

2.7.1 UDP 套接字编程	105	编程作业	195
2.7.2 TCP 套接字编程	109	Wireshark 实验: 探究 TCP	196
2.8 小结	112	Wireshark 实验: 探究 UDP	196
课后习题和问题	113	人物专访	196
复习题	113	第4章 网络层: 数据平面	198
习题	114	4.1 网络层概述	198
套接字编程作业	118	4.1.1 转发和路由选择: 数据平面和 控制平面	199
Wireshark 实验: HTTP	119	4.1.2 网络服务模型	202
Wireshark 实验: DNS	120	4.2 路由器工作原理	203
人物专访	120	4.2.1 输入端口处理和基于目的地 转发	205
第3章 运输层	121	4.2.2 交换	207
3.1 概述和运输层服务	121	4.2.3 输出端口处理	209
3.1.1 运输层和网络层的关系	122	4.2.4 何处出现排队	209
3.1.2 因特网运输层概述	123	4.2.5 分组调度	211
3.2 多路复用与多路分解	125	4.3 网际协议: IPv4、寻址、IPv6 及 其他	214
3.3 无连接运输: UDP	130	4.3.1 IPv4 数据报格式	214
3.3.1 UDP 报文段结构	132	4.3.2 IPv4 数据报分片	216
3.3.2 UDP 检验和	133	4.3.3 IPv4 编址	217
3.4 可靠数据传输原理	134	4.3.4 网络地址转换	225
3.4.1 构造可靠数据传输协议	135	4.3.5 IPv6	227
3.4.2 流水线可靠数据传输协议	143	4.4 通用转发和 SDN	231
3.4.3 回退 N 步	145	4.4.1 匹配	233
3.4.4 选择重传	148	4.4.2 动作	234
3.5 面向连接的运输: TCP	152	4.4.3 匹配加动作操作中的 OpenFlow 例子	234
3.5.1 TCP 连接	152	4.5 小结	236
3.5.2 TCP 报文段结构	154	课后习题和问题	236
3.5.3 往返时间的估计与超时	157	复习题	236
3.5.4 可靠数据传输	159	习题	237
3.5.5 流量控制	164	Wireshark 实验	240
3.5.6 TCP 连接管理	166	人物专访	241
3.6 拥塞控制原理	170	第5章 网络层: 控制平面	242
3.6.1 拥塞原因与代价	171	5.1 概述	242
3.6.2 拥塞控制方法	175	5.2 路由选择算法	244
3.7 TCP 拥塞控制	176	5.2.1 链路状态路由选择算法	246
3.7.1 公平性	183	5.2.2 距离向量路由选择算法	248
3.7.2 明确拥塞通告: 网络辅助拥塞 控制	184	5.3 因特网中自治系统内部的路由 选择: OSPF	254
3.8 小结	185		
课后习题和问题	187		
复习题	187		
习题	189		

5.4	ISP 之间的路由选择: BGP	256
5.4.1	BGP 的作用	257
5.4.2	通告 BGP 路由信息	257
5.4.3	确定最好的路由	259
5.4.4	IP 任播	261
5.4.5	路由选择策略	262
5.4.6	拼装在一起: 在因特网中 呈现	264
5.5	SDN 控制平面	265
5.5.1	SDN 控制平面: SDN 控制器和 SDN 网络控制应用程序	266
5.5.2	OpenFlow 协议	267
5.5.3	数据平面和控制平面交互的 例子	269
5.5.4	SDN 的过去与未来	270
5.6	ICMP: 因特网控制报文协议	272
5.7	网络管理和 SNMP	274
5.7.1	网络管理框架	274
5.7.2	简单网络管理协议	275
5.8	小结	277
	课后习题和问题	278
	复习题	278
	习题	279
	套接字编程作业	281
	编程作业	282
	Wireshark 实验	282
	人物专访	283
第 6 章	链路层和局域网	285
6.1	链路层概述	285
6.1.1	链路层提供的服务	287
6.1.2	链路层在何处实现	287
6.2	差错检测和纠正技术	288
6.2.1	奇偶校验	289
6.2.2	检验和方法	290
6.2.3	循环冗余检测	291
6.3	多路访问链路和协议	292
6.3.1	信道划分协议	294
6.3.2	随机接入协议	295
6.3.3	轮流协议	301
6.3.4	DOCSIS: 用于电缆因特网接入的 链路层协议	301
6.4	交换局域网	302
6.4.1	链路层寻址和 ARP	303
6.4.2	以太网	308
6.4.3	链路层交换机	312
6.4.4	虚拟局域网	317
6.5	链路虚拟化: 网络作为链路层	319
6.6	数据中心网络	322
6.7	回顾: Web 页面请求的历程	326
6.7.1	准备: DHCP、UDP、IP 和 以太网	326
6.7.2	仍在准备: DNS 和 ARP	327
6.7.3	仍在准备: 域内路由选择到 DNS 服务器	328
6.7.4	Web 客户 - 服务器交互: TCP 和 HTTP	329
6.8	小结	330
	课后习题和问题	331
	复习题	331
	习题	331
	Wireshark 实验	335
	人物专访	336
第 7 章	无线网络和移动网络	338
7.1	概述	339
7.2	无线链路和网络特征	341
7.3	WiFi: 802.11 无线 LAN	346
7.3.1	802.11 体系结构	347
7.3.2	802.11 MAC 协议	350
7.3.3	IEEE 802.11 帧	353
7.3.4	在相同的 IP 子网中的 移动性	355
7.3.5	802.11 中的高级特色	356
7.3.6	个人域网络: 蓝牙和 ZigBee	357
7.4	蜂窝因特网接入	358
7.4.1	蜂窝网体系结构概述	359
7.4.2	3G 蜂窝数据网: 将因特网扩展 到蜂窝用户	360
7.4.3	走向 4G: LTE	362
7.5	移动管理: 原理	364
7.5.1	寻址	367
7.5.2	路由选择到移动节点	367
7.6	移动 IP	371

7.7 管理蜂窝网中的移动性	374	8.9 运行安全性：防火墙和入侵检测系统	424
7.7.1 对移动用户呼叫的路由选择 ...	375	8.9.1 防火墙	424
7.7.2 GSM 中的切换	376	8.9.2 入侵检测系统	429
7.8 无线和移动性：对高层协议的影响	378	8.10 小结	431
7.9 小结	380	课后习题和问题	432
课后习题和问题	380	复习题	432
复习题	380	习题	434
习题	381	Wireshark 实验	437
Wireshark 实验	383	IPsec 实验	437
人物专访	383	人物专访	438
第 8 章 计算机网络中的安全	385	第 9 章 多媒体网络	439
8.1 什么是网络安全	385	9.1 多媒体网络应用	439
8.2 密码学的原则	387	9.1.1 视频的性质	439
8.2.1 对称密钥密码体制	388	9.1.2 音频的性质	440
8.2.2 公开密钥加密	392	9.1.3 多媒体网络应用的类型	441
8.3 报文完整性和数字签名	396	9.2 流式存储视频	443
8.3.1 密码散列函数	397	9.2.1 UDP 流	444
8.3.2 报文鉴别码	398	9.2.2 HTTP 流	444
8.3.3 数字签名	399	9.3 IP 语音	447
8.4 端点鉴别	404	9.3.1 尽力而为服务的限制	448
8.4.1 鉴别协议 ap1.0	404	9.3.2 在接收方消除音频的时延抖动	449
8.4.2 鉴别协议 ap2.0	405	9.3.3 从丢包中恢复	451
8.4.3 鉴别协议 ap3.0	405	9.3.4 学习案例：使用 Skype 的 VoIP	453
8.4.4 鉴别协议 ap3.1	406	9.4 实时会话式应用的协议	455
8.4.5 鉴别协议 ap4.0	406	9.4.1 RTP	455
8.5 安全电子邮件	407	9.4.2 SIP	457
8.5.1 安全电子邮件	407	9.5 支持多媒体的网络	461
8.5.2 PGP	409	9.5.1 定制尽力而为网络	462
8.6 使 TCP 连接安全：SSL	410	9.5.2 提供多种类型的服务	463
8.6.1 宏观描述	411	9.5.3 区分服务	468
8.6.2 更完整的描述	413	9.5.4 每连接服务质量保证：资源预约和呼叫准入	470
8.7 网络层安全性：IPsec 和虚拟专用网	415	9.6 小结	472
8.7.1 IPsec 和虚拟专用网	415	课后习题和问题	473
8.7.2 AH 协议和 ESP 协议	416	复习题	473
8.7.3 安全关联	416	习题	473
8.7.4 IPsec 数据报	417	编程作业	477
8.7.5 IKE：IPsec 中的密钥管理	420	人物专访	478
8.8 使无线 LAN 安全	420	参考文献	480
8.8.1 有线等效保密	421		
8.8.2 IEEE 802.11i	422		

计算机网络和因特网

今天的因特网无疑是有史以来由人类创造、精心设计的最大系统，该系统具有数以亿计的相连的计算机、通信链路和交换机，有数十亿的用便携计算机、平板电脑和智能手机连接的用户，并且还有一批与因特网连接的“物品”，包括游戏机、监视系统、手表、眼镜、温度调节装置、体重计和汽车。面对如此巨大并且具有如此众多不同组件和用户的因特网，是否能够理解它的工作原理？是否存在某些指导原则和结构，能够作为理解这种规模和复杂程度惊人的系统的基础？这样的话，能让学习计算机网络成为既引人入胜又趣味盎然的事吗？幸运的是，对所有这些问题都是响亮的肯定答复。本书的目的的确是向读者介绍计算机网络这个动态领域的最新知识，使他们深入地理解网络的原则和实践，使读者不仅能理解今天的网络，而且能理解明天的网络。

第1章概述了计算机网络和因特网。这一章的目标是从整体上粗线条地勾勒出计算机网络的概貌，并且描述本书内容的框架。这一章包括大量的背景知识，讨论大量的计算机网络构件，而且将它们放在整个网络的大环境中进行讨论。

本章将以如下方式组织计算机网络的概述：在介绍了某些基本术语和概念后，将首先查看构成网络的基本硬件和软件组件。我们从网络的边缘开始，考察在网络中运行的端系统和网络应用；接下来探究计算机网络的核心，查看传输数据的链路和交换机，以及将端系统与网络核心相连接的接入网和物理媒体。我们将了解因特网是“网络的网络”，并将得知这些网络是怎样彼此连接起来的。

在浏览完计算机网络的边缘和核心之后，本章的后半部分将从更广泛、更抽象的角度来考察计算机网络。我们将考察在计算机网络中数据的时延、丢包和吞吐量，给出一个端到端吞吐量和时延的简单定量模型：该模型兼顾了传输、传播和排队时延等因素。接下来，我们将介绍计算机联网时的一些关键的体系结构原则，如协议分层和服务模型。我们还将了解到计算机网络对于许多不同类型的攻击来说是脆弱的，将回顾其中的某些攻击并且考虑使计算机网络更为安全的方法。最后，我们将以计算机网络的简要历史结束本章的学习。

1.1 什么是因特网

在本书中，我们使用一种特定的计算机网络，即公共因特网，作为讨论计算机网络及其协议的主要载体。但什么是因特网？回答这个问题有两种方式：其一，我们能够描述因特网的具体构成，即构成因特网的基本硬件和软件组件；其二，我们能够根据为分布式应用提供服务的联网基础设施来描述因特网。我们先从描述因特网的具体构成开始，并用图1-1举例说明我们的讨论。

1.1.1 具体构成描述

因特网是一个世界范围的计算机网络，即它是一个互联了遍及全世界数十亿计算设备的网络。在不久前，这些计算设备多数是传统的桌面PC、Linux工作站以及所谓的服务器

(它们用于存储和传输 Web 页面和电子邮件报文等信息)。然而,越来越多的非传统的因特网“物品”(如便携机、智能手机、平板电脑、电视、游戏机、温度调节装置、家用安全系统、家用电器、手表、眼镜、汽车、运输控制系统等)正在与因特网相连。的确,在许多非传统设备连接到因特网的情况下,计算机网络 (computer network) 这个术语开始听起来有些过时了。用因特网术语来说,所有这些设备都称为主机 (host) 或端系统 (end system)。据估计,到 2015 年有大约 50 亿台设备与因特网连接,而到 2020 年该数字将达 250 亿 [Gartner 2014]。2015 年,全世界已有超过 32 亿因特网用户,接近世界人口的 40% [ITU 2015]。

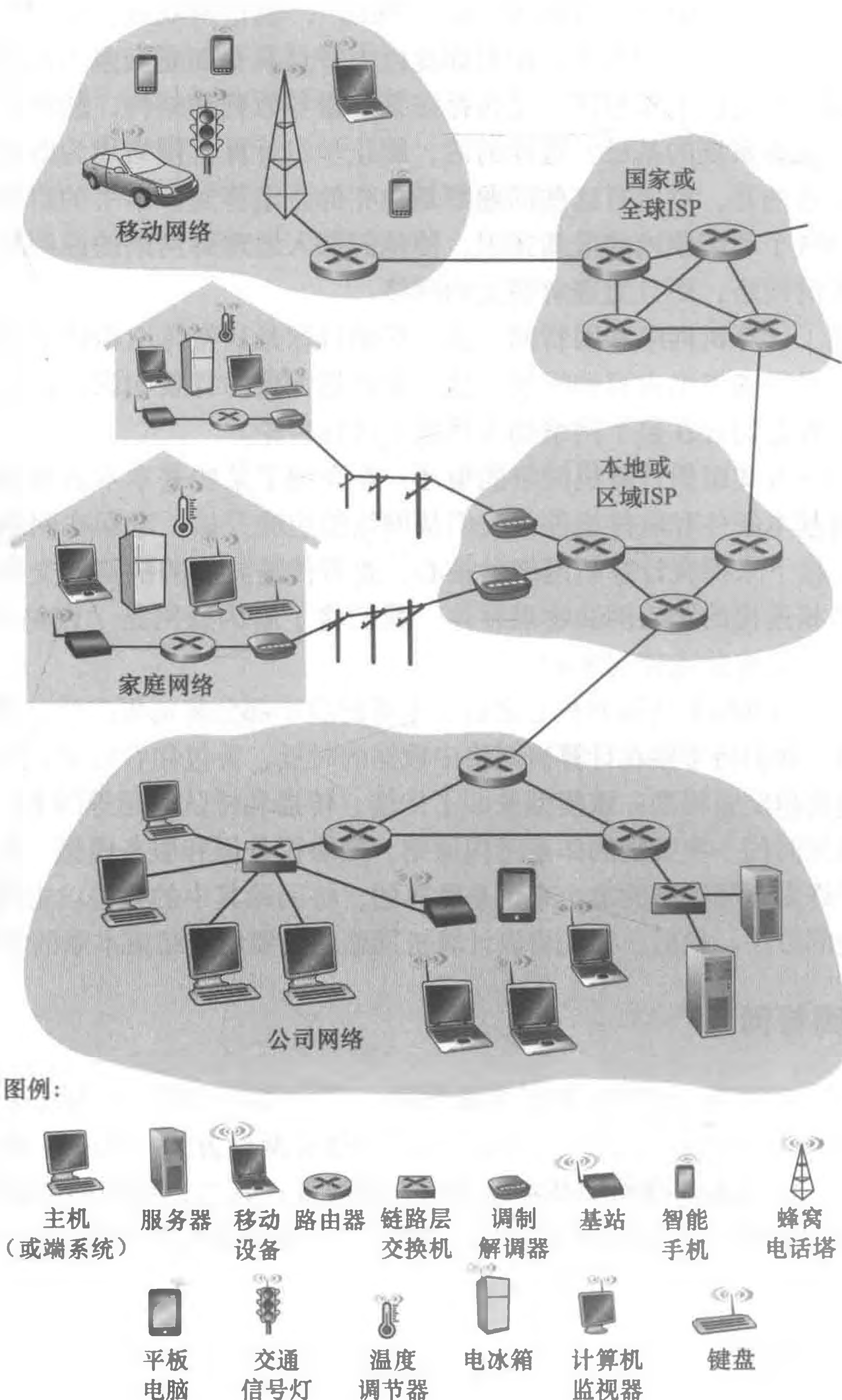


图 1-1 因特网的一些部件

端系统通过通信链路 (communication link) 和分组交换机 (packet switch) 连接到一起。在 1.2 节中, 我们将介绍许多类型的通信链路, 它们由不同类型的物理媒体组成。这些物理媒体包括同轴电缆、铜线、光纤和无线电频谱。不同的链路能够以不同的速率传输数据, 链路的传输速率 (transmission rate) 以比特/秒 (bit/s, 或 bps) 度量。当一台端系统要向另一台端系统发送数据时, 发送端系统将数据分段, 并为每段加上首部字节。由此形成的信息包用计算机网络的术语来说称为分组 (packet)。这些分组通过网络发送到目的端系统, 在那里被装配成初始数据。

分组交换机从它的一条入通信链路接收到达的分组, 并从它的一条出通信链路转发该分组。市面上流行着各种类型、各具特色的分组交换机, 但在当今的因特网中, 两种最著名的类型是路由器 (router) 和链路层交换机 (link-layer switch)。这两种类型的交换机朝着最终目的地转发分组。链路层交换机通常用于接入网中, 而路由器通常用于网络核心中。从发送端系统到接收端系统, 一个分组所经历的一系列通信链路和分组交换机称为通过该网络的路径 (route 或 path)。思科公司估计到 2016 年年底全球年度 IP 流量将超过泽字节 (zettabyte, 10^{21} 字节), 到 2019 年将达每年 2 泽字节 [Cisco VNI 2015]。

用于传送分组的分组交换网络在许多方面类似于承载运输车辆的运输网络, 该网络包括了高速公路、公路和交叉口。例如, 考虑下列情况, 一个工厂需要将大量货物搬运到数千公里以外的某个目的地仓库。在工厂中, 货物要分开并装上卡车车队。然后, 每辆卡车独立地通过高速公路、公路和立交桥组成的网络向仓库运送货物。在目的地仓库, 卸下这些货物, 并且与一起装载的同一批货物的其余部分堆放在一起。因此, 在许多方面, 分组类似于卡车, 通信链路类似于高速公路和公路, 分组交换机类似于交叉口, 而端系统类似于建筑物。就像卡车选取运输网络的一条路径前行一样, 分组则选取计算机网络的一条路径前行。

端系统通过因特网服务提供商 (Internet Service Provider, ISP) 接入因特网, 包括如本地电缆或电话公司那样的住宅区 ISP、公司 ISP、大学 ISP, 在机场、旅馆、咖啡店和其他公共场所提供 WiFi 接入的 ISP, 以及为智能手机和其他设备提供移动接入的蜂窝数据 ISP。每个 ISP 自身就是一个由多台分组交换机和多段通信链路组成的网络。各 ISP 为端系统提供了各种不同类型的网络接入, 包括如线缆调制解调器或 DSL 那样的住宅宽带接入、高速局域网接入和移动无线接入。ISP 也为内容提供者提供因特网接入服务, 将 Web 站点和视频服务器直接连入因特网。因特网就是将端系统彼此互联, 因此为端系统提供接入的 ISP 也必须互联。较低层的 ISP 通过国家的、国际的较高层 ISP (如 Level 3 Communications、AT&T、Sprint 和 NTT) 互联起来。较高层 ISP 是由通过高速光纤链路互联的高速路由器组成的。无论是较高层还是较低层 ISP 网络, 它们每个都是独立管理的, 运行着 IP 协议 (详情见后), 遵从一定的命名和地址规则。我们将在 1.3 节中更为详细地考察 ISP 及其互联的情况。

端系统、分组交换机和其他因特网部件都要运行一系列协议 (protocol), 这些协议控制因特网中信息的接收和发送。TCP (Transmission Control Protocol, 传输控制协议) 和 IP (Internet Protocol, 网际协议) 是因特网中两个最为重要的协议。IP 协议定义了路由器和端系统之间发送和接收的分组格式。因特网的主要协议统称为 TCP/IP。我们在这一章中就开始接触这些协议。但这仅仅是个开始, 本书的许多地方与计算机网络协议有关。

鉴于因特网协议的重要性, 每个人就各个协议及其作用取得一致认识是很重要的, 这样人们就能够创造协同工作的系统和产品。这正是标准发挥作用的地方。因特网标准

(Internet standard) 由因特网工程任务组 (Internet Engineering Task Force, IETF) [IETF 2016] 研发。IETF 的标准文档称为请求评论 (Request For Comment, RFC)。RFC 最初只是普通的请求评论 (因此而得名), 目的是解决因特网先驱者们面临的网络和协议问题 [Allman 2011]。RFC 文档往往是技术性很强并相当详细的。它们定义了 TCP、IP、HTTP (用于 Web) 和 SMTP (用于电子邮件) 等协议。目前已经有将近 7000 个 RFC。其他组织也在制定用于网络组件的标准, 最引人注目的是针对网络链路的标准。例如, IEEE 802 LAN/MAN 标准化委员会 [IEEE 802 2016] 制定了以太网和无线 WiFi 的标准。

1.1.2 服务描述

前面的讨论已经辨识了构成因特网的许多部件。但是我们也能从一个完全不同的角度, 即从为应用程序提供服务的基础设施的角度来描述因特网。除了诸如电子邮件和 Web 冲浪等传统应用外, 因特网应用还包括移动智能手机和平板电脑应用程序, 其中包括即时讯息、与实时道路流量信息的映射、来自云的音乐流、电影和电视流、在线社交网络、视频会议、多人游戏以及基于位置的推荐系统。因为这些应用程序涉及多个相互交换数据的端系统, 故它们被称为分布式应用程序 (distributed application)。重要的是, 因特网应用程序运行在端系统上, 即它们并不运行在网络核心中的分组交换机中。尽管分组交换机能够加速端系统之间的数据交换, 但它们并不在意作为数据的源或宿的应用程序。

我们稍深入地探讨一下为应用程序提供服务的基础设施的含义。为此, 假定你对某种分布式因特网应用有一个激动人心的新想法, 它可能大大地造福于人类, 或者它可能直接使你名利双收。你将如何把这种想法转换成一种实际的因特网应用呢? 因为应用程序运行在端系统上, 所以你需要编写运行在端系统上的一些软件。例如, 你可能用 Java、C 或 Python 编写软件。此时, 因为你在研发一种分布式因特网应用程序, 运行在不同端系统上的软件将需要互相发送数据。此时我们碰到一个核心问题, 这个问题导致了另一种描述因特网的方法, 即将因特网描述为应用程序的平台。运行在一个端系统上的应用程序怎样才能指令因特网向运行在另一个端系统上的软件发送数据呢?

与因特网相连的端系统提供了一个套接字接口 (socket interface), 该接口规定了运行在一个端系统上的程序请求因特网基础设施向运行在另一个端系统上的特定目的地程序交付数据的方式。因特网套接字接口是一套发送程序必须遵循的规则集合, 因此因特网能够将数据交付给目的地。我们将在第 2 章详细讨论因特网套接字接口。此时, 我们做一个简单的类比, 在本书中我们将经常使用这个类比。假定 Alice 使用邮政服务向 Bob 发一封信。当然, Alice 不能只是写了这封信 (相关数据) 然后把该信丢出窗外。相反, 邮政服务要求 Alice 将信放入一个信封中; 在信封的中间写上 Bob 的全名、地址和邮政编码; 封上信封; 在信封的右上角贴上邮票; 最后将该信封丢进一个邮局的邮政服务信箱中。因此, 该邮政服务有自己的“邮政服务接口”或一套规则, 这是 Alice 必须遵循的, 这样邮政服务才能将她的信件交付给 Bob。同理, 因特网也有一个发送数据的程序必须遵循的套接字接口, 使因特网向接收数据的程序交付数据。

当然, 邮政服务向顾客提供了多种服务, 如特快专递、挂号、普通服务等。同样, 因特网向应用程序提供了多种服务。当你研发一种因特网应用程序时, 也必须为你的应用程序选择其中的一种因特网服务。我们将在第 2 章中描述因特网服务。

我们已经给出了因特网的两种描述方法: 一种是根据它的硬件和软件组件来描述, 另

一种是根据基础设施向分布式应用程序提供的服务来描述。但是，你也许还是对什么是因特网感到困惑。什么是分组交换和 TCP/IP？什么是路由器？因特网中正在使用什么样的通信链路？什么是分布式应用？一个恒温调节器或人体秤如何与因特网相连？如果你现在还对这些心存疑惑，请不要担心。这本书除了向你介绍因特网的具体构成外，还要介绍支配因特网的工作原理以及它工作的来龙去脉。我们将在后续章节中解释这些重要的术语和问题。

1.1.3 什么是协议

既然我们已经对因特网是什么有了一点印象，那么下面考虑计算机网络中另一个重要的时髦术语：协议（protocol）。什么是协议？协议是用来干什么的？

1. 人类活动的类比

也许理解计算机网络协议这一概念的一个最容易办法是，先与某些人类活动进行类比，因为我们人类无时无刻不在执行协议。考虑当你想要向某人询问时间时将要怎样做。图 1-2 中显示了一种典型的交互过程。人类协议（至少说是好的行为方式）要求一方首先进行问候（图 1-2 中的第一个“你好”），以开始与另一个人的通信。对“你好”的典型响应是返回一个“你好”报文。此人用一个热情的“你好”进行响应，隐含着一种指示，表明能够继续向那人询问时间了。对最初的“你好”的不同响应（例如“不要烦我！”，或“我不会说英语”，或某些不合时宜的回答）也许表明了一个勉强的或不能进行的通信。在此情况下，按照人类协议，发话者就不能询问时间了。有时，问的问题根本得不到任何回答，在此情况下，发话者通常会放弃向这个人询问时间。注意在我们人类协议中，有我们发送的特定报文，也有我们根据接收到的应答报文或其他事件（例如在某个给定的时间内没有回答）采取的动作。显然，发送和接收的报文，以及这些报文发送和接收或其他事件出现时所采取的动作，这些在一个人类协议中起到了核心作用。如果人们使用不同

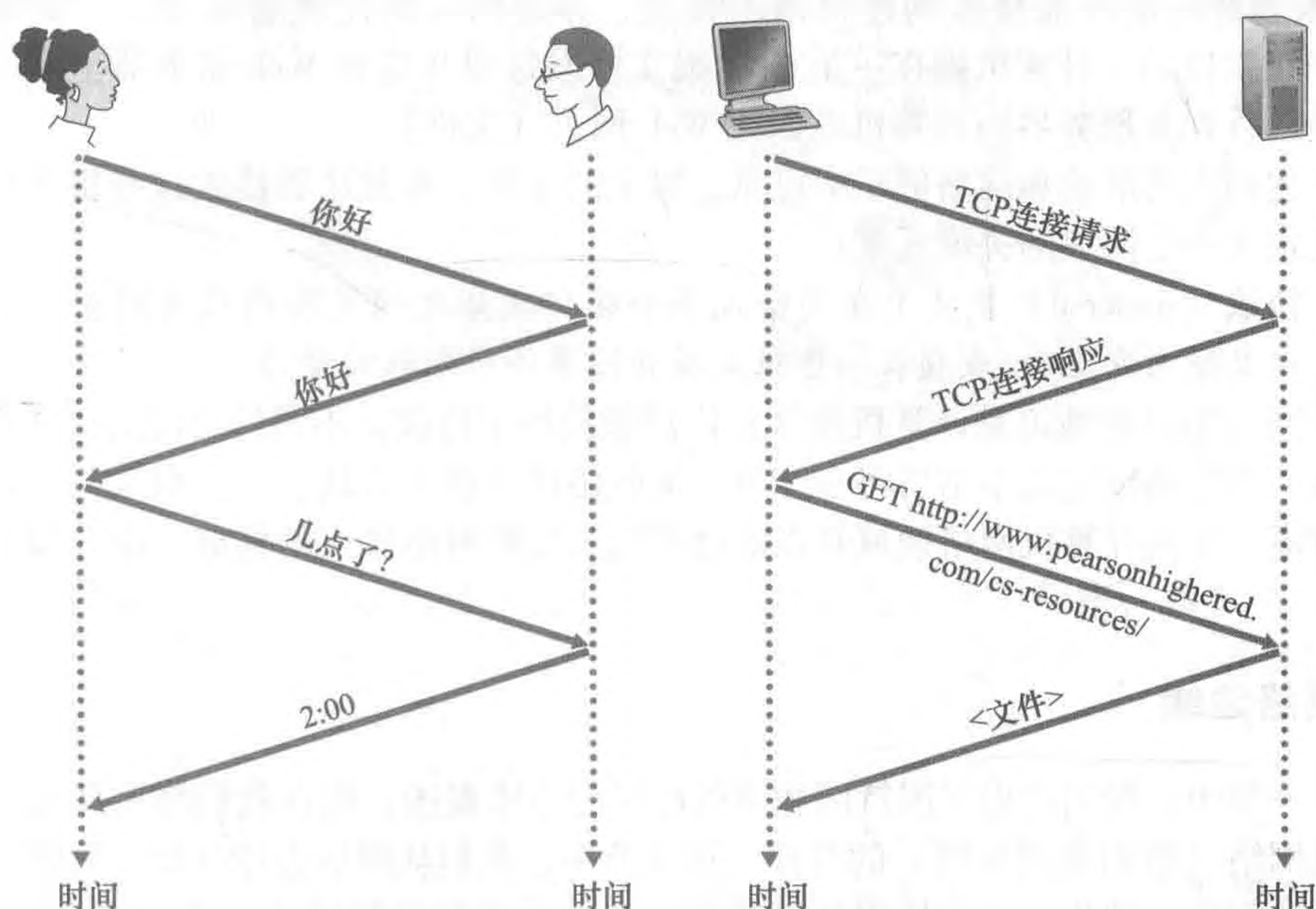


图 1-2 人类协议和计算机网络协议

的协议（例如，如果一个人讲礼貌，而另一人不讲礼貌，或一个人明白时间的概念，而另一人却不理解），这些协议就不能交互，因而不能完成有用的工作。在网络中这个道理同样成立。即为了完成一项工作，要求两个（或多个）通信实体运行相同的协议。

我们再考虑第二个人类类比的例子。假定你正在大学课堂里上课（例如上的是计算机网络课程）。教师正在唠唠叨叨地讲述协议，而你困惑不解。这名教师停下来问：“同学们有什么问题吗？”（教师发送出一个报文，该报文被所有没有睡觉的学生接收到了。）你举起了手（向教师发送了一个隐含的报文）。这位教师面带微笑地示意你说：“请讲……”（教师发出的这个报文鼓励你提出问题，教师喜欢被问问题。）接着你就问了问题（向该教师传输了你的报文）。教师听取了你的问题（接收了你的问题报文）并加以回答（向你传输了回答报文）。我们再一次看到了报文的发送和接收，以及这些报文发送和接收时所采取的一系列约定俗成的动作，这些是这个“提问与回答”协议的关键所在。

2. 网络协议

网络协议类似于人类协议，除了交换报文和采取动作的实体是某些设备（可以是计算机、智能手机、平板电脑、路由器或其他具有网络能力的设备）的硬件或软件组件。在因特网中，涉及两个或多个远程通信实体的所有活动都受协议的制约。例如，在两台物理上连接的计算机中，硬件实现的协议控制了在这两块网络接口卡间的“线上”的比特流；在端系统中，拥塞控制协议控制了在这发送方和接收方之间传输的分组发送的速率；路由器中的协议决定了分组从源到目的地的路径。在因特网中协议运行无处不在，因此本书的大量篇幅都与计算机网络协议有关。

以大家可能熟悉的一个计算机网络协议为例，考虑当你向一个 Web 服务器发出请求（即你在 Web 浏览器中键入一个 Web 网页的 URL）时所发生的情况。图 1-2 右半部分显示了这种情形。首先，你的计算机将向该 Web 服务器发送一条连接请求报文，并等待回答。该 Web 服务器将最终能接收到连接请求报文，并返回一条连接响应报文。得知请求该 Web 文档正常以后，计算机则在一条 GET 报文中发送要从这台 Web 服务器上取回的网页名字。最后，Web 服务器向计算机返回该 Web 网页（文件）。

从上述的人类活动和网络例子中可见，报文的交换以及发送和接收这些报文时所采取的动作是定义一个协议的关键元素：

协议（protocol）定义了在两个或多个通信实体之间交换的报文的格式和顺序，以及报文发送和/或接收一条报文或其他事件所采取的动作。

因特网（更一般地说是计算机网络）广泛地使用了协议。不同的协议用于完成不同的通信任务。当你阅读完这本书后将会知道，某些协议简单而直截了当，而某些协议则复杂且晦涩难懂。掌握计算机网络领域知识的过程就是理解网络协议的构成、原理和工作方式的过程。

1.2 网络边缘

在上一节中，我们给出了因特网和网络协议的总体概述。现在我们将更深入一些来探究计算机网络（特别是因特网）的部件。在本节中，我们从网络边缘开始，观察一下我们更为熟悉的部件，即我们日常使用的计算机、智能手机和其他设备。在接下来的一节中，我们将从网络边缘向网络核心推进，查看计算机网络中的交换和选路。

回想前一节中计算机网络的术语，通常把与因特网相连的计算机和其他设备称为端系统。如图 1-3 所示，因为它们位于因特网的边缘，故而被称为端系统。因特网的端系统包括了桌面计算机（例如，桌面 PC、Mac 和 Linux 设备）、服务器（例如，Web 和电子邮件服务器）和移动计算机（例如，便携机、智能手机和平板电脑）。此外，越来越多的非传统物品正被作为端系统与因特网相连（参见“历史事件”）。

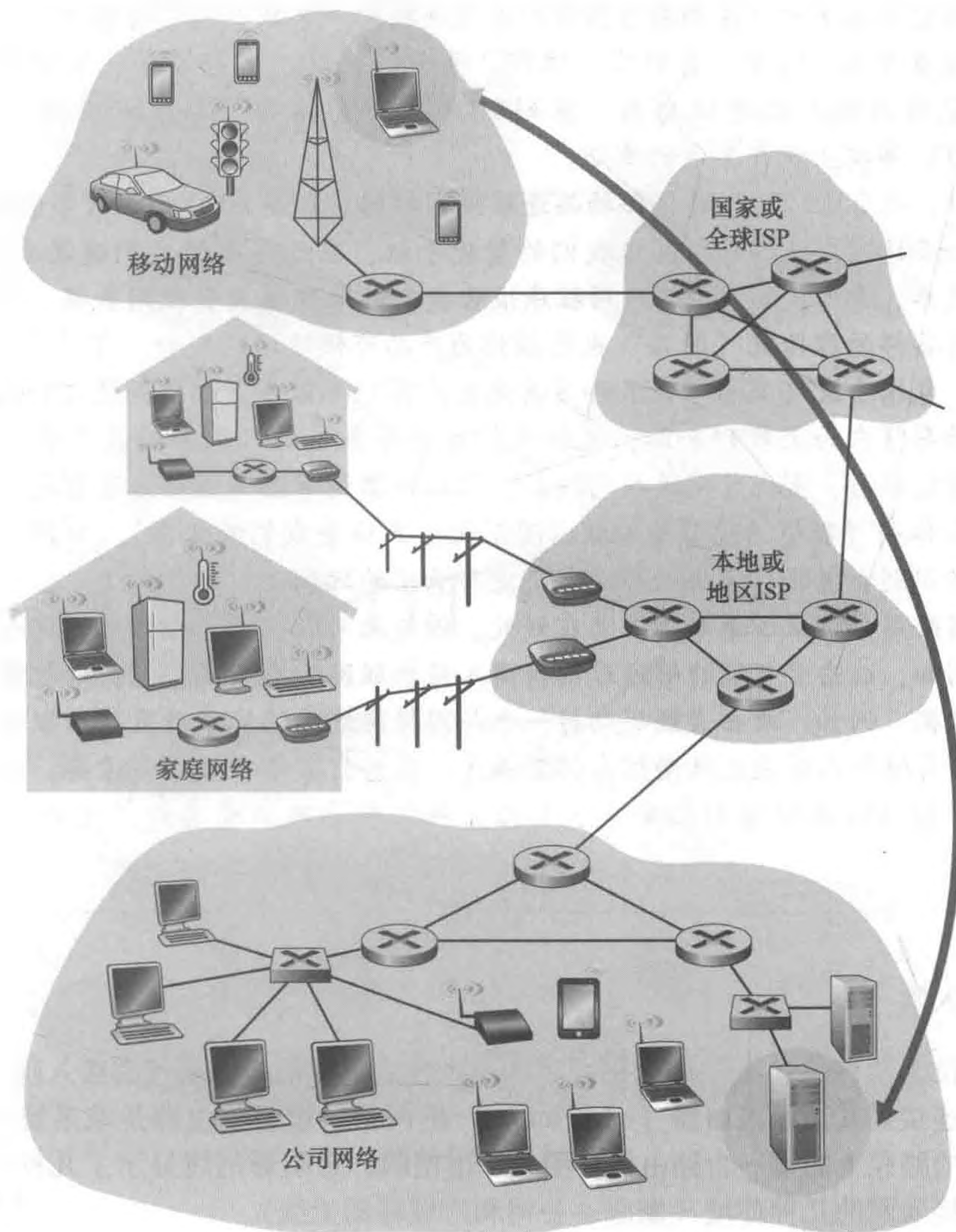


图 1-3 端系统交互

端系统也称为主机（host），因为它们容纳（即运行）应用程序，如 Web 浏览器程序、Web 服务器程序、电子邮件客户程序或电子邮件服务器程序等。本书通篇将交替使用主机和端系统这两个术语，即主机 = 端系统。主机有时又被进一步划分为两类：客户（client）和服务器（server）。客户通常是桌面 PC、移动 PC 和智能手机等，而服务器通常是更为强大的机器，用于存储和发布 Web 页面、流视频、中继电子邮件等。今天，大部分提供搜索结果、电子邮件、Web 页面和视频的服务器都属于大型数据中心（data center）。例如，谷歌公司（Google）拥有 50 ~ 100 个数据中心，其中 15 个大型数据中

心每个都有 10 万台以上的服务器。

历史事件

物联网

你能够想象几乎每样东西都与因特网无线连接的一个世界吗？能够想象在这个世界上连接了大多数人、汽车、自行车、眼镜、手表、玩具、医疗设备、家用传感器、教室、视频监控系統、大气传感器、陈列架产品和宠物吗？这种物联网（Internet of Things, IoT）事实上就在不远的某处。

据估计，在 2015 年有 50 亿件物品连接到因特网，到了 2020 年该数字能够达到 250 亿 [Gartner 2014]。这些物品包括我们的智能手机，它已经跟随我们遍及我们的家庭、办公室和汽车，向 ISP 和因特网应用程序报告我们的地理位置和使用数据。而除了智能手机，各种各样的非传统“物品”也已经作为产品可供使用。例如，有因特网连接的可穿戴设备，包括手表（来自苹果和许多其他生产商）和眼镜。因特网连接的眼镜能够让我们看到的每样东西上载到云端，允许我们与世界各地的人们实时共享我们的视觉体验。对于智能家居，有因特网连接的物品，包括恒温调节器（能够通过智能手机远程控制）以及人体秤（能够通过智能手机以图形化方式审查我们的饮食）。有因特网连接的玩具，包括识别和解释孩子讲话并适当地进行响应的玩偶。

物联网给用户带来了革命性的潜在好处。但与此同时，也存在巨大的安全风险和隐私风险。例如，攻击者可能能够通过因特网入侵物联网设备或服务器，以收集来自物联网设备的数据。例如，攻击者能够劫持一个与因特网连接的玩偶并直接与孩子交谈，或者能够入侵存储个人健康或活动信息的数据库，这些信息来自可穿戴设备。这些安全和隐私问题可能会使消费者渐渐地失去信心，并可能导致消费者的广泛接受程度下降 [FTC 2015]。

1.2.1 接入网

考虑了位于“网络边缘”的应用程序和端系统后，我们接下来考虑接入网，这是指将端系统物理连接到其边缘路由器（edge router）的网络。边缘路由器是端系统到任何其他远程端系统的路径上的第一台路由器。图 1-4 用粗的、带阴影的线显示了几种类型的接入链路和使用接入网的几种环境（家庭、公司和广域移动无线）。

1. 家庭接入：DSL、电缆、FTTH、拨号和卫星

2014 年，在发达国家中，实现因特网接入的家庭超过 78%，名列前茅的韩国、荷兰、芬兰和瑞典实现因特网接入的家庭则超过 80%，几乎所有家庭都是经过高速宽带连接 [ITU 2015]。考虑到家庭对网络接入的强烈兴趣，我们从家庭与因特网的连接开始介绍接入网概况。

今天，宽带住宅接入有两种最流行的类型：数字用户线（Digital Subscriber Line, DSL）和电缆。住户通常从提供本地电话接入的本地电话公司处获得 DSL 因特网接入。因此，当使用 DSL 时，用户的本地电话公司也是它的 ISP。如图 1-5 所示，每个用户的 DSL 调制解调器使用现有的电话线（即双绞铜线，将在 1.2.2 节中讨论它）与位于电话公司的

本地中心局（CO）中的数字用户线接入复用器（DSLAM）交换数据。家庭的 DSL 调制解调器得到数字数据后将其转换为高频音，以通过电话线传输给本地中心局；来自许多家庭的模拟信号在 DSLAM 处被转换回数字形式。

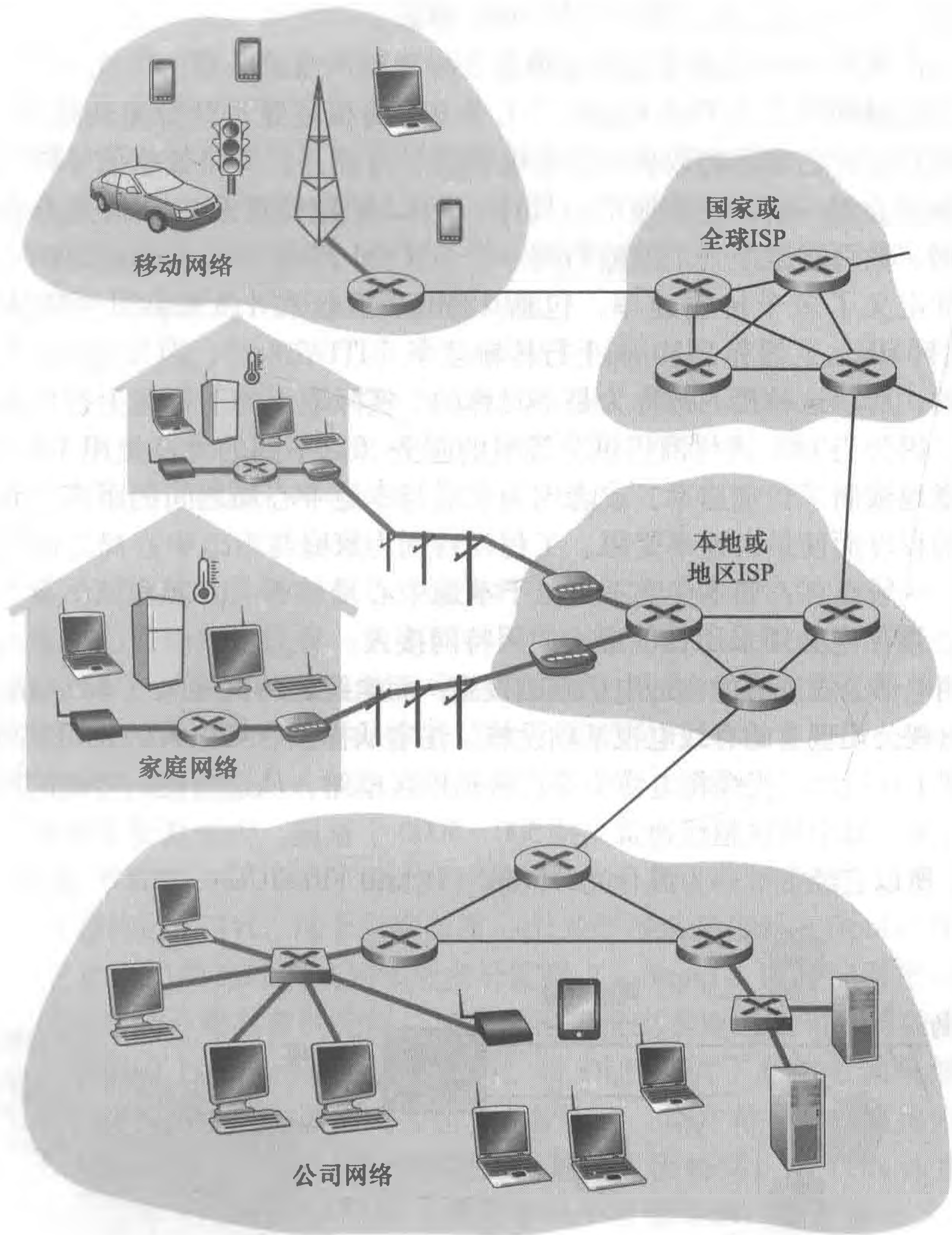


图 1-4 接入网

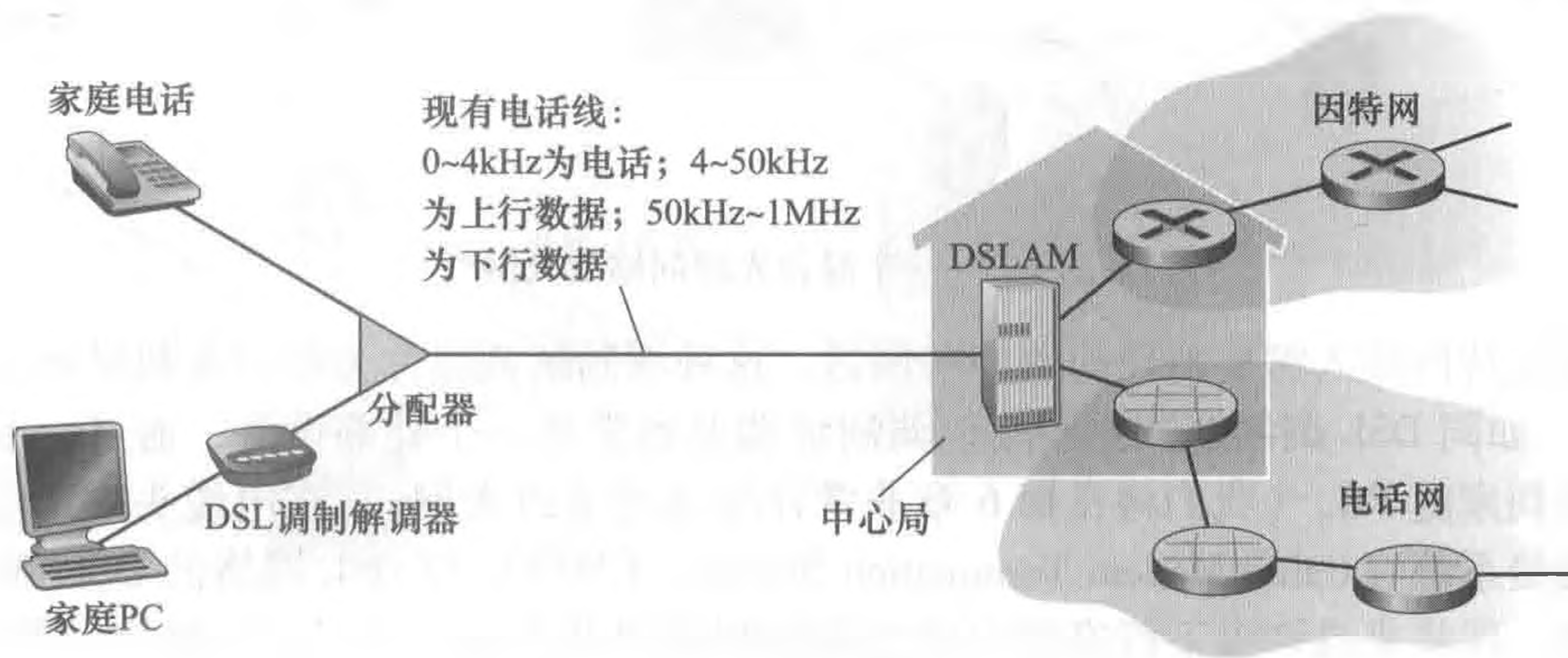


图 1-5 DSL 因特网接入

家庭电话线同时承载了数据和传统的电话信号，它们用不同的频率进行编码：

- 高速下行信道，位于 50kHz 到 1MHz 频段；
- 中速上行信道，位于 4kHz 到 50kHz 频段；
- 普通的双向电话信道，位于 0 到 4kHz 频段。

这种方法使单根 DSL 线路看起来就像有 3 根单独的线路一样，因此一个电话呼叫和一个因特网连接能够同时共享 DSL 链路。（1.3.1 节将描述这种频分复用技术。）在用户一侧，一个分配器把到达家庭的数据信号和电话信号分隔开，并将数据信号转发给 DSL 调制解调器。在电话公司一侧，在本地中心局中，DSLAM 把数据和电话信号分隔开，并将数据送往因特网。数百甚至上千个家庭与同一个 DSLAM 相连 [Dischinger 2007]。

DSL 标准定义了多个传输速率，包括 12Mbps 下行和 1.8Mbps 上行传输速率 [ITU 1999]，以及 55Mbps 下行和 15Mbps 上行传输速率 [ITU 2006]。因为这些上行速率和下行速率是不同的，所以这种接入被称为是不对称的。实际取得的下行和上行传输速率也许小于上述速率，因为当 DSL 提供商提供分等级的服务（以不同的价格使用不同的速率）时，他们也许有意地限制了住宅速率，或者因为家庭与本地中心局之间的距离、双绞线的规格和电气干扰的程度而使最大速率受限。工程师特别为家庭与本地中心局之间的短距离接入设计了 DSL；一般而言，如果住宅不是位于本地中心局的 5 ~ 10 英里（1 英里 = 1609.344 米）范围内，该住宅必须采用其他形式的因特网接入。

DSL 利用电话公司现有的本地电话基础设施，而**电缆因特网接入**（cable Internet access）利用了有线电视公司现有的有线电视基础设施。住宅从提供有线电视的公司获得了电缆因特网接入。如图 1-6 所示，光缆将电缆头端连接到地区枢纽，从这里使用传统的同轴电缆到达各家各户和公寓。每个地区枢纽通常支持 500 ~ 5000 个家庭。因为在这个系统中应用了光纤和同轴电缆，所以它经常被称为混合光纤同轴（Hybrid Fiber Coax, HFC）系统。

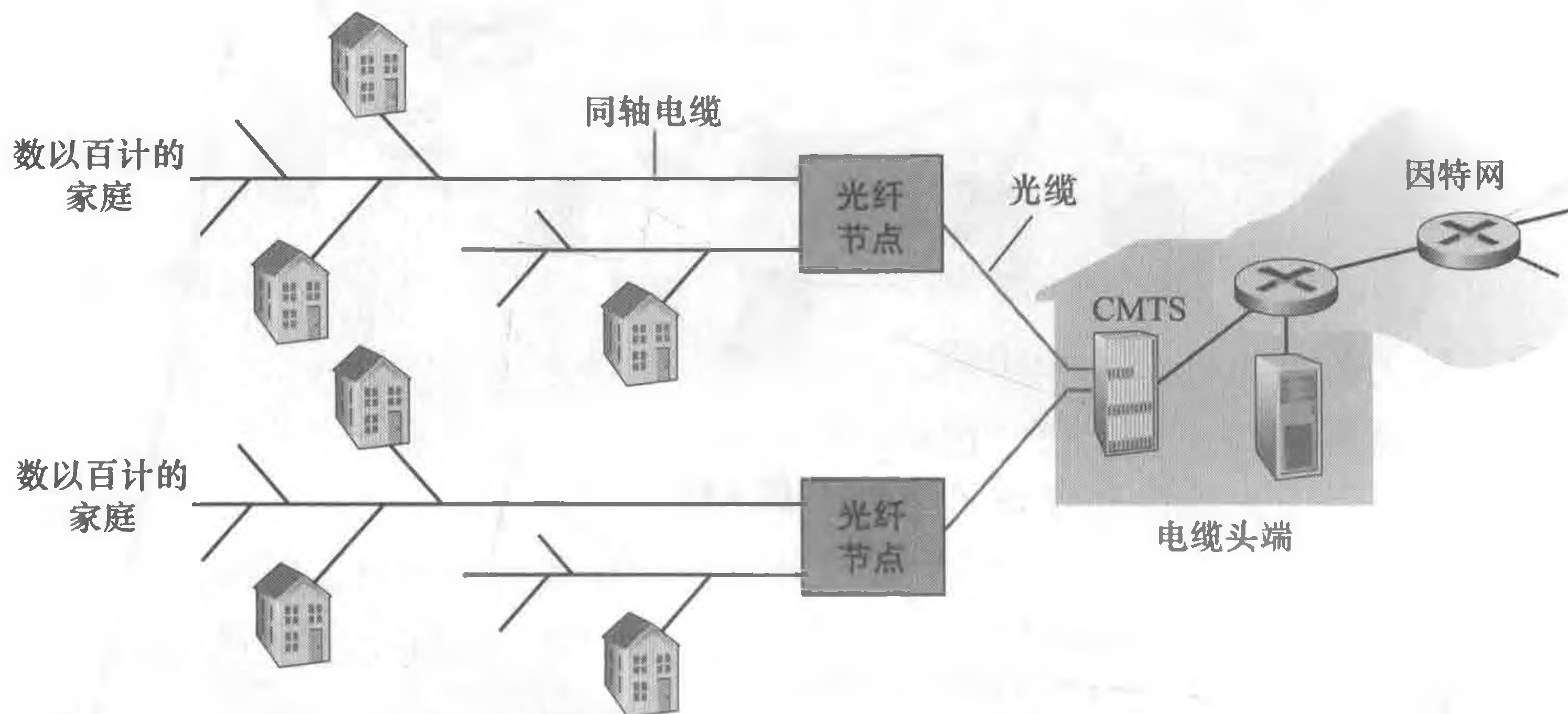


图 1-6 一个混合光纤同轴接入网

电缆因特网接入需要特殊的调制解调器，这种调制解调器称为**电缆调制解调器**（cable modem）。如同 DSL 调制解调器，电缆调制解调器通常是一个外部设备，通过一个以太网端口连接到家庭 PC。（我们将在第 6 章非常详细地讨论以太网。）在电缆头端，电缆调制解调器端接系统（Cable Modem Termination System, CMTS）与 DSL 网络的 DSLAM 具有类似的功能，即将来自许多下行家庭中的电缆调制解调器发送的模拟信号转换回数字形式。电缆调制解调器将 HFC 网络划分为下行和上行两个信道。如同 DSL，接入通常是不对称

的，下行信道分配的传输速率通常比上行信道的高。DOCSIS 2.0 标准定义了高达 42.8Mbps 的下行速率和高达 30.7Mbps 的上行速率。如在 DSL 网络中的情况那样，由于较低的数据率或媒体损耗，可能不一定能达到最大可取得的速率。

电缆因特网接入的一个重要特征是共享广播媒体。特别是，由头端发送的每个分组向下行经每段链路到每个家庭；每个家庭发送的每个分组经上行信道向头端传输。因此，如果几个用户同时经下行信道下载一个视频文件，每个用户接收视频文件的实际速率将大大低于电缆总计的下行速率。而另一方面，如果仅有很少的活跃用户在进行 Web 冲浪，则每个用户都可以以全部的下行速率接收 Web 网页，因为用户们很少在完全相同的时刻请求网页。因为上行信道也是共享的，需要一个分布式多路访问协议来协调传输和避免碰撞。（我们将在第 6 章中更为详细地讨论碰撞问题。）

尽管 DSL 和电缆网络当前代表了超过 85% 的美国住宅宽带接入，但一种提供更高速率的新兴技术是光纤到户（Fiber To The Home, FTTH）[FTTH Council 2016]。顾名思义，FTTH 概念简单，从本地中心局直接到家庭提供了一条光纤路径。现在，许多国家或地区，包括阿拉伯联合酋长国、韩国、中国香港、日本、新加坡、中国台湾、立陶宛和瑞典，入户率已经超过了 30% [FTTH Council 2016]。

从本地中心局到家庭有几种有竞争性的光纤分布方案。最简单的光纤分布网络称为直接光纤，从本地中心局到每户设置一根光纤。更为一般的是，从中心局出来的每根光纤实际上由许多家庭共享，直到相对接近这些家庭的位置，该光纤才分成每户一根光纤。进行这种分配有两种有竞争性的光纤分布体系结构：主动光纤网络（Active Optical Network, AON）和被动光纤网络（Passive Optical Network, PON）。AON 本质上就是交换因特网，我们将在第 6 章讨论它。

这里，我们简要讨论一下 PON，该技术用于 Verizon 的 FIOS 服务中。图 1-7 显示了使用 PON 分布体系结构的 FTTH。每个家庭具有一个光纤网络端接器（Optical Network Terminator, ONT），它由专门的光纤连接到邻近的分配器（splitter）。该分配器把一些家庭（通常少于 100 个）集结到一根共享的光纤，该光纤再连接到本地电话和公司的中心局中的光纤线路端接器（Optical Line Terminator, OLT）。该 OLT 提供了光信号和电信号之间的转换，经过本地电话公司路由器与因特网相连。在家庭中，用户将一台家庭路由器（通常是无线路由器）与 ONT 相连，并经过这台家庭路由器接入因特网。在 PON 体系结构中，所有从 OLT 发送到分配器的分组在分配器（类似于一个电缆头端）处复制。

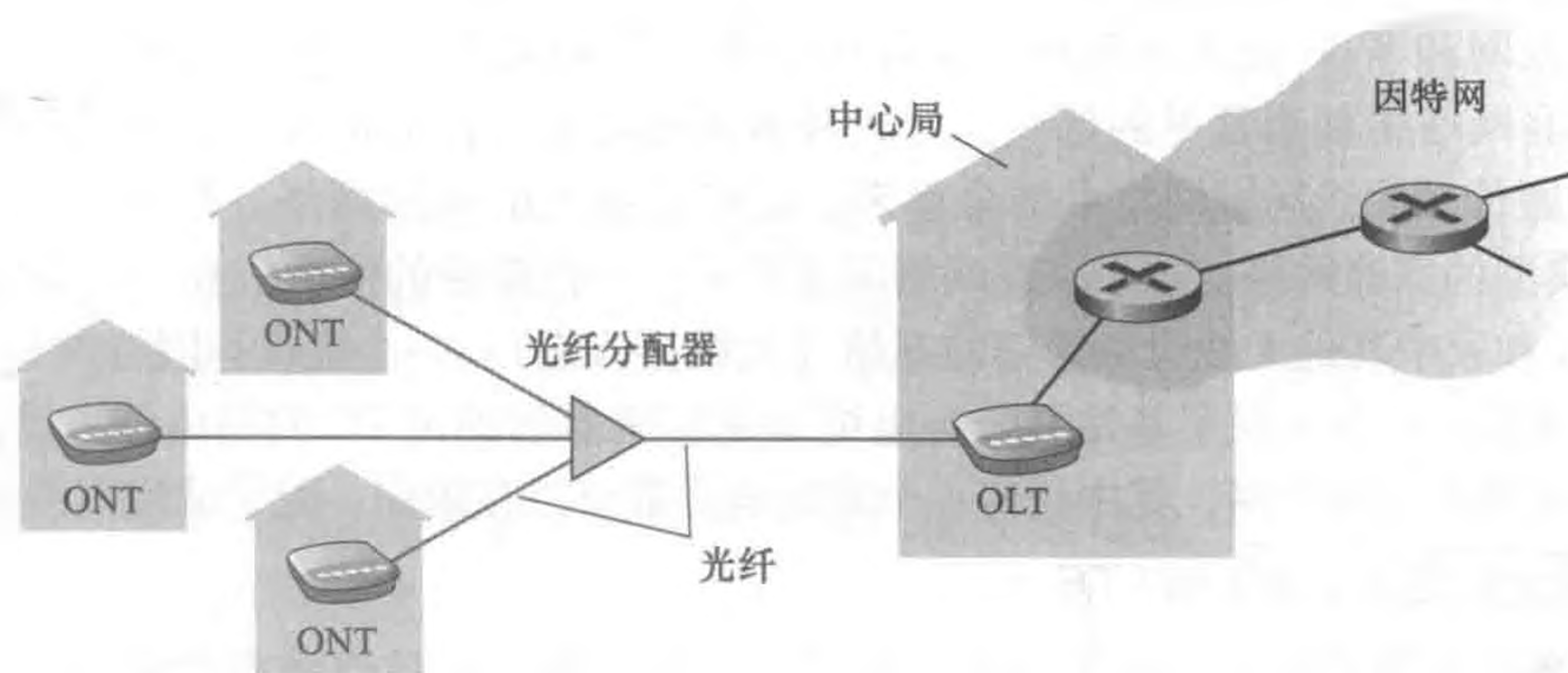


图 1-7 FTTH 因特网接入

FTTH 有潜力提供每秒千兆比特范围的因特网接入速率。然而，大多数 FTTH ISP 提供

多种不同速率选择，较高的速率自然花费更多。美国 FTTH 用户在 2011 年的平均下行速率大约为 20Mbps（与 13Mbps 电缆接入网相当，而小于 5Mbps DSL）[FTTH Council 2011b]。

还可采用另外两种接入网技术为家庭提供因特网接入。在无法提供 DSL、电缆和 FTTH 的地方（例如在某些乡村环境），能够使用卫星链路将住宅以超过 1Mbps 的速率与因特网相连。StarBand 和 HughesNet 是两家这样的卫星接入提供商。使用传统电话线的拨号接入与 DSL 基于相同的模式：家庭的调制解调器经过电话线连接到 ISP 的调制解调器。与 DSL 和其他宽带接入网相比，拨号接入 56kbps 的慢速率是令人痛苦的。

2. 企业（和家庭）接入：以太网和 WiFi

在公司和大学校园以及越来越多的家庭环境中，使用局域网（LAN）将端系统连接到边缘路由器。尽管有许多不同类型的局域网技术，但是以太网到目前为止是公司、大学和家庭网络中最为流行的接入技术。如图 1-8 中所示，以太网用户使用双绞铜线与一台以太网交换机相连，第 6 章中将详细讨论该技术。以太网交换机或这样相连的交换机网络，则再与更大的因特网相连。使用以太网接入，用户通常以 100Mbps 或 1Gbps 速率接入以太网交换机，而服务器可能具有 1Gbps 甚至 10Gbps 的接入速率。

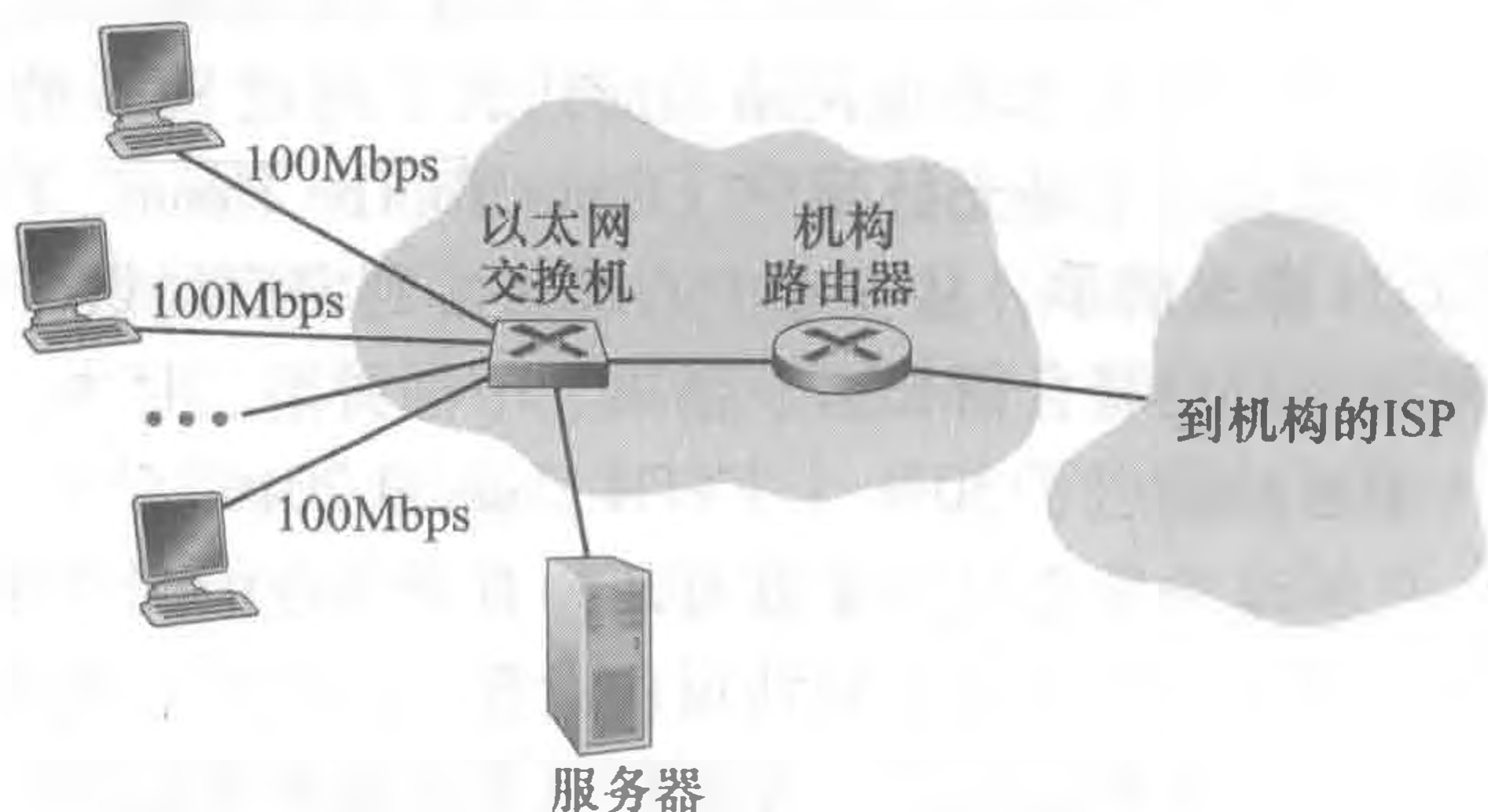


图 1-8 以太网因特网接入

然而，越来越多的人从便携机、智能手机、平板电脑和其他物品无线接入因特网（参见前面的插入内容“物联网”）。在无线 LAN 环境中，无线用户从/到一个接入点发送/接收分组，该接入点与企业网连接（很可能使用了有线以太网），企业网再与有线因特网相连。一个无线 LAN 用户通常必须位于接入点的几十米范围内。基于 IEEE 802.11 技术的无线 LAN 接入，更通俗地称为 WiFi，目前几乎无所不在，如大学、商业办公室、咖啡厅、机场、家庭，甚至在飞机上。在许多城市，人们能够站在街头而位于 10 或 20 个基站范围内（对于可浏览的 802.11 基站全局图，参见 [wicle.net 2016]，这些基站由那些非常热心做这种事情的人发现并记载在 Web 站点上）。如在第 7 章详细讨论的那样，802.11 今天提供了高达 100Mbps 的共享传输速率。

虽然以太网和 WiFi 接入网最初是设置在企业（公司或大学）环境中的，但它们近来已经成为家庭网络中相当常见的部件。今天许多家庭将宽带住宅接入（即电缆调制解调器或 DSL）与廉价的无线局域网技术结合起来，以产生强大的家用网络 [Edwards 2011]。图 1-9 显示了典型的家庭网络。这个家庭网络组成如下：一台漫游的便携机和一台有线 PC；一个与无线 PC 和家中其他无线设备通信的基站（无线接入点）；一个提供与因特网宽带接入的电缆调制解调器；一台互联了基站及带有电缆调制解调器的固定 PC 的路由器。该网络允许家庭成员经宽带接入因特网，其中任何一个家庭成员都可以在厨房、院子或卧室漫游上网。

3. 广域无线接入：3G 和 LTE

iPhone 和安卓等设备越来越多地用来在移动中发信息、在社交网络中分享照片、观看视频和放音乐。这些设备应用了与蜂窝移动电话相同的无线基础设施，通过蜂窝网提供商运营的基站来发送和接收分组。与 WiFi 不同的是，一个用户仅需要位于基站的数千米