

排版厂	京润天源		
天：26 地：18 订：20 切：20			
印 色	单黑		
字数行数	38*35	版 心	196*146
文前页码	16	正文页码	500
校 次	2	提校日期	12.30
书 号	72149		

计算机类专业系统能力培养系列教材

计算机网络

黄传河 杜瑞颖 张 健 张沪寅 牛晓光 编著



机械工业出版社
China Machine Press

图书在版编目 (CIP) 数据

计算机网络 / 黄传河等编著. —北京: 机械工业出版社, 2022.11

计算机类专业系统能力培养系列教材

ISBN 978-7-111-72149-9

I. ①计… II. ①黄… III. ①计算机网络 - 高等学校 - 教材 IV. ① TP393

中国版本图书馆 CIP 数据核字 (2022) 第 229990 号

本书从思维能力、创新能力、系统能力培养出发, 以建构型、设计型、创新型思维引导学生从无到有设计计算机网络。本书首先从需求出发, 明确计算机网络应满足的要求和约束条件以及可用的设计方法, 然后将计算机连接成小型网络, 进而将小型网络互联成大型网络, 实现任意计算机及应用程序之间数据的可靠传输; 之后设计计算机网络服务和应用, 并设计服务质量保证机制、网络安全机制和网络自动管理机制; 最后讨论面向未来的新型网络技术。

本书适合作为研究型大学计算机相关专业计算机网络课程的教材, 也适合计算机网络领域的技术人员参考。

出版发行: 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: 曲 熠

印 刷:

开 本: 186mm×240mm 1/16

书 号: ISBN 978-7-111-72149-9

责任校对: 张亚楠

版 次: 2022 年 月第 1 版第 1 次印刷

印 张:

定 价: 元

客服电话: (010) 88361066 68326294

版权所有·侵权必究

封底无防伪标均为盗版

编委会名单

- 荣誉主任：吴建平
- 主 任：周兴社

王志英

武永卫
- 副 主 任：马殿富

陈 钟

古天龙

温莉芳
- 委 员：金 海

李宣东

庄越挺

臧斌宇

吴功宜
- 陈文光

袁春风

安 虹

包云岗

章 毅
- 毛新军

姚 新

陈云霄

陈向群

向 勇
- 陈莉君

孟小峰

于 戈

张 昱

王宏志
- 汤 庸

朱 敏

卢 鹏

明 仲

王晓阳
- 单 征

陈卫卫

F O R E W O R D

丛书序言

人工智能、大数据、云计算、物联网、移动互联网以及区块链等新一代信息技术及其融合发展是当代智能科技的主要体现，并形成智能时代在当前以及未来一个时期的鲜明技术特征。智能时代来临之际，面对全球范围内以智能科技为代表的新技术革命，高等教育也处于重要的变革时期。目前，全世界高等教育的改革正呈现出结构的多样化、课程内容的综合化、教育模式的产学研一体化、教育协作的国际化以及教育的终身化等趋势。在这些背景下，计算机专业教育面临着重要的挑战与变化，以新型计算技术为核心并快速发展的智能科技正在引发我国计算机专业教育的变革。

计算机专业教育既要凝练计算技术发展中的“不变要素”，也要更好地体现时代变化引发的教育内容的更新；既要突出计算机科学与技术专业的核心地位与基础作用，也需兼顾新设专业对专业知识结构所带来的影响。适应智能时代需求的计算机类高素质人才，除了应具备科学思维、创新素养、敏锐感知、协同意识、终身学习和持续发展等综合素养与能力外，还应具有深厚的数理理论基础、扎实的计算思维与系统思维、新型计算系统创新设计以及智能应用系统综合研发等专业素养和能力。

智能时代计算机类专业教育计算机类专业系统能力培养 2.0 研究组在分析计算机科学技术及其应用发展特征、创新人才素养与能力需求的基础上，重构和优化了计算机类专业在数理基础、计算平台、算法与软件以及应用共性各层面的知识结构，形成了计算与系统思维、新型系统设计创新实践等能力体系，并将所提出的智能时代计算机类人才专业素养及综合能力培养融于专业教育的各个环节之中，构建了适应时代的计算机类专业教育主流模式。

自 2008 年开始，教育部计算机类专业教学指导委员会就组织专家组开展计算机系统能力培养的研究、实践和推广，以注重计算系统硬件与软件有机融合、强化系统设计与优化能力为主体，取得了很好的成效。2018 年以来，为了适应智能时代计算机教育的重要变化，计算机类专业教学指导委员会及时扩充了专家组成员，继续实施和深化智能时代计算机类专业

教育的研究与实践工作，并基于这些工作形成计算机类专业系统能力培养 2.0。

本系列教材就是依据智能时代计算机类专业教育研究结果而组织编写并出版的。其中的教材在智能时代计算机专业教育研究组起草的指导大纲框架下，形成不同风格，各有重点与侧重。其中多数将在已有优秀教材的基础上，依据智能时代计算机类专业教育改革与发展需求，优化结构、重组知识，既注重不变要素凝练，又体现内容适时更新；有的对现有计算机专业知识结构依据智能时代发展需求进行有机组合与重新构建；有的打破已有教材内容格局，支持更为科学合理的知识单元与知识点群，方便在有效教学时间范围内实施高效的教学；有的依据新型计算理论与技术或新型领域应用发展而新编，注重新型计算模型的变化，体现新型系统结构，强化新型软件开发方法，反映新型应用形态。

本系列教材在编写与出版过程中，十分关注计算机专业教育与新一代信息技术应用的深度融合，将实施教材出版与 MOOC 模式的深度结合、教学内容与新型试验平台的有机结合，以及教学效果评价与智能教育发展的紧密结合。

本系列教材的出版，将支撑和服务智能时代我国计算机类专业教育，期望得到广大计算机教育界同人的关注与支持，恳请提出建议与意见。期望我国广大计算机教育界同人同心协力，努力培养适应智能时代的高素质创新人才，以推动我国智能科技的发展以及相关领域的综合应用，为实现教育强国和国家发展目标做出贡献。

智能时代计算机类专业教育计算机类专业系统能力培养 2.0

系列教材编委会

2020 年 1 月

P R E F A C E

前言

目前主流的计算机网络教材，大多沿用计算机网络出现初期大学刚开设计算机网络课程时的教学思想，即对计算机网络——主要是因特网——进行解剖，将知识点、技术要点呈现给学生，希望学生学习并掌握这些知识和技术，侧重知识点的学习。为叙述方便，这些教材通常按体系结构模型的层次，逐层介绍每层涉及的知识、方法和技术。国际上著名的教材都采用这样的体例，遵从相似的内容架构，区别只是有的教材自底向上逐层介绍，有的教材自顶向下逐层介绍。这种基于解剖型思维的教材，优点是易于学生了解知识点，但其不足是容易使学生缺乏系统观、整体观，如果授课教师没有高屋建瓴地统筹教材内容，就不利于培养学生解决整体问题的系统能力。

研究型大学，特别是建立了计算机科学基础学科拔尖学生培养基地或计算机类一流本科专业建设点的大学，其目标是培养未来的高端人才、顶端人才，因而应以思维能力、创新能力、系统能力培养为目标。然而，传统的解剖型思维教学不利于这一目标的达成。

本书从思维能力、创新能力、系统能力培养出发，以建构型、设计型、创新型思维引导学生从无到有设计计算机网络。本书尽量摒弃单纯分析已有网络、按层次结构解剖的方式，而是从需求和任务出发，研究相关系统和问题的解决方法，据此构造计算机网络系统，并把已有的网络作为构造、设计网络的示例以及检验创新设计的例证。这也符合大型公司进行创新的模式。希望借此模式，帮助学生养成创新思维习惯，提升创新思维能力，建立系统观，培养解决复杂问题的能力，引导学生从注重学习知识点，转向注重培养系统能力和创新能力。

本书的总体逻辑是：首先，从需求出发，明确计算机网络应满足的要求和约束条件以及计算机网络可用的设计方法，基于此，设计将计算机连接成小型网络的方法，解决连接所涉及的基本理论、方法和技术问题；其次，设计将小型网络互联以组成大型网络的方法，实现任意计算机及应用程序之间数据的可靠传输，在此基础上，设计计算机网络服务和应用，并设计服务质量保证机制、网络安全机制、网络自动管理机制；最后，讨论面向未来的一些网

络技术，引导学生开拓新的研究视野。

对每类问题，本书在讲解原理和方法之后，将具体的网络技术或产品作为示例进行介绍。对具体的网络技术，如果现在已淘汰或基本上不再使用，本书原则上不介绍，对那些还在使用但从发展趋势上将会被代替的技术，做少量介绍。本书主要介绍的是基础性的共性技术、当前的主流核心技术或即将使用的新技术，对于相关技术的具体实现方法，大多以华为公司的方案为主进行介绍，这也是目前国际领先的技术。

本书适合研究型大学计算机相关专业的计算机网络课程教学使用，建议按顺序进行教学，第8章和第9章的顺序可互换。课程建议教学总时长为48学时，各章建议学时分别为第1章3学时、第2章14学时、第3章10学时、第4章3学时、第5章7学时、第6章1学时、第7章3学时、第8章3学时、第9章2学时、第10章2学时。有些学校单独开设“网络安全”必修课，这样可压缩第8章的课时，将其课时分摊到第6章和第7章。

本书由黄传河规划统筹，其他人员参与撰写了部分内容，其中，杜瑞颖撰写了第3章的部分内容和第8章的部分内容，张健撰写了第5章的部分内容，张沪寅撰写了第9章的部分内容，牛晓光撰写了第6章（刘丹丹、吕慧、周浩参与撰写），其余内容由黄传河撰写。

书中有关新的无线网络、交换机与路由器的实现技术以及有关网络管理系统的内容，华为公司的黄明祥、骆兰军、王生月、陈乐、王义博、崔云龙等专家撰写了翔实的技术资料。有关光网络（OTN、PON等）的内容，华为公司的何慧、卫年丽、彭玮、马建刚、曹慈卓、吕霄云等专家撰写了翔实的技术资料。限于篇幅，本书只选用了其中少量内容，在此对这些专家表示诚挚的感谢！

本书引用了其他书籍、论文、网络上的一些资料，未一一标注，将其一并列在“参考文献”部分，对这些文献的作者一并表示诚挚的感谢，若有遗漏之处，深表歉意。

目 录

丛书序言

前言

第 1 章 计算机网络设计方法学 1

1.1 计算机网络设计需求 2

1.1.1 计算机网络应用模式 2

1.1.2 计算机网络设计需求与
约束 41.1.3 计算机网络主要性能
指标 5

1.1.4 设计的综合平衡 7

1.2 网络服务模型与传输模式设计 7

1.2.1 网络服务模型 7

1.2.2 数据传输模式 8

1.3 网络连接方式设计 9

1.3.1 直接连接 9

1.3.2 多跳转发连接 10

1.4 网络拓扑结构设计 10

1.4.1 拓扑结构设计 10

1.4.2 网络覆盖性设计 12

1.4.3 通信模式设计 13

1.5 数据交换方式设计 13

1.5.1 电路交换 13

1.5.2 包交换 13

1.5.3 广播交换 14

1.5.4 包交换的实现方式 14

1.6 网络设计模式 18

1.6.1 网络体系结构分层设计
模式 18

1.6.2 协议、接口与服务 19

1.6.3 数据的封装与传输 22

1.6.4 ISO 分层模型示例——
OSI 参考模型 231.6.5 因特网分层模型示例——
TCP/IP 参考模型 24

1.6.6 虚拟层次模型 25

1.6.7 从单网到多网互联的
增量设计 26

1.7 计算机网络构成要素设计 26

1.7.1 计算机网络硬件 26

1.7.2 计算机网络软件 27

1.7.3 计算机网络应用 27

习题 28

第 2 章 基本网络设计 29

2.1 传输介质 29

2.1.1	传输介质的类别与特性	29	2.5.10	800Ge 以太网	92
2.1.2	有线传输介质	30	2.5.11	网卡设计	92
2.1.3	无线传输介质	34	2.5.12	交换机设计	94
2.1.4	传输介质的传输能力	36	2.5.13	VLAN	100
2.1.5	Wi-Fi 无线信号的功率	37	2.6	无线局域网设计	101
2.2	数据通信系统与物理层	38	2.6.1	无线局域网设计要点	101
2.2.1	数据通信系统模型	38	2.6.2	无线局域网示例—— Wi-Fi	103
2.2.2	数据传输模式	38	2.6.3	Wi-Fi 6 如何实现高 性能	108
2.2.3	信号编码与调制	41	2.6.4	Wi-Fi 7 前瞻	114
2.3	保证传输正确性的方法与数据链 路层	49	2.6.5	无线 AP 设计	114
2.3.1	介质接入控制	49	2.6.6	LiFi	116
2.3.2	组帧	53	2.7	个域网与体域网设计	116
2.3.3	检错与纠错	55	2.7.1	个域网与体域网的一般 结构	116
2.3.4	可靠性与应答机制	57	2.7.2	个域网示例——蓝牙 网络	117
2.3.5	同步方式	65	2.8	广域网设计	119
2.4	提高传输性能的方法	68	2.8.1	广域网设计思想	119
2.4.1	多路复用	68	2.8.2	广域网示例——SDH	120
2.4.2	认知无线电	72	2.8.3	广域网示例——OTN	124
2.4.3	网络编码	73	2.8.4	广域网示例——5G/6G 移动网络	134
2.5	局域网设计	73	2.9	接入网设计	137
2.5.1	局域网设计思想	73	2.9.1	接入网的一般结构	137
2.5.2	局域网示例——以太网	74	2.9.2	接入网技术示例—— PON	137
2.5.3	百兆以太网	80	2.9.3	接入网应用示例—— FTTH/FTTR	141
2.5.4	千兆以太网	83	习题		142
2.5.5	万兆以太网	87			
2.5.6	40Ge/100Ge 以太网	90			
2.5.7	25Ge 以太网	91			
2.5.8	50Ge/100Ge/200Ge 以太网	91			
2.5.9	200Ge/400Ge 以太网	91			

第 3 章 网络互联机制设计 143

3.1 网络互联的方法 143

3.1.1 网络互联的需求 143

3.1.2 网络互联的一般方法 144

3.2 路由算法 146

3.2.1 研究方法——图论方法 147

3.2.2 路由算法分类 148

3.2.3 随机路由算法 149

3.2.4 洪泛路由算法 149

3.2.5 热土豆路由算法 150

3.2.6 距离向量路由算法 150

3.2.7 链路状态路由算法 153

3.2.8 广播路由算法 154

3.2.9 组播路由算法 155

3.2.10 泛播路由算法 155

3.2.11 层次路由选择算法 156

3.2.12 多目标 QoS 路由 算法 156

3.3 网络互联协议示例——IPv4 及 相关协议 157

3.3.1 IP 数据报 158

3.3.2 IP 数据报的封装与分片 160

3.3.3 IP 地址与子网 161

3.3.4 IP 包传递流程 164

3.3.5 IP 路由协议——RIP 165

3.3.6 IP 路由协议——OSPF 168

3.3.7 边界网关协议——BGP 171

3.3.8 组播协议——PIM 与 MOSPF 173

3.3.9 地址解析与反向地址 解析——ARP 与 RARP 174

3.3.10 控制报文协议——

ICMP 176

3.3.11 自动地址分配——

DHCP 178

3.3.12 地址不足的解决方案之一 ——无分类域间路由 181

3.3.13 地址不足的解决方案之二 ——网络地址转换 182

3.4 网络互联协议示例——IPv6 及 相关协议 183

3.4.1 IPv6 的设计理念 184

3.4.2 IPv6 包格式 184

3.4.3 IPv6 地址 188

3.4.4 IPv6 地址自动配置 192

3.4.5 邻节点发现过程 193

3.4.6 IPv6 控制报文协议—— ICMPv6 194

3.4.7 IPv6 包转发过程 195

3.4.8 IPv4 向 IPv6 过渡 196

3.5 移动互联 198

3.5.1 移动性管理 198

3.5.2 移动 IPv4 199

3.5.3 移动 IPv6 202

3.6 分段路由——SRv6 205

3.6.1 IPv4/IPv6 的困境 205

3.6.2 SRv6 解决方案 205

3.7 路由器设计 207

3.7.1 路由器设计需求 207

3.7.2 路由器的结构 209

3.7.3 路由器处理流程 212

3.7.4 路由器关键技术 213

习题 213

第 4 章 端到端数据传输方法

设计 215

4.1 端到端的数据传输服务 215

4.1.1 为什么需要端到端传输
服务 215

4.1.2 传输服务类别及实现策略 216

4.1.3 端到端传输服务的一般
模式 216

4.2 端到端不可靠传输协议示例——
UDP 217

4.2.1 UDP 的设计目标 217

4.2.2 UDP 应用和端口 218

4.2.3 UDP 报文格式 218

4.2.4 UDP 工作流程 219

4.2.5 UDP 校验和的计算 219

4.3 端到端可靠传输服务示例——
TCP 219

4.3.1 TCP 的设计目标 220

4.3.2 端口与套接字 221

4.3.3 TCP 报文格式 221

4.3.4 TCP 建立与终止连接
机制 225

4.3.5 TCP 定时管理自适应
重传机制 229

4.3.6 滑动窗口机制 230

4.3.7 TCP 优先数据传送 232

4.3.8 移动 TCP 优化 232

4.3.9 星际 TCP 优化 232

4.3.10 多路 TCP 233

4.3.11 取代 TCP 的 QUIC 233

4.4 实时传输协议示例——RTP 234

4.4.1 RTP 的功能与需求 234

4.4.2 RTP 报文格式 234

4.4.3 实时传输控制协议——
RTCP 235

习题 235

第 5 章 网络公共服务设计 237

5.1 网络公共服务 237

5.1.1 公共服务分类 237

5.1.2 公共服务的提供形式 238

5.1.3 因特网公共服务 239

5.2 DNS 239

5.2.1 DNS 的功能及原理 239

5.2.2 因特网的域名结构 239

5.2.3 资源记录 241

5.2.4 域名服务器 242

5.2.5 域名解析——递归解析 244

5.2.6 域名解析——迭代查询 245

5.2.7 利用缓存提高查询速度 246

5.2.8 DNS 报文格式 246

5.3 文件传输协议 247

5.3.1 FTP 的工作过程 248

5.3.2 FTP 的工作模式 250

5.3.3 FTP 的登录鉴别与匿名
访问 251

5.3.4 FTP 的命令与应答 251

5.3.5 FTP 的实现与工具 255

5.3.6 FTP 工具的改进 255

5.4 Web 与 HTTP 255

5.4.1 Web 资源组织方式	256	6.1.1 搜索引擎的设计思想	292
5.4.2 Web 文档格式与 HTML	258	6.1.2 信息收集与存储	293
5.4.3 HTTP 概述	261	6.1.3 信息的快速检索	294
5.4.4 HTTP 通信过程	263	6.1.4 信息推荐	294
5.4.5 HTTP 报文	264	6.2 即时通信系统	295
5.4.6 状态保存	267	6.2.1 即时通信系统与系统 组成	295
5.4.7 提高应答速度的方法	268	6.2.2 IM 示例——QQ	298
5.4.8 HTTP 2.0	270	6.3 网络电视与视频服务	302
5.4.9 HTTP 3.0	273	6.3.1 音视频编码	302
5.4.10 浏览器	273	6.3.2 流式音视频传输	304
5.4.11 Web 的发展	274	6.3.3 应用层组播	306
5.5 电子邮件协议	275	6.3.4 网络电视	307
5.5.1 电子邮件系统架构	275	6.4 电子商务	308
5.5.2 邮件格式	276	6.4.1 电子商务系统架构	308
5.5.3 邮件传输协议—— SMTP	277	6.4.2 电子支付协议	309
5.5.4 邮件接收协议—— POP3	281	6.4.3 移动支付	309
5.5.5 邮件访问协议—— IMAP4	283	6.5 工业互联网	310
5.5.6 POP3 与 IMAP4 的比较	285	习题	311
5.6 远程终端协议 Telnet	285	第 7 章 网络服务质量	312
5.6.1 Telnet 的设计目的	285	7.1 网络性能评价	312
5.6.2 Telnet 的基本内容	286	7.1.1 综合性能评价	312
5.6.3 Telnet 的关键方案	287	7.1.2 排队论模型	313
5.6.4 Telnet 的重要规定	288	7.1.3 利用排队论分析网络 系统的性能	316
5.6.5 基于 Telnet 的工具	291	7.2 网络资源优化分配	318
习题	291	7.2.1 资源类别与分配原则	318
第 6 章 网络应用设计	292	7.2.2 典型的优化方法	318
6.1 搜索引擎	292	7.3 拥塞控制的一般方法	319
		7.3.1 拥塞控制的目标	319

7.3.2 拥塞控制的原理	320	8.2.4 恶意程序攻击	354
7.3.3 拥塞控制的方法	321	8.2.5 程序漏洞攻击	355
7.4 拥塞控制示例——TCP 与 IP		8.2.6 渗透攻击与欺骗攻击	358
拥塞控制方案	325	8.3 网络防护	359
7.4.1 加增乘减	325	8.3.1 加密	359
7.4.2 慢启动和拥塞避免	327	8.3.2 防篡改	362
7.4.3 快重传和快恢复	328	8.3.3 数字签名	363
7.4.4 CUBIC	330	8.3.4 认证	363
7.4.5 TCP 的性能优化	332	8.3.5 访问控制	366
7.4.6 IP 层随机早期检测	332	8.3.6 防火墙	372
7.4.7 因特网拥塞控制机制		8.3.7 入侵检测与入侵防护	379
讨论	333	8.3.8 VPN 与 NVo3	384
7.5 服务质量与体验质量	334	8.3.9 物理网络安全	390
7.5.1 服务质量概述	334	8.4 网络安全通信协议示例	392
7.5.2 QoS 示例——IntServ	335	8.4.1 HTTPS	392
7.5.3 QoS 示例——DiffServ	337	8.4.2 TLS 与 SSL	394
7.5.4 QoS 示例——MPLS	338	8.4.3 IPSec	399
7.6 流量工程	340	8.4.4 电子邮件安全协议	399
习题	341	8.4.5 安全外壳协议	400
第 8 章 网络安全	342	8.4.6 无线安全协议	402
8.1 网络安全风险	342	8.5 网络系统安全保证体系	403
8.1.1 网络面临的主要安全风险	342	8.5.1 网络安全评估标准	403
8.1.2 网络安全风险的根源	345	8.5.2 公钥基础设施	407
8.1.3 网络安全的目标	345	8.6 区块链	409
8.1.4 网络攻击与防御的主要		8.6.1 区块链的原理	409
技术	346	8.6.2 区块链的应用	411
8.2 网络攻击	346	8.7 暗网	411
8.2.1 网络探测	346	习题	412
8.2.2 拒绝服务攻击	349	第 9 章 网络管理	413
8.2.3 缓冲区溢出攻击	352	9.1 网络管理需求	413

9.1.1 为什么需要网络管理	413	9.7.2 iMaster NCE 的系统 组成	444
9.1.2 网络管理的功能需求	414	习题	448
9.1.3 网络管理的性能需求	417	第 10 章 新型网络技术	449
9.1.4 网络管理体系结构	418	10.1 软件定义网络	449
9.2 管理信息的组织与表示	422	10.1.1 SDN 的设计思想	449
9.2.1 信息组织方案示例—— SMI	422	10.1.2 OpenFlow 协议	452
9.2.2 信息表示方案示例—— MIB	424	10.1.3 OpenFlow 交换机	458
9.3 网络管理协议——SNMP	428	10.1.4 OpenFlow 控制器	460
9.3.1 SNMP 的设计理念	428	10.1.5 SDN 开源系统	461
9.3.2 SNMPv1	429	10.1.6 SD-WAN	463
9.3.3 SNMPv2	432	10.2 网络功能虚拟化	464
9.3.4 SNMPv3	433	10.2.1 NFV 原理	464
9.3.5 SNMP 的应用	433	10.2.2 网络切片	466
9.4 网络管理协议——NETCONF	434	10.2.3 FlexE	468
9.4.1 NETCONF 的设计理念	434	10.3 数据中心网络	470
9.4.2 NETCONF 的主要内容	435	10.3.1 数据中心网络的设计 需求	470
9.4.3 YANG 模型	439	10.3.2 数据中心网络架构	470
9.5 网络管理协议——OpenConfig	439	10.3.3 InfiniBand 网络	472
9.5.1 OpenConfig 的设计理念	439	10.3.4 RoCE 网络	473
9.5.2 OpenConfig 的主要内容	440	10.4 物联网	473
9.6 网络测量	440	10.4.1 物联网架构	474
9.6.1 网络拓扑测量	440	10.4.2 物品联网方法与典型 协议	476
9.6.2 网络性能测量	442	10.4.3 对象名称服务 / 标识 解析服务	476
9.6.3 网络流量测量	443	10.4.4 物联网应用示例—— VANET	477
9.7 网络管理系统示例——iMaster NCE	444	10.5 时间敏感网络	477
9.7.1 iMaster NCE 的设计 思想	444	10.5.1 TSN 的设计要求	477

10.5.2 TSN 的主要技术	478	10.7 命名数据网络	486
10.6 卫星互联网	481	10.7.1 命名数据网络的体系 结构	486
10.6.1 卫星互联网的轨道	481	10.7.2 内容命名与缓存方法	489
10.6.2 卫星互联网的星座	481	10.7.3 内容路由方法	490
10.6.3 卫星互联网的关键 技术	482	10.8 网络中的人工智能方法	490
10.6.4 卫星互联网的局限	483	习题	492
10.6.5 主要的卫星互联网 系统	483	参考文献	493
10.6.6 空天地一体化网络	485	缩略词	495

供學生臨時使用

计算机网络设计方法学

计算机能自动化、智能化地处理数据，但如何将一台计算机上的数据自动传输到另一台计算机上，自动实现数据在多台计算机之间的共享与处理呢？计算机的普及自然而然地催生了这一应用需求。对于满足这种需求的理论、技术及系统，我们统称其为计算机网络。现在，连接到计算机网络上的不仅有计算机，还有各种智能设备，包括具有通信功能的物品（智慧物品）。

计算机网络是一个复杂的系统，研究、设计计算机网络时应从哪里着手？目标是什么？有哪些设计素材可供利用？

从需求开始研究、设计、构建计算机网络，需要明晰并解决以下关键问题：

- 计算机网络的应用模式是什么？为此需要的服务模型是什么？据此可支撑哪些类别的应用？
- 计算机网络应满足哪些主要设计需求与约束条件？
- 组成计算机网络的典型要素是什么？
- 如何设计良好的网络形态及连接方式以方便、高性价比地将计算机连接成网络并使网络高效运行？
- 如何通过良好的设计方法、网络架构以降低设计复杂性，并提高网络性能、可靠性、可用性、可维护性、可扩展性和安全性？

对这些问题的理论和技术回答，构成了计算机网络设计方法学的基本内容。从无到有设计计算机网络并非一切从零开始，而是要充分利用前人的研究成果，比如已有的通信介质、数据通信理论和技术、计算机系统等。

本书所说的计算机网络设计不是计算机网络工程设计。网络工程设计是指按照工程规范购买已有的网络设备、系统等，把它们高效地集成到一起，以实现用户业务需求。本书所说的网络设计是指从理论出发，构想、设计网络的理论、算法、协议、设备、基础软件等，使其能协同工作，以体现网络特性和功能，它们是网络工程的前提，是网络工程可供使用的原材料。

在开始本章内容之前，我们首先介绍几个术语。

- 计算机网络：顾名思义，计算机网络就是将计算机连接起来实现信息共享的网络。这里的计算机是广义的，包括各种类型的计算机，如 PC、超级计算机、平板计算机等，也包括具有计算功能的智能设备，如智能手机、智能穿戴设备等，还包括智能家电设备、智能机器人、工厂智能生产设备、卫星设备等。因此，将今天的计算机网络称为信息网络也许更恰当，将其称为物联网更符合未来的发展趋势。
 - 设备：泛指计算机网络中的计算机、网络设备（如交换机、路由器）等。
 - 节点：泛指计算机网络中的任意设备，有时特指交换机、路由器等设备。
 - 链路：指直接连接两个相邻节点的传输介质。
 - 路径：指任意两个节点之间由一系列首尾相连的节点及链路组成的通路，极端情况下就是一条链路。
 - 跳（hop）：指一段链路的距离，一段链路为一跳，多段链路为多跳，不考虑链路的物理长度。 N 跳的路径从起点到终点共包含 $N+1$ 个节点。
- 除非需要显式地进行区分，本书将混合使用计算机、设备、节点等概念。

1.1 计算机网络设计需求

1.1.1 计算机网络应用模式

计算机网络应用系统利用多台计算机之间的协作完成用户所期望的功能。与其他所有系统一样，在进行网络设计之前需要尽可能清晰地定义应用需求，设定运行模式，界定边界条件。

应用各种各样，无法一一枚举。按照应用的功能，大致可把应用分为信息分享类应用、商务交易类应用、交互通信类应用、在线娱乐类应用、分布计算类应用五类。这不是唯一的分类方法，不同的分类标准会得到不同的分类结果。

（1）信息分享类应用

信息分享类应用将所持有的信息发送给其他一个或一批用户（或被其他用户主动读取）以实现信息共享。按照信息传输和处理的方式信息分享类应用可细分为信息展示类应用、文件传输类应用、流媒体传输类应用等类别。

信息展示类应用的典型例子是信息网站类应用（如门户网站）。其模式是用户端向网络（对应的是网站服务器）提出查看内容的请求，服务器将所请求查看的内容发送给用户端，用户端接收、解释并显示相关内容，所请求的内容以网页形式呈现，如图 1-1 所示，图中的 ■ 表示数据块。此类应用的关键是服务器端的并发应答与服务机制。

文件传输类应用是信息展示类应用的一种形式。其模式是用户端向网络（文件服务器）请求下载所需的文件，双方协同将服务器上的文件传送到用户端计算机上保存，如

图 1-2 所示。与下载相反的是,用户将文件上传到文件服务器以供其他用户共享或保存。此类应用的关键是文件传输的正确性保证机制。

(2) 商务交易类应用

商务交易类应用的典型例子是电子商务,借助交易平台实现商品的买卖,其中包括商品交易、资金清算、物流信息管理等重要信息的处理,涉及买方、卖方、银行、物流公司等交易的参与方。此类应用的原理如图 1-3 所示,其关键是多方参与、交易数据的完整性与安全性保障机制。

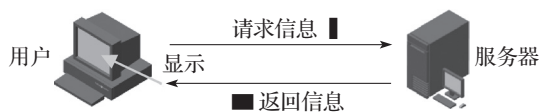


图 1-1 信息展示类应用原理示意图

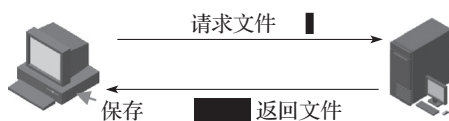


图 1-2 文件传输类应用原理示意图

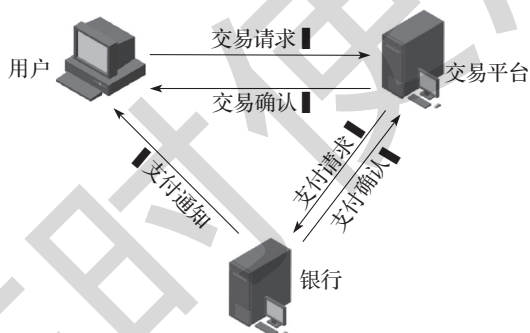


图 1-3 商务交易类应用原理示意图

(3) 交互通信类应用

交互通信类应用实现用户之间的实时交互通信,典型代表是 QQ、微信等即时通信工具。此类应用的原理如图 1-4 所示,其关键是服务器端的并发处理与存储以及通信双方/多方之间的同步机制。

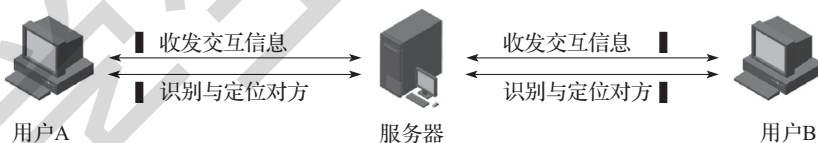


图 1-4 交互通信类应用原理示意图

可以将流媒体传输类应用看成交互式应用的一种特例,典型代表是音视频通话服务。用户向服务器请求流媒体内容(音频、视频等)或一方主动分发流媒体内容(如电视、视频会议等)。流媒体并非一次性传输完整的内容,而是将内容分解为片段并连续传输片段,接收方接收、解析、播放/展示所接收的流媒体内容。此类应用的原理如图 1-5 所示,其关键是流式传输机制和实



图 1-5 流媒体传输类应用原理示意图

时性保证机制。在线视频通话是不同于视频下载的双向音视频传输。

(4) 在线娱乐类应用

在线娱乐类应用的典型代表是在线游戏，提供游戏场景以模拟现实游戏环境，供多方同时游戏。此类应用的原理如图 1-6 所示，其关键是游戏场景的逼真性、游戏数据的快速多向传输与同步机制。

(5) 分布计算类应用

分布计算类应用将用户的大型计算任务调度到多台计算机或单一高性能计算机（不属于用户所有）上进行协同和快速计算，云计算、边缘计算是这类应用的典型代表。此类应用的原理如图 1-7 所示，其关键是高效任务调度和高速计算性能保障机制。

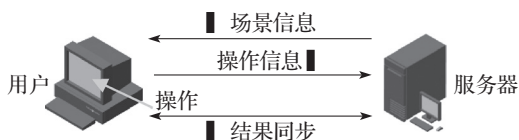


图 1-6 在线娱乐类应用原理示意图

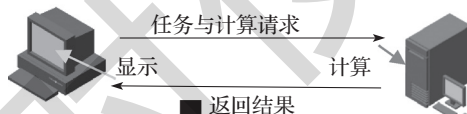


图 1-7 分布计算类应用原理示意图

除此之外，其他常见应用模式（比如电子邮件）也可归为上述某类应用模式或其组合模式。

了解了网络应用的特征及传输需求，就明确了网络设计的基本前提和目标。

1.1.2 计算机网络设计需求与约束

1. 设计需求

计算机网络应满足如下主要需求。

(1) 可扩展的连通性

网络中的任意两台设备之间应该至少存在一条连通的路径，能够彼此通信，且通常情况下网络中任意两台设备之间的距离不随着网络规模的扩大而快速增加，尤其不随着规模的扩大而呈线性或指数增加。设备连网的方式不因网络规模的变化而变化，设备接入网络的物理连接性能（不是指实际获得的使用性能）尽可能不因网络规模的扩大而显著降低，以使网络具有良好的可扩展性。

(2) 高性价比的资源共享机制

计算机网络是实现资源共享的系统，共享的资源包括数据资源、硬件资源、软件资源、通信资源等。

- 数据资源包括各类静态文件、动态数据、音视频数据等。
- 硬件资源包括服务器、存储设备、打印机等。
- 软件资源包括显式共享的软件，如云计算软件，以及隐式共享的软件，如数据库管理系统、网站服务软件等。

- 通信资源包括共享的通信线路（带宽）、网络通信设备（交换机、路由器等）、网络安全设备等。

对于数量庞大的资源，需要设计良好的、性价比高的共享机制，才能有效支持共享和网络的有效运行。

（3）公共服务机制的普适性

计算机网络提供的服务种类繁多，现在已发展到几乎无所不能的程度。我们不可能为每种功能、服务单独设计一种机制或工作模式。这就要求对网络的功能、服务进行深入研究，抽象出公共特征，设计出具有广泛适应性的公共服务机制，支持各种网络功能的执行，降低网络设计、实现的难度和工作量。

（4）规模无关的可管理性

网络分布范围广、功能复杂，不可能依靠人工方式进行维护和管理。在网络设计之初，就需要设计一套能自动运行的网络管理机制，实现网络的自我管理，并且该机制具有规模自适应性，不受网络规模的限制，在规模发生变化时依然保持鲁棒性、高效性。

（5）业务适配的安全性

网络空间中存在各种各样的不安全因素，其中既有设计缺陷、漏洞等技术因素，也有故意破坏等人为因素。为保证网络提供的功能和服务是安全可信的，需要将安全性作为网络固有的需求，与网络功能同时设计，使其具有内生安全性。同时，安全性机制还应具有进化适应性和业务适配性，在网络规模扩展、技术迭代、业务需求升级等多种条件下依然有效。

2. 设计约束

在进行计算机网络设计时，由于一些客观因素的限制，会受到一些条件的约束。应在满足这些约束条件的前提下，尽量遵循前述设计需求。一些典型的约束条件包括：

- 方法普适性。设计方法应具有普适性和前瞻性，所依据的理论应具有普适性，不能完全受限于自身水平和当前技术条件，应能满足未来发展的需要。
- 技术可行性。在具体实现时，在坚持先进性的同时，必须考虑技术上的可行性和可实现性。
- 性能可及性。一方面性能应在现实条件下达到尽可能高的程度，另一方面应尽可能达到应用级用户所期望和能接受的程度。
- 经济有效性。计算机网络是一种信息基础设施，与其他基础设施一样，在设计之初也应考虑其实现的经济性、用户的可及性。

1.1.3 计算机网络主要性能指标

评价计算机网络的优劣有多种维度、多种指标。依据不同的目的，评价指标也会不同。本节介绍几种常用的指标。

1. 数据率

数据率 (data rate) 指数据的传送速率, 也叫比特率, 也经常被称为速率。数据率表示单位时间内能将数据发送到传输介质上的速率 (不等于传输到目的地的速率), 由于网络中单一传输信道上的数据通常都是按串行方式一位一位传输的, 因此该数值也表示单位时间内信道传输数据的能力。数据率一般用 bit/s 表示 (为简化表达, 本书使用 bps 表示), 在数值很大时, 通常用 kbps、Mbps、Gbps、Tbps 等表示。K、M、G、T 的数据值通常取为 10^3 、 10^6 、 10^9 、 10^{12} , 但其实际值也可能分别为 2^{10} 、 2^{20} 、 2^{30} 、 2^{40} , 此时也可用 Kibps、Mibps、Gibps、Tibps 表示。具体取哪种值应根据上下文确定。

需指出的是, 数据率是指信号级的传输速率, 主要由所采用的编码或调制技术决定, 不等同于具体应用级的性能。

2. 带宽

通俗地说, 带宽 (band width) 是指频带的宽度, 即信道允许通过的信号频带宽度, 代表线路传输信号的能力, 一般用 Hz 表示, 但在日常生活中, 通常用转换后的数值即 bps (bit/s) 来表示。网络带宽更多是针对网络而言的, 更通俗地说, 大家习惯用其表示“网速”。带宽与数据率的区别是, 数据率主要指单一信道上的数据传送速率, 带宽更多指网络级 (可能有多条信道) 的数据传送能力。在特定条件下, 二者的数值可以是相等的。

第 2 章会给出带宽的准确含义。需要指出的是, 带宽并不代表单位时间内一定能把相应数量的用户数据传送到目的地。例如, 带宽为 100Mbps 的网络, 并不能保证在 1s 内把用户的 100Mb (12.5MB) 数据传输到目的地, 因为网络传输时还有其他各种开销 (如同寄快递还要外加包装、经历多次分拣中转, 包装就是一种开销, 分拣中转也是开销), 此外, 网络上还可能因出现拥塞而导致等待或因出现差错而需要进行纠错等。

3. 延迟

延迟 (delay) 也称时延, 是指从发送数据开始到数据被收到为止所经历的时间。延迟通常由四部分构成。

- 发送延迟 (transmission delay): 把一定量的数据从计算机发送到传输介质所经历的时间。在网络上发送数据通常是串行进行的 (一位一位地发送), 因此发送延迟跟数据率、数据量有关, 即

$$\text{发送延迟} = \text{发送的数据量} / \text{数据率}$$

- 传播延迟 (propagation delay): 信号从网络 (或介质) 的一端传输到另一端所经历的时间。理论上, 信号都是电磁波 (除量子信息、生物信息之外), 其传播速度为光速, 因此, 传播延迟跟传输介质的长度有关, 即

$$\text{传播延迟} = \text{介质长度} / \text{信号传播速度}$$

受多种因素的影响, 信号在介质中的传播速度不一定是光速。比如, 信号在非均匀介质中的传播速度约为 $2c/3$ (光速的 $2/3$)。

- 等待延迟 (queuing delay): 数据在网络上排队等待发送 (转发) 的等待时间。这个时间长度跟网络负载、网络性能、调度策略等因素相关, 没有固定的值, 变化范围较大, 但通常可以用排队论模型 (第7章介绍) 进行估算。
- 处理延迟 (processing delay): 在计算机或网络设备上对数据进行处理所需要的时间, 包括分配缓冲区的时间、计算校验和判断数据正确性的时间、封装 / 解封数据包的时间等。通常, 这部分时间相对于前述三部分时间而言比较短, 所以, 在要求不是特别严格时, 可忽略这部分延迟。

4. 延迟抖动

延迟抖动 (jitter) 是指两个节点之间传输数据时连续两次传输的延迟时间差。该参数表征网络性能的稳定性, 对在线音视频的传输和播放 (包括视频会议、网络电话等) 质量具有重要的影响。

5. 服务质量 (QoS) 与体验质量 (QoE)

QoS 与 QoE 是用户直观感受到的网络质量。QoS 从网络角度定义网络能为用户提供的服务质量, 比如响应速度、可靠性、安全性、丢包率等, 可以用量化的指标进行评价。QoE 侧重用户的感受, 比如用户感受到的网络速度、流畅度、等待时间等。这些质量也可通过综合的定量指标进行定义。

1.1.4 设计的综合平衡

通常, 网络的相关方包括网络设计者、网络运维者、网络使用者三方。这三方关注的重点有所不同。

- 网络设计者: 通常更关注技术先进性、系统可实现性、网络可靠性。
- 网络运维者: 通常更关注网络可靠性、管理方便性、技术先进性。
- 网络使用者: 通常更关注功能全面性、使用方便性、价格合理性。

作为网络的设计者, 不能只关注自己重视的那些方面, 还要站在其他两方的角度对设计方案进行全面平衡, 对一些暂时不能完全满足各方需求的内容, 应进行适当折中, 最终给出性价比高的方案。

1.2 网络服务模型与传输模式设计

1.2.1 网络服务模型

用户向网络提出服务请求, 网络应答用户的请求, 提供用户所需要的数据或信息, 或者执行规定的操作, 这种功能的实现方式简称为网络服务模型。依据网络执行服务请求方式的不同, 宏观上可以将网络服务模型分为 C/S 和 P2P 两种基本模型。

1. C/S 模型

C/S 的全称是 Client/Server, 即客户机 / 服务器。C/S 模型的含义是, 客户机向特定

的服务器发出请求，服务器接收请求、分析请求、执行请求并返回执行结果。C/S 模型的理论基础是 RPC（远程过程调用，相对于单机上的本地过程调用），其原理如图 1-8 所示。图 1-8a 表示基本 C/S 模型，图 1-8b 表示一对多 C/S 模型（一项功能需要多个服务器共同完成），图 1-8c 表示多级 C/S 模型。这些模型还可组合使用，也可以一个请求返回一连串的结果（例如音视频播放，在一个请求后持续地返回音视频片段）。

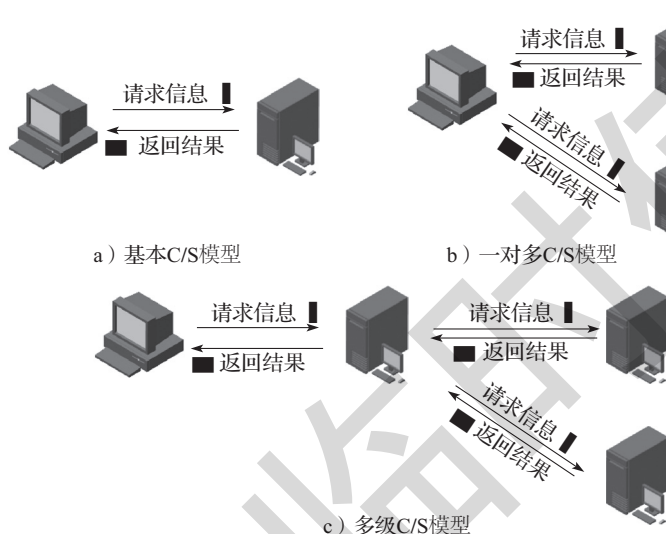


图 1-8 C/S 模型示意图

C/S 的一种特殊形式是 B/S (Browser/Client)，即浏览器 / 服务器模式，客户端发送给服务器的请求信息是一个网页 ID，浏览器返回的信息是一个网页（文件）。当然，B/S 请求与返回的也可以不是网页文件，而只是一个参数，比如请求的是用户 ID，返回的是用户的姓名。

2. P2P 模型

P2P 的全称是 Peer to Peer，即对等模型，其原理是所有客户机都是对等的，没有固定的服务器，客户机发出的请求不针对特定的服务提供者，其他客户机都可接收请求、评估请求并决定是否应答该客户机的请求。如果符合条件，则执行请求并返回请求的结果。其原理如图 1-9 所示。

究竟采用哪种服务模型，要根据所设计网络的功能及其服务需求确定。

1.2.2 数据传输模式

将网络中一台设备上的数据传输到另一台设备（或服务器）时，有两种基本的传

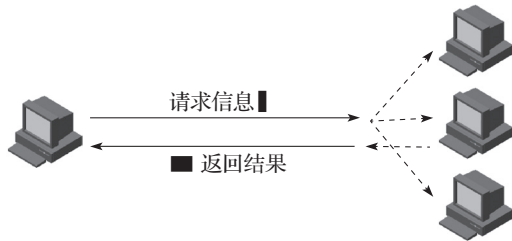


图 1-9 P2P 模型示意图

输模式。所有数据都可以这两种传输模式之一为基础实现数据的传输。

- 请求 / 应答 (request/reply) 模式。当一台设备要向另一台设备发送或从另一台设备请求获取数据时, 先向对方发出请求 (请求对方接收或请求对方发送), 对方收到请求后给出应答 (响应), 然后双方之间完成一次数据传输。重复此过程便可完成大量数据的传输。利用此模式可以实现众多类型的应用。
- 流式 (stream) 模式。当一台设备要向另一台设备发送或从另一台设备请求获取数据时, 先向对方发出请求 (请求对方接收或请求对方发送), 对方收到请求后给出应答 (同意), 然后双方之间连续传输一批数据 (多次传输), 而不需要像请求 / 应答模式那样逐次请求 / 应答。这种模式主要用于流媒体类应用, 也可以用于批量数据传输类应用 (例如文件传输)。

1.3 网络连接方式设计

用什么方法将计算机连接在一起组成网络呢? 从计算机之间的关联关系看, 有两种主要的方法, 即直接连接和多跳转发连接 (间接连接)。

1.3.1 直接连接

直接连接是指所有的计算机都直接连接到共享的物理介质上以组成网络, 可将所有的计算机都看成彼此直接相连。介质可以有有线介质 (如光纤、电缆), 也可以是无线介质 (如无线电、微波), 所有设备直接共享传输介质。这种类型的网络通常被称为广播型网络, 主要采用广播方式发送信息。当一个设备要发送信息时, 把信息广播出去, 其他设备都可以接收信息。信息从源节点传送到目的节点只有一条路径, 不需要进行路由选择。在这种类型的计算机网络中, 可以将任意两个节点看作相邻的节点, 因此, 信息只需经过一跳 (hop, 即一段链路) 即可到达目的节点。

直接连接的网络的典型形态有简单直连网络 (如图 1-10a 所示)、总线型网络 (如图 1-10b 所示)、环形网络 (如图 1-10c 所示) 等。

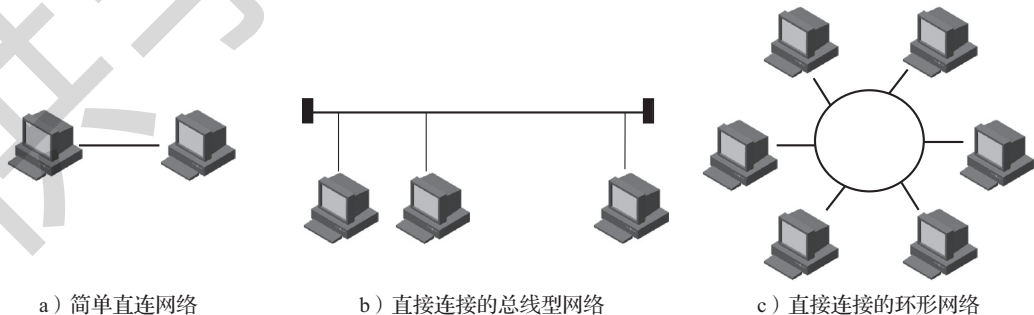


图 1-10 直接连接型网络

1.3.2 多跳转发连接

很多情况下, 计算机不能直接连接在一个介质上形成直接相连的网络, 这就需要间接连接、中间节点转发这种连接方式。

多跳转发网络是指两台计算机之间一般没有直接的链路相连, 从而借助其他节点间接相连形成的网络。这样从源计算机发送的信息需要经过多个中间节点转发, 经过多跳才能到达目的计算机。中间的过渡节点称为转发节点。可以设计专用的转发节点, 也可让一些用户计算机作为转发节点。图 1-11a 所示是一种多跳转发型有线网络, 图 1-11b 所示是一种多跳转发型无线网络。

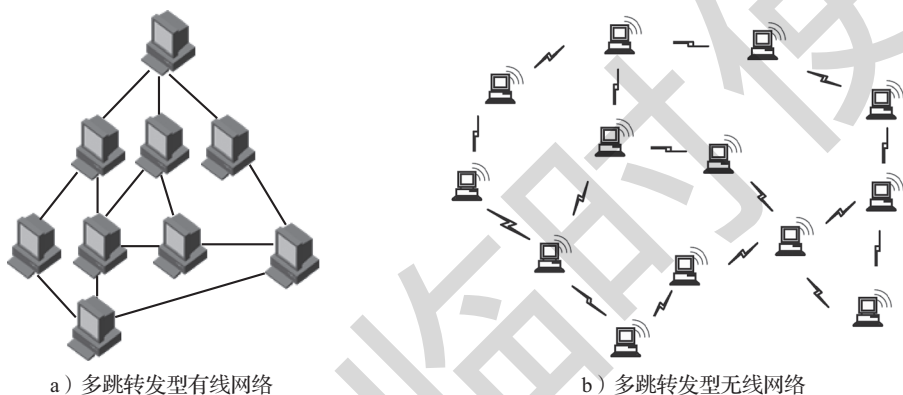


图 1-11 多跳转发网络

多跳转发网络的结构通常较复杂, 任意两个节点之间一般都存在多条可选路径, 这样, 信息传输前一般需要先进行路径选择。

1.4 网络拓扑结构设计

1.4.1 拓扑结构设计

网络的拓扑结构是指网络的布局结构, 通俗地说就是网络的构型。不同的拓扑结构具有不同的特性, 对网络的性能、可靠性、实现成本甚至安全性等具有较大的影响。理论上, 可选用的拓扑结构形式主要有以下几种。

- 总线型拓扑: 用总线把各计算机连接起来, 如图 1-12 所示。总线型网络的优点是建网容易, 增减节点方便, 节省线路。其缺点是重负载时通信效率不高, 因为所有节点共享一条通信线路。
- 星形拓扑: 每台计算机都通过单独 (专用) 的线路与中心设备 (如交换机) 相连, 如图 1-13、图 1-14 所示。星形网络的优点是结构简单、建网容易、延迟小、便于管理, 缺点是线路较多、中心设备对故障敏感。

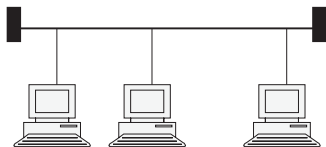


图 1-12 总线型网络

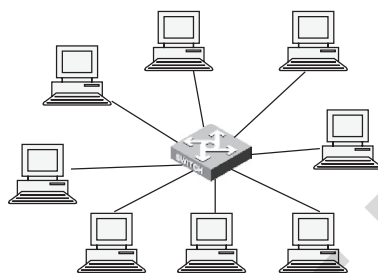


图 1-13 有线星形网络

- 环形拓扑：所有计算机通过环接口设备连接成一个环，如图 1-15 所示。环可以是单环也可以是双环，环中信号是单向传输的。双环网络中两个环上信号的传输方向相反，一般具备自愈功能。
- 树形拓扑：节点组织成树状结构，具有层次性，如图 1-16 所示。

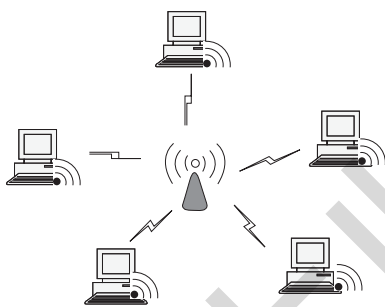


图 1-14 无线星形网络

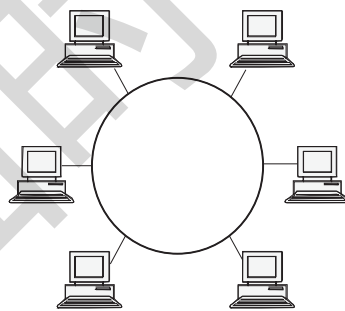


图 1-15 环形网络

- 网格形拓扑：一般情况下，每个节点至少要有两条路径与其他节点相连，如图 1-17 所示，可进一步细分为规则型和非规则型。网格形拓扑的优点是可靠性高，缺点是控制复杂、线路成本高。

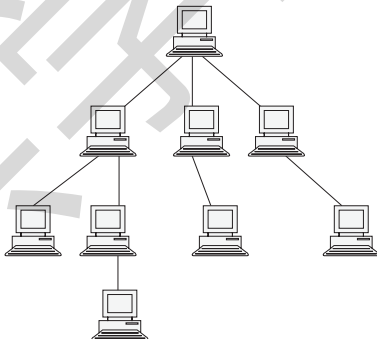


图 1-16 树形网络

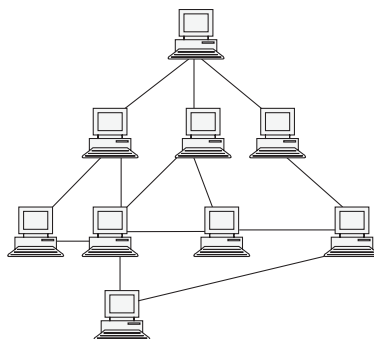


图 1-17 网格形网络

可以将这些基本型拓扑组合起来，组成更为复杂的网络。

这些类型的网络有时只能覆盖较小的范围、构建规模较小的网络，当需要构建大型网络时，需要联合使用多种方案。

不管是哪种网络，首要问题都是要解决发送权的分配问题，当有多个节点同时发送信息时，就有可能产生冲突，导致发送失败，因此，需要一种机制来保证任何时刻只有一个节点发送，或者无冲突地同时发送。解决这一问题的机制称为访问控制机制（或称接入控制机制），将在后续章节中具体介绍。

1.4.2 网络覆盖性设计

为适应不同用途、不同分布范围的需求，可以有针对性地设计更加有效的网络。通常，如果范围不同，比如几十米范围的网络与几千千米范围的网络，直觉上应采用不同的技术。为简化设计，可以根据不同的用途和分布范围采用不同的设计方法，先设计出不同类型的网络，再设计集成和互联方法，将其融合为统一的网络。

- 局域网。局域网（Local Area Network, LAN）一般分布在几十米到几千米的范围，连网设备数在几十到几百范围内。局域网要求数据率高、延迟小、建网成本低，可采用基于广播的通信方式。当使用无线传输介质时，称为无线局域网（WLAN），是局域网的一种特殊形式。
- 广域网。广域网（Wide Area Network, WAN）一般分布在数十千米的区域。连网设备数不受限制。广域网要求更高的可靠性措施、更复杂的拓扑结构、更复杂的路由机制。常规广域网一般不能采用广播式通信。
- 城域网。城域网（Metropolitan Area Network, MAN）一般分布在一个城区，通常使用广域网的技术，可以把它看作一个较小的广域网。因为范围比广域网小，所以其速度与局域网接近。实际上，城域网不是一种独立的网络类型。
- 个域网。个域网（Personal Area Network, PAN）一般指家庭内甚至是个人随身携带的网络，通常分布在几米至十几米范围内，用于将家用计算机设备、家用电器、消费电子设备、各类感知设备、可管理的物联化设备连接成一个小网络，以采用无线通信方式为主。依据具体应用的不同，有些个域网要求很高的速度（例如家庭内的高清电视节目传输），有的对速度要求不高但要求信号有很强的穿透能力、抗干扰能力（例如智能家居控制）。
- 体域网。体域网（Body Area Network, BAN）一般指个人随身携带的网络，通常分布在几米甚至几十厘米范围内，用于将随身携带（包括植入身体）的电子设备连接成一个微型网络，以采用无线通信方式为主。例如，智能手表、健康监测设备、手环、脑机接口设备、物联化设备等组成的微型网络。有时需要为体域网设计一个对外联络的设备，称其为网关。
- 空间信息网。空间信息网（Spatial Information Network/Space Information Network, SIN）指跨越太空的网络，通常分布范围在几万千米，用于将地球上的设备、空中

飞行器、海上舰船、卫星及太空中的其他人造设备、发射到其他星球上的设备连接成一个超远距离的网络，采用无线通信方式。该类网络因通信链路长、延迟时间大、干扰因素多，需要设计特别的网络技术才能保证网络的连通性、正确性和有效性。

1.4.3 通信模式设计

通信技术为计算机网络提供了通信基础。按照通信模式的不同，主要有有线通信和无线通信两种。

- 有线通信借助有线通信介质实现数据的传输。有线介质主要有光纤、双绞线、同轴电缆等，其特点是带宽较高，缺点是需要布设有线介质。
- 无线介质借助空间传输的电磁波实现数据传输。无线介质主要有无线电、微波、激光、红外线等，其特点是无须布设通信介质，缺点是一般情况下带宽低于有线介质。

1.5 数据交换方式设计

计算机网络中设备之间的信息交换需要遵循共同的规则，包括数据格式、转发方式等，这些规则在原理上统称为交换方式，在实现时被称为协议。

可利用的交换方式主要有电路交换（circuit switch）、包交换（也称分组交换，packet switch）、广播交换（broadcast switch）等。

1.5.1 电路交换

两个节点开始正式通信之前，在源节点和目的节点之间建立一个专用的通信连接，激活并独占通信线路用于数据传送。其在整个数据传输期间一直独占线路，通信结束后释放已建立的通信连接。

电路交换有两大优点。一是传输延迟小，主要的延迟是信号的传播延迟；二是一旦线路接通，便不会产生竞争和冲突，因为通信双方独占线路。

电路交换也有三个缺点。一是建立连接需要额外的时间。二是线路独占会造成资源浪费，因为线路一旦因建立连接而被激活，即便空闲也不能被其他用户所用。三是不能进行差错控制，源节点通常直接发送原始数据，且经过中间节点时也不停顿，即使在传输过程中出错，也没有相应的措施可发现并纠正错误。

传统的电话网络使用的就是电路交换方式，但今天的电话系统在骨干网上一般不再使用这种方式，而改用包交换方式。

1.5.2 包交换

包交换也称分组交换，类似于邮政寄送信件，在源节点，将用户待传送的数据分割成等长的数据块，加上源地址、目的地址、长度、校验码等辅助信息并封装成包（packet），然后发送给下一个节点。下一个节点收到后先暂存包，待输出线路空闲时再转

发给下一个节点，重复这一过程直到到达目的节点。每个包可单独选择到达目的节点的路径。这种方式称为存储转发方式。

包交换具有如下优点。

- 可以充分利用线路容量。可以利用多路复用技术，将多路数据复用后传送，也可以利用空闲时间传送其他用户的数据。
- 可以实现一对多、多对一的访问。在电路交换方式下，节点也是独占的，一个节点只能为一个用户服务，显然无法满足多对一、一对多的访问需求。
- 可以实现差错控制。由于对用户数据进行了封装，加入了校验信息，因此一旦传输出错，可借助校验信息发现错误并纠正错误。
- 便于实现格式转换。由于采用存储转发方式，因此每个中间节点收到包后可以根据需要进行格式转换后再发送出去，比如加密、压缩、编码格式转换等。
- 缓冲区易于管理，因为包的长度是固定的。
- 包的平均延迟更小，网络中占用的平均缓冲区更少。这是相对于非固定长度的报文交换而言的，利用排队论理论可以证明这一结论。
- 对应用的适应性更好。

包交换方式也有如下缺点。

- 加资源开销。例如增加辅助信息会导致时间和存储资源的开销，辅助信息的传送会增加带宽资源的开销。
- 增加缓冲延迟，同时因要额外传送辅助信息，也增加了传送时间。
- 多个包的顺序可能发生变更，需要额外的顺序控制机制。

包交换是一种通用的交换方式，可作为网络的基本交换方式。

1.5.3 广播交换

广播交换是指发送节点向接收节点发送信息时采用广播方式向网上所有节点发送信息，哪个（哪些）节点应该接收信息则根据数据包的地址信息决定。有的物理网络本身就是广播型的，采用广播交换具有优越性，比如无线网络、总线型网络、卫星网络等。广播交换一般与包交换联合使用。

此外，研究人员还提出过其他的交换方式，如长度更小的信元交换（信元长度为53B，用于ATM网络）、长度不固定的报文交换等，它们已经被淘汰，不再使用。

1.5.4 包交换的实现方式

包交换在具体实现时，有两种基本方式，即数据报方式和虚电路方式。

1. 数据报方式

（1）数据报方式的原理

数据报传输方式，将每个包作为一个独立的信息单元传送，在发送前不需要与目的

节点之间建立连接，在传输过程中也不考虑它与前面已发出的包以及与后面将要发出的包之间的顺序关系。包每经过一个中继节点时，都根据当时当地情况按照一定的算法为其选择一条最佳的传输路由。也就是说没有建立连接和释放连接的开销，每个包独立地到达目的节点，不保证可靠性和有序性，纠错和排序功能由上层完成，不保证服务质量，网络尽最大能力把包交付给目的节点。

如图 1-18 所示，主机 1 有四个包需要传输到主机 2，主机 1 将四个包发送到节点 A。节点根据目的主机的地址利用路由算法选择最佳的后继节点 B，并将包 P₀ 发送给节点 B。同理，节点 B 将 P₀ 发送给节点 D。这样构成了一条路径：主机 1 → A → B → D → 主机 2，同理，包 P₁ 也沿这条路径传输。当节点 A 发送完两个包后，由于网络状态发生变化，节点 B 拒绝接收新的包，这时节点 A 需要重新选择新的路径继续传输剩余的包。节点 A 选择节点 C，节点 C 选择节点 E，节点 E 选择节点 D，即可通过路径：主机 1 → A → C → E → D → 主机 2，传输剩余的两个包（P₂、P₃）。由于包沿不同的路径传输、节点处理能力不同，以及网络状态是不断变化的，使得节点 A 先发送的包有可能后到达，也有可能某个包在传输过程中丢失。

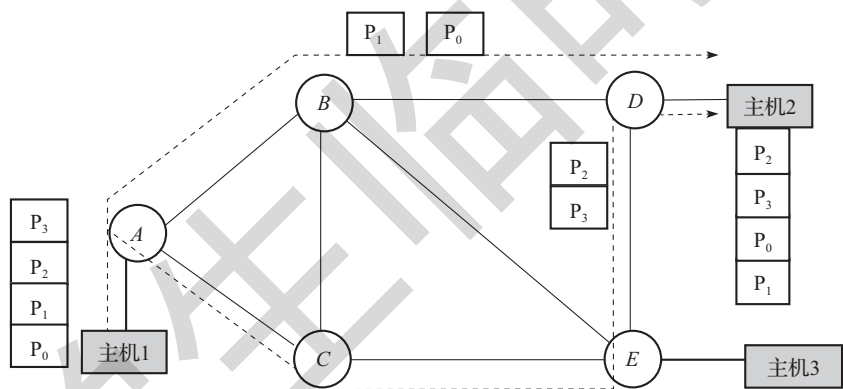


图 1-18 数据报传输方式

(2) 数据报方式的实现

每一个包都必须含有全称目的地址。网络的每个节点都保存一张由本节点到达其他节点的输出线选择表，该表的内容由路由选择算法填写。

例如，若对于某节点 X，从该节点可到达节点 A、B、C，其输出线选择表可能为

目的节点	输出线
A	线 1
B	线 2
C	线 1

数据传输过程可描述为：当节点收到一个带有全称目的地址的包时，根据包中的目的地址，查找本节点的输出线选择表，找出一条输出线将包转发出去。输出线选择表中只

记录一条输出路径,当存在多条转发路径时,由路由选择算法来确定使用哪条转发路径。

2. 虚电路方式

(1) 虚电路方式的原理

在源主机与目标主机通信之前,应首先在两者之间建立一个网络连接,即为虚电路(简称 VC)所有发往该目的主机的数据都将沿着所建立的虚电路传送。当通信结束时,释放该虚电路。

虚电路是相对电路而言的,在电路交换中面向连接的服务所请求到的是一条物理的电路连接,它以独占的方式使用物理电路连接。而虚电路则是在一条物理电路连接上实现复用,是一条逻辑电路,在一条物理电路上可以建立许多虚电路。

虚电路可分为永久虚电路和临时虚电路两种类型。永久虚电路是两个节点之间永久性连接的虚电路,不论发送还是接收包时,都不需要建立虚电路和释放虚电路的操作。因此,永久虚电路类似点到点的专线。临时虚电路也叫交换虚电路,是在源节点和目的节点之间暂时性连接的虚电路。它根据源节点的呼叫请求而建立,通信结束时释放。

图 1-19 表示虚电路的工作过程。假设主机 1 要传输四个包到目的主机 2。主机 1 首先发送与主机 2 建立连接的请求给节点 A,节点 A 将在网络中转发主机 1 的连接请求,并建立一条路径:主机 1 $\rightarrow A \rightarrow B \rightarrow D \rightarrow$ 主机 2,这条路径即为虚电路。主机 1 则将要发送的四个包沿这条虚电路按序传输,这四个包将按序到达主机 2。传输完成后,释放这条虚电路。这个过程类似于在电路交换网上的通信。但与电路交换的不同之处在于,包交换网采用的是存储-转发机制,每个包在通过网络的节点交换机时有一定的处理延迟。这条路径中所经过的链路并不是被主机 1 和主机 2 的通信独占,而是链路的复用。也就是说链路在为主机 1 与主机 2 提供服务的同时,也可为其他主机之间的通信提供服务,例如,当主机 3 与主机 2 之间进行通信时,节点 B 和 D 之间的链路被复用。

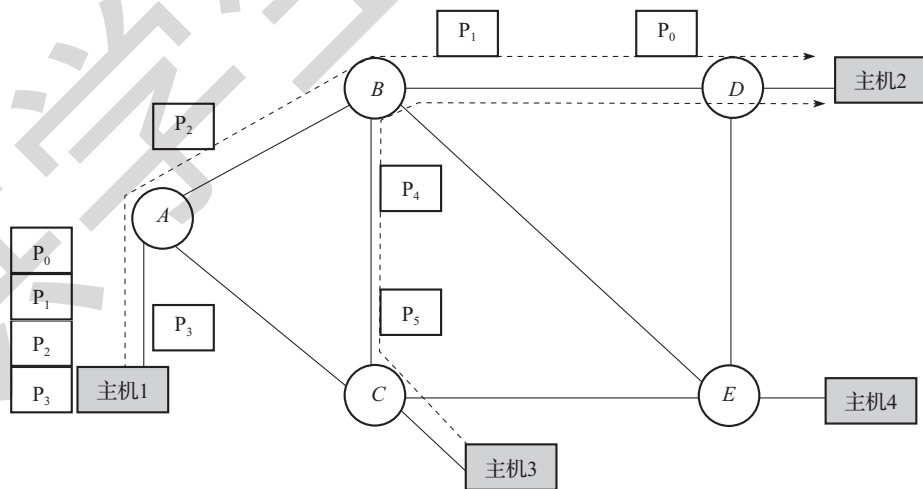


图 1-19 虚电路传输方式

在虚电路建立后，网络向用户提供的服务就好像在两个主机之间建立了一条穿过网络的通信管道。所发送的包都按发送的顺序进入管道，然后按先进先出的原则沿着此管道传送到目的主机。这样到达目的主机的包顺序与发送顺序一致，网络提供的虚电路服务对服务质量有较好的保证。

(2) 虚电路方式的实现

虚电路服务过程包括建立连接、数据传送和释放连接三个阶段。通过建立连接，在两个物理节点之间的一条线路上形成一个逻辑信道。在一条物理线路上允许建立多条逻辑信道，彼此之间通过缓冲区来区分。

为实现虚电路功能，要求所传送的包中包括一个虚电路号字段，同时每个节点需要保存一张虚电路入口、出口表，该表在虚电路建立过程中填写。经过节点 *X* 的入口、出口表可能为表 1-1 所示的内容。其中每行表示经过该节点的一条虚电路的入口和出口信息。

表 1-1 虚电路入口、出口表

入口		出口	
输入线	虚电路号	输出线	虚电路号
<i>A</i>	0	<i>B</i>	0
<i>A</i>	1	<i>C</i>	1

另外，每个节点应保存一张输出线选择表以供建立虚电路时路由选择之用。虚电路的实现过程如下：

- ① 接收到传输请求后，源节点发送一个呼叫请求包，其中包括初始虚电路号、源节点、目的节点等信息，并根据输出线选择表选择一条输出线将请求包发送出去。
- ② 呼叫请求包经过的每个节点，都在其入口、出口表上登记该虚电路的入口和出口，其中出口信息根据路由选择算法确定的输出线进行填写；输出虚电路号选择本节点当前未用的最小号码作为其值。
- ③ 在呼叫请求包到达目的节点后，由目的节点发送一个呼叫接受包给源节点，该包沿呼叫请求包经过的路线反向传送。源节点收到呼叫接受包后即认为虚电路已经建立，可进行数据传送。
- ④ 源节点封装数据包时，在每个数据包上填上虚电路号，并从虚电路入口、出口表中找到相应的输出线来发送数据包。
- ⑤ 数据包经过的每个节点在收到数据包后，根据数据包来自哪条输入线及其中的虚电路号查找虚电路入口、出口表，找出输出虚电路号，修改数据包中的虚电路号，并转发至对应的下一节点。
- ⑥ 数据发送完后，源节点发送一个呼叫清除包给目的节点，目的节点返回一个确认包。呼叫清除包经过的每一节点都清除相应虚电路的入口、出口信息并释放所分配的资源。

根据上述方法所建立的虚电路，允许虚电路号在虚电路的各链路上不同。例如，对于图 1-20 所示的网络，假定有通信请求 $A \rightarrow E$ 和 $B \rightarrow F$ ，分别建立了虚电路 $A-C-D-E$ 和 $B-C-D-F$ ，并假定 A 和 B 在建立虚电路时都选 0 作为虚电路号，则 C 的入口、出口表为 $A/0 \rightarrow D/0$ 和 $B/0 \rightarrow D/0$ 。 D 的入口表本应该有两条信息，分别对应两条虚电路，但由于现在入口信息都为 $C/0$ 而无法区分数据包是从哪条虚电路来的，也就无法确定转发路径。解决这一问题的办法是对于任一节点，在建立虚电路时，总选择当前未使用的最小号码作为输出虚电路号。采用该方法后，对于上述例子， C 的入口、出口表就变为 $A/0 \rightarrow D/0$ 和 $B/0 \rightarrow D/1$ ， D 的入口、出口表为 $C/0 \rightarrow E/0$ 和 $C/1 \rightarrow F/0$ 。经过公共节点 C 、 D 的多条虚电路即可相互区分。

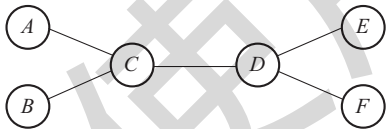


图 1-20 实现虚电路的简单网络

3. 虚电路与数据报方式的比较

虚电路服务与数据报服务的特性差异如表 1-2 所示。

表 1-2 虚电路与数据报的比较

特点	虚电路	数据报
目的地址	开始建立时需要	每个包都需要
错误处理	网络负责	主机负责
流量控制	网络负责	主机负责
拥塞控制	网络实现	难
路由选择	只需在建立时进行一次	每个包都需独立进行
包顺序	按发送顺序到达	到达顺序不确定
建立与释放连接	需要	不需要
服务方式	面向连接	无连接
适用条件	数据量大、实时性要求较低、可靠性高的系统	数据少（多为 1 个包的情况）、实时性高、可靠性低的系统

1.6 网络设计模式

1.6.1 网络体系结构分层设计模式

从前几节的内容明显可以看出，设计功能完善、性能良好、适应广泛的计算机网络具有高度的复杂性。为降低设计与实现难度，提高网络的通用性、正确性、有效性和性能，必须采用有效的设计方法，而网络体系结构就是最有效的方法。

网络体系结构是指构成计算机网络的各组成部分及计算机网络本身所必须实现的功能的完整定义。由于体系结构通常都具有可分层的特性，因此网络体系结构也可转化成网络层次结构。这样，网络体系结构就更加直观，是指计算机网络中的层次划分、各层

协议以及层间接口定义的集合。

计算机网络分层的基本原则是：

- 各层之间的接口清晰自然，易于理解，层间通信尽可能少。
- 各层功能的定义独立于具体的实现方法。
- 保持下层对上层的独立性，单向使用下层提供的服务。

网络体系结构规定：相邻层次之间才能发生交互，不能跨层进行交互。但是，针对新型网络的研究表明，只有相邻层次之间可交互的规定在一定程度上限制了新方法的应用和网络性能的改善。例如，在光网络中，通信路径不仅取决于是否存在光纤，还取决于光纤上是否有可供使用的波长，而这两个属性分属不同的层次，只有联合设计才能获得更优的结果，这种研究方法称为跨层设计。对无线网络，同样应将寻找传输路径（无线信号能覆盖）与分配频道（频率）联合进行跨层设计，才能取得更优的结果。分层带来的优点是易于理解、易于实现、易于更新（可只更新单个层次）和易于抽象。

在分层时应考虑层次的清晰程度与运行效率间的折中和层次数量的折中。层次越多，每层的定义就可能越清晰，其实现就可能越容易，但其总体运行效率可能就越低。

依据一定的原则，通常将网络分成多个层次，从最低层到最高层依次称为第1层，第2层，…，第 n 层，通常还为每层起一个特定的名称，如第1层称为物理层、第2层称为数据链路层等。每层由完成给定功能的实体（硬件或软件）组成，第 n 层的实体可记为 n 实体。

下层对上层提供服务的接口称为服务访问点（SAP），它也是上层调用下层功能的接口， n 层实体与 $n-1$ 层实体之间的服务访问点记为 $(n-1)$ SAP，其对应关系如图1-21所示。

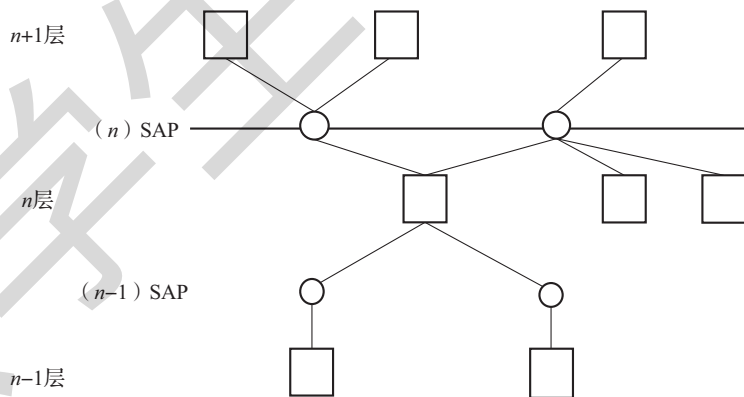


图 1-21 层次、实体与服务访问点的对应关系

1.6.2 协议、接口与服务

1. 协议、接口与服务设计

接口是指同一系统内部两个相邻层次之间的交互规则。在传统研究中，每一层只为

与其紧相邻的层次定义接口，不能跨层定义接口。可将接口看作上面所说的服务访问点。

服务是下层对上层提供的功能调用或实现。

协议是指通信双方（不同系统）实现相同功能的对应层之间的交互规则。协议与接口的关系如图 1-22 所示。协议由语法、语义和时序三部分构成。语法规定传输数据的格式，语义规定所要完成的功能，时序规定执行各种操作的条件、顺序关系等。

一个完整的网络协议通常应具有如下功能：

- 线路管理：建立与释放连接。
- 寻址：标识与识别主体。
- 差错控制：对传输过程中可能出现的差错进行识别和处理。
- 流量控制：控制网络流量，使其与传输能力相匹配。
- 路由选择：在源节点与目的节点之间选择传输信息的路径。
- 同步控制：在不同的节点之间实现位、字符、帧、状态的同步。
- 数据分段与装配：将用户的数据分解为较小的传输单元或进行相反的操作。
- 排序：对多个传输单元进行顺序控制。
- 数据转换：按需进行数据格式变换。
- 安全管理、计费管理等。

2. 服务模式设计

服务是指为紧相邻的上层提供的功能调用或实现，每层只能调用紧相邻的下层提供的服务。服务的提供与调用如图 1-23 所示。

计算机网络的服务特性一般在三个维度进行定义：通信双方在通信前是否先建立联系（连接）、在通信后是否给出应答、在通信过程中是否提供可靠性机制。

（1）面向连接的服务与无连接的服务

面向连接的服务是指在通信之前，通信双方需先建立连接，然后才能开始传送数据，数据传送完成后需释放连接。建立连接时可能需要分配相应的资源，如缓冲区、带宽等（依具体任务而定），以保证通信能正常进行。比如，打电话就是面向连接的服务。

无连接的服务是指双方通信前不事先建立连接，直接发送数据。比如，平常写信交由邮局投递的过程就是无连接的服务。

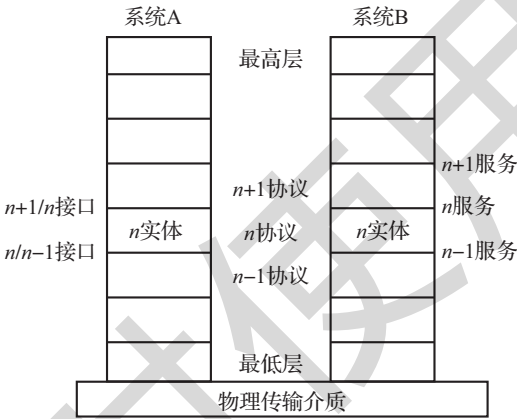


图 1-22 协议与接口的关系

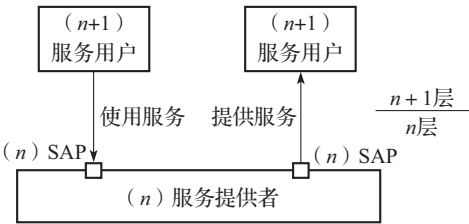


图 1-23 服务的提供与调用

(2) 有应答服务与无应答服务

有应答服务是指接收方在收到数据后向发送方给出相应的应答, 该应答由传输系统内部自动实现, 而不是由上层用户实现。所发送的应答可以是肯定应答(确认), 也可以是否定应答(否认), 通常在接收到的数据是错误时发送否定应答。例如, 文件传输服务就是一种有应答服务。

无应答服务是指接收方传输系统收到数据后不自动给出应答。若需应答, 由高层用户给出。

有无应答的区分取决于传输系统自身, 而不是整个网络功能是否有应答。举一个日常的例子, 我们常见的邮递系统就是一个无应答的系统, 发件人判断所发送的邮件是否被收到, 要看收件人的回信。如果将邮递系统改为有应答的系统, 其工作过程就是: 发件人将邮件交给邮递机构, 邮递机构一步一步地传递邮件直到最后将其交给收件人, 收件人签收后, 邮递机构再沿相反的方向逐个报告说邮件已经送到, 直到最后发件人接到邮递机构的报告, 发件人据此知道邮件是否发送成功, 而不需要收件人发送确认通知。

(3) 可靠服务与不可靠服务

可靠服务是指网络具有检错、纠错、应答等机制, 能保证数据被正确、可靠地传送到目的地。

不可靠服务是指网络不能保证数据被正确、可靠地传送到目的地, 网络只是尽量正确、可靠地传输数据, 是一种尽力而为的服务。

对用户而言, 如果网络是不可靠的, 则很难放心使用, 尤其是一些关键应用, 如银行业务。而网络本身由于各种因素, 在传输过程中出错总是难免的, 因此对于提供不可靠服务的网络, 其网络的可靠性、正确性就要由应用或用户来保障。例如, 用户收到信息后要判断信息的正确性, 如果不正确, 用户应把出错信息报告给信息的发送者, 以便双方协同纠正错误。通过应用或用户的这些措施, 把不可靠的服务变成可靠的网络。

其实, 不可靠服务也是有存在价值的。例如, 通过网络看电影时, 可能每秒传送 30 帧数据, 只要正确率比较高, 偶尔传错或丢失一帧数据, 并不会影响我们对电影的理解。如果要保证数据绝对正确, 可能需要很多保证机制, 这将大大增加网络的负担, 显著降低网络的效率, 甚至导致播放不流畅。

3. 服务质量

网络服务质量 (Quality of Service, QoS) 是指网络提供优先服务的能力。网络服务质量通常用一些参数来衡量, 如延迟、延迟抖动、带宽、出错率、丢包率等。不同的应用对网络服务质量的要求不尽相同。例如, IP 电话需要延迟时间小, 同时延迟抖动小, 对于网络电视, 延迟抖动是非常关键的参数。

现在的因特网基本上不能提供服务质量保证, 它所实现的是一种尽力而为 (best effort) 的服务, 即网络尽量以最好的质量传送数据, 但没有任何承诺和保证。因此, 因特网不能满足很多应用的要求, 这就需要提供额外的服务质量保证机制以满足用户的需求。

除可量化的 QoS 外，用户实际感受到的网络质量是一种更直观的评价，称为体验质量（Quality of Experience, QoE）。QoE 一般通过网络服务的延迟（等待时间）、是否卡顿、播放连续性、画面质量等直观体验对网络进行评价。

1.6.3 数据的封装与传输

在网络中，为每一层定义传送数据的单位，其名称、大小和含义各有不同。几个主要的术语及其含义如下。

- 服务数据单元（SDU）：为完成用户所要求的功能而应传送的数据，即该层的输入。第 n 层的服务数据单元记为 n -SDU。
- 协议控制信息（PCI）：控制协议操作的信息，是为完成数据传输而附加的信息，类似于信封。第 n 层的协议控制信息记为 n -PCI。
- 协议数据单元（PDU）：协议交换的数据单位，即实际传送的数据，该层的输出。第 n 层的协议数据单元记为 n -PDU。

三者之间的关系为 n -SDU + n -PCI = n -PDU = $(n-1)$ SDU，其变换过程如图 1-24 所示。其中，H4 表示第 4 层的协议控制信息，俗称头部信息（或称首部信息），H3 等以此类推。

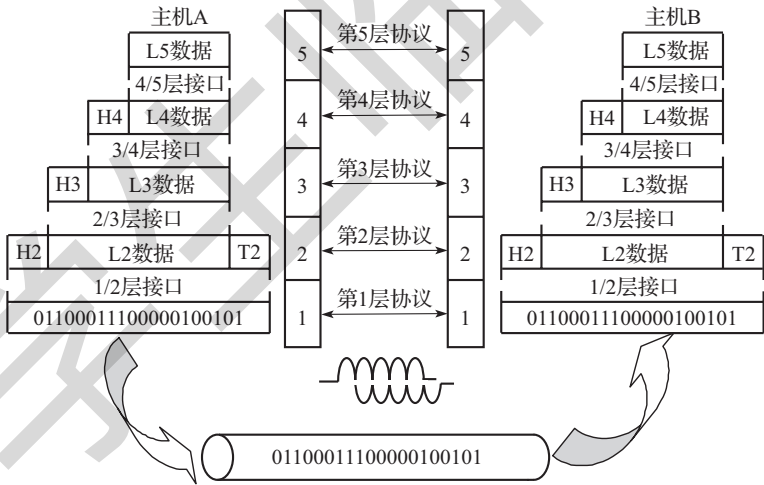


图 1-24 数据单元关系

在发送端，每一层都对从上一层接收到的数据进行一次封装（加上协议控制信息，俗称报文头部信息），层层封装，直到物理层将整个报文进行适当的编码后发送出去。在接收端，每一层都对从下一层递交上来的数据进行解封，去掉本层的协议控制信息（报文头部信息），提交给上一层，直到到达最上层。层层封装是层次结构的一种典型设计方法，虽然增加了开销，但有利于分层实现。

1.6.4 ISO 分层模型示例——OSI 参考模型

国际标准化组织（ISO）于1978年提出了一个网络体系结构模型，称为开放系统互连（Open System Interconnection, OSI）参考模型。OSI参考模型有7层，从低到高依次称为物理层、数据链路层、网络层、传输层、会话层、表示层和应用层，如图1-25所示。用户终端设备（简称端设备，指计算机等）包括全部7层，网络设备（路由器等，属于通信子网）包括较低的1~3层。

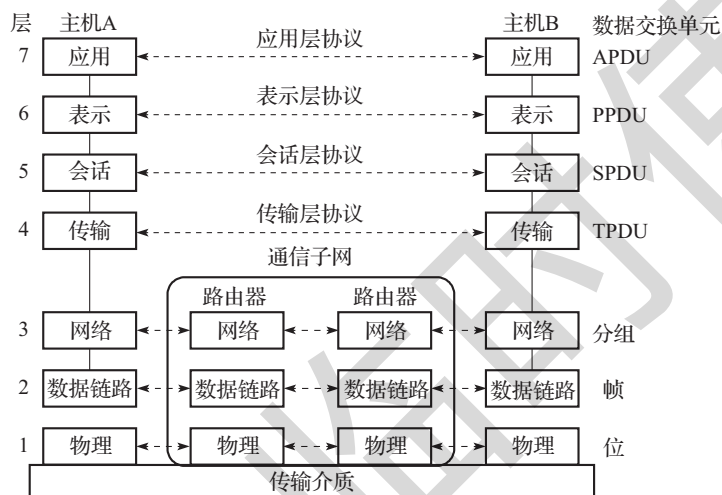


图 1-25 OSI 参考模型的层次结构

下面介绍 OSI 参考模型中各层的主要功能。

- 物理层：在链路上透明地传输位。需要完成的工作包括线路配置、确定数据传输模式、确定信号形式、对信号进行编码、连接传输介质。为此定义了建立、维护和拆除物理链路所需要的机械特性、电气特性、功能特性以及规程特性。
 - 机械特性：规定接口部件的形状、大小、引脚数目及其排列方式。
 - 电气特性：规定信号的表示方式及其参数、信号间的相互关系及运行要求。比如，用什么信号形式（脉冲、正弦信号等）、电压值及其含义（几伏表示1/0等）、信号时长（宽度）等。
 - 功能特性：规定每根引脚（线）的功能。
 - 规程特性：规定为完成指定的功能应执行的操作及其时序。
- 数据链路层：把不可靠的物理信道变为可靠的逻辑信道。为此将一连串的数据位组成帧，在链路上提供点到点的帧传输，并进行差错控制、流量控制等。
- 网络层：在源节点和目的节点之间进行路由选择、拥塞控制、顺序控制并传送包，保证上层报文（分解成多个包）的正确性。网络层控制通信子网（即在计算机等用户设备之外为实现数据传输而附加的功能系统）的运行，因而它又被称为通信子网层。

- 传输层（也可称为传送层、传达层、运输层）：提供端 - 端（应用进程）间可靠的、透明的数据传输，保证报文顺序的正确性和数据的完整性。
- 会话层（也可称为会议层、会晤层）：为完成一次通信而进行的一次联系过程称为会话（session）。建立通信进程的逻辑名字与物理名字之间的联系，提供进程之间建立、管理和终止会话的方法，处理同步与恢复问题。
- 表示层（也可称为表达层）：实现数据转换（包括格式转换、压缩、加密等），提供标准的应用接口、公用的通信服务、公共数据表示方法。
- 应用层：对用户不透明的各种服务，如 Email、FTP。

OSI 参考模型是一种基于学术研究的体系结构模型，比较完整，但也非常复杂，冗余较多。除了第 1 层和第 2 层有实现外，其余层次基本没有实现。OSI 参考模型现在已不使用，只作为评价网络的一种参考。

1.6.5 因特网分层模型示例——TCP/IP 参考模型

美国国防部高级研究规划署（ARPA）1969 年在研究 ARPANET 时提出了一个网络实现模型，并在 1970 年提出了网络控制协议（NCP）。后经多年发展，ARPA 于 1974 年提出了 TCP，1981 年提出了 IP（从 TCP 中分离出来），1983 年对 ARPANET 通信协议进行标准化，从 NCP 切换到 TCP/IP，1984 年美国国防部将 TCP/IP 确定为网络标准。1984 年，美国国家科学基金会（NSF）决定使用 TCP/IP 建设 NSFNET，主要是民用，同时将其与 ARPANET 连通。ARPANET 于 1989 年被关闭，1990 年被撤销，之后 NSFNET 成为因特网的骨干。

网络数据传输过程的管控包括两种基本理念：一种是以网络为中心，提供可靠、完善的机制，减少用户端的参与；一种是以用户端为中心，网络只提供基本的、尽力而为的机制，降低网络的复杂度，让网络的管控更多在用户端完成。TCP/IP 采用的是第二种理念。网络执行简单的网际数据传输，由互联网层（IP）实现，而将与应用相关的功能、传输控制等都放在用户端（应用层、传输层）上实现。

TCP/IP 是逐步发展起来的，现在所说的 TCP/IP 参考模型是发展较为成熟的架构，各层从低到高依次为网络主机接口层（简称网络接口层）、互联网层、传输层、应用层，如图 1-26 所示。

应用层、传输层、互联网层都定义了相应的协议和功能，但网络接口层没有明确地定义其功能、协议和实现方式，可由用户自主实现或利用其他已有的实现。

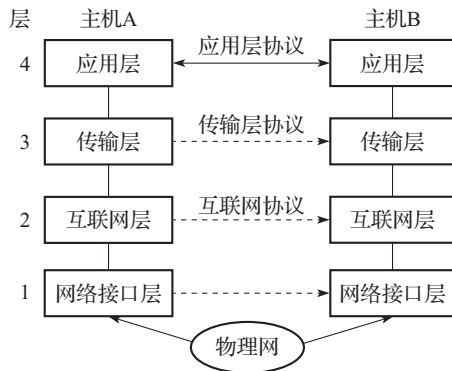


图 1-26 TCP/IP 层次结构

- 应用层：定义与应用相关的共性功能，主要功能有域名解析、WWW 服务、邮件服务、文件传输、网络管理等。
- 传输层：定义端到端数据传输功能，主要有传输控制协议（TCP，实现端－端可靠数据传输）、用户数据报协议（UDP，实现端－端不可靠数据传输）。
- 互联网层：定义网际数据传输功能，主要有网际协议（IP，实现主机－主机数据报传输）、控制报文协议（ICMP，实现网络内部管理）等。
- 网络主机接口层：没有明确的定义。

TCP/IP 是一种基于技术实现的网络体系结构模型，它并不完整（网络主机接口层都未定义，由用户自主决定），某些功能的层次归属也不尽合理，后续增加的有些功能甚至未遵守上下层之间的调用规则。

TCP/IP 最关键的功能协议及其关系如图 1-27 所示，所有信息都要用 IP 进行传输。TCP/IP 体系结构也被称为沙漏结构、细腰结构。

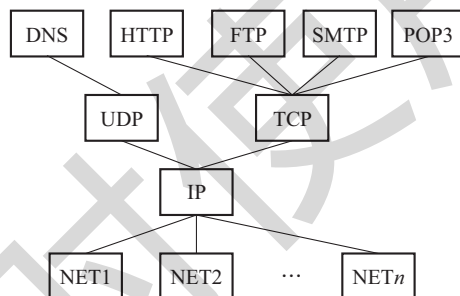


图 1-27 TCP/IP 体系结构

基于 TCP/IP 层次模型的网络也是层次结构，由本地网络、地区骨干网、国家骨干网、国际互联网等多级组成，称为网络的网络，如图 1-28 所示。

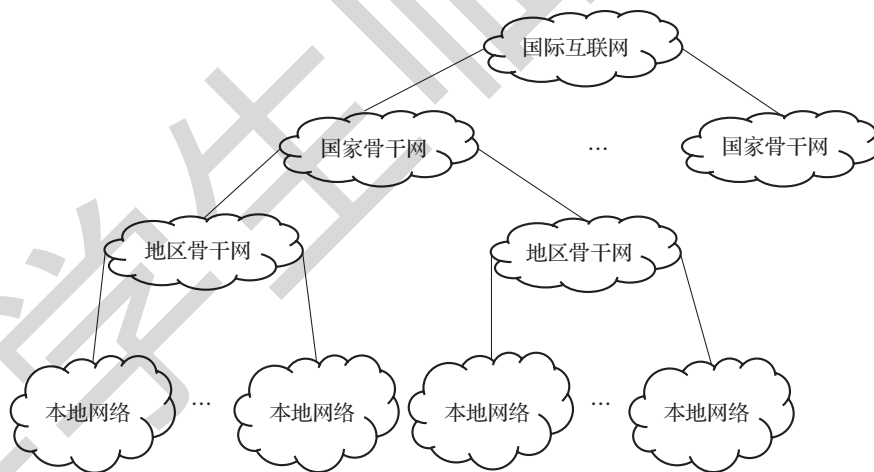


图 1-28 网络的网络

1.6.6 虚拟层次模型

上述两个模型存在差异，为了方便，现在几乎所有的书籍在介绍网络时都采用一种结合上述两种模型的虚拟模型，该模型将网络分为 5 个层次，依次为：

- 第 5 层：应用层。

- 第4层：传输层。
- 第3层：网络层。
- 第2层：数据链路层。
- 第1层：物理层。

下面约定各层传输的数据单元名称的定义。

- 应用层：统称用户数据或用户消息，也称报文（message）或消息。
- 传输层：报文段（segment），通常简称为报文。报文段的长度不超过 64KB，当用户数据太长时，需要分成多个报文段传送。
- 网络层：包（packet）或分组。包的长度较小，比较典型的长度是 1KB 左右（随网络和协议的不同而不同，比如 IPv4 包的默认长度为 576B，IPv6 包的默认长度为 1280B）。一个报文段可能会被分成多个包发送。包在 IPv4 中被称为 IP 数据报（datagram）。
- 数据链路层：帧（frame）。帧与包之间基本上是一一对应的（包被分片时除外）。
- 物理层：位或比特（bit）。

1.6.7 从单网到多网互联的增量设计

网络设计的另一种方法是先设计简单的小网络，然后将小网络互联起来组成覆盖范围广、接入用户多的网络，最终的网络是一种层次化的、网络的网络（network of networks）。对小网络的设计，遵循前述的网络体系结构模型方式。

这样的设计模式从最开始就要同步设计可扩展性方案。

1.7 计算机网络构成要素设计

不论从网络设计者的角度看，还是从网络使用者、网络运维者的角度看，计算机网络都是由网络硬件、网络软件和网络应用三大类要素构成的。因此，设计计算机网络就是设计这三大类要素并将其有机地组合在一起。

1.7.1 计算机网络硬件

计算机网络硬件是组成计算机网络的基础，包括各类网络设备。为实现不同的连网需求和功能，需要设计不同的网络设备。网络设备泛指组成通信子网的各类设备，常见的网络设备包括网卡、交换机、路由器、无线接入点（无线路由器）、无线控制器、光调制解调器等基本的网络设备，也包括防火墙等网络安全设备，不包括计算机、服务器等属于资源子网的信息处理设备。

交换机（switch）是实现计算机连网的基本设备，它将多台计算机连接在一起组成基本的计算机网络。将多个网络互联在一起组成更大的网络时，通常需要使用路由器

(router)。交换机是在数据链路层实现互联的设备(三层交换机也能在网络层实现互联),路由器是在网络层实现互联的设备。

随着需求的变化和新的理论、技术的出现,需要研究、设计、制造与之相适应的新网络设备。

1.7.2 计算机网络软件

网络软件是实现网络功能时不可缺少的软件环境,是在网络环境下运行、使用、控制和管理网络工作以及通信双方传输信息的核心组件。这里所说的网络软件特指控制和管理网络运行、提供网络通信、管理和维护共享资源的基础软件,包括具有网络功能的网络操作系统、网络协议软件、支撑网络应用的共性基础软件、网络管控软件和网络编程平台软件等,也包括嵌入网络设备中的一些嵌入式软件。

- 具有网络功能的网络操作系统。早期的操作系统没有网络功能,但现在的操作系统都集成了网络功能,TCP/IP已集成到操作系统内核中。除了网络通信功能外,还有大量网络服务功能,如Web服务器、FTP等。
- 网络协议软件。协议是网络的核心和灵魂,协议最终主要通过各种软件实现,实现各种协议的软件统称为协议软件。网络体系结构的每一层都有自己的协议,相应地都有对应的协议软件。研究协议软件及其高效实现技术是网络研究的重要内容。
- 支撑网络应用的共性基础软件。共性基础软件为各种网络应用提供支撑,使应用开发人员不必开发每一项网络功能,只需聚焦应用本身。共性基础软件主要包括各种API、加解密软件包、编解码函数等。
- 网络管控软件。网络管控软件是对网络运行状况进行信息统计、监视、警告和报告,对网络进行配置和管理的软件系统。网络管控软件既在管理计算机上运行,也分布在各类网络设备(包括嵌入式设备和软件)中。
- 网络编程平台软件。网络编程平台软件主要用于开发各种其他网络系统软件和网络应用,主要是指支持网络软件开发的各种开发工具,如Socket API、RPC等,也包括应用开发工具,如HTML、JavaScript、JavaEE、Eclipse、Xcode、Android Studio等。

1.7.3 计算机网络应用

网络应用是指为某一应用目的而开发的网络应用软件,它为用户提供一些实际的应用功能。网络应用软件既可用于管理和维护网络本身,也可用于某一个业务领域。

网络应用软件可以根据网络的组建目的和业务的需要进行研制、开发。网络应用软件有通用和专用之分。通用网络应用软件适用于较广泛的领域和行业,如网络版文字处理软件、即时通信软件、视频会议软件、人事管理软件等。专用网络应用软件只适用于

特定的行业和领域，如银行核算软件、铁路控制软件、军事指挥软件、天气信息软件等。除了在计算机上运行的各种应用软件之外，还有在智能终端（如智能手机、平板计算机、智能手表、手环等）上运行的各种应用软件。另外，还有各类云计算软件、边缘计算软件等。

一个实用的、具有较大效益的计算机网络，除了配置上述各种软件外，通常还应在网络协议软件与网络应用系统之间建立一个完善的网络应用支撑平台，为网络用户创造良好的运行环境和开发环境，比如各种 SDK、IDE 等。

习题

- 1-1 分别从计算机系统设计 and 软件设计的角度，尽量清晰地给出计算机网络的设计需求。
- 1-2 分析 C/S 模式在各种情况下的性能。达到极限性能的条件是什么？
- 1-3 除了教材中所列类型之外，请给出计算机网络更多可行的拓扑结构，并分析其优缺点。
- 1-4 分析请求 / 应答模式与流式模式在保证传输正确性方面的差异。
- 1-5 分析 OSI 参考模型的成功和失败之处。
- 1-6 分析 TCP/IP 参考模型的主要不足。
- 1-7 如果网络上的数据像日常的包裹一样只封装一次，试分析其可行性及优缺点。
- 1-8 在多条转发网络中，假定在每条链路（含两端节点）上出错的概率为 p ，数据要经过 n 条链路到达目的节点，计算数据包被正确传到目的节点的概率。如果应答信息出错的概率相同，计算双方都认为信息被成功传输到目的节点的概率。
- 1-9 中国天宫空间站轨道高度约为 400km，下行（空间站到地球）的数据率为 1.2Gbps，上行（地球到空间站）的数据率为 6Mbps，假设空间站借助天链卫星可持续与地球通信，天地间采用包交换方式发送数据，每个包的载荷最大为 1KB，请问：空间站确信 4GB 的文件成功发送到地球并被正确接收大概需要多长时间？

基本网络设计

要把分散的计算机连接成可共享信息的简单计算机网络，应利用已有的技术，并研究尚缺少的新理论、新技术、新设备。根据第 1 章所确定的基本目标和原则，确定了基本的网络构型和功能后，需要研究以下问题：

- 有什么传输介质能用于连接计算机？选用标准和使用条件是什么？
- 在传输介质上如何表示数据？
- 如何构成基本的数据传输系统，以实现数据在计算机之间的传输？
- 如何保证在传输系统上正确地传输数据？如何纠正可能出现的错误？
- 用什么方法可以提高传输性能以满足不断增长的性能需求？
- 可设计哪些具体形态的网络，使其更好地满足应用需求？要为这些网络设计什么核心技术？解决哪些关键问题？

2.1 传输介质

2.1.1 传输介质的类别与特性

传输介质是数据传输系统中在发送设备和接收设备之间的物理链路，数据都是通过电磁信号（电磁波）的形式在介质上传输。按照传输介质的形态，可以将其分为有线传输介质和无线传输介质两大类。有线传输介质主要有双绞线、同轴电缆、光纤、光电混合缆等。无线传输介质主要有微波、无线电、红外线、激光等。

研究表明，不同介质允许在其上通过的电磁波的频率是不同的，表现出的性能也就不同。图 2-1 是不同介质上能传输的电磁信号的频率范围（频谱）。计算机网络应根据数据传输需求选择合适的传输介质。

无线频率在各个国家都是政府管控的紧缺资源，除个别频段外，都需要政府授权才能使用。

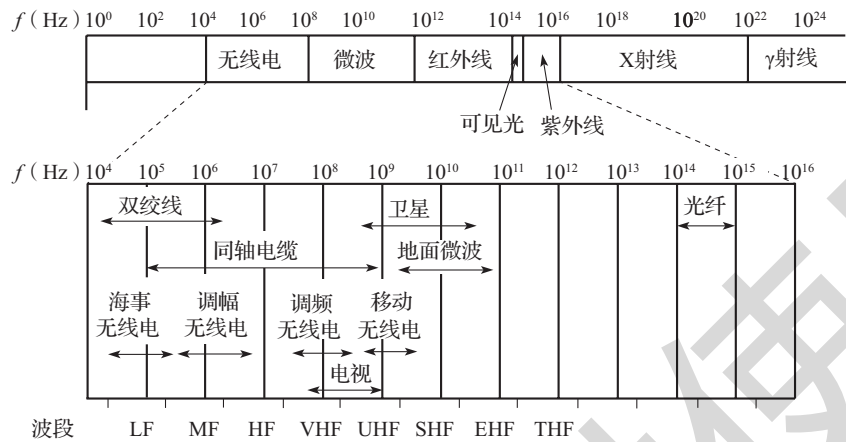


图 2-1 传输介质与频谱

2.1.2 有线传输介质

有线传输介质属于导向介质 (guided media), 主要有双绞线、同轴电缆、光纤和光电混合纤。

1. 双绞线

把两根包有绝缘层的铜导线绞合在一起就构成了双绞线。绞合 (而非平行的两根导体) 可减少相邻导线的电磁干扰, 降低因平行导体之间形成分布电容所造成的充放电影响, 其结构如图 2-2a 所示。

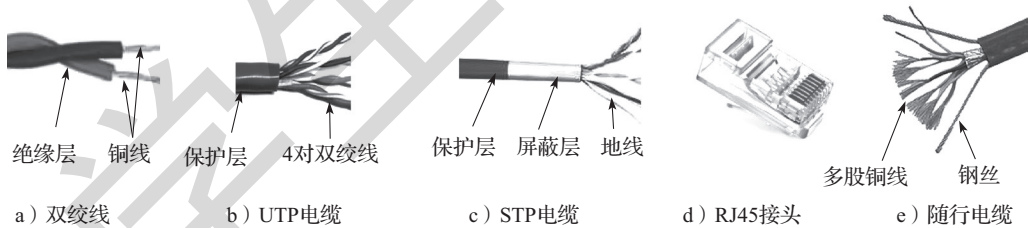


图 2-2 双绞线结构

计算机网络不宜直接使用简单的双绞线作为传输介质, 而是使用双绞线电缆作为传输介质。这样的双绞线电缆由 4 对双绞线构成, 如图 2-2b 所示, 称为非屏蔽式双绞线 (UnShielded Twisted Pair, UTP) 电缆。为了提高双绞线电缆的抗电磁干扰性能, 可以在双绞线的外面再加上一个金属屏蔽层 (金属丝编织网或金属薄膜), 再加一根金属导线作为地线, 同时增加抗拉强度, 称为屏蔽式双绞线 (Shielded Twisted Pair, STP) 电缆, 如图 2-2c 所示。双绞线电缆使用 RJ45 接头 (如图 2-2d 所示) 连接电缆和设备。图 2-2e 所示为随行电缆, 用于电梯等特殊环境。

计算机网络可使用双绞线电缆以点到点的连接方式，构成星形网络。

1991年，美国电气工业协会（EIA）和电信工业协会（TIA）联合发布了标准 EIA/TIA-568，这个标准规定用于室内传送数据的非屏蔽双绞线和屏蔽双绞线的标准。随着网络上数据传送速率的不断提高，EIA/TIA 在 1995 年将布线标准更新为 EIA/TIA-568A，此标准规定了从 1 类线到 5 类线的 5 个种类的 UTP 标准，2001 年发布标准 TIA/EIA-568B。568A 规定了两台计算机互连或 ADSL MODEM 与 HUB 连接，568B 规定了计算机与交换机或 HUB 相连，其连接时的线序也不同，568A 的排线顺序为白绿、绿、白橙、蓝、白蓝、橙、白棕、棕，568B 的排线顺序为白橙、橙、白绿、蓝、白蓝、绿、白棕、棕（现在工程布线用的标准），即 1 与 2、3 与 6 分别配对，分别用于发送与接收。2009 年发布 ANSI/TIA-568C。5 类线和 3 类线相似，但绞合螺距更小（3 类线约为 7.5 ~ 10cm，5 类线约为 0.6 ~ 0.85cm），交互感应更少，更适用于高速计算机网络通信。主要的双绞线电缆的性能如表 2-1 所示。

表 2-1 双绞线电缆的性能

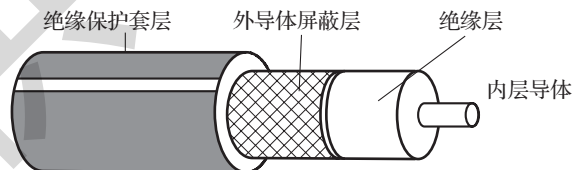
类别	带宽	数据率	类别	带宽	数据率
3 类	16MHz	10Mbps	超 5 类	100MHz	1Gbps
4 类	20MHz	16Mbps	6 类	250MHz	10Gbps
5 类	100MHz	100Mbps	7 类	600MHz	10Gbps

STP 和 UTP 电缆主要用于星形网络，每段长度小于等于 100m，最多可级联四级（5 段），因此用双绞线电缆组建计算机网络，最远距离一般不超过 500m。

超 5 类及以上的双绞线，可以同时为 POE 设备（如 Wi-Fi 的 AP）供电，模式 A 规定使用 4/5（正）、7/8（负）供电，模式 B 规定使用 1/2（正）、3/6（负）供电，电压为直流 48V。

2. 同轴电缆

同轴电缆由内层导体、绝缘层、外导体屏蔽层以及绝缘保护套层所组成，如图 2-3 所示。内层导体可以是铜线、铜质镀银线、银线、铝质镀铜线等。由于外导体屏蔽层的作用，同轴电缆具有很好的抗干扰特性。



a) 同轴电缆的结构



b) BNC连接器

图 2-3 同轴电缆及其连接器

同轴电缆可以传输各种信号，可以按点点串联方式连接，也可以按多点并联方式连接。同轴电缆使用 BNC 连接器（如图 2-3b）进行连接。

按阻抗数值的不同，使用较多的同轴电缆有 50Ω 电缆、 75Ω 电缆等。

50Ω 是指在组网时需要在电缆两端的内层导体和外导体屏蔽层之间连接一个 50Ω 的电阻。 50Ω 同轴电缆主要用于在数据通信中传送基带数字信号，又称为基带同轴电缆，曾在早期局域网中得到广泛应用。

根据线径的不同，又分为粗电缆（直径为 1.016cm ）和细电缆（直径为 0.508cm ）。

细电缆用于计算机局域网组网时单段理论长度小于 300m ，工程上实际小于 185m ，每段上可连接的计算机不超过 30 台，最多 5 段，因此用细电缆的网络最远 925m 。细电缆使用 BNC 连接器连接计算机，组成总线型网络。

粗电缆用于计算机局域网组网时单段理论长度小于 500m ，每段上可连接的计算机不超过 100 台，最多 5 段，因此用粗电缆的网络最远 2500m 。不论细电缆还是粗电缆，都是用中继器将多段电缆串联以增大网络的覆盖范围。

75Ω 、 93Ω 、 150Ω 等同轴电缆主要用于模拟传输系统，是早期有线电视系统 CATV 使用的传输电缆。现在国际上对 CATV 电缆的最新研究将大幅度提升其带宽，有可能拓展其用途，延续其生命。

3. 光纤

光纤就是能导光的玻璃纤维，利用光纤传递光脉冲进行通信就是光纤通信，可以假定有光脉冲表示 1，无光脉冲表示 0。光纤结构如图 2-4 所示。

光纤由三层组成，内层是导光的玻璃纤维，中间层是一层隔离光的隔离层，外层是保护层。内层和中间层的折射率不同，使得在内层传输的光几乎不会发生折射而进入中间层，而以全反射的方式沿内层纤维传播，外面的光也不会进入内层。目前光纤使用的波长主要有三个波长段： $0.85\mu\text{m}$ 、 $1.3\mu\text{m}$ 、 $1.55\mu\text{m}$ 。

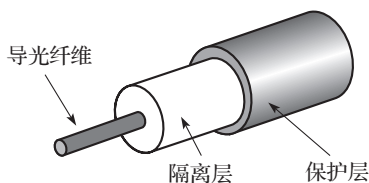


图 2-4 光纤结构

光纤具有如下显著特点：

- 带宽大、传输速率高。光纤的频带约为 $10^{14} \sim 10^{15}\text{Hz}$ ，按奈奎斯特定律，其极限数据率可达约 1800Tbps 。光纤的传输能力还没有被完全开发出来，目前的理论和技术已经使光纤的数据率超过了 560Tbps 。
- 传输损耗小，传输距离长，理论上可达 300km ，适用于远距离传输。
- 不受电磁干扰、防腐和不会锈蚀。
- 无串音干扰，保密性好。

光纤主要用于点到点的串联连接。光纤按传输模式可分为多模光纤和单模光纤。

(1) 多模光纤

多模光纤利用光的全反射特性来导光。多模光纤的直径较大，光源发出的可见光定

向性稍差，光能以不同的角度进入纤芯，因此，存在许多条不同角度入射的光线（多个模）在一根光纤中传输，如图 2-5 所示。光脉冲在多模光纤中传输时会逐渐展宽，造成失真，因此多模光纤只适合于近距离传输，通常用作室内的传输介质。多模光纤的内径主要有 $50\mu\text{m}$ 、 $62.5\mu\text{m}$ 、 $100\mu\text{m}$ 等规格，外径有 $125\mu\text{m}$ 、 $140\mu\text{m}$ 等规格。

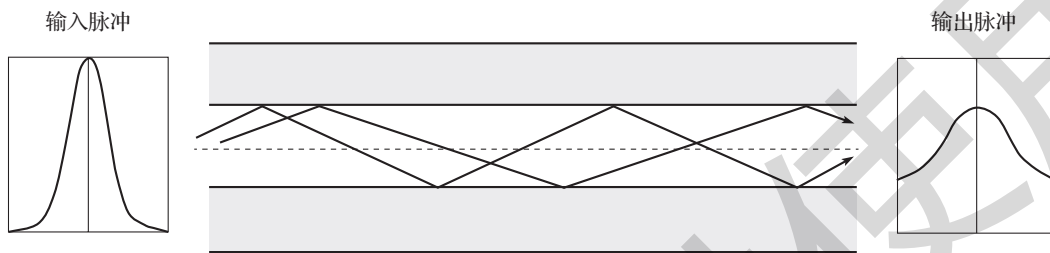


图 2-5 多模光纤

(2) 单模光纤

如果光纤的直径减小到只有一个光的波长大小，则光纤就像一根波导那样，它可使光线沿直线传播，而不会产生多次反射。单模光纤就是按这样的原理制成的，如图 2-6 所示。单模光纤的纤芯很细，直径小到只有几微米，制造成本较高。同时，单模光纤的光源使用定向性很好的激光二极管。单模光纤的损耗较小，传输距离较远，传输速率比多模光纤高，一般用作骨干网的传输介质。单模光纤的内径主要有 $5.0\mu\text{m}$ 、 $8.1\mu\text{m}$ 等规格，外径有 $85\mu\text{m}$ 、 $125\mu\text{m}$ 等规格。

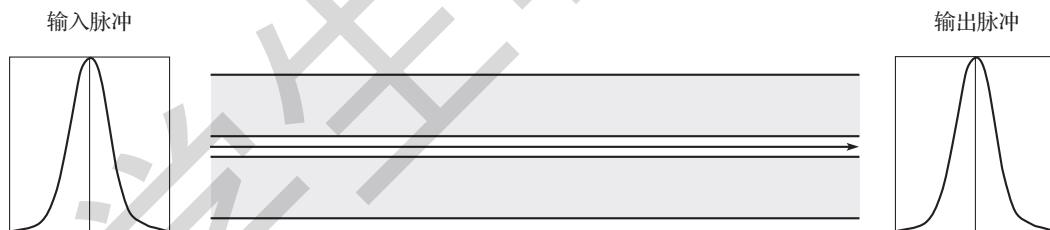


图 2-6 单模光纤

由于光纤很细，因此必须将光纤做成很结实的光缆。一根光缆少则只有两根光纤（称为二芯光缆），多则可包括数十至数百根光纤，再加上加强元件和填充物就可以大大增加其机械强度，最后加上外护套，就可以大大增加其抗拉强度，以满足工程施工的强度要求。

光缆在布设时在转弯处不能简单折成直角（否则光不能顺畅通过）而要保持圆弧状，多余的光缆通常盘成圆圈状。为此，在光缆中会加一根硬度很大的钢丝，既增加抗拉强度，又使其不易被折弯。

4. 光电混合缆

光电混合缆是一种集成了光纤和铜导线的混合形式线缆，光纤用于数据传输，铜导

线用于为设备供电,适用于各类网络系统的综合布线,可以优先降低建设成本。

2.1.3 无线传输介质

1. 陆地微波

微波是指频率在 0.3 ~ 300GHz 范围的电磁波,陆地微波通信就是利用此频段的电磁波来传递信息的,目前主要是使用 2 ~ 40GHz 的频率范围,其中 2.4GHz 频段是 ITU-R 定义的,称为 ISM (工业、科学、医学) 频段,无须授权,也是 Wi-Fi 使用的频段。

陆地微波系统的主要用途是完成远程通信服务和在楼宇间建立短距离的点对点通信。与其他传输介质相比,陆地微波具有如下特点:

- 传输距离受限。微波波长短,接近于光波,在空间主要是直线传播,而地球表面是一个曲面,微波会穿透电离层而进入宇宙空间,因此传播距离受到限制,必须设立中继站以增大传输距离。微波无法穿透障碍物,因此相邻微波站之间必须直视,距离不会太远,地球表面两点之间能直视的最远距离约为 82km。
- 容量大。微波频率高,频段范围也很宽,因此通信信道的容量大。
- 质量高。因为工业干扰和天电干扰的主要频谱成分比微波频率低得多,所以微波传输质量较高。
- 方向性强。由于波长短,因此天线尺寸可做得很小,通常做成面式天线,增益高,方向性强。
- 成本低。与相同容量和长度的电缆通信比较,微波接力通信建设投资少,见效快,特别是在人口密度小、布设光缆成本高的地区。

微波的损耗与距离、波长有关,可由下式估算:

$$L \approx 10 \lg \left(\frac{4\pi d}{\lambda} \right)^{2 \sim 6} \quad (\text{dB})$$

其中 d 是传输距离, λ 是波长,损耗跟电磁环境密切相关,通常取二次方进行估算。

微波的传播有时会受到恶劣气候的影响。

2. 卫星微波

卫星微波是陆地微波的发展,利用人造卫星作为中继站转发微波信号,在多个地面站之间进行数据传输。

卫星从上行链路接收自地面站传输来的信号,将其放大或再生,再从下行链路发送到其他地面站。移动卫星在空中移动,卫星落下水平线后,通信就不能进行,直到它重新在另一个水平线上出现。地球同步卫星能保证持续地进行传输,因为地球同步卫星与地球保持相对固定的位置,它位于赤道上空,离地面约 35 784km。三颗相隔 120° 的同步卫星几乎能覆盖整个地球表面,基本实现全球通信。

与陆地微波相比,卫星微波具有以下特点:

- 卫星通信的距离远。

- 卫星微波具有广播性质。
- 卫星信道的传播延迟较大。地球同步卫星的单程传播延迟为 $0.25 \sim 0.27\text{s}$ 。

目前卫星通信主要使用三个波段：

- 4/6 波段：上行 $5.925 \sim 6.425\text{GHz}$ ，下行 $3.7 \sim 4.2\text{GHz}$ ，目前已分配完。
- 12/14 波段：上行 $14 \sim 14.5\text{GHz}$ ，下行 $11.7 \sim 12.2\text{GHz}$ ，目前已基本分配完。
- 19/29 波段：上行 $27.5 \sim 31\text{GHz}$ ，下行 $17.7 \sim 21.2\text{GHz}$ 。

卫星通信的最佳波段在 10GHz 以下，超过 10GHz 有一个明显的缺点——电波容易被雨水吸收，因此易受天气影响。

卫星通信还受到不可避免的日凌干扰。日凌是指地面站、卫星、太阳在一条线上，由于太阳发出的各类粒子能量太强，因此地面站接收不到卫星的信号，导致通信中断。地球上的地点每年都可能约有约一周的时间出现日凌现象，每次持续的时间为 $10 \sim 30\text{min}$ 。

3. 无线电

无线电波很容易产生，可以传播很远，容易穿过建筑物，可以被电离层反射，因此被广泛用于通信，不管是室内还是室外。无线电波同时还是全方向传播的（没有方向性），因此发射和接收装置不必在物理上对准，同时，无线电通信本身具有广播特性，因此很容易实现广播通信。无线通信频谱是稀缺资源，世界上所有的政府都控制频谱的分配权。

无线电通信一般使用 $30\text{MHz} \sim 1\text{GHz}$ 频段。无线电波的特性与频率有关。在较低频率上（波长较长），无线电波能轻易地通过障碍物，但是到达接收处的能量随着与信号源距离的增大而急剧减小。在高频上，无线电波趋于直线传播并受障碍物的阻挡，还会被雨水吸收。在所有频率上，无线电波易受发送机和其他电子设备的干扰。

无线电波的传输距离可以用下述经验公式估算：

$$d \approx 7.14 \times \sqrt{Kh} \text{ (km)}$$

其中， K 为系数，通常取经验值 $4/3$ ， h 为天线高度（单位为 m ）。

由于无线电波能传得很远，用户间的相互串扰就是个大问题。

4. 红外线

红外线的主要特点是不能穿透物体，这意味着一间房屋里的红外系统不会对其他房间里的系统产生干扰，其防窃听的安全性要比无线电系统高。所以使用红外系统不需要政府授权。

红外通信使用调制非相干红外线光的收发机进行，收发机互相置于视线内对准，直接或经房间天花板的浅色表面的反射传递信息，被广泛用于短距离通信。

5. 激光

这里所说的激光指大气激光，利用大气作为传输介质实现通信。

激光通信的优点是：

- 通信容量大。

- 保密性强。可采用不可见光，不易被发现、截获。
- 设备经济。

激光通信的缺点包括：

- 通信距离限于视距，易受天气影响，云、雨、雾、雪、尘埃（霾）等都可能影响通信的正常进行。
- 瞄准困难。激光束有极高的方向性，这给发射和接收点之间的瞄准带来一定难度。

现在，利用激光通信实现卫星与地面站之间的高速数据传输，是实现卫星数据高速下载的一种重要方式。

2.1.4 传输介质的传输能力

任何介质的传输能力都是有限的。根据信息论的有关理论，传输速率（数据率）与带宽、噪声等都有关系，奈奎斯特、香农分别给出了相应条件下的数据率计算公式。

下面所说的信道（介质）的带宽是指允许通过的最高频率与最低频率之间的宽度（差值），可以将数据率简单地理解为在介质的截面上单位时间内通过的数据量。

1. 奈奎斯特定律

1924 年，奈奎斯特（Nyquist）推导出在理想低通信道下的最高码元（即信号，比如正弦信号、脉冲信号等）传输速率的公式，称为奈奎斯特定律（也称为奈奎斯特准则）。该定律为：

在带宽为 W （Hz）的无噪声信道上传输信号，假定每个信号取 V 个离散电平值，则信道的极限数据率（比特率）为

$$B = 2W \times \log_2 V \text{ (bps)}$$

奈奎斯特定律的另一种等效情况是：无噪声信道具有无限带宽，其上传输的信号带宽为 W ，假定每个信号取 V 个离散电平值，则传输该信号的极限数据率为

$$B = 2W \times \log_2 V \text{ (bps)}$$

通过奈奎斯特定律建立了带宽与数据率之间的对应关系，带宽越宽，数据率就越高，传输能力就越大。这也是人们常把数据率高的网络称为宽带网的原因。

如果不考虑每个信号所取的电平数，只考虑码元数，则其值为码元速率，

$$\text{理想低通信道的最高码元传输速率} = 2W$$

称为奈奎斯特信号速率，也称为波特率（单位时间内传送的信号数）。

若码元的传输速率超过了奈奎斯特定律给出的数值，则将出现码元之间的相互干扰，以致接收方无法正确识别发送方所发送的码元。

由于码元的传输速率受奈奎斯特定律的制约，因此要提高数据的传输速率，就必须设法使每个码元能携带更多个比特的信息量，这就需要采用多元制的编码方法。

2. 香农定理

奈奎斯特定律考虑的信道是没有噪声的，但实际信道都是有噪声的，这样奈奎斯

特定律就不适用。

1948年, 香农(Shannon)用信息论的理论推导出了有噪声信道上带宽与数据率的关系, 称为香农定理。该定理为:

在带宽为 W (Hz) 的有噪声信道上传输信号, 假定信噪比为 S/N (信号与噪声的功率比), 则信道的极限数据率为

$$B = W \times \log_2(1 + S/N) \text{ (bps)}$$

香农定理的另一种等效情况是: 信道具有无限带宽, 其上传输的信号带宽为 W , 信噪比为 S/N , 则传输该信号的极限数据率为

$$B = W \times \log_2(1 + S/N) \text{ (bps)}$$

香农定理表明, 信道的带宽越宽或信道中的信噪比越大, 则信道的极限传输速率就越高。更重要的是, 香农定理指出: 只要数据传输速率低于信道的极限数据传输速率, 就一定能找到某种方法来实现无差错的传输。

由于信噪比 S/N 本身的物理和数学特性, 实际使用时都是取对数 $\lg(S/N)$, 单位为贝(B), 但贝的单位太大, 要再乘 10, 把单位变换成分贝(dB)。

【例题】带宽为 3kHz 的信道传输二进制信号, 如果没有噪声, 按奈奎斯特定律, 极限数据率为 6kbps。如果有噪声且信噪比为 30dB, 极限数据率为 $3000 \times \log_2(1 + 1000) \approx 30\text{kbps}$ (信噪比为 30dB, $10\lg(S/N) = 30$, $S/N = 1000$)。

这一结果并不是说有噪声信道比无噪声信道更好, 其原因是:

- 两个公式计算出来的都是极限数据率, 都是一个上限值, 有的上限值与实际值差得更多一点。
- 奈奎斯特定律只考虑二进制, 即一个信号只取两种离散电平值(0、1), 而香农定理考虑最大可能情况, 比如一个信号取 N 个离散电平值, 而 $N \rightarrow \infty$ 。

从奈奎斯特定律和香农定理可以知道, 介质的带宽越宽, 其传输能力(单位时间传输的数据量)就越强。这也就是社会上用宽带网来指称性能好、传输能力强的网络的原因。要想提高数据率, 要么增大带宽, 要么增大信噪比, 要么增大单个信号(也称为符号)的离散电平数(1个信号表示的位数), 或同时增大三者, 这就为提高数据率指明了方向。

2.1.5 Wi-Fi 无线信号的功率

Wi-Fi 使用无线通信时, 无线信号的功率对接收性能有较大影响。信号太强, 信噪比高, 易提升接收性能, 但也容易造成大范围干扰; 信号太弱, 导致无法识别而接收失败。对于 Wi-Fi, 接收端信号的功率单位是 dBm (分贝毫瓦)。换算公式是接收端 $\text{dBm} = 10\lg(P_{\text{rx}}/P_{\text{tx}})$, 其中 P_{rx} 表示信号被接收到时的功率, P_{tx} 表示信号被发送时的功率, 因为 $P_{\text{rx}} < P_{\text{tx}}$, 所以 dBm 值都是负数。通常, 取 $P_{\text{tx}} = 1\text{mW}$, 则接收端信号 $\text{dBm} = 10\lg(P_{\text{rx}})$, 因为 $P_{\text{rx}} < 1\text{mW}$, 所以其值是负数。比如, -40dBm 对应接收到的功率是 10^{-4}mW ,

-60dBm 对应接收到的功率是 10^{-6}mW 。dBm 的负数值越大（绝对值越小）表示信号越强。Wi-Fi 的 dBm 值正常范围是 $-40 \sim -85\text{dBm}$ ，低于 -90dBm 时基本不能接收。

其他无线通信系统接收端的信号功率表示方法类似，但具体数值不同。

2.2 数据通信系统与物理层

2.2.1 数据通信系统模型

将计算机或终端与数据传输线路连接起来，达到数据传输、收集、分配、存储、处理目的的系统，统称为数据通信系统。数据通信系统一般包括发送端、接收端以及收发两端之间的信道三个部分，如图 2-7 所示。

信息源（简称信源）是信息或信息序列的产生源，它泛指一切发信者，可以是人，也可以是机器设备，能够产生声音、数据、文字、图像、代码等信号。信息源发出信息的形式可以是连续的，也可以是离散的。

发送设备是把信息源发出的信息转换成适合传输的信号形式（码元），使之适应于信道传输特性的要求并送入信道的各种设备。发送设备是一个整体概念，可能包括许多的电路、器件与系统，比如把声音转换为电信号的声音采集设备（最简单的就是麦克风）、把基带信号转换成频带信号的调制器等。

信道是指传输信号的通道，传输介质是直接、简单的信道（复杂信道可以是一条网络通路）。信道中会有噪声，可能是进入信道的各种外部噪声，也可能是通信系统中各种电路、器件或设备自身产生的内部噪声。

接收设备接收从信道传输过来的信息，并将该信息转换成信息宿（简称为信宿）便于接收的形式，其功能与发送设备的功能刚好相反。接收设备也是一个整体概念，可能包括许多的电路、器件与系统，比如把模拟信号转换为数字信号的模/数转换器等。

信息宿是接收发送端信息的对象，它可以是人，也可以是机器设备。

网卡是典型的发送和接收设备，其原理将在后续章节介绍。

按照信道中所传输信号形式的不同，通信系统可以进一步分为模拟通信系统和数字通信系统。

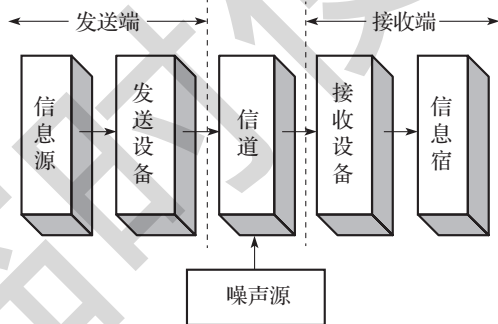


图 2-7 数据通信系统的模型

2.2.2 数据传输模式

1. 信号的频域与时域特性

为便于理解和叙述，我们定义：数据是设备中保存的表征客观事物的静态属性，信

号是数据在介质上传输时的动态存在形式。

笼统地说,数据都要变换为信号在介质上传输,而不同形式的信号具有不同的特性和不同的传输性能。

研究表明,任何信号都可以被看成由无穷多个不同频率的信号分量合成,其中最小的那个频率称为基频,是合成后信号呈现出的频率。这一特性称为信号的频域特性。例如,

$$x = \sum_{k=1}^{\infty} \frac{1}{2k-1} \sin 2\pi(2k-1)ft$$

是由频率为 $f, 3f, 5f, \dots$ 的正弦信号合成的,其合成结果为方波,合成后的方波信号频率为 f ,如图2-8所示。我们平时说无线信号的频率是某个值,其实是指合成后的信号的频率,它事实上包含大量不同频率的信号分量,这也是不同频率的信号之间也存在干扰的原因之一。

周期性信号随着时间而重复,可以表示为 $f(t) = f(t+T)$, T 为周期,这被称为信号的时域特性。

2. 数字传输与模拟传输

按承载数据的信号形式的不同,传输可分为模拟传输和数字传输。

- 模拟传输是指以模拟信号来传输数据的传输方式,与信号所代表的数据的形式无关。
- 数字传输是指用数字信号来传输数据的传输方式,与信号所代表的数据的形式无关。

不论是数字数据还是模拟数据,都可以采用两种传输方式之一进行传输:

- 数字数据(二进制序列) → 编码为数字信号 → 数字传输
- 数字数据(二进制序列) → 调制为模拟信号 → 模拟传输
- 模拟数据(连续值) → 编码为数字信号 → 数字传输
- 模拟数据(连续值) → 调制为模拟信号 → 模拟传输

数字信号的典型形式是脉冲波,模拟信号的典型形式是正弦波。

通常数字传输的距离不及模拟传输的距离远。为了增大传输距离,两种传输采用了不同的方式。

- 数字传输采用中继方式,由中继器完成。其原理是:接收信号,进行整形,还原为脉冲信号(称为再生),然后发送出去。每次中继,都把信号还原,由于去掉了噪声,因此传输质量高。
- 模拟传输采用放大方式,由放大器完成。其原理是:接收信号,按比例放大信号的各部分,然后发送出去。显然,放大方式把噪声一并放大、传送到了接收方。

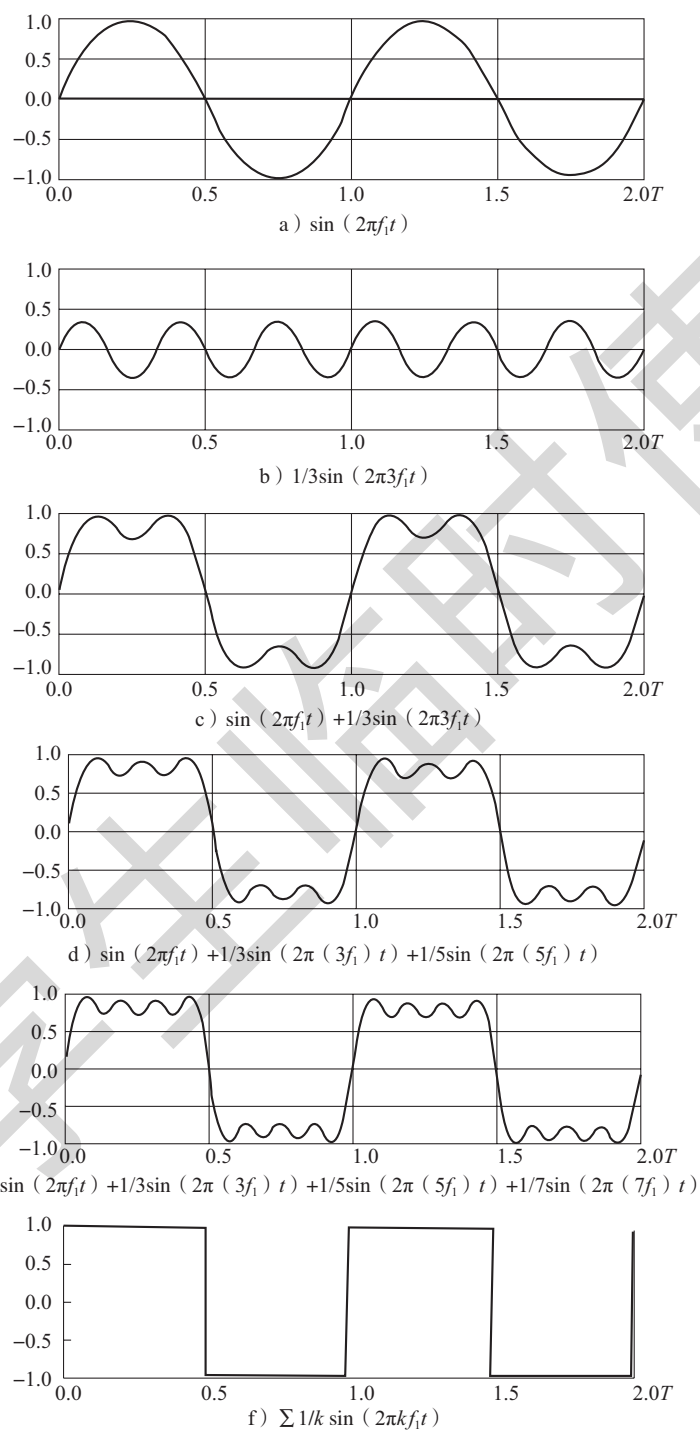


图 2-8 多个正弦信号合成后为方波信号

数字传输的优点是：

- 设备简单，只需处理 0 和 1，可优化信道设计。
- 传输质量高，出错率低。
- 通信能力强。
- 便于统一多种网络。利用数字化技术，可以把所有网络都改造成数字传输，这样可把各种网络统一成一个公共传输网络。

数字传输的最大缺点是传输距离近，比如数字传输现在还不能用于远距离卫星通信。

3. 基带传输与频带传输

基带是指未经调制变换的信号所占的频带。基带信号是指高限频率与低限频率之比远大于 1 的信号。比如声音信号的频率范围是 $300 \sim 3400\text{Hz}$ ，其高限频率为 3400Hz ，低限频率为 300Hz ，高低限比值超过 11。基带数字信号的频谱从 0 开始。

基带传输是指不改变基带信号频谱的传输方式。采用这种信号传输技术的通信系统称为基带传输通信系统，简称基带系统。

有时基带传输是很难实现的，比如声音，如果直接将 $300 \sim 3400\text{Hz}$ 的信号进行传输，将需要几千米高且直径为几千米的天线才可能发送，这显然不现实。所以，为了能够方便、经济地传输，通常需要对基带信号进行变换。

频带传输是指改变原始信号的频带，使其在适合的通信频带上传输的传输方式。采用这种信号传输技术的通信系统称为频带传输通信系统，简称频带系统。例如，调频通信，将 $300 \sim 3400\text{Hz}$ 的信号变换到 $88\text{MHz} + (300 \sim 3400)\text{Hz}$ 的频带上，高频与低频的比值，即 $(88 + 3400 \times 10^{-6}) / (88 + 300 \times 10^{-6})$ ，非常接近 1。

2.2.3 信号编码与调制

数据究竟应当是数字的还是模拟的，是由所产生的数据的性质决定的。但数据必须转换成信号才能在网络介质上传输。2.2.2 节已经介绍，模拟数据和数字数据都可以转换为模拟信号或数字信号进行传输。表示数据元素一般用位或比特，表示信号元素一般用脉冲或波形，称为码元。

1. 数字数据编码为数字信号

(1) 非归零制编码

对于数字数据，最普遍且最容易的编码方法就是用两个电压来表示二进制的两个状态。例如，无电压表示 0，有电压表示 1。但是没有电压表示 0 与没有传输信号（也没有电压）无法区分，因此需要对此加以改善。通信中，用正电压表示 1，用负电压表示 0，这种编码方式就是非归零制编码。

但是非归零制传输也有缺点：它难以决定一位的结束和另一位开始，特别是有多个连续 1 或连续 0 的情况，需要用某种方法使发送器和接收器进行定时或同步，以识别有多少个 1 或 0。

(2) 双极 AMI

AMI 码的编码规律如图 2-9 所示, 数据序列中的“0”码仍为“0”, 数据序列中的“1”码则交替编码为 +1 和 -1。由于“1”码极性交替, 如果接收端发现极性不是交替出现就一定出现了传输错误。但码流中连续 0 过多时, 依然不利于识别。

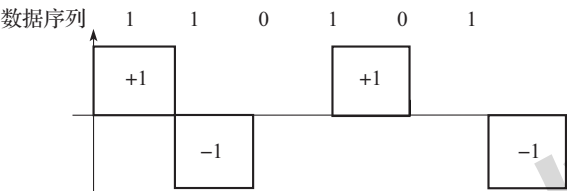


图 2-9 双极 AMI 码的编码规律

(3) 曼彻斯特编码

曼彻斯特编码方法是每一个码元再分成两个相等部分, 前一部分与后一部分电压高低相反。可规定前一个部分为高电平而后一部分为低电平表示 0, 而 1 则刚好相反, 如图 2-10 所示。这种编码的优点就是可以保证在每一个码元的正中间出现一次电平的变化, 该变化既作为同步信号, 又作为信号的取值, 有利于接收方提取位同步信号。但是它的缺点是编码效率不高, 曼彻斯特编码相当于用两个脉冲表示一个位。10Mbps 以太网用的就是这种编码方式。

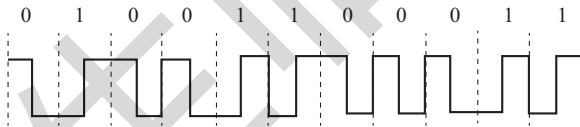


图 2-10 曼彻斯特编码

(4) nB/mB 码

nB/mB 码 ($m > n$) 用 m 位编码表示 n 位数据, 因此实际的码组有 2^m 种, 冗余码组有 $2^m - 2^n$, 编码效率为 n/m 。

目前应用较为广泛的有 4B/5B、5B/6B、8B/10B、64B/66B、256B/257B。其中 100Mbps 以太网使用的是 4B/5B 编码方式, 1Gbps (千兆) 以太网使用的是 8B/10B 编码方式, 10Gbps/40Gbps/100Gbps 以太网使用的是 64B/66B 编码方式, 200Gbps/400Gbps 以太网使用的是 256B/257B 编码方式。

下面以 5B/6B 为例说明 nB/mB 码的编码方式。5B/6B 码从 6 位二进制的 64 种组合中选出 32 个码组对数据进行编码, 其编码表如表 2-2 所示。该编码表列有正模式和负模式两种, 使用时成对选择, 以使码序列中“1”和“0”的个数趋于平衡。5B/6B 码具有如下特性:

- 编码后的序列中最大的连 0 和连 1 个数为 5。

- 累积的 1、0 码个数的差值（称为数字和）在 $-3 \sim +3$ 范围内变化，这一特性可用于误码检测。
- 各码组的数字和没有 ± 1 的值，可用于码组同步。

表 2-2 5B/6B 编码表

输入二进制码组	输出二进制码组		输入二进制码组	输出二进制码组	
	正模式	负模式		正模式	负模式
00000	110010	110010	10000	110001	110001
00001	110011	100001	10001	111001	010001
00010	110110	100010	10010	111010	010010
00011	100011	100011	10011	010011	010011
00100	110101	100100	10100	110100	110100
00101	100101	100101	10101	010101	010101
00110	100110	100110	10110	010110	010110
00111	100111	000111	10111	010111	010100
01000	101011	101000	11000	111000	011000
01001	101001	101001	11001	011001	011001
01010	101010	101010	11010	011010	011010
01011	001011	001011	11011	011011	001010
01100	101100	101100	11100	011100	011100
01101	101101	000101	11101	011101	001001
01110	101110	000110	11110	011110	001100
01111	001110	001110	11111	001101	001101

(5) PAM4 码

PAM4 是脉冲幅度调制的一种形式，在 Wi-Fi 6 中使用，用 4 种电平值编码 2 位（比特）。例如，用四个电平值 $-1V$ 、 $-0.5V$ 、 $+0.5V$ 、 $+1V$ 分别表示 00、01、10、11。以此类推能否设计出 PAM8 呢？理论上可行，但对器件性能要求很高。

2. 数字数据调制为模拟信号

模拟信号发送的基础是载波信号，可用 $x = A \cdot \sin(\omega t + \phi)$ 表示，其中有振幅、频率、相位三个基本参数。基本思想是选取某一频率的正弦信号作为载波用以运载所要传送的数字数据。用待传送的数字数据改变载波的幅值、频率或相位或其组合，将改变后的信号发送出去，到达目的地后进行分离。

把数字数据加到载波上的过程称为调制，从载波上取出数字数据的过程称为解调。完成调制功能的设备叫调制器，完成解调功能的设备叫解调器。通常，调制器和解调器集成在一起，称为调制解调器，缩写为 MODEM。

(1) 幅移键控调制方法

幅移键控（Amplitude Shift Keying, ASK）的基本思想是利用待传数据改变载波的振

幅，而载波的频率和相位都保持不变。例如，如果待传数据为 0，则载波的振幅变为 0；如果待传数据为 1，则载波的振幅不变。该方法的特点是实现简单，但抗干扰性能差。

设 R 为要求达到的数据率， r 为系数 ($0 < r < 1$)， L 为离散电平数量，则调制后模拟信号要求的带宽为 $B_T = (1+r)R/\log_2 L$ 。

(2) 频移键控调制方法

频移键控 (Frequency Shift Keying, FSK) 的基本思想是利用待传数据改变载波的频率，而载波的振幅和相位不变。例如，如果待传数据为 0，则载波的频率不变；如果待传数据为 1，则载波的频率加倍。与 ASK 相比，该方法的抗干扰性能更好。

通常的方法是，待传数据 0、1 分别将载波的频率 f_c 改为 f_0 、 f_1 ，三个频率满足关系 $f_c - f_0 = f_1 - f_c$ 。

设 R 为要求达到的数据率， r 为系数 ($0 < r < 1$)， L 为离散电平数量，则调制后模拟信号要求的带宽为 $B_T = 2\Delta F + (1+r)R/\log_2 L$ ， $\Delta F = f_1 - f_c = f_c - f_0$ 。

(3) 相移键控调制方法

相移键控 (Phase Shift Keying, PSK) 的基本思想是利用待传数据改变载波的相位，而载波的振幅和频率不变。

PSK 有两种基本实现方法，即绝对相移和相对相移。绝对相移方法是用待传的数据 0 将载波的相位变为某一固定相位值，而用待传的数据 1 将载波的相位变为另一个固定相位值。例如， 0° 表示 1， 180° 表示 0。相对相移是指用前后信号有无相位变化来表示 0 和 1。

设 R 为数据率， r 为系数 ($0 < r < 1$)， L 为离散电平数量，则模拟信号要求的带宽为 $B_T = (1+r)R/\log_2 L$ 。

二进制幅移键控、频移键控和相移键控的例子如图 2-11 所示。ASK 中用载波有幅度和幅度为 0 分别表示数字数据的“1”和“0”；FSK 中用两种不同的频率分别表示数字数据的“1”和“0”；PSK 中用 0° 相位和 180° 相位分别表示数字数据的“1”和“0”。

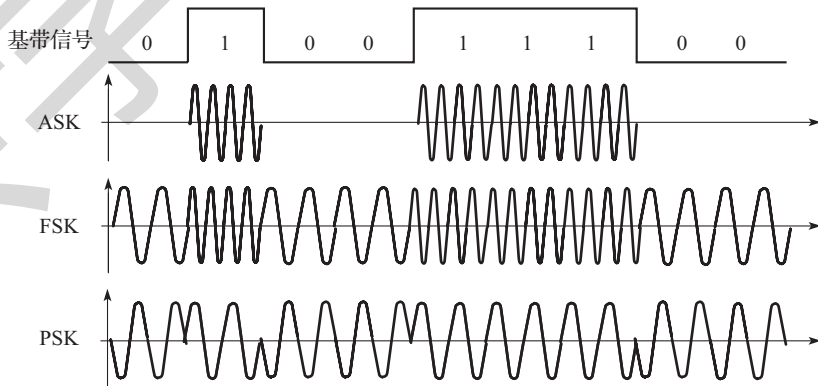


图 2-11 二进制 ASK、FSK、PSK

FSK 信号的抗噪性能优于 ASK 信号, PSK 信号抗噪性能与 ASK、FSK 相比是最优的。

(4) 正交振幅调制

为更加高效地进行调制(编码),可以联合使用上述方法。正交振幅调制(Quadrature Amplitude Modulation, QAM)就是一种常用的技术,它将幅移键控和相移键控结合在一起,把两个频率相同的模拟信号叠加在一起,一个对应正弦函数,一个对应余弦函数。M 进制的正交振幅调制可简记为 MQAM,其信号可表示为

$$S(t) = C\sin(\omega t) + D\cos(\omega t)$$

更高阶的 QAM 的一个信号能传送更多的二进制位数据。图 2-12 是 16QAM 的例子,一个信号能传送 4 位数据。

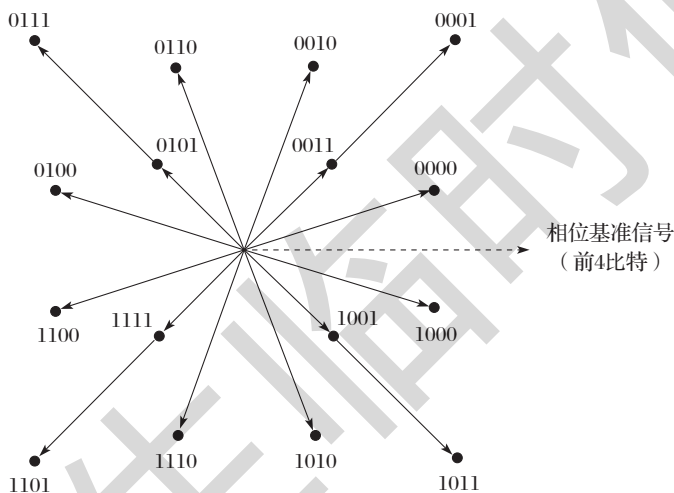


图 2-12 16QAM

可以以低阶的 QAM 为基础生成高阶 QAM。例如,4 个 64QAM 组成一个 256QAM,4 个 256QAM 组成一个 1024QAM。最新的调制方式已经提升到 4096QAM,将在下一代 Wi-Fi 中得到应用。

3. 模拟数据编码为数字信号

模拟数据编码为数字信号包括两步:第一步是将模拟数据变成数字数据,称为数字化,第二步是将数字数据编码成数字信号。其中第二步可以使用前述方法编码,这里只介绍第一步的方法。

通常,模拟信号离散化是用一个周期为 T 的脉冲信号控制采样电路对模拟信号 $f(t)$ 实施采样,得到样值序列 $f_s(t)$ 。如果取出的样值足够多,这个样值序列就能逼近原始的连续信号。但采样周期 T 取多大才能满足用样值序列 $f_s(t)$ 代表模拟信号 $f(t)$ 的要求呢?

奈奎斯特采样定理为:如果一个模拟信号 $f(t)$ 的最高频率分量为 f_m ,当满足采样频

率 $f_s \geq 2f_m$ ($f_s = 1/T$) 时, 所获得的样值序列 $f_s(t)$ 就可以完全代表原模拟信号 $f(t)$ 。

$f_s(t)$ 的值可能包含多位小数, 这样的取样值与原始信号的误差较小, 但在实际实现时, 一般要四舍五入取整数。实现数字化的主要方法有 PCM、DM、DPCM 等。

(1) PCM

实现数字化的最直接方法是 PCM (Pulse Code Modulation, 脉码调制) 方法。PCM 包括采样、量化、编码三个步骤。

1) 采样: 按一定的采样频率采样, 得到样值序列 $f_s(t)$ 。

2) 量化: 量化是将样值离散化 (变为整数) 的过程。普遍使用的是均匀量化。所谓均匀量化是指以等间隔对任意信号值来量化, 即将信号样值幅度的变化范围 $[-U, +U]$ 等分成 N 个量化级, 记作 Δ , 则

$$\Delta = \frac{2U}{N}$$

根据量化的规则, 样值幅度落在某一量化级区间内, 则由该级的中心值来量化。

3) 编码: 获得量化值后, 再用 n 位二进制码对其进行编码, 码组的长度 n 与量化级数 N 之间的关系为

$$N = 2^n$$

量化为整数时存在量化噪声 (误差)。量化噪声的估算公式为

$$S/N = 6n - a \quad (0 < a < 1, \text{取 } 1.8\text{dB})$$

【例题】一个模拟数据的电压变化范围为 $[-1\text{V}, 1\text{V}]$, 采样值为 0.33V , 采用线性 PCM 将其编码为 3 位二进制, 解码后误差为多少?

首先求出 $N = 2^3 = 8$, $\Delta = 2 \times 1/8 = 0.25$; 然后设计出量化编码表, 如表 2-3 所示。

表 2-3 量化编码表

变化区间	量化值	编码
$[-1\text{V}, -0.75\text{V})$	-0.875V	000
$[-0.75\text{V}, -0.5\text{V})$	-0.625V	001
$[-0.5\text{V}, -0.25\text{V})$	-0.375V	010
$[-0.25\text{V}, 0\text{V})$	-0.125V	011
$[0\text{V}, 0.25\text{V})$	0.125V	100
$[0.25\text{V}, 0.5\text{V})$	0.375V	101
$[0.5\text{V}, 0.75\text{V})$	0.625V	110
$[0.75\text{V}, 1\text{V}]$	0.875V	111

根据表 2-3 可知, 0.33V 编码为 101。接收方收到 101 后, 也根据表 2-3 解码为 0.375V , 误差为 $0.375\text{V} - 0.33\text{V} = 0.045\text{V}$ 。

【例题】利用 PCM 方式实现数字电话。因音频范围为 $300 \sim 3400\text{Hz}$, 通常取最高频率为 4kHz , 按采样定理, 采样频率为 8kHz 。假定每次采样用 8 位编码, 则传输一路数

数字化声音需要的带宽为 64kbps。

(2) DM

PCM 用固定长度的位数对样值进行编码,但在实际应用中,连续采样的样值之间可能变化较小,用较长的位数编码、传送会造成较大的浪费。DM (Delta 调制,或称增量调制)的目的是减少冗余信息。

DM 的原理是:首次采样用 n 位编码,随后每次采样,只用 1 位进行编码,用 1 表示相对于上次的样值增大,用 0 表示相对于上次的样值没有增大(相等或减小)。DM 需要约定每次增大或减小量的数值。

DM 的一个明显缺点是,对于变化缓慢的原始信号(曲线平坦),因为每次采样几乎没有变化,但实际判断为每次减少,所以累计误差将会很大。

(3) DPCM

DPCM (差分 PCM) 是对 PCM 和 DM 的一种折中,首次采样用 n 位编码,随后每次采样用 m 位编码 ($m < n$),其中有 1 位符号位, $m-1$ 位数据位,这样每次编码可节省 $n-m$ 位,传输效率提高了 $(n-m)/n$ 。比如,对声音,设 $n=8$ 、 $m=4$,其中 1 位符号(表示正/负),3 位表示增/减的数值,这样对带宽的需求降低了一半。

4. 模拟数据调制为模拟信号

模拟数据经由模拟信号传输时不需要进行变换,但是模拟数据本身的频率不高,考虑到发送的难度问题,模拟形式的输入数据也需要在高频下进行调制,其输出信号是一种带有输入数据的频率极高的模拟信号。模拟数据调制为模拟信号有三种典型的调制技术:调幅 (Amplitude Modulation, AM)、调频 (Frequency Modulation, FM) 与调相 (Phase Modulation, PM)。调频和调相调制信号的频谱都是调制信号频谱的非线性变化,而且二者的已调信号都反映出载波向量角度的变化,所以统称为角度调制。

(1) 调幅

调幅是一种使高频载波的幅度随着原始模拟数据的幅度变化而变化的技术。载波的幅度会在整个调制过程中变动,而载波的频率是不变的。调幅调制信号的表达式为

$$s_{AM}(t) = s(t) \cos \omega t$$

式中, $\cos \omega t$ 为载波, $s(t)$ 是要进行调制的基带信号。

(2) 调频

调频是一种使高频载波的频率随着原始模拟数据的幅度变化而变化的技术。因此,载波的频率在整个调制过程中波动,而载波的幅度是相同的。调频调制信号的表达式为

$$s_{FM}(t) = A \cos \left[\omega t + \int_{-\infty}^t K_F m(t) dt \right]$$

式中, K_F 代表调频器的灵敏度, $m(t)$ 为待调制信号。

(3) 调相

调相是一种使高频载波的相位随着原始模拟数据的幅度变化而变化的技术。载波的

相位在整个调制过程中变动，而载波的幅度是相同的。调相调制信号的表达式为

$$s_{\text{PM}}(t) = A \cos[\omega t + K_p m(t)]$$

式中， K_p 代表调相器的灵敏度， $m(t)$ 为待调制信号。

5. 扩频通信

在无线通信中，为了提高通信系统的抗干扰性能，往往需要从调制和编码多方面进行改进，以提高通信质量，扩频通信就是方法之一。由于扩频通信利用了扩展频谱技术，在接收端对干扰频谱能量加以扩散，对信号频谱能量压缩集中，因此在输出端就得到了信噪比的增益。

扩频通信是指系统占用的频带宽度远大于要传输的原始信号的带宽（或信息比特率），且与原始信号带宽无关。如果信息带宽为 B ，扩频信号带宽为 f_{ss} ，则扩频信号带宽与信息带宽之比 f_{ss}/B 称为扩频因子。

- 当 $f_{\text{ss}}/B = 1 \sim 2$ ，即扩频信号带宽略大于信息带宽时，称为窄带通信。
- 当 $f_{\text{ss}}/B \geq 50$ ，即扩频信号带宽大于信息带宽时，称为宽带通信。
- 当 $f_{\text{ss}}/B \geq 100$ ，即扩频信号带宽远大于信息带宽时，称为扩频通信。

扩频通信系统可以分为以下几种基本形式。

(1) 直接序列扩频

直接序列（Direct Sequencing, DS）扩频方式中，在发送端直接用扩频码（伪随机码）序列去扩展信号的频谱，即将 1 位变成多位。在接收端，用相同的扩频码序列进行解扩，将展宽的频谱扩展信号还原成原始信号，即将多位还原为 1 位。图 2-13 所示是 DS 扩频的一个实现例子，其中扩频码只有收发双方知道，这样扩频后的信号也具有保密性。

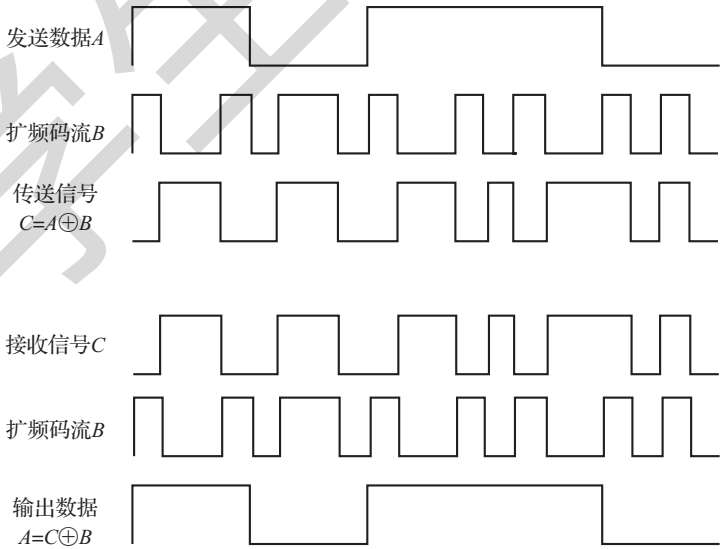


图 2-13 直接序列扩频例子