一个读者能够从本书中得到他们在标准的 Cisco 手册中所需的配置信息。
- 路由协议的实施：本书的附录包含了在案例学习时所用协议的基本原理，但这并不是本书的主要目的。
- 路由协议的选择：所有先进的路由器协议都有优点和弱点。我们并不是帮你确定哪一种是最好的路由协议，但是我们可以帮你确定哪一种最适合你的网络（尽管最常选用的是静态的路由）。
- 路由选择信息协议（RIP）和内部网关路由选择协议(IGRP)：这些是比较旧的协议，所以我们认为它们不适合大规模网络的设计，它们可能在本书中被提及，但没有深入地涉及。
- 根据负载选择合适的路由器：这些专门的实施细节最好留给别的书来讲，因为可供参考的关于此类主题的书很多。
- 局域网（LAN）或广域网（WAN）传输介质的选择，电路速度或其它物理层的要求：它们虽不与 IP 网络规划直接相关，但对网络的可扩展性非常重要，这些内容也包含在其它关于分别由物理层及数据链路层建立网络的各类书中。

本书包含 OSPF，IS-IS，EIGRP 和 BGP 协议，因为它们是先进的协议并且各有千秋，如今在大规模网络上广泛而有效地使用着。毫无疑问将来会制订出其它的协议。

本书的重点在好的网络规划上，因为有了好的网络规划基础，就可以不去理会链路速度、物理技术、交换技术、交换速度或所使用的路由协议。你不会从安装出色的、新的数据链路层交换机或出色的、新超级快速路由器中获得网络稳定性。

你也无法从变换不同的先进路由协议中获得网络稳定性（除非你设计的网络工作不正常）。网络稳定性甚至不是通过确保没人接触路由器而获得的（虽然这样有时会有所帮助）。

只有建立在被时间和经验验证过的一致原则基础上的，规划良好的网络，你才可以高枕无忧。

本书的目的

写本书的主要目的是我们无法找到其它涉及到此类主题的书，还有就是因为我们相信在网络领域里，网络中网络层的规划是一个非常重要但又极少涉及的主题，我们希望读者能喜欢这本书，并且在日后把它作为参考书使用。

接着，请你坐回到舒适的椅子上，继续看上几页。你就能告诉老板：你正在掌握网络！

目 录

第一部分 稳定性的基础：分层的网络

第一章 分层规划的原则	3	实例学习：网络地址转换.....	28
从哪开始	3	复习	30
正确的拓扑结构	4	第三章 冗余	32
网络核心层	5	冗余的要点及策略.....	32
分布层	8	核心层冗余	33
访问层	9	分布冗余	39
与公共服务的连接	11	访问层冗余	41
小结	12	与公共服务的连接.....	43
实例学习：分层结构在交换网络中		小结	45
重要吗	12	实例学习：什么是最佳路径.....	46
复习	13	实例学习：使用交换在数据链路层	
第二章 地址分配与聚和	15	进行冗余	47
聚和	15	实例学习：单个路由器的拨号备用链接.....	48
成功分配地址的策略	20	实例学习：两个路由器的拨号备用.....	51
IPv6 地址方案	23	复习	54
地址分配的一般原则	24	第四章 网络规划原则的应用	55
小结	25	改造一个不稳定的网络.....	55
实例学习：接口的默认路径	26	复习	66

第二部分 内部网关协议

第五章 OSPF 网络规划	71	衡量 IS-IS 的其它因素	102
为 OSPF 实施而划分网络	71	IS-IS 邻机的故障排除	104
实例学习：OSPF 相邻体问题的		实例学习：单个区域的选项.....	106
疑难解答	85	实例学习：两层的网络.....	107
实例学习：该网络应该在哪个区域里	87	复习	109
实例学习：确定将链接放在哪个区域里	89	第七章 EIGRP 网络规划	110
实例学习：拨号备份	91	分析网络核心层	110
实例学习：OSPF 外部路径和下一跳	92	分析网络分布层的聚和	113
复习	93	分析网络访问层中的路由	116
第六章 IS-IS 网络规划	95	分析外部连接的路径	118
划分网络	95	分析到公共服务区的路径	120
分析与外部相连的 DMZ 路由器	99	分析到拨号接入客户的路径	122

EIGRP 网络规划的聚和	124
实例学习：聚和方法	125
实例学习：控制查询传播	126
实例学习：过多的拓扑结构表条目	127
实例学习：EIGRP 网络邻居关系的 疑难解答	129
实例学习：保持激活状态路径的 疑难解答	132

实例学习：重新分布	136
实例学习：EIGRP/IGRP 重新分布	140
实例学习：重传输和 SIA	142
实例学习：多重 EIGRP ASs	143
复习	146

第三部分 网络扩展

第八章 BGP 核心层和网络的可扩展性 .. 149

核心层里的 BGP	150
实例学习：迁移示例	150
核心层外的扩展	153
将网络划分为段	154
BGP 网络扩展的困难	155
实例学习：路由反射器和路由服务器	160
实例学习：BGP 相邻关系的疑难解答	161
实例学习：根据情况发布	164

实例学习：双归接入 Internet	165
实例学习：路由衰减 (Dampening)	170
复习	172

第九章 其它大规模核心层 .. 173

NHRP	173
实例学习：ATM 网里的 NHRP	174
MPLS	176
复习	181

第四部分 附录

附录 A OSPF 基础 .. 185

OSPF 是怎样工作的	185
路由器标识符 (Router ID)	185
LSA 类型	185
LSA 的可靠扩散	186
建立邻居 (Adjacencies)	186
多路存取网络上的邻居	188
OSPF 和无广播多路存取网络	189
区域 (areas)	191
外部路由注入	192
虚拟链接	193
按需路由	193

附录 B IS-IS 基础 .. 195

IS-IS 如何工作	195
端系统和中间系统	195
CLNS 地址	195

IS-IS 网络里的路由	196
IS-IS 中的量度 (Metrics) 和外部路径	197
建立邻居	197
LSP 扩散和 SPF 重算计时器	198
邻居丢失和 LSP 重生成	198
IP 集成到 IS-IS 里	199
多重 net 语句	199

附录 C EIGRP 基础 .. 200

扩散更新算法 (DUAL) 操作	200
在一个 EIGRP 网里建立邻居关系	201
EIGRP 网络里的量度	203
EIGRP 网络里的无环路路由	204
EIGRP 中的水平分割 (Split-Horizon)	206
在 EIGRP 网络中清除拓扑表和查询邻居	206
保持激活的路由	208
EIGRP 网的边界查询	208

EIGRP 聚和.....	208	邻居关系.....	214
为可靠性传输而在 EIGRP 里更改量度.....	209	BGP 里的路径过滤.....	217
EIGRP 网里的负载平衡.....	209	iBGP 同步.....	219
附录 D BGP 基础.....	211	BGP 聚和.....	219
路径向量协议机制.....	211	附录 E 复习题答案.....	221
路径判断.....	212	词汇表.....	232
公共字符串.....	213		

第一部分

稳定性的基础：分层的网络

第一章 分层规划的原则

第二章 地址分配与聚和

第三章 冗余

第四章 网络规划原则的应用

第一章 分层规划的原则

老板走进你的办公室，把一张购买订单扔在你的桌上说：“看，订单已经签了，采购说一千个路由器要在这占很大的地方，让你的人尽快走。现在开始干吧！”这是一场梦还是一场梦魇？

这当然不是真的——实际的网络只需两个路由器和一个链接就可以工作了，无需一千个路由器的订单。但即使是一个接有十个路由器的网络还是太小以致于网络规划不是个问题，这样说对吗？错，任何时候开始规划你的网络以后将如何扩展都不会太早。

从哪开始

好，你已决定开始考虑网络规划了。当规划一个网络时，最好从位于底部的物理层开始。物理层规划的最重要部分是比特和字节，如何正确排列一个链接，使用哪种传输介质，使用什么信号传输方式进行收发？

这些都很重要，因为你必须有稳定的物理链接以便信号在网络上传输。不稳定的物理链接会引起网络中路由器必须适应的变化，但网络的拓扑结构即网络布局对网络的稳定性的影响，比是否使用 ATM 或帧中继用作广域连接都大。

设计优良的拓扑结构是所有稳定网络的基础

为了理解这种说法，请考虑如下问题：为什么网络会消失？简单的答案是网络的消失是因为路由协议没有收敛。因为当所有的路由协议在收敛时产生路径循环，并且，当处于传输状态时，路由协议无法提供正确的转移信息，因此网络系统中发生变化后尽快地进行收敛是很重要的。

收敛路由协议所需要的时间依赖于以下两个因素：

- 参与收敛的路由器数量
- 它们必须处理的信息量

参与收敛的路由器数量取决于于拓扑结构变化必需传送的区域。聚和隐藏了来自路由器的信息，并且当通往目的地的路径改变了或不能达到时，那些不知道指定目的地的路由器不必再重新计算它们的路由表。

路由器为寻找到达任何目的地的最佳路径所必须处理的信息的量取决于可以达到指定目的地的路径数。同时，当网络结构的拓扑结构改变时，聚和也会减少路由器必须处理的信息的量。

所以，聚和是减少参与收敛的路由器数量和收敛时路由器所必须处理的信息量的关键。聚和依靠地址分配表进行，地址分配的好就会有好的聚和点。好的地址分配总是依靠一个好的潜在的拓扑结构。

为了聚和的发生，很难在一个结构拙劣的网络中进行地址分配。尽管许多人试图通过

更强大的路由器，把地址分配表调整得更好，或用更大、更好的路由协议来解决由拙劣的拓扑结构和地址分配方案所产生的问题，没有什么能够取代一个精心设计的拓扑结构。

正确的拓扑结构

那什么是正确的拓扑结构呢？比较容易解决问题的方法是，把它分成一些小部分加以解决，大规模网络也不例外。你可以将一个大规模网络系统分为几个可以分别处理的较小的部分。大多数成功的大规模网络系统是被分级设计的，或者说是分层的。分层产生了分开的问题区域，每层规划的重点放在单个目标或一组目标上。

这个概念类似于，将计算机之间的对话过程分为层，而每一层有不同的规划目标及标准的 OSI 模型。每一层必须尽可能地实现自身的规划目标；试图给某层增加太多的功能通常最终会造成记录和维护的困难。

通常分层结构的网络中定义了三层，如图 1-1 所示，每一层都有其自身的规划目标：网络核心层高速运送流量，其设备的主要工作是交换数据包。

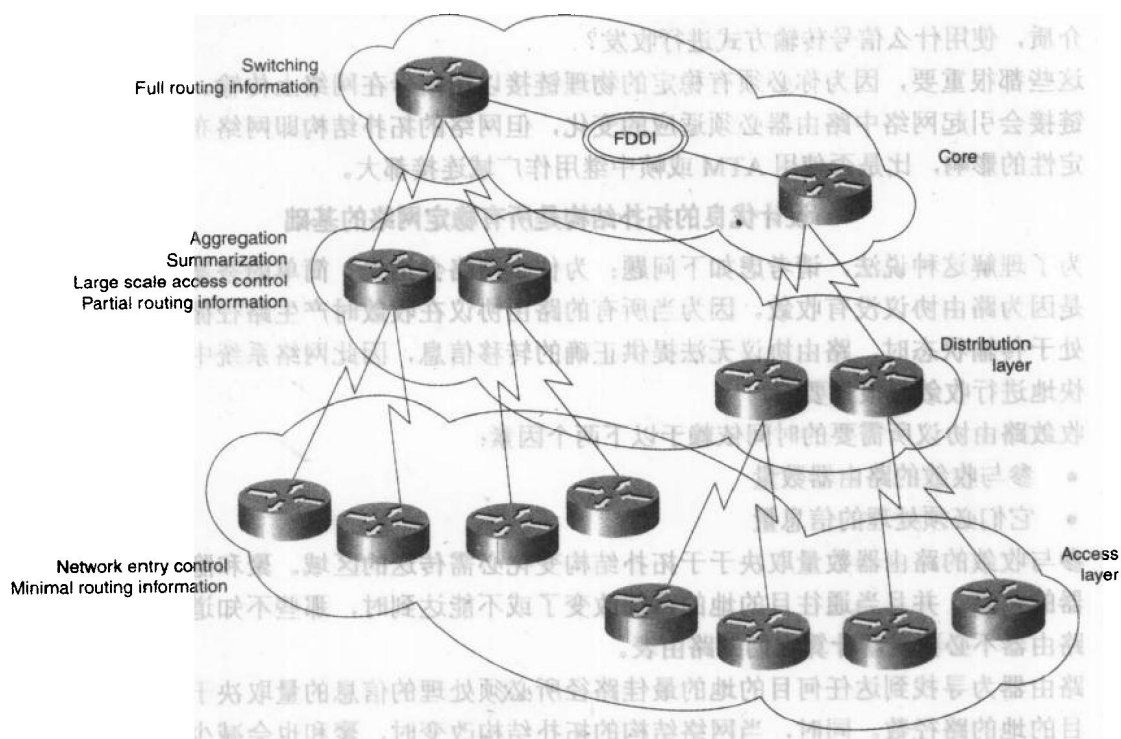


图 1-1 分层网络设计

分布层负责聚合路由路径，并且收敛数据流量。

访问层将流量馈入网络，执行网络访问控制，并且提供其它边缘服务。

知道了各层的名称后，现在回头看看它们与前面提及的基本规划原则之间的关系。下面是重新叙述的两点基本规划原则。下一步任务是看它们是否符合分层模型。

- 网络中受拓扑结构改变影响的区域应该被限制到最小的程度。
- 路由器（及其它网络设备）应传输尽量少的信息。

你可以通过聚和实现上述两个目标，并且聚和在分布层中进行。所以你通常希望将收敛区限定在分布层。

例如，一个失败的访问层链接不应该影响核心层的路由表，并且一个失败的核心层链接应对访问层的路由表产生最小的影响。

分层结构的网络中，流量在从访问层流向核心层时，被收敛在有较高速度的链接上，并且在从核心层流向访问层时，被分离到较小的链接上，如图 1-2 所示。这不仅意味着访问层路由器可以是一些较小的设备，也意味着它们交换数据包需要较少的时间，因此，就具备了更强的用来执行网络策略的处理能力。

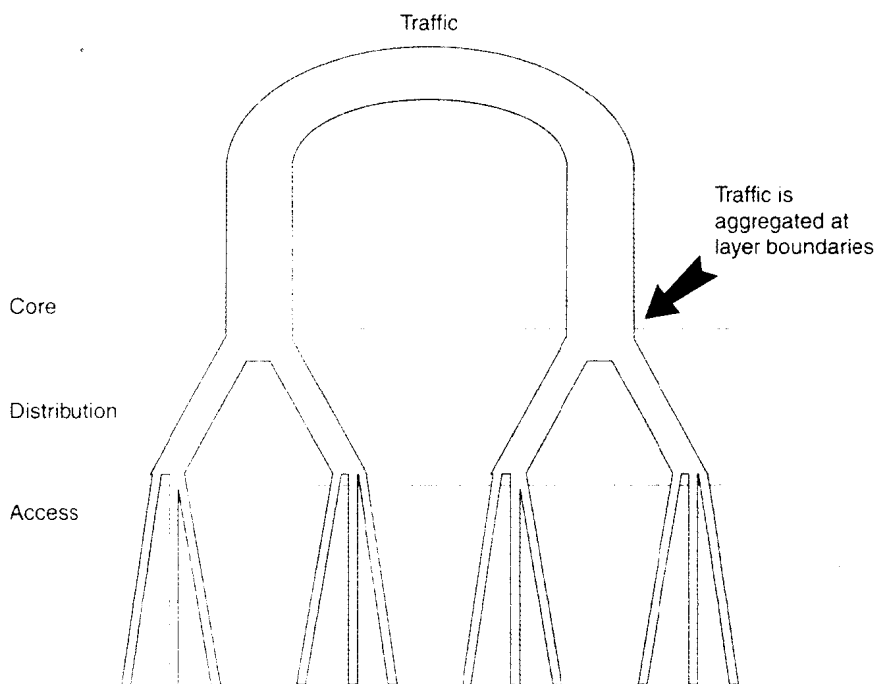


图 1-2 层体中的流量收敛和路由总汇

分层网络中一个固有的主要缺陷是它在物理层内隐含（或者产生）单个故障点。分层模型越强壮，就越有可能发现某一设备或某一失败的链接能导致网络遭到严重损坏的地方。当然如果你不愿意碰到这种麻烦，你的网络就必须用些冗余的手段来补偿这些缺陷。该主题详见第三章——“冗余”。

网络核心层

网络核心层有一个目的，即交换数据包。就像开足马力的机器一样，核心层设备应该得到充分的支持并以峰值性能运行。以下两个基本策略将有助于实现此目标：

- 不要在网络核心层执行网络策略。
- 核心层的所有设备应对网络中的每个目的地具备充分的可达性。

不执行网络策略

任何形式的策略必须在核心层外执行；数据包的过滤和策略路由是两个极好的例子。即使核心层设备可以高速率地过滤和策略路由数据包，核心层也不是执行这些任务的合适地方。网络核心层的任务是交换数据包，严禁采用任何降低核心层设备处理能力或是增加数据包交换延迟时间的方法。

此外,应尽量避免增加核心层路由器配置的复杂程度。它是网络边界策略执行出错并且导致用户组间失去连通性的可能因素，但是当核心层执行策略变化时出错将会导致整个网络瘫痪。

将网络策略的执行放在访问层的边界设备上，或者在某些情况下，可放在访问层与分布层之间的边界上。只有在某些特殊情况下，你才可以将其放在核心层中或核心层与分布层之间。

案例学习：基于策略的路由

一般而言，路由器根据最终目的地的地址发送信息流量，但某些情况下，你希望路由器基于源地址、流量类型或其它标准做出主动的决定。这些基于某一标准或由系统管理员配置好策略的主动的决定称为基于策略的路由。

路由器可配置为基于以下方法确定发送方式：

- 源地址
- 源/目的地地址对
- 目的地地址
- IP 数据包类型(TCP,UDP,ICMP 等等)
- 服务类型(Telnet,FTP,SMTP)
- IP 包头中的优先位

典型地说，基于策略的路由包括以下三步：

1. 建立过滤器，将需要应用特殊策略的流量从一般流量中分离出来。
2. 建立一个策略。
3. 执行该策略。

在 Cisco 路由器中，一个策略是通过使用路由映射建立，并且用接口命令来实现的。

如图 1-3 所示的管理员已经决定的系统，最佳方案是用较低速的帧中继链接来发送 Telnet，并通过卫星线路链接传送其余流量。

为应用该策略,网络管理员可对两个路由器都进行如下配置：

1. 为分离流量建立过滤器。

```
access-list 150 permit any eq telnet any
```

```
access-list 150 permit any any eq telnet
```

在第一行中选择发往 Telnet 端口的任意 TCP 流量；第二行以 Telnet 端口为源选择任意的 TCP 流量。

2. 建立策略。

```
route-map telnetthroughframe permit 10
```

```
match ip address 150
```

```
set ip next-hop 192.168.10.x
```

这几行建立了一个匹配前面选择的任一数据包的路由映射（所有数据包都源自或发往 TCP Telnet 端口），接着将这些数据包转发到帧中继链接的另一端路由器的 IP 地址上。

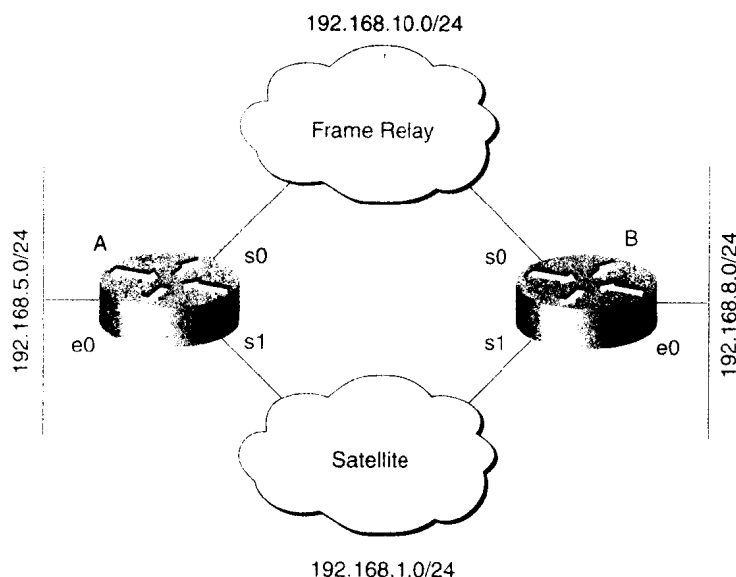


图 1-3 访问控制过滤器

3. 应用策略。

```
interface ethernet 0
ip policy route-map telnetthroughframe
```

- 最后，告诉路由器每个在以太网 0 接口接收的数据包都需要运用第三步中的策略。按策略路由的数据包以所有的直到 11.3 版的 IOS（因特网操作系统，Internet Operation System）版本的形式被交换。交换数据包的处理过程对路由器有极大的负冲击。每个需要交换的数据包都必须是有计划的交换，而不是通过一个优化了的交换路径来交换。

充分的可达性

核心层设备应该具有足够的路由信息来智能地交换发往网络中任意端设备的数据包；核心层的路由器不应该使用默认的路径到达内部的目的。但这并不意味着该层上的路由器应该有到达该网每个角落中独立子网的路径。聚和路径能够也应该用来减少核心层路由表的大小。默认路径应该用来到达外部的目的地，例如因特网上的主机。

采用没有默认路径的策略有三条原因：

- 便于核心层的冗余
- 减少次优化的路径
- 阻止路由循环

每次交换都计算的话，核心层的流量处于最大的状态。次优化路径在此种环境下是不