

Advanced IP Routing in Cisco Networks
Second Edition

✓ Cisco

网络高级 IP 路由技术

(第2版)

(美) Terry Slattery Bill Burton 著
达达翻译组 译



机械工业出版社
China Machine Press



Education

Cisco 专业技术丛书

Cisco 网络高级 IP 路由技术

(第2版)

(美) Terry Slattery 著

Bill Burton

达达翻译组 译



机械工业出版社
China Machine Press

本书详细介绍TCP/IP的工作原理以及如何在Cisco路由器上配置TCP/IP。涉及基本配置、静态路由、有类路由、无类路由、内部路由和外部路由以及各种Internet连接选项等方面内容。每章都有Cisco路由器的配置实例，并配有说明。本书既可以作为基于IP的网络日常操作的实用参考书，也可以作为深入理解Cisco路由器工作原理的技术资料，还适合于从CCNA到CCIE的各种考试的应试者。

Terry Slattery & Bill Burton: Advanced IP Routing in Cisco Networks, Second Edition (ISBN 0-07-212591-8).

Original edition copyright © 2000 by McGraw-Hill. All rights reserved.

Chinese edition copyright © 2001 by China Machine Press. All rights reserved.

本书中文简体字版由美国麦格劳-希尔公司授权机械工业出版社独家出版。未经出版者书面许可，不得以任何方式复制或抄袭本书内容。

版权所有，侵权必究。

本书版权登记号：图字：01-2001-0487

图书在版编目(CIP)数据

Cisco网络高级IP路由技术：第二版/（美）斯莱特里（Slattery, T.），（美）博极（Burton, B.）著；达达翻译组译. -北京：机械工业出版社，2001.7

（Cisco专业技术丛书）

书名原文：Advanced IP Routing in Cisco Networks.

ISBN 7-111-07201-4

I. C… II. ①斯… ②博… ③达… III. 计算机网络—路由选择 IV. TP393

中国版本图书馆CIP数据核字（2001）第26654号

机械工业出版社（北京市西城区百万庄大街22号 邮政编码 100037）

责任编辑：张鸿斌 胡郁林

北京牛山世兴印刷厂印刷·新华书店北京发行所发行

2001年7月第2版第1次印刷

787mm × 1092mm 1/16 · 32.75印张

印数：0 001 ~ 5 000册

定价：55.00元

凡购本书，如有倒页、脱页、缺页，由本社发行部调换

序

Cisco路由器越来越广泛地应用于家庭、中小型企业、大型企业和ISP骨干网，其功能也日趋完善。Cisco的命令非常多。每个Cisco路由器都需要经过一定的配置，方能正常工作。十分有趣的是，可以有多种命令组合来完成同样的工作。基于彻底理解Cisco IOS而产生的配置将使你能更好地完成任务，并使网络取得更高的性能和可靠性。

管理和配置Cisco路由器的人员都希望能够得到一些适合于他们的特定环境的“正确的”配置范例。Cisco的配置指南及培训教程就提供了此类信息。然而，这些资料可能无法满足路由器管理员寻求深入信息以及相应解释的要求。此时人们通常就开始询问新闻组、友好的路由器向导和Cisco指导者，以求得到正确答案。

如果你曾经有此境遇，想要深入地理解Cisco IOS，那么本书最适合于你。本书引导读者从基本配置、静态路由、有类路由、无类路由、内部路由和外部路由到各种Internet连接选项等方面，了解整个IP协议及路由。每章都有Cisco路由器的配置实例，并配有说明。本书既可以作为参考书，也可以逐章阅读以深入理解Cisco路由器是如何工作的。本书还适合于从CCNA到CCIE的各种考试的应试者。

Cisco系统世界培训中心、发行服务部高级经理

Anthony Wolfenden

前言

感谢你选择了本书。我们在这本书里融入了多年的实践经验以及培训知识，希望它有助于你们学习如何在网络中应用Cisco路由器。这本书的独特之处在于，它是一本含有大量网络实例与相应的Cisco配置的书籍。我们讨论配置中的相关部分以便让你明白所发生的事情。列出完整配置的目的是让你掌握实例的完整配置。

本书的目的之一是为你提供一个指南，帮助你详细了解TCP/IP如何操作以及如何在Cisco路由器上配置TCP/IP。通过阅读以及理解本书中的概念和实例，你能够有更加充分的准备来承担运行大型复杂的IP网络的任务，或挑战Cisco认证互联网专家(CCIE)考试。(有关CCIE考试的进一步消息，请查看Cisco的网页：www.cisco.com。)

我们还想让这本书成为帮助你进行基于IP的网络的日常操作的实用参考指南。TCP/IP成为首选协议的主要原因是Internet及WWW的爆炸性增长。但是有多种TCP/IP路由选择协议，本书有助于你为实际应用选择一个最好的路由选择协议并能够进行实际配置。

无论你是一个专家或者是一个小公司的网络管理员，我们相信本书对你的日常工作将是有益的。你应该将它放在手边。我们欢迎你如何将改进本书的下一版的意见反馈到：www.intelligent.com。

内容

当撰写像本书这样的专业书籍时，我们必须假设读者应该具备一定的知识水准。我们并不想写一本基本的Cisco路由器配置指南，因为Chris Lewis已经在他的专著《Cisco TCP/IP路由管理专业参考》(Lewis 1998，已由机械工业出版社引进出版)中做了这个工作。实际问题是我們是否应该讲解IP是如何工作的。经过讨论之后，我们依据在培训班上收到的问题做出决策：用一章介绍关于TCP/IP如何工作的基本知识，将使我们的书更加系统并有更好的叙述基础。参考其他资源将有助于弥补读者所缺乏的TCP/IP知识。如果你对TCP/IP的完整细节感兴趣，我们推荐《TCP/IP Illustrated, Volumn 1》(Stevens著，1994)。

本书不是一本有关TCP/IP应用程序的书籍。如果你想知道有关TFTP、SMTP、SNMP以及使用TCP/IP的其他协议的更多知识，我们向你推荐Paul Simoneau所著的《Hands-on TCP/IP》(1997)。该书也是一本有关TCP/IP操作的好参考书。

如果你有很坚实的TCP/IP背景知识，可以跳过第1章“TCP/IP概述”。尽管这是基本知识，但我们发现有许多人不了解TCP/IP协议如何操作，所以包含了此章。在该章中，我们描述了TCP/IP操作以及端口号如何用于数据流的多路复用及多路分解。通过采用与邮政系统的类比，使读者了解TCP/IP在实际世界中是如何工作的。

第2章描述了路由器如何决定报文到达目的地的最佳路径，还描述了如何决定网络掩码及其用法。最后我们详细解释了对许多网管员都具有挑战性的概念，即可变长子网掩码。

当你理解了路由器如何寻找路由之后，第3章向你介绍静态路由。静态路由是手工输入路由器的路由，有些地方要用到这样的路由，我们会说明哪些地方应该使用静态路由，也强调了有些地方不应该用静态路由。

在原先的Internet中，所有路由选择都是基于IP地址类的(A、B、C或D类)。在第4章中，将介绍老的IP路由选择协议，即RIP及IGRP，它们都用于有类路由网络。

然后，在第5章中，我们阐述无类编址及路由选择，这是目前Internet所基于的路由选择机制。通过使用无类路由选择，我们可以更充分地利用已有的IP地址空间（以往的地址浪费是惊人的）。IPv6也使用无类路由选择，所以理解无类路由也有助于向下一代IP过渡。

以无类路由选择为背景，你可以学到几种无类路由选择协议，这些协议通过搜索路由选择表查找目的地地址的最佳匹配路由，以选择最佳路径。这些协议包括RIPv2、OSPF及Cisco专用的EIGRP。

另一种无类路由选择协议受到主要的大型ISP的欢迎，即IS-IS协议。它与OSPF类似，但也有不少差别，并且应用得较广，所以我们认为有必要用一章的篇幅（第6章）来介绍它。

许多公司采用一种以上的协议，为此需要以某种方式实现不同路由选择协议间的路由信息交换。这是非常有趣并且容易出现问题的地方。在第7章，你会看到如何集成多种协议，使得它们能够协调工作，并避免路由循环。

回顾一下Internet在过去18年内的发展将是十分有趣的。那时许多公司会说“我们决不打算与Internet连接，那太危险了”。诞生了万维网(WWW)后，加入WWW就会有商业机会。在接下来的几章里，我们将讨论支持Internet连接所需的各种技术。

公司在加入Internet前，通常随便选择一些IP地址供内部使用。一旦决定连入Internet，这家公司面临的问题是：要么重新对成千上万的设备编址，要么在连入Internet处安装网络地址转换(NAT)网关。在第8章中，你将学习如何利用Cisco路由器完成这种功能（你会需要这种技术）。你还可以学到私用地址的知识，私用地址是保留给公司内部使用的，用于分离内部地址与Internet地址，有助于提高安全性、提供服务提供商之间的可移植性等；使用私用地址也可能是因为公司不打算与Internet连接。

有多种连入Internet的方法，最简单的方法并不运行路由选择协议，而采用静态路由简化连接，并且避免在通往ISP的链路上运行路由选择协议的开销。第9章讨论连接到一个或多个服务提供商时，如何使用静态路由。

有些公司具有复杂的网络及Internet连接，他们需要采用外部协议。在第10章中，你将学习边界网关协议，这是服务提供商及大型公司所采用的标准协议。

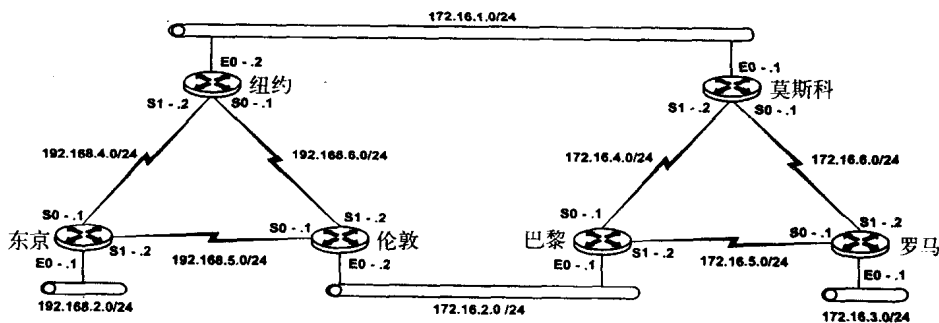
在本书的第1版印刷后，我们发现许多组织在配置即时拨号路由选择方面遇到了问题。第11章专门讨论即时拨号路由选择，这一章有助于配置路由器间的间歇的拨号链路或拨号备份机制。

最后，在第12章中，我们将讨论有关故障排除问题。你将看到如何用TTCP、ping、trace等命令来高效地排除TCP/IP问题。先前的每一章都包含一些与所讨论专题相关的故障排除提示，但有些时候需要额外的信息来找出问题。在解决复杂的网络问题时，正确地运用这些工具将会帮助你快速决定必须详细地检查什么以及可以忽略什么。

动手实验

配置Cisco路由器主要是一个动手的过程。为了真正学到东西，必须不断练习并试验各种不同的实现方法。如果你正在准备CCIE，做实验和练习是十分必要的。我们发现，许多网络技术是清晰明了的，但将练习转换为可以实际工作的配置比预计的要费时得多。

本书的每章都含有若干个动手实验练习。实验中使用三个路由器组成的小组。两组路由器用于需要较复杂的拓扑的实验。图I-1显示两组路由器的“工作组”，每个串行连接采用“背对背”电缆，一端是DCE(数据通信装置)设备，另一端是DTE(数据终端装置)设备。DCE通常是串行通信链路的时钟源。每个路由器的串口1是DCE连接，需要增加clock rate配置命令。这是我们的实例中很重要的部分。



图I-1 两组路由器的工作组

你可以建立自己的路由器环境，也可以访问MentorLabs的虚拟实验室（vLabSM）的一个或多个路由器环境，只是需要支付少许费用。从Internet上，你可以访问多个可供进行实验的路由器的控制台。对你可以使用的路由器配置命令没有限制，并且你可以将你的配置存储到一个配置存储区域中供以后使用。详细信息请访问MentorLabs的站点www.mentorlabs.com。

我们用一个城市名称给每个路由器命名，以便在文字中引用。当发出涉及该路由器的命令时，我们将用IP地址而不是主机名。大部分Cisco配置命令都使用IP地址，主机名只是用于ping及trace等场合，以便于阅读。我们发现错误的主机名表往往是网络问题的起因。如果你喜欢使用主机名，完全可以把必要的主机名表加入到路由器配置中，只是小心地址变化可能会造成主机名表无效。

书中路由器组所用的IP地址并不是一样的，这是我们故意设计的。网络问题很少会每次都出现在同一地方。你应该习惯在处理不同的网络问题时看到不同的IP地址。这种练习对实际工作及CCIE考试均有好处，在这些场合，你必须能够很快地识别网络拓扑及编址。

当我们使用地址时，我们常常会使用私有地址范围内的地址，这在RFC1918(Rekhter等，1996)中有定义：

```
10.0.0.0-10.255.255.255 (10/8 前缀)
172.16.0.0-172.31.255.255 (172.16/12 前缀)
192.168.0.0-192.168.255.255 (192.168/16 前缀)
```

这些地址决不应该出现在Internet上，它们被保留供私用。通过使用这些地址(即使在应该使

用分配的地址的地方), 可以防止偶而引发其他路由器问题。

示例

本书含有大量配置实例。我们曾见过由于不完整的实例或遗漏重要的项目而导致严重问题。通常情况下, 产生的问题比能解决的要多得多。为了解决这个问题, 我们的详细配置以及输出实例中均有行号, 重要的行还有详细的解释 (见程序清单I-1)。

程序清单I-1 原型示例

Prototype example

```
1 NewYork#sh cont s 1
2 HD unit 1, idb = 0xDAA44, driver structure at 0xDEBE0
3 buffer size 1524 HD unit 1, V.35 DCE cable
4 cpb = 0x63, eda = 0x2940, cda = 0x2800
5 RX ring with 16 entries at 0x632800
```

第3行显示了接到串口1的电缆提供DCE接口, 用clock rate命令定义连接到这个串口的DTE设备的时钟频率。

本书例子所用的路由器全部运行IOS 11.3, 它的文档(Cisco 1998b)在www.cisco.com上, 大多数实例和实验也可以在旧版本上执行。如果你的路由器运行的是其他版本, 那么可能会得到不同的输出, 在告之我们你所发现的问题之前, 请先检查一下版本差异。

路由器配置含有许多重复信息(如主机表、vty配置、口令等), 有些例子中可能含有这些信息, 以便使你有整体概念, 而有些实例中去掉了无关信息。

具体配置实例中会含有你所需的所有详细信息, 以便你可以将我们的实现方案复制到你的环境中。我们鼓励通过完成我们的实例以及相关实验来进行自己的实验。没有显示的接口都是不重要的, 可以关闭掉。

CCIE准备

Cisco认证网络互连专家(CCIE)的认证过程是网络专业人员遇到的专业性最强的挑战之一。在每一章里, 你都会发现有一节描述为了成功地通过考试所应该了解的内容。如果你花些时间完成所有例子及实验, 你会更有把握通过笔试及动手考试。

请注意本书的CCIE部分不是直接来源于CCIE考试, Cisco也没有认可这些章节的权威性。我们提供的材料覆盖了基本知识, 这是所有参加CCIE考试的人都应该知道的、应该动手完成的。毫无疑问, 有些考试内容会没有这些内容, 但是只要动手做了这些例子及实验, 你就有了更好的准备。

约定

许多图及正文中出现了配置实例。为了便于阅读, 代码按定宽字体显示, 与你在计算机终端上看到的一样。命令用粗体印刷。配置命令及某些命令输出有行号是为了便于阅读, 在实际的Cisco路由器的配置及输出中没有行号。

VIII

组合键用尖括号括起，这意味着你必须同时按下这些键。例如，<CTRL-Z>表示你应该在按住Ctrl的同时，按下Z键；更复杂的组合如<CTRL-SHIFT-6>，这表示必须同时按住Ctrl及Shift键，再按6键。

最后，我们再次感谢你选择了这本书。希望你觉得这本书有用并从中受益，我们欢迎你与我们共同讨论本书的相关问题，并对本书提出宝贵意见。

Terry Slattery, CCIE#1026

William E. Burton, CCIE#1119

目 录

序	
前言	
第1章 TCP/IP 概述	1
1.1 OSI参考模型	1
1.1.1 类比邮政系统	2
1.1.2 邮局中的路由	5
1.1.3 封装与拆封	6
1.1.4 OSI模型的优势	6
1.2 TCP/IP协议栈	7
1.2.1 总体结构	7
1.2.2 协议栈多路复用	8
1.2.3 TCP	11
1.2.4 UDP	16
1.2.5 IP	18
1.3 其他协议	22
1.3.1 地址解析协议	22
1.3.2 网际控制消息协议	25
1.4 性能	30
1.5 CCIE提示	35
1.5.1 故障排除	35
1.5.2 开环路由	38
1.5.3 端口号	38
1.6 IP的未来	39
1.7 小结	39
第2章 路由器工作原理	41
2.1 路由器	41
2.2 寻找IP目的地址	42
2.3 网络掩码	43
2.4 路由选择过程	45
2.5 路由选择表	46
2.6 管理距离	47
2.7 路由选择表查找	49
2.8 在Cisco路由器间交换路径	49
2.9 多路径操作	51
2.10 子网化	53
2.11 子网化举例	54
2.12 使用什么样的掩码	57
2.13 可变长子网化	58
2.14 无类路由选择	61
2.14.1 基于VLSM的示例网络	62
2.14.2 VLSM路由选择协议	64
2.14.3 路由查找	67
2.15 CCIE提示	68
2.16 小结	69
第3章 静态路由	70
3.1 桩网络	70
3.2 缺省路由	72
3.2.1 何时使用缺省路由	72
3.2.2 如何配置缺省路由	73
3.3 浮动静态路由	74
3.3.1 实验练习	77
3.3.2 触发备份链路	77
3.4 缺省网络	81
3.5 多个缺省路由	86
3.6 调试	89
3.7 CCIE提示	90
3.8 小结	90
第4章 有类内部路由	92
4.1 距离-向量协议	92
4.2 RIP协议	92
4.2.1 配置	93
4.2.2 RIP协议的更新	95

4.2.3 RIP协议的度量值	96	5.4.1 度量值	135
4.2.4 RIP协议的计数到无穷大	96	5.4.2 基本配置	135
4.2.5 距离-向量协议的水平分割	98	5.4.3 操作	138
4.2.6 毒性逆转	99	5.4.4 EIGRP协议小结	145
4.2.7 瞬间更新	99	5.5 OSPF协议	145
4.2.8 阻止机制	100	5.5.1 操作	146
4.2.9 定时器	100	5.5.2 基本配置	146
4.2.10 报文格式	102	5.5.3 度量值	149
4.2.11 有限的RIP协议更新	103	5.5.4 OSPF协议区域	151
4.2.12 RIP协议的网络设计	105	5.5.5 虚链路、路由器标识和环回接口	156
4.2.13 调试和显示命令	108	5.5.6 外部路由	162
4.2.14 RIP协议小结	110	5.5.7 再分布缺省路由	165
4.3 IGRP协议	110	5.5.8 桩和完全桩区域	167
4.3.1 配置	110	5.5.9 非桩区域	171
4.3.2 度量值	114	5.5.10 非广播多访问网络	176
4.3.3 IGRP协议的多路径操作	116	5.5.11 即时链路	181
4.3.4 更新、无效、阻止、刷新定时器	116	5.5.12 hello报文和相邻路由器的邻接 关系	182
4.3.5 水平分割、毒性逆转和瞬间更新	117	5.5.13 路由过滤	184
4.3.6 报文格式	117	5.5.14 OSPF协议的故障排除	184
4.3.7 传送	118	5.5.15 OSPF小结	185
4.3.8 IGRP协议的网络设计	118	5.6 CCIE提示	185
4.3.9 调试和显示命令	119	5.7 小结	186
4.3.10 IGRP协议小结	121	第6章 用IS-IS进行IP路由选择	187
4.4 有类网络中的路由选择	121	6.1 IS-IS区域	187
4.5 路由聚合	122	6.1.1 网络实体标题和区域ID	191
4.6 不连续的子网	123	6.1.2 区域分割	191
4.7 CCIE提示	125	6.1.3 区域限制	192
4.8 小结	126	6.2 路径选择	192
第5章 无类路由	127	6.3 度量值	195
5.1 路由器的工作过程	127	6.4 协议概述	196
5.2 无类路由和路径前缀	128	6.4.1 Hello报文	196
5.3 RIP协议版本2	129	6.4.2 链路状态和序号PDU	197
5.3.1 与RIPv1协议的不同之处	129	6.5 故障排除	197
5.3.2 报文格式	129	6.6 CCIE提示	198
5.3.3 与RIPv1协议的对比	130	6.7 小结	198
5.3.4 RIP协议小结	134		
5.4 EIGRP协议	134		

第7章 集成多个路由选择协议	199	9.1 静态路由为Internet互连提供哪些功能	361
7.1 路由汇总	199	9.2 静态路由不能为Internet互连提供哪些功能	361
7.1.1 RIP协议	199	9.3 CCIE提示	382
7.1.2 RIP协议版本2	200	9.4 小结	382
7.1.3 IGRP协议	200	9.5 问题与答案	383
7.1.4 OSPF协议	200	第10章 边界网关协议Internet连接	384
7.1.5 EIGRP协议	200	10.1 BGP概况	384
7.2 路由映射	201	10.1.1 自治系统定义	384
7.3 路由环	201	10.1.2 BGP背景	384
7.4 路由再分布	202	10.1.3 BGP操作	385
7.5 ip classless命令	202	10.1.4 BGP路由的确定	385
7.5.1 路由再分布使用的命令	202	10.1.5 连接到ISP	386
7.5.2 内部协议间的再分布	202	10.2 示例	386
7.6 CCIE提示	288	10.2.1 例1: 单出口BGP Internet连接	386
7.7 小结	288	10.2.2 BGP对等体连通性的调试	393
7.8 问题与解答	288	10.2.3 例2: 多出口BGP Internet连接	396
7.9 练习解答	289	10.2.4 例3: 更多的BGP特性	429
第8章 网络地址转换	312	10.3 CCIE提示	462
8.1 NAT能提供的功能	312	10.4 小结	462
8.2 NAT不能提供的功能	312	第11章 即时拨号路由选择	463
8.3 术语	312	11.1 DDR与路由选择协议交互	463
8.3.1 内部局部地址	313	11.1.1 例1: 距离-向量协议的快速路由选择	463
8.3.2 内部全局地址	313	11.1.2 例2: OSPF拨号连接	472
8.3.3 外部局部地址	313	11.1.3 例3: 用EIGRP进行即时路由选择	481
8.3.4 外部全局地址	313	11.2 CCIE提示	487
8.3.5 简单转换	313	11.3 小结	487
8.3.6 地址超载	313	第12章 故障排除	488
8.4 特定情况	314	12.1 方法	488
8.4.1 专用地址	314	12.2 丢失路由	489
8.4.2 更换ISP	314	12.3 开环路由	490
8.4.3 TCP负载共享	314	12.4 黑洞路由	491
8.5 配置地址转换	314	12.5 拥塞	492
8.6 CCIE提示	359	12.6 配置Cisco路由器	493
8.7 实验	359	12.7 ping命令	493
8.8 小结	360		
第9章 静态Internet互连	361		

12.8 trace命令	494	12.10.2 show ip route	499
12.9 TTCP	495	12.10.3 debug	499
12.9.1 Cisco路由器上的TTCP.....	495	12.11 CCIE提示	500
12.9.2 TTCP分析	498	12.12 小结	500
12.9.3 TTCP小结	498	附录A 10进制、16进制和2进制换算	501
12.10 路由器命令	498	附录B Mentor技术子网计算器.....	505
12.10.1 show interfaces	499	附录C 缩写词表	507

第1章 TCP/IP概述

TCP/IP是传输控制协议/网际协议（Transmission Control Protocol/Internet Protocol）的缩写，这两个协议是目前Internet上应用最广泛的协议。TCP/IP协议簇实际上是一个协议集，TCP和IP是最著名的两个协议，此协议簇中的其他协议包括用户数据报协议(UDP)、网际控制消息协议(ICMP)以及地址解析协议(ARP)。整个协议簇称为TCP/IP。

如果你认为已经掌握了全部TCP/IP，你是否能够解释发生在协议栈各层的多路复用？你是否了解大多数TCP实现所用的最大窗口长度，以及这如何影响长延迟路径(或者是很高带宽的管道)的吞吐量？

在本章中，我们将回顾TCP/IP协议簇中的各个协议的重点，如果你希望更多地了解TCP/IP协议本身或是使用TCP/IP的应用程序的操作，我们推荐参考Stallings(1997)、Stevens(1994)、Comer(1995)或Simoneau(1997)。

1.1 OSI参考模型

每本网络书几乎都要从OSI模型的描述开始（请参考图1-1）。这不是因为每个人都喜欢谈论它，而是通过描述此模型可以使得我们在讨论中有共同的术语。擅长故障排除的人往往也用它来查找网络问题。例如，确认数据链路层是否正常，从而将精力放在寻找更高层上的问题。

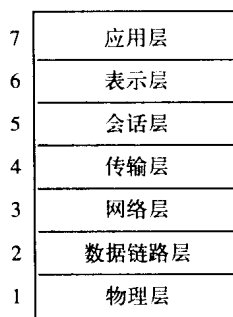


图1-1 OSI模型

正确的网络模型对学习技术也是十分重要的，它给出一个概念的框架。使你知道什么概念应该在什么地方，就好像知道轮胎及传送带在汽车设计与运行中的作用一样。

但是OSI模型只是一个模型。它不是所有东西都要遵循的药方。作为一个模型，它给我们一个框架，我们可以在其中构架、讨论和对比各种网络协议。

OSI模型由国际标准化组织(ISO)制定。OSI代表开放式系统互联。它原先是一个模型，后来演变成一个与之同名的协议簇。OSI协议簇本身没有在世界范围内广泛应用，但它的发展极大地促进了我们对网络协议的理解，并有助于了解如何更好地开发它们。

开发这个模型和协议簇的主要原因是制订国际标准，以便每个人都能够用它互操作并进行网络设备间的通信。实际上，由于OSI协议的开发时间太长，以致于许多公司采用了TCP/IP。随着越来越多的公司采用TCP/IP，对OSI协议的需求也就相应减弱，最后结果是TCP/IP达到了OSI想达到的目标。

1.1.1 类比邮政系统

通过把设计及建造网络协议的复杂任务分解为几层（也称为“协议栈”），我们就能够很容易地完成任务。图1-2说明七层OSI模型与邮政系统的类比。

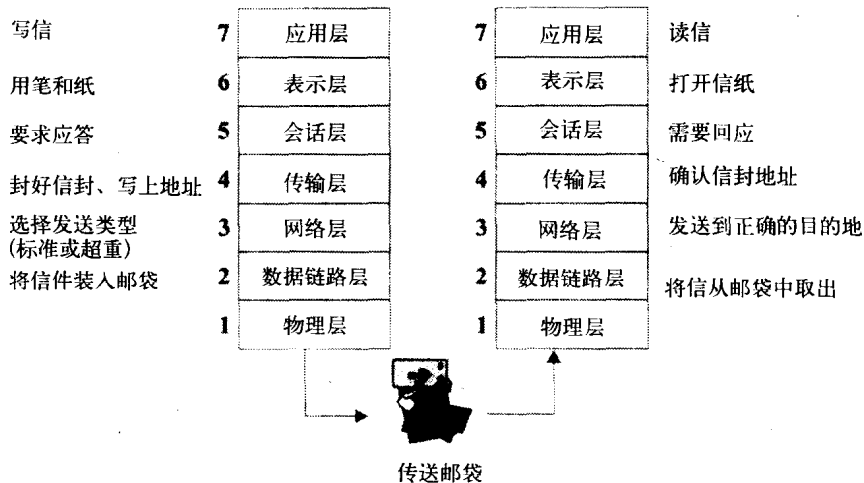


图1-2 邮政类比

下面，我们开始比较OSI模型与邮政系统（这个类比非常合适）。

提示 我们把网络协议的多个层称为“协议栈”，因为一层在另一层上。每层的操作依赖于下层所提供的服务。例如，传输层依赖于网络层的尽力而为的发送，网络层依赖于数据链路层以将数据打包进适合于用于传递数据的物理介质规范的帧（如以太网帧、令牌环网帧、FDDI帧等）。

1. 应用层

应用层(application)是应用程序(如Ftp、Telnet等)与网络协议栈之间的接口，它定义了应用程序为了在网络上传输数据所必须进行的程序调用。

类比邮政系统中，应用可能是与笔友进行交流。接口选择钢笔与纸、铅笔与纸、打印机与纸，或计算机与文字处理程序以生成数据(如图1-3所示)。这一层生成的所有数据被递交给表示层。

2. 表示层

表示层(presentation)处理所有远程通信所需的数据转换。所

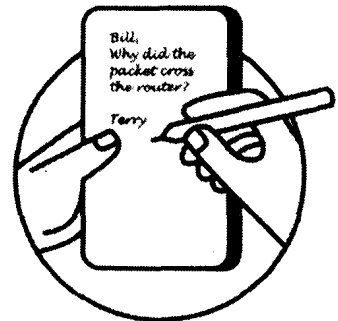


图1-3 应用层——写一封信

有网络协议都执行数据转换,将数据转换为“网络标准格式”。对于Telnet及FTP,它们使用网上定义好的格式,每个系统必须将内部数据表示转换为网络标准格式。采用EBCDIC字符集的计算机必须将它们的数据转为ASCII字符。数据加密及压缩可能在这一层进行。

如果你写信而内容需要保密,可以把要发送的数据转换为一种加密的形式。简单的方法是将字母表轮换13次。A字母变成N, B变成O, C变成P等(如图1-4所示)。这个转换称为Rot-13, 因为字母表中的字母轮换了13个字母。

接下来,转换后的数据被传到会话层。

3. 会话层

会话层(session)维持与远程系统的会话。对于Telnet这样的网络应用,它涉及确定用户终端类型,以便屏幕编辑器正确工作。它也可能需要每个字符的协商,或者是每次一行用户输入数据传送的协商。

与邮政系统相比,这类似于决定采用什么语言、是否修改字母、或是有什么要说就发送什么。会话头部表明这封信是一封完整的信或者在一定的时间内需要有一个回答(如图1-5所示)。然后,数据被交给传输层。

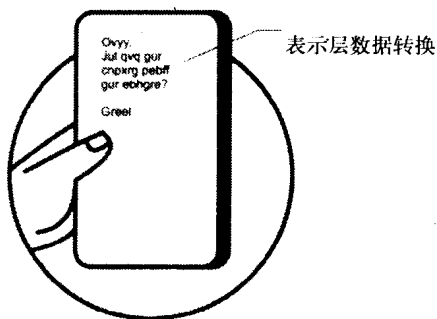


图1-4 表示层——数据转换

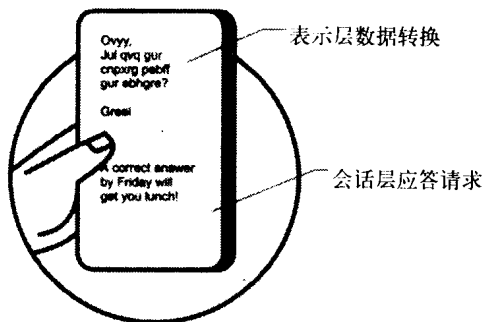


图1-5 会话层——组织会话

4. 传输层

传输层(transport)提供一种方法,使得一个系统上运行的计算机程序能够与另一个系统上运行的程序进行数据通信。大多数协议栈提供两种传输机制。

一种传输机制是可靠发送协议,它重传丢失或受到破坏的数据,使程序不必为此操心;另一种是数据报服务,开销小,但是不保证可靠发送。由程序员决定采用哪一种传输协议(可靠数据流或不可靠数据报)最适合正在开发的程序。

程序员要考虑的最重要因素之一是,通信是面向数据流的(如文件传输)还是面向事务的(如数据库查询)。另一个因素是内在可靠性。对于可靠传输协议,如果发送失败,传输层执行重传功能;对于数据报协议,必须由程序本身重传丢失或受破坏的数据。当在本章后面介绍TCP/IP

和UDP时, 你会学到这两种传输协议的其他差别。

在传输层, 数据被封装在邮寄信封内, 目的地逻辑名用于查找目的地址。类比如邮政系统, 这包括检查笔友的地址以及为此信件添加一个序号。你的笔友将会知道信的顺序, 并能够按照正确的顺序装订这些信件(因为信件可能未按次序投递), 如果出现错误, 将请求重传。我们看到这一层的数据(包括会话控制信息)被封装在某种传输介质之中, 例如信封(如图1-6所示)。

接下来, 传输层将它的数据递交给网络层。

提示 将数据放在容器中的过程称为。可以把信封(或TCP报文)理解为容器, 用它在网络上传送数据。其外面有源地址及目的地址, 以便网络层决定将信件(或报文)转发到正确的目的地的最佳转发方法。

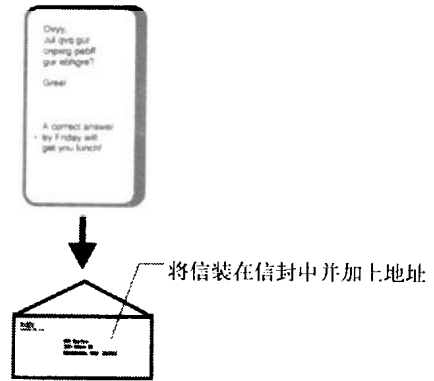


图1-6 传输层——逻辑寻址与可靠性

相反, 从容器中取出东西(数据)的过程被称为拆封。在网络世界中, 我们经常看到封装及拆封发生在传输层、网络层及数据链路层的过程。这些层都以自己的头部描述其自身的内容。

5. 网络层

网络层(network)尝试将数据传递到目的地, 这一层采用“尽力而为”方法进行传递。它检查目的地址并决定最佳路径。知道了最佳路径, 网络层就可以决定使用哪一个数据链路层来传送数据到下一跳。然后, 网络层将数据分为若干段, 段的大小将适合于用于到达下一跳的传输机制。TCP/IP中所用的网络地址与用于连接系统的物理层技术无关。这也是路由器工作的基本条件。

在邮政系统中, 我们可以选择不同的邮递机制, 这类类似于不同的网络技术。我们可以用传真、邮政系统、联邦快递或其他方法传递信件。决定到达目的地的出发路径是路由过程的第一步。在这种情况下, 我们已经写了一封适合于单个信封的信件, 没有投递优先级要求, 所以选择邮局邮递, 如图1-7所示。

然后, 信封被交给数据链路层。

6. 数据链路层

数据链路层(data-link)只关心将信件送往目的地的下一跳地址。在网络世界中, 我们关心介质访问(公平地访问介质, 以保证没有一个站长期独占传输介质)。我们也关心将数据格式化成为适合于介质传输的格式。大多数网络有一套独特的编址方法。在网络层采用一种与数据链路层编址无关的高层地址, 这样可以方便地建立一个基于多种WAN及LAN技术的网络。

类比如邮政系统, 这相当于将信件交给邮递员, 可能是通过邮箱完成。信封必须符合一定的规格(大小、

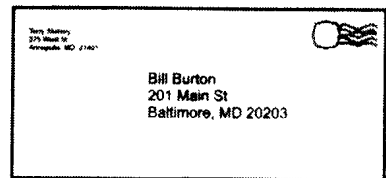


图1-7 网络层——决策路径

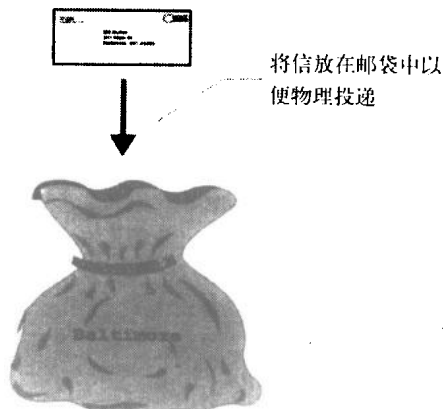


图1-8 物理层——准备通过物理介质发送