



H3CIE认证
参考用书

一部全面剖析**IP路由技术**的中文典籍
一部让初学者迅速掌握**网络核心技术**的秘籍



IP路由

技术详解与配置实践

尹光成 编著

从路由基础开始，全面介绍**RIP、OSPF、IS-IS、BGP、MBGP**等常见路由协议及路由引入、路由策略、策略路由、**BFD、GR**等路由技术，用例丰富，便于理解



清华大学出版社

IP 路由技术详解与配置实践

尹光成 编著

清华大学出版社
北 京

内 容 简 介

本书是一本详细、完整、全面介绍 TCP/IP 路由技术的专业书籍, 堪称有关路由技术方面不可多得的中文经典。本书共分为 4 个部分: 第一部分(IP 路由基础)主要介绍路由基础知识和静态路由, 包括路由技术概述、路由控制与转发、路由负载分担与备份、路由聚合与 CIDR、路由协议、静态路由等; 第二部分是本书的精华, 这部分详细、深入地描述了各种常用的 IGP 协议(如 RIP、OSPF、IS-IS 等), 除了对协议的实现原理、配置参数做了详细描述外, 还通过大量的配置示例演示了如何在 H3C 设备上部署这些协议, 以帮助读者获取大量解决实际问题的专业技能; 第三部分(EGP 路由协议)详细介绍了 BGP 协议和 MBGP 协议, 并通过大量配置示例使读者对 EGP 协议有更深入的认识; 第四部分(路由高级特性)详细讲述了路由引入、路由策略、策略路由、BFD 及 GR 等路由技术。本书对协议的深入分析和大量配置示例的演示相信会让读者受益匪浅。

本书语言通俗易懂、形象生动、易于理解, 并配备了大量的配置示例, 便于读者掌握高深复杂的路由技术知识。

本书可以作为 H3CIE 备考的参考书, 也可以作为深入理解 IP 路由技术的专业资料, 更适用于高校教师及工程技术人员。

本书封面贴有清华大学出版社防伪标签, 无标签者不得销售。

版权所有, 侵权必究。侵权举报电话: 010-62782989 13701121933

图书在版编目(CIP)数据

IP 路由技术详解与配置实践/尹光成编著. --北京: 清华大学出版社, 2012.3
ISBN 978-7-302-27985-3

I. ①I… II. ①尹… III. ①计算机网络—通信协议—路由选择 IV. ①TN915.04

中国版本图书馆 CIP 数据核字(2012)第 018698 号

责任编辑: 汤涌涛

封面设计: 杨玉兰

责任校对: 王 晖

责任印制: 杨 艳

出版发行: 清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址: 北京清华大学学研大厦 A 座 邮 编: 100084

社总机: 010-62770175 邮 购: 010-62786544

投稿与读者服务: 010-62776969, c-service@tup.tsinghua.edu.cn

质 量 反 馈: 010-62772015, zhiliang@tup.tsinghua.edu.cn

课 件 下 载: <http://www.tup.com.cn>, 010-62791865

印 刷 者: 北京密云胶印厂

装 订 者: 北京市密云县京文制本装订厂

经 销: 全国新华书店

开 本: 190mm×260mm

印 张: 31.75

字 数: 768 千字

版 次: 2012 年 3 月第 1 版

印 次: 2012 年 3 月第 1 次印刷

印 数: 1~4000

定 价: 60.00 元

产品编号: 041236-01

推荐序

——天边的云，脚下的路

从 IT 技术领域来看，互联网正在深刻地影响着人们生活的方方面面，无论你是否做好了准备，我们都已经进入网络改变生活的快车道。在这个快车道上，各种思想、火花、创意不断涌现，在计算机、智能终端所组成的“网络互联社会”中碰撞、激荡，从而产生更多的需求、应用、解决方案。近几年，伴随数据中心、虚拟化等应用而生的“云计算”、“物联网”，也都是这些解决方案的最新表现形式，而促使这一切快速变化的重要物质基础就是 IP 网络产品技术。

以虚拟化技术为代表的“云”已经升起在天边，对所有从事 IP 网络建设和维护的技术人员而言，都想掌握它的规律，那么脚下的路在那里？

IP 网络产品技术虽然涵盖门类繁多，应用发展迅速，但路由技术仍是 IP 网络技术中最为重要的基础和核心。因此，从路由入手，是快速掌握 IP 网络技术的必由之路，也是运维好“云”的必然选择。

H3C 公司作为 IP 领域的领导厂商，其所推出的 H3C 认证已经成为中国 IP 技术认证领域最具含金量的认证品牌，在 H3C 认证体系中 H3CIE 作为整个认证的最高端品牌，也一直引领着 IP 网络技术培训领域的方向。尹光成作为 H3C 认证培训领域的“终身金牌荣誉讲师”和首批 H3CIE，依托他多年网络实践的总结和授课经验的提炼，推出了这本《IP 路由技术详解与配置实践》，恰恰为我们提供了一本优秀的参考书。通过这本书，读者可以从路由基础入手，对常见路由协议原理、应用特点、配置实践等方面进行全面的了解，还可以通过作者亲自验证和编排的各个案例体会路由技术应用的特点，既可以作为 IP 技术专业人士之间经验的分享，也可以作为 H3CIE 等认证备考人员的学习参考书。

“不积跬步，无以至千里；不积小流，无以成江海”，正是来自光成自己平时一点一滴的技术积累和努力，才有了这本书。也感谢光成为 IP 网络技术推广所做的努力，高兴的看到，它不是 RFC 的简单堆砌，也不是国外资料的生硬照搬，是国人自己原创的努力。所有这些努力，最终让我们脚下的路走得更踏实，而让天边的“云”更实际。

李 林

杭州华三通信技术有限公司
全球技术服务部 副总裁
认证培训开发委员会 主任

前言

今天，随着无线城市、物联网、云计算的提出，以及触屏技术、平板电脑的发展，网络已经融入了人们的日常生活；据中国互联网络信息中心 CNNIC 统计，截至 2011 年 6 月，中国的网民数量达到 4.85 亿，域名总数为 786 万个，互联网普及率达到 36.2%。在很多行业，网络作为一个平台成为了支撑业务运行不可或缺的一部分，如银行跨行转账、证券实时交易、电力系统调度、政府电子政务等，所以我们相信网络能连接整个世界，实现无数的可能、创新和发明。

而这一切，在 24 年前是多么的不可思议。1987 年 9 月 20 日，钱天白教授利用了分组交换网络发出了我国第一封电子邮件，揭开了中国人使用互联网的序幕。但美国却借口安全问题，不对中国开放互联网。经各方不懈努力，1994 年 4 月 20 日，在中科院计算机网络信息中心（也就是今天的中国互联网络信息中心 CNNIC），中国实现了与国际互联网的全功能连接，成为国际互联网大家庭的第 77 个成员。就这么一连，给中国带来了巨大的资源，特别是全球互联网发展的动态，对中国科研事业的发展促进很大，从此，中国进入了互联网时代。

当今，IT 业务 IP 化，已经成为了不可否认的事实。而云计算和移动互联网两大趋势，必将推动 IT(Information Technology, 信息技术)和 CT(Communication Technology, 通信技术)的融合，它们将使全球范围内的数据流量、新服务和用户出现巨大增长。未来 10 年，这些趋势将给技术行业，特别是网络创新带来深远影响。

而这一切的发展和变化离不开底层网络平台的驱动和承载。在众多网络技术中，路由是最核心的一块，它可以使全球物理互联变为逻辑业务上的互通，可以指引数据到达互联网上的任意一个地方。没有路由就没有网络的互通，没有路由就没有今天的互联网。可以说，路由是网络技术的核心，更是互联网络的核心。互联网的飞跃发展，势必需要大量的网络技术人员来服务于这个行业并推动这个行业的发展。但并不是每一个网络技术人员都能熟练地掌握路由这门技术！

据权威部门统计，国内有近 2000 所高校开设的计算机院系中设有网络技术专业，在每年 600 多万的毕业生中，有近 20 万人来自网络技术专业，加上社会上的各类培训学校培训出的网络技术人员，每年涌入网络技术行业的人员有近 30 万人。大浪淘沙，实际上每年能真正在网络技术行业留下来的却不到 5 万人。究其缘由，绝大多数都因只是在网络技术边缘徘徊，没有掌握核心网络技术，没能登堂入室，而最终被市场所淘汰。所以，要想继续留在推动未来社会革新发展的网络技术行业，就必须掌握核心的技术，站在金字塔的塔尖，只有这样，你才会觉得这个行业是多么的美好。

纵观网络技术，大体可以为四大块：广域网、交换技术、路由技术和 VPN 技术。交换技术简单，但种类繁多，实现局域网/数据中心的互联；广域网将不同地域的局域网彼此互联起来，技术单一，受制于线路运营商；VPN 实现数据在公网上的安全传输或资源上的隔离(MPLS/VPN)，配置实现复杂，大多需安全设备辅助；路由则实现跨网段或跨网络的互通，原理复杂、种类繁多且应用场合各有不同，堪称网络技术核心之核心。故要站在网络技术金字塔

的塔尖，就必须熟练掌握路由技术。

10 年前华为发布了中国第一家建立国际规范的完整网络技术认证体系，笔者便进入这个行业，期间不断地努力寻找到达金字塔塔尖的秘籍，无奈，得到的要么是没有一个汉字的文献典籍，要么是甚难理会的中文翻译，个中苦涩，可想而知，作为网络技术的传播者，写一本通俗易懂、深入浅出的网络核心书籍是我多年的心愿。

现在，华三继承并发扬了华为的 IT 认证体系，秉承“专业务实，学以致用”的理念，于 2006 年 9 月 27 日在北京嘉里中心发布顶级认证——H3CIE(H3C Certified Internetwork Expert,H3C 认证互联网络专家)。自发布以来因考试难度大，参考资料缺乏让很多人望而却步，通过者非常之少，我很幸运，在 2007 年 5 月，第一个一次性通过了笔试认证，随后取得了 NO.00200 的 H3CIE 编号。在随后的几年里，一直有朋友建议我将备考的资料和心得加以整理，出版成书，供广大备考人员参考。几经思考，要写一本关于 H3CIE 的完整资料，难度很大，精力也有限，一时无法完成。而在 H3CIE 考试中，最难而最变幻莫测的莫过于路由技术，一旦掌握了路由技术，也就掌握了通过 H3CIE 的法宝，当然也掌握了网络技术的核心，自然也就到达了金字塔的顶峰。

本书共分为四个部分：第一部分是 IP 路由基础，主要介绍路由基础知识和静态路由，其中包括路由技术概述、路由控制与转发、路由负载分担与备份、路由聚合与 CIDR、路由协议基本、静态路由等；第二部分是本书的精华，详细、深入地描述各种常用的 IGP 协议，如 RIP、OSPF、IS-IS 等，除了对协议的实现原理、配置参数做了详细描述外，还通过大量的配置示例演示如何在 H3C 设备上部署这些协议，以帮助读者获取大量解决实际问题的专业技能；第三部分是 EGP 路由协议，详细介绍 BGP 协议和 MBGP 协议，通过大量配置示例让读者对 EGP 协议有更深入的认识；第四部分是路由高级特性，详细讲述路由引入、路由策略、策略路由及 BFD、GR 等最新的路由技术。

本书力求以通俗易懂、形象生动的语言来描述复杂的理论，并配备大量的配置示例，以帮助读者快速掌握高深复杂的路由技术。

本书不仅可以作为 H3CIE 的备考的专业参考书，也可以作为深入理解 IP 路由技术的专业参考资料，同样适用于高校教师及工程技术人员。本书对协议的深入分析和大量配置示例的演示相信会让读者受益匪浅。

本书历经整整一年的创作，在编写过程中得到了清华大学出版社的大力支持和帮助，得到了华三培训中心领导李林(H3C 全球技服副总裁)、刘凤敏(H3C 培训管理部部长)、陈喆(H3C 培训开发部部长)、朱冬光(H3C 培训市场部部长)的大力支持，并授权将本书为“H3CIE 认证参考用书”，同时也得到了深圳齐普生公司乔东斌总经理、高勇副总经理等领导等诸多同事的支持和帮助，在此，表示衷心的感谢。当然，还要感谢我的妻子，是她照顾家庭及刚出生不久的儿子，让我无后顾之忧，在业余时间能安心的创作。

由于作者水平有限，书中难免存在不足之处，恳请读者指正，如有任何建议或意见，请发至电子邮箱 ygc424@126.com。

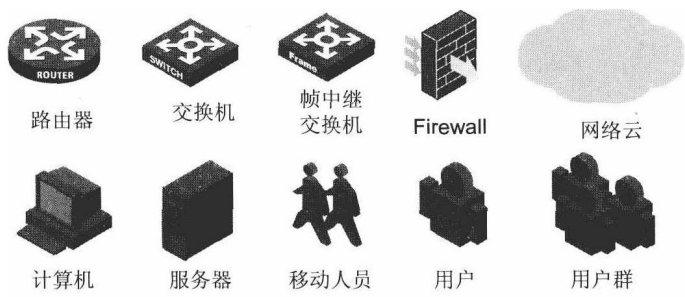
编者
于郑州

内容约定

为便于读者阅读本书，对本书中的图标、命令格式、设备及版本、设备及端口命名等做了一些约定。

1) 图标

本书中使用的图标及其含义如下所示。



2) 命令格式

本书所用命令及参数格式与 H3C Comware 操作系统命令格式完全一致，命令格式约定如下表所示。

格 式	意 义
斜体	命令行参数(命令中必须由实际值进行替代的部分)采用斜体表示
[]	表示用[]括起来的部分在命令配置时是可选的
{ x y ... }	表示从两个或多个选项选取一个
[x y ...]	表示从两个或多个选项选取一个或者不选
{ x y ... } *	表示从两个或多个选项选取多个，最少选取一个，最多选取所有选项
[x y ...] *	表示从两个或多个选项选取多个或者不选
&<1~n>	表示符号&前面的参数可以重复输入 1~n 次
#	由#号开始的行表示为注释行

3) 设备及版本

本书典型配置举例中所使用的路由器均为 H3C MSR 20 系列路由器，软件版本为 Comware 5.20, Release 2104, Standard。

```
<H3C> display version
H3C Comware Platform Software
Comware software, Version 5.20, Release 2104, Standard
Copyright (c) 2004-2010 Hangzhou H3C Tech Co., Ltd. All rights reserved.
H3C MSR20-40 uptime is 0 week, 0 day, 0 hour, 20 minutes
```

4) 设备命名

- 路由器命名为：RT-1、RT-2 及 RT-3，依此类推。
- 交换机命名为：LS-1、LS-2 及 LS-3，依此类推。
- PC 命名为：PC-1、PC-2 及 PC-3，依此类推。

5) 端口命名

(1) 路由器端口。路由器广域网口 Serial 简写为 S，端口号按实际编写如 S1/0；百兆以太网口 Ethernet 简写为 E，端口号按实际编写如 E1/0。

(2) 交换机端口。交换机百兆以太网口 Ethernet 简写为 E，端口号按实际编写如 E1/0/1。

目录

第 1 章 IP 路由概述1	
1.1 路由技术概述.....1	
1.1.1 什么是路由.....1	
1.1.2 什么是路由表.....1	
1.1.3 路由器对数据的转发操作.....3	
1.1.4 路由的来源.....5	
1.1.5 路由的度量值.....6	
1.1.6 路由优先级.....6	
1.1.7 路由环路.....7	
1.1.8 如何查看路由表.....8	
1.2 路由控制与转发.....14	
1.2.1 路由的控制与转发平面.....14	
1.2.2 FIB 表.....16	
1.2.3 快速转发表.....18	
1.3 路由负载分担与备份.....20	
1.3.1 负载分担.....20	
1.3.2 路由备份.....21	
1.4 路由聚合与 CIDR.....23	
1.4.1 路由聚合.....23	
1.4.2 路由聚合中环路的产生及避免.....23	
1.4.3 VLSM 与 CIDR.....25	
1.5 路由协议基础.....26	
1.5.1 路由协议分类.....26	
1.5.2 动态路由协议.....29	
1.5.3 路由信息的共享.....30	
1.5.4 衡量路由协议的主要指标.....30	
1.5.5 各路由协议比较.....31	
1.5.6 路由选择原则.....31	
第 2 章 静态路由33	
2.1 静态路由.....33	
2.1.1 静态路由的配置.....33	
2.1.2 静态路由配置示例.....34	
2.1.3 用静态路由实现路由备份和负载分担.....36	
2.1.4 静态黑洞路由.....40	
2.1.5 静态路由应用.....40	
2.2 缺省路由.....40	
2.2.1 配置缺省路由.....40	
2.2.2 缺省路由的两种生成方式.....41	
2.2.3 缺省路由应用.....42	
2.3 IPv6 静态路由.....42	
2.3.1 IPv6 静态路由简介.....42	
2.3.2 配置 IPv6 静态路由.....42	
2.3.3 IPv6 静态路由显示与维护.....42	
2.3.4 IPv6 静态路由配置示例.....42	
2.4 知识扩展.....44	
2.4.1 Null 接口.....44	
2.4.2 ping 命令详解.....44	
第 3 章 RIP 协议48	
3.1 RIP 协议的原理.....48	
3.1.1 RIP 协议由来.....48	
3.1.2 RIP 协议概述.....48	
3.1.3 RIP 路由表的初始化.....49	
3.1.4 RIP 路由表的更新.....49	
3.1.5 RIP 路由表的维护.....51	
3.1.6 路由扩散与收敛.....52	
3.1.7 单路径网络中路由环路的产生与避免.....52	
3.1.8 多路径网络中路由环路的产生与避免.....56	
3.1.9 水平分割与 NBMA 网络.....59	
3.1.10 RIP 的版本.....60	
3.1.11 RIP 的报文格式.....61	
3.2 RIP 的配置.....63	

3.2.1 启动 RIP 及配置接口 运行 RIP	63	4.1.7 OSPF 的邻居状态迁移	107
3.2.2 配置 RIP 的接口工作状态	64	4.1.8 OSPF LSDB 更新	108
3.2.3 配置 RIP 版本	64	4.1.9 OSPF 的区域划分	109
3.2.4 配置接口附加度量值	65	4.1.10 OSPF 的区域间通信	110
3.2.5 配置 RIPv2 路由聚合及 禁止接收主机路由	65	4.1.11 与自治系统外部通信	112
3.2.6 配置 RIP 发布缺省路由	66	4.1.12 OSPF 协议的 Cost 值	114
3.2.7 配置 RIP 路由过滤	66	4.1.13 OSPF LSA 类型	114
3.2.8 配置 RIP 优先级	67	4.2 OSPF 的高级特性	119
3.2.9 配置 RIP 引入外部路由	67	4.2.1 OSPF Stub 区域	119
3.2.10 调整和优化 RIP	68	4.2.2 OSPF NSSA 区域	120
3.2.11 RIP 显示和维护	70	4.2.3 OSPF 的路由聚合	122
3.3 RIP 典型配置实践	71	4.2.4 OSPF 的缺省路由	122
3.3.1 RIPv1 配置示例	71	4.2.5 OSPF 的安全特性	124
3.3.2 RIPv2 配置示例	74	4.2.6 OSPF 的路由过滤	125
3.3.3 RIPv2 认证配置示例	76	4.2.7 OSPF 的路由选路	126
3.3.4 RIP 接口 Cost 值配置示例	78	4.2.8 OSPF 的附录 E 问题	127
3.3.5 RIP 发布聚合路由配置示例	80	4.2.9 Forwarding Address	129
3.3.6 RIP 在 NBMA 网络中的应用	82	4.3 OSPF 的配置	130
3.4 RIPng 的原理与配置	85	4.3.1 配置 OSPF 基本功能	130
3.4.1 RIPng 的工作机制	85	4.3.2 配置 OSPF 区域	131
3.4.2 RIPng 报文格式	85	4.3.3 配置 OSPF 网络类型	134
3.4.3 RIPng 报文处理过程	87	4.3.4 配置 OSPF 路由信息控制	136
3.4.4 RIPng 的配置与显示	87	4.3.5 配置 OSPF 网络调整优化	143
3.4.5 RIPng 的配置示例	87	4.3.6 OSPF 的显示和维护	150
3.5 RIP 的故障诊断与排除	88	4.4 OSPF 典型配置实践	151
3.5.1 收不到邻居发送的更新报文	88	4.4.1 OSPF 单区域配置示例	151
3.5.2 RIP 网络发生路由震荡	89	4.4.2 OSPF 多区域配置示例	159
第 4 章 OSPF 协议	90	4.4.3 DR/BDR 选举配置示例	165
4.1 OSPF 的基本原理	90	4.4.4 NBMA 网络类型配置示例	168
4.1.1 OSPF 协议概述	90	4.4.5 OSPF 虚连接配置示例	172
4.1.2 OSPF 分层结构	93	4.4.6 OSPF 验证配置示例	180
4.1.3 Router ID 与网络类型	95	4.4.7 OSPF 多进程配置示例	184
4.1.4 OSPF 的报文类型	96	4.4.8 OSPF 引入外部路由 配置示例	189
4.1.5 邻居与邻接关系	101	4.4.9 OSPF 缺省路由引入 配置示例	192
4.1.6 DR/BDR 选举	104	4.4.10 OSPF Cost 值配置示例	193
		4.4.11 Stub 区域配置示例	196

4.4.12	Totally Stub 区域配置 示例.....	200	5.3.5	配置 IS-IS 主机名映射	276
4.4.13	NSSA 区域配置示例	201	5.3.6	IS-IS 显示和维护	277
4.4.14	Totally NSSA 区域 配置示例.....	205	5.4	IS-IS 典型配置实践	278
4.4.15	OSPF 路由聚合配置示例	206	5.4.1	IS-IS 单区域配置示例	278
4.4.16	OSPF 路由过滤配置示例	209	5.4.2	IS-IS 多区域配置示例	285
4.4.17	OSPF 静默接口配置示例	212	5.4.3	IS-IS DIS 选择配置示例	290
4.4.18	OSPF 选路配置示例	212	5.4.4	IS-IS 层次划分配置示例	293
4.4.19	OSPF 日常网络中的 应用配置示例.....	221	5.4.5	IS-IS 路由引入配置示例	297
4.4.20	OSPF 的其他应用 注意事项.....	222	5.4.6	IS-IS 验证配置示例	299
4.5	OSPFv3	226	5.4.7	IS-IS 路由聚合配置示例	303
4.5.1	OSPFv3 简介	226	5.5	IS-IS 对 IPv6 的支持.....	305
4.5.2	OSPFv3 配置	228	5.5.1	IS-IS IPv6 简介.....	305
4.5.3	OSPFv3 典型配置举例	228	5.5.2	IS-IS IPv6 的基本配置.....	306
4.6	OSPF FAQ.....	231	5.5.3	IS-IS IPv6 基本配置示例.....	306
第 5 章	IS-IS 协议	243	第 6 章	BGP 协议.....	309
5.1	IS-IS 的基本概念	243	6.1	BGP 协议的基本原理	309
5.1.1	IS-IS 概述	243	6.1.1	BGP 协议概述.....	309
5.1.2	IS-IS 协议的基本概念	244	6.1.2	BGP 协议的特性	309
5.1.3	IS-IS 分层网络	246	6.1.3	BGP 的基本术语	310
5.1.4	IS-IS 与 OSPF 的对比	248	6.1.4	BGP 的同步	312
5.2	IS-IS 协议的基本原理	248	6.1.5	BGP 消息.....	313
5.2.1	OSI 地址	248	6.1.6	BGP 的状态机.....	319
5.2.2	IS-IS 协议报文	251	6.1.7	BGP 的路由属性介绍	320
5.2.3	IS-IS 网络类型	256	6.1.8	BGP 对路由的处理流程.....	321
5.2.4	邻居关系的建立.....	256	6.1.9	BGP 的路由选路策略	322
5.2.5	邻接关系的建立.....	257	6.1.10	BGP 负载分担时的 选路策略.....	323
5.2.6	LSDB 的同步	259	6.1.11	BGP 的路由发布策略.....	323
5.2.7	拓扑结算与 IP 路由的形成	262	6.2	BGP 的路由属性	324
5.2.8	路由渗透.....	264	6.2.1	Origin 属性	324
5.3	IS-IS 的配置	264	6.2.2	AS_Path 属性	324
5.3.1	配置 IS-IS 基本功能	264	6.2.3	AS_Path 属性的作用	325
5.3.2	配置 IS-IS 路由信息控制	266	6.2.4	Next_Hop 属性	326
5.3.3	配置 IS-IS 调整与优化	270	6.2.5	MED 属性.....	327
5.3.4	提高 IS-IS 网络的安全性	275	6.2.6	Local_Pref 属性.....	328
			6.2.7	Community 属性.....	328
			6.3	BGP 的基本配置	329

6.3.1	配置 BGP 连接.....	329	7.2.2	MP-BGP for MPLS/VPN 基本配置.....	406
6.3.2	控制路由信息的生成.....	331	7.3	MP-BGP 协议典型配置实践.....	407
6.3.3	控制路由信息的 发布与接收.....	332	7.3.1	MP-BGP 在 IPv6 中的 配置示例.....	407
6.3.4	配置 BGP 的路由属性.....	335	7.3.2	MP-BGP 在 MPLS/VPN 中的 配置示例.....	409
6.3.5	调整和优化 BGP 网络.....	338	第 8 章	路由引入.....	415
6.3.6	BGP 显示与维护.....	340	8.1	路由引入.....	415
6.4	控制 BGP 路由.....	341	8.1.1	多路由协议网络与 路由引入.....	415
6.4.1	控制 BGP 路由概述.....	341	8.1.2	单向路由引入.....	417
6.4.2	利用 BGP 属性控制 BGP 路由.....	341	8.1.3	多向路由引入.....	418
6.4.3	利用过滤器控制 BGP 路由....	345	8.1.4	路由引入产生路由环路及 解决方法.....	418
6.5	BGP 在大规模网络中的应用.....	351	8.1.5	路由引入产生次优路由及 解决方法.....	419
6.5.1	大规模 BGP 网络概述.....	351	8.2	路由引入配置.....	420
6.5.2	配置 BGP 对等体组.....	352	8.2.1	配置 RIP 协议引入外部 路由.....	420
6.5.3	配置 BGP 团体属性.....	353	8.2.2	配置 OSPF 协议引入 外部路由.....	421
6.5.4	配置 BGP 聚合.....	354	8.2.3	配置 IS-IS 协议引入 外部路由.....	421
6.5.5	配置 BGP 反射.....	355	8.2.4	配置 BGP 协议引入 外部路由.....	422
6.5.6	配置 BGP 联盟.....	359	8.3	路由引入典型配置实践.....	422
6.5.7	配置 BGP 衰减.....	361	8.3.1	RIP 协议路由引入配置 示例.....	422
6.5.8	部署多出口 BGP 网络.....	362	8.3.2	OSPF 协议路由引入 配置示例.....	423
6.6	BGP 典型配置实践.....	365	8.3.3	IS-IS 协议路由引入 配置示例.....	425
6.6.1	BGP 基本配置示例.....	365	8.3.4	BGP 协议路由引入 配置示例.....	427
6.6.2	BGP 与 IGP 交互配置示例....	371	第 9 章	路由策略.....	432
6.6.3	BGP 负载分担配置示例.....	376	9.1	路由策略概述.....	432
6.6.4	BGP 反射配置示例.....	380	9.1.1	路由策略的作用.....	432
6.6.5	BGP 联盟配置示例.....	383			
6.6.6	BGP 路由聚合配置示例.....	386			
6.6.7	BGP 选路配置示例.....	389			
6.7	BGP FAQ.....	394			
第 7 章	MP-BGP 协议.....	399			
7.1	MP-BGP 协议原理.....	399			
7.1.1	MP-BGP 概述.....	399			
7.1.2	MP-BGP 的扩展属性.....	399			
7.1.3	MP-BGP 的应用.....	401			
7.2	MP-BGP 协议配置.....	405			
7.2.1	MP-BGP for IPv6 基本配置 ...	405			

9.1.2 路由过滤方法.....	433	10.3.3 基于报文地址的策略 路由配置示例.....	461
9.1.3 路由过滤工具.....	434		
9.2 使用静默接口过滤路由.....	435	第 11 章 BFD 技术.....	464
9.3 地址前缀列表.....	436	11.1 BFD 技术的原理.....	464
9.3.1 地址前缀列表概述.....	436	11.1.1 BFD 简介.....	464
9.3.2 地址前缀列表配置.....	437	11.1.2 BFD 的功能.....	464
9.3.3 地址前缀列表的应用.....	438	11.1.3 BFD 的工作机制.....	464
9.4 Filter-policy.....	438	11.1.4 BFD 的检测方式.....	466
9.4.1 Filter-policy 概述.....	438	11.1.5 BFD 会话的工作方式.....	467
9.4.2 Filter-policy 配置.....	440	11.1.6 BFD 的运行模式.....	467
9.5 Route-policy.....	440	11.1.7 BFD 的认证模式.....	467
9.5.1 Route-policy 概述.....	440	11.1.8 BFD 的报文格式.....	468
9.5.2 Route-policy 的组成.....	441	11.1.9 各协议与 BFD 的关系.....	469
9.5.3 Route-policy 的原理.....	441	11.2 BFD 的配置.....	469
9.5.4 Route-policy 的配置.....	442	11.2.1 BFD 基本功能配置.....	470
9.6 路由策略典型配置实践.....	444	11.2.2 静态路由 BFD 配置.....	470
9.6.1 使用路由策略过滤发布 路由配置示例.....	444	11.2.3 OSPF 路由协议 BFD 配置 ..	471
9.6.2 使用路由策略过滤接收 路由配置示例.....	446	11.3 BFD 与 NQA 的联动应用	471
9.6.3 使用路由策略过滤并设置 引入路由属性配置示例.....	448	11.4 BFD 典型配置实践.....	473
9.6.4 使用静默接口过滤路由 配置示例.....	450	11.4.1 静态路由与 BFD 联动 配置示例.....	473
第 10 章 策略路由 PBR.....	453	11.4.2 OSPF 与 BFD 联动 配置示例.....	475
10.1 策略路由原理.....	453	第 12 章 GR 技术概述.....	479
10.1.1 策略路由概述.....	453	12.1 GR 简介.....	479
10.1.2 策略路由的应用.....	454	12.1.1 GR 简介.....	479
10.2 策略路由配置.....	455	12.1.2 GR 的作用.....	479
10.2.1 策略路由的配置.....	455	12.2 GR 的基本概念.....	479
10.2.2 策略路由的显示与维护.....	457	12.2.1 GR Restarter.....	479
10.3 策略路由典型配置实践.....	457	12.2.2 GR Helper.....	479
10.3.1 基于协议类型的本地策略 路由配置示例.....	457	12.2.3 GR Session.....	479
10.3.2 基于报文大小的策略 路由配置示例.....	459	12.2.4 GR Time.....	479
		12.3 GR 的通信过程.....	480
		12.3.1 在 GR Restarter 和 GR Helper 之间建立 GR Session.....	480
		12.3.2 GR Restarter 重新启动.....	480

12.3.3 GR Restarter 向 GR Helper 发送信号.....	481	12.5 GR 典型配置实践	484
12.3.4 GR Restarter 从 GR Helper 获取拓扑或路由信息:.....	481	12.5.1 OSPF 的 GR 典型 配置举例.....	484
12.4 各协议的 GR 机制.....	481	12.5.2 BGP 的 GR 典型 配置举例.....	487
12.4.1 OSPF 的 GR 机制	482	附录 A 缩略语	489
12.4.2 IS-IS 的 GR 机制.....	483	附录 B H3CIE 认证简介	492
12.4.3 BGP 的 GR 机制	484		

IP 路由概述

1.1 路由技术概述

1.1.1 什么是路由

在 OSI(Open System Interconnect, 开放式系统互联)参考模型中, 位于网络层的设备——路由器——提供了将不同网络互连的机制, 实现将报文从一个网络转发到另一个网络, 从而实现不同网络之间的通信。

路由器转发报文时必须有一个依据, 这个依据就是路由。也就是说, 路由就是指导 IP 报文转发的路径信息。这好比要去一个陌生的地方, 一开始只知道目的地, 而不知道如何到达该目的地, 这时需要去查看地图, 地图上指示首先要到达 A 地, 在 A 地如何走才能到达 B 地, 到达 B 地后我们同样查找地图, 看如何到达 C 地; 依次类推, 最终便可到达该目的地。在这过程中从始发地到目的地的路径信息就相当于路由, 它指导人们每一步该如何走。

在因特网中进行路由选择需要使用路由器, 路由器根据所收到的 IP 报文中的目的地址选择一条合适的路径(也就是路由), 并将报文转发到在此路径上的下一个路由器。路径中最后的路由器负责将报文送交目的主机。报文在网络上的传递就像体育比赛中的接力赛一样, 报文就是接力棒, 路由器就是运动员, 每一个路由器只负责将报文在本站通过最优的路径转发, 通过多个路由器一站一站地接力, 最终将报文通过最优路径转发到目的地。当然也有一些例外情况, 比如有些网络中部署了路由策略, 从而使报文通过的路径并不一定是最优的。

路由器的转发特点是逐跳转发, 在如图 1-1 所示的网络拓扑中, Network A 的 IP 报文要想发送给 Network B, 首先发给 RT-1, RT-1 收到后根据报文的目的 IP 地址查找路由并将报文转发给 RT-2, RT-2 收到后根据报文的目的 IP 地址查找路由并将报文转发给 RT-3, RT-3 收到后将报文转发给 Network B。这就是路由转发的逐跳性, 即路由只指导本地转发, 而不影响其他设备转发, 设备之间的转发是相互独立的。

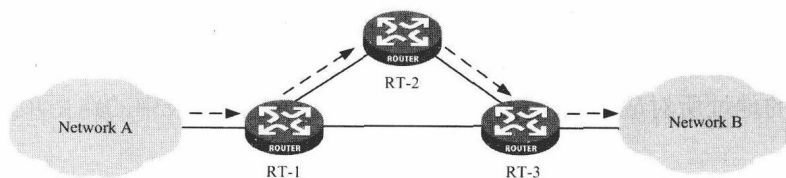


图 1-1 路由转发示意图

1.1.2 什么是路由表

路由器转发 IP 报文的关键是路由表。每个路由器中都保存着一张路由表, 表中每条路由项都指明了要到达某个子网或某台主机的 IP 报文应通过路由器的哪个物理接口发送, 才能到

达该路径上的下一个路由器，或者无须再经过别的路由器便可转发到直接相连的网络中的目的主机。

在如图 1-2 所示的拓扑中，RT-1 和 RT-2 连接了 1.1.1.0/24 和 3.3.3.0/24 两个网络，1.1.1.0/24 网络中的报文要想转发到 3.3.3.0/24 网络中，需要先将报文发送给 RT-1，RT-1 按照如表 1-1 所示的路由信息表将报文转发给 RT-2，RT-2 再转发给 3.3.3.0/24 网络。

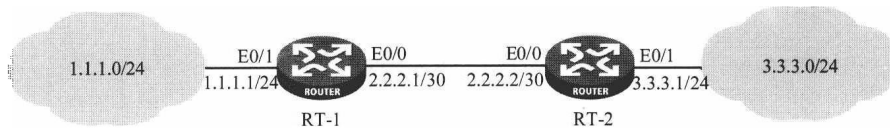


图 1-2 路由拓扑

表 1-1 路由信息表

路 由 器	目的地址	子网掩码	下 一 跳	出 接 口	度 量 值	优 先 级
RT-1	3.3.3.0	255.255.255.0	2.2.2.2	E0/0	10	60
RT-2	1.1.1.0	255.255.255.0	2.2.2.1	E0/0	5	10

如表 1-1 所示的路由表中包含了下列关键项。

- 目的地址：用来标识 IP 数据报的目的地址或目的网络。
- 子网掩码：与目的地址一起来标识目的主机或路由器所在的网段的地址。将目的地址和网络掩码“逻辑与”后可得到目的主机或路由器所在网段的地址。例如，目的地址为 3.3.3.1、掩码为 255.255.255.0 的主机或路由器所在网段的地址为 3.3.3.0。掩码由若干个连续“1”构成，既可以用点分十进制法表示，也可以用掩码中连续“1”的个数来表示。
- 下一跳：更接近目的网络的下一个路由器地址。如果只配置了出接口，那么下一跳 IP 地址是出接口的地址。
- 出接口：指明 IP 报文将从该路由器的哪个接口转发。
- 度量值：说明 IP 报文需要花费多大的代价才能到达目标。主要作用是当网络存在到达目的网络的多个路径时，路由器可依据度量值来选择一条较优的路径以转发 IP 报文，从而使 IP 报文能够更快、更好地到达目的地。
- 优先级：对于同一目的地，可能存在若干条不同下一跳的路由，这些不同的路由可能是由不同的路由协议发现的，也可能是手工配置的静态路由。优先级高(数值小)的路由将成为当前的最优路由。

根据路由目的地址所对应的子网掩码长度不同，可以把路由表中的路由项划分为以下 3 种类型。

- 子网路由：目的地为子网，掩码长度大于 0 但小于 32，表明匹配某一个子网。
- 主机路由：目的地为主机，掩码长度为 32 位的路由，表明匹配某一台主机。
- 默认路由：掩码长度为 0，表明此路由匹配全部的 IP 地址。

另外，根据目的地与该路由器是否直接相连，又可将路由表中的路由项划分两种。

- 直接路由：目的地所在网络与路由器直接相连。
- 间接路由：目的地所在网络与路由器非直接相连。

为了不使路由表变得过于庞大，可以设置一条缺省路由。凡数据报文查找路由表失败，便根据缺省路由转发。

1.1.3 路由器对数据的转发操作

路由器转发 IP 报文的依据是路由表，通过匹配路由表里的路由项来实现对 IP 报文的转发。

如图 1-3 所示，当路由器收到一个 IP 报文的时候，将报文中的目的 IP 地址提取出来，然后与路由表中路由表项包含的目的地址进行比较。如果与某路由项中的目的地址相同，则认为与此路由项匹配；如果没有路由项能够匹配，则丢弃该 IP 报文。

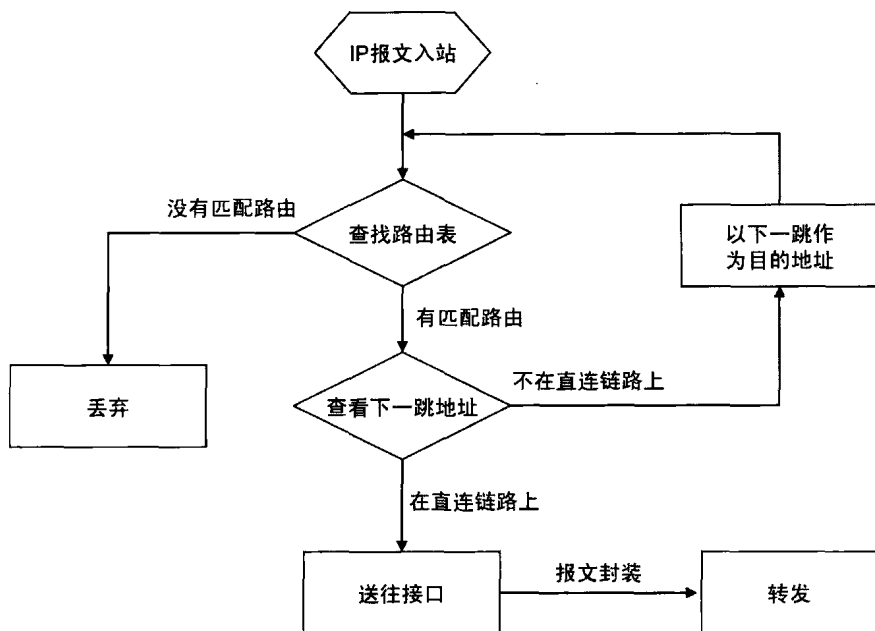


图 1-3 路由器多跳转发流程图

IP 报文中的目的 IP 地址往往是主机地址，而路由表中的目的地址往往为网络地址，怎么让二者匹配呢？这里面有个底层的操作：首先将 IP 报文中的目的地址和路由表项中的子网掩码进行“逻辑与”操作，得到一个网络地址，然后拿此网络地址与路由项中的网络地址做比较，如果一致就认为匹配，否则认为不匹配。

如果路由项匹配，则路由器查看所匹配的路由项的下一跳地址是否在直连的链路上。如果在直连的链路上，则根据此下一跳转发；如果不在直连的链路上，则需要在路由表中再次查找此下一跳地址所匹配的路由项。

确定了下一跳地址后，路由器将此报文送往对应的接口，接口进行相应的地址解析，解析出对应的链路层地址后，对 IP 报文进行数据封装并转发。

1. 路由最长匹配原则

当路由表中存在多个路由项可以同时匹配目的 IP 地址时，路由查找进程会选择其中掩码最长的路由项进行转发，掩码越长表明其代表的网络范围越小，匹配的程度就越精确。这就是所谓的最长匹配原则。