



普通高等教育“十一五”国家级规划教材

# 路由器原理与技术

张宏科 苏 伟 编著



高等教育出版社  
HIGHER EDUCATION PRESS

普通高等教育“十一五”国家级规划教材

# 路由器原理与技术

Luyouqi Yuanli yu Jishu

张宏科 苏 伟 编著



高等教育出版社·北京  
HIGHER EDUCATION PRESS BEIJING

## 内容提要

本书是普通高等教育“十一五”国家级规划教材。本书在介绍互联网基本概念的基础上,对路由器工作原理、路由器硬件体系结构、三大路由协议(RIP、OSPF和BGP)、移动路由技术、组播路由技术、网络安全和管理协议、互联网工程设计与应用以及未来互联网技术等进行全面而系统的介绍。

本书取材新颖、内容丰富、实用性强,反映了国内外路由器技术的现状与未来,适合于从事通信、计算机技术开发与研究的广大工程技术人员阅读,也可作为大专院校通信、计算机等专业和相关培训班的教材或教学参考书。

## 图书在版编目(CIP)数据

路由器原理与技术/张宏科,苏伟编著. —北京:高等教育出版社,2010.9

ISBN 978-7-04-030327-8

I. ①路… II. ①张…②苏… III. ①计算机网络-路由选择-高等学校-教材 IV. ①TN915.05

中国版本图书馆CIP数据核字(2010)第141044号

策划编辑 吴陈滨 责任编辑 曲文利 封面设计 于文燕 责任绘图 尹莉  
版式设计 王莹 责任校对 刘莉 责任印制 毛斯璐

出版发行 高等教育出版社  
社 址 北京市西城区德外大街4号  
邮政编码 100120

经 销 蓝色畅想图书发行有限公司  
印 刷 北京嘉实印刷有限公司

开 本 787×1092 1/16  
印 张 20.25  
字 数 490 000

购书热线 010-58581118  
咨询电话 400-810-0598  
网 址 <http://www.hep.edu.cn>  
<http://www.hep.com.cn>  
网上订购 <http://www.landaco.com>  
<http://www.landaco.com.cn>  
畅想教育 <http://www.widedu.com>

版 次 2010年9月第1版  
印 次 2010年9月第1次印刷  
定 价 29.70元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有 侵权必究

物料号 30327-00

# 前 言

当前,人类社会已经进入一个全球信息化的崭新时代,而信息化则离不开信息网络的支持。互联网作为现有信息网络的重要组成部分,在全球信息化的进程中发挥出越来越重要的作用。而路由器是互联网的核心和基础,掌握路由器原理与技术对研究、开发及使用新一代信息网络具有十分重要的意义。

为了推动国内信息网络技术的发展,跟踪世界新一代网络先进技术,满足广大科研人员和工程技术人员的需要,我们在多年学习、研究与工作实践的基础上,先后编写了《路由器原理与技术》(第一、二版)、《IPv6 路由协议栈原理与技术》等书籍。这些书自出版以来,受到了广大读者的好评和支持。为了更好地为读者服务,我们决定结合近几年来科学研究和工作实践成果,并采纳很多读者提出的宝贵建议,编写一部较为全面地介绍互联网关键技术,特别是路由器原理与技术方面的书籍。相对于以前的几本书而言,这次编写除对三大路由协议的工作机理、设计思路及实现方案等作进一步的完善外,还增加了对移动和组播路由技术、网络安全和管理技术的阐述,并且介绍了当前新一代互联网技术领域的最新研究进展和成果等。

本书主要内容包括:第1章主要使读者对互联网有基本了解,叙述互联网的基本概念、组成和发展历史等;第2章主要介绍路由器的工作原理,包括路由选择的基本概念、互联网地址结构、路由转发原理、路由算法以及选路策略等;第3章介绍路由器的硬件设计与实现;第4章主要介绍路由信息协议 RIP、特别是 RIP 在 IPv6 下的版本——RIPng 的基本原理与实现;第5章主要讲述开放最短路径优先协议 OSPF 在 IPv6 下的版本——OSPFv3 的基本原理与实现,重点描述 OSPFv3 的工作过程、设计与实现方案等;第6章系统地描述边界网关协议 BGP4 的基本原理与设计实现等;第7章则重点讲述移动路由技术,包括移动 IPv4、移动 IPv6 以及移动子网的基本原理;第8章主要介绍组播路由技术的基本原理,以及当前主要的组播路由协议和组管理协议等;第9章对互联网上使用的网络安全协议 IPSec 进行系统描述;第10章则简要介绍网络管理协议 SNMP;第11章主要介绍互联网的设计与工程应用;第12章则讲述新一代互联网技术的研究现状和发展趋势等,以期使读者掌握当前互联网领域的最新研究动态。

本书是普通高等教育“十一五”国家级规划教材,在编著的过程中,得到了国家“973”项目“一体化可信网络与普适服务基础研究”(资助号 2007CB307100)、国家自然科学基金重点项目“未来互联网体系理论及关键技术研究”(资助号 60833002)、国家自然科学基金青年科学基金项目“未来互联网的认知服务与感知路由基础研究”(资助号 60903150)等课题的资助。在编著 RIPng、OSPFv3、BGP4 +、移动路由等技术的设计与实现方案时,参考了国家“863”重大项目“高性能 IPv6 路由器协议栈软件”的大量技术资料,这些资料都是项目组老师与同学多年辛勤劳动的结晶;此外,编写“组播路由技术”一章时还参考了李盛和宋晓峰两位同学的硕士学位论文,在此一并向他们表示谢意。另外,对本书所参考的所有文献的作者表示诚挚的感谢。

张宏科教授负责本书的统筹规划,并编写了其中的第1、12章;苏伟老师编写了第2、3、4、5、7、10、11章,郜帅老师编写了第6章,董平老师编写了第8章,刘颖老师编写了第9章,刘琪博士

为本书的文字校阅做了大量工作,并绘制了书中的大量插图。

北京交通大学张思东教授审阅了全书,并提出许多宝贵意见,在此表示衷心的感谢。高等教育出版社对本书的出版给予了大力支持,在此致以诚挚的谢意。

由于信息网络发展极为迅速,路由原理与技术仍在发展和完善之中,加之编写时间有限,书中难免存在不妥之处,敬请广大读者指正。

作者

2010年6月于北京交通大学

## 郑重声明

高等教育出版社依法对本书享有专有出版权。任何未经许可的复制、销售行为均违反《中华人民共和国著作权法》，其行为人将承担相应的民事责任和行政责任，构成犯罪的，将被依法追究刑事责任。为了维护市场秩序，保护读者的合法权益，避免读者误用盗版书造成不良后果，我社将配合行政执法部门和司法机关对违法犯罪的单位和个人给予严厉打击。社会各界人士如发现上述侵权行为，希望及时举报，本社将奖励举报有功人员。

**反盗版举报电话：**(010) 58581897/58581896/58581879

**反盗版举报传真：**(010) 82086060

**E - mail：**dd@hep.com.cn

**通信地址：**北京市西城区德外大街4号

高等教育出版社打击盗版办公室

**邮 编：**100120

**购书请拨打电话：**(010)58581118

# 目 录

|  |    |
|--|----|
| <b>第 1 章 绪论</b> .....                    | 1  |
| 1.1 互联网的基本概念 .....                       | 1  |
| 1.2 互联网体系结构模型 .....                      | 2  |
| 1.3 互联网的发展历史 .....                       | 4  |
| 1.3.1 基于 IPv4 的互联网 .....                 | 4  |
| 1.3.2 基于 IPv6 的互联网 .....                 | 4  |
| <b>第 2 章 路由器工作原理</b> .....               | 10 |
| 2.1 路由选择的基本概念 .....                      | 10 |
| 2.2 互联网地址结构 .....                        | 14 |
| 2.2.1 IPv4 地址结构 .....                    | 14 |
| 2.2.2 IPv6 地址结构 .....                    | 19 |
| 2.3 路由转发原理 .....                         | 25 |
| 2.3.1 路由器的基本工作原理 .....                   | 25 |
| 2.3.2 IP 包转发的物理过程 .....                  | 26 |
| 2.3.3 IPv6 路由转发原理 .....                  | 29 |
| 2.4 路由选择算法 .....                         | 33 |
| 2.4.1 距离矢量路由算法 .....                     | 33 |
| 2.4.2 链路状态路由算法 .....                     | 34 |
| 2.5 路由协议概述 .....                         | 36 |
| <b>第 3 章 路由器的硬件体系结构</b> .....            | 39 |
| 3.1 路由器硬件体系结构的基本组成 .....                 | 39 |
| 3.2 路由器硬件体系结构的类型 .....                   | 40 |
| 3.3 基于网络处理器的路由器硬件设计<br>方案 .....          | 44 |
| 3.3.1 网络处理器简介 .....                      | 44 |
| 3.3.2 IXP 2400 介绍 .....                  | 49 |
| 3.3.3 基于 IXP 2400 的 IPv6 路由器<br>设计 ..... | 54 |
| 3.4 路由查询和流分类简介 .....                     | 57 |
| 3.4.1 路由查询算法简介 .....                     | 57 |
| 3.4.2 流分类算法简介 .....                      | 63 |
| <b>第 4 章 路由信息协议</b> .....                | 65 |
| 4.1 RIP 的发展历程 .....                      | 65 |

|                                 |     |
|---------------------------------|-----|
| 4.2 RIPng 的工作原理 .....           | 66  |
| 4.2.1 RIPng 的基本工作过程 .....       | 66  |
| 4.2.2 RIPng 路由表的建立过程 .....      | 73  |
| 4.2.3 RIPng 路由表的维护过程 .....      | 74  |
| 4.2.4 RIPng 定时器 .....           | 80  |
| 4.2.5 RIPng 的报文格式 .....         | 81  |
| 4.2.6 RIPng 与 RIPv2 的区别 .....   | 84  |
| 4.2.7 RIP 的主要缺陷 .....           | 85  |
| 4.3 RIPng 的设计与实现 .....          | 85  |
| 4.3.1 线程化实现思想 .....             | 85  |
| 4.3.2 RIPng 的总体框架设计 .....       | 86  |
| 4.3.3 RIPng 的启动 .....           | 87  |
| 4.3.4 输入处理 .....                | 87  |
| 4.3.5 输出处理 .....                | 90  |
| 4.3.6 路由操作 .....                | 92  |
| 4.3.7 定时器处理 .....               | 93  |
| 4.3.8 RIPng 的关闭 .....           | 93  |
| <b>第 5 章 开放最短路径优先协议</b> .....   | 95  |
| 5.1 OSPF 概述 .....               | 95  |
| 5.2 OSPFv3 的工作原理 .....          | 96  |
| 5.2.1 区域划分 .....                | 97  |
| 5.2.2 OSPFv3 路由选择分类 .....       | 99  |
| 5.2.3 接口状态机和邻居状态机 .....         | 102 |
| 5.2.4 呼叫协议 .....                | 109 |
| 5.2.5 近邻关系建立过程 .....            | 111 |
| 5.2.6 可靠泛洪过程 .....              | 113 |
| 5.2.7 路由计算 .....                | 114 |
| 5.2.8 OSPFv3 分组结构 .....         | 118 |
| 5.2.9 OSPFv3 与 OSPFv2 的比较 ..... | 129 |
| 5.2.10 OSPFv3 的特点 .....         | 130 |
| 5.3 OSPFv3 的设计与实现 .....         | 130 |
| 5.3.1 OSPFv3 的总体框架设计 .....      | 131 |
| 5.3.2 全局模块和域模块 .....            | 132 |
| 5.3.3 LSDB 模块 .....             | 133 |
| 5.3.4 LSA 模块 .....              | 133 |

|              |                        |            |               |                  |            |
|--------------|------------------------|------------|---------------|------------------|------------|
| 5.3.5        | 消息发送和接收模块              | 138        | 7.2.4         | 移动 IPv4 协议的主要弊端  | 205        |
| 5.3.6        | 数据库交互模块                | 141        | 7.3           | 移动 IPv6 的工作原理    | 206        |
| 5.3.7        | 接口与接口状态机               | 144        | 7.3.1         | 移动 IPv6 技术概述     | 206        |
| 5.3.8        | 邻居与邻居状态机               | 146        | 7.3.2         | 移动 IPv6 的基本概念和术语 | 206        |
| 5.3.9        | 路由处理模块                 | 148        | 7.3.3         | 移动 IPv6 的工作过程    | 208        |
| 5.3.10       | 公共功能模块接口               | 153        | 7.4           | 移动子网的工作原理        | 212        |
| <b>第 6 章</b> | <b>边界网关协议</b>          | <b>154</b> | 7.4.1         | 移动子网技术概述         | 212        |
| 6.1          | BGP 概述                 | 154        | 7.4.2         | 移动子网的结构和术语       | 213        |
| 6.2          | BGP 的工作原理              | 155        | 7.4.3         | 移动子网的基本工作原理      | 215        |
| 6.2.1        | BGP 的基本工作过程            | 155        | 7.4.4         | 移动子网技术面临的问题和挑战   | 216        |
| 6.2.2        | BGP4 分组格式              | 155        | 7.5           | 移动 IPv6 的设计方案    | 216        |
| 6.2.3        | BGP 对等体协商有限状态机         | 160        | 7.6           | 移动互联网的发展趋势       | 218        |
| 6.2.4        | BGP4 路径属性              | 161        | 7.6.1         | 移动互联网面临的问题和挑战    | 218        |
| 6.2.5        | 建立 BGP 对等体对话           | 165        | 7.6.2         | 移动互联网技术的发展趋势     | 219        |
| 6.2.6        | BGP4 选路及决策过程           | 167        | <b>第 8 章</b>  | <b>组播路由技术</b>    | <b>222</b> |
| 6.2.7        | 自治系统联盟                 | 169        | 8.1           | 组播路由技术概述         | 222        |
| 6.2.8        | 路由反射器                  | 169        | 8.2           | 组播路由的基本概念        | 223        |
| 6.2.9        | 路由振荡衰减器                | 171        | 8.2.1         | IP 组播地址          | 223        |
| 6.2.10       | 权能通告                   | 173        | 8.2.2         | 组播分布树            | 226        |
| 6.3          | 面向多协议的 BGP4 + 扩展       | 174        | 8.2.3         | 组播转发原理           | 227        |
| 6.3.1        | BGP4 + 引入的新路径属性        | 175        | 8.3           | 组播路由协议           | 228        |
| 6.3.2        | BGP4 + 工作机制            | 175        | 8.3.1         | 组播路由协议分类         | 228        |
| 6.3.3        | 基于 BGP4 + 传递 IPv6 路由信息 | 176        | 8.3.2         | PIM - SM 工作原理    | 231        |
| 6.4          | BGP4 + 设计与实现           | 176        | 8.4           | 组管理协议            | 239        |
| 6.4.1        | BGP4 + 总体框架设计          | 176        | 8.4.1         | 组管理协议分类          | 239        |
| 6.4.2        | BGP4 + 初始化模块           | 177        | 8.4.2         | MLDv2 的工作原理      | 240        |
| 6.4.3        | 线程管理器模块                | 178        | <b>第 9 章</b>  | <b>网络安全协议</b>    | <b>245</b> |
| 6.4.4        | 有限状态机模块                | 180        | 9.1           | IPSec 协议概述       | 245        |
| 6.4.5        | 消息处理模块                 | 181        | 9.2           | IPSec 的工作原理      | 246        |
| 6.4.6        | 路由信息库模块                | 184        | 9.2.1         | IPSec 的基本框架      | 246        |
| 6.4.7        | 消息封装模块                 | 195        | 9.2.2         | AH 和 ESP         | 247        |
| <b>第 7 章</b> | <b>移动路由技术</b>          | <b>198</b> | 9.2.3         | SADB 和 SPDB      | 249        |
| 7.1          | 移动路由技术概述               | 198        | 9.2.4         | IKE 协议           | 251        |
| 7.2          | 移动 IPv4 的工作原理          | 200        | 9.3           | IPSec 的设计方案      | 252        |
| 7.2.1        | 移动 IPv4 技术概述           | 200        | <b>第 10 章</b> | <b>简单网络管理协议</b>  | <b>257</b> |
| 7.2.2        | 移动 IPv4 的基本概念和术语       | 201        | 10.1          | SNMP 概述          | 257        |
| 7.2.3        | 移动 IPv4 的工作过程          | 203        |               |                  |            |



---

|   |            |  |            |
|---|------------|--|------------|
| 10.2 SNMP 的工作原理 .....                   | 258        | 11.2.4 FTP 分析及在 IPv6 下的<br>实现 .....    | 290        |
| 10.2.1 SNMP 协议基础及 SNMPv1/v2<br>简介 ..... | 258        | 11.2.5 Telnet 分析及在 IPv6 下的<br>实现 ..... | 295        |
| 10.2.2 SNMPv3 的体系结构 .....               | 261        | 11.3 IPv6 网络典型应用 .....                 | 297        |
| 10.2.3 SNMPv3 的消息格式 .....               | 263        | <b>第 12 章 未来互联网技术 .....</b>            | <b>299</b> |
| 10.3 基于 IPv6 的 SNMPv3 设计方案 .....        | 264        | 12.1 国内外研究现状 .....                     | 299        |
| <b>第 11 章 互联网工程设计与应用 .....</b>          | <b>266</b> | 12.2 一体化标识网络体系概述 .....                 | 300        |
| 11.1 互联网工程设计举例 .....                    | 266        | 12.3 基础设施层模型与理论 .....                  | 301        |
| 11.1.1 IP 网络编址设计原则 .....                | 266        | 12.4 普适服务模型与理论 .....                   | 302        |
| 11.1.2 VLSM 设计举例 .....                  | 268        | 12.5 新网络体系下原型系统研制与<br>验证 .....         | 304        |
| 11.2 典型应用及 IPv6 下的实现 .....              | 273        | <b>常用缩略语汇编 .....</b>                   | <b>307</b> |
| 11.2.1 概述 .....                         | 273        | <b>主要参考文献 .....</b>                    | <b>311</b> |
| 11.2.2 传输层协议简介 .....                    | 275        |  |            |
| 11.2.3 套接口编程简介 .....                    | 279        |  |            |

# 第 1 章 绪 论

随着科学技术的发展,信息已成为推动社会向前发展的巨大动力,信息网络在当今世界各国经济与社会发展中起着决定性作用,信息领域的竞争也成为世界经济竞争的焦点之一。现有信息网络大致可以分为互联网、电信网和广播电视网三大类。其中,诞生于上世纪六、七十年代的互联网出现最晚,发展也最快,已经充分融入人们的生产和生活之中。

在当今互联网的体系结构中,路由器处于非常重要的地位,它负责对网络中的数据进行交换和转发,事实上,正是由于路由器的工作,才把世界各地大大小小的网络连接在一起,构成了一个全球范围内的 Internet。另一方面,IPv6 作为下一代互联网的关键技术,日益受到人们的重视,研究 IPv6 相关理论与技术具有深刻的现实意义和长远的学术意义。基于上述原因,本书主要讲述路由器、特别是 IPv6 路由的基本原理与技术,包括路由器的工作原理,路由器的硬件体系结构,三大路由协议 RIP(Routing Information Protocol,路由信息协议)、OSPF(Open Shortest Path First,开放最短路径优先)、BGP(Border Gateway Routing Protocol,边界网关路由协议)等。此外,网络安全和管理是互联网正常运行的重要支撑技术,而随着互联网的发展,移动和组播等业务也日益受到人们的重视,因此,本书还将讲述移动路由技术、组播路由技术、网络安全协议和网络管理协议等。最后,为了使读者了解互联网的最新动态,本书还将结合编者在国家“973”项目研究中的成果,介绍未来互联网的发展趋势等。

为了使读者对互联网有一个更加清楚的认识,本章首先介绍互联网的基本概念,然后介绍互联网的体系结构模型,最后简要回顾互联网的发展历程。

## 1.1 互联网的基本概念

互联网从本质上讲,是一种基于 IP 协议、提供面向无连接服务的分组交换网。这里的 IP 是计算机网络中网络层(网络分层的概念见 1.2 节)的一种协议标准,1981 年的 RFC(Request for Comments,建议请求)791 对 IP 作了定义。下面分别介绍分组交换和面向无连接的概念。

分组交换(也称为包交换)是在 20 世纪 60 年代发展起来的,当时主要应用于军事环境下的通信基础设施。不同于电路交换,在采用分组交换的网络中,用于通信的计算机把它们的数据划分为若干个分组,也就是包,然后将这些分组通过分组交换网传输。分组交换技术有很多,如 X.25、帧中继(Frame Relay,FR)和 IP 等。

计算机在通信时为什么要采用分组交换呢?为什么不在计算机之间建立一条临时的“电话呼叫”呢?这种差别主要来源于电话网和计算机网通信特性的不同。在电话网中,用户在通话之前首先要发送一个呼叫,一旦呼叫成功,通话双方之间就建立了一条物理通路,这条通路直到通话结束才被释放,在整个通话过程中,它是一直被占用的。由于在电话网中很强调话音的质量,而且一次通话会持续较长的时间,其呼叫建立时间相对于通话的总时间来说是很短的,即呼叫建立的开销在整个通话中所占的比例不是很大,所以在电话网中采用电路交换还是比较合算的。

计算机通信就不一样了,大多数计算机通信常常是短暂的,一般在秒的量级,而且经常是“突发性”的。计算机可能在很短的时间之内向多个对等体(层)发送一段较短的数据比特,在每秒内建立多个连接,以实现多个对等计算机之间的“会话”。很显然,如果在计算机通信中采用电路交换,那么通信的效率会大大降低。因此要采用分组交换的方式。一般说来,分组交换有两种基本类型,即“面向连接”和“面向无连接”。

“面向连接”服务在网络层也被称为“虚电路”,它在每个虚呼叫的开头需要建立一条虚连接,以后的通信就在这条虚连接上进行。虚电路与传统的电路交换是有很区别的,在电路交换中,一个呼叫在其存在期间一直占用一条端到端的物理信道;而虚电路建立在分组交换的基础之上,尽管我们感觉是占用了一条端到端的物理信道,但实际上只是断续地占用一段又一段的链路。在虚电路中,没有必要让每个包都携带目的地址,只要建立连接时的数据包携带目的地址就可以了,其他后续的包使用短的虚电路号。虚电路的优点是可靠性比较高,但效率较低。

“面向无连接”是分组交换的另一种主要形式,IP 协议即采用这种服务。在面向无连接的模式中,数据发送之前并不建立一条可靠的连接,而是让每个包都携带完整的目的地址,即只管将包发送到网络中,至于如何将包转发到它所指明的目的地,完全依赖于中间的每个分组交换设备。面向无连接的服务不提供可靠性,它只是尽最大能力交付,这样可以达到效率的最大化,至于可靠性则由上层协议来负责。从目前互联网的发展趋势来看,在网络层提供面向无连接的服务是非常成功的,IP 得到广泛应用就是一个很好的例证。

在分组交换网中,分组交换设备是一个很关键的部件,它负责数据包的转发工作。互联网也不例外。互联网中的分组交换设备被称为“路由器”,如图 1-1 所示。路由器可以在其任何接口上接收数据包,并通过不同的接口把数据包转发出去,将其发送到目的地。图中的双向箭头用以指示数据包既可以从该接口进入路由器,也可以从该接口离开路由器,路由器是互联网的基本和核心部件。

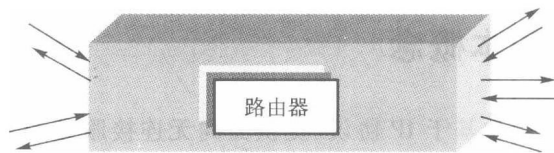


图 1-1 路由器的基本功能

在互联网上流动的元素就是 IP 包。就 IP 包所携带的数据而言,可以是传统的计算机数据,也可以是在 IP 上使用的数字化音频和视频流。处理之后的语音和视频流其实也是数据,但是,与传统的数据传输(例如文件传输)不同,它们在传输上具有特殊的需求,因为语音和视频流对时间十分敏感,并且对延迟的容忍也十分有限。不过,语音和视频流可以丢失一些数据而不会产生听觉或者视觉上的失真,但数据业务却必须保证传输的正确性,也就是说需要花费时间来重传丢失或者损坏的数据包以保证整个传输完好无损。

## 1.2 互联网体系结构模型

现有信息网络基本上都采用了分层的体系结构,即将其协议体系划分为若干个层次,每个层

次完成特定的功能,这样,各个层次综合在一起,就可以完成一个完整的系统功能。这种分层结构实际上体现了模块化设计的思想,具有设计简单、互换性好、标准化等优点。例如,OSI(Open System Interconnect,开放系统互联)参考模型是7层的垂直结构;互联网采用的是4层垂直结构。图1-2中给出了7层OSI参考模型和互联网的4层体系模型的结构比较。在这两种模型中,每层的功能都由一个以上的实体来执行。一个层中的实体与紧接着它的下一层中的实体直接相互作用,并为上一层提供服务。

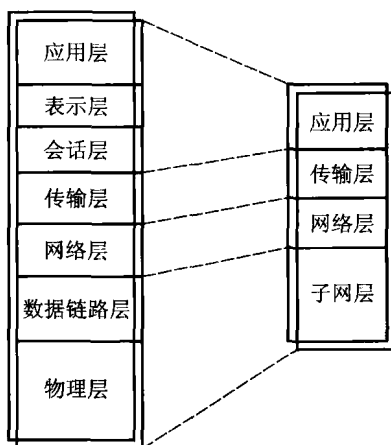


图1-2 OSI模型与互联网体系结构模型比较

由图1-2可以看出,互联网的体系结构包括子网层、网络层、传输层和应用层等4层。其中,重点是网络层、传输层和应用层。

子网层一般又称网络接口层,它大致相当于OSI参考模型中的物理层和数据链路层,负责从网络层接收IP报文并向物理网络发送,或从网络上接收物理帧,取出IP数据报并提交给网络层。互联网协议栈本身对子网层没有作太多的定义,只是要求能够使用某种协议来传输网络层的分组(即IP分组)。事实上,当今互联网的底层基本上是按照OSI的物理层和数据链路层建立的。

网络层负责处理分组在网络中的活动,提供跨越多个网络的选路功能,并对上层屏蔽底层具体子网技术的细节。在互联网中,一般常把网络层称为IP层。对IPv4协议栈来说,该层的协议主要包括IP、ICMP(Internet Control Messages Protocol,因特网控制报文协议)以及IGMP(Internet Group Management Protocol,因特网组管理协议)等;对IPv6协议栈来说,该层的协议主要包括IPv6、ICMPv6(Internet Control Messages Protocol version 6,因特网控制报文协议版本6)、MIPv6(Mobile IP version 6,移动IPv6)等。

传输层主要为两台主机上的应用程序提供端到端的通信。在互联网中,主要有两个互不相同的传输协议:TCP(Transmission Control Protocol,传输控制协议)和UDP(User Datagram Protocol,用户数据报协议)。TCP是面向连接的传输控制协议,用于为应用层的数据传输提供可靠连接;UDP则为应用层提供无连接的服务。

应用层负责处理特定的应用程序细节,为用户完成各种网络服务,它所包含的主要协议有:Telnet(远程登录),FTP(File Transfer Protocol,文件传输协议),SMTP(Simple Mail Transfer

Protocol,简单邮件传送协议)、SNMP(Simple Network Management Protocol,简单网络管理协议)和HTTP(HyperText Transfer Protocol,超文本传输协议)等。本书所讲述的各种路由协议(RIP和BGP等工作在应用层,而OSPF则直接位于网络层之上)大部分也工作在应用层,并为网络层提供服务。

在互联网体系结构的各个协议中,最重要的有两个:即TCP和IP。因此,人们经常用“TCP/IP”来表示互联网的体系结构或者整个TCP/IP协议族。

## 1.3 互联网的发展历史

### 1.3.1 基于IPv4的互联网

互联网从诞生到现在不过几十年的历史,它最早可以追溯到1969年的ARPANET。当时,美国国防部高级研究计划署资助建立了一个将4所大学的计算机主机连接在一起的网络,该网络采用分组交换技术,各节点通过专用的交换设备和通信线路进行互联。可以说,ARPANET就是互联网的前身和雏形。

在ARPANET诞生后5年,即1974年,Vinton Cerf和Robert Kahn提出了TCP/IP协议族的概念,用以解决不同计算机网络的互联问题。1981年,IP的协议规范RFC 791和TCP的协议规范RFC 793出现,并在1983年成为ARPANET的正式标准,这使ARPANET的规模迅速扩大。美国国防部也向全世界公开了TCP/IP的技术细节。这样,就出现了各种各样的网络,并利用TCP/IP协议连接在一起,从而导致了Internet的诞生。

20世纪80年代中期,美国国家自然科学委员会也开始筹建自己的网络,即NSFNET,这个网络连接了大学和政府职能部门的许多超级计算机,并通过不断的升级来提高网络的能力,最初是56 kbps的网络线路,后来提高到T-1(1.544 Mbps),直到最后为T-3(43 Mbps)。到20世纪90年代初,NSFNET已经连接了3 000多个子网,并取代旧有的ARPANET成为Internet的正式骨干网,随着计算机网络在全球的拓展与应用,美国之外的网络也逐渐加入NSFNET或者其子网,使Internet的规模不断扩大。1995年,NFSNET被一个更有竞争力、商业化更强的骨干网代替,它减少了将主机连向Internet的限制,从而将Internet向商业用户开放。

值得一提的是,1992年,欧洲粒子物理实验室提出了一个称为WWW(World Wide Web)的概念,随后一年,发布了称为Mosaic的WWW客户程序。这是Internet发展史上一个划时代的事件,因为它使得Internet从一个由科学家和研究人员使用的文本工具转变为普通人就可以使用的图形工具,对Internet的大规模推广和应用起到了非常重要的作用,最终使Internet成为全球范围内最大的互联网。

### 1.3.2 基于IPv6的互联网

近年来,随着世界各国信息化进程的推进,以TCP/IP协议体系为基础技术支撑的互联网在取得巨大成功的同时,也面临着越来越多的挑战和问题。这些挑战和问题产生的直接原因是互联网规模的迅速膨胀和各种新业务的不断出现;而其根本原因则是现有互联网使用的IP协议——IPv4存在着诸多设计上的缺陷。当然,把这些问题归咎于互联网的最初设计者是不公正

的。客观地讲, TCP/IP 协议体系的早期设计者们, 谁也没有想到互联网能在全世界范围内得到如此广泛的应用, 当然更不会预见到由此而产生的各种问题。

归纳起来, 现有 IPv4 协议在以下几方面存在着明显的缺陷。

#### (1) 地址空间不足。

IPv4 可提供 32 位(b) 的 IP 地址, 虽然理论上  $2^{32} \approx 43$  亿个地址, 但在实际使用中, 将这些地址划分成五类, D、E 类保留作特殊用途, 只有 A、B、C 三类用于 IP 网络的地址分配; 另外, 还要去除网络地址、广播地址、划分子网的开销、路由器地址、保留地址等, 最后有效的地址数目比总数要低得多; 再加上早期对 IP 地址缺乏长远规划, 使得有效地址的使用率较低。虽然一些缓解地址需求的新技术开始出现, 其中包括 CIDR (Classless InterDomain Routing, 无类别域间路由)、NAT (Network Address Translation, 网络地址转换) 等, 这些技术一定程度上缓解了地址空间被耗尽的危机, 但为基于 IP 的网络增加了复杂性, 并且破坏了 IP 协议的一些核心特性, 比如端到端原则等, 因此不能从根本上解决 IPv4 当前面临的困难。

#### (2) 配置复杂。

对于 IPv4 节点的配置一直比较复杂, 而网络管理员和用户更喜欢“即插即用”, 即: 将计算机插在网络上就可以开始使用。而 IPv4 主要采用手工配置的方法来给用户分配地址, 这不仅增加了管理的费用, 而且无法为那些需要移动的 IP 用户提供更好的服务。虽然 IPv4 也可以采用 DHCP (Dynamic Host Configuration Protocol, 动态主机配置协议) 允许系统在启动或者只在需要时才通过服务器获取其 IP 网络配置, 但这样需要一个额外的服务器, 因此, 也需要额外的操作和维护。

#### (3) 服务质量差。

IPv4 在最初设计时, 主要是面向数据业务, 没有考虑对传输语音、视频等业务的支持。随着互联网业务的发展, 人们希望能够在互联网上传输多种不同类型的业务, 并作了很多有益的尝试, 但由于 IPv4 最初设计时没有考虑服务质量问题 (尽管 IPv4 包头中有一个 TOS 字段, 但一般的路由器都会忽略它), 所以现有的 IPv4 网络很难承载各种对服务质量要求比较高的业务。

#### (4) 安全性不高。

早期的互联网主要用于科学研究, 安全问题不突出。随着互联网的商用化, 现有 IPv4 网络暴露出越来越多的安全缺陷, 各种网络安全事件层出不穷。其中一个重要原因是: 在 IPv4 时代, 人们认为安全性在网络协议栈的底层并不重要, 安全性的责任应交给应用层。在这种情况下, 安全性就意味着只对净荷数据加密。即使应用层数据本身是加密的, 携带它的 IP 数据仍会泄露给其他参与处理的进程和系统, 这样就使得 IP 数据包容易受到诸如信息包探测、连接截获等手段的攻击。需要说明的是, 尽管用于网络层加密与认证的 IPSec (IP Security, IP 安全) 协议可以应用于 IPv4 中, 保护 IPv4 网络层数据的安全, 但 IPSec 只是作为 IPv4 中的一个可选项, 没有任何强制性措施用以保证 IPSec 在 IPv4 中的实施。

#### (5) 移动性支持差。

IPv4 诞生时, 互联网的结构还是以固定、有线为主, 所以 IPv4 没有考虑对移动性的支持。但到了 20 世纪 90 年代中期, 各种无线、移动业务的发展要求互联网能够提供对移动性的支持。因此, 研究人员提出移动 IPv4 来解决这些问题。但由于 IPv4 本身的缺陷, 造成移动 IPv4 存在着诸

多弊端,如三角路由问题、安全问题、源路由过滤问题、转交地址分配问题等。事实上,移动 IPv4 没有得到大规模应用也是由这些问题造成的。

鉴于 IPv4 存在的上述缺陷在某种程度上限制了互联网的进一步发展,因此,IETF(Internet Engineering Task Force,因特网工程任务组)从 20 世纪 90 年代就着手制订下一代网际协议——IPv6,以解决这些问题。到目前为止,IPv6 已经形成了比较完善的协议体系,并得到广泛的部署和应用。

IPv6 的协议标准并不是一下子就提出来的,它也经历了一个长期的发展过程。事实上,到 1994 年,已经有三个主要提案提交给 IETF,它们分别是 SIPP(简单增强 IP)、TUBA(更多地址的 TCP 和 UDP)和 TP/IX。

SIPP 整合了两个工作组——SIP(简单 IP)和 PIP(增强 IP)的工作。其中,SIP 是针对 IPv4 提出了一个最小幅度的修改建议,采用 64 位地址和一个不同的首部格式;PIP 采用了更大的、可变长度的和有层次结构的地址(支持以 16 位为单位的变长地址,地址间通过标识符进行区分,能够支持选择承载商、移动性和高效转发),而且首部格式也不相同。后来这两个工作组意识到 PIP 中大量的先进特征可以由 SIP 在不修改基本的 SIP 协议的前提下完成,从而在本质上使得各种协议的最佳特征都可以被保持。基于这一点,两个工作组合并为一个新的 SIPP 组。

TUBA 基于 OSI 的 CLNP(Connectionless Network Protocol,无连接网络协议),它是一个与 IP 类似的 OSI 协议,可以提供大得多的地址空间,可变长度达 20 字节(B)。RFC 1347 提供了 TUBA 的有关细节。

TP/IX 采用了 64 位的地址,此外,它还改变了 TCP 和 UDP 的格式:两个协议均为 32 位的端口号,64 位的序列号,64 位的确认号,以及 TCP 的 32 位窗口。RFC 1475 对它进行了描述。

1995 年,IETF 发布了 RFC 1752,描述了下一代 IP 协议的特征,并对上述三个提案做了评价。最后,IETF 以 SIPP 为基础,对 SIPP 作了适当的修改(如将其中的地址长度扩展为 128 位),同时吸纳了其他两个提案的优点,制订了下一代互联网的协议——IPv6。1996 年,描述 IPv6 及其支持协议的 RFC 出现(RFC 1883 ~ 1887,这些 RFC 基本上都已经被新的标准所代替)。1998 年,新的描述 IPv6 的协议标准 RFC 2460 取代了旧的 RFC 1883,新的描述 IPv6 地址结构的 RFC 2373 代替了 RFC 1884,而这个 RFC 在 2003 年又被 RFC 3513 所代替,2006 年的 RFC 4291 又代替了 RFC 3513。此外,IPv6 的相关协议标准如邻居发现等也相继被提出。到目前为止,IPv6 已经形成了比较完善的协议体系。

如前所述,提出 IPv6 的背景是传统 IPv4 网络存在着各种缺陷,难以满足互联网发展的需要。那么,IPv6 有哪些优点,或者说,相对于 IPv4,它有哪些改进呢?归纳起来,大致有以下几点。

#### (1) 地址空间的扩展。

IPv6 将地址空间从 IPv4 的 32 位扩展到了 128 位,可以满足相当长一段时间内对 IP 地址的需求。此外,还对主机可能获得的不同类型地址作了一些调整,如取消了广播地址,而用组播地址来代替,增加了任意播地址等。关于 IPv6 的地址在本书的第 2 章还会详细介绍。

#### (2) 首部格式的简化。

IPv6 首部包括总长为 40 字节的 8 个字段,而 IPv4 首部包含至少 12 个不同字段,且长度在没有选项时为 20 字节,但在包含选项时可达 60 字节。IPv6 使用了固定格式的首部,减少了需要检

查和处理的字段数量,这将使得选路的效率更高。

(3) 引入扩展首部,更好地支持移动性和安全性。

在 IPv4 中可以在 IP 首部的尾部加入选项,与此不同,IPv6 把选项加在单独的扩展首部中。通过这种方法,选项头只有在必要的时候才需要检查和处理。更为重要的是,通过引入扩展首部,可以大大增强 IPv6 协议的可扩展性,更好地支持移动性和安全性等。例如,在移动 IPv6 中,通过引入了新的扩展首部,移动设备就能利用家乡地址和转交地址之间的匹配转换把 IPv6 包透明地路由给移动节点,对通信节点和转交地址之间的路由进行优化,从而使网络的利用率更高;另外 IPv6 丰富的地址空间也使转交地址的大量分配成为可能。在安全性支持方面,IPv6 协议族定义了有关安全性的基本信息,如封装安全负荷(Encapsulated Security Protocol, ESP)和鉴别头(Authentication Header, AH)。这些功能均是 Internet 中的安全性通信过程所必需的。

(4) 引入了流标签等机制,以支持服务质量。

IPv6 协议中的流标签机制可保证路由器识别端到端的分组流,为服务质量提供一定的保证。

(5) 支持地址的自动配置。

IPv6 支持主机地址的无状态自动配置,可方便无 TCP/IP 经验的用户接入 Internet,即实现插即用的功能。这样用户不论在数据链路层的任何接入点接入网络都能与 Internet 上的其他接入点进行通信,而不需要额外的 DHCP 服务器的支持。

IPv6 最早是由 IETF 提出来的,并同时由它负责制订和完善。在 IPv6 技术作为下一代互联网网络层技术几乎已成定局的情况下,人们对 IPv6 技术的研究十分重视。下面简单分析一下 IPv6 技术的研究现状和发展趋势。

应该说,1996 年 IETF 建立的全球范围的 IPv6 实验床 6Bone(IPv6 BackBone)为 IPv6 的研究创造了一个良好的条件,因为 6Bone 为 IPv6 研究、开发和实践提供了一个必要的平台。到 1998 年初,IPv6 协议的基本框架已经逐步成熟,在越来越广泛的范围内得到实践。例如:1998 年,面向实用的全球性 IPv6 研究和教育网(IPv6 Research and Education Network, 6REN)启动,建立了物理上以 ATM 为中心的 IPv6 洲际网;2000 年 5 月,3G 标准化组织 3GPP(Third Generation Partnership Project)采纳 IPv6 为多媒体服务的必选协议。另外,由 Cisco、Nortel、Microsoft、Lucent、Nokia、3Com 等公司联合发起成立的 IPv6 论坛也极大地推动了 IPv6 的发展。

从世界各国研究的情况来看,目前在 IPv6 的研究和应用方面比较领先的主要是美国、欧洲和日本等发达国家和地区,都启动了一系列与 IPv6 相关的计划或项目。这些计划为 IPv6 在全球范围内的部署和应用起到了重大的推动作用。

我国在 IPv6 的研究方面起步不算早,但近几年发展特别快。事实上,我国是全球最关心 IPv6 发展的国家之一,原因就在于 IPv6 将给中国信息网络建设带来新的契机。因此,政府、各科研院所和企业对 IPv6 的研究十分重视,投入了大量的财力和人力,也取得了很大的成就。标志性的事件有:

- ① 1998 年 6 月,中国国家教育科研网 CERNET 加入 6Bone,建立了 IPv6 实验床。
- ② 1999 年 4 月,CERNET 启动 IPv6 实验床项目,极大地推动了中国 IPv6 技术的研究。
- ③ 北京交通大学下一代互联网互联设备国家工程实验室(前身是下一代互联网研究中心)



在2000年下半年在国内首家推出了支持IPv6协议的路由器,并在2002年1月通过了北京市科技成果鉴定;该实验室承担的国家“863”重大课题“高性能IPv6路由器协议栈软件”在2004年8月通过验收,研究成果中的RIPng、OSPFv3和BGP4+等IPv6动态路由协议软件达到了国际先进水平。

④ 2002年4月,中国科学技术部高新技术发展及产业化司决定对“863”计划信息技术领域的多个课题进行公开招标,其中多个课题是与IPv6技术相关的,它们分别是高性能IPv6路由器基础平台及实验系统、高性能IPv6路由器协议栈软件和IPv6协议测试技术等,这一事件表明我国以进入全面开展IPv6高端核心路由器技术的研究阶段。

⑤ 2003年3Tnet重大专项启动了IPv6/v4实验床和设备的研制项目。国家发改委、国家自然科学基金委员会、中国科学院等部门和单位也分别启动了IPv6研究和实施计划。

⑥ 2003年底,中国宣布实施名为“中国下一代互联网示范工程”(China Next Generation Internet, CNGI)的新一代互联网计划,经过几年的努力,到2009年已经建成了世界上最大的IPv6骨干网。

⑦ 2004年8月,北京交通大学下一代互联网互联设备国家工程实验室自主研制的BJTU IPv6无线/移动路由器通过了教育部组织的科技成果鉴定,与会专家一致认为该产品填补了国内空白,达到国际先进水平;2005年12月,该中心自主研制的IPv6微型传感路由器通过科技成果鉴定,为IPv6在传感器网络领域的应用创造了一个很好的开端。

⑧ 2009年2月温家宝总理主持召开国务院常务会议,审议并原则通过电子信息产业调整振兴规划。规划提出加大投入开展下一代互联网商业应用。

目前,我国在IPv6的研究上取得了很大的成绩,比起IPv4阶段已经有了长足的进步,但也存在一些问题,主要有:IPv6的相关标准很少有我国提出的,在标准制订方面我们尚需努力;在新一轮的IPv6地址分配中,我国仍然处在与人口不相配的劣势地位等。图1-3给出了我国在IPv4和IPv6地址分配中的基本情况。

| 名次 | 国家 | IPv4地址数目      | 百分比    | 名次 | 国家   | IPv6地址块(/48) | 百分比    |
|----|----|---------------|--------|----|------|--------------|--------|
| 1  | 美国 | 1 461 113 344 | 55.08% | 1  | 美国   | 984 100 015  | 22.40% |
| 2  | 中国 | 182 928 128   | 6.90%  | 2  | 德国   | 645 791 777  | 14.70% |
| 3  | 日本 | 151 575 040   | 5.71%  | 3  | 日本   | 547 307 530  | 12.46% |
| 4  | 英国 | 86 535 256    | 3.26%  | 4  | 法国   | 546 570 249  | 12.44% |
| 5  | 德国 | 81 763 256    | 3.08%  | 5  | 澳大利亚 | 542 187 540  | 12.34% |
|    |    |               |        | 18 | 中国   | 4 128 771    | 0.09%  |

2009年1月2009年12月

图1-3 我国在IPv4和IPv6地址分配中的地位

另外,需要指出的是,尽管在IPv6的研究上已经取得了巨大的成就,但目前IPv6还缺乏杀手铜级的应用,这实际上也是制约IPv6大规模应用的一个关键因素,并且决定了在很长一段时间内,将是IPv4和IPv6并存,而不太可能是IPv6迅速代替IPv4。因此,除了要研究IPv6的关键理论、技术与设备外,对IPv6应用的研究也非常重要。当前物联网的迅速发展也许会给IPv6的部署提供一个巨大的契机。