现代IP技术丛书

MODERN IP TECHNOLOGY

# IPv6 技术揭密

李振强　赵晓宇　马　严　等　编著

人民邮电出版社
POSTS & TELECOM PRESS

现代 IP 技术丛书

# IPv6 技术揭密

李振强　赵晓宇　马　严　等编著

人 民 邮 电 出 版 社

**图书在版编目（CIP）数据**

IPv6 技术揭密 / 李振强，赵晓宇，马严等编著. —北京：人民邮电出版社，2006.4
（现代 IP 技术丛书）

ISBN 7-115-14549-0

Ⅰ. I… Ⅱ. ①李… ②赵… ③马… Ⅲ. 计算机网络—传输控制协议 Ⅳ. TN915.04

中国版本图书馆 CIP 数据核字（2006）第 012137 号

## 内 容 提 要

本书是一本面向中高级读者的进阶读物，特别适合熟悉 IPv4、对 IPv6 有一定了解的读者进一步系统学习 IPv6 时使用。

本书内容丰富，既有 IPv6 协议的基本介绍，也有和 IPv6 相关的高级议题。特别的，本书部分内容来自作者单位多年来承担国家科研项目的经验总结。全书共分 10 章，分别介绍了计算机网络的基础知识和基本概念，IPv6 的由来，IPv6 的新特性，IPv6 基本协议，IPv6 路由技术，IPv6 过渡技术，IPv6 网络管理技术，IPv6 网络安全，移动 IPv6 技术，IPv6 在国内外的最新进展，以及 IPv6 在 Windows、Linux、freeBSD 和 Cisco IOS 等上的配置和实践。本书在每章的最后都给出了和该章内容相关的参考文献，以供读者进一步学习时参考和查阅。

本书可以作为大学本科高年级学生和研究生学习 IPv6 技术的参考教材。对于正在从事 IPv6 相关研究和开发的研究人员和工程技术人员，本书也具有很高的参考价值。

现代 IP 技术丛书

**IPv6 技术揭密**

◆ 编　　著　李振强　赵晓宇　马　严　等
责任编辑　陈万寿

◆ 人民邮电出版社出版发行　　北京市崇文区夕照寺街 14 号
邮编　100061　　电子函件　315@ptpress.com.cn
网址　http://www.ptpress.com.cn
北京通州太中印刷厂印刷
新华书店总店北京发行所经销

◆ 开本：787×1092　1/16
印张：26.5
字数：644 千字　　2006 年 4 月第 1 版
印数：1 – 3 600 册　　2006 年 4 月北京第 1 次印刷

ISBN 7-115-14549-0/TN • 2743

定价：49.00 元

**读者服务热线：(010)67129258　印装质量热线：(010)67129223**

# 前　言

北京邮电大学信息网络中心作为中国教育和科研计算机网（CERNET）的主节点之一，一直专注于计算机网络技术的研究工作。

北京邮电大学信息网络中心自 1999 年开始进行 IPv6 研究至今，先后承担了中国高技术研究发展计划（863 计划）、国家重点基础研究发展规划（973 计划）、国家自然科学基金委、国家计算机网络与信息安全中心、信息产业部等单位立项的多项科研课题，并同中国移动公司研发中心、Nokia、Cisco、KDDI Lab、NTT、Siemens、France Telecom 等国内外知名企业的研发机构合作进行了多项 IPv6 相关研究，研发内容涉及 IPv6 网络管理、IPv4 向 IPv6 的过渡技术、IPv6 网络安全、移动 IPv6、IPv6 域名管理、IPv6 QoS、IPv6 部署策略等。经过多年的探索和积累，北京邮电大学信息网络中心在 IPv6 研发和研究生教育方面积累了较多的经验，取得了一定的成果。我们一直期望能把这些成果、经验，甚至一些教训总结出来供大家参考，恰逢人民邮电出版社约稿，于是开始了本书的写作。

本书在兼顾 IPv6 初学者的同时，更多的面向有一定技术基础的 IPv6 研发人员，结合我们的研究心得对 IPv6 的一些关键技术进行重点介绍。本书写作时尤其注意了相关参考资料的收集和整理工作，尽量做到每个章节所讲述的内容都能在本章的参考文献中找到更加详细和权威的来源。因此，我们也希望本书成为广大 IPv6 研发人员的一本工具书，为大家的研发工作提供尽可能多的直接或间接的帮助。

本书由李振强、赵晓宇和马严组织编写，是信息网络中心集体智慧的结晶，是所有参与编著的老师和同学共同努力的结果。本书内容及其主要编著者如下：第 1 章，IPv6 概述，由李振强编写；第 2 章，IPv6 的新特性，由莫辉编写；第 3 章，IPv6 技术详解，由孔令宇编写；第 4 章，IPv6 路由技术，由张岩编写；第 5 章，IPv6 过渡技术，由张海涛编写；第 6 章，IPv6 网络管理技术，由马瑞编写；第 7 章，IPv6 网络安全，由孙国学编写；第 8 章，移动 IPv6，由王洋编写；第 9 章，IPv6 最新研究进展，由徐朝锋和汪守峰编写；第 10 章，IPv6 配置与实践，由田朝文编写。全书最后由李振强统稿，并整理常用缩略语对照表。

在本书的写作过程中，北京邮电大学信息网络中心的各位老师为我们提供了很多的帮助，在此向他们无私的帮助表示深深的感谢！

由于时间和能力所限，本书肯定会存在一些不够清晰、不够明确，甚至是不够正确的地方，恳请读者提出宝贵意见，以帮助我们改进和改正。

作　者

2006 年 1 月于北京邮电大学

# 目　录


**第 1 章　IPv6 概述** …… 1

1.1　计算机网络体系结构和参考模型 …… 1

1.1.1　协议、服务、分层和封装 …… 1

1.1.2　OSI 7 层参考模型 …… 5

1.1.3　TCP/IP 协议栈 …… 11

1.1.4　OSI 参考模型和 TCP/IP 协议栈的比较 …… 14

1.2　IPv4 简介 …… 15

1.2.1　互联网的历史和发展 …… 15

1.2.2　IPv4 协议简介 …… 17

1.2.3　IPv4 的局限性 …… 20

1.3　IPv6 的由来和国内外的发展状况 …… 22

1.3.1　IPv6 的由来 …… 22

1.3.2　IPv6 在国外的发展 …… 24

1.3.3　IPv6 在国内的发展 …… 26

1.4　参考文献 …… 28

**第 2 章　IPv6 的新特性** …… 29

2.1　地址及寻址 …… 29

2.1.1　128 位超大地址空间 …… 30

2.1.2　可聚合层次化地址结构 …… 31

2.2　高效报头 …… 32

2.2.1　简洁性 …… 33

2.2.2　高扩展性 …… 33

2.3　端点分片 …… 34

2.4　即插即用 …… 34

2.4.1　地址自动配置 …… 34

2.4.2　重新编址 …… 35

2.5　服务质量 …… 36

2.5.1　DSCP …… 36

2.5.2　Flow Label …… 37

2.6　安全性 …… 38

2.6.1　IPsec …… 38

2.6.2 AH ····· 39
2.6.3 ESP ····· 39
2.7 移动性 ····· 40
2.7.1 路由头和家乡地址目的地选项 ····· 41
2.7.2 三角路由优化 ····· 41
2.7.3 更好的安全性 ····· 43
2.8 其他 ····· 43
2.9 参考文献 ····· 44

**第 3 章 IPv6 技术详解** ····· 45

3.1 IPv6 地址 ····· 45
3.1.1 IPv6 地址空间及语法表示 ····· 45
3.1.2 IPv6 地址分类及分配状况 ····· 46
3.2 IPv6 基本报头（IPv6 Header） ····· 49
3.3 IPv6 扩展报头（IPv6 Extension Header） ····· 50
3.3.1 逐跳选项头（Hop-by-Hop Options Header） ····· 51
3.3.2 路由头（Routing Header） ····· 51
3.3.3 分段头（Fragment Header） ····· 53
3.3.4 认证头（Authentication Header） ····· 55
3.3.5 封装安全载荷头（ESP Header） ····· 56
3.3.6 目的选项头（Destination Options Header） ····· 58
3.3.7 选项格式及扩展 ····· 58
3.3.8 扩展报头的顺序 ····· 60
3.4 ICMPv6 ····· 60
3.4.1 ICMPv6 概述 ····· 60
3.4.2 ICMPv6 错误报文 ····· 62
3.4.3 ICMPv6 信息报文 ····· 64
3.5 邻居发现（ND） ····· 65
3.5.1 邻居发现概述 ····· 65
3.5.2 地址解析（Address Resolution） ····· 73
3.5.3 邻居可达性检测（Neighbor Unreachability Detection） ····· 76
3.5.4 路由器和前缀发现 ····· 78
3.5.5 重定向（Redirect） ····· 79
3.6 多播侦听者发现（MLD） ····· 81
3.6.1 多播侦听者发现概述 ····· 81
3.6.2 多播侦听者报文格式 ····· 82
3.6.3 多播侦听者状态转移图 ····· 87
3.7 地址自动配置 ····· 91
3.7.1 DHCPv6 ····· 92

3.7.2　无状态地址自动配置 …… 96
3.8　域名解析系统 …… 97
3.8.1　IPv6 域名系统 …… 97
3.8.2　自动域名更新 …… 99
3.9　参考文献 …… 100

**第 4 章　IPv6 路由技术** …… 101

4.1　IPv6 静态路由 …… 101
4.1.1　静态路由概述 …… 101
4.1.2　IPv6 静态路由在 Windows 平台上的设置 …… 102
4.1.3　IPv6 静态路由在 Linux 平台上的设置 …… 103
4.1.4　IPv6 静态路由在 IOS 平台上的设置 …… 104
4.2　IPv6 动态 IGP 路由 …… 104
4.2.1　RIPng …… 106
4.2.2　OSPFv3 …… 114
4.2.3　IS-ISv6 …… 140
4.3　IPv6 动态 EGP 路由 …… 151
4.3.1　EGP 路由协议简介 …… 151
4.3.2　BGP4+ …… 152
4.4　参考文献 …… 159

**第 5 章　IPv6 过渡技术** …… 160

5.1　过渡问题的提出 …… 160
5.1.1　向 IPv6 过渡的概念 …… 160
5.1.2　向 IPv6 过渡期存在的原因 …… 161
5.1.3　过渡到 IPv6 的意义 …… 161
5.1.4　过渡需要解决的问题 …… 162
5.2　过渡技术概述 …… 163
5.2.1　基本过渡技术介绍 …… 163
5.2.2　各种过渡技术的应用场景 …… 165
5.3　双栈技术 …… 166
5.3.1　基本双栈技术 …… 166
5.3.2　双协议栈过渡机制 DSTM …… 168
5.4　隧道技术 …… 171
5.4.1　隧道技术概述 …… 171
5.4.2　手工隧道 …… 174
5.4.3　自动隧道 …… 174
5.4.4　隧道代理 …… 178
5.5　翻译技术 …… 179

5.5.1 无状态 IP/ICMP 翻译 SIIT ····· 180
5.5.2 NAT-PT ····· 181
5.5.3 应用层网关 ALG ····· 183
5.5.4 BIS ····· 187
5.5.5 其他 ····· 188
5.6 参考文献 ····· 190

**第 6 章 IPv6 网络管理技术** ····· 191

6.1 网络管理概述 ····· 191
6.1.1 网络管理的概念 ····· 191
6.1.2 网络管理的功能 ····· 191
6.1.3 网络管理的体系结构 ····· 192
6.1.4 简单网络管理协议（SNMP） ····· 193
6.2 IPv6 网络管理 ····· 203
6.2.1 IPv6 网管概述 ····· 203
6.2.2 IPv6 MIB ····· 208
6.2.3 IPv6 拓扑发现 ····· 212
6.3 网络管理新技术 ····· 215
6.3.1 基于 XML 的网络管理 ····· 215
6.3.2 基于移动代理的网络管理 ····· 219
6.4 参考文献 ····· 223

**第 7 章 IPv6 网络安全** ····· 225

7.1 网络安全概述 ····· 225
7.2 IPsec 协议 ····· 225
7.2.1 IPsec 简介 ····· 225
7.2.2 IPsec 策略 ····· 230
7.2.3 IPsec 在实际部署中遇到的问题 ····· 232
7.3 IKE 协议 ····· 234
7.3.1 IKE 简介 ····· 234
7.3.2 IKE 认证和协商 ····· 241
7.3.3 IKE 新进展 ····· 251
7.4 IPv6 安全脆弱性 ····· 251
7.4.1 IPv6 安全脆弱性分类 ····· 251
7.4.2 过渡时期的安全问题 ····· 252
7.4.3 IPv6 特有的安全问题 ····· 254
7.5 IPv6 防火墙 ····· 256
7.5.1 防火墙的基本原理 ····· 256
7.5.2 防火墙的实现技术 ····· 258

7.5.3 防火墙的安全策略 …… 259
7.5.4 针对 IPv6 的防火墙设计 …… 259
7.6 IPv6 入侵检测系统 …… 262
7.6.1 SNORT 入侵检测系统原理 …… 264
7.6.2 IPv6 入侵检测系统关键问题研究 …… 265
7.6.3 入侵检测系统与防火墙的联动 …… 274
7.7 参考文献 …… 279

**第 8 章 移动 IPv6** …… 281

8.1 移动 IPv6 概述 …… 281
8.2 移动 IPv6 对 IPv6 的扩展 …… 282
8.2.1 移动头（Mobility Header） …… 283
8.2.2 移动选项（Mobility Options） …… 288
8.2.3 家乡地址选项（Home Address Option） …… 290
8.2.4 第二类路由头（Type 2 Routing Header） …… 291
8.2.5 对 ICMPv6 的扩展 …… 291
8.3 移动 IPv6 协议详解 …… 293
8.3.1 移动 IPv6 协议实体和主要术语 …… 293
8.3.2 移动 IPv6 操作过程 …… 294
8.3.3 通信对端操作 …… 298
8.3.4 移动节点操作 …… 300
8.3.5 家乡代理操作 …… 305
8.4 移动 IPv6 高级议题 …… 308
8.4.1 快速切换技术 …… 308
8.4.2 移动 IPv6 的安全问题 …… 312
8.4.3 移动 IPv6 性能 …… 318
8.4.4 网络移动 …… 321
8.5 参考文献 …… 322

**第 9 章 IPv6 最新研究进展** …… 323

9.1 IPv6 标准研究情况 …… 323
9.1.1 IPv6 工作组 …… 323
9.1.2 v6ops 工作组 …… 323
9.2 IPv6 网络的发展现状 …… 324
9.2.1 IPv6 在美国 …… 324
9.2.2 IPv6 在欧洲 …… 326
9.2.3 IPv6 在韩国 …… 328
9.2.4 IPv6 在日本 …… 329
9.2.5 IPv6 在中国 …… 331

9.3 IPv6 的最新应用 …… 337
9.3.1 移动 IPv6 …… 337
9.3.2 网络视频监控技术 …… 341
9.3.3 “面对面”的在线游戏 …… 342
9.3.4 家庭网络 …… 342
9.3.5 IPTV …… 343
9.3.6 传感器网络 …… 345
9.3.7 RFID …… 345
9.3.8 VoIP …… 345
9.4 参考文献 …… 346

**第 10 章 IPv6 配置与实践** …… 347

10.1 软件支持情况 …… 347
10.1.1 Linux …… 347
10.1.2 FreeBSD …… 348
10.1.3 Microsoft Windows …… 348
10.1.4 Cisco IOS …… 348
10.1.5 Juniper JUNOS …… 349
10.1.6 其他 …… 349
10.2 软件配置命令 …… 350
10.2.1 Linux …… 350
10.2.2 FreeBSD …… 360
10.2.3 Microsoft Windows …… 366
10.2.4 Cisco IOS …… 382
10.3 案例分析 …… 399
10.4 参考文献 …… 410

**附录 常用缩略语** …… 411

# 第1章 IPv6概述

IPv6和IPv4一样，是网络层的一个协议，IPv6是网际协议版本6，IPv4是网际协议版本4。IPv6和IPv4相比，既有相同之处，也有不同之处，关于它们异同的比较和介绍会贯穿本书。为了能够更好地理解IPv6，特别是能够更好地掌握IPv6相比IPv4增强改进的地方，本章做一些背景知识的介绍和交待。本章内容包括理解计算机网络运行机理和网络协议工作原理必须知道的计算机网络体系结构和参考模型，包括协议、服务、分层和封装的概念；然后介绍现在互联网的基础协议——IPv4；最后介绍IPv6的由来及其在国内外的发展部署情况。

## 1.1 计算机网络体系结构和参考模型

本节首先抽象介绍协议、服务、分层和封装的概念，引出计算机网络体系结构和参考模型的定义；然后分别介绍两个著名的计算机网络体系结构和参考模型：OSI 7层参考模型和TCP/IP 4层参考模型；最后是这两个参考模型的异同比较和总结。

### 1.1.1 协议、服务、分层和封装

1．通信协议

计算机网络是通过通信线路和设备将分布在不同地理位置的、各自独立的计算机系统或者智能设备互联起来，从而实现数据通信、信息和资源共享的网络。两台计算机之间要进行通信，必须使用相同的信息交换规则。在计算机网络中用于规定信息的格式以及如何发送和接收信息的一套规则称为网络协议（Network Protocol）或者通信协议（Communication Protocol）。

2．协议分层

通信协议必须能够处理计算机网络中可能发生的各种情况：硬件故障，网络拥塞，数据包延迟、错序、重复、错误，甚至丢失，等等。这使得设计一个单一、庞大的协议来处理所有这些情况非常困难。为了降低网络协议设计的复杂性，网络协议设计者并不是在一个通信协议中规定完整的细节，而是采用“分而治之”的方法，将计算机网络按照信息流动和逻辑功能划分成若干层，然后为每个层次设计一个或者多个单独的协议。这样每个协议的设计、分析、实现和测试都比较容易。这种方法称为协议分层。分层后，每一层都建立在它的下一层之上，每一层都要完成一定的功能，并为它的上一层提供一定的服务，某层功能的实现细节对其上一层加以屏蔽。不同的网络，其层次的数量、各层的名称、内容和功能可以不尽相同。

图 1.1 所示是一个 5 层协议。假定用户 A 通过计算机 A 要将一段信息传递给使用计算机 B 的用户 B。要传递的信息并不是直接从计算机 A 的第 5 层传送到计算机 B 的第 5 层。在计算机 A，传递的信息从上到下要依次经过第 5 层、第 4 层、第 3 层、第 2 层和第 1 层，最后经过物理通信线路到达计算机 B。在计算机 B，接收到的信息从下到上要经历一个相反的过程，经过第 1 层、第 2 层、第 3 层、第 4 层和第 5 层，最后用户 B 才能看到用户 A 传来的信息。可见，有了协议分层，协议实际上运行在不同机器的对等层（Peer）之间，是对等层之间的虚拟通信（图中的虚线箭头）；同一机器的相邻层之间通过接口进行交互：下层通过接口为上层提供服务，上层通过接口调用下层提供的服务（图中的实线箭头）。

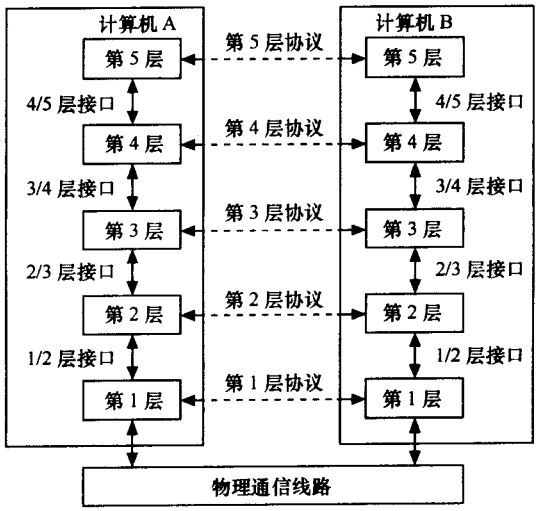


图 1.1　分层、协议和接口

需要特别指出的是，计算机网络中协议分层不同于程序设计中的模块化。在程序设计中，各个模块之间的耦合性越低越好，它们甚至可以相互独立、任意拼装或者并行执行；而协议分层是依信息的流动而产生的，往往隐含有上下层，或者服务和被服务的关系。

3．服务和接口

在协议分层中已经提到了服务和接口，下面给出它们的定义和详细说明。

服务是各层向它的上一层提供的一组原语（Primitive）或者一组操作。服务隐含有上下级的关系，第 $n$ 层提供的服务为 $n$+1 层所利用。这种情况下，第 $n$ 层称为服务提供者（Service Provider），第 $n$+1 层称为服务消费者或者服务用户（Service User）。在分层模型中，某层可能既是服务提供者，又是服务用户。例如图 1.1 中的第 3 层，相对于第 4 层，它要提供服务，是服务提供者；同时它又要利用第 2 层提供的服务，因此，相对于第 2 层它又是服务用户。

某层要为它的上一层提供一定的服务就必须完成和实现一定的功能，但某层提供的服务又不同于某层所完成的功能。$n$ 层服务是 $n$ 层实体向外部用户环境的直接作用，是用户可以使用和看得见的；而 $n$ 层所实现和完成的功能则是 $n$ 层实体为向 $n$+1 层提供服务所进行的一切活动，$n$ 层功能的实现保证了 $n$ 层服务得以向上一层提供，但 $n$ 层服务用户只能看到 $n$ 层服务，却看不到 $n$ 层的全部功能。因此，并非在 $n$ 层内完成的全部功能都称为 $n$ 层服务，只有那些能够被高一层看得见和利用的功能才是服务。

接口是相邻两层之间的界面，是相邻层之间交换信息的连接点。$n$+1 层通过它和 $n$ 层之间的接口调用 $n$ 层提供的服务，$n$ 层通过它和 $n$+1 层之间的接口为 $n$+1 层提供服务。接口被抽象表示为服务访问点 SAP（Service Access Point），如图 1.2 所示。

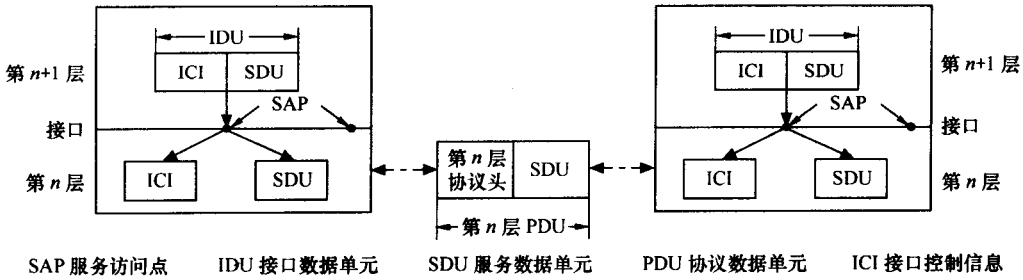


图 1.2　接口及相邻两层间的关系

相邻层之间要交换信息，对接口必须有一致同意的规则。在接口上，$n$+1 层实体通过服务访问点把一个接口数据单元 IDU（Interface Data Unit）传递给 $n$ 层实体。IDU 由服务数据单元 SDU（Service Data Unit）和一些控制信息组成。这些控制信息就是图 1.2 中的接口控制信息 ICI（Interface Control Information），用于帮助下一层完成任务，它本身不是要传递数据的一部分。SDU 是将要跨过网络传递给对等实体，然后向上交给 $n$+1 层的信息。为了传递 SDU，$n$ 层实体可能将 SDU 分成几段，每一段加上一个协议头后作为独立的协议数据单元 PDU（Protocol Data Unit）送出。对等层之间的协议就是用于规定协议头的格式以及如何发送和接收 PDU 的规则，就如同人和人之间进行交流时使用的共同的语言。

在计算机网络中，下层通过接口向上层提供两种不同类型的服务：面向连接的服务（Connection-oriented Service）和无连接服务（Connectionless Service）。面向连接的服务和电话系统相类似。使用电话系统和某人通话时，要先拿起电话，然后拨号，再进行通话，最后通话结束挂断电话。同样，使用面向连接的服务时，也要经历建立连接，使用连接，然后释放连接的过程。这种类型的服务可靠性高，但因为要有连接的建立和拆除，往往比较复杂，开销也较大，比较适合文件传输。

无连接服务和邮政系统相类似。每个报文（相当于邮政系统中的信件）都带有完整的目的地址和源地址（相当于信封上的信息），并且每个报文都独立于其他报文在网络中进行传输。到达同一目的地址的报文可能经历不同的传输路线，并且先发出的报文，可能因为传输延误，而比后发出的报文到达得还要晚。这就意味着，使用无连接服务的用户必须解决所有这些无连接服务可能带来的问题。无连接服务虽然可靠性较低，但简单高效，特别适用于实时业务或者多媒体业务的传输。

4．封装

要完全理解协议分层中各层间的交互方法并实现无缝互操作，除掌握服务和接口的概念外，还必须理解封装（Encapsulation）的概念。在某种意义上，如果一块数据以某种方式打包以便传输，这时就发生了封装。理解封装在网络模型中工作方式的最好办法是简单地跟踪协议栈中的流程。下面以 5 层网络中的一个客户机向服务器的一次查询为例来加以说明。

如图 1.3 所示，客户端用户输入一个查询 Q，在第 5 层要对 Q 进行封装，构成第 5 层的协议数据单元 PDU。该 PDU 中包含了查询数据 Q，并用如何处理 Q 的信息将 Q “包起来”。

这些信息包括：远端主机上目的应用的逻辑名、地址或其他指针，以及下一层（第 4 层）正确处理该包所需的必要信息。

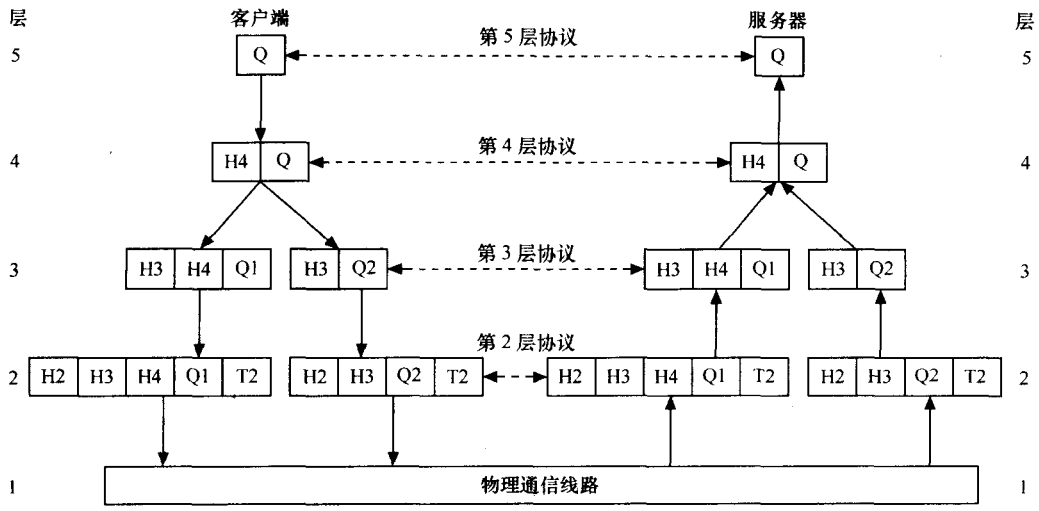


图 1.3　封装及虚拟通信

第 4 层简单地将从第 5 层传递来的包作为位串，在其前面加上用以正确识别该消息的报头（Header）后交给第 3 层进行传输。报头包括控制消息，如序列号等，以使服务器上的第 4 层能在下层未保持数据报顺序时能够按正确的顺序递交。

第 4 层往往对消息的长度没有限制，对消息长度的限制一般在第 3 层。因此，第 3 层实体可能根据需要将要传送的消息分成若干较小的单元，称为分组，并在每个分组的前面加上第 3 层报头。图 1.3 中的查询 Q 被分成了两个分组，Q1 和 Q2。

第 3 层实体决定使用哪条线路来传输该报文，并在第 3 层报头中加上客户机和服务器的第 3 层地址等内容，然后将整个包交给协议栈中的下一层——第 2 层处理。第 2 层不仅给每段信息加上报头信息，而且还在报文的尾部加上了控制信息，如用于错误检测的循环冗余校验 CRC（Cyclic Redundancy Check）等。最后，第 2 层封装的结果到达第 1 层，通过实际的物理通信线路（例如光纤）传送到服务器。

当第 2 层报文到达其目的地——服务器时，报文依次向上传递，每传递一层，该层的报头就被去掉。决不会出现把带有 *n* 层以下报头的报文交给 *n* 层的情况。最后，查询消息 Q 到达服务器的第 5 层。第 5 层在去掉封装后对查询进行处理和响应。在数据离开发送方之后直至到达接收方之前，低层操作的协议不对数据中的净荷 Q 进行处理。虽然可能有这样那样的完整性检查，除了高层提供的封装信息（头/尾信息）之外，低层协议无需查看其他部分的数据。

只要所有的中间系统能正常操作，且两个端系统使用的应用软件可以互操作，封装机制就使得连接在不同网络上的不同主机可以进行无缝互操作，而不管是什么样的系统类型、网络结构和物理接口。

5．计算机网络体系结构

有了以上这些概念，现在给出计算机网络体系结构的定义。计算机网络体系结构就是计

算机网络中的分层模型以及各层功能和所提供服务的精确定义，如图 1.4 所示。对网络体系结构的描述必须包括足够的信息，使得协议的实现者根据网络体系结构的描述就可以进行硬件设计或者软件编码，而不会引起歧义，造成互通问题。但是，网络协议实现的细节并不属于网络体系结构的内容，网络体系结构只进行功能的定义和接口的描述。

计算机网络体系结构 = {层，协议，接口}

图 1.4　计算机网络体系结构

不同的网络可以具有不同的网络体系结构，其层次的数量、各层的名字、功能、对等层之间的协议、相邻层之间的接口都可以不同。下面两节分别介绍两个具体的计算机网络体系结构——OSI 7 层参考模型和 TCP/IP 4 层协议栈。

### 1.1.2　OSI 7 层参考模型

国际标准化组织 ISO（International Standardization Organization）的开放系统互连 OSI（Open System Interconnection）参考模型（Reference Model）共规定了 7 个层次，如图 1.5 所示。图中的实线表示实际通信，虚线表示对等层间的虚拟通信。OSI 参考模型的 7 个层次就像洋葱的层次一样，每一层都将其下面的层遮起来。在上面的层里，看不到下面层次的细节。下面从最底层——物理层开始一一介绍 OSI 参考模型的 7 个层次及其功能和协议。

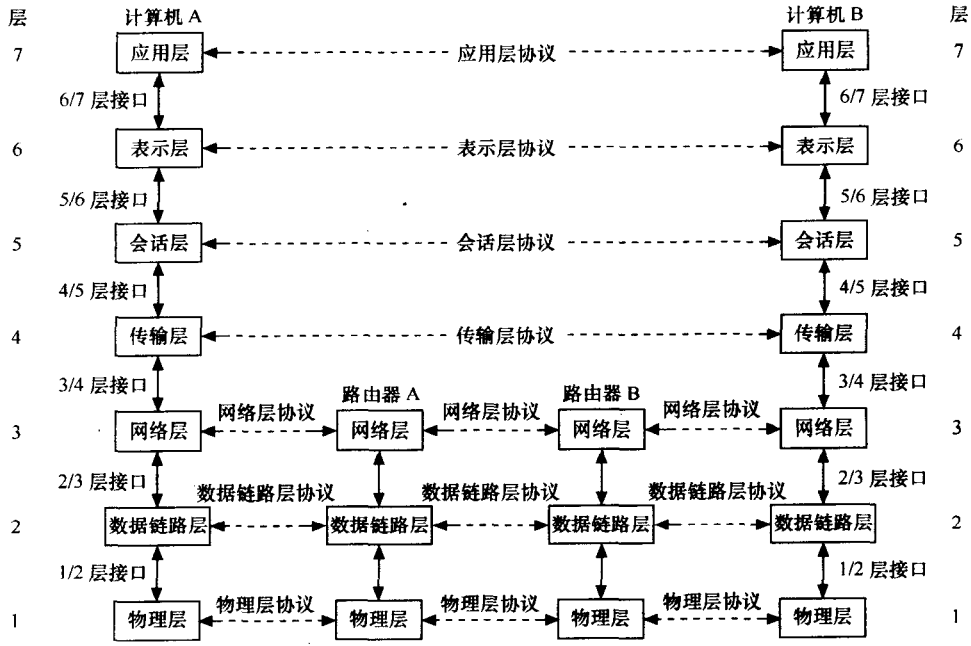


图 1.5　OSI 7 层参考模型

1．物理层

物理层是 OSI 参考模型的第 1 层，它虽然处于最底层，却是整个开放系统互连的基础。物理层为设备之间的数据通信提供传输媒体及互连设备，为数据传输提供可靠的环境。但物理层并不是指连接计算机的具体的物理设备或具体的传输介质，而是对有关连接接口的标准

和特性的描述。物理层为数据链路层提供一个物理连接，构造一个透明通信信道以传输各种数据的比特流。网络中的通信设备和传输介质种类繁多，通信方式和通信手段也多种多样，物理层的作用正是要屏蔽掉这些差异，使上层的数据链路层看不见这些差异，也不必考虑具体的网络通信设备和传输介质。

物理层主要包括三个基本功能。（1）物理连接的建立、维持和拆除。当数据链路层实体有建立连接的请求时，物理层实体使用有关的接口协议来完成连接的建立，并在数据传输过程中维持这个连接。传输结束后，对连接进行拆除。（2）数据传输。在物理连接即物理链路存在期间，在物理层协议控制下完成物理层数据单元——比特的传输。传输方式根据需要可以采用同步方式，也可以采用异步方式。（3）物理层管理。对物理层内的一切活动进行管理，具体内容和物理层协议相关。

物理层规定数据终端设备 DTE（Data Terminal Equipment）和数据线路设备 DCE（Data Circuit Equipment）接口标准的四个特性：（1）机械特性，规定物理接口的几何尺寸、排列方式、引脚数量和适用范围等；（2）电气特性，规定电气接口的连接方式和电气参数，如信号电平的范围和意义、驱动器的输入和输出阻抗、最大传输速率和传输距离等；（3）功能特性，规定物理层连接器各芯线的含义、功能以及各信号之间的对应关系等；（4）规程特性，规定物理层接口界面上进行信号传输的控制过程和控制步骤。

DTE 和 DCE 的概念模型见图 1.6。

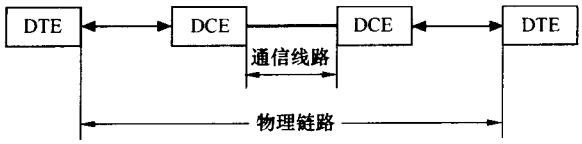


图 1.6　DTE/DCE 概念模型

物理层的标准和协议很多，下面列出常见和常用的几种：（1）EIA RS-232 标准。EIA RS-232 标准是美国电子工业协会颁布的，最初用于通过电话网实现两个 DTE 的通信，是 DTE 和调制解调器（相当于 DCE）之间的物理层接口标准。现在，几乎每台网络设备都配备了 RS-232 接口，EIA RS-232 已经成为事实上的串行通信接口标准。（2）ISO2110、ISO2593、ISO4902 和 ISO4903。ISO2110 和 ISO2593 分别是 25 芯和 34 芯 DTE/DCE 接口连接器和插针分配标准。ISO4902 和 ISO4903 分别是 37 芯和 15 芯 DTE/DEC 接口连接器和插针分配标准。（3）CCITT V 系列建议。该系列物理层建议主要为以电话网为基础的模拟通信设备而制定的，是 DTE 和调制解调器之间的接口。（4）CCITT X 系列建议。该系列物理层建议是为以公共数据网为基础的数字通信设备而制定的数字通信标准。（5）CCITT I 系列建议。该系列物理层协议是针对综合业务数字网 ISDN 而制定的。

2．数据链路层

OSI 参考模型的第二层是数据链路层，它位于网络层和物理层之间，对上它要为网络层提供服务，利用物理层提供的服务把网络层传来的数据包传走，并把物理层接收到的数据解封装后送给网络层。数据链路层能够加强物理层传输原始比特流的功能，使物理层提供的不可靠的物理链路对网络层呈现为逻辑上无差错的数据链路。

数据链路层必须完成数据链路层实体间二进制信息块（称为帧）的正确传输，其功能包

括组帧、帧同步、差错控制、流量控制、数据链路管理等。

（1）组帧。帧是数据链路层进行信息传递的基本单元，网络层传来的数据都要被封装成帧，然后才能在数据链路上传送。帧中往往不仅包含网络层的数据，还会在头尾加上数据链路层的控制信息。不同的数据链路层协议具有不同的帧格式，图 1.7 所示是高级数据链路控制协议 HDLC(High-level Data Link Control)的帧格式。有了帧结构后，数据传输发生错误时，只有错误的帧需要重新传送。

| F | A | C | I | FCS | F |
|---|---|---|---|---|---|
| 帧起始标志<br>01111110 | 目的站地址<br>8bit | 控制字段<br>8bit | 信息字段<br>*N* bit | 帧校验序列<br>16bit | 帧结束标志<br>01111110 |

图 1.7　HDLC 帧格式

（2）帧同步。帧同步是指帧结构的设计应当使得接收方能够明确地从物理层收到的比特流中识别出帧的开始与结束。图 1.7 中的帧起始标志和帧结束标志就是 HDLC 中的帧同步字段。

（3）差错控制。数据链路层的差错控制不同于网络层和传输层的差错控制，它只保证相邻节点间的传输差错控制在所允许的最小范围内。常见的差错控制方法有纠错法和检错重传法。

（4）流量控制。数据链路层只进行点到点的流量控制，在传输层还会提到流量控制，传输层进行的流量控制是端到端的流量控制。点到点和端到端的概念如图 1.8 所示。由于收发双方所使用的设备的工作速率和缓存空间的差异，可能会出现发送方的传送能力大于接收方的接收能力的情况，这时如果不对发送方的传送速率进行控制，接收方就会来不及处理而被“淹没”，即后面传来的数据帧被丢弃或者后面传来的数据帧覆盖先前收到但还没有来得及处理的帧。数据链路层的流量控制就是要在点到点避免这种情况的发生。流量控制可以通过硬件来实现，也可以通过软件来实现。

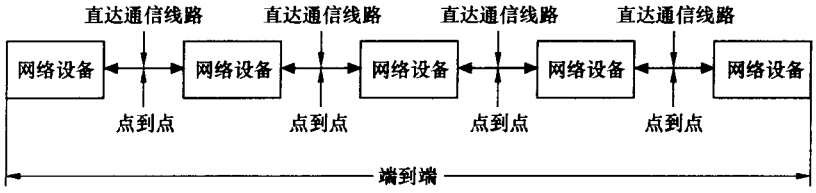


图 1.8　点到点与端到端

（5）数据链路管理。数据链路管理主要是指数据链路连接的建立、维持和释放，以及对共享物理信道的仲裁和合理分配等。

3．网络层

计算机网络可以分为通信子网和用户资源子网两部分，见图 1.9。网络层是通信子网的最高层，它提供协议使得系统之间可以通信，它把系统而不仅仅是网络接口连接到一起。正是在这一层，通信被认为发生在系统间而不只是在网络接口间。这一层需要考虑如何在位于两个不同网络的两个不同节点间传送数据，它向传输层提供一个透明的端到端（参见图 1.8）通路，使传输层能够完成端到端的透明数据传输。