

IPv6 技术

——新一代网络技术

王相林 编著



机械工业出版社
CHINA MACHINE PRESS

2008

王相林 编著

机械工业出版社

本书突出 IPv6 技术理论和应用,说明从 IPv4 向 IPv6 过渡中需要注意的问题,透彻分析 IPv6 技术的各个主题;注重开拓 IPv6 技术深层次的内容,指出解决 IPv6 技术问题的思路和途径。

本书共 9 章,主要内容包括:IPv6 研究的历程、IPv6 的制订依据、IPv6 技术新特性、IPv6 技术的推广和部署、IPv6 技术标准研究、IPv6 的结构、IPv6 与相邻层协议的关系、IPv6 的地址分类、IPv6 地址配置技术、ICMPv6、IPv6 邻居发现技术、IPv6 路由技术 RIPng、IPv6 的 OSPFv3、IPv6 的 BGP-4、IPv6 安全技术、IPv6 的安全要素、IPv6 中的加密、IPv6 中的认证、密钥交换协议、IPv6 过渡的技术、过渡需要采取的措施、IPv6 与底层网络技术、移动 IPv6 技术等。

本书适合计算机科学与技术领域的科研人员、研究生和高年级本科生,以及从事计算机网络、IPv6 网络技术和下一代网络研究和应用的 IT 专业人员阅读。对需要了解下一代因特网(NGI)和下一代网络(NGN)的核心技术 IPv6 的读者,本书也是有益的读物。

图书在版编目(CIP)数据

IPv6 技术:新一代网络技术/王相林编著. —北京:机械工业出版社, 2008.3

ISBN 978-7-111-23468-5

I. I… II. 王… III. 因特网-协议(计算机) IV. TN915.04

中国版本图书馆 CIP 数据核字(2008)第 018051 号

机械工业出版社(北京市百万庄大街 22 号 邮政编码 100037)

责任编辑:王保家 责任校对:李秋荣

封面设计:王伟光 责任印制:杨 曦

北京机工印刷厂印刷(兴文装订厂装订)

2008 年 3 月第 1 版第 1 次印刷

184mm×260mm·15.75 印张·384 千字

标准书号:ISBN 978-7-111-23468-5

定价:26.00 元

凡购本书,如有缺页、倒页、脱页,由本社发行部调换

销售服务热线电话:(010) 68326294

购书热线电话:(010) 88379639 88379641 88379643

编辑热线电话:(010) 88379711

封面无防伪标均为盗版

前言

计算机网络技术的掌握和应用水平,反映了一个国家的科学技术水平,而构成信息社会的基础设施的计算机网络,对国民经济的发展起到至关重要的作用,对下一代因特网(NGI)和下一代网络(NGN)中的核心技术IPv6的研究和了解,已经成为21世纪我国国家建设和发展的重要内容。NGI、NGN已经确定其核心技术为IPv6,IPv6技术的知识和学习成为计算机网络研究与应用的热点。国家发展与改革委员会在2003年已经立项,投入大量经费开展IPv6技术的研究。本书可以较好地配合IPv6技术的学习,满足对IPv6核心技术学习的基本要求。

本书涵盖了IPv6技术的主要内容,包括IPv6技术发展历史、IPv4与IPv6的比较、IPv6技术核心的内容和IPv6技术部署中的措施等。本书内容循序渐进、脉络清晰,清楚地讲解了IPv6核心技术是什么、为什么、怎样用,可以使对IPv4比较熟悉,又对IPv6有一定了解的读者看到IPv6技术的广阔性和多样性,比较容易学习和掌握从IPv4技术过渡到IPv6技术的知识。

本书共9章,主要内容包括:IPv6研究的历程、IPv6的制订依据、IPv6技术新特性、IPv6技术的推广和部署、IPv6技术标准研究、IPv6的结构、IPv6与相邻层协议的关系、IPv6的地址分类、IPv6地址配置技术、ICMPv6、IPv6邻居发现技术、IPv6路由技术RIPng、IPv6的OSPFv3、IPv6的BGP-4、IPv6安全技术、IPv6的安全要素、IPv6中的加密、IPv6中的认证、密钥交换协议、IPv6过渡的技术、过渡需要采取的措施、IPv6与底层网络技术、移动IPv6技术等。

本书突出IPv6技术的主要内容,说明了从IPv4向IPv6过渡中,在理论、应用和工程方面需要注意的问题,透彻地分析了IPv6技术的各个主题,注重开拓IPv6技术深层次的内容,指出了解决IPv6技术问题的思路和途径,使IT专业人员以及从事计算机网络研究和应用的工程技术人员,可以很容易地通过阅读本书,掌握和了解IPv6技术的主要内容。

本书是在IPv6技术研究和教学工作基础上完成的,讲述清晰易懂,除告诉读者IPv6应用中需注意的问题外,还结合网络应用,讲解了IPv6核心技术的基本理论、来龙去脉,以及需要注意和说明的问题,使读者能做到举一反三、触类旁通。

由于有关IPv6技术的书籍较少,而计算机网络技术的建设对IPv6知识的学习又是十分迫切的,因此,希望本书可以为下一代网络技术的研究和应用提供很好的学习途径和参考。

本书适合计算机科学与技术领域的科研人员、研究生和高年级本科生,以及从事计算机网络、IPv6网络技术和下一代网络(NGN)研究和应用的IT专业人员阅读。对需要了解下一代因特网(NGI)和下一代网络(NGN)的核心技术IPv6的读者,本书也是有益的读物。

吴阿明、王源参加了书稿修改工作;卢庆菲、李蓓蕾参加了书稿校对工作。

由于作者水平有限,书中难免存在不妥之处,恳请读者批评指正。

作者的联系方式:wangedu@163.com。

作者

目 录

前言

第1章 IPv6 概述 1

1.1 计算机网络体系结构 1

1.1.1 计算机网络体系结构概述 1

1.1.2 计算机网络协议 3

1.1.3 OSI/RM 参考模型 4

1.1.4 TCP/IP 模型 5

1.1.5 网络中的基本术语 5

1.2 IPv4 的局限性 6

1.2.1 IPv4 的基本知识 6

1.2.2 IPv4 存在的问题 8

1.3 IPv6 基本知识 10

1.3.1 IPv6 的目标 10

1.3.2 IPv6 研究的历程 10

1.3.3 IPv6 的制订 12

1.3.4 IPv6 报头的特点 12

1.3.5 IPv4 与 IPv6 的比较 13

1.4 IPv6 技术标准研究 13

1.4.1 与 IPv6 技术有关的国际标准组织 13

1.4.2 中国 IPv6 标准化工作的开展 16

1.4.3 IPv6 技术标准 18

1.4.4 支持 IPv6 技术的厂商 19

1.5 IPv6 技术的推广和部署 20

1.5.1 IPv6 在国外的推广和部署 20

1.5.2 IPv6 在国内的推广和部署 21

1.5.3 IPv6 是下一代网络的核心技术 22

1.5.4 IPv6 技术推广和部署时面临的问题 23

1.5.5 正在研究的 IPv6 关键技术 24

思考题和练习题 26

第2章 IPv6 结构 28

2.1 IPv6 分组结构 28

2.1.1 IPv6 基本术语 28

2.1.2 IPv6 协议数据单元 29

2.1.3 IPv6 首部与 IPv4 首部的比较 31

2.2 IPv6 的扩展首部 32

2.2.1 IPv6 扩展首部的基本知识 32

2.2.2 逐跳选项扩展首部 35

2.2.3 路由选择扩展首部 36

2.2.4 分段扩展首部 37

2.2.5 身份认证扩展首部 39

2.2.6 封装安全载荷扩展首部 40

2.2.7 目的站选项扩展首部 42

2.3 IPv6 与相邻层协议的关系 42

2.3.1 上层校验和计算 42

2.3.2 报文生成时间和最大上层协议载荷 42

2.4 IPv6 的特性 43

2.4.1 IPv6 具有层次化的地址结构 43

2.4.2 即插即用的联网方式 44

2.4.3 网络层的认证与加密 45

2.4.4 服务质量的支持 45

2.4.5 对移动通信更好的支持 46

思考题和练习题 46

第3章 IPv6 地址技术 47

3.1 IPv6 地址技术概述 47

3.1.1 IPv6 地址表示方式 47

3.1.2 IPv6 的地址空间和地址前缀 48

3.1.3 IPv6 寻址模型 49

3.2 IPv6 地址分类 50

3.2.1 IPv6 地址分类概述 50

3.2.2 IPv6 单播地址 52

3.2.3 IPv6 多播地址 55

3.2.4 IPv6 任播地址 57

3.3 IPv6 地址配置技术 58

3.3.1 地址的手工配置和检测 59

3.3.2 地址自动配置 59

3.3.3 DHCPv6 61

3.3.4 IPv6 域名系统 64

思考题和练习题 66

第4章 ICMPv6 及相关协议 67

4.1 ICMPv6 概述 67

4.1.1 ICMPv6 的功用 67

4.1.2 ICMPv6 与 ICMPv4 的比较	67	5.3.7 OSPFv3 链路状态请求报文	114
4.2 ICMPv6 格式	68	5.3.8 OSPFv3 链路状态更新报文	114
4.2.1 ICMPv6 报文的类型	68	5.3.9 OSPFv3 链路状态确认报文	115
4.2.2 ICMPv6 错误报文	71	5.3.10 OSPFv3 链路状态通告	116
4.2.3 ICMPv6 信息报文	74	5.3.11 OSPFv3 路由表的计算	118
4.2.4 ICMPv6 处理规则	75	5.3.12 OSPFv3 中的其他技术分析	122
4.3 邻居发现协议	75	5.4 IPv6 的 BGP-4	125
4.3.1 邻居发现协议概述	75	5.4.1 BGP 概述	125
4.3.2 邻居发现协议的功能	77	5.4.2 BGP 的连接建立和路由存储	126
4.3.3 路由器请求和路由器通告	77	5.4.3 BGP4 + 简介	128
4.3.4 邻居请求和邻居通告	79	5.4.4 BGP 报文首部	128
4.3.5 ICMP 重定向报文	80	5.4.5 OPEN 报文	129
4.3.6 邻居发现选项	81	5.4.6 UPDATE 报文	130
4.3.7 邻居缓存和目的地缓存	81	5.4.7 BGP 的属性	131
4.4 IPv6 地址解析技术	82	5.4.8 通知和生命期报文	132
4.4.1 主机的数据结构	82	5.4.9 IPv6 的 BGP 扩展	132
4.4.2 主机数据包的发送算法	83	思考题和练习题	134
4.4.3 邻居发现协议与 ARP 的比较	84	第 6 章 IPv6 安全技术	136
4.4.4 地址可达性检测	86	6.1 IPv6 安全问题	136
4.4.5 重定向技术	87	6.1.1 IPv6 安全问题概述	136
4.5 多播监听者发现协议	88	6.1.2 网络安全面临的威胁	137
4.5.1 多播监听者发现协议概述	88	6.1.3 基本的安全需求和技术	138
4.5.2 多播监听者发现报文格式	88	6.2 Internet 的安全技术	142
思考题和练习题	90	6.2.1 数据包过滤和防火墙	143
第 5 章 IPv6 路由技术	92	6.2.2 运输层保护	143
5.1 IPv6 路由协议概述	92	6.2.3 应用层安全	144
5.1.1 IPv6 路由协议基本知识	92	6.2.4 Internet 安全的开放性	145
5.1.2 IPv6 路由协议	92	6.3 IPv6 的安全要素	146
5.2 RIPng	94	6.3.1 IPsec 的功能	146
5.2.1 RIPng 概述	94	6.3.2 IPsec 框架	146
5.2.2 RIPng 路由更新的规则	95	6.3.3 IPsec 安全关联	147
5.2.3 RIPng 的局限性及解决方法	96	6.3.4 IPsec 安全策略	149
5.2.4 RIPng 报文格式	98	6.3.5 IPsec 部署	150
5.2.5 RIPng 下一跳字段和默认路由	100	6.3.6 IPsec 存在的问题	152
5.2.6 RIPng 的工作原理	100	6.4 IPv6 中的认证	152
5.3 IPv6 的 OSPFv3	103	6.4.1 认证的内容和方法	152
5.3.1 IPv6 OSPFv3 概述	103	6.4.2 IPv6 认证	154
5.3.2 OSPFv3 涉及的技术	104	6.5 IPv6 中的加密	155
5.3.3 IPv6 的 OSPF 和 IPv4 的 OSPF 的 比较	108	6.5.1 IPv6 中的加密概述	155
5.3.4 IPv6 的 OSPFv3 报文格式	109	6.5.2 IPv6 中的加密模式	156
5.3.5 OSPFv3 的 Hello 报文	110	6.6 密钥交换协议	157
5.3.6 OSPFv3 数据库描述报文	112	6.6.1 密钥交换协议概述	157
		6.6.2 ISAKMP	158

6.6.3 密钥交换协议 OAKLEY 和 SKEME	159
6.6.4 Internet 密钥交换	160
6.6.5 IPsec 与其他技术的联系	161
思考题和练习题	162
第 7 章 IPv6 过渡技术	164
7.1 IPv6 过渡技术概述	164
7.1.1 IPv6 过渡期的特点	164
7.1.2 过渡需要采取的措施	165
7.1.3 过渡面临的问题	165
7.1.4 过渡时期采用技术的选择	167
7.2 双栈技术	167
7.2.1 双栈技术工作原理	167
7.2.2 基本双栈和有限双栈技术	168
7.2.3 双栈机制 DSTM	169
7.3 隧道技术	170
7.3.1 隧道技术概述	170
7.3.2 隧道技术工作原理	171
7.3.3 手工隧道	172
7.3.4 基本的自动隧道技术	173
7.3.5 隧道 IPv6	174
7.3.6 6to4 机制	176
7.3.7 ISATAP	178
7.4 协议转换技术	178
7.4.1 NAT	179
7.4.2 NAT	179
7.4.3 NAT-PT	180
7.4.4 无状态 IP/ICMP 转换	181
7.4.5 IPv4 转换至 IPv6	182
7.4.6 ICMPv4 转换至 ICMPv6	183
7.4.7 IPv6 转换至 IPv4	184
7.4.8 ICMPv6 转换至 ICMPv4	185
7.5 过渡技术分析比较	186
7.5.1 几种转换技术的比较	186
7.5.2 如何选择合适的过渡机制	187
思考题和练习题	188
第 8 章 IPv6 与底层网络技术	189
8.1 IPv6 对底层网络的支持	189
8.1.1 IPv6 技术与网络中的第 2 层	189
8.1.2 IPv6 与 Ethernet	189
8.1.3 用于点对点协议的 IPv6CP	191
8.1.4 IPv6 与 ATM	192
8.1.5 IPv6 与帧中继	193
8.2 IPv6 与多播	194
8.2.1 多播的选路方法	194
8.2.2 IPv6 中多播实现的机制	195
8.3 简单 IPv6 网络设计思路	195
8.3.1 底层网络之上的 IPv6 网络构成	195
8.3.2 简单 IPv6 网络设计例子	196
思考题和练习题	197
第 9 章 移动 IPv6 技术	198
9.1 移动 IPv6 概述	198
9.1.1 移动 IP 技术的基本概念	198
9.1.2 移动 IP 技术涉及的术语	199
9.1.3 移动 IP 技术的发展历程	200
9.1.4 移动 IPv6 与移动 IPv4 的比较	201
9.1.5 移动 IP 的工作过程	203
9.2 移动 IPv6 的组成和特征	204
9.2.1 移动 IPv6 的技术要求	204
9.2.2 移动节点具有的 3 种功能	204
9.2.3 移动 IPv6 的组成	205
9.2.4 移动节点和对端节点之间的通信模式	205
9.2.5 移动 IPv6 增加的新协议及内容	206
9.3 移动 IPv6 报文和选项格式	207
9.3.1 移动 IPv6 首部格式	207
9.3.2 绑定更新请求报文	207
9.3.3 家乡测试初始报文	208
9.3.4 转交测试初始报文	208
9.3.5 家乡测试报文	209
9.3.6 转交测试报文	209
9.3.7 绑定更新报文	210
9.3.8 绑定确认报文	211
9.3.9 绑定错误报文	212
9.4 移动选项	213
9.4.1 移动选项格式	213
9.4.2 Pad1 和 Pad N	213
9.4.3 绑定更新建议选项	214
9.4.4 备用转交地址选项	214
9.4.5 随机数索引选项	214
9.4.6 绑定授权数据选项	215
9.5 家乡地址选项和第 2 类路由首部	215
9.5.1 家乡地址选项	215
9.5.2 第 2 类路由首部	216

9.6 移动 IPv6 对 ICMPv6 的扩展	217	9.7.4 移动 IPv6 性能分析	222
9.6.1 ICMP 家乡代理地址发现请求报 文	217	9.7.5 移动 IPv6 的安全特性	223
9.6.2 ICMP 家乡代理地址发现应答报 文	217	9.7.6 移动 IPv6 的服务质量支持	223
9.6.3 ICMP 移动前缀请求报文	218	思考题与练习题	224
9.6.4 ICMP 移动前缀通告报文	219	附录	226
9.7 移动 IPv6 技术分析	219	附录 A 书中英文缩写词	226
9.7.1 移动 IPv6 工作原理	219	附录 B RFC 文档	230
9.7.2 移动 IPv6 机制实现的例子	220	附录 C IPv6 参数	235
9.7.3 移动 IPv6 快速切换	221	附录 D 常用网址	240
		参考文献	241

第 1 章 IPv6 概述

1.1 计算机网络体系结构

计算机网络是计算机技术和通信技术相结合的技术，计算机网络技术还涉及到微电子技术、光通信技术和智能控制技术。计算机网络是通过通信协议和传输介质把分散在不同地点的计算机设备连接起来，实现资源共享和数据传输的系统。计算机网络是一个复杂的系统，对计算机网络系统的设计采用分层的方法，把计算机网络的功能分散到每一层，可以简化复杂问题的处理，也便于对计算机网络系统的维护和扩充。

与人类日常生活中的通信需要使用协议一样，计算机网络采用网络协议进行通信，协议是通信双方彼此遵循的规则和约定。计算机网络的设计采用开放的体系结构，所谓开放，是指不同厂商生产的计算机网络硬件、软件设备采用大家遵循的计算机网络体系结构参考模型，使所生产的设备可以方便地实现互连。计算机网络采用分组交换，把需要传输的数据分割成小数据段，加上网络协议首部，形成协议数据单元（PDU，Protocol Data Unit）在网络中传输，这些协议首部是通信传输的规则和约定，是冗余的，但是不可缺少的。

计算机网络中两台计算机之间的通信，类似于人类社会中的两个人之间的讲话，一次成功的通信一般可以分为 3 个阶段：连接建立阶段、数据传输阶段、连接释放阶段。计算机网络中通信的计算机首先要找到对方，要实现这一点，就需要设计和建立网络中计算机的连接标识，也就是说要建立计算机网络中的寻址机制。Internet 中采用 IP 地址标识网络中一个节点接口的连接。计算机网络中数据分组在经过路由器时，路由器根据分组中目的 IP 地址在路由表中查找出适合该分组的转发路径。

1.1.1 计算机网络体系结构概述

计算机网络体系结构是计算机网络中的分层和协议的集合。计算机网络体系结构描述了计算机网络设计时应该遵循的层次功能划分、每一层协议的标识和格式，并涉及到层与层之间的联系、层与层之间接口的实现方法，以及层与层之间服务的关系和对等层的概念。

计算机网络中的层次应划分为多少层为好呢？层次数不能太多，也不能太少。层数太多会使网络体系结构过于复杂，层之间的接口过多，给网络设计带来困难。层数过少会造成网络功能的界面不清楚，每一层实现的内容太多，给层功能实现带来不便。早期的计算机网络体系结构层次一般多为 7 层或 8 层，例如 IBM 公司的系统网络体系结构（SNA，System Network Architecture）和美国数字设备公司（DEC，Digital Equipment Corporation）的网络体系结构 DecNET，以及国际标准化组织（ISO，International Standardization Organization）给出的开放系统互连参考模型（OSI，Open System Interconnection）。

伴随着计算机网络技术的发展，计算机网络体系结构的描述也发生了变化，目前，计算机网络体系结构的层次划分一般采用 5 层结构，自顶向下依次为：应用层；运输层；网络层；数据链路层、物理层。5 层计算机网络体系结构如图 1-1 所示。

计算机网络中的地址分为逻辑地址和物理地址，物理地址为硬件地址，逻辑地址为软件地址。物理地址有网卡地址，有时也称为 MAC 地址。逻辑地址有前面提到的 IP 地址，主要用来标识计算机网络中一个节点中网络接口的连接，逻辑地址还有端口地址，用来标识应用进程。域名地址也是逻辑地址，与 IP 地址对应，主要是为了便于记忆和使用，在应用时通过域名解析系统

转换为对应的 IP 地址。要找到计算机网络中的某一台主机，最后必须通过物理地址。之所以采用逻辑地址，是为了进行网络中计算机设备的连接、实现网络互连，简化计算机网络中的寻址处理，以及区分和标识不同的应用。计算机网络中的地址，以及地址与层次的关系如图 1-2 所示。从图中可以看出，数据链路层及其以下层的地址为硬件地址，网络层及其以上层的地址为软件地址。

计算机网络中的体系结构有时也称为“洋葱头”式的体系结构，在发送方，需要传输的应用数据，沿着自顶向下依次经过应用层、运输层、网络层、数据链路层和物理层，在经过每一层时，都要加上该层的协议首部信息，构成每一层的协议数据单元 PDU。在数据链路层不仅要加上协议首部，还要加上协议尾，组成数据链路层的协议数据单元帧。这一过程叫做封装，也称为打包和协议封装。网络中传输的协议数据单元，在经过网络中的路由器和交换机这些中间节点时，需要根据中间节点的层次结构，进行拆封和再封装。到达接收方后，自底向上依次经过物理层、数据链路层、网络层、运输层和应用层进行拆封（拆封也称为拆包），最后，把应用数据交给应用进程。

需要说明的是，这里说的协议首部就是每一层要实现的网络控制协议，也是每一层协议数据单元中的控制部分，每一层通过协议实现该层的功能。有时协议首部也称为协议报头。不同系统中的同一层称为对等层，在打包和拆包过程中实现了对等层之间的通信，用虚线标识，好像是两个对等层之间在通信一样，而实际的信号通路在发送端系统是垂直向下、经过网络传输介质连接到通信的另一端。对等协议的打包、拆包通信过程如图 1-3 所示。

人们会问，为什么要加各层的协议信息？这些协议信息是为了实现应用数据的可靠传输，分别解决计算机网络中的数据在不同网络部位传输时需要处理的问题。当然，这些协议信息是冗余信息，但又是必不可少的。

计算机网络体系结构和组成，也分为资源子网和通信子网两个部分，又称为网络边缘和网络核心。资源子网负责计算机网络中数据的发送和接收，是数据源端点和数据目的端点。资源子网的设备包括上网的计算机设备，例如计算机、网络打印机、手持移动设备等，用作计算机网络中的端节点。通信子网负责数据的传输、交换和连接，以及通信控制，例如路由

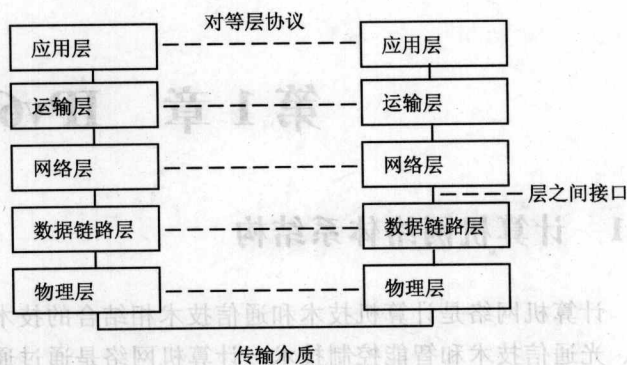


图 1-1 5层计算机网络的协议体系结构

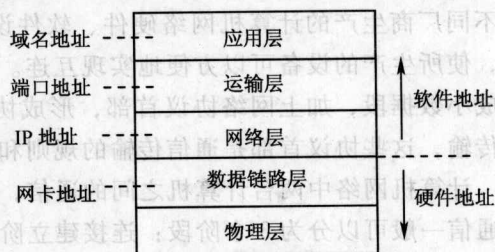


图 1-2 计算机网络中的地址及与层次的对应

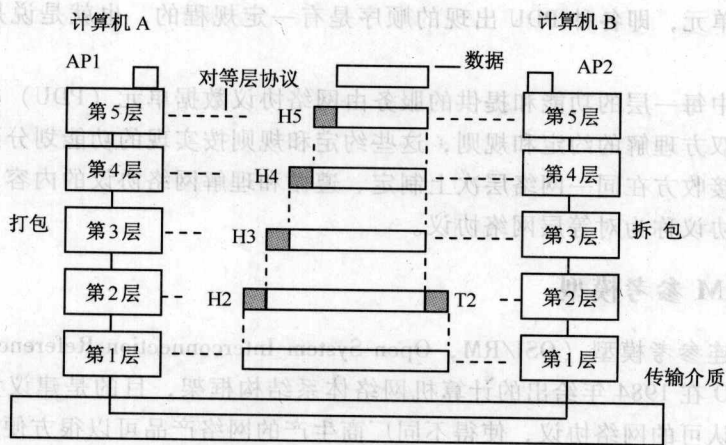


图 1-3 对等协议的打包、拆包通信过程

器、交换机等，用作计算机网络中的交换节点和访问节点。计算机网络中两级子网的概念如图 1.4 所示。

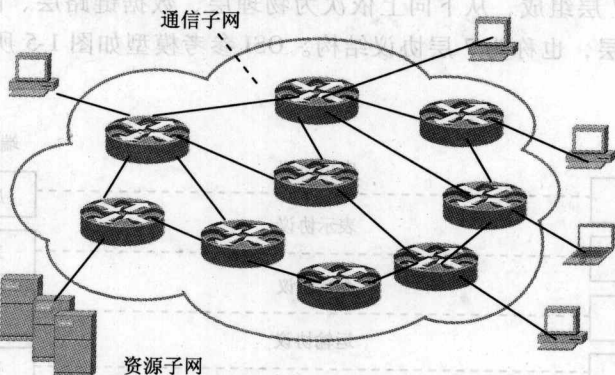


图 1.4 计算机网络中两级子网的概念

1.1.2 计算机网络协议

计算机网络协议是计算机网络中通信设备之间遵循的规则和约定。协议的 3 个基本要素是：

- (1) 语法 控制信息和数据信息的格式和结构。
- (2) 语义，用若干二进制位表示的具体含义，例如是何种控制信息、给出何种动作或响应。
- (3) 同步，通信过程中各种动作的先后顺序，有预先规定好的通信步骤。同步是可以设计的。

计算机和计算机网络中传输的是二进制位 (bit) 流。计算机网络中计算机设备之间通信采用的是二进制语言，这些二进制位流携带着控制信息和数据信息，用计算机网络协议将这些二进制位流构成协议数据单元 (PDU)。PDU 有一定的语法格式。若干二进制位组成一个字段，字段值有确定的语义，代表通信双方彼此可以理解的含义。同步用来协调哪些协议数据单元先传输，哪些协议数据单元后传输，以及在一些协议数据单元传输后，下面应该传

输哪些协议数据单元,即各种 PDU 出现的顺序是有一定规程的,也就是说是可预知的,这就是同步。

计算机网络中每一层的功能和提供的服务由网络协议数据单元(PDU)描述和实现。网络协议涉及通信双方理解的约定和规则,这些约定和规则按实现的功能划分到不同的层次上实现。发送方和接收方在同一网络层次上制定、遵循和理解网络协议的内容。不同系统中同一层次上的网络协议称为对等层网络协议。

1.1.3 OSI/RM 参考模型

开放系统互连参考模型(OSI/RM, Open System Interconnection Reference Model),是国际标准化组织 ISO 在 1984 年给出的计算机网络体系结构框架,目的是建议大家遵循一致的网络层次结构和认可的网络协议,使得不同厂商生产的网络产品可以很方便地互连,实现彼此开放。OSI/RM 的国际标准编号为 ISO 7498。OSI 主要针对计算机网络的发展迫切需要解决多个厂商的网络产品不能互连的问题。与开放对应的是封闭,封闭系统指的是一个厂商的网络产品只能在自己的产品之间互连,而不能与其他厂商的产品互连和兼容。

OSI 参考模型由 7 层组成,从下向上依次为物理层、数据链路层、网络层、运输层、会话层、表示层和应用层,也称为 7 层协议结构。OSI 参考模型如图 1-5 所示。

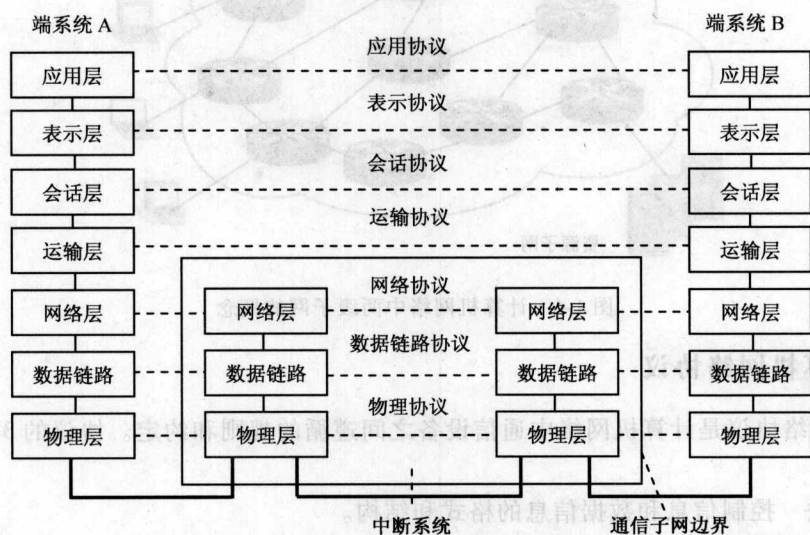


图 1-5 OSI 参考模型

这里需要说明的是,OSI 给出的是一个计算机网络体系结构框架,提供在设计计算机网络时参考的模型,以及研制计算机网络硬件和软件时遵循的理念。一个具体的计算机网络实现过程,并不要求完全按照 OSI/RM 给出的层次协议结构来设计。可以类比房屋建筑的系统结构来理解 OSI 的作用。例如,明代风格的建筑,是从房屋的外形和内在结构讲的,是讲这些房屋具有明代房屋的特征和外貌,而一所具体的明代房屋则要涉及到房屋的具体设计、房屋的位置等属性。

1.1.4 TCP/IP 模型

TCP/IP 是 Internet 采用的网络协议, TCP/IP 是 Internet 的语言。TCP/IP 是事实上的工业标准, 由于种种原因, 国际标准化组织 ISO 并没有把 TCP/IP 纳为国际标准。TCP/IP 的研究是在 1974 年开始的, 1983 年 TCP/IP 用于 ARPANET 网络。1990 年 ARPANET 网络退出使用, 演变为今天的 Internet 因特网。

TCP/IP 是一个协议簇的标识, TCP/IP 的层次结构如图 1-6 所示。

在 TCP/IP 中没有定义相应的物理层和数据链路层, 而是结合这两层的特性给出一个网络接口层, 从技术的层面上讲, 这是 TCP/IP 很重要的设计思想。TCP/IP 不规定物理层和数据链路层的内容, 只要能够把 IP 分组作为数据封装到这些底层网络的帧中传输, 就可以与所有类型的

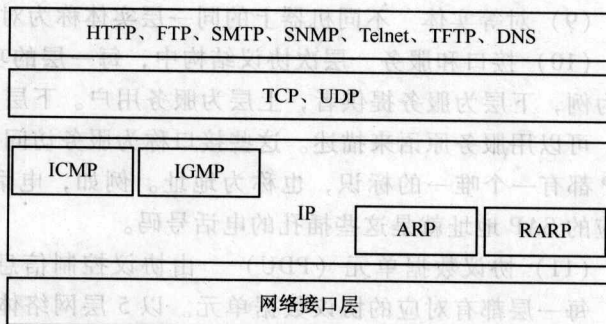


图 1-6 TCP/IP 层次结构

通信子网进行网络协议捆绑, 这些底层网络可以是局域网、广域网、无线网等。这种设计思想真正实现了不同网络的互连, 是 TCP/IP 得以长期应用和发展的核心技术。

TCP/IP 中最重要的两个协议是 IP 和 TCP。IP 实现计算机网络的互连和网络中计算机的寻址, 提供无连接的数据报服务, 是尽力交付的服务。目前使用较多的 IP 协议的版本号为 4, 也称为 IPv4。现在正在研究和推广使用的网络协议是 IPv6。TCP 是面向连接的网络协议, 实现端节点到端节点的可靠数据传输, 是一个复杂的运输层协议, 以弥补 IP 网络服务存在的缺陷, 为应用层进程提供可靠的运输服务。

1.1.5 网络中的基本术语

计算机网络是综合的技术, 涉及许多理论概念和技术知识, 对计算机网络中基本术语进行规范和了解对学习计算机网络很重要。

(1) 节点 一般指网络中的主机 (计算机) 或路由器, 又分为访问节点、交换节点和端节点, 前两者的例子是路由器, 后者的例子是主机。端节点用作信源和信宿, 访问节点直接连接信源和信宿, 交换节点与其他节点连接, 主要用于传输、交换和通信控制。节点能够发送、接收和转发网络中的协议数据单元 (PDU)。

(2) 信源 (源) 发送数据的计算机。

(3) 信宿 (目的) 接收数据的计算机。

(4) 传输介质 提供网络和计算机之间传输的信道, 即提供信号传输的通路。传输介质分为有线传输介质和无线传输介质。常用的有线传输介质有双绞线和光缆, 常用的无线传输介质有红外线、微波、无线电和激光。无线传输介质是利用电磁波频谱中一些波段的特性进行信号的传输。

(5) 信号的传输方向 以 A 与 B 之间传输为例, 可以分为 3 种情况: 单向传输, 即 A 只能传给 B; 双向交替传输, 在某一时刻 A 可以传给 B, 在另一个时刻, B 可以传给 A; 双

向同时传输,在某一时刻,A可以传给B,B也可以传给A。对应上述3种情况,以前术语分别为单工、半双工、全双工通信。

(6) 面向连接 通信的双方在通信时,需要建立连接。

(7) 无连接 通信的双方在通信时,不需要建立连接。

(8) 网络实体(Entity) 实体是网络协议层次中的活动元素,实体可以是软件实体或硬件实体,例如一个进程,或一个芯片。

(9) 对等实体 不同机器上的同一层实体称为对等实体(Peer Entity)。

(10) 接口和服务 层次协议结构中,每一层的功能是为它的上层提供服务。以两个层次为例,下层为服务提供者,上层为服务用户。下层为上层提供的服务通过层之间的接口进行,可以用服务原语来描述。这些接口称为服务访问点(SAP, Service Access Point),每个SAP都有一个唯一的标识,也称为地址。例如,电话系统中的SAP可以是标准电话插孔,对应的SAP地址就是这些插孔的电话号码。

(11) 协议数据单元(PDU) 由协议控制信息和数据信息组成,与网络中的层次联系,每一层都有对应的协议数据单元。以5层网络体系结构为例,自顶向下各层次PDU的名称:应用层是报文(Message);运输层是报文段(Segment);网络层为分组(Packet);数据链路层为帧(Frame);物理层为位流(Bits)。PDU用于对等实体之间执行它们的同等层协议。

(12) 服务质量(QoS, Quality of Service) 用于评价每种服务的特性,计算机网络中的服务质量涉及信道的容量、延时、带宽、数据丢失等参数。

(13) 服务与协议的关系 服务定义了两层之间的接口,协议定义了同等层实体之间的协议数据单元(PDU)。实体利用协议来实现它们定义的服务,只要提供的服务可以满足,协议的格式可以不予限制。网络服务需要靠网络协议来实现。

1.2 IPv4 的局限性

1.2.1 IPv4 的基本知识

IPv4是在1974年开始研制的,最初用于ARPANET,目标是在网络硬件受到损坏后,尽量减少对整个网络的影响,把网络中复杂的可靠性问题留到网络边缘解决。IPv4实现了提供尽力交付的服务,采用IP地址这一逻辑地址实现了网络的互连,以及网络中计算机设备网络接口的连接标识,为不同网络和网络中计算机设备的互连起到重要作用。IPv4的格式如图1-7所示。

在TCP/IP中,各种数据格式常用4个字节、32位为单位描述。现在的Internet是基于IPv4的,IPv4的最大的特点是简单易用。

IPv4格式分为固定部分和可变部分。固定部分有20个字节,含有12个字段,其中有3个字段用于分组的分片处理。可变部分含有可选字段,长度是可以变化的,有时需要填充位,符合以32位为一个单位的描述要求。

IP数据报固定首部包含有12个字段,下面介绍每个字段的含义。

1) 版本VER,占4位,记录IP的版本。这里使用的IP版本号为4,即IPv4。通信双方

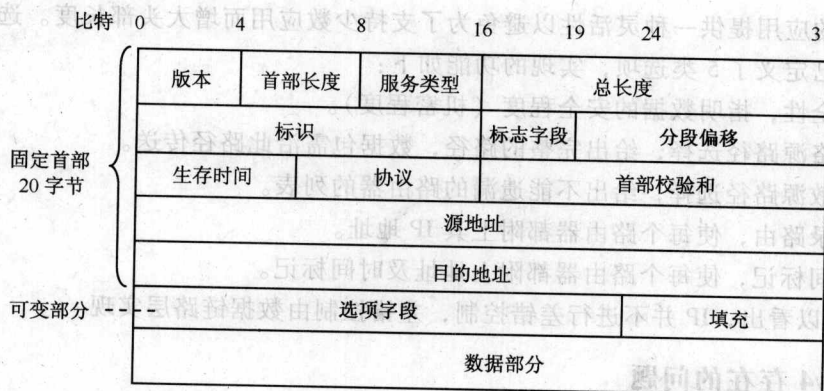


图 1-7 IPv4 格式

使用的 IP 版本要一致，该字段的值为 0100。

2) 首部（报头）长度 IHL，占 4 位，表示 IP 数据报首部的长度（以组为单位，一个单位 4 个字节），取值为 5~15，默认值为 5，即首部固定部分为 20 字节。IP 数据报首部中包括了选项部分，选项占据的长度由该字段指明。

3) 服务类型 TOS，占 8 位，其格式为 PPPDTRC0，其中 PPP 三个位特定义了 8 个优先级，D 位为延迟，T 位为吞吐量，R 位为可靠性，C 位是新增的，表示选择代价更小的路由，最后一位未使用。优先级、延迟、吞吐量、可靠性、更小代价路由可任意组合。主机通过此字段告诉子网它所要求的服务。该字段早期很少使用，1998 年以后随着多媒体信息传输需求的增加，对该字段开始引起重视。

4) 总长度，占 16 位，指包含首部和数据部分以字节为单位的总长度。最大 65535 字节。

5) 标识，占 16 位，是一个计数器，产生数据报的编号。当把一个数据报分成多个分段（IP 数据报）时，属于同一个数据报的多个 IP 分段的该字段值相同。

6) 标志字段，占 3 位，其中，DF 位表示数据报是否可分段，该位为 1 时表示不能分段，因为目的端可能不能重装配分段，为 0 时表示可以分段；MF 位表示是否为最后一个段，为 0 时表示是最后分段，为 1 时表示后面还有分段；当有 N 个段时，前 N-1 个数据分段的 M 位都为 1。

7) 分段偏移，占 13 位，指出本分段数据相对于其所属数据报起点的偏移量，以 8 字节为单位，即计算时偏移值应乘以 8。每个分段的长度是 8 字节（64 位）的整数倍。

8) 生存时间，占 8 位，用来限制 IP 数据报在网络中存在的时间，以秒（s）为单位，每经过一个节点都递减 1，在等待时可加倍递减。当该字段值减为 0 时就将其丢弃，以防止数据报在网中无限制地传输。

9) 协议，占 8 位，该字段指明 IP 数据报应传送给哪个传输进程，用编号表示，如 TCP 为 6，UDP 为 17 等，具体编号在 RFC 1700 中定义。

10) 首部校验和，IP 数据报首部的校验和。该字段在每个节点都必须重新计算，因为生命周期、标志、片偏移等字段的值经过每个节点时都发生了变化。

11) 源地址和目的地址，各占 32 位，分别表示源节点、目的节点地址。

12) 选项字段，为可选内容，最多 10 组共 40 字节，为后续版本提供新的功能而预留，

或者为特定的应用提供一种灵活性以避免为了支持少数应用而增大头部长度的。选项字段很少使用。目前已定义了 5 类选项,实现的功能如下:

- ① 安全性,指明数据的安全程度(机密程度)。
- ② 严格源路径选择,给出完整的路径,数据包需沿此路径传送。
- ③ 松散源路径选择,给出不能遗漏的路由器的列表。
- ④ 记录路由,使每个路由器都附上其 IP 地址。
- ⑤ 时间标记,使每个路由器都附上地址及时间标记。

从 IP 可以看出,IP 并不进行差错控制,差错控制由数据链路层实现。

1.2.2 IPv4 存在的问题

Internet 经历了快速膨胀的发展。由于历史的原因,IPv4 在设计时存在局限性。该协议的最初目标是用在军用网络中,认为使用网络的人都是可靠的,没有考虑到网络安全问题。当初只考虑在网络中传输的是一些计算数据和文本信息,没有考虑对多种服务的支持,不提供服务质量(QoS)保证。设计时用 32 位(bit)标识 IP 地址空间,认为这是一个巨大的数字,对网络地址的连接标识是足够用了。

TCP/IP 的工程师和设计人员早在 20 世纪 80 年代初期就意识到了 IPv4 升级的需求,因为当时已经发现 IP 地址空间随着 Internet 的发展只能支持很短的时间。对于 IPv4 节点的配置一直比较复杂,而网络管理员与用户则更喜欢“即插即用”,将计算机插在网络上然后就可以开始使用。IPv4 主机移动性的增强也要求当主机在不同网络间移动和使用不同的网络接入点时能提供更好的配置支持。

IPv4 存在的主要问题有地址空间匮乏、存在网络安全隐患、不提供服务质量保证、IP 地址配置复杂、缺少移动性支持等。IPv4 必须升级的原因以及可以同时改进之处包括:

(1) 地址空间的局限性 IPv4 地址为 32 位长,经常以 4 个两位十六进制数字表示,也常常以 4 个 0~255 间的十进制数字表示,数字之间用小数点间隔。每个 IP 主机地址包括两部分:①网络地址,用于指出该主机属于哪一个网络(属于同一个网络的主机使用同样的网络地址);②主机地址,它唯一地定义了某一个网络中的主机。

由于 IPv4 的地址空间可能具有多于 40 亿的地址编码,有人可能会认为 Internet 很容易容纳数以亿计的主机。但是这只适用于 IP 地址以顺序化分布的情况,即第一台主机的地址为 1,第二台主机的地址为 2,依此类推。而 IPv4 地址采用分类的层次结构划分地址,造成地址空间浪费严重。

(2) 增强安全性 长期以来,人们认为安全问题在网络协议的底层并不重要,都是把网络安全问题交给高层处理,例如安全套接字层(SSL, Security Socket Layer)和安全超文本传输协议(SHTTP, Security Hyper Text Transfer Protocol),就是分别在运输层和应用层增强网络的安全性。这些技术均不能从根本上解决网络安全问题,IP 数据仍然会泄露网络应用进程信息。

(3) 支持自动配置 IPv4 网络的节点配置比较复杂,在将计算机接入网络时一般需要专业人员的指导和帮助,需要设置 IP 地址、子网掩码、网关地址、域名系统(DNS, Domain Name System)地址,或者使用动态主机配置协议(DHCP, Dynamic Host Configuration Protocol),以及路由的配置等。人们希望使用计算机网络应像使用移动电话那样方便,只要

拨打对方的电话号码就可以通话。在计算机网络中要求网络设备、终端设备可以即插即用,实现设备的自动识别和自动配置。尤其是移动设备应用的普及,移动设备在不同的网络之间移动和使用不同的接入点时,需能够提供自动配置支持。

(4) 提供服务质量支持 随着 Internet 应用的普及和发展,多媒体业务的需求成倍增加,音频信息和视频信息逐渐成为网络中主要的传输数据。IPv4 网络提供尽力交付的服务,但对计算机网络中涌现的新型业务缺乏有效的支持,不提供服务质量保证,例如带宽、时延、误码率和抖动等。尽管已经提出了资源预留协议 (RSVP, Resource reSerVation Protocol)、综合服务 InteServ、区分服务 DiffServ、支持实时传输的实时传输协议 (RTP, Real Transfer Protocol) 和实时传输通信协议 (RTCP, Real Transfer Communication Protocol),但是这些附加的协议又增加了构建网络的复杂性和成本。

(5) 支持移动性 移动 IPv4 中必须有外地代理,需要使用更多的 IP 地址。移动 IPv4 中存在三角路由问题,移动节点离开本地网络后,需要通过本地(家乡)代理转发发给该移动节点的数据,会对本地网络和家乡代理带来很大影响,并有可能引起单点故障。移动 IPv4 还存在入境过滤问题,不支持节点快速移动。

(6) IP 路由问题 在 Internet 或内联网上传输的 IPv4 包必须从一个网络选路到另一个网络以到达其目的地。选路协议可以使用动态机制来确定路由,但是所有选路最终依赖于某个路由器查看不同路由的列表并确定正确的路由。选路表包含网络的列表和连接到这些网络的接口的列表。路由器查看 IP 包,确定 IP 包所在的网络(或该网络可能在的网络),然后把包发送到适当的网络接口。

现在的關鍵問題在於路由表的長度將隨著網絡數量的增加而變長。而路由表越長,路由器在表中查詢正確路由的時間就越長。如果只需要了解 10 個、100 個或 1000 個網絡,這不是問題,但是現在的 Internet 擁有大量的網絡,在骨干路由器上通常攜帶超過 11 萬個不同網絡地址的顯式路由,此時查詢路由的延時就會增加許多。

查詢路由的延時影響到網絡的性能,這種影響遠比地址空間的匱乏更緊迫。必須使用分級地址尋址來匯聚和簡化選路,否則 Internet 的性能可能在最近甚至現在就變得不可接受。

(7) 網絡地址翻譯 (NAT) 隨著專用 IP 網絡的發展,為避免過快減少可分配的 IP 地址,有一組 IP 地址被拿出來用於內部 IP 網絡。任何一個內部 IP 網絡均可以使用包括一個 A 類地址 (10.0.0.0)、16 個 B 類地址 (從 172.16.0.0 到 172.31.0.0) 和 256 個 C 類地址 (從 192.168.0.0 到 192.168.255.0) 在內的任何地址。內部網絡 IP 地址在 RFC1918 中定義,把這些內部網絡連接到公用網絡的路由器不轉發該內部 IP 網絡上的任何數據。

網絡地址翻譯 (NAT) 在內部網絡和公用網絡之間的接口實現,該系統 (一般是防火牆或路由器) 了解內部網絡上所有主機的地址,並將其翻譯為可訪問的公用網絡地址,這樣所有的內部主機就可以與外部主機通信。

NAT 為一些小型機構提供了一種自己管理其地址空間的簡單方法,無需依賴於地址授權機構為它們現在及將來的需要來分配足夠的地址空間。NAT 還使得一些機構可以非常快速和靈活地定義臨時地址或真正的內部網絡地址。與 CIDR 不同, NAT 確實提供了一種可以真正減少 IP 地址需求的辦法,儘管它使用起來有很大隨意性,並且在重新對內部 IP 網絡編址時將花費較長的時間和昂貴的代價。

雖然 NAT 辦法對於提高 IP 地址的分配效率有所幫助,但是網絡設計人員在決定一個網