第1章　概　　述

本章是全书的概要。在本章的开始，先介绍计算机网络在信息时代的作用。接着对互联网进行概述，包括互联网基础结构发展的三个阶段以及今后的发展趋势。然后，讨论互联网组成的边缘部分和核心部分。在简单介绍计算机网络在我国的发展以及计算机网络的类别后，讨论计算机网络的性能指标。最后，论述整个课程都要用到的重要概念——计算机网络的体系结构。

本章最重要的内容是：

(1) 互联网边缘部分和核心部分的作用，其中包含分组交换的概念。

(2) 计算机网络的性能指标。

(3) 计算机网络分层次的体系结构，包含协议和服务的概念。这部分内容比较抽象。在没有了解具体的计算机网络之前，很难完全掌握这些很抽象的概念。但这些抽象的概念又能够指导后续的学习，因此也必须先从这些概念学起。建议读者在学习到后续章节时，经常再复习一下本章中的基本概念。这对掌握好整个计算机网络的概念是有益的。

1.1　计算机网络在信息时代中的作用

我们知道，21 世纪的一些重要特征就是**数字化**、**网络化**和**信息化**，它是一个**以网络为核心的信息时代**。要实现信息化就必须依靠完善的网络，因为网络可以非常迅速地传递信息。网络现在已经成为信息社会的命脉和发展知识经济的重要基础。网络对社会生活和经济发展的很多方面已经产生了不可估量的影响。

有三大类大家很熟悉的网络，即**电信网络**、**有线电视网络**和**计算机网络**。按照最初的服务分工，电信网络向用户提供电话、电报及传真等服务；有线电视网络向用户传送各种电视节目；计算机网络则使用户能够在计算机之间传送数据文件。这三种网络在信息化过程中都起着十分重要的作用，但其中发展最快并起着核心作用的则是计算机网络，而这正是本书所要讨论的内容。

随着技术的发展，电信网络和有线电视网络都逐渐融入了现代计算机网络的技术，扩大了原有的服务范围，而计算机网络也能够向用户提供电话通信、视频通信以及传送视频节目的服务。从理论上讲，把上述三种网络融合成一种网络就能够提供所有的上述服务，这就是很早以前就提出来的“三网融合”。然而事实并不如此简单，因为这涉及各方面的经济利益和行政管辖权的问题。

20 世纪 90 年代以后，以 Internet 为代表的计算机网络得到了飞速的发展，已从最初的仅供美国人使用的免费教育科研网络逐步发展成为供全球使用的商业网络（有偿使用），成为全球最大的和最重要的计算机网络。可以毫不夸大地说，Internet 是人类自印刷术发明以来在存储和交换信息领域的最大变革。

Internet 的中文译名并不统一。现有的 Internet 译名有两种：

(1) **因特网**，这个译名是全国科学技术名词审定委员会推荐的。虽然因特网这个译名较为准确，但却长期未得到推广。

(2) **互联网**，这是目前流行最广的、事实上的标准译名。现在我国的各种报刊杂志、政府文件以及电视节目中都毫无例外地使用这个译名。Internet 是**由数量极大的各种计算机网络互连起来的**，采用互联网这个译名能够体现出 Internet 最主要的特征。

也有些人愿意直接使用英文名词 Internet，而不使用中文译名。这避免了译名的误解。但编者认为，在中文教科书中，常用的重要名词应当使用中文的。当然，对国际通用的英文缩写词，我们还是要尽量多使用。例如，直接使用更简洁的“TCP”，比使用冗长的中文译名“传输控制协议”要方便得多。这样做也更加便于阅读外文技术资料。

曾有人把 Internet 译为国际互联网。其实互联网本来就是**覆盖全球的**，因此“国际”二字显然是多余的。

对于仅在**局部范围**互连起来的计算机网络，只能称之为互连网，而不是互联网 Internet。但是我们在教材中讨论互联网的原理时，不可能画出整个互联网的连接图。因此，即使只有几个计算机相互连接起来的网络，也常常称为互联网。

有时，我们往往使用更加简洁的方式表示互联网，这就是只用一个 “**网**”字。例如，“上网”就是表示使用某个电子设备连接到互联网，而不是连接到其他的网络上。还有如网民、网吧、网银（网上银行）、网购（网上购物），等等。这里的“网”，一般都不是指电信网或有线电视网，而是指当今世界上最大的计算机网络 Internet——互联网。

那么，什么是互联网呢？很难用几句话说清楚。但我们可以从两个不同的方面来认识互联网。这就是互联网的应用和互联网的工作原理。

绝大多数人认识互联网都是从接触互联网的应用开始的。现在很小的孩子就会上网玩游戏、看网上视频，或和朋友在社交软件上聊天。人们经常利用互联网的社交电子邮件相互通信（包括传送各种照片和视频文件），这就使得传统的邮政信函的业务量大大减少。许多商品现在都可以在互联网上购买，既方便又经济实惠，改变了必须去实体商店的传统购物方式。在互联网上购买机票、火车票或预订酒店都非常方便，可以节省大量出行排队的时间。过去各银行的大厅内往往拥挤不堪，但现在这种情况已经得到了明显的改善，因为原来人们必须到银行进行的业务，基本上都可以改为在家中的网上银行进行操作。现在只要携带一个接入到互联网的智能手机就可以非常方便地付款，而不必使用现金或刷卡支付。必须指出，现在互联网的应用范围早已大大超过当初设计互联网时的几种简单的应用，并且各种意想不到的新的应用总是不断地在出现。本书不可能详细地介绍互联网的各种应用，这需要有另一本专门的书。

互联网之所以能够向用户提供许多服务，就是因为互联网具有两个重要基本特点，即**连通性**和**共享**。

所谓连通性(connectivity)，就是互联网使上网用户之间，不管相距多远（例如，相距数千公里），都可以非常便捷、非常经济地（在很多情况下甚至是免费的）交换各种信息（数据，以及各种音频、视频），**好像**这些用户终端都彼此直接连通一样。这与使用传统的电信网络有着很大的区别。我们知道，传统的电信网向用户提供的最重要的服务就是人与人之间的电话通信，因此电信网也具有连通性这个特点。但使用电信网的电话用户，往往要为此向电信网的运营商缴纳相当昂贵的费用，特别是长距离的越洋通信。但应注意，互联网具有虚拟的特点。例如，当你从互联网上收到一封电子邮件时，你可能无法准确知道对方是谁（朋友还是骗子），也无法知道发信人的地点（在附近，还是在地球对面）。

所谓共享就是指**资源共享**。资源共享的含义是多方面的，可以是信息共享、软件共享，也可以是硬件共享。例如，互联网上有许多服务器（就是一种专用的计算机）存储了大量有价值的电子文档（包括音频和视频文件），可供上网的用户很方便地读取或下载（无偿或有偿）。由于网络的存在，这些资源**好像**就在用户身边一样，使用非常方便。

现在人们的生活、工作、学习和交往都已离不开互联网。设想一下，某一天我们所在城市的互联网突然瘫痪不能工作了，会出现什么结果呢？这时，我们既不能上网查询有关的资料，也无法使用微信或电子邮件与朋友及时交流信息，网上购物也将完全停顿。我们无法购买机票或火车票，因为即使是在售票处工作的售票员，也无法通过互联网得知目前还有多少余票可供出售。我们不能到银行存钱或取钱，无法交纳水电费和煤气费等。股市交易也将停顿。在图书馆我们检索不到所需要的图书和资料。可见，人们的生活越依赖互联网，互联网的可靠性也就越重要。现在互联网已经成为社会最为重要的基础设施之一。

现在常常可以看到一种新的提法，即“互联网+”。它的意思就是“互联网 + 各个传统行业”，因此可以利用信息通信技术和互联网平台来创造新的发展生态。实际上“互联网+”代表一种新的经济形态，其特点就是把互联网的创新成果深度融合于经济社会各领域之中，这就大大地提升了实体经济的创新力和生产力。我们也必须看到互联网的各种应用对各行各业的巨大冲击。例如，电子邮件迫使传统的电报业务退出市场，网络电话的普及使得传统的长途电话（尤其是国际长途电话）的通信量急剧下降，对日用商品快捷方便的网购造成了不少实体商店的停业，网约车的问世对出租车行业产生了巨大的冲击，网上支付的飞速发展使得很多银行的营业厅变得冷冷清清。

互联网也给人们带来了一些负面影响。有人肆意利用互联网传播计算机病毒，破坏互联网上数据的正常传送和交换；有的犯罪分子甚至利用互联网窃取国家机密和盗窃银行或储户的钱财；网上欺诈或在网上肆意散布谣言、不良信息和播放不健康的视频节目也时有发生；有的青少年弃学而沉溺于网吧的网络游戏中；等等。

虽然如此，但互联网的负面影响毕竟还是次要的。随着对互联网管理的加强，我们可以使互联网给社会带来正面积极的作用成为互联网的主流。

由于互联网已经成为世界上最大的计算机网络，因此下面我们先对互联网进行一下概述，包括互联网的主要构件，这样就可以对计算机网络有一个最初步的了解。

1.2 互联网概述

1.2.1 网络的网络

起源于美国的互联网[①]现已发展成为世界上最大的覆盖全球的计算机网络。

我们先给出关于网络、互连网、互联网（因特网）的一些最基本的概念。

请读者注意：为了方便，在本书中，“网络”往往就是“计算机网络”的简称，而不是表示电信网或有线电视网。

计算机网络（简称为**网络**）由若干**节点**(node)[②]和连接这些节点的**链路**(link)组成。网络中的节点可以是计算机、集线器、交换机或路由器等（在后续的两章我们将会介绍集线器、交换机和路由器等设备的作用）。图 1-1(a)给出了一个具有四个节点和三条链路的网络。我们看

① 注：1994 年全国自然科学名词审定委员会公布的名词中，interconnection 是“互连”，interconnection network 是“互连网络”，internetworking 是“网际互连”。但 1997 年 8 月全国科学技术名词审定委员会在其推荐名（一）中，将 internet, internetwork, interconnection network 的译名均推荐为“互联网”，而在注释中说“又称互连网”，即“互联网”与“互连网”这两个名词均可使用，但请注意，“联”和“连”并非同义字。术语“互连”和“互联”并不等同。“连接”和“联接”也不等同。

② 注：根据《计算机科学技术名词》第 112 页，名词 node 的标准译名是：节点 08.078, 结点 12.023。再查一下 12.023 这一节是**计算机网络**，因此，在计算机网络领域，node 显然应当译为结点，而不是节点。但在网络领域，标准译名“结点”一直未得到推广。考虑到与大多数中文资料中的 node 译名保持一致，本书从第 4 版起也采用“节点”这个非标准译名。

到，有三台计算机通过三条链路连接到一个集线器上。这是一个非常简单的计算机网络（可简称为网络）。又如，在图 1-1(b)中，有多个网络通过一些路由器相互连接起来，构成了一个覆盖范围更大的计算机网络。这样的网络称为**互连网**(internetwork 或 internet)。因此互连网是**“网络的网络”**(network of networks)。用一朵云表示一个网络的好处，就是可以先不考虑每一个网络中的细节，而是集中精力讨论与这个互连网有关的一些问题。

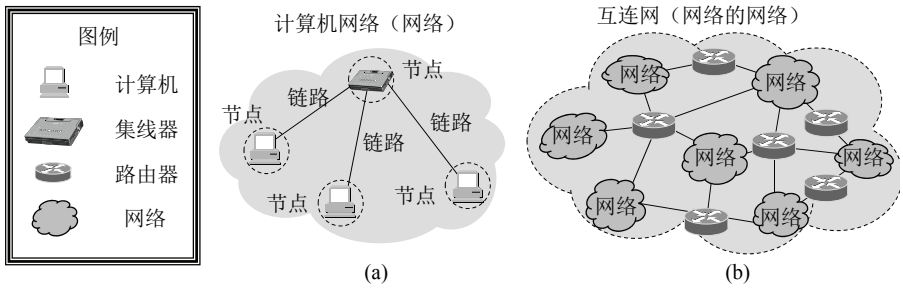


图 1-1　简单的网络(a)和由网络构成的互连网(b)

请读者注意，当我们使用一朵云来表示网络时，可能会有两种不同的情况。一种情况如图 1-1(a)所示，用云表示的网络已经包含了网络中的计算机。但有时为了讨论问题的方便（例如，要讨论几个计算机之间如何进行通信），也可以把有关的计算机画在云的外面，如图 1-2 所示。习惯上，与网络相连的计算机常称为主机(host)。在互连网中不可缺少的路由器，是一种特殊的计算机（有中央处理器、存储器、操作系统等），但不能称为主机。

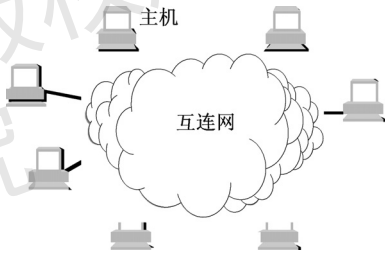


图 1-2　互连网与所连接的主机

这样，我们初步建立了下面的基本概念：

网络把许多计算机连接在一起，而互连网则把许多网络通过一些路由器连接在一起。与网络相连的计算机常称为主机。

还有一点也必须注意：网络互连并不仅仅是把计算机简单地在物理上连接起来，因为这样做并不能达到计算机之间能够相互交换信息的目的。我们还必须在计算机上安装许多使计算机能够交换信息的软件才行。因此当我们谈到网络互连时，就隐含地表示在这些计算机上已经安装了可正常工作的适当软件，在计算机之间可以通过网络交换信息。

现在使用智能手机上网已非常普遍。由于智能手机包含中央处理器、存储器以及操作系统，因此，从计算机网络的角度看，连接在计算机网络上的智能手机也相当于一个主机。实际上，智能手机已远远不是个单一功能的设备，它既是电话机，同时也是计算机、照相机、摄像机、电视机、导航仪等综合多种功能于一体的智能机器。同理，连接在计算机网络上的智能电视机，也是计算机网络上的主机。

1.2.2　互联网基础结构发展的三个阶段

互联网的基础结构大体上经历了三个阶段的演进。但这三个阶段在时间划分上并非截然分开而是有部分重叠的，这是因为**网络的演进是逐步的**，而并非在某个日期发生了突变。

第一阶段是从单个网络 ARPANET 向互连网发展的过程。1969 年美国国防部创建的第一个分组交换网 ARPANET 最初只是一个单个的分组交换网。所有要连接在 ARPANET 上的主机都直接与就近的节点交换机相连。但到了 20 世纪 70 年代中期，人们已认识到不可能仅使用一个单独的网络来满足所有的通信需求。于是 ARPA 开始研究多种网络（如分组无线电网络）

互连的技术，这就导致了互连网络的出现，成为现今**互联网**(Internet)的雏形。1983 年 TCP/IP 协议成为 ARPANET 上的标准协议，使得所有使用 TCP/IP 协议的计算机都能利用互连网相互通信，因而人们就把 1983 年作为互联网的诞生时间。1990 年 ARPANET 正式宣布关闭，因为它的实验任务已经完成。

请读者注意以下两个意思相差很大的英文名词 internet 和 Internet：

以小写字母 i 开始的 **internet（互连网）是一个通用名词，它泛指由多个计算机网络互连而成的计算机网络**。在这些网络之间的通信协议（即通信规则）可以任意选择，不一定非要使用 TCP/IP 协议。

以大写字母 I 开始的 **Internet（互联网，或因特网）则是一个专用名词，它指当前全球最大的、开放的、由众多网络相互连接而成的特定互连网，它采用 TCP/IP 协议族作为通信的规则，且其前身是美国的 ARPANET**。

可见，任意把几个计算机网络互连起来（不管采用什么协议），并能够相互通信，这样构成的网络称为互连网(internet)，而不是互联网(Internet)。

第二阶段的特点是建成了**三级结构的互联网**。从 1985 年起，美国国家科学基金会 NSF (National Science Foundation)就围绕六个大型计算机中心建设计算机网络，即国家科学基金网 NSFNET。它是一个三级计算机网络，分为**主干网**、**地区网**和**校园网**（或**企业网**）。这种三级计算机网络成为互联网中的主要组成部分。1991 年，NSF 和美国的其他政府机构开始认识到，互联网必将扩大其使用范围，不应仅限于大学和研究机构。世界上的许多公司纷纷接入到互联网，网络上的通信量急剧增大，使互联网的容量已满足不了需要。于是美国政府决定把互联网的主干网转交给私人公司来经营，并开始对接入互联网的单位收费。

第三阶段的特点是逐渐形成了**全球范围的多层次 ISP 结构的互联网**。从 1993 年开始，由美国政府资助的 NSFNET 逐渐被若干个商用的**互联网主干网**替代。这样就出现了一个新的名词：**互联网服务提供者** ISP (Internet Service Provider)。在许多情况下，互联网服务提供者 ISP 就是一个进行商业活动的公司，因此 ISP 又常译为**互联网服务提供商**。例如，中国电信、中国联通和中国移动等公司都是我国最有名的 ISP。

互联网服务提供者 ISP 可以从互联网管理机构申请到很多 IP 地址（互联网上的主机都必须有 IP 地址才能上网），同时拥有通信线路（大 ISP 自己建造通信线路，小 ISP 则向电信公司租用通信线路）以及路由器等连网设备，因此任何机构和个人只要向某个 ISP 交纳规定的费用，就可从该 ISP 获取所需 IP 地址的租用权，并可通过该 ISP 接入互联网。所谓“上网”就是指“（通过某 ISP 获得的 IP 地址）接入互联网”。IP 地址的管理机构不会把单个的 IP 地址零星地分配给单个用户，而是把整块的 IP 地址有偿租赁给经审查合格的 ISP。由此可见，现在的互联网已不是某个单个组织所拥有而是全世界无数大大小小的 ISP 所共同拥有的，这就是互联网也称为“**网络的网络**”的原因。

根据提供服务的覆盖面积大小以及所拥有的 IP 地址数目的不同，ISP 也分为不同层次的 ISP：主干 ISP、地区 ISP 和本地 ISP。目前已经覆盖全球的互联网，其主干 ISP 只有十几个，但本地 ISP 有好几十万个。

主干 ISP 由几个专门的公司创建和维护，服务面积最大（一般都能够覆盖国家范围），并且还拥有高速主干网（例如 10 Gbit/s 或更高）。不同的网络运营商都有自己的主干 ISP 网络，并且可以彼此互通。

地区 ISP 是一些较小的 ISP。这些地区 ISP 通过一个或多个主干 ISP 连接起来。它们位于等级中的第二层，数据率也低一些。

本地 ISP 给用户提供直接的服务（这些用户有时也称为**端用户**，强调是末端的用户）。本

地 ISP 可以连接到地区 ISP，也可直接连接到主干 ISP。绝大多数的用户都是连接到本地 ISP 的。本地 ISP 可以是一个仅仅提供互联网服务的公司，也可以是一个拥有网络并向自己的雇员提供服务的企业，或者是一个运行自己的网络的非营利机构（如学院或大学）。

图 1-3 是具有三层 ISP 结构的互联网的概念示意图，但这种示意图并不表示各 ISP 的地理位置关系。图中给出了主机 A 经过许多不同层次的 ISP 与主机 B 通信的示意图。

随着互联网上数据流量的急剧增长，为了更快地转发分组和更加有效和更加经济地利用网络资源，**互联网交换点** IXP (Internet eXchange Point)就应运而生了。

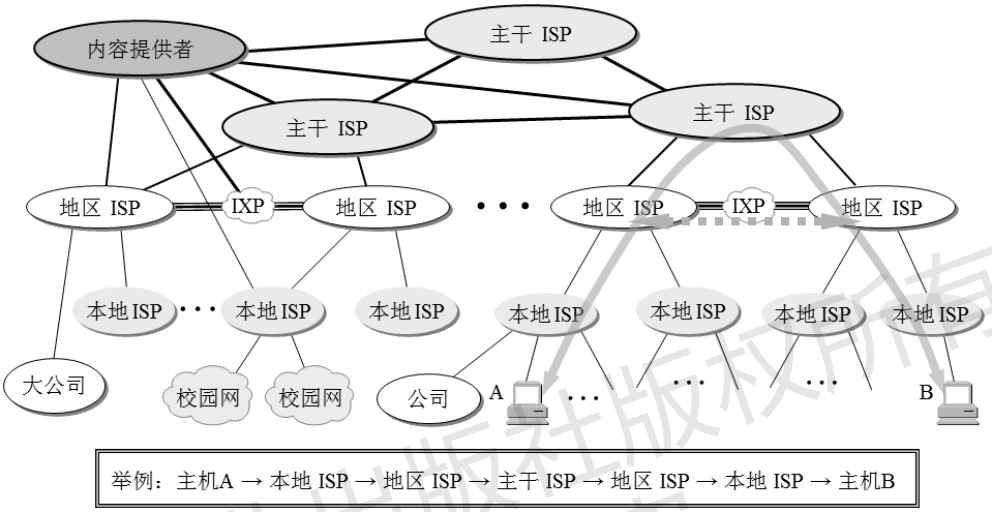


图 1-3　具有三层 ISP 结构的互联网的概念示意图

互联网交换点 IXP 的主要作用就是允许两个网络直接相连并交换分组，而不需要再通过第三个网络来转发分组。例如，在图 1-3 中右方的两个地区 ISP 通过一个 IXP 连接起来了。这样，主机 A 和主机 B 交换分组时，就不必再经过最上层的主干 ISP，而是直接在两个地区 ISP 之间用高速链路对等地交换分组。这样就使互联网上的数据流量分布更加合理，同时也减少了分组转发的迟延时间，降低了分组转发的费用。现在许多 IXP 在进行对等交换分组时，都互相不收费。但本地 ISP 或地区 ISP 通过 IXP 向高层的 IXP 转发分组时，则需要交纳一定的费用。IXP 的结构非常复杂。典型的 IXP 由一个或多个网络交换机组成，许多 ISP 再连接到这些网络交换机的相关端口上。IXP 常采用工作在数据链路层的网络交换机，这些网络交换机都用局域网互连起来。

这里特别要指出的是，当前互联网上最主要的流量就是视频文件的传送。图 1-3 中左上角所示的**内容提供者**(content provider)是在互联网上向所有用户提供视频文件的公司。这种公司和前面提到的 ISP 不同，因为他们并不向用户提供互联网的转接服务，而是提供视频内容的服务。由于传送视频文件产生的流量非常大，为了提高数据传送的效率，这些公司都有独立于互联网的专门网络（仅承载出入该公司的服务器的流量），并且能够和各级 ISP 以及 IXP 相连。这就使得互联网上的所有用户能够更加方便地观看网上的各种视频节目。现在许多 ISP 已不仅向用户提供互联网的接入服务，而且还提供信息服务和一些增值服务。

互联网已经成为世界上规模最大和增长速度最快的计算机网络，没有人能够准确说出互联网究竟有多大。互联网的迅猛发展始于 20 世纪 90 年代。由欧洲原子核研究组织 CERN 开发的**万维网 WWW** (World Wide Web)被广泛使用在互联网上，大大方便了广大非网络专业人员对网络的使用，成为互联网的这种指数级增长的主要驱动力。万维网的站点数目也急剧增长。互联网上准确的通信量是很难估计的，但有文献介绍，互联网上的数据通信量每月约增加

10%。在 2005 年互联网的用户数超过了 10 亿，在 2010 年超过了 20 亿，在 2014 年已接近 30 亿，截止到 2021 年 1 月底，互联网的用户数已达到 46.6 亿。

1.2.3 互联网的标准化工作

互联网的标准化工作对互联网的发展起到了非常重要的作用。我们知道，标准化工作的好坏对一种技术的发展有着很大的影响。缺乏国际标准将会使技术的发展处于比较混乱的状态，而盲目自由竞争的结果很可能形成多种技术体制并存且互不兼容的状态（如过去形成的彩电三大制式），给用户带来较大的不方便。但国际标准的制定又是一个非常复杂的问题，包括不同厂商之间经济利益的争夺等。标准制定的时机也很重要。标准制定得过早，由于技术还没有发展到成熟水平，会使技术比较陈旧的标准限制了产品的技术水平。其结果是以后不得不再次修订标准，造成浪费。反之，若标准制定得太迟，也会使技术的发展无章可循，造成产品的互不兼容，也不利于产品的推广。

1992 年由于互联网不再归美国政府管辖，因此成立了一个国际性组织叫作**互联网协会** (Internet Society，简称为 ISOC)，以便对互联网进行全面管理以及在世界范围内促进其发展和使用。ISOC 下面有一个技术组织叫作**互联网体系结构委员会 IAB** (Internet Architecture Board)[①]，负责管理互联网有关协议的开发。IAB 下面又设有两个工程部：

(1) **互联网工程部 IETF** (Internet Engineering Task Force)

IETF 是由许多**工作组** WG (Working Group)组成的论坛 (forum)，具体工作由**互联网工程指导小组** IESG (Internet Engineering Steering Group)管理。这些工作组划分为若干个领域(area)，每个领域集中研究某一特定的短期和中期的工程问题，主要针对协议的开发和标准化。

(2) **互联网研究部 IRTF** (Internet Research Task Force)

IRTF 是由一些**研究组** RG (Research Group)组成的论坛，具体工作由**互联网研究指导小组** IRSG (Internet Research Steering Group)管理。IRTF 的任务是研究一些需要长期考虑的问题，包括互联网的一些协议、应用、体系结构等。

互联网在制定其标准上很有特色，其中的一个很大的特点是面向公众。所有的互联网标准都是以 RFC 的形式在互联网上发表的。RFC (Request For Comments)的意思就是“请求评论”。所有的 RFC 文档都可从互联网上免费下载，而且任何人都可以用电子邮件随时发表对某个文档的意见或建议。这种开放方式对互联网的迅速发展影响很大。但应注意，并非所有的 RFC 文档都是互联网标准。互联网标准的制定往往要花费漫长的时间，并且是一件非常慎重的工作。只有很少部分的 RFC 文档最后才能变成互联网标准。RFC 文档按发表时间的先后编上序号（即 RFC xxxx，这里的 xxxx 是阿拉伯数字）。一个 RFC 文档更新后就使用一个新的编号，并在文档中指出原来老编号的 RFC 文档已成为陈旧的或被更新，但陈旧的 RFC 文档并不会被删除，而是永远保留着，供用户参考。

制定互联网的正式标准要经过以下三个阶段：

(1) **互联网草案**(Internet Draft)——互联网草案的有效期只有 6 个月。在这个阶段还不能算是 RFC 文档。

(2) **建议标准**(Proposed Standard)——从这个阶段开始就成为 RFC 文档。

(3) **互联网标准**(Internet Standard)——如果经过长期的检验，证明了某个建议标准可以成为互联网标准时，就给它分配一个标准编号，记为 STDxx，这里 STD 是“Standard”的英文

① 注：最初的 IAB 中的 A 曾经代表 Activities（活动）。在一些旧的 RFC 中使用的是这个旧名词。

缩写，而“xx”是标准的编号（有时也写成 4 位数编号，如 STD0005）。一个互联网标准可以和多个 RFC 文档关联。

原先制定互联网标准的过程是：“建议标准”→“草案标准”→“互联网标准”。现在制定互联网标准的过程简化为：“建议标准”→“互联网标准”。

除了建议标准和互联网标准这两种 RFC 文档，还有三种 RFC 文档，即历史的、实验的和提供信息的 RFC 文档。历史的 RFC 文档或者是被后来的规约所取代，或者是从未达到必要的成熟等级因而始终未变成为互联网标准。实验的 RFC 文档表示其工作处于正在实验的情况，而不能够在任何实用的互联网服务中进行实现。提供信息的 RFC 文档包括与互联网有关的一般的、历史的或指导的信息。

1.3 互联网的组成

互联网的拓扑结构虽然非常复杂，并且在地理上覆盖了全球，但从其工作方式上看，可以划分为以下两大块：

(1) **边缘部分** 由所有连接在互联网上的主机组成。这部分是**用户直接使用的**，用来进行通信（传送数据、音频或视频）和资源共享。

(2) **核心部分** 由大量网络和连接这些网络的路由器组成。这部分是**为边缘部分提供服务的**（提供连通性和交换）。

图 1-4 给出了这两部分的示意图。下面分别讨论这两部分的作用和工作方式。

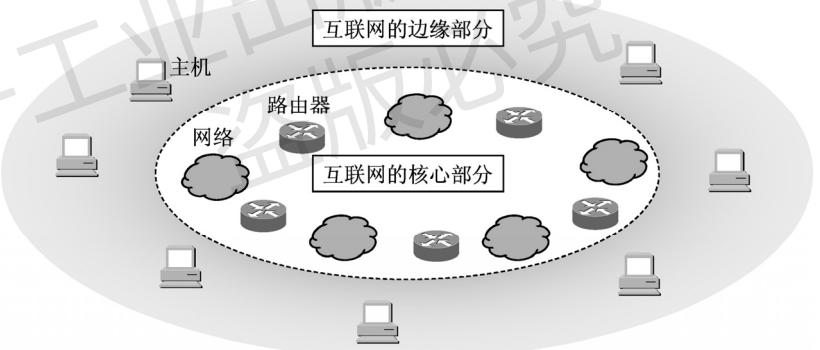


图 1-4 互联网的边缘部分与核心部分

1.3.1 互联网的边缘部分

处在互联网边缘的部分就是连接在互联网上的所有的主机。这些主机又称为**端系统**(end system)，“端”就是“末端”的意思（即互联网的末端）。端系统在功能上可能有很大的差别，小的端系统可以是一台普通个人电脑（包括笔记本电脑或平板电脑）和具有上网功能的智能手机，甚至是一个很小的网络摄像头（可监视当地的天气或交通情况，并在互联网上实时发布），而大的端系统则可以是一台非常昂贵的大型计算机（这样的计算机通常称为**服务器** server）。端系统的拥有者可以是个人，也可以是单位（如学校、企业、政府机关等），当然也可以是某个 ISP（即 ISP 不仅仅是向端系统提供服务，它也可以拥有一些端系统）。边缘部分利用核心部分所提供的服务，使众多主机之间能够互相通信并交换或共享信息。值得注意的是，现今大部分能够向网民提供信息检索、万维网浏览以及视频播放等功能的服务器，都不再

是一个孤立的服务器，而是属于某个大型数据中心。例如，谷歌公司(Google)拥有上百个数据中心，而其中的 15 个大型数据中心的每一个都拥有 10 万台以上的服务器。又如，我国的百度公司在山西阳泉建造的数据中心拥有 16 万台服务器。

我们先要明确下面的概念。我们说：“主机 A 和主机 B 进行通信”，实际上是指：“运行在主机 A 上的某个程序和运行在主机 B 上的另一个程序进行通信”。由于“进程”就是“运行着的程序”，因此这也就是指：**“主机 A 的某个进程和主机 B 上的另一个进程进行通信”**。这种比较严密的说法通常可以**简称为“计算机之间通信”**。

在网络边缘的端系统之间的通信方式通常可划分为两大类：客户–服务器方式（C/S 方式）和对等方式（P2P 方式）[①]。下面分别对这两种方式进行介绍。

1. 客户–服务器方式

这种方式在互联网上是最常用的，也是传统的方式。我们在上网发送电子邮件或在网站上查找资料时，都使用客户–服务器方式（有时写为客户/服务器方式）。

当我们打电话时，电话机的振铃声使被叫用户知道现在有一个电话呼叫。计算机通信的对象是应用层中的应用进程，显然不能用响铃的办法来通知所要找的对方的应用进程。然而采用客户-服务器方式可以使两个应用进程能够进行通信。

客户(client)和**服务器**(server)都是指通信中所涉及的两个应用进程。客户-服务器方式所描述的是进程之间服务和被服务的关系。在图 1-5 中，主机 A 运行客户程序而主机 B 运行服务器程序。在这种情况下，A 是客户而 B 是服务器。客户 A 向服务器 B 发出请求服务，而服务器 B 向客户 A 提供服务。这里最主要的特征就是：

客户是服务请求方，服务器是服务提供方。

服务请求方和服务提供方都要使用网络核心部分所提供的服务。

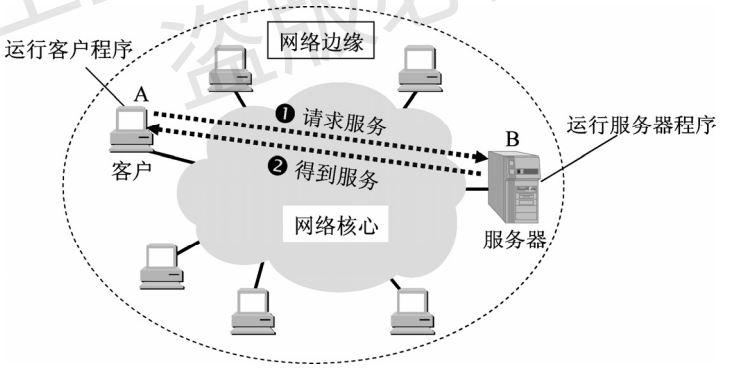


图 1-5　客户-服务器工作方式

在实际应用中，客户程序和服务器程序通常还具有以下一些主要特点。

客户程序：

(1) 被用户调用后运行，在通信时主动向远地服务器发起通信（请求服务）。因此，客户程序必须知道服务器程序的地址。

(2) 不需要特殊的硬件和很复杂的操作系统。

服务器程序：

(1) 是一种专门用来提供某种服务的程序，**可同时处理**多个远地或本地客户的请求。

① 注：C/S 方式表示 Client/Server 方式，P2P 方式表示 Peer-to-Peer 方式。有时还可看到另外一种叫作浏览器–服务器方式，即 B/S 方式（Browser/Server 方式），但这仍然是 C/S 方式的一种特例。

(2) 系统启动后即一直不断地运行着，被动地等待并接受来自各地的客户的通信请求。因此，服务器程序不需要知道客户程序的地址。

(3) 一般需要有强大的硬件和高级的操作系统支持。

客户与服务器的通信关系建立后，通信可以是双向的，客户和服务器都可发送和接收数据。

顺便要说一下，上面所说的**客户和服务器本来都指的是计算机进程（软件）**。使用计算机的人是计算机的“用户”(user)而不是“客户”(client)。但在许多国外文献中，经常也把运行客户程序的**机器**称为 client（在这种情况下也可把 client 译为“客户机”），把运行服务器程序的**机器**也称为 server。因此我们应当根据上下文来判断 client 或 server 是指软件还是硬件。在本书中，在表示机器时，我们也使用“客户端”（或“客户机”）或“服务器端”（或服务器）来表示“运行客户程序的机器”或“运行服务器程序的机器”。

2. 对等连接方式

对等连接（peer-to-peer，简写为 P2P。这里使用数字 2 是因为英文的 2 是 two，其读音与 to 同，因此英文的 to 常写为数字 2）是指两台主机在通信时，并不区分哪一个是服务请求方和哪一个是服务提供方。只要两台主机都运行了对等连接软件（P2P 软件），它们就可以进行平等的对等连接通信。这时，双方都可以下载对方已经存储在硬盘中的共享文档。因此这种工作方式也称为 **P2P 方式**。在图 1-6 中，主机 C, D, E 和 F 都运行了 P2P 程序，因此这几台主机都可进行对等通信（如 C 和 D，E 和 F，以及 C 和 F）。实际上，对等连接方式从本质上看仍然使用客户-服务器方式，只是对等连接中的每一台主机既是客户同时又是服务器。例如主机 C，当 C 请求 D 的服务时，C 是客户，D 是服务器。但如果 C 又同时向 F 提供服务，那么 C 又同时起着服务器的作用。

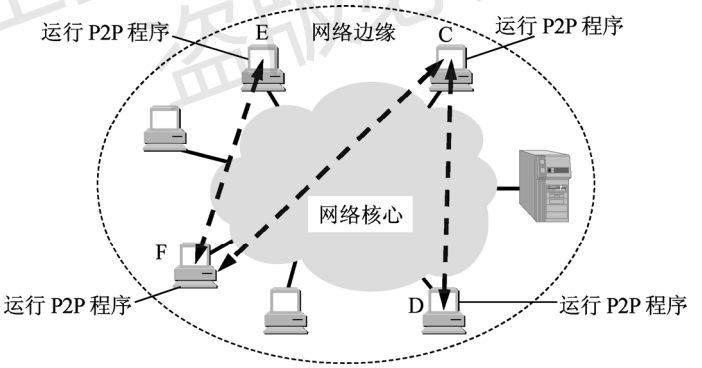


图 1-6　对等连接工作方式（P2P 方式）

对等连接工作方式可支持大量对等用户（如上百万个）同时工作。关于这种工作方式我们将在后面第 6 章的 6.6 节进一步讨论。

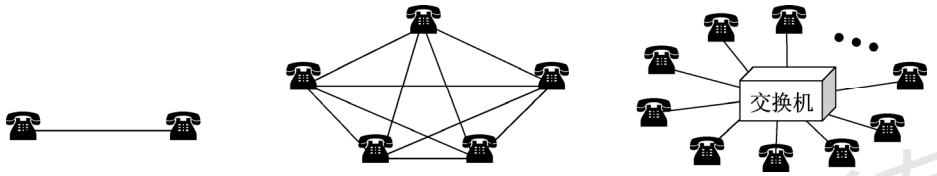
1.3.2　互联网的核心部分

网络核心部分是互联网中最复杂的部分，因为网络中的核心部分要向网络边缘部分中的大量主机提供连通性，使边缘部分中的任何一台主机都能够与其他主机通信。

在网络核心部分起特殊作用的是**路由器**(router)，它是一种专用计算机（但不叫作主机）。路由器是实现**分组交换**(packet switching)的关键构件，其任务是**转发收到的分组**，这是网络核心部分最重要的功能。为了弄清分组交换，下面先介绍电路交换的基本概念。

1. 电路交换的主要特点

在电话问世后不久，人们就发现，要让所有的电话机都两两相连接是不现实的。图 1-7(a)表示两部电话只需要用一对电线就能够互相连接起来。但若有 5 部电话要两两相连，则需要 10 对电线，如图 1-7(b)所示。显然，若 N 部电话要两两相连，就需要 $N(N-1)/2$ 对电线。当电话机的数量很大时，这种连接方法需要的电线数量就太大了（与电话机的数量的平方成正比）。于是人们认识到，要使得每一部电话能够很方便地和另一部电话进行通信，就应当使用电话交换机将这些电话连接起来，如图 1-7(c)所示。每一部电话都连接到交换机上，而交换机使用交换的方法，让电话用户彼此之间可以很方便地通信。电话发明后的一百多年来，电话交换机虽然经过多次更新换代，但交换的方式一直都是电路交换(circuit switching)。



(a) 两部电话直接相连　(b) 5 部电话两两直接相连　(c) 用交换机连接许多部电话

图 1-7　电话机的不同连接方法

当电话机的数量增多时，就要使用很多彼此连接起来的交换机来完成全网的交换任务。用这样的方法，就构成了覆盖全世界的电信网。

从通信资源的分配角度来看，**交换**(switching)就是按照某种方式动态分配传输线路的资源。在使用电路交换打电话之前，必须先拨号请求建立连接。当被叫用户听到交换机送来的振铃音并摘机后，从主叫端到被叫端就建立了一条连接，即一条**专用的物理通路**。这条连接保证了双方通话时所需的通信资源，而这些资源在双方通信时不会被其他用户占用。此后主叫和被叫双方就能互相通电话。通话完毕挂机后，交换机释放刚才使用的这条专用的物理通路（即把刚才占用的所有通信资源归还给电信网）。这种必须经过“**建立连接**（占用通信资源）→**通话**（一直占用通信资源）→**释放连接**（归还通信资源）”三个步骤的交换方式称为**电路交换**[①]。如果用户在拨号呼叫时电信网的资源已不足以支持这次的呼叫，则主叫用户会听到忙音，表示电信网不接受用户的呼叫，用户必须挂机，等待一段时间后再重新拨号。

图 1-8 为电路交换的示意图。为简单起见，图中没有区分市话交换机和长途电话交换机。应当注意的是，用户线是电话用户到所连接的市话交换机的连接线路，是用户独占的传送模拟信号的专用线路，而交换机之间拥有大量话路的中继线（这些传输线路早已都数字化了）则是许多用户共享的，正在通话的用户只占用了中继线里面的一个话路。电路交换的一个重要特点就是**在通话的全部时间内，通话的两个用户始终占用端到端的通信资源**。

当使用电路交换来传送计算机数据时，其**线路的传输效率往往很低**。这是因为计算机数据是突发式地出现在传输线路上的，因此线路上真正用来传送数据的时间往往不到 10%甚至 1%。已被用户占用的通信线路资源在绝大部分时间里都是空闲的。例如，当用户阅读终端屏幕上的信息或用键盘输入和编辑一份文件时，或计算机正在进行处理而结果尚未返回时，宝贵的通信线路资源并未被利用而是被白白浪费了。

① 注：电路交换最初指的是连接电话机的双绞线对在交换机上进行的交换（交换机有人工的、步进的和程控的，等等）。后来随着技术的进步，采用了多路复用技术，出现了频分多路、时分多路、码分多路等，这时电路交换的概念就扩展到在双绞线、铜缆、光纤、无线媒体中多路信号中的某一路（某个频率、某个时隙、某个码序等）和另一路的交换。

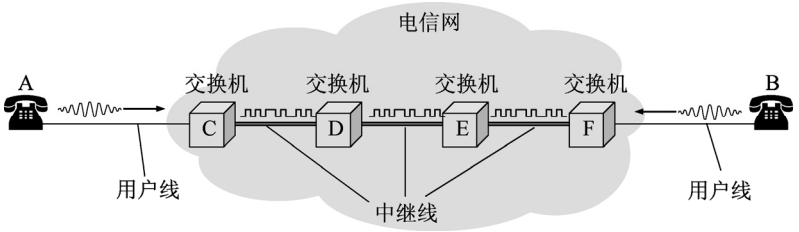


图 1-8　电路交换的用户始终占用端到端的通信资源

2. 分组交换的主要特点

分组交换则采用**存储转发**技术[①]。图 1-9 表示把一个报文划分为几个分组后再进行传送。通常我们把要发送的整块数据称为一个**报文**(message)。在发送报文之前，先把较长的报文划分为一个个更小的等长数据段，例如，每个数据段为 1024 bit[②]。在每一个数据段前面，加上一些必要的控制信息组成的**首部**(header)后，就构成了一个**分组**(packet)。分组又称为“**包**”，而分组的首部也可称为“**包头**”。分组是在互联网中传送的数据单元。分组中的“首部”是非常重要的，正是由于分组的首部包含了诸如目的地址和源地址等重要控制信息，每一个分组才能在互联网中独立地选择传输路径，并被正确地交付到分组传输的终点。

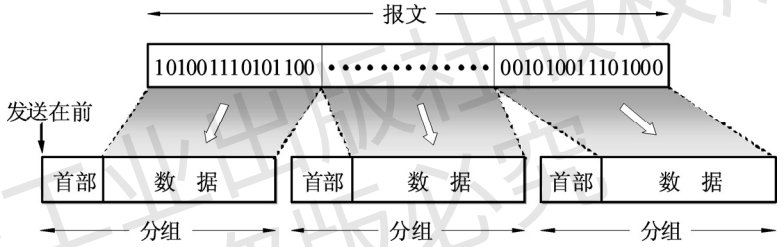


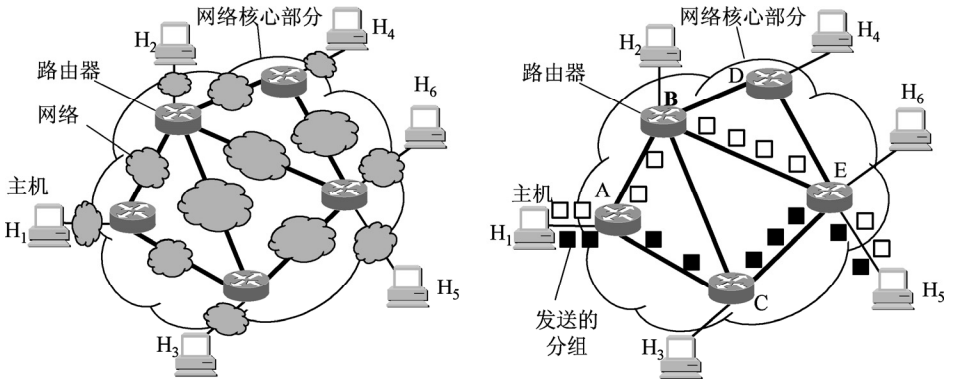
图 1-9　以分组为基本单位在网络中传送

图 1-10(a)强调互联网的核心部分是由许多网络和把它们互连起来的路由器组成的，而主机处在互联网的边缘部分。在互联网核心部分的路由器之间一般都用高速链路相连接，而在网络边缘部分的主机接入到核心部分则通常以相对较低速率的链路相连接。

位于网络边缘部分的主机和位于网络核心部分的路由器都是计算机，但它们的作用却很不一样。**主机是为用户进行信息处理的**，并且可以和其他主机通过网络交换信息。**路由器则用来转发分组，即进行分组交换**。路由器收到一个分组，先暂时存储一下，检查其首部，查找转发表，按照首部中的目的地址，找到合适的接口转发出去，把分组交给下一个路由器。这样一步一步地（有时会经过几十个不同的路由器）以存储转发的方式，把分组交付最终的目的主机。各路由器之间必须经常交换彼此掌握的路由信息，以便创建和动态维护路由器中的转发表，使得转发表能够在整个网络拓扑发生变化时及时更新。

① 注：存储转发的概念最初是于 1964 年 8 月由巴兰(Baran)在美国兰德(Rand)公司的“论分布式通信”的研究报告中提出的。在 1962—1965 年间，美国国防部远景研究规划局 DARPA 和英国国家物理实验室 NPL 都在对新型的计算机通信网进行研究。1966 年 6 月，NPL 的戴维斯(Davies)首次提出“分组”(packet)这一名词。1969 年 12 月，美国的分组交换网 ARPANET（当时仅有 4 个节点）投入运行。从此，计算机网络的发展就进入了一个崭新的纪元。1973 年英国国家物理实验室 NPL 也开通了分组交换试验网。现在大家都公认 ARPANET 为分组交换网之父。除英美两国外，法国也在 1973 年开通其分组交换网 CYCLADES。

② 注：在本书中，bit 表示“比特”。在计算机领域中，bit 常译为“位”。在许多情况下，“比特”和“位”可以通用。在使用“位”作为单位时，请根据上下文特别注意是二进制的“位”还是十进制的“位”。请注意，bit 在表示信息量（比特）或信息传输速率（比特/秒）时不能译为“位”。



(a) 核心部分中的路由器把许多网络互连起来　　(b) 核心部分中的网络可用一条链路表示

图 1-10　分组交换的示意图

当我们讨论互联网的核心部分中的路由器转发分组的过程时，往往把单个的网络简化成一条**链路**，而路由器成为核心部分的**节点**，如图 1-10(b)所示。这种简化图看起来可以更加突出重点，因为在转发分组时最重要的就是要知道路由器之间是怎样连接起来的。

现在假定图 1-10(b)中的主机 H_1 向主机 H_5 发送数据。主机 H_1 先将分组逐个地发往与它直接相连的路由器 A。此时，除链路 H_1–A 外，其他通信链路并不被目前通信的双方所占用。需要注意的是，即使是链路 H_1–A，也只是当分组正在此链路上传送时才被占用。在各分组传送之间的空闲时间，链路 H_1–A 仍可为其他主机发送的分组使用。

路由器 A 把主机 H_1 发来的分组放入缓存。假定从路由器 A 的转发表中查出应把该分组转发到链路 A–C。于是分组就传送到路由器 C。当分组正在链路 A–C 传送时，该分组并不占用网络其他部分的资源。

路由器 C 继续按上述方式查找转发表，假定查出应转发到路由器 E。当分组到达路由器 E 后，路由器 E 就最后把分组直接交给主机 H_5。

假定在某一个分组的传送过程中，链路 A–C 的通信量太大，那么路由器 A 可以把分组沿另一个路由传送，即先转发到路由器 B，再转发到路由器 E，最后把分组送到主机 H_5。在网络中可同时有多台主机进行通信，如主机 H_2 也可以经过路由器 B 和 E 与主机 H_6 通信。

这里要注意，路由器暂时存储的是一个个短分组，而不是整个的长报文。短分组是暂存在路由器的存储器（即内存）中而不是存储在磁盘中的。这就保证了较高的交换速率。

在图中只画了一对主机 H_1 和 H_5 在进行通信。实际上，互联网可以容许非常多的主机同时进行通信，而一台主机中的多个进程（即正在运行中的多道程序）也可以各自和不同主机中的不同进程进行通信。

应当注意，分组交换在传送数据之前不必先占用一条端到端的通信资源。分组在哪段链路上传送才占用那段链路的通信资源。分组到达一个路由器后，先暂时存储下来，查找转发表，然后从另一条合适的链路转发出去。分组在传输时就这样逐段地断续占用通信资源，而且还省去了建立连接和释放连接的开销，因而数据的传输效率更高。

互联网采取了专门的措施，保证了数据的传送具有非常高的可靠性（在第 5 章 5.4 节介绍运输层协议时要着重讨论这个问题）。当网络中的某些节点或链路突然出故障时，在各路由器中运行的路由选择**协议**(protocol)能够自动找到转发分组最合适的路径。这些将在第 4 章 4.5 节中详细讨论。

从以上所述可知，采用存储转发的分组交换，实质上是采用了在数据通信的过程中断续（或动态）分配传输带宽的策略（关于带宽的进一步讨论见后面的 1.6 节）。这对传送突发式的

计算机数据非常合适，使得通信线路的利用率大大提高了。

为了提高分组交换网的可靠性，互联网的核心部分常采用网状拓扑结构，使得当发生网络拥塞或少数节点、链路出现故障时，路由器可灵活地改变转发路由而不致引起通信的中断或全网的瘫痪。此外，通信网络的主干线路往往由一些高速链路构成，这样就可以较高的数据率迅速地传送计算机数据。

综上所述，分组交换的主要优点可归纳如表 1-1 所示。

表 1-1　分组交换的优点

优　　点	所采用的手段
高效	在分组传输的过程中动态分配传输带宽，对通信链路逐段占用
灵活	为每一个分组独立地选择最合适的转发路由
迅速	以分组作为传送单位，不先建立连接就能向其他主机发送分组
可靠	保证可靠性的网络协议；分布式多路由的分组交换网，使网络有很好的生存性

分组交换也带来一些新的问题。例如，分组在各路由器存储转发时需要排队，这就会造成一定的**时延**。因此，必须尽量设法减少这种时延。此外，由于分组交换不像电路交换那样通过建立连接来保证通信时所需的各种资源，因而无法确保通信时端到端所需的带宽。

分组交换带来的另一个问题是各分组必须携带的控制信息也造成了一定的**开销**(overhead)。整个分组交换网还需要专门的管理和控制机制。

应当指出，从本质上讲，这种断续分配传输带宽的存储转发原理并非是全新的概念。自古代就有的邮政通信，就其本质来说也属于存储转发方式。而在 20 世纪 40 年代，电报通信也采用了基于存储转发原理的**报文交换**(message switching)。在报文交换中心，一份份电报被接收下来，并穿成纸带。操作员以每份报文为单位，撕下纸带，根据报文的目的站地址，拿到相应的发报机转发出去。这种报文交换的时延较长，从几分钟到几小时不等。现在报文交换已不使用了。分组交换虽然也采用存储转发原理，但由于使用了计算机进行处理，因此分组的转发非常迅速。例如，ARPANET 建网初期的经验表明，在正常的网络负荷下，当时横跨美国东西海岸的端到端平均时延小于 0.1 秒。这样，分组交换虽然采用了某些古老的交换原理，但实际上已变成了一种崭新的交换技术。

图 1-11 显示了电路交换、报文交换和分组交换的主要区别。图中的 A 和 D 分别是源点和终点，而 B 和 C 是在 A 和 D 之间的中间节点。图的最下方归纳了三种交换方式在数据传送阶段的主要特点：

电路交换——整个报文的比特流连续地从源点直达终点，好像在一个管道中传送。

报文交换——整个报文先传送到相邻节点，全部存储下来后查找转发表，转发到下一个节点。

分组交换——单个分组（这只是整个报文的一部分）传送到相邻节点，存储下来后查找转发表，转发到下一个节点。

从图 1-11 可看出，若要连续传送大量的数据，且其传送时间远大于连接建立时间，则电路交换的传输速率较快。报文交换和分组交换不需要预先分配传输带宽，在传送突发数据时可提高整个网络的信道[①]利用率。由于一个分组的长度往往远小于整个报文的长度，因此分组交换比报文交换的时延小，同时也具有更好的灵活性。

在过去很长的时期，人们都有这样的概念：电路交换适合于话音通信，而分组交换则适合

① 注：信道(channel)是指以传输媒体为基础的信号通路（包括有线或无线电线路），其作用是传输信号。

于数据通信。然而随着蜂窝移动通信的发展，这种概念已经发生了根本的变化。从第四代蜂窝移动通信网开始，无论是话音通信还是数据通信，都要采用分组交换（见第 9 章 9.3 节有关蜂窝移动通信网的讨论）。

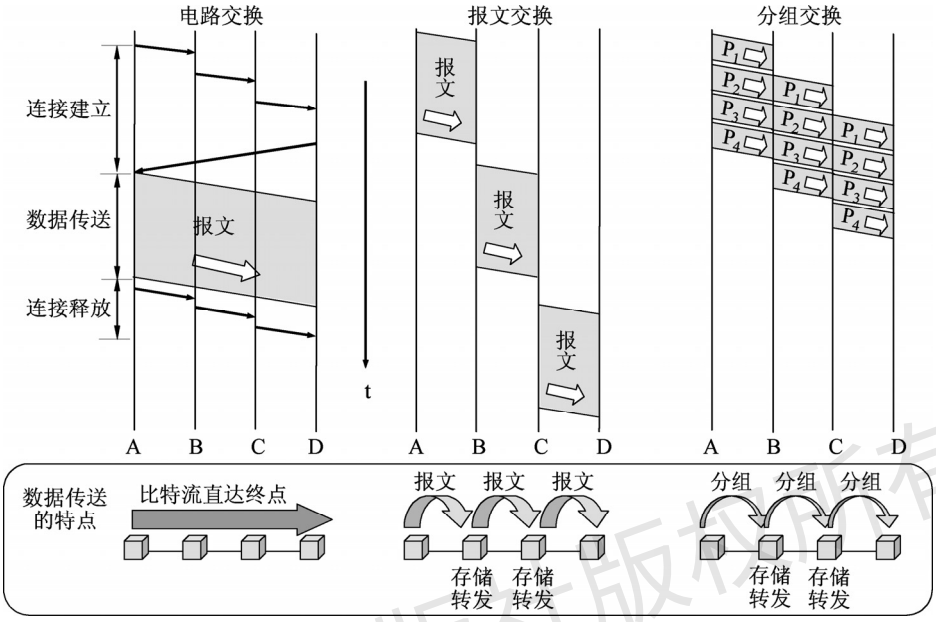


图 1-11　三种交换的比较。电路交换、报文交换、分组交换，P_1～P_4 表示 4 个分组

1.4　计算机网络在我国的发展

下面简单介绍一下计算机网络在我国的发展情况。

最早着手建设专用计算机广域网的是铁道部。铁道部在 1980 年即开始进行计算机联网实验。1989 年 11 月我国第一个公用分组交换网 CNPAC 建成运行。在 20 世纪 80 年代后期，公安、银行、军队以及其他一些部门也相继建立了各自的专用计算机广域网。这对迅速传递重要的数据信息起着重要的作用。另一方面，从 20 世纪 80 年代起，国内的许多单位相继安装了大量的局域网。局域网的价格便宜，其所有权和使用权都属于本单位，因此便于开发、管理和维护。局域网的发展很快，对各行各业的管理现代化和办公自动化起到了积极的作用。

这里应当特别提到的是，1994 年 4 月 20 日我国用 64 bit/s 专线正式连入互联网。从此，我国被国际上正式承认为接入互联网的国家。同年 5 月中国科学院高能物理研究所设立了我国的第一个万维网服务器。同年 9 月中国公用计算机互联网 CHINANET 正式启动。到目前为止，我国陆续建造了基于互联网技术并能够和互联网互连的多个全国范围的公用计算机网络，其中规模最大的就是下面这五个：

(1) 中国电信互联网 CHINANET（也就是原来的中国公用计算机互联网）

(2) 中国联通互联网 UNINET

(3) 中国移动互联网 CMNET

(4) 中国教育和科研计算机网 CERNET

(5) 中国科学技术网 CSTNET

2004 年 2 月，我国的第一个下一代互联网 CNGI 的主干网 CERNET2 试验网正式开通，

并提供服务。试验网以 2.5～10 Gbit/s 的速率连接北京、上海和广州三个 CERNET 核心节点，并与国际下一代互联网相连接。这标志着中国在互联网的发展过程中，已逐渐达到与国际先进水平同步。

中国互联网络信息中心 CNNIC (China Network Information Center)每年两次公布我国互联网的发展情况。读者可在其网站上查到最新的和过去的历史文档。CNNIC 把过去半年内使用过互联网的 6 周岁及以上的中国居民称为**网民**。根据 2021 年 9 月 CNNIC 发表的《中国互联网络发展状况统计报告》，截至 2021 年 6 月，我国网民已达到 10.11 亿，互联网普及率已达到 71.6%。

现在微博和网络视频的用户明显增多。移动互联网营销发展迅速，当前网民最主要的网络应用就是即时通信（例如微信）、搜索引擎（即在互联网上使用搜索引擎来查找所需的信息）、网络音乐、网络新闻和博客等。此外，更多的经济活动已步入了互联网时代。网上购物、网上支付和网上银行的使用率也迅速提升。截至 2021 年 6 月，我国网络支付用户规模达 8.72 亿，占网民整体的 86.3%。

对我国互联网事业发展影响较大的人物和事件不少，限于篇幅，下面仅列举几个例子。

1996 年，张朝阳创立了中国第一家以风险投资资金建立的互联网公司——爱特信公司。两年后，爱特信公司推出“搜狐”产品，并更名为搜狐公司(Sohu)。搜狐公司最主要的产品就是搜狐网站(Sohu.com)，是中国首家大型分类查询搜索引擎。1999 年，搜狐网站增加了新闻及内容频道，成为一个综合门户网站。

1997 年，丁磊创立了网易公司(NetEase)，推出了中国第一个中文全文搜索引擎。网易公司开发的超大容量免费邮箱（如 163 和 126 等），由于具有高效的杀毒和拦截垃圾邮件的功能，安全性很好，已成为国内最受欢迎的中文邮箱。网易网站现在也是全国出名的综合门户网站。

1998 年，王志东创立新浪网站(Sina.com)，该网站现已成为全球最大的中文综合门户网站。新浪的微博是全球使用最多的微博之一。

同年，马化腾、张志东创立了腾讯公司(Tencent)。1999 年腾讯推出了用在个人电脑上的即时通信软件 OICQ，后改名为 QQ。QQ 的功能不断更新，现在已成为一款集话音、短信、文章、音乐、图片和视频于一体的网络沟通交流工具，成为几乎所有网民都在电脑中安装的软件，腾讯也因此成为中国最大的互联网综合服务提供商之一。

2011 年，腾讯推出了专门供智能手机使用的即时通信软件“微信”(国外版的微信叫作 WeChat，在功能上有些差别)。这个软件是在张小龙（著名的电子邮件客户端软件 Foxmail 的作者）领导下成功研发的。微信能够通过互联网快速发送话音短信、视频、图片和文字，并且支持多人视频会议。由于微信能在各种不同操作系统的智能手机中运行，因此目前几乎所有的智能手机用户都在使用微信。微信的功能也在不断更新。装有微信软件的智能手机，已从简单的社交工具演变成一个具有支付能力的全能钱包。几乎所有使用智能手机的人，都离不开微信。

2000 年，李彦宏和徐勇创建了百度网站(Baidu.com)，现在已成为全球最大的中文搜索引擎。自谷歌于 2010 年退出中国后，中国最大的搜索引擎无疑就是百度了。现在，百度网站也可以用主题分类的方法进行查找，非常便于网民对各种信息的浏览。

1999 年，马云创建了阿里巴巴网站(Alibaba.com)，这是一个企业对企业的网上贸易市场平台。2003 年，马云创立了个人网上贸易市场平台——淘宝网(Taobao.com)。2004 年，阿里巴巴集团创立了第三方支付平台——支付宝(Alipay.com)，为中国电子商务提供了简单、安全、快速的在线支付手段。

上述的一些事件对互联网应用在我国的推广普及，起着非常积极的作用。

1.5 计算机网络的类别

1.5.1 计算机网络的定义

计算机网络的精确定义并未统一。

计算机网络较好的定义是这样的：计算机网络主要是由一些通用的、可编程的硬件互连而成的，而这些硬件并非专门用来实现某一特定目的（例如，传送数据或视频信号）。这些可编程的硬件能够用来传送多种不同类型的数据，并能支持广泛的和日益增长的应用。

根据这个定义：(1) 计算机网络所连接的硬件，并不限于一般的计算机，而是包括了智能手机或智能电视机；(2) 计算机网络并非专门用来传送数据，而是能够支持很多种应用（包括今后可能出现的各种应用）。当然，没有数据的传送，这些应用是无法实现的。

请注意，上述的“可编程的硬件”表明这种硬件一定包含有中央处理器 CPU。

我们知道，起初，计算机网络是用来传送数据的。但随着网络技术的发展，计算机网络的应用范围不断增大，不仅能够传送音频和视频文件，而且应用的范围已经远远超过一般通信的范畴。

有时我们也能见到“计算机通信网”这一名词，但这个名词容易使人误认为这是一种专门为了通信而设计的计算机网络。计算机网络显然应具有通信的功能，但这种通信功能并非计算机网络最主要的功能。因此本书不使用“计算机通信网”这一名词。

1.5.2 几种不同类别的计算机网络

计算机网络有多种类别，下面进行简单的介绍。

1. 按照网络的作用范围进行分类

(1) **广域网 WAN** (Wide Area Network) 广域网的作用范围通常为几十到几千公里，因而有时也称为**远程网**(long haul network)。广域网是互联网的核心部分，其任务是长距离（例如，跨越不同的国家）运送主机所发送的数据。连接广域网各节点交换机的链路一般都是高速链路，具有较大的通信容量。本书不专门讨论广域网。

(2) **城域网 MAN** (Metropolitan Area Network) 城域网的作用范围一般是一个城市，可跨越几个街区甚至整个城市，其作用距离约为 5～50 km。城域网可以为一个或几个单位所拥有，也可以是一种公用设施，用来将多个局域网进行互连。目前很多城域网采用的是以太网技术，因此有时也常并入局域网的范围进行讨论。

(3) **局域网 LAN** (Local Area Network) 局域网一般用微型计算机或工作站通过高速通信线路相连（速率通常在 10 Mbit/s 以上），但地理上则局限在较小的范围（如 1 km 左右）。在局域网发展的初期，一个学校或工厂往往只拥有一个局域网，但现在局域网已非常广泛地使用，学校或企业大都拥有许多个互连的局域网（这样的网络常称为**校园网**或**企业网**）。我们将在第 3 章 3.3 至 3.5 节详细讨论局域网。

(4) **个人区域网** PAN (Personal Area Network) 个人区域网就是在个人工作的地方把属于个人使用的电子设备（如便携式电脑等）用无线技术连接起来的网络，因此也常称为**无线个人区域网** WPAN (Wireless PAN)，其范围很小，大约在 10 m 左右。我们将在第 9 章 9.2 节对这种网络进行简单的介绍。

顺便指出，若中央处理机之间的距离非常近（如仅 1 米的数量级或更小些），则一般就称之为**多处理机系统**而不称它为计算机网络。

2. 按照网络的使用者进行分类

(1) **公用网**(public network)　这是指电信公司（国有或私有）出资建造的大型网络。“公用”的意思就是所有愿意按电信公司的规定交纳费用的人都可以使用这种网络。因此公用网也可称为**公众网**。

(2) **专用网**(private network)　这是某个部门为满足本单位的特殊业务工作的需要而建造的网络。这种网络不向本单位以外的人提供服务。例如，军队、铁路、银行、电力等系统均有本系统的专用网。

公用网和专用网都可以传送多种业务。如传送的是计算机数据，则分别是公用计算机网络和专用计算机网络。

3. 用来把用户接入到互联网的网络

这种网络就是**接入网** AN (Access Network)，它又称为**本地接入网**或**居民接入网**。这是一类比较特殊的计算机网络。我们在前面的 1.2.2 节已经介绍了用户必须通过本地 ISP 才能接入到互联网。本地 ISP 可以使用多种接入网技术把用户的端系统连接到互联网。接入网实际上就是本地 ISP 所拥有的网络，它既不是互联网的核心部分，也不是互联网的边缘部分。接入网由某个端系统连接到本地 ISP 的第一个路由器（也称为边缘路由器）之间的一些物理链路所组成。从覆盖的范围看，其长度在几百米到几公里之间。很多接入网还是属于局域网。从作用上看，接入网只是起到让用户能够与互联网连接的“桥梁”作用。在互联网发展初期，用户多用电话线拨号接入互联网，速率很低（每秒几千比特到几十千比特），因此那时并没有使用接入网这个名词。直到最近，由于出现了多种宽带接入技术，宽带接入网才成为互联网领域中的一个热门课题。我们将在第 2 章 2.6 节讨论宽带接入技术。

1.6　计算机网络的性能指标

性能指标从不同的方面来度量计算机网络的性能。下面介绍常用的 7 个性能指标。

1. 速率

我们知道，计算机发送的信号都是数字形式的。**比特**(bit)来源于 binary digit，意思是一个“**二进制数字**”，因此一个比特就是二进制数字中的一个 1 或 0。比特也是信息论中使用的**信息量的单位**。网络技术中的**速率**指的是**数据的传送速率**，它也称为**数据率**(data rate)或**比特率**(bit rate)。速率是计算机网络中最重要的一个性能指标。速率的单位是 bit/s（比特每秒）（或 b/s，有时也写为 bps，即 bit per second）。当数据率较高时，就常常在 bit/s 的前面加上一个字母。例如，k (kilo) = 10^3 = 千，M (Mega) = 10^6 = 兆，G (Giga) = 10^9 = 吉，T (Tera) = 10^{12} = 太，P (Peta) = 10^{15} = 拍，E (Exa) = 10^{18} = 艾，Z (Zetta) = 10^{21} = 泽，Y (Yotta) = 10^{24} = 尧[①]。这样，4 × 10^{10} bit/s 的数据率就记为 40 Gbit/s。现在人们在谈到网络速率时，常省略速率单位

① 注：在计算机领域，数的计算使用二进制。因此习惯上，千 = K = 2^{10} = 1024，兆 = M = 2^{20}，吉 = G = 2^{30}，太 = T = 2^{40}，拍 = P = 2^{50}，艾 = E = 2^{60}，泽 = Z = 2^{70}，尧 = Y = 2^{80}。此外，计算机中的数据量往往用字节 B 作为度量的单位（B 代表 byte）。通常 1 B = 8 bit。例如，15 GB 数据块的大小是 15 × 2^{30} × 8 bit，而不是 15 × 10^9 × 8 bit。但 10 Gbit/s 的速率则表示 10 × 10^9 bit/s。在计算机领域中，所有的这些单位都使用大写字母，但在通信领域中，只有“1000”使用小写“k”，其余的也都用大写。请注意，也有的书不这样严格区分，大写 K 既可表示 1000，又可表示 1024，因此这时要特别小心，不要弄错。

中应有的 bit/s，而使用不太正确的说法，如“40 G 的速率”。另外要注意的是，当提到网络的速率时，往往指的是**额定速率**或**标称速率**，而并非网络实际上运行的速率。

2. 带宽

“**带宽**”(bandwidth)有以下两种不同的意义：

(1) 带宽本来是指某个**信号具有的频带宽度**。信号的带宽是指该信号所包含的各种不同频率成分所占据的频率范围。例如，在传统的通信线路上传送的电话信号的标准带宽是 3.1 kHz（从 300 Hz 到 3.4 kHz，即话音的主要成分的频率范围）。这种意义的**带宽的单位**是**赫**（或**千赫**、**兆赫**、**吉赫**等）。在过去很长的一段时间，通信的主干线路传送的是模拟信号（即连续变化的信号）。因此，表示某信道允许通过的信号频带范围就称为该信道的**带宽**（或**通频带**）。

(2) 在计算机网络中，带宽用来表示网络中某**通道**传送数据的能力，因此网络带宽表示在单位时间内网络中的某信道所能通过的**“最高数据率”**。在本书中提到“带宽”时，主要是指这个意思。这种意义的**带宽的单位**就是**数据率的单位** bit/s，**是“比特每秒”**。

在“带宽”的上述两种表述中，前者为**频域**称谓，而后者为**时域**称谓，其本质是相同的。也就是说，一条通信链路的“带宽”越宽，其所能传输的“最高数据率”也越高。

3. 吞吐量

吞吐量(throughput)表示在单位时间内通过某个网络（或信道、接口）的实际数据量。吞吐量更经常地用于对现实世界中的网络的一种测量，以便知道实际上到底有多少数据量能够通过网络。显然，吞吐量受网络带宽或网络额定速率的限制。例如，对于一个 1 Gbit/s 的以太网，就是说其额定速率是 1 Gbit/s，那么这个数值也是该以太网的吞吐量的绝对上限值。因此，对 1 Gbit/s 的以太网，其实际的吞吐量可能只有 100 Mbit/s，甚至更低，并没有达到其额定速率。请注意，有时吞吐量还可用每秒传送的字节数或帧数来表示。

接入到互联网的主机的实际吞吐量，取决于互联网的具体情况。假定主机 A 和服务器 B 接入到互联网的链路速率分别是 100 Mbit/s 和 1 Gbit/s。如果互联网的各链路的容量都足够大，那么当 A 和 B 交换数据时，其吞吐量显然应当是 100 Mbit/s。这是因为，尽管服务器 B 能够以超过 100 Mbit/s 的速率发送数据，但主机 A 最高只能以 100 Mbit/s 的速率接收数据。现在假定有 100 个用户同时连接到服务器 B（例如，同时观看服务器 B 发送的视频节目）。在这种情况下，服务器 B 连接到互联网的链路容量被 100 个用户平分，每个用户平均只能分到 10 Mbit/s 的带宽。这时，主机 A 连接到服务器 B 的吞吐量就只有 10 Mbit/s 了。

最糟糕的情况就是如果互联网的某处发生了严重的拥塞，则可能导致主机 A 暂时收不到服务器发来的视频数据，因而使主机 A 的吞吐量下降到零！主机 A 的用户或许会想，我已经向运营商的 ISP 交了速率为 100 Mbit/s 的宽带接入费用，怎么现在不能保证这个速率呢？其实你交的宽带费用，只是保证了从你家里到运营商 ISP 的某个路由器之间的数据传输速率。再往后的速率就取决于整个互联网的流量分布了，这是任何单个用户都无法控制的。了解这一点，对理解互联网的吞吐量是有帮助的。

4. 时延

时延(delay 或 latency)是指数据（一个报文或分组，甚至比特）从网络（或链路）的一端传送到另一端所需的时间。时延是个很重要的性能指标，它有时也称为**延迟**或**迟延**。

需要注意的是，网络中的时延是由以下几个不同的部分组成的：

(1) **发送时延**　**发送时延**(transmission delay)**是主机或路由器发送数据帧所需要的时间，**也就是从发送数据帧的第一个比特算起，到该帧的最后一个比特发送完毕所需的时间。因此发送

时延也叫作**传输时延**（我们尽量不采用传输时延这个名词，因为它很容易和下面要讲到的传播时延弄混）。发送时延的计算公式是：

$$\text{传播时延} = \frac{\text{信道长度(m)}}{\text{电磁波在信道上的传播速率(m/s)}} \tag{1-1}$$

由此可见，对于一定的网络，发送时延并非固定不变，而是与发送的帧长（单位是比特）成正比，与发送速率成反比。

(2) **传播时延**　传播时延(propagation delay)**是电磁波在信道中传播一定的距离需要花费的时间**。传播时延的计算公式是：

$$\text{传播时延} = \frac{\text{信道长度(m)}}{\text{电磁波在信道上的传播速率(m/s)}} \tag{1-2}$$

电磁波在自由空间的传播速率是光速，即 3.0×10^5 km/s。电磁波在网络传输媒体中的传播速率比在自由空间要略低一些：在铜线电缆中的传播速率约为 2.3×10^5 km/s，在光纤中的传播速率约为 2.0×10^5 km/s。例如，1000 km 长的光纤线路产生的传播时延大约为 5 ms。

以上两种时延有本质上的不同。但只要理解这两种时延发生的地方就不会把它们弄混。发送时延发生在机器内部的发送器中（一般就是发生在网络适配器中，见第 3 章 3.3.1 节），**与传输信道的长度（或信号传送的距离）没有任何关系**。但传播时延则发生在机器外部的传输信道媒体上，而**与信号的发送速率无关。信号传送的距离越远，传播时延就越大**。可以用一个简单的比喻来说明。假定有 10 辆车按顺序从公路收费站入口出发到相距 50 公里的目的地。再假定每一辆车过收费站要花费 6 秒，而车速是每小时 100 公里。现在可以算出这 10 辆车从收费站到目的地总共要花费的时间：发车时间共需 60 秒（相当于网络中的发送时延），在公路上的行车时间需要 30 分钟（相当于网络中的传播时延）。因此从第一辆车到收费站开始计算，到最后一辆车到达目的地为止，总共花费的时间是二者之和，即 31 分钟。

下面还有两种时延也需要考虑，但比较容易理解。

(3) **处理时延**　主机或路由器在收到分组时要花费一定的时间进行处理，例如分析分组的首部、从分组中提取数据部分、进行差错检验或查找转发表等，这就产生了处理时延。

(4) **排队时延**　分组在经过网络传输时，要经过许多路由器。但分组在进入路由器后要先在输入队列中排队等待处理。在路由器确定了转发接口后，还要在输出队列中排队等待转发。这就产生了排队时延。排队时延的长短往往取决于网络当时的通信量。当网络的通信量很大时会发生队列溢出，使分组丢失，这相当于排队时延为无穷大。

这样，数据在网络中经历的总时延就是以上四种时延之和：

$$\text{总时延} = \text{发送时延} + \text{传播时延} + \text{处理时延} + \text{排队时延} \tag{1-3}$$

一般说来，小时延的网络要优于大时延的网络。在某些情况下，一个低速率、小时延的网络很可能要优于一个高速率但大时延的网络。

图 1-12 画出了这几种时延所产生的地方，希望读者能够更好地分清这几种时延。

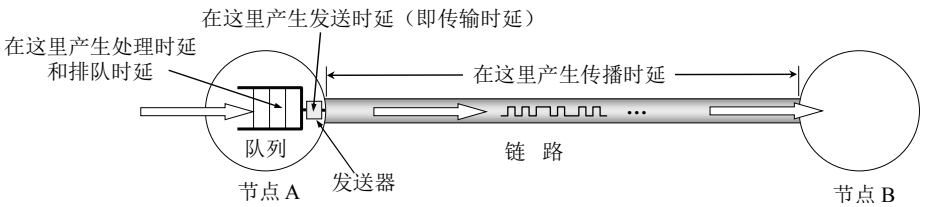


图 1-12　几种时延产生的地方不一样

必须指出，在总时延中，究竟是哪一种时延占主导地位，必须具体分析。下面举个例子。

现在我们暂时忽略处理时延和排队时延[①]。假定有一个长度为 100 MB 的数据块（这里的 M 显然不是指 10^6 而是指 2^{20}。B 是字节，1 字节 = 8 比特）。在带宽为 1 Mbit/s 的信道上（这里的 M 显然是 10^6）连续发送（即发送速率为 1 Mbit/s），其发送时延是

$$100 \times 2^{20} \times 8 \div 10^6 = 838.9 \text{ s}$$

现在把这个数据块用光纤传送到 1000 km 远的计算机。由于在 1000 km 的光纤上的传播时延约为 5 ms，因此在这种情况下，发送 100 MB 的数据块的总时延 = 838.9 + 0.005 ≈ 838.9 s。可见对于这种情况，发送时延决定了总时延的数值。

如果我们把发送速率提高到 100 倍，即提高到 100 Mbit/s，那么总时延就变为 8.389 + 0.005 = 8.394 s，缩小到原有数值的 1/100。

但是，并非在任何情况下，提高发送速率就能减小总时延。例如，要传送的数据仅有 1 个字节（如键盘上键入的一个字符，共 8 bit）。当发送速率为 1 Mbit/s 时，发送时延是

$$8 \div 10^6 = 8 \times 10^{-6} \text{ s} = 8 \ \mu\text{s}$$

若传播时延仍为 5 ms，则总时延为 5.008 ms。在这种情况下，传播时延决定了总时延。如果我们把数据率提高到 1000 倍（即将数据的发送速率提高到 1 Gbit/s），不难算出，总时延基本上仍是 5 ms，并没有明显减小。这个例子告诉我们，不能笼统地认为：“数据的发送速率越高，其传送的总时延就越小”。这是因为数据传送的总时延是由公式(1-3)右端的四项时延组成的，不能仅考虑发送时延一项。

如果上述概念没有弄清楚，就很容易产生这样错误的概念：“在高速链路（或高带宽链路）上，比特会传送得更快些”。但这是不对的。我们知道，汽车在路面质量很好的高速公路上可明显地提高行驶速率。然而**对于高速网络链路，我们提高的仅仅是数据的发送速率而不是比特在链路上的传播速率**。荷载信息的电磁波在通信线路上的传播速率（这是光速的数量级）取决于通信线路的介质材料，而与数据的发送速率并无关系。**提高数据的发送速率只是减小了数据的发送时延**。还有一点也应当注意，就是数据的发送速率的单位是每秒发送多少个比特，这是指在**某个点**或**某个接口上**的发送速率。而传播速率的单位是每秒传播多少公里，是指在**某一段传输线路上**比特的传播速率。因此，通常所说的“光纤信道的传输速率高”是指可以用很高的速率向光纤信道发送数据，而光纤信道的传播速率实际上还要比铜线的传播速率略低一点。这是因为经过测量得知，光在光纤中的传播速率约为每秒 20.5 万公里，它比电磁波在铜线（如 5 类线）中的传播速率（每秒 23.1 万公里）略低一些。

上述的重要概念请读者务必弄清。

5. 利用率

利用率有信道利用率和网络利用率两种。信道利用率指出某信道有百分之几的时间是被利用的（有数据通过）。完全空闲的信道的利用率是零。网络利用率则是全网络的信道利用率的加权平均值。信道利用率并非越高越好。这是因为，根据排队论的理论，当某信道的利用率增大时，该信道引起的时延也就迅速增加。这和高速公路的情况有些相似。当高速公路上的车流量很大时，由于在公路上的某些地方会出现堵塞，因此行车所需的时间就会变长。网络也有类似的情况。当网络的通信量很少时，网络产生的时延并不大。但在网络通信量不断增大的情况下，由于分组在网络节点（路由器或节点交换机）进行处理时需要排队等候，因此网络引起的时延就会增大。如果令 D_0 表示网络空闲时的时延，D 表示网络当前的时延（设现在的网络利用率为 U），那么在适当的假定条件下，可以用下面的简单公式(1-4)来表示 D 与 D_0 以及利用

① 注：当计算机网络中的通信量过大时，网络中的许多路由器的处理时延和排队时延将会大大增加，因而处理时延和排队时延有可能在总时延中占据主要成分。这时整个网络的性能就变差了。

率 U 之间的关系：

$$D=\frac{D_0}{1-U} \tag{1-4}$$

这里 U 是网络利用率，数值在 0 到 1 之间。当网络利用率达到其容量的 1/2 时，时延就要加倍。特别值得注意的就是：当网络利用率接近最大值 1 时，网络产生的时延就趋于无穷大。因此我们必须有这样的概念：**信道利用率或网络利用率过高就会产生非常大的时延**。图 1-13 给出了上述概念的示意图。因此，一些拥有较大主干网的 ISP 通常控制信道利用率不超过 50%。如果超过了就要准备扩容，增大线路的带宽。

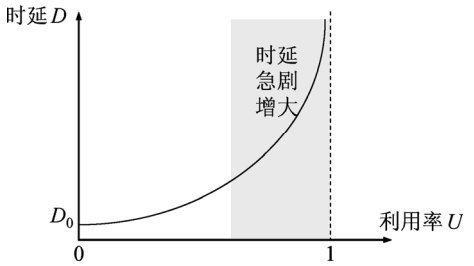


图 1-13　时延与利用率的关系

1.7　计算机网络体系结构

在计算机网络的基本概念中，分层次的**体系结构**（或**架构**）是最基本的。计算机网络体系结构的抽象概念较多，在学习时要多思考。这些概念对后面的学习很有帮助。

1.7.1　计算机网络体系结构的形成

计算机网络是个非常复杂的系统。为了说明这一点，可以设想一种最简单的情况：连接在网络上的两台计算机要互相传送文件。

显然，在这两台计算机之间必须有一条传送数据的通路。但这还远远不够。至少还有以下几项工作需要去完成：

(1) 发起通信的计算机必须发出一些信令，保证要传送的计算机数据能在这条通路上正确发送和接收。

(2) 告诉网络如何识别接收数据的计算机。

(3) 发起通信的计算机必须查明对方计算机是否已开机，并且与网络连接正常。

(4) 发起通信的计算机中的应用程序必须弄清楚，在对方计算机中的文件管理程序是否已做好接收文件和存储文件的准备工作。

(5) 若计算机的文件格式不兼容，则至少其中一台计算机应完成格式转换功能。

(6) 对出现的各种差错和意外事故，如数据传送错误、重复或丢失，网络中某个节点交换机出现故障等，应当有可靠的措施保证对方计算机最终能够收到正确的文件。

还可以列举一些要做的其他工作。由此可见，相互通信的两个计算机系统必须高度协调工作才行，而这种“协调”是相当复杂的。为了设计这样复杂的计算机网络，早在最初的 ARPANET 设计时即提出了分层的方法。“**分层**”可将庞大而复杂的问题，转化为若干较小的局部问题，而这些较小的局部问题就比较易于研究和处理。

1974 年，美国的 IBM 公司宣布了**系统网络体系结构** SNA (System Network Architecture)。这个著名的网络标准就是按照分层的方法制定的。现在用 IBM 大型机构建的专用网络仍在使用 SNA。不久后，其他一些公司也相继推出自己公司的具有不同名称的体系结构。

不同的网络体系结构出现后，使用同一个公司生产的各种设备都能够很容易地互连成网。这种情况显然有利于一个公司垄断市场。但由于网络体系结构的不同，不同公司的设备很难互相连通。

然而，全球经济的发展使得不同网络体系结构的用户迫切要求能够互相交换信息。为了使不同体系结构的计算机网络都能互连，国际标准化组织 ISO 于 1977 年成立了专门机构研究该问题。他们提出了一个试图使各种计算机在世界范围内互连成网的标准框架，即著名的**开放系统互连基本参考模型** OSI/RM (Open Systems Interconnection Reference Model)，简称为 OSI。OSI/RM 是个抽象的概念。在 1983 年形成了开放系统互连基本参考模型的正式文件，即所谓的七层协议的体系结构。

OSI 试图达到一种理想境界，即全球计算机网络都遵循这个统一标准，因而全球的计算机将能够很方便地进行互连和交换数据。在 20 世纪 80 年代，许多大公司甚至一些国家的政府机构纷纷表示支持 OSI。当时看来似乎在不久的将来全世界一定会按照 OSI 制定的标准来构造自己的计算机网络。然而到了 20 世纪 90 年代初期，虽然整套的 OSI 国际标准都已经制定出来了，但由于基于 TCP/IP 的互联网已抢先在全球相当大的范围成功地运行了，而与此同时却几乎找不到有什么厂家生产出符合 OSI 标准的商用产品。因此人们得出这样的结论：OSI 只获得了一些理论研究的成果，但在市场化方面则事与愿违地失败了。现今规模最大的、覆盖全球的、基于 TCP/IP 的互联网并未使用 OSI 标准。OSI 失败的原因可归纳为：

(1) OSI 的专家们缺乏实际经验，他们在完成 OSI 标准时缺乏商业驱动力；

(2) OSI 的协议实现起来过分复杂，而且运行效率很低；

(3) OSI 标准的制定周期太长，因而使得按 OSI 标准生产的设备无法及时进入市场；

(4) OSI 的层次划分不太合理，有些功能在多个层次中重复出现。

按照一般的概念，网络技术和设备只有符合有关的国际标准才能大范围地获得工程上的应用。但现在情况却反过来了。得到最广泛应用的不是**法律上的国际标准** OSI，而是非国际标准 TCP/IP。这样，TCP/IP 就常被称为是**事实上的国际标准**。从这种意义上说，能够占领市场的就是标准。在过去制定标准的组织中往往以专家、学者为主。但现在许多公司都纷纷加入各种标准化组织，使得技术标准具有浓厚的商业气息。一个新标准的出现，有时不一定反映其技术水平是最先进的，而是往往有着一定的市场背景。

1.7.2 协议与划分层次

在计算机网络中要做到有条不紊地交换数据，就必须遵守一些事先约定好的规则。**这些规则明确规定了所交换的数据的格式以及有关的同步问题**。这里所说的**同步**不是狭义的（即同频或同频同相）而是广义的，即在一定的条件下应当发生什么事件（例如，应当发送一个应答信息），因而**同步含有时序的意思**。这些**为进行网络中的数据交换而建立的规则、标准或约定**称为**网络协议**(network protocol)。

网络协议是计算机网络不可缺少的组成部分。实际上，只要我们想让连接在网络上的另一台计算机做点什么事情（例如，从网络上的某台主机下载文件），都需要有协议。但是当我们经常在自己的个人电脑上进行文件存盘操作时，就**不需要任何网络协议**，除非这个用来存储文件的磁盘是网络上的某个文件服务器的磁盘。

协议通常有两种不同的形式。一种是使用便于人来阅读和理解的文字描述。另一种是使用让计算机能够理解的程序代码。这两种不同形式的协议都必须能够对网络上的信息交换过程做出精确的解释。

ARPANET 的研制经验表明，对于非常复杂的计算机网络协议，其结构应该是层次式的。我们可以举一个简单的例子来说明划分层次的概念。

现在假定我们在主机 1 和主机 2 之间通过一个通信网络传送文件。这是一项比较复杂的工

作，因为需要做不少的工作。

我们可以将要做的工作划分为三类。第一类工作与传送文件直接有关。例如，发送端的文件传送应用程序应当确信接收端的文件管理程序已做好接收和存储文件的准备。若两台主机所用的文件格式不一样，则至少其中的一台主机应完成文件格式的转换。这两项工作可用一个文件传送模块来完成。这样，两台主机可将文件传送模块作为最高的一层（如图 1-14 所示）。在这两个模块之间的虚线表示两台主机系统交换文件和一些有关文件交换的命令。

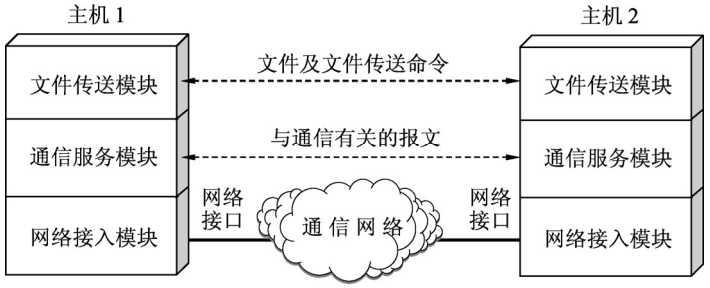


图 1-14　划分层次举例

但是，我们并不想让文件传送模块完成全部工作的细节，这样会使文件传送模块过于复杂。可以再设立一个通信服务模块，用来保证文件和文件传送命令可靠地在两个系统之间交换。也就是说，让位于上面的文件传送模块利用下面的通信服务模块所提供的服务。我们还可以看出，如果将位于上面的文件传送模块换成电子邮件模块，那么电子邮件模块同样可以利用在它下面的通信服务模块所提供的可靠通信的服务。

同理，我们可以再构造一个网络接入模块，让这个模块负责做与网络接口细节有关的工作，并向上层提供服务，使上面的通信服务模块能够完成可靠通信的任务。

从上述的简单例子可以更好地理解分层能带来很多好处，如：

(1) **各层之间是独立的**。某一层并不需要知道它的下一层是如何实现的，而仅仅需要知道该层通过层间的接口（即界面）所提供的服务。由于每一层只实现一种相对独立的功能，因而可将一个难以处理的复杂问题分解为若干个较容易处理的更小一些的问题。这样，整个问题的复杂程度就下降了。

(2) **灵活性好**。当任何一层发生变化时（例如由于技术的变化），只要层间接口关系保持不变，那么在这层以上或以下各层均不受影响。此外，对某一层提供的服务还可进行修改。当不再需要某层提供的服务时，甚至可以将这层取消。

(3) **结构上可分割开**。各层都可以采用最合适的技术来实现。

(4) **易于实现和维护**。这种结构使得实现和调试一个庞大而又复杂的系统变得更加容易，因为整个系统已被分解为若干个相对独立的子系统。

(5) **能促进标准化工作**。因为每一层的功能及其所提供的服务都已有了精确的说明。

分层时应注意使每一层的功能非常明确。若层数太少，就会使每一层的协议太复杂；但层数太多又会在描述和综合各层功能的系统工程任务时遇到较多的困难。通常各层所要完成的功能主要有以下一些（可以只包括一种，也可以包括多种）：

① **差错控制**　使相应层次对等方的通信更加可靠。

② **流量控制**　发送端的发送速率必须使接收端来得及接收，不要太快。

③ **分段和重装**　发送端将要发送的数据块划分为更小的单位，在接收端将其还原。

④ **复用和分用**　发送端几个高层会话复用一条低层的连接，在接收端再进行分用。

⑤ **连接建立和释放**　交换数据前先建立一条逻辑连接，数据传送结束后释放连接。

分层当然也有一些缺点，例如，有些功能会在不同的层次中重复出现，因而产生额外开销。

计算机网络的各层及其协议的集合就是网络的**体系结构**(architecture)。换种说法，**计算机网络的体系结构就是这个计算机网络及其构件所应完成的功能的精确定义**。需要强调的是：这些功能究竟是用何种硬件或软件完成的，则是一个遵循这种体系结构的**实现**(implementation)的问题。体系结构的英文名词 architecture 的原意是建筑学或建筑的设计和风格。它和一个具体的建筑物的概念很不相同。例如，我们可以走进一个明代的建筑物中，但却不能走进一个明代的建筑风格之中。同理，我们也不能把一个具体的计算机网络说成是一个抽象的网络体系结构。总之，**体系结构是抽象的，而实现则是具体的，是真正在运行的计算机硬件和软件**。

1.7.3 具有五层协议的体系结构

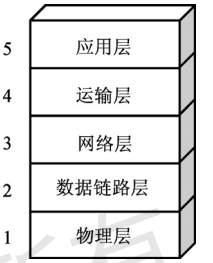


图 1-15 具有五层协议计算机网络体系结构

OSI 的七层协议体系结构的概念清楚，理论也较完整，但它既复杂又不实用。TCP/IP 体系结构则不同，它现在得到了非常广泛的应用。TCP/IP 是一个四层的体系结构，它包含应用层、运输层、网际层和链路层（网络接口层）。用网际层这个名字是强调本层解决不同网络的互连问题。从实质上讲，TCP/IP 只有最上面的三层，因为最下面的链路层并没有属于 TCP/IP 体系的具体协议。链路层所使用的各种局域网标准，并非由 IETF 而是由 IEEE 的 802 委员会下属的各工作组负责制定的。在讲授计算机网络原理时往往采取另外的办法，即综合 OSI 和 TCP/IP 的优点，采用如图 1-15 所示的五层协议的体系结构，这对阐述计算机网络的原理是十分方便的。

现在结合互联网的情况，自上而下地、非常简要地介绍一下各层的主要功能。实际上，只有认真学习完本书各章的协议后才能真正弄清各层的作用。

(1) **应用层**(application layer)

应用层是体系结构中的最高层。应用层的任务是**通过应用进程间的交互来完成特定网络应用**。应用层协议定义的是**应用进程间通信和交互的规则**。这里的**进程**就是指主机中**正在运行的程序**。对于不同的网络应用需要有不同的应用层协议。互联网中的应用层协议很多，如域名系统 DNS、支持万维网应用的协议 HTTP、支持电子邮件的协议 SMTP，等等。我们把应用层交互的数据单元称为**报文**(message)。

(2) **运输层**(transport layer)

运输层的任务就是负责向**两台主机中进程之间的通信**提供**通用的数据传输**服务。应用进程利用该服务传送应用层报文。所谓“通用的”，是指并不针对某个特定网络应用，而是多种应用可以使用同一个运输层服务。由于一台主机可同时运行多个进程，因此运输层有复用和分用的功能。复用就是多个应用层进程可同时使用下面运输层的服务，分用和复用相反，是运输层把收到的信息分别交付上面应用层中的相应进程。

运输层主要使用以下两种协议：

- **传输控制协议** TCP (Transmission Control Protocol)——提供面向连接的、可靠的数据传输服务，其数据传输的单位是**报文段**(segment)。
- **用户数据报协议** UDP (User Datagram Protocol)——提供无连接的**尽最大努力** (best-effort)的数据传输服务（不保证数据传输的可靠性），其数据传输的单位是**用户数据报**。

顺便指出，现在很多人愿意把这一层称为**传输层**，因为这一层的重要协议 TCP 的 T 是传输的意思。但互联网标准 RFC 1122 采用的名词是 Transport 而不是 Transmission，因此本书采

用“运输层”作为 Transport 的译名可能较为准确。

(3) **网络层**(network layer)

网络层负责为分组交换网上的不同**主机**提供通信服务。在发送数据时，网络层把运输层产生的报文段或用户数据报封装成**分组**或**包**进行传送。在 TCP/IP 体系中，由于网络层使用协议 IP，因此分组也叫作 **IP 数据报**，或简称为**数据报**。**本书把“分组”和“数据报”作为同义词使用**。

请注意：不要将运输层的“用户数据报协议 UDP”和网络层的“IP 数据报”弄混。此外，**无论在哪一层传送的数据单元，都可笼统地用“分组”来表示**。

网络层的具体任务有两个。第一个任务是通过一定的算法，在互联网中的每一个路由器上生成一个用来转发分组的转发表。第二个任务较为简单，就是每一个路由器在接收到一个分组时，依据转发表中指明的路径把分组转发到下一个路由器。这样就可以使源主机运输层所传下来的分组，能够通过合适的路由最终到达目的主机。

这里要强调指出，网络层中的“**网络**”二字，已不是我们通常谈到的具体网络，而是在计算机网络体系结构模型中的第 3 层的名称。

互联网是由大量的**异构**(heterogeneous)网络通过**路由器**(router)相互连接起来的。互联网使用的网络层协议是无连接的**网际协议** IP (Internet Protocol)和许多种路由选择协议，因此互联网的网络层也叫作**网际层**或 **IP 层**。在本书中，网络层、网际层和 IP 层都是同义语。

(4) **数据链路层**(data link layer)

数据链路层常简称为**链路层**。我们知道，两台主机之间的数据传输，总是在一段一段的链路上传送的，这就需要使用专门的链路层的协议。在两个相邻节点之间传送数据时，数据链路层将网络层交下来的 IP 数据报**组装成帧**(framing)，在两个相邻节点间的链路上传送**帧**(frame)。每一帧包括数据和必要的**控制信息**（如同步信息、地址信息、差错控制等）。

在接收数据时，控制信息使接收端能够知道一个帧从哪个比特开始和到哪个比特结束。这样，数据链路层在收到一个帧后，就可从中提取出数据部分，上交给网络层。

控制信息还使接收端能够检测到所收到的帧中有无差错。如发现有差错，数据链路层就简单地**丢弃**这个出了差错的帧，以免继续在网络中传送下去白白浪费网络资源。如果需要改正数据在数据链路层传输时出现的差错（这就是说，数据链路层不仅要检错，而且要纠错），那么就要采用可靠传输协议来纠正出现的差错。这种方法会使数据链路层的协议复杂些。

(5) **物理层**(physical layer)

在物理层上所传数据的单位是**比特**。发送方发送 1（或 0）时，接收方应当收到 1（或 0）而不是 0（或 1）。因此物理层要考虑用多大的电压代表“1”或“0”，以及接收方如何识别出发送方所发送的比特。物理层还要确定连接电缆的插头应当有多少根引脚以及各引脚应如何连接。当然，解释比特代表的意思，不是物理层的任务。请注意，传递信息所利用的一些物理传输媒体，如双绞线、同轴电缆、光缆、无线信道等，并不在物理层协议之内，而是在物理层协议的下面。因此也有人把物理层下面的物理传输媒体当作第 0 层。

在互联网所使用的各种协议中，最重要的和最著名的就是 TCP 和 IP 两个协议。现在人们经常提到的 TCP/IP 并不一定是单指 TCP 和 IP 这两个具体的协议，而往往是表示互联网所使用的整个 **TCP/IP 协议族**(protocol suite)[①]。

图 1-16 说明的是应用进程的数据在各层之间的传递过程中所经历的变化。这里为简单起见，假定两台主机通过一台路由器连接起来。

① 注：请注意 suite 这个词的特殊读音/swi:t/，不要读错。

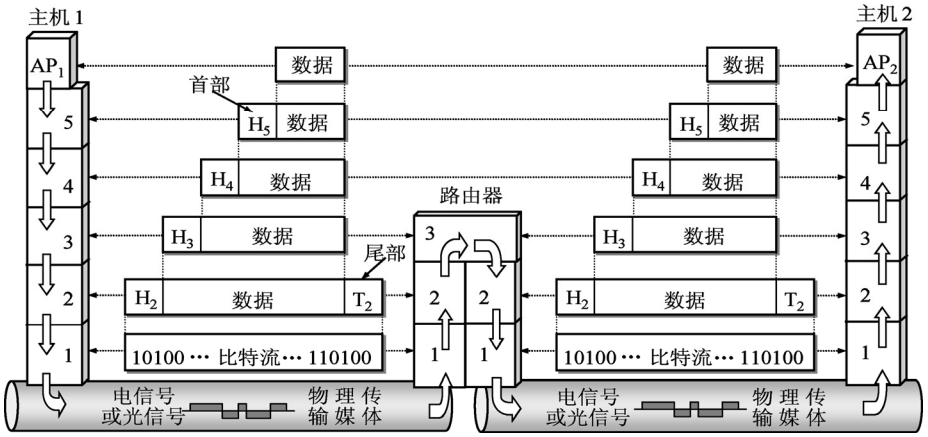


图 1-16　数据在各层之间的传递过程

假定主机 1 的应用进程 AP_1 向主机 2 的应用进程 AP_2 传送数据。AP_1 先将其数据交给本主机的第 5 层（应用层）。第 5 层加上必要的控制信息 H_5 就变成了下一层的数据单元。第 4 层（运输层）收到这个数据单元后，加上本层的控制信息 H_4，再交给第 3 层（网络层），成为第 3 层的数据单元。依此类推。不过到了第 2 层（数据链路层）后，控制信息被分成两部分，分别加到本层数据单元的首部（H_2）和尾部（T_2）；而第 1 层（物理层）由于是比特流的传送，所以不再加上控制信息。请注意，传送比特流时应从首部开始传送。

OSI 参考模型把对等层次之间传送的数据单位称为该层的**协议数据单元** PDU (Protocol Data Unit)。这个名词现已被许多非 OSI 标准采用。

当这一串的比特流离开主机 1 经网络的物理传输媒体传送到路由器时，就从路由器的第 1 层依次上升到第 3 层。每一层都根据控制信息进行必要的操作，然后将控制信息剥去，将该层剩下的数据单元上交给更高的一层。当分组上升到了第 3 层网络层时，就根据首部中的目的地址查找路由器中的转发表，找出转发分组的接口，然后往下传送到第 2 层，加上新的首部和尾部后，再到最下面的第 1 层，然后在物理传输媒体上把每一个比特发送出去。

当这一串的比特流离开路由器到达目的站主机 2 时，就从主机 2 的第 1 层按照上面讲过的方式，依次上升到第 5 层。最后，把应用进程 AP_1 发送的数据交给目的站的应用进程 AP_2。

可以用一个简单例子来比喻上述过程。有一封信从最高层向下传。每经过一层就包上一个新的信封，写上必要的地址信息。包有多个信封的信件传送到目的站后，从第 1 层起，每层拆开一个信封后就把信封中的信交给它的上一层。传到最高层后，取出发信人所发的信交给收信人。

虽然应用进程数据要经过如图 1-16 所示的复杂过程才能送到终点的应用进程，但这些复杂过程对用户屏蔽掉了，以致应用进程 AP_1 觉得好像是直接把数据交给了应用进程 AP_2。同理，任何两个同样的层次（例如在两个系统的第 4 层）之间，也好像如同图 1-17 中的水平虚线所示的那样，把数据（即数据单元加上控制信息）通过水平虚线直接传递给对方。这就是所谓的“**对等层**”(peer layers)之间的通信。我们以前经常提到的各层协议，实际上就是在各个对等层之间传递数据时的各项规定。

在文献中也还可以见到术语“**协议栈**”(protocol stack)。这是因为几个层次画在一起很像一个**栈**(stack)的结构。

1.7.4　实体、协议、服务和服务访问点

当研究开放系统中的信息交换时，往往使用**实体**(entity)这一较为抽象的名词表示**任何可发**

送或接收信息的硬件或软件进程。在许多情况下，实体就是一个特定的软件模块。

协议是控制两个对等实体（或多个实体）进行通信的规则的集合。协议的语法方面的规则定义了所交换的信息的格式，而协议的语义方面的规则就定义了发送者或接收者所要完成的操作，例如，在何种条件下，数据必须重传或丢弃。

在协议的控制下，两个对等实体间的通信使得本层能够向上一层提供服务。要实现本层协议，还需要使用下面一层所提供的服务。

一定要弄清楚，协议和服务在概念上是很不一样的。

首先，协议的实现保证了能够向上一层提供服务。**使用本层服务的实体只能看见服务而无法看见下面的协议**。也就是说，**下面的协议对上面的实体是透明的**。

其次，**协议是“水平的”**，即协议是控制对等实体之间通信的规则。但**服务是“垂直的”**，即服务是由下层向上层通过层间接口提供的。另外，并非在一个层内完成的全部功能都称为服务。只有那些能够被高一层实体“**看得见**”的功能才能称之为“服务”。

在同一系统中相邻两层的实体进行交互（即交换信息）的地方，通常称为**服务访问点** SAP (Service Access Point)。服务访问点 SAP 是一个抽象的概念，它实际上就是一个逻辑接口，有点像邮政信箱（可以把邮件放入信箱和从信箱中取走邮件），但这种层间的逻辑接口和两个设备之间的硬件接口（并行的或串行的）是不一样的。

计算机网络的协议还有一个很重要的特点，就是协议必须把**所有**不利的条件事先都估计到，而**不能假定一切都是正常的和非常理想的**。例如，两个朋友在电话中约好，下午 3 时在某公园门口碰头，并且约定“不见不散”。这就是一个很不科学的协议，因为任何一方临时有急事来不了而又无法通知对方时（如对方的电话或手机都无法接通），则另一方按照协议就必须永远等待下去。因此，看一个计算机网络协议是否正确，不能只看在正常情况下是否正确，还必须**非常仔细地检查协议能否应付任何一种出现概率极小的异常情况**。

下面是一个有关网络协议的非常著名的例子。

【例 1-1】 占据东、西两个山顶的蓝军 1 和蓝军 2 与驻扎在山谷的白军作战。其力量对比是：单独的蓝军 1 或蓝军 2 打不过白军，但蓝军 1 和蓝军 2 协同作战则可战胜白军。现蓝军 1 拟于次日正午向白军发起攻击。于是用计算机发送电文给蓝军 2。但通信线路很不好，电文出错或丢失的可能性较大（没有电话可使用）。因此要求收到电文的友军必须送回一个确认电文。但此确认电文也可能出错或丢失。试问能否设计出一种协议使得蓝军 1 和蓝军 2 能够实现协同作战因而一定（即 100 %而不是 99.999…%）取得胜利？

【解】 蓝军 1 先发送：“拟于明日正午向白军发起攻击。请协同作战和确认。”

假定蓝军 2 收到电文后发回了确认。

然而现在蓝军 1 和蓝军 2 都不敢下决心进攻。因为，蓝军 2 不知道此确认电文对方是否正确地收到了。如未正确收到，则蓝军 1 必定不敢贸然进攻。在此情况下，自己单方面发起进攻就肯定要失败。因此，必须等待蓝军 1 发送“对确认的确认”。

假定蓝军 2 收到了蓝军 1 发来的确认。但蓝军 1 同样关心自己发出的确认是否已被对方正确地收到。因此还要等待蓝军 2 的“对确认的确认的确认”。

这样无限循环下去，蓝军 1 和蓝军 2 都始终无法确定自己最后发出的电文对方是否已经收到（如图 1-17 所示）。因此，在本例题给出的条件下，没有一种协议可以使蓝军 1 和蓝军 2 能够 100%地确保胜利。

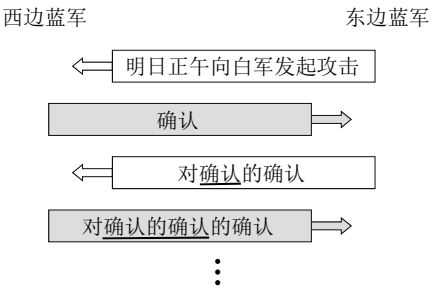


图 1-17　无限循环的协议

这个例子告诉我们，看似非常简单的协议，设计起来要考虑的问题还是比较多的。

1.7.5 TCP/IP 的体系结构

前面已经说过，TCP/IP 的体系结构比较简单，它只有四层。图 1-18 给出了这种四层协议表示方法的例子。

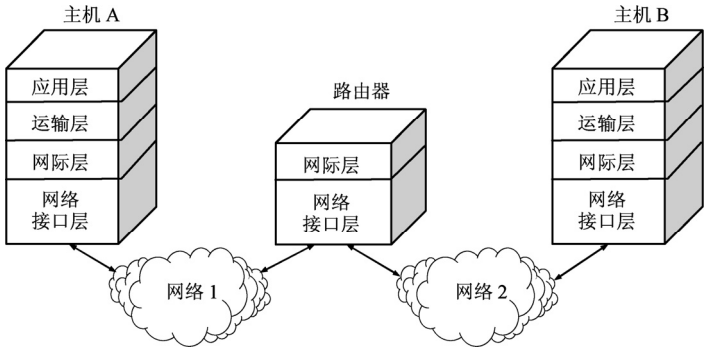


图 1-18　TCP/IP 四层协议的表示方法举例

还有另一种方法用来表示 TCP/IP 协议族（如图 1-19 所示），它的特点是上下两头大而中间小：应用层和网络接口层都有多种协议，而中间的 IP 层很小，上层的各种协议都向下汇聚到一个协议 IP 中。这种很像沙漏计时器形状的 TCP/IP 协议族表明：IP 层可以支持多种运输层协议（虽然这里只画出了最主要的两种），而不同的运输层协议上面又可以有多种应用层协议（所谓的 everything over IP），同时协议 IP 也可以在多种类型的网络上运行（所谓的 IP over everything）。正因为如此，互联网才会发展到今天的这种全球规模。从图 1-19 不难看出 IP 协议在互联网中的核心作用。

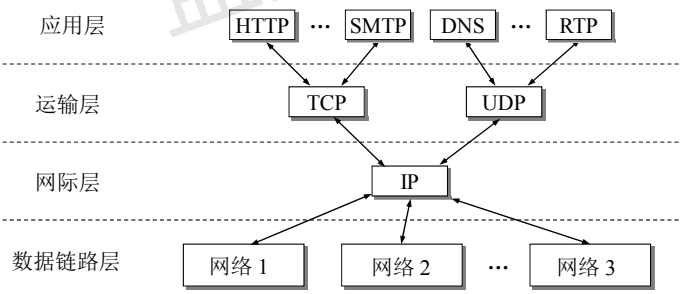


图 1-19　沙漏计时器形状的 TCP/IP 协议族示意

实际上，图 1-19 还反映出互联网的一个十分重要的设计理念，这就是网络的核心部分越简单越好，把一切复杂的部分让网络的边缘部分去实现。

【例 1-2】 利用协议栈的概念，说明在互联网中常用的客户–服务器工作方式。

【解】 图 1-20 中的主机 A 和主机 B 都各有自己的协议栈。主机 A 中的应用进程（即客户进程）的位置在最高的应用层。这个客户进程向主机 B 应用层的服务器进程发出请求，请求建立连接（图中的❶）。然后，主机 B 中的服务器进程接受 A 的客户进程发来的请求（图中的❷）。所有这些通信，实际上都需要使用下面各层所提供的服务。但若仅仅考虑客户进程和服务器进程的交互，则可把它们之间的交互看成是图 1-20 中的水平虚线所示的那样。

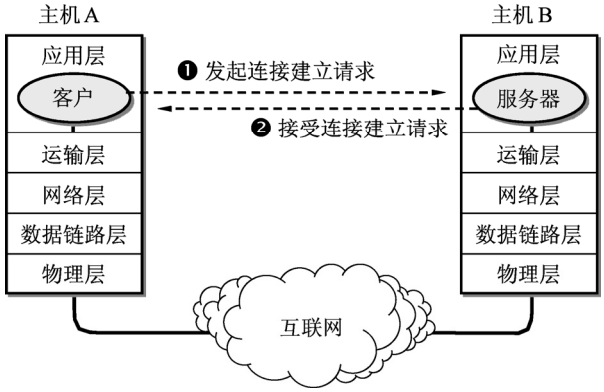


图 1-20　在应用层的客户进程和服务器进程的交互

图 1-21 画出了三台主机的协议栈。主机 C 的应用层中同时有两个服务器进程在通信。服务器 1 在和主机 A 中的客户 1 通信，而服务器 2 在和主机 B 中的客户 2 通信。有的服务器进程可以同时向几百个或更多的客户进程提供服务。

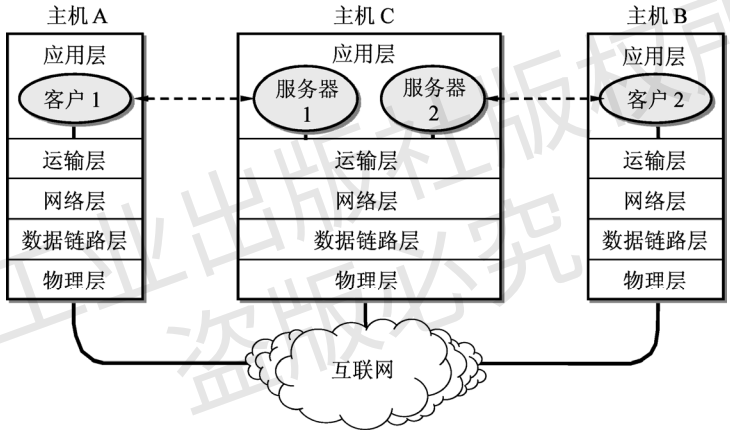


图 1-21　主机 C 的两个服务器进程分别向 A 和 B 的客户进程提供服务

本章的重要概念

- 计算机网络（可简称为网络）把许多计算机连接在一起，而互连网则把许多网络连接在一起，是网络的网络。
- 以小写字母 i 开始的 internet（互连网）是通用名词，它泛指由多个计算机网络互连而成的网络。在这些网络之间的通信协议（即通信规则）可以是任意的。
- 以大写字母 I 开始的 Internet（互联网）是专用名词，它指当前全球最大的、开放的、由众多网络相互连接而成的特定计算机网络，并采用 TCP/IP 协议族作为通信规则，且其前身是美国的 ARPANET。Internet 的推荐译名是“因特网”，但很少被使用。
- 互联网现在采用存储转发的分组交换技术以及三层 ISP 结构。
- 互联网按工作方式可划分为边缘部分与核心部分。主机在网络的边缘部分，其作用是进行信息处理。路由器在网络的核心部分，其作用是按存储转发方式进行分组交换。
- 计算机通信是计算机中的进程（即运行着的程序）之间的通信。计算机网络采用的通信方式是客户–服务器方式和对等连接方式（P2P 方式）。

- 客户和服务器都是指通信中所涉及的应用进程。客户是服务请求方，服务器是服务提供方。
- 按作用范围的不同，计算机网络分为广域网 WAN、城域网 MAN、局域网 LAN 和个人区域网 PAN。
- 计算机网络最常用的性能指标是：速率、带宽、吞吐量、时延（发送时延、传播时延、处理时延、排队时延）和信道（或网络）利用率。
- 网络协议即协议，是为进行网络中的数据交换而建立的规则。计算机网络的各层及其协议的集合，称为网络的体系结构。
- 五层协议的体系结构由应用层、运输层、网络层（或网际层）、数据链路层和物理层组成。运输层最重要的协议是 TCP 和 UDP，而网络层最重要的协议是协议 IP。

习题

1-01　计算机网络可以向用户提供哪些服务？

1-02　试简述分组交换的要点。

1-03　试从多个方面比较电路交换、报文交换和分组交换的主要优缺点。

1-04　为什么说互联网是自印刷术发明以来人类在存储和交换信息领域的最大变革？

1-05　互联网基础结构的发展大致分为哪几个阶段？请指出这几个阶段最主要的特点。

1-06　简述互联网标准制定的几个阶段。

1-07　小写和大写开头的英文名字 internet 和 Internet 在意思上有何重要区别？

1-08　计算机网络都有哪些类别？各种类别的网络都有哪些特点？

1-09　计算机网络中的主干网和本地接入网的主要区别是什么？

1-10　互联网的两大组成部分（边缘部分与核心部分）的特点是什么？它们的工作方式各有什么特点？

1-11　客户-服务器方式与 P2P 对等通信方式的主要区别是什么？有没有相同的地方？

1-12　计算机网络有哪些常用的性能指标？

1-13　假定网络的利用率达到了 90%。试估算一下现在的网络时延是它的最小值的多少倍？

1-14　收发两端之间的传输距离为 1000 km，信号在媒体上的传播速率为 2×10^8 m/s。试计算以下两种情况的发送时延和传播时延：

(1) 数据长度为 10^7 bit，数据发送速率为 100 kbit/s。

(2) 数据长度为 10^3 bit，数据发送速率为 1 Gbit/s。

从以上计算结果可得出什么结论？

1-15　假设信号在媒体上的传播速率为 2.3×10^8 m/s。媒体长度 l 分别为：

(1) 10 cm（网络接口卡）

(2) 100 m（局域网）

(3) 100 km（城域网）

(4) 5000 km（广域网）

现在连续传送数据，数据率分别为 1 Mbit/s 和 10 Gbit/s。试计算每一种情况下在媒体中的比特数。（提示：媒体中的比特数实际上无法使用仪表测量。本题是假想我们能够看见媒体中正在传播的比特，能够给媒体中的比特拍个快照。媒体中的比特数取决于媒体的长度和数据率。）

1-16　网络体系结构为什么要采用分层次的结构？试举出一些与分层体系结构的思想相似的日常生活的例子。

1-17　协议与服务有何区别？有何关系？

1-18　为什么一个网络协议必须把各种不利的情况都考虑到？

1-19　试述具有五层协议的网络体系结构的要点，包括各层的主要功能。

1-20　试举出日常生活中有关“透明”这一名词的例子。

1-21　试解释以下名词：协议栈、实体、对等层、协议数据单元、服务访问点、客户、服务器、客户-服务器方式。

1-22　试解释 everything over IP 和 IP over everything 的含义。