

ICT 认证系列丛书

HCIIE

路由交换学习指南

泰克教育集团◎主编
刘大伟 陈亮 丁琳琦◎编著



中国工信出版集团



人民邮电出版社
POSTS & TELECOM PRESS

ICT 认证系列丛书

HCI
HCIE

路由交换学习指南

泰克教育集团◎主编

刘大伟 陈亮 丁琳琦◎编著

人民邮电出版社
北京

图书在版编目 (C I P) 数据

HCIE路由交换学习指南 / 泰克教育集团主编 ; 刘大伟, 陈亮, 丁琳琦编著. -- 北京 : 人民邮电出版社, 2017.8

(ICT认证系列丛书)

ISBN 978-7-115-45000-5

I. ①H… II. ①泰… ②刘… ③陈… ④丁… III. ①计算机网络—路由选择 IV. ①TN915.05

中国版本图书馆CIP数据核字(2017)第045860号

内 容 提 要

本书是专注于讲解HCIE路由交换技术的学习指导书,书中对路由交换技术涉及的大部分协议都做了全面细致的讲解,针对重要协议原理及实现细节都结合实例加以系统的分析和归纳总结,通过学习本书读者能全面掌握路由交换理论中的知识。

本书共分十章,第一章从路由基础讲起,涉及IPv4和IPv6的基本概念;其后第二章到第五章分别阐述RIP、OSPF、IS-IS及路由控制技术;第六章和第七章重点阐述大型网络中BGP协议、组播路由协议(IGMP及PIM等)原理;第八章对MPLS及MPLS VPN技术做了深入阐述;第九章则对交换技术,包括VLAN技术、生成树技术及WAN技术加以阐述;第十章对华为的QoS实现做了探讨。

本书适合准备参加HCIE路由交换考试的考生或希望深入了解华为路由交换技术的工程技术人员学习使用,同时也适合对网络技术感兴趣的人员作参考学习之用。

-
- ◆ 主 编 泰克教育集团
编 著 刘大伟 陈 亮 丁琳琦
责任编辑 李 静
执行编辑 王国霞
责任印制 彭志环
 - ◆ 人民邮电出版社出版发行 北京市丰台区成寿寺路11号
邮编 100164 电子邮件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
大厂聚鑫印刷有限责任公司印刷
 - ◆ 开本: 787×1092 1/16
印张: 50 2017年8月第1版
字数: 1 167千字 2017年8月河北第1次印刷
-

定价: 128.00元

读者服务热线: (010) 81055488 印装质量热线: (010) 81055316

反盗版热线: (010) 81055315

序 一

斗转星移，物是人非，多少英雄尘与土，IT 江湖代有才人出！

都说“IT”就是“挨踢”的代名词，“网管员”已经成“网吧小伙计”的专属称谓。倒退 30 年，如果你说是搞计算机的，在女孩儿的心目中，那你可是“搞高科技”的优质男。

倒退 20 年，如果你说是搞互联网的，在女孩儿的心目中，你俨然又成了“无所不能”的神秘男。

倒退 10 年，如果你说是搞云计算的，在女孩儿的心目中，你依然可以得到“高看一眼”的待遇。

这曾经是几代 IT 男赖以“显摆”的经典招数之一，可当时间定格到 2017 年，新一代的 IT 男还能有类似的金字招牌招摇吗？

要回答这个问题，让我们看看现在的市场环境，“自主可控”上升到国家战略这几年，国内 IT 品牌市场份额不断刷新历史记录，“一带一路”更将为国内品牌走向世界强力助推。伴随着这种趋势，需要大量的 IT 人才满足各方面的信息化建设与使用需求，作为早就国际化的国内优秀 IT 认证品牌，华为 HCIE 认证越来越得到广大客户和学员的青睐，有望扛起这个大旗。

因业务机缘，自己有幸参与过华为企业业务 IT 认证体系设计的初期调研，知道为什么要在传统的笔试、实验的基础上，特别加上答辩的环节。几年的实践证明，这个决策是非常正确的，正是这个环节的严格把关，避免了简单的“啃教材、背版本、拼机时”套路，成为真金不怕火炼的试金石。

这种严苛，就得靠自学、自悟，很难提高学员的通过率，于是我们看到有的学员因不得要领而折戟，有的学员为得到某个高师的指点而四处求教，有的学员干脆放弃了努力……

确实，要靠真本事，就要学到真经，一本融理论与实践精华，深入浅出的 HCIE 辅导教材就成为学员们学习的关键。这里，非常感谢刘大伟、陈亮、丁琳琦三位老师的辛勤努力，终于给大家烹煮出了这道教学知识大餐！这三位老师都是业界教学的大拿，培训、培养了大量的 HCIE 人才，现在能够把多年的实践经验贡献出来，实属难能可贵，相信每个学员都能从中找到茅塞顿开的感觉。

如果你是新一代的 IT 男，那么就以此书为利剑，在 HCIE 的比武台上，一展英姿吧！

高洪福
华为本部技术中心总经理
神州数码集团股份有限公司
2017-5-18

序 二

2012年12月，华为公司推出了一项新的职业认证计划——HCIE（华为认证互联网专家）认证，该认证是为了适应全球对更多的ICT（信息与通信技术）技术专家的需求而推出的，是目前业界最权威的技术认证之一，同时该认证也是业界考取难度超大、含金量超高的认证。

《HCIE路由交换学习指南》由华为公司核心教育合作伙伴——泰克教育集团主编，由刘大伟、陈亮、丁琳琦三位泰克专家讲师基于上千名HCIE人才的培养经验编写而成。

本书编写的主要目的是帮助读者通过HCIE认证。本书覆盖了HCIE所有重要知识，无论你是要成为HCIE，还是为了对数据网络设计、故障诊断等有更深入的了解，阅读本书都会使你获益匪浅。

本书为泰克教育集团计划出版ICT系列教程中的其中一本，泰克将通过推出现有和未来的教科书的方式，来助力实现华为全球培训认证部的主要目标：培训华为用户群体ICT专业知识，使得该群体能够构建和维护可靠、弹性的数据信息网络，并通过严格培训来满足更多用户的学习需求。

最后希望本书对您开卷有益。

林康平
泰克教育集团 CEO
2017年5月

前 言

华为 HCIE 认证是目前 IT 界最为权威的认证考试之一,该认证考试包含 Lab 考试和面试考试两部分,其中的面试考试因考核的内容多、追问细而使 HCIE 认证成为目前 IT 界较难通过的考试,其含金量不言而喻。

华为 HCIE 路由交换方向的考试认证包含多种路由技术及交换技术,协议较多,原理较复杂,为了帮助考生丰富其理论知识并准备面试,泰克网络实验室的老师集中精力编写了这本《HCIE 路由交换学习指南》指导书,希望能助在准备 HCIE 路由交换认证的考生一臂之力,更重要的是,我们也希望考生通过本书的学习,能增进理论知识,提高分析问题及解决问题的能力,当然这也是华为推出 HCIE 路由交换认证考试的目的。

近些年来,网络技术更新的速度非常快,新的协议、新的技术层出不穷,但任何协议技术都是以路由交换技术为基础的,很多协议甚至在设计开发时都是相互借鉴,原理都彼此相通,所以学好路由交换技术除了能帮助我们提高理解能力之外,还有助于我们掌握其他新技术。

华为 HCIE 路由交换技术所涉及的内容非常多,本书在内容选择上,仅选取了一些较为重要的知识模块加以阐述,并没有覆盖到路由交换的全部内容,选取的内容包括 IGP、BGP、组播、MPLS L3VPN、交换及 QoS 技术。

书中内容按照如下顺序来组织。

第一章: IP 路由基础,包含 IPv4 及 IPv6 路由的基本概念。

第二章: RIP 及 RIPng 的协议原理。

第三章: 阐述 OSPF 协议原理及 OSPF 与 OSPFv3 协议的区别。

第四章: IS-IS 路由协议原理及其对 IPv6 的支持。

第五章: Route-control 技术,列举各种路由控制技术。

第六章: BGP 协议原理。

第七章: IP 骨干网络中组播路由协议,包括 IGMP、PIM、MSDP 及 mBGP。

第八章: 域内及域间 MPLS L3VPN 技术。

第九章: 交换技术,包括 VLAN、STP 及 WAN 技术。

第十章: QoS 实现,包括标记、队列、整形、监管及限速等。

本书在原理阐述过程中尽量结合实例或场景加以分析,力图透彻详尽地把概念阐述清楚。

本书在编写过程中,得到各地泰克老师的大力支持,在此一并表示诚挚的感谢。由于作者水平所限及时间仓促,书中难免存在一些谬误和不足之处,敬请读者批评指正,也欢迎读者或考生直接到泰克和我们一起探讨,我们会在本书的后续版本中对其加以

更新。

本书适合于准备华为 HCIE 路由交换认证的考生参考学习之用，本书同样适合于已通过 HCIE 认证的考生及对路由交换技术感兴趣的工程师学习使用，也可作为其他厂商的技术工程师熟悉华为技术的一本技术参考书籍。

作者

2017年5月

目 录

第一章 IP 路由基础	1
1.1 IPv4 地址规划	2
1.1.1 IPv4 地址	2
1.1.2 主网及子网划分	2
1.1.3 IPv4 报文	3
1.2 IP 路由表和 FIB	5
1.2.1 IP 路由转发和标签交换	5
1.2.2 IP 路由表和 FIB 表	5
1.2.3 路径选择	9
1.2.4 路由的度量	17
1.2.5 下一跳	17
1.3 有类和无类路由	18
1.3.1 有类路由协议与无类路由协议	18
1.3.2 有类查找方式和无类查找方式	18
1.4 静态路由	19
1.4.1 概述	19
1.4.2 静态路由与动态路由的区别	19
1.4.3 静态路由	19
1.5 动态路由协议	20
1.5.1 概述	20
1.5.2 动态路由协议的分类	20
1.5.3 矢量路由协议	21
1.5.4 链路状态路由协议	21
1.6 IPv6	23
1.6.1 IPv6 概述	23
1.6.2 IPv6 地址	25
1.6.3 IPv6 的报文格式	29

1.6.4 ICMPv6	38
1.7 NDP	45
1.7.1 NDP 概述	45
1.7.2 无状态自动配置	46
1.7.3 路由器发现	46
1.7.4 地址解析	47
1.7.5 DAD (重复地址检测)	49
1.7.6 重定向原理	51
1.8 思考题	51
第二章 路由信息协议	53
2.1 RIP 的基本知识	54
2.1.1 RIP 的基本原理	54
2.1.2 RIP 的定时器	54
2.1.3 距离矢量协议的问题	55
2.2 RIP 的报文及版本	56
2.2.1 RIP 的工作过程	56
2.2.2 RIPv1 的报文格式	56
2.2.3 RIPv2 的报文格式	57
2.2.4 RIPv1 和 RIPv2 的比较	58
2.2.5 RIP 的兼容版本	59
2.2.6 RIPv1 路由收发规则	60
2.2.7 案例研究: 运行不同版本能否正常收发路由	60
2.2.8 案例研究: 运行 RIPv1 路由能否正常收发路由	62
2.2.9 案例研究: 运行 RIPv1 传递路由的问题	64
2.3 RIP 的防环机制	66
2.3.1 水平分割	67
2.3.2 毒性逆转	67
2.3.3 最大计数器	68
2.3.4 触发更新	68
2.4 RIP 的高级特性	68
2.4.1 silent-interface 和 peer ip-address	68
2.4.2 metricin 和 metricout	69
2.4.3 RIPv2 的认证	69

2.4.4	路由聚合	71
2.4.5	案例研究: 自动路由聚合造成不连续子网问题	72
2.4.6	案例研究: 手动聚合造成的路由环路问题	73
2.5	RIPng 协议	73
2.5.1	RIPng 协议介绍	73
2.5.2	RIPng 报文格式	74
2.5.3	RIPng 协议特性	74
2.5.4	RIPng 下一跳	75
2.5.5	RIPng 配置实例	76
2.6	案例分析	82
2.6.1	案例 1: 接口故障造成的 RIP 收敛问题	82
2.6.2	案例 2: RIPv2 中 Next-hop 的作用	84
2.7	思考题	86
第三章 OSPFv2 及 OSPFv3		87
3.1	OSPF 邻居关系的建立及握手过程	88
3.1.1	概述	88
3.1.2	邻居发现	88
3.1.3	邻居关系建立过程	89
3.1.4	邻接关系建立过程	91
3.1.5	影响邻居关系及邻接关系建立的问题	93
3.1.6	OSPF 网络类型	95
3.2	数据库同步及泛洪机制	97
3.2.1	OSPF 报文结构	97
3.2.2	LSA 格式及类型	101
3.2.3	泛洪机制	102
3.3	拓扑描述及路由计算	105
3.3.1	LSA 1 和 LSA 2 内容分析	105
3.3.2	DR 及虚节点作用	110
3.3.3	SPF 路由计算	112
3.4	区域结构设计	115
3.4.1	OSPF 区域结构及防环设计	116
3.4.2	LSA3 及区域间路由通告	116

3.4.3	ABR 定义及其路由通告	119
3.4.4	Area 类型及特殊类型区域	120
3.4.5	区域分割	122
3.4.6	Vlink 原理	123
3.5	LSA4 和 LSA5	129
3.5.1	LSA4	129
3.5.2	LSA5	130
3.5.3	Forwarding-Address 作用	131
3.5.4	场景分析 1: 外部路由的访问路径	136
3.5.5	场景分析 2: 外部路由计算	137
3.6	NSSA 及 LSA7	138
3.6.1	NSSA 区域	138
3.6.2	NSSA ABR=ASBR	139
3.6.3	LSA7 翻译	140
3.6.4	场景分析 1: 负载分担及次优路径	141
3.6.5	场景分析 2: OSPF 外部路由所引起的环路	142
3.7	选路规则及路由控制	144
3.7.1	OSPF 选路规则	144
3.7.2	OSPF 矢量特性	145
3.7.3	OSPF 路由控制	146
3.7.4	路由聚合	156
3.7.5	默认路由	159
3.8	OSPFv3	161
3.8.1	OSPFv3 与 OSPFv2 的相同点	162
3.8.2	OSPFv3 与 OSPFv2 的不同点	162
3.8.3	OSPFv3 中 LSA 的定义	165
3.8.4	OSPFv3 中新 LSA 及其对收敛的影响	166
3.8.5	OSPFv3 配置实例	168
3.9	思考题	176
第四章	IS-IS	178
4.1	IS-IS 基础	179
4.1.1	历史及 OSI 模型	179
4.1.2	IS-IS 地址	179

4.1.3 IS-IS 报文结构	180
4.1.4 IS-IS 区域及路由器角色	182
4.1.5 IS-IS 网络类型	185
4.2 IS-IS 邻接关系	185
4.2.1 握手报文	185
4.2.2 邻接关系的建立	187
4.3 IS-IS 链路状态数据库	197
4.3.1 概述	197
4.3.2 IS-IS 链路状态数据库同步使用的报文	198
4.3.3 泛洪机制	208
4.3.4 链路状态数据库同步过程	210
4.4 LSP 分片扩展 (Fragmentation)	212
4.4.1 原理概述	212
4.4.2 基本概念	212
4.4.3 LSP 分片方式	213
4.5 IS-IS 的收敛和扩展特性	214
4.5.1 增量 SPF 算法 (I-SPF)	214
4.5.2 部分路由计算 (PRC)	215
4.5.3 智能定时器	216
4.5.4 LSP 的快速扩散	217
4.5.5 IS-IS 认证	217
4.5.6 IS-IS 度量及扩展	218
4.6 IS-IS GR	220
4.6.1 GR 技术概述	220
4.6.2 基本术语	220
4.6.3 IS-IS GR TLV 与计时器	220
4.6.4 IS-IS GR 的运行机制	222
4.7 IS-IS 网络案例	222
4.7.1 案例 1: 路由泄露	222
4.7.2 案例 2: IS-IS 路由聚合	225
4.7.3 案例 3: 路由通告和过滤	226
4.8 IS-IS for IPv6	230
4.8.1 概述	230

4.8.2	IS-ISv6 使用的 TLV	230
4.8.3	IS-IS 多拓扑	232
4.8.4	IS-IS 在 IPv6 网络中的配置案例	234
4.9	OSPF 与 IS-IS 的对比	238
4.10	思考题	239
第五章	路由控制	241
5.1	ACL 和 ip-prefix	242
5.1.1	ACL	242
5.1.2	前缀列表 ip-prefix	250
5.2	路由策略与策略路由	253
5.2.1	路由策略与策略路由介绍	253
5.2.2	路由策略工具 1: filter-policy	253
5.2.3	路由策略工具 2: route-policy	259
5.2.4	策略路由 PBR	265
5.3	聚合路由和默认路由	268
5.3.1	聚合路由	268
5.3.2	默认路由	271
5.4	路由引入	273
5.4.1	路由引入	273
5.4.2	路由引入的问题	282
5.5	综合案例: 多协议之间的路由引入	286
5.6	思考题	289
第六章	BGP	290
6.1	BGP 基础	291
6.1.1	BGP 概述	291
6.1.2	路径矢量路由协议	291
6.1.3	BGP 与 IGP 协议对比	292
6.1.4	BGP 接入	293
6.1.5	BGP 邻居建立	294
6.1.6	BGP 报文结构	299
6.2	BGP 路径属性	304

6.2.1 属性分类	304
6.2.2 属性的用法	306
6.2.3 BGP 工作原理	314
6.2.4 BGP 路由及默认路由	323
6.2.5 案例分析	325
6.3 BGP 的路由控制	330
6.3.1 控制策略	330
6.3.2 案例分析	337
6.4 BGP 的应用与优化	348
6.4.1 大型的 BGP	348
6.4.2 路由反射器	348
6.4.3 BGP 联盟	359
6.4.4 BGP 对等体组	372
6.4.5 路由衰减	374
6.4.6 BGP 设计场景	375
6.5 BGP 的特性	385
6.5.1 BGP 的安全性	385
6.5.2 BGP 的可靠性	386
6.5.3 MP-BGP	391
6.6 思考题	395
第七章 Multicast	396
7.1 组播基础	397
7.1.1 组播应用	397
7.1.2 组播工作原理	397
7.1.3 组播模型	401
7.1.4 组播地址	401
7.1.5 组播转发过程	403
7.2 IGMP	406
7.2.1 IGMP version 1	407
7.2.2 IGMP version 2	409
7.2.3 IGMP version 3	413
7.2.4 IGMP Proxy	421
7.2.5 IGMP 路由表	423

7.3	IGMP Snooping	426
7.3.1	IGMP Snooping 原理描述	426
7.3.2	IGMP Snooping Proxy	430
7.4	PIM 组播路由协议基础	433
7.4.1	协议概述	433
7.4.2	邻居发现及协商	433
7.4.3	PIM 邻居协商	433
7.4.4	PIM 报文种类及格式	436
7.5	PIM DenseMode	437
7.5.1	扩散 (Flooding) 及剪枝机制	437
7.5.2	嫁接机制	443
7.5.3	状态刷新机制	444
7.5.4	剪枝否决机制及剪枝延迟 (PrunePending) 计时器	448
7.5.5	断言 (Assert) 机制	449
7.6	PIM Sparse Mode	456
7.6.1	SPT 和 RPT 树对比	457
7.6.2	RP 控制	457
7.6.3	动态 RP 部署	460
7.6.4	PIM SM 建树过程	466
7.6.5	场景示例 1: 共享网段上掉头 (turnaround) 机制	482
7.6.6	场景示例 2: 分析组播转发过程	486
7.7	PIM SSM	490
7.7.1	PIM SSM 简介	490
7.7.2	SSM 的工作机制	491
7.7.3	IGMP SSM Mapping	493
7.8	MSDP 及 MBGP	495
7.8.1	MSDP 原理描述	495
7.8.2	MSDP 路由器的类型及部署结构	496
7.8.3	MSDP 工作机制	497
7.8.4	AnycastRP 应用	500
7.8.5	SA 过滤控制	505
7.8.6	MBGP 简介	506
7.8.7	MBGP 及 MSDP 案例	508
7.9	思考题	513

第八章 MPLS 及 MPLS VPN	515
8.1 MPLS 基础	516
8.1.1 什么是 MPLS	516
8.1.2 MPLS 的产生背景	516
8.1.3 MPLS 网络结构及术语解释	516
8.1.4 MPLS 标签格式	518
8.1.5 LSP 的建立	519
8.2 LDP 协议	522
8.2.1 LDP 的基本概念	522
8.2.2 LDP 的报文	524
8.2.3 LDP 标签的发布和管理	527
8.2.4 LDP 会话的建立过程	532
8.2.5 LDP 的环路检测	535
8.2.6 LDP 和 IGP 的同步	536
8.2.7 LDP 案例分析	538
8.3 MPLS VPN	546
8.3.1 VPN 基础	546
8.3.2 MPLS VPN 的网络结构	548
8.3.3 MPLS VPN 基本概念	548
8.3.4 路由发布过程	553
8.3.5 数据转发过程	561
8.3.6 常见问题	567
8.3.7 PE-CE 之间的路由协议	567
8.3.8 MCE	584
8.3.9 Internet 接入	588
8.3.10 跨域的 MPLS VPN 解决方案介绍	592
8.4 思考题	609
第九章 交换技术	610
9.1 VLAN	611
9.1.1 VLAN 基础	611
9.1.2 VLAN 原理	612
9.1.3 划分 VLAN 的方式	614

9.1.4	VLAN 扩展应用	621
9.2	STP	648
9.2.1	STP 概述	648
9.2.2	STP 基本概念	649
9.2.3	STP 报文类型	655
9.2.4	拓扑收敛计算	657
9.2.5	拓扑变化通知	659
9.2.6	STP 设计建议及不足	660
9.3	RSTP	661
9.3.1	RSTP 概述	661
9.3.2	RSTP 相比于 STP	661
9.3.3	快速收敛机制	664
9.3.4	拓扑变化通知机制	669
9.3.5	扩展性及优化 RSTP 与 STP 兼容	671
9.4	MSTP	671
9.4.1	MSTP 和 STP/RSTP 的比较	671
9.4.2	MST 原理	672
9.4.3	MST 拓扑计算	683
9.5	STP 其他保护技术	692
9.5.1	BPDU 保护 (BPDU protection)	692
9.5.2	根保护 (Root Protection)	693
9.5.3	环路保护 (Loop Protection)	693
9.5.4	TC 防护 (TC Protection)	694
9.5.5	BPDU 过滤 (BPDU filter)	695
9.6	PPP	695
9.6.1	PPP 简介	695
9.6.2	PPP 认证	698
9.6.3	MP 简介	700
9.7	PPPoE 简介	701
9.7.1	PPPoE 工作原理	702
9.7.2	PPPoE 验证过程	703
9.7.3	PPPoE 报文	704
9.8	思考题	706