

1

模块 1

网络基础概论

▶ 本模块学习前的思考点

- 什么是计算机网络?
- 计算机网络是如何产生的?
- 计算机网络由哪些部分组成?
- 如何通过软件来画网络拓扑图?
- 网络新的发展方向有哪些?

1.1 必备知识

1.1.1 认知计算机网络

计算机网络是现代通信技术与计算机技术相结合的产物。它的发展不仅促进了人类社会的信息化和全球经济一体化,也改变了人们的生活习惯。

对“计算机网络”这个概念的理解和定义,随着计算机网络本身的发展,人们提出了各种不同的观点。

关于计算机网络的最简单定义是:一些相互连接的、以共享资源为目的的、自治的计算机的集合。

另外,从逻辑功能上看,计算机网络是以传输信息为基础目的,用通信线路将多个计算机连接起来的计算机系统的集合。一个计算机网络组成包括传输介质和通信设备。

从用户角度看,计算机网络是这样定义的:存在一个能为用户自动管理的网络操作系统。由它调用完成用户需要调用的资源,而整个网络像一个大的计算机系统一样,对用户是透明的。

综上所述,把计算机网络定义为:把分布在不同地理区域的计算机与专门的外部设备用

通信线路互联成一个规模大、功能强的系统，从而使众多的计算机可以方便地互相传递信息，共享硬件、软件、数据信息等资源。简单来说，计算机网络就是由通信线路互相连接的许多自主工作的计算机构成的集合体。

从定义中看出计算机网络包含的如下几个核心问题。

- (1) 至少两台计算机互连。
- (2) 通信设备与线路介质。
- (3) 网络软件、通信协议和网络操作系统。

1.1.2 了解计算机网络的发展过程

为了更好地了解网络的概念，这里首先了解计算机网络的演变。在计算机的概念中我们了解到，计算机网络是现代通信技术与计算机技术相结合的产物，一直以来它们紧密结合，相互促进，相互影响，共同推进计算机网络的发展。图 1-1 所示为计算机网络发展演进图。

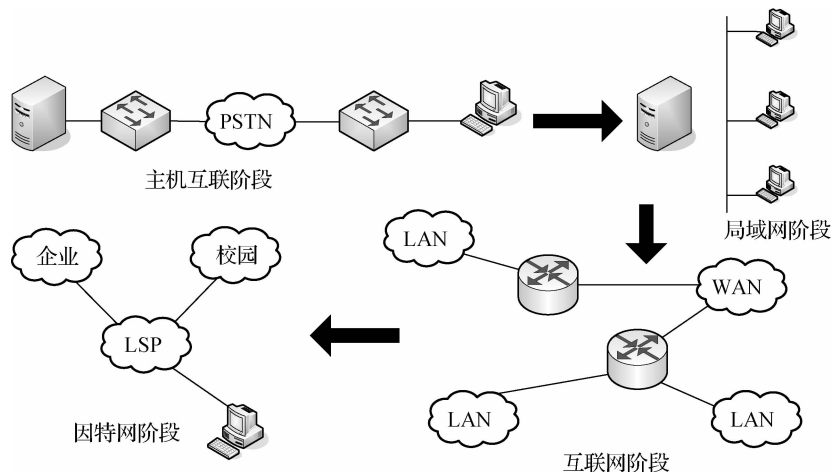


图 1-1 计算机网络发展演进图

1. 主机互联阶段

这种产生于 20 世纪 60 年代初期，基于主机之间的低速串行连接的联机系统是计算机网络的雏形。在这种早期的网络中，终端借助电话线路访问计算机，由于计算机发送/接收的为数字信号，电话线传输的是模拟信号，这就要求在终端和主机间加入调制解调器（modem，俗称“猫”），进行数/模间的转换。

在这种连接中，计算机是网络的中心，同时也是控制者。这是一种非常原始的计算机网络。它的主要任务是通过远程终端与计算机的连接，提供应用程序执行、远程打印和数据服务等功能。

2. 局域网阶段

20 世纪 70 年代，随着计算机体积减小、价格下降，出现了以个人计算机为主的商业计算模式。商业计算的复杂性要求大量终端设备的资源共享和协同操作，导致了对本地大量计算机设备进行网络化连接的需求，局域网由此产生了。

局域网技术——以太网就是在此时期产生的。1973年, Xerox(施乐)公司的 Robert Metcalfe 博士(以太网之父)提出并实现了最初的以太网。后来 DEC、Intel 和 Xerox 合作制定了一个产品标准,该标准最初以这3家公司名称的首字母命名,称为 DIX 以太网。其他流行的 LAN 技术还有 IBM 的令牌环技术等。

3. 互联网阶段

由于单一的局域网无法满足对网络的多样性要求,20世纪70年代后期,广域网技术逐渐发展起来,以便将分布在不同地域的局域网互相连接起来。1983年,ARPANET 采用传输控制协议(transmission control protocol, TCP)和因特网协议(Internet protocol, IP)作为其主要的协议簇,使大范围的网络互联成为可能。

4. 因特网阶段

20世纪80年代到90年代是网络互联发展时期。在这一时期,ARPANET 网络的规模不断扩大,将全球无数的公司、校园、因特网服务提供商(Internet service provider, ISP)和个人用户联系起来,最终演变成今天的几乎延伸到全球每一个角落的 Internet。1990年 ARPANET 正式被 Internet 取代,退出了历史舞台。越来越多的机构、个人参与到 Internet 中来,使得 Internet 获得了高速发展。

而到现在,网络几乎联系着所有国家和地区。2018全球互联网发展报告显示:互联网用户为40.12亿,同比增加7%;社交媒体用户数为31.96亿,同比增加13%。上面的数字表明,网络已经成为名副其实的世界信息资源最丰富的信息资源库。网络被认为是未来全球信息高速公路的雏形。

1.1.3 认识计算机网络的组成

由于网络是计算机技术和通信技术相互结合而成的,因而网络组成与通信技术和计算机技术都有联系。除此之外,网络的组成还必须匹配相应的网络软件系统。

1. 典型计算机网络的组成

典型的计算机网络由计算机系统、数据通信系统、网络软件及协议三大部分组成。

(1)计算机系统。计算机系统是网络的基本模块,它作为网络中的一个节点,为网络内的其他计算机提供共享资源。在网络中,按照计算机系统的用途可分为服务器和客户机。

①服务器。服务器(server)是网络环境中的高性能计算机,它侦听网络上的其他计算机(客户机)提交的服务请求,并提供相应的服务。为此,服务器必须具有承担服务并且保障服务的能力。相对于普通 PC 来说,服务器在稳定性、安全性、性能等方面都有更高的要求,因此服务器的 CPU、芯片组、内存、磁盘系统、网络等硬件和普通 PC 有所不同。

②客户机。简单来说,客户机(client)就是用户使用的计算机,它在网络中数量大、分布广。

在网络中对服务器和客户机没有特别的区分,对于一台计算机来说,如果作为信息的提供者,那就是服务器;如果为信息的使用者,就是客户机。

(2)数据通信系统。计算机网络中,数据通信系统的任务是把数据源计算机所产生的数据迅速、可靠、准确地传输到目的计算机或专用外设。

从计算机网络技术的组成来看,一个完整的数据通信系统一般由数据终端设备、通信控

制器、通信信道和网络互连设备几部分组成。

①数据终端设备:数据的生成者和使用者,它根据协议控制通信的功能。最常用的数据终端设备就是网络中的计算机。当然随着网络的发展,数据终端设备还可以是网络中的手机、PDA等。

②通信控制器:它除能进行通信状态的连接、监控和拆除等操作外,还可接收来自多个数据终端设备的信息,并转换信息格式。如最常见的网卡就是通信控制器。

③通信信道:通信信道是信息在信号变换器之间传输的通道。如电话线路等模拟通信信道、专用数字通信信道、宽带电缆和光纤等。

④网络互连设备:网络互连设备就是在物理上把两种网络连接起来。实现一种网络与另一种网络互访与通信,解决它们之间协议方面的差别,处理速率与带宽的差别,处理数据信号的变换等功能。主要包括中继器、网桥、路由器、桥由器、网关、集线器、交换机和调制解调器。

(3)网络软件及协议。网络软件和协议指在计算机网络环境中用于支持数据通信和各种网络活动的软件。通常根据系统本身的特点、能力和服务对象,为连入计算机网络的系统配置不同的网络应用系统。软件和协议的目的是为了本机用户共享网络中其他系统的资源,或是为了把本机系统的功能和资源提供给网络中其他用户使用。为此,每个计算机网络都制定了一套全网共同遵守的网络协议,并要求网络中每个主机系统配置相应的协议软件,以确保网络中不同系统之间能够可靠、有效地相互通信和合作。

2. 计算机网络的逻辑组成

从计算机系统功能上看,网络主要完成网络通信与资源共享。通常,人们把负责网络通信的部分称为通信子网,负责实现资源共享的部分称为资源子网。因此,可以把网络看成由通信子网与资源子网组成,如图 1-2 所示。

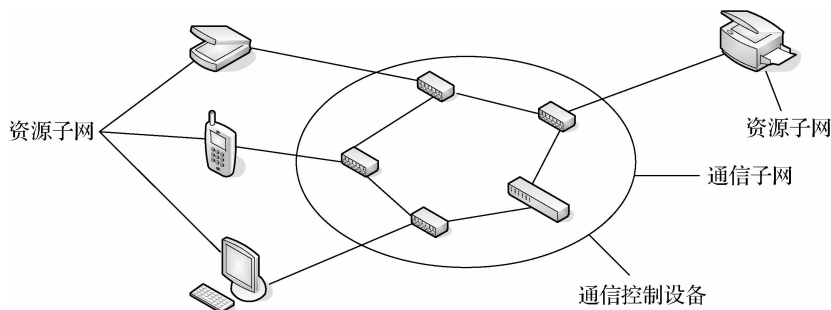


图 1-2 计算机网络的资源子网和通信子网

(1)资源子网。资源子网主要由服务器、工作站、共享设备(打印机、网络硬盘)、各种软件资源与信息资源组成。资源子网负责全网的数据处理,向网络用户提供各种网络资源与各种网络服务。

(2)通信子网。通信子网主要由网络适配器(网卡)、集线器、交换机、路由器、传输介质及相关软件组成,主要完成网络数据传输、转发等通信处理任务。

若用户只访问本地计算机,则只在资源子网内部进行,无须通过通信子网。若用户要访问异地计算机资源,则必须通过通信子网。

1.1.4 了解计算机网络的结构

网络拓扑结构是指用传输媒体互联各种设备的物理布局。将参与网络工作的各种设备用媒体互联在一起有多种方法,但实际上只有几种方式能适合网络的工作。

如果一个网络只连接几台设备,最简单的方法是将它们都直接连在一起,这种连接称为点对点连接。用这种方式形成的网络称为全互连网络,如图 1-3 所示。

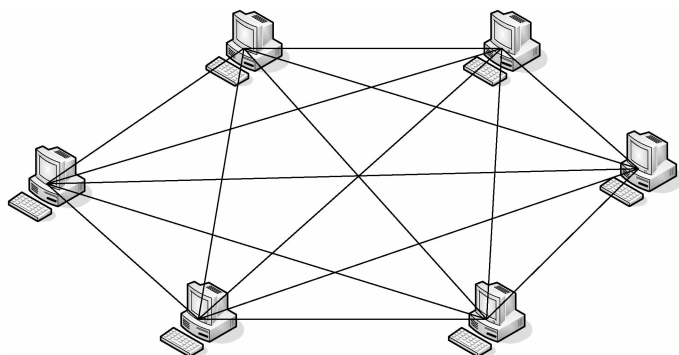


图 1-3 全互连网络拓扑图

图中有 6 个设备,在全互连情况下,需要 15 条传输线路。如果要连接的设备有 n 个,则所需线路将达到 $n(n-1)/2$ 条。显而易见,这种方式只有在涉及地理范围不大、设备数很少的条件下才有使用的可能。在其他情况下使用则会花费大量人力和物力。常见的拓扑结构是当需要通过互联设备(如路由器)互联多个网络时,将有可能遇到的互联技术。目前大多数网络使用的拓扑结构有星型拓扑结构、环型拓扑结构和总线型拓扑结构 3 种。下面分别介绍这三种拓扑结构。

1. 星型拓扑结构

星型拓扑结构是最古老的一种连接方式,大家每天都使用的电话就属于这种结构,如图 1-4 所示。其中,图 1-4(a)所示为电话网的星型结构,图 1-4(b)所示为目前使用最普遍的以太网(Ethernet)星型结构,处于中心位置的网络设备称为集线器,英文名称为 Hub。

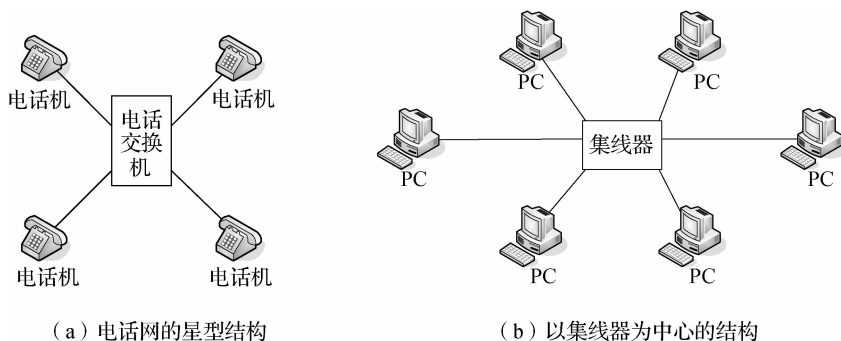


图 1-4 星型拓扑结构

这种结构是由中央节点和通过点到点通信链路连接到中央节点的各个站点组成的,中



动画
星型拓扑结构

央节点控制全网的通信,其中任何两个节点之间的通信都必须通过中央节点。星型结构的中央节点一般是交换机或集线器。中央节点执行集中式通信控制策略,因此,中央节点相当复杂,而各个站点的通信负担都比较小。

(1)星型拓扑结构的优点。

①控制简单。任何一个站点只和中央节点相连,因此介质访问控制很简单,访问协议也非常简单。

②故障诊断和隔离容易。中央节点可以对线路进行逐一隔离来进行故障检测和定位,单个连接点故障不影响整个网络。

③配置方便。中央节点可以方便为各个站点提供服务,或者重新配置网络。

(2)星型拓扑结构的缺点。

①电缆长度和安装费用高。因为每个站点直接连接到中央节点,所以这种拓扑结构需要大量电缆。电缆维护、安装等会产生高额费用。

②扩展困难。如果要增加新的站点,就要增加到中央节点的连接。

③过于依赖中央节点。若中央节点产生故障,则全网不能工作,所以对中央节点设备的可靠性和冗余度要求非常高。

2. 环型拓扑结构

环型结构在网络中使用较多。这种结构中的传输媒体从一个站点到另一个站点,直到将所有站点连成环型,如图 1-5 所示。这种结构显然消除了站点通信时对中心系统的依赖性。

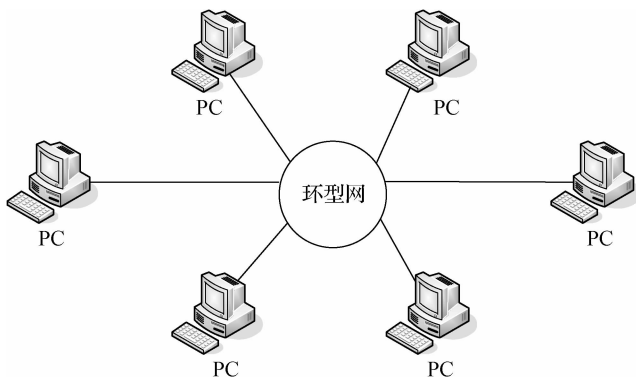


图 1-5 环型拓扑结构

环型拓扑结构的优点如下。

(1)电缆长度短。因为所有的站点都连接到一个公共数据通路,因此,只需要很短的电缆,减少了安装费用,易于布线和维护。

(2)可用光纤。光纤的传输速度快,环形拓扑是单方向传输,光纤传输介质十分适用。

3. 总线型拓扑结构

总线型结构是使用同一媒体或电缆连接所有站点的一种方式,也就是说,连接站点的物理媒体由所有设备共享,如图 1-6 所示。使用这种结构必须解决的一个问题是确保站点使用媒体发送数据时不能出现冲突。在点到点链路配置时,这是相当简单的。如果这条链路

是半双工操作,只需使用很简单的机制便可保证两个站点轮流工作。在一点到多点方式中,对线路的访问依靠控制端的探询来确定。然而,在网络环境下,由于所有数据站都是平等的,因而不能采取上述机制。对此,人们研究出一种在总线共享型网络中使用的媒体访问方法:带有碰撞检测的载波侦听多路访问,英文缩写为 CSMA/CD。

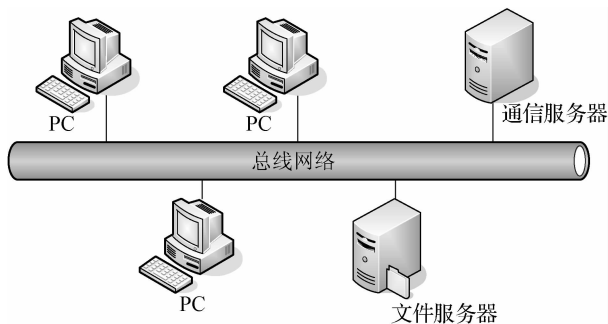


图 1-6 总线型拓扑结构



动画
总线型拓扑
结构

(1) 总线型拓扑结构的优点。

① 电缆长度短,容易布线。总线型拓扑结构和环型拓扑结构相似,所用电缆比星型拓扑结构要短得多。

② 可靠性高。总线的结构简单,又是无源元件,从硬件的观点看,十分可靠。

③ 易于扩充。当需要增加新的站点时,只需要在总线的任何节点处接入,如需要增加长度,可通过中继器扩展。

(2) 总线型拓扑结构的缺点。

① 所有的数据都需经过总线传送,总线成为整个网络的“瓶颈”。

② 总线的传输距离有限,通信范围受影响。

在实际网络设置过程中,经常需要把几种拓扑结构综合在一起运用。

1.1.5 熟悉计算机网络的分类

计算机网络可按不同的标准进行分类。

(1) 从网络节点分布来看,计算机网络可分为局域网(local area network, LAN)、广域网(wide area network, WAN)和城域网(metropolitan area network, MAN)。

① 局域网是一种在小范围内实现的计算机网络,一般在一个建筑物内,或一个工厂、一个事业单位内部,为单位独有。局域网距离可在十几千米以内,信道传输速率可达 1~20 Mbps,结构简单,布线容易。

② 广域网范围很广,可以分布在一个省内、一个国家或几个国家。广域网信道传输速率较低,一般小于 0.1 Mbps,结构比较复杂。

③ 城域网是在一个城市内部组建的计算机信息网络,提供全市的信息服务。目前,我国许多城市正在建设城域网。

(2) 计算机网络按交换方式可分为线路交换网络(circuit switching)、报文交换网络(message switching)和分组交换网络(packet switching)。

① 线路交换最早出现在电话系统中,早期的计算机网络就是采用此方式来传输数据的,

数字信号经过变换成为模拟信号后才能在线路上传输。

②报文交换是一种数字化网络。当通信开始时,源机发出的一个报文被存储在交换器里,交换器根据报文的地址选择合适的路径发送报文,这种方式称为存储—转发方式。

③分组交换也采用报文传输,但它不是以不定长的报文作为传输的基本单位,而是将一个长的报文划分为许多定长的报文分组,以分组作为传输的基本单位。这不仅大大简化了对计算机存储器的管理,而且也加速了信息在网络中的传播速度。由于分组交换优于线路交换和报文交换,具有许多优点,因此它已成为计算机网络的主流。

(3)计算机网络按服务方式可分为客户机/服务器网络、对等网。

①客户机/服务器网络就是一种基于服务器的网络,基于服务器的网络提供了更好的运行性能,并且可靠性也有所提高。在基于服务器的网络中,不必将工作站计算机的硬盘与人共享。当需要共享时,就必须把文件复制到服务器的硬盘上,这样别人才可以访问。它的典型应用是数据库的应用。

②对等网是采用分散管理的方式,网络中的每台计算机既可作为客户机,又可作为服务器来工作,每个用户要管理自己机器上的资源。

(4)计算机网络按传输介质可分为同轴电缆网络、双绞线网络、卫星网络和无线网络。网络传输介质就是通信线路。目前常用同轴电缆、双绞线、光纤、卫星、微波等有线或无线传输介质,相应的网络就分别称为同轴电缆网、双绞线网、卫星网、无线网等。

(5)计算机网络按网络拓扑结构可分为星型网络、总线型网络、环型网络。

1.1.6 网络图绘制工具 Visio 简介

绘制网络拓扑结构是学习网络时必备的一项重要内容,拓扑图的绘制对于对网络的理解及自身的动手能力的提升都非常重要。有很多种软件以及方法可以用来绘制网络拓扑图,而在实际应用中,Microsoft Office Visio 是一个不错的选择。该软件易学、易懂、实用,使用方便,是一款对网络设计人员非常实用的工具。

Visio 公司的创始人是来自阿图斯公司的几个程序员。他们于 1990 年在西雅图成立了 Shapeware 公司(1995 年,公司更名为 Visio),并于 1992 年发布了 Visio 1.0,该程序立即取得了成功,并相继开发了 2.0、3.0、4.0 和 5.0。1999 年,微软公司并购了 Visio 公司,差不多在同一时间发布了 Visio 2000 软件产品,该产品与微软紧密结合,支持多种微软件技术。它和 Microsoft Word、Microsoft Excel 等系列产品很相像。

Microsoft Office Visio 2010 可以创建具有专业外观的图表,以便理解、记录和分析信息、数据、系统和过程。便于 IT 和商务专业人员就复杂信息、系统和流程进行可视化处理、分析和交流。使用具有专业外观的 Microsoft Office Visio 2010 图表,可以促进对系统和流程的了解,深入了解复杂信息并利用这些知识做出更好的业务决策。

安装和激活 Microsoft Office Visio 2010 非常简单。这里我们以安装 Visio 简化版为例进行讲解。

(1)将 Visio 2010 安装光盘插入光驱中。

(2)运行光盘,选中“我接受此协议的条款”复选框,然后单击“继续”按钮,如图 1-7 所示。

(3)在弹出的“选择所需的安装”界面中,用户可自主选择“立即安装”或“自定义”安装,

如图 1-8 所示。

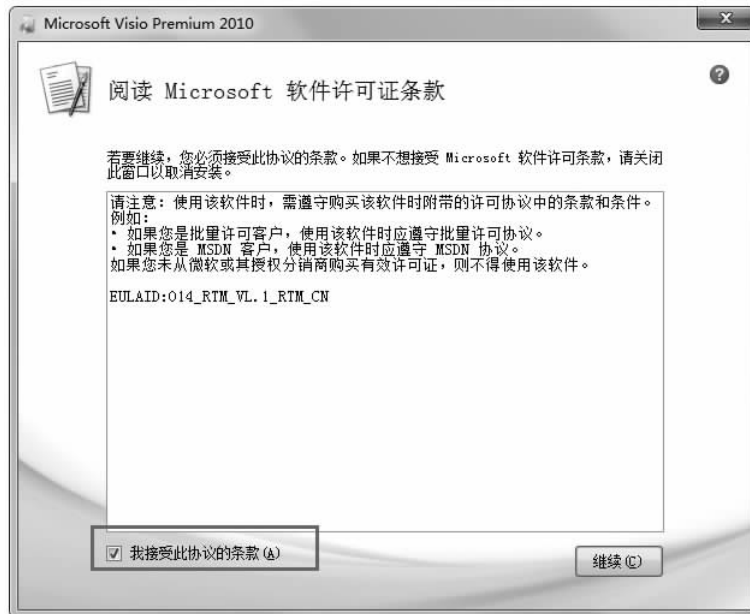


图 1-7 软件许可证条款界面



图 1-8 自主选择安装界面

(4) 这里单击“立即安装”按钮, 软件将开始安装进程, 并进入“安装进度”界面, 如图 1-9 所示。



图 1-9 “安装进度”界面

(5)安装完成后单击“关闭”按钮。

1.2 扩展知识

1.2.1 计算机网络热点问题

随着网络的发展,近年的网络有许多热点,如物联网、云计算、VR/AR 等,下面就来介绍这些新兴技术热点。

1. 物联网

1)物联网技术的基本定义

物联网是新一代信息技术的重要组成部分,其英文名称是 the Internet of things。顾名思义,物联网就是物物相连的互联网。这里有两层意思:第一,物联网的核心和基础仍然是互联网,是在互联网基础上的延伸和扩展的网络;第二,其用户端延伸和扩展到了任何物品与物品之间,进行信息交换和通信。因此,物联网的定义是通过射频识别(RFID)、红外感应器、全球定位系统、激光扫描器等信息传感设备,按约定的协议把任何物品与互联网相连接,进行信息交换和通信,以实现物品的智能化识别、定位、跟踪、监控和管理的一种网络,如图 1-10 所示。

(7)在世界网络中有可被识别的唯一编号。

4)物联网的应用

物联网用途广泛,遍及智能交通、环境保护、政府工作、公共安全、平安家居、智能消防、工业监测、环境监测、老人护理、个人健康、花卉栽培、水系监测、食品溯源、敌情侦查和情报收集等多个领域。

在这里通过一个简单的路灯智能管理系统来介绍物联网在实际工作中的应用。当路灯突然出现故障不能亮启时,传感器接收到电流的异常变化,这时它就发出一个异常变化的信号,并连同自己的网络地址通过网络传输物联网平台。

在平台中,平台判断异常变化的情况给出结论显现出来,这时操作员根据给出的结论来对路灯进行处理,如图 1-11 所示。

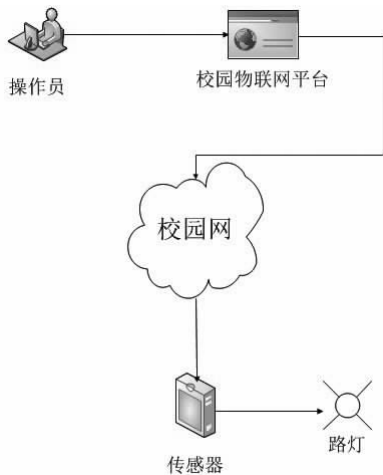


图 1-11 智能路灯工作图

登录平台之后,进入“校园管理”→“路灯管理”界面,如图 1-12 所示。

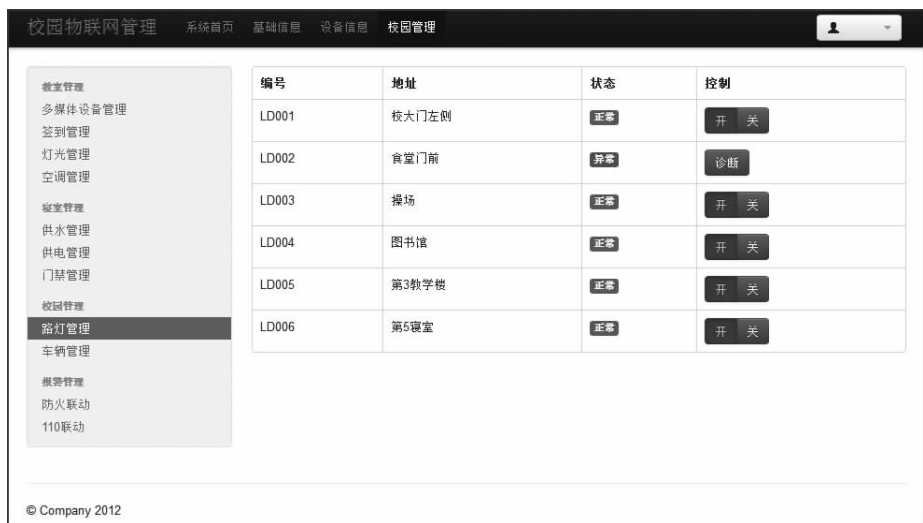


图 1-12 “路灯管理”界面

校园网内所有路灯都显示在页面:第一列是路灯编号;第二列是路灯的地址,便于发现问题之后找到问题路灯,进行维护;第三列是当前路灯状态,表示正常工作,而则表示路灯需要维护;第四列可以对正常的路灯进行开关控制,而异常的路灯则给出诊断按钮,用于诊断异常信息。

这个案例只是一个最简单的物联网技术的应用,在物联网时代的人类生活的图景如下:当司机出现操作失误时汽车会自动报警,公文包会提醒主人忘带了什么东西,衣服会“告诉”洗衣机对颜色和水温的要求等。物联网在物流领域内的应用也会广泛。例如,一家物流公司应用了物联网系统的货车,当装载超重时,汽车会自动告诉相关人员超载了,并且超载多少,但空间还有剩余,告诉相关人员轻重货怎样搭配;当搬运人员卸货时,一只货物包装可能会大叫“你扔疼我了”,或者说“亲爱的,请你不要太野蛮,可以吗”;当司机在和别人扯闲话时,货车会装作老板的声音怒吼:“笨蛋,该发车了!”

物联网把新一代 IT 技术充分运用在各行各业之中,具体地说,就是把感应器嵌入和装备到电网、铁路、桥梁、隧道、公路、建筑、供水系统、大坝、油气管道等各种物体中,然后将物联网与现有的互联网整合起来,实现人类社会与物理系统的整合。在这个整合的网络当中,存在能力超级强大的中心计算机群,能够对整合网络内的人员、机器、设备和基础设施实施实时的管理和控制。在此基础上,人类可以以更加精细和动态的方式管理生产与生活,达到“智慧”状态,提高资源利用率和生产力水平,改善人与自然间的关系。

毫无疑问,如果物联网时代来临,人们的日常生活将发生翻天覆地的变化。然而,不谈什么隐私权和辐射问题,单就把所有物品都植入识别芯片这一点来说,现在看来还不太现实。人们正走向物联网时代,但这个过程可能需要很长的时间。

2. 云计算

1) 云计算的基本定义

云计算(cloud computing)是基于互联网的相关服务的增加、使用和交付模式,通常涉及通过互联网来提供动态易扩展且经常是虚拟化的资源。云是网络、互联网的一种比喻说法。过去在网络图中往往用云来表示电信网,后来也用它来表示互联网和底层基础设施的抽象。

狭义的云计算是指 IT 基础设施的交付和使用模式,通过网络以按需、易扩展的方式获得所需资源;广义的云计算是指服务的交付和使用模式,通过网络以按需、易扩展的方式获得所需服务。这种服务可以是 IT 和软件、互联网相关,也可是其他服务。它意味着计算能力也可作为一种商品通过互联网进行流通。

2) 云计算的基本特点

互联网上的云计算服务特征和自然界的云、水循环具有一定的相似性,因此,云是一个相当贴切的比喻。根据美国国家标准和技术研究院的定义,云计算服务应该具备以下几个特征。

- (1) 按需分配的自助服务。
- (2) 宽带网络访问。
- (3) 资源池化。
- (4) 快速弹性。
- (5) 可评测的服务。

一般认为,云计算还有如下特征。

- (1)基于虚拟化技术快速部署资源或获得服务。
- (2)减少用户终端的处理负担。
- (3)降低了用户对于 IT 专业知识的依赖。

3)云计算的应用

云计算的应用主要有以下几点。

(1)云物联。随着物联网业务量的增加,对数据存储和计算量的需求将带来对“云计算”能力的要求。

(2)云安全。云安全(cloud security)是一个从“云计算”演变而来的新名词。云安全的策略构想是:使用者越多,每个使用者就越安全,因为如此庞大的用户群,足以覆盖互联网的每个角落,只要某个网站被“挂马”或某个新木马病毒出现,就会立刻被截获。“云安全”通过网状的大量客户端对网络中软件行为的异常监测,获取互联网中木马、恶意程序的最新信息,推送到服务器端进行自动分析和处理,再把病毒和木马的解决方案分发到每一个客户端。

(3)云存储。云存储是在云计算概念上延伸和发展出来的一个新的概念,是指通过集群应用、网格技术或分布式文件系统等功能,将网络中大量各种不同类型的存储设备通过应用软件集合起来协同工作,共同对外提供数据存储和业务访问功能的一个系统。当云计算系统运算和处理的核心是大量数据的存储和管理时,云计算系统中就需要配置大量的存储设备,那么云计算系统就转变成为一个云存储系统,所以云存储是一个以数据存储和管理为核心的云计算系统。

4)云计算的主要问题

云技术要求大量用户参与,也不可避免地出现了隐私问题。用户参与即要收集某些用户数据,从而引发了对用户数据安全的担心。很多用户担心自己的隐私会被云技术收集。正因如此,在加入云计划时很多厂商都承诺尽量避免收集到用户隐私,即使收集到也不会泄露或使用。但不少人还是怀疑厂商的承诺,他们的怀疑也不是没有道理的,不少知名厂商都被指责有可能泄露用户隐私,并且泄露事件也确实时有发生,例如,2011年包括 CSDN 在内的多家互联网公司数据库泄密造成大量用户数据泄密。

3. VR/AR

1)虚拟实境

虚拟实境(virtual reality,VR)简称虚拟技术,也称虚拟环境,是利用计算机模拟产生一个三维空间的虚拟世界,提供用户关于视觉等感官的模拟,让用户感觉仿佛身历其境,可以及时、没有限制地观察三维空间内的事物。用户进行位置移动时,计算机可以立即进行复杂的运算,将精确的三维世界视频传回产生临场感。该技术集成了计算机图形、计算机仿真、人工智能、感应、显示及网络并行处理等技术的最新发展成果,是一种由计算机技术辅助生成的高技术模拟系统。

(1)虚拟实境的概念。从技术的角度来说,虚拟现实系统具有下面三个基本特征:三个“I”,即 immersion-interaction-imagination(沉浸-交互-构想),它强调了在虚拟系统中的人的主导作用。从过去人只能从计算机系统的外部去观测处理的结果,到人能够沉浸到计算机系统所创建的环境中,从过去人只能通过键盘、鼠标与计算环境中的单维数字信息发生作用,到人能够用多种传感器与多维信息的环境发生交互作用;从过去的人只能以定量计算为主的结果中启发从而加深对事物的认识,到人有可能从定性和定量综合集成的环境中得到

感知和理性的认识从而深化概念和萌发新意。总之,在未来的虚拟系统中,人们的目的是使这个由计算机及其他传感器所组成的信息处理系统去尽量“满足”人的需要,而不是强迫人去“凑合”那些不是很亲切的计算机系统。

现在的大部分虚拟现实技术都是视觉体验,一般是通过计算机屏幕、特殊显示设备或立体显示设备获得的,不过一些仿真中还包含了其他感觉处理,如从音响和耳机中获得声音效果。在一些高级的触觉系统中还包含了触觉信息,也称为力反馈,在医学和游戏领域有这样的应用。人们与虚拟环境交互,要么通过使用标准装置,如一套键盘与鼠标,要么通过仿真装置,如一只有线手套,要么通过情景手臂和/或全方位踏车。虚拟环境是可以和现实世界类似的,如飞行仿真和作战训练,也可以和现实世界有明显差异,如虚拟现实游戏等。就目前的实际情况来说,它还很难形成一个高逼真的虚拟现实环境,这主要是技术上的限制造成的,这些限制来自计算机处理能力、图像分辨率和通信带宽。然而,随着时间的推移,处理器、图像和数据通信技术变得更加强大,并具有成本效益,这些限制将最终被克服。

(2)相关设备。一般的虚拟现实设备至少包含一个屏幕、一组感测器及一组计算组件,这些东西被组装在这个设备中,图 1-13 所示为 Google 的原始 VR 眼镜。屏幕用来显示仿真的视频,投射在用户的视网膜上,感测器用来感知用户的旋转角度,计算组件收集感测器的数据,决定屏幕显示的画面。



图 1-13 Google 的原始 VR 眼镜

额外的设备可能包括一台高级计算机,用以补充计算组件的不足,也可能有一对把手及定位器,用以侦测用户的位置。

对屏幕最重要的要求是反应时间,目前计算机、手机等设备所使用的屏幕多为 TFT-LCD,反应时间太长难以满足虚拟现实的要求,这也是行动虚拟现实最大的挑战。当前所知效果最好的屏幕为 OLED,但全世界能够达到足够质量要求的公司寥寥无几。

(3)典型应用——淘宝 Buy+VR 购物。2016 年 4 月 1 日下午,淘宝推出全新购物方式 Buy+。Buy+使用虚拟现实技术,利用计算机图形系统和辅助传感器,生成可交互的三维购物环境。Buy+将突破时间和空间的限制,真正实现各地商场随便逛,各类商品随便试。该功能在 2016 年 11 月 1 日正式试运营。

Buy+通过 VR 技术可以 100%还原真实场景,也就是说,使用 Buy+,身在广州的家中,戴上 VR 眼镜,进入 VR 版淘宝,可以选择去逛纽约第五大道,也可以选择到英国复古集市。

让你身临其境去购物,到全世界去买买买。例如,在选择一款沙发时,用户再也不用因为不太确定沙发的尺寸而纠结。戴上 VR 眼镜,直接将这款沙发放在家里,尺寸颜色是否合适,一目了然。

Buy+利用 TMC 三维动作捕捉技术捕捉消费者的动作并触发虚拟环境的反馈,最终实现虚拟现实中的互动。简单来说,你可以直接与虚拟世界中的人和物进行交互。甚至将现实生活中的场景虚拟化,成为一个可以互动的商品,如图 1-14 所示。



图 1-14 淘宝 Buy+VR 购物图

例如,利用带有动作捕捉的 VR 设备,你眼前的香蕉、书籍在 Buy+中可以化身为架子鼓,利用这种互动形式,让你在购买商品的过程中拥有更多体验。

除了以上,Buy+产品视频里还有一个有意思场景。例如,当你去给女朋友买内衣时,再也不用如此尴尬,可以直接查看内衣详情,甚至内衣上身效果。

Buy+能够大大加码线上商品的真实感,甚至通过虚拟技术能拥有实体店所没有的惊喜和体验。

2)增强现实

增强现实(augmented reality,AR)是一种实时地计算摄影机影像的位置及角度并加上相应图像的技术,这种技术的目标是在屏幕上把虚拟世界套在现实世界并进行互动。这种技术估计由 1990 年提出。随着随身电子产品运算能力的提升,预期增强现实的用途将会越来越广。

(1)相关设备。已上市的 AR 硬件包含光学投影系统、监视器、行动装置、头戴式显示器、抬头显示器、计算机。在开发中还有仿生隐形眼镜等。

头戴式显示器透过眼罩或头盔的形式,将显示屏幕贴近用户的眼睛,目前有以下几家公司推出了 AR 头戴式显示器:uSens、Gestigon、Meta、Google 眼镜。

目前在 iPhone 手机和 Google Android 手机上,已经出现不少的增强现实的应用。

(2)AR 的应用。

①百度地图的 AR 功能。打开百度地图,在搜索栏输入目的地,进入步行导航界面后,用户即可在原有的“跟我走”导引左侧看到新增的“AR 导航”。点击“AR 导航”,百度地图将自动打开用户手机中的相机功能,快速呈现全实景路线,如图 1-15 所示。



图 1-15 百度地图 AR 导航

在导航过程中,用户不必反复查看地图路线信息,只需结合语音导航、真实街景即可轻松找到方向,顺利抵达目的地。步行导航与 AR 技术的连接,完美融合了虚拟与现实,在起点、转向、目的地等场景中,步行结合 AR 第一视角进行 POI 导览、辅助决策、指引导航等服务,极大提高地图的读图效率,提升步行的趣味性和科技感。同时,用户可自由切换步行 AR 导航与普通导航两种模式,使用普通导航的用户只需点击页面右下角的缩略图,即可一键切换到 AR 导航模式。

此外,步行 AR 导航的界面极具丰富性和生动性,包含罗盘、路线、转向标、途经点和终点气泡、滚轴诱导提示、路线全览等元素,顶部的诱导提示与步行导航保持一致,辅助用户了解具体的路线距离,有效解决“路痴”们找不到方向和目的地的尴尬。

步行 AR 导航是百度地图利用 AR 技术将虚拟嫁接于现实之上,继 AR 游戏之后从娱乐性转向实用性的大胆尝试。除该功能外,百度地图 9.5.5 新版还发布最新跨城交通解决方案,提供智能、完整的一站式长途出行方案。在 2017 年 4 月举行的全球移动互联网大会上,百度地图事业部总经理李东旻曾表示,通过不断的技术创新让地图更简单好用,是百度地图的梦想,百度地图在过去 1 年中平均每 3 天就会上线 1 个创新功能,全年申请了超过 300 项专利。依托于大数据与人工智能领域的技术优势,百度地图让虚拟地图无限贴近真实世界,加速开启智能出行新时代,业已成为“基于大数据的人工智能出行平台”。

②QQ 红包雨。2017 年 1 月,QQ LBS+AR 红包正式推出,玩法和支付宝、Pokemon GO 都有些类似,也是发起者在指定地点发出红包,用户开启摄像头,根据定位前往领取,如图 1-16 所示。



图 1-16 QQ 红包雨

除了最基础的个人方式派发红包,QQ 还联合屈臣氏、百事可乐等线下的商家,以及 Angelababy 等明星在全国几百万个地理位置上任意放置一批红包,用户只要在以红包放置点为中心的半径 120 米范围之内点击红包开启摄像头,就可以参与到抢红包的活动。

4. 大数据

近年来,随着社交网络对人们生活的渗透,使人们从其中的数据中观察到更多的人类社会的复杂行为模式。大量的信息汇集、分析的第一手资料,产生了重要的数据资产。这些数据资产产生了巨大的经济价值与社会价值。人类历史迎来第四次革命,大数据的产生使得从前孤立的数据具有关联性,使得人们发现新的机遇,创造新的价值。

1) 大数据的定义

作为全球咨询行业的标杆,麦肯锡公司俨然成为大数据的研究先驱。在 2011 年麦肯锡的报告中给出的关于大数据的定义是:大数据(big data, mega data)或称巨量资料,指的是需要新处理模式才能具有更强的决策力、洞察力和流程优化能力的海量、高增长率和多样化的信息资产。大数据的“大”的界定范畴是动态的,从前的 GB 就是数据类的巨大范畴,但是大数据出现后,物理、基因等很多领域,TB 级的数据已普遍,更有 PB 级,甚至 EB 级。数据的类型有很多种,主要分为结构化数据、半结构化数据和非结构化数据 3 种。因此,数据量的不断增长及数据类型的多样化,都给大数据系统的存储和计算带来了不小的挑战。

2) 大数据的价值

大数据并不是简单的大量数据。大数据的 3 个明显的特征分别是数据量大,数据实时性强,具有数据多样性。

量大是大数据的必要条件,但并不充分,因为大数据的数据增长是爆发性的,因此计量单位一般是 PB、EB,甚至 ZB。

大数据具有多变性,它包括结构化数据也包括非结构化数据,与传统结构化的数据相比,呈多样性。人们必须运用大数据的技术从海量的非结构化数据中提炼、分析数据,因此非结构化的数据在数据储存、收集上更为复杂。

大数据的实时性代表着这些数据的分析与使用是随时的、实时的,与传统数据不同,动态性是大数据的显著特征。

传统的数据时代,商业决策的做出主要依靠历史数据与经验总结,无可避免地出现由于信息滞后造成的决策效果不佳,大数据时代,依据在线的、实时的数据收集与分析,实现精准营销,极大地提高决策实效性。

在大数据时代,个人计算机和手机移动端全面普及,使每个人随时随地都在提供数据。各种各样的行为,如清晨搭车、点击网上商品、刷卡购物、使用手机玩游戏,都会产生专属于每个人的数据痕迹,然后形成大数据被记录下来,每个人的年龄段、性别、消费偏好、喜欢的商品类型、出行习惯等信息都被记录成数据,商家可以提取有效的商业信息,根据客户的习惯和偏好,精准营销。

大数据使得每个人从中受益,生物领域的专家在对基因信息、遗传物质的等信息分析的基础上,结合每个人特有的健康数据、身体功能指标、既往病史、过敏史得出研究结果。医疗研发机构,根据互联网采集的病人数据基础,推进慢性病医疗方面的服务,探索慢性疾病的信息管理和新型的医疗方式,反过来,互联网借助医疗机构的治疗数据,构建起慢性疾病患者的大数据。

大数据的时代拥有更便捷的方式来甄选有效真实的数据。大数据的多样性使来自不同数据源、不同维度的数据相互之间产生一定程度的关联性,这种关联性可以通过多种方式交互验证。例如,某厂的生产产量少报一半,目的少报税,但是它的生产电力和能耗却超过对应的指标一半,这种虚假数据就能及时被大数据系统甄别,大数据能根据各种关联性的明细数据,综合判断出企业真实的盈利能力,并能形成成熟的数据信息,生成更多更有价值的

数据。

数据作为现代社会的资源之一,不同于物质性的资源,那些资源缺乏可再生性,无法共享,但是数据资源却能反复使用,并且产生不同的价值。这种良性的资源使用,使得大数据贡献,能产生出多赢的局面。

大数据因其背后的价值,被比喻为新世纪的黄金,被看作新兴起的经济元素,大数据不仅本身可以看作重要的生产要素,其对产品的形成过程也起到至关重要的作用。大数据的主要价值如下。

(1)大数据是新时代信息技术的关键支撑。大数据的热潮在全球的盛行,顺应了现代信息技术发展的趋势。互联网时代为大数据的普及和发展打下了坚实的基础,人们能随时通过移动端使用互联网,伴随着物联网、网上购物、交友网站、云计算的兴起,每个人的数据无处不在,随时随地产生。作为信息技术时代的产物,大数据的应用又反作用于信息技术的发展,促进物联网、云计算等技术的革新,大数据作为融合现代信息技术的关键支撑,为物联网、云计算等现代信息技术的发展提供了依托的平台。

(2)大数据是促进现代社会经济发展的推动力。大数据本身隐含着巨大的经济价值和

社会价值。大数据行业的爆发式发展,将带来一批针对大数据市场的新的商业理念、新的营销服务、新的产品和新的技术,推动现代信息产业的加速发展。在国内的城市建设,民生发展等领域,大数据也起着举足轻重的作用。目前我国着力推行智慧城市的建设、大数据的应用能将城市中的方方面面数据联合起来,分析提取有效数据,依靠它们做出智慧决策。在红绿灯的时间设置上,可以依照不同的时间段、该条道路的车行流量、拥堵状况数据分析,来合理设置红绿灯的时间,缓解交通难题。伴随着智慧城市在我国不断建设完善,大数据的出色表现在提升地方政务能力和社会管理能力方面发挥着重要作用,使之形成充满各地特色的、新兴的智能领域应用。

大数据帮助企业深度挖掘客户喜好,助力企业智能决策。大数据对企业洞察用户提供了有力的武器,满足企业针对客户的个性化营销需求,为企业做出正确的市场决策提供更多维度。大数据出现以前,市场调查是通过人工获取方式,采用调研和营销实现,这样的数据具有明显的“人工计划”特征,在市场调查之前,收集数据的样板、调研方式、分析方式和获取数据的目的都有一个清晰的规划,因此,这些产生的数据是“结构化”的。依靠互联网产生的大数据,其来源是互联网用户行为,包括网页检索、页面浏览、网络交易和网络社交行为等,它并不受人工计划,因此数据的产生、分析过程具有不确定性,这样的大数据是多维度的,360°全方位接近用户,让决策的依据更科学。

(3)大数据将成为科技创新的引擎。人工数据时代,信息化的滞后使得大量的数据彼此分离,闲置在各自的系统空间,技术的落后让传统的信息处理方式无法满足科技发展的需求。新兴的大数据在整合数据、分析数据、储存数据、处理数据、应用数据,解决系统实时性的、并发性的问题,包括云储存,数据价值分析等方面都颠覆了传统。大数据成为各个领域科技创新的引擎。例如,大型家电生产厂家,在产品生产线安装上传感器采集数据,这些生产信息的分析和价值挖掘,能实时提高产品合格率。在电力领域,智能电表的数据采集同样发挥着不可忽视的作用。不仅作为电费收取的依据,还扮演着判断房屋空置与否的角色,延伸开来,可作为城市房价定位的参考指标。再者,电网所采集的耗电量数据可以判断出该部分地区的商业发展情况。在未来,不论是国家政府,还是金融商业,各个数据集中的领域,大数据将成为各企业、各单位提升竞争力、占领市场的核心,加速企业从“业务驱动”向“数据驱动”转型升级,为企业提高利润、增强实力、研发产品带来新的机遇。

3)大数据的特点

如图 1-17 所示,大数据具有四大特点:volume(容量),代表海量的数据规模;variety(种类),代表数据类型的多样性;value(价值),代表深度的数据价值;velocity(速度),代表数据流转的迅速与体系的动态性。



图 1-17 大数据的特点

(1) volume(数据体量巨大)。目前,人类社会所生产的印刷材料总和的数据量是200 PB(1 PB=1 024 TB),人类说过的语言的总和数据量大约是5 EB(1 EB=1 024 PB)。数据的体量决定了它背后的信息价值,随着各种移动端的流行,云存储技术的发展,现代社会的人类活动都可以被记录下来,因此产生了海量的数据。发送的微博,自拍的照片,穿戴的运动手环等,通过互联网上传到云端,各种数据聚集到特定地点的存储系统,如政府等机构,最终形成了体量巨大的大数据。

(2) variety(数据类型呈多样性)。数据类型的多样性使得数据被分为结构化数据与非结构化数据两种。而互联网将网络通过各种移动端形成了整体,人们不仅可以通过互联网获取数据,也是数据的传播者,相对于过去,以文本为主的结构化数据往往是便于存储的,随着非结构化的数据越来越多,如网络小说、拍摄的视频、录制的音频、共享的地理位置等,这些多样性的数据使得对数据处理的能力要求更高,需要对数据进行加工、清洗、分析等步骤,将它们变为易于存储的结构化数据。这需要在海量的数据之间发现它们之间的关联性,把看似毫无关系的数据联系起来,形成有价值的信息。

(3) velocity(处理迅速)。这是大数据区别于传统数据挖掘的最显著特征,大数据具有实时性。例如,人们出去吃饭,导航餐厅,用移动端的地图查询位置,选择避免堵车的路线,还会从网络上查看餐厅的口味评价如何,吃饭后,也许会拍下食物和餐厅的照片上传到微博,因此,各种网络的链接会带来大量的数据交换,对速度的要求更高,要以实时的方式传达给用户。

(4) value(数据价值)。大数据的应用,在物联网、云计算、大数据挖掘等技术的迅速发展地带动下,呈现出它的过程:把数据源的信号转换为数据,再把大数据加工成信息,通过获取的信息来做决策。因此,大数据价值的挖掘过程就像大浪淘沙,数据的体量越大,相对有价值的信息就越少。

大数据的价值密度实际是比较低的,因为数据采集并非都是及时的,样本的数量有限,数据不完全连续,但是,当数据的体量越来越大,就能从海量数据中提取有价值的信息,为决策做支撑。

4) 大数据经典案例

在大数据时代,如何通过大数据的应用来实现企业价值,是很多企业思考的问题,下面介绍一些利用大数据创造价值的成功案例。

(1) 啤酒与尿布。全球零售业巨头沃尔玛在对消费者购物行为分析时发现,男性客户在购买婴儿尿布时,常常会顺便搭配几瓶啤酒来犒劳自己,于是尝试推出了将啤酒和尿布摆在一起的促销手段。没想到这个举措居然使尿布和啤酒的销量都大幅增加了。如今,“啤酒+尿布”的数据分析成果早已成了大数据技术应用的经典案例,被人津津乐道。

(2) 数据新闻让英国撤军。2010年10月23日,《卫报》利用维基解密的数据做了一篇“数据新闻”。将伊拉克战争中所有的人员伤亡情况均标注于地图之上。地图上一个红点便代表一次死伤事件,点击红点后弹出的窗口则有详细的说明:伤亡人数、时间,造成伤亡的具体原因。密布的红点多达39万,显得格外触目惊心。一经刊出立即引起朝野震动,推动英国最终做出撤出驻伊拉克军队的决定。

(3) Google成功预测冬季流感。2009年,Google通过分析5 000万条美国人最频繁检索的词汇,将之和美国疾病中心在2003年到2008年间季节性流感传播时期的数据进行比

较,并建立一个特定的数学模型。最终,Google 成功预测了 2009 年冬季流感的传播,甚至可以具体到特定的地区和州。

(4)大数据与乔布斯癌症治疗。乔布斯是世界上第一个对自身所有 DNA 和肿瘤 DNA 进行排序的人。为此,他支付了高达几十万美元的费用。他得到的不是样本,而是包括整个基因的数据文档。医生按照所有基因按需下药,最终这种方式帮助乔布斯延长了好几年的生命。

(5)奥巴马大选连任成功。2012 年 11 月奥巴马大选连任成功的胜利果实也被归功于大数据,因为他的竞选团队进行了大规模的深入的数据挖掘。时代杂志更是断言,依靠直觉与经验进行决策的优势急剧下降,在政治领域,大数据的时代已经到来;各种媒体、论坛、专家铺天盖地的宣传让人们对于大数据时代的来临兴奋不已,无数公司和创业者都纷纷跳进了这支狂欢队伍。

(6)微软大数据成功预测奥斯卡 21 项大奖。2013 年,微软纽约研究院的经济学家大卫·罗斯柴尔德(David Rothschild)利用大数据成功预测 24 个奥斯卡奖项中的 19 个,被人们津津乐道。后来,罗斯柴尔德再接再厉,成功预测第 86 届奥斯卡金像奖颁奖典礼 24 个奖项中的 21 个,继续向人们展示现代科技的神奇魔力。

5) 大数据的处理过程

大数据的处理过程包含大数据的采集、大数据的导入和预处理、大数据的统计和分析、大数据的挖掘 4 个步骤。

(1)大数据的采集。在“大数据时代”的今天,数据似乎被提到一个前所未有的高度。无论是小企业还是大公司,网上销售还是线下营销都意识到数据的重要性。随着大数据越来越被重视,数据采集的挑战也变得尤为突出。

很多人不清楚需要收集什么样的数据;也有的不清楚通过什么渠道来收集数据;还有大部分不清楚收集整理的数据如何去分析,进而也就不清楚怎么去利用这些数据。所以,很多数据也就仅仅只是数字,无法去转化和为公司利益服务,成了摆设。

下面介绍 3 类将数据做成摆设的类型。

①重视数据但不清楚如何收集,这是“被数据”类型。表现为对数据处于模糊了解状态,明确公司和企业做事与计划要靠数据来支撑。由于缺乏专业的相关数据人员,公司该做哪些数据,通过什么渠道来收集整理,处于一知半解的状态,通过网上学习东拼西凑而成的数据,这样的数据自然就真的只是摆设了。

②了解所需数据但来源不规范,这是“误数据”类型。表现为对数据比较了解,大概明确需要什么数据。由于缺乏专业的数据人员,对于数据的来源和制作并不规范,数据采集也可能存在误差。因此,采集的数据就可能失真,数据价值较小。

③会做数据但不会解读分析,这是“低估数据”类型。表现为对数据清楚了解,并有准确的数据来源和较明确的数据需求,但是坐拥金矿却不会利用,只是简单的收集整理,把数据形成可视化的报表,这种简单化的采集方式使得数据的价值被低估。

数据背后的意义是什么?怎样去解读数据来为公司和个人创造价值?去利用数据来规避可能存在的风险?这些才是数据采集的意义。

数据的采集系统是基于计算机或测试平台的测量系统。常见的采集工具有很多,如麦克风、摄像头等,数据的采集技术应用广泛。

大数据的采集一般分为以下两个层次。

①大数据智能感知层。包括传感适配体系、网络通信系统、智能识别体系、数据传感体系和软硬件资源接入系统,用来完成对不同类型的数据结构的智能识别、清洗、接入、信号转换、监控、处理和管理等。

②大数据基础支撑层。一种虚拟的服务器,是大数据服务平台所必需的,提供包含各种类型数据结构的数据库和物联网等支撑环境。

在大数据的采集过程中,现存难点是并发数高,也许存在无数的用户在同一时间访问同一个页面的情况,在并发数高峰期,访问量峰值高达每分钟百万次,必须在数据库之间进行负载均衡与分片,同时在采集端衔接大量数据库才能支撑。

(2)大数据的导入和预处理。要实现对海量数据的有效分析,需要将数据岛屿到集中的分布式数据库,或者导入到分布式存储集群,之后需要对数据库进行简单的预处理和清洗。如果对业务有实时需求,可以在导入时使用 Storm 对数据进行流式计算。

(3)大数据的统计和分析。随着技术的更新,大数据分析越来越多地在医疗、建设智慧城市等方面发挥了积极的作用。在商业应用方面,不少企业对大数据分析的需求上升,迫切需要引进专业的数据分析人员,或者与大数据分析服务机构合作,以挖掘数据价值,为企业科学的运营决策做支撑。

运用好大数据的统计与分析技术,能协助企业精准定位客户喜好、优化资源配置、定制营销。目前,在发达国家,大数据分析行业已进入蓬勃发展期,专业的数据分析机构和数据分析人员的规模也不断扩大,大数据分析广泛应用于发达国家的各个行业,如电商、金融、零售、通信等领域。

大数据的统计与分析主要是利用分布式计算集群,或者分布式数据库来对数据进行分类和汇总。在企业的实时性需求方面,可以用 Oracle 的 Exadata、EMC 的 GreenPlum、基于 MySQL 的列式存储 Infobright 等。对于批处理或者半结构化数据的需求,则可以使用 Hadoop。

(4)大数据的挖掘。人们需要从海量的数据中发现有用的数据价值,进而将数据价值转化为决策依据,这需要一些合适的工具来进行这项工作,因此产生了大数据的挖掘。数据挖掘的是一个新生的、动态的领域,是人们从数据时代迈入信息时代必不可少的步骤。

人们每天都在百度搜索引擎进行查询,每天可达数亿次查询,假如人们的查询都被看作一个任务,人们通过关键词描述任务需求,那么,日积月累,搜索引擎能从海量的查询中学到什么?这里有一个发现,在海量的查询中,有些查询模式能呈现出大量的知识,而这些知识却不能通过仅仅读取单个人的查询数据得到。例如,百度的飞行时间查询,使用这个搜索项作为航班飞行活动的指示,它呈现出搜索飞行时间相关信息的人数与正在候机的人数之间的联系。当与飞行时间相关的搜索都汇总在一起,即产生了一种模式。使用这种汇聚的搜索数据,百度的飞行时间能比传统的系统早几个小时对航班准点率做出评估。这样的实例表示,大数据的挖掘能把数据集转换成信息,帮助人们得到找寻的答案。与统计和分析过程相区别的是,大数据的挖掘通常没有预先设定的主题,而是在现有数据的基础上进行计算,来实现预测的结果,用以满足高级别的分析需求。常见的算法有 Kmeans(用于聚类)、SVM(用于统计)、NaiveBayes(用于分类)等。大数据的挖掘通常使用的算法以单线程为主,因其计算的数据量大。

6) 大数据技术的特征

大数据技术具有以下几个特征。

(1) 数据进行全面分析。大数据技术的数据分析是全面的,而不是随机抽样进行的。在大数据技术之前,因缺乏对全体样本进行抽取的技术,对待样本的抽取方式,都是从小样本进行随机抽取。在理论上曾认为,随机抽取的样本,能代表整体样本的多样性,但这种方法费力且费时。在大数据出现后,在云计算和数据库的基础上,大数据技术能获得足够大的样本,并能储存到数据库中。所有的数据都存储于数据仓库里,因此不需要以随机抽样的方法对数据进行分析。获取大数据本身并不是人们最终的目的,如果能用小数据解决人们的疑惑,就不需要使用大数据进行分析。牛顿发现的力学定律、人类发现的行星定律等都是通过小数据分析的。典型的,人脑就是通过小数据学习来获取知识。

(2) 强化数据的复杂性。对于小数据来说,收集的样本是有限的,因此需要尽可能使保存的数据精准。例如,采用抽样方法时,要求在运算时精准,在1万只羊中采取随机抽取的方式,抽取100只羊,如果在100只羊的样本上计算有误,放大至1万只羊,偏差就会扩大。

相对于小数据,最重要的是减少差错以保证质量,而大数据更重视数据的复杂性。

在小数据的情况下,为了避免放大时造成的偏差,要求得到样本的精准计算结果,但需要耗费很多的时间,在大数据的情况下,样本等于总体,能迅速获得总体的特点和趋势,这比精准性更为重要。

大数据的算法简单,但比小数据有效,因此对大数据不必要求精准性。

(3) 重视数据的相关性。变量A与变量B有关联,变量A与变量B的变化存在一定的联系,表明两者具有相关性,但相关性不代表因果关系,不能说变量A是变量B变化的原因。

例如,淘宝网运用它的大数据技术算法,根据消费者的历史购买记录或浏览记录来推送给该消费者可能喜欢的商品,这种算法并不能说明该消费者喜欢推送商品的原因,也不能说明消费者如果购买了A之后又购买了B,购买A就是购买B的原因,只能说购买两者具有相关性,或存在一定的概率。大数据技术知道是“什么”,但不知道“为什么”,在大数据技术下,通过相关性查找数据,比小数据时代更便捷、迅速。

大数据系统依赖相关性,而非因果性,相关性表明发生的可能性,而不是发生的原因,通过大数据技术分析,查询到现象的关系物,更快、更迅速且不易受到偏见的影响。建立起技术分析法的预测是大数据的内在要求。

(4) 算法复杂度高。大数据是一种综合交叉的科学,具有不同于一般统计学的计算方法,处理海量的数据需要更智能、更简单的操作的方法和求解方式。因此,对于算法的要求更高,不仅仅是简单算法的集合,更为复杂。

7) 大数据的关键问题和关键技术

(1) 大数据的关键问题。大数据的数据源来源广泛,且数据类型呈多样性,数据计算时,读取和分析的数据量大,要求数据服务具有高效性。

① 半结构化和非结构化的数据处理。大数据中,结构化数据只占15%左右,其余的85%都是半结构和非结构化的数据,它们大量存在于互联网和电子商务等各个领域。如果把系统通过分析数据得到的信息的过程称为“一次挖掘”过程,将得到的信息再结合人们的

主观知识。例如,具体的经验、常识、本能、情境知识和用户偏好,而产生“智能知识”过程则叫作“二次挖掘”。从“一次挖掘”到“二次挖掘”类似事物“量”到“质”的飞跃。

由于大数据所具有的半结构化和非结构化的特点,经过大数据的一次挖掘后的结构化的“粗糙知识”(潜在模式)产生出一些新的特征。“一次挖掘”后的结构化粗糙知识可以被主观知识加工处理并转化,生成半结构化和非结构化的智能知识。寻求“智能知识”反映了大数据研究的核心价值。

②大数据的系统建模与其复杂性。这一问题的突破是利用大数据转化为知识的基础和重点。就目前而言,由于大数据的数据个体复杂且随机,这种数据特征将促使大数据形成自己的数学结构,有利于建立并完善大数据的统一理论。现在,研究界倡导发展一种适应大数据交叉应用的、一般性的结构化数据和半结构化、非结构化数据之间的转化原则。管理学的理论将在实现这种一般性原则和建立大数据规律中发挥关键的作用。

实践中的大数据处理问题是非常繁复的,很难运用单一的计算模式,解决各种不同的大数据计算需求。

大数据的复杂形式促使产生了很多对“粗糙知识”的量化和评估的相关研究。已知的最优化、数据包络分析、期望理论、管理科学中的效用理论可以被应用到研究如何将主观知识与“二次挖掘”过程相融合。这里人机交互将起到至关重要的作用。

③大数据的异构性与决策异构性影响知识发展。大数据本身的复杂性使得传统的数据挖掘理论和技术无法适应大数据的需求。在大数据条件下,管理决策迎来了挑战,面临两个“异构性”问题:“数据异构性”和“决策异构性”。传统的管理决策基于对自身的知识构建和过往的业务经验,而数据分析又是管理决策的基础。

大数据改变了传统的管理决策结构的模式。决策结构的变化要求人们去探讨如何通过“二次挖掘”获得的知识去支撑管理决策。无论大数据带来了哪种数据异构性,大数据中的“粗糙知识”仍可被看作“一次挖掘”的范畴。通过寻找“二次挖掘”产生的“智能知识”来作为数据异构性和决策异构性之间的桥梁是十分必要的。

大数据是具有隐秘规则的人造森林,获寻大数据的科学模式是人们的挑战也是机遇,如果人们找到了将非结构化、半结构化数据转化成结构化数据的规则,已知的数据挖掘方法将成为大数据挖掘的工具。

(2)大数据的关键技术。大数据的关键性技术主要分为流处理、并行化、可视化和摘要索引四种。

①流处理。随着公司的业务处理流程越发复杂,流式数据处理技术已成为大数据的重要处理技术,能满足实时处理的数据处理需求,随时发生数据流的架构,随时处理。

例如,传统的方法只能计算已经给出具体数据的一组数据的平均值,如果数据是移动的,这样的平均值计算则需要大数据的流处理方法,更高效,创建一个数据流统计集,逐步添加数据块,进行移动平均值计算。

②并行化。小数据的存储能力通常不到10 GB,中数据的存储能力不到1 TB,大数据的存储则是分布于多台机器上,多达PB为单位的数据,在分布式数据条件下,需要在极短的时间内处理数据,需要并行化处理。

③可视化。数据可视化分为信息可视化和科学可视化两种。可视化工具是实现可视化的必要手段,常见的可视化工具有两类。

a. 管理决策者或者数据分析师可以利用探索性可视化工具找出数据之间的关联性,这是可视化工具的洞察力作用,如 Tableau、TIBCO、QlikeView。

b. 叙事性可视化工具挖掘数据的方式较为独特。例如,需要用叙事性可视化工具查看某个时间段内某企业的营销数据,可视化格式将预先被创建,数据会按照时间点被逐年显示,并按照设定的条件排序。

④摘要索引。摘要索引是加速查询数据的预计算摘要的过程,这个预计算摘要会被预先创建。摘要索引的作用是为将要进行的查询做计划。现在摘要索引尚没有一个明确的规则,但随着大数据技术的发展,这一问题将会得到解决。

5. 移动互联网

1) 移动互联网的概念

移动互联网是一种通过智能移动终端,采用移动无线通信方式获取业务和服务的新兴业务。移动互联网包含终端、软件和应用 3 个层面。终端层包括智能手机、平板电脑、电子书、MID 等,软件层包括操作系统、中间件、数据库和安全软件等。应用层包括休闲娱乐类、工具媒体类、商务财经类等不同应用与服务。随着技术和产业的发展,未来,长期演进(long term evolution, LTE)和移动支付的支撑技术——近场通信(near field communication, NFC)等网络传输层关键技术也将被纳入移动互联网的范畴。

随着宽带无线接入技术和移动终端技术的飞速发展,人们迫切希望能够随时随地都能方便地从互联网获取信息和服务,由此移动互联网应运而生并迅猛发展。然而,移动互联网在移动终端、接入网络、应用服务、安全与隐私保护等方面还面临着一系列的挑战。其基础理论与关键技术的研究,对于国家信息产业整体发展具有重要的现实意义。

2) 移动互联网的特点

移动互联网业务的特点不仅体现在移动性上,可以随时、随地、随心地享受互联网业务带来的便捷,还表现在更丰富的业务种类、个性化的服务和更高服务质量的保证上。当然,移动互联网在网络和终端方面也受到一定的限制。其特点概括起来主要有以下几个方面。

(1)终端的移动性。移动互联网业务使得用户可以在移动状态下接入和使用互联网服务,移动的终端便于用户随身携带和随时使用。

(2)终端和网络的局限性。移动互联网业务在使用便携的同时,也受到了来自网络能力和终端能力的限制:在网络能力方面,受到无线网络传输环境、技术能力等因素的限制;在终端能力方面,受到终端大小、处理能力、电池容量等的限制。

(3)业务与终端、网络的强关联性。由于移动互联网业务受到了网络及终端能力的限制,因而其业务内容和形式也需要适合特定的网络技术规格和终端类型。

(4)业务使用的私密性。在使用移动互联网业务时,其使用的内容和服务更私密,如手机支付业务等。

3) 移动互联网的发展现状

(1)移动化逐渐完成。随着移动互联网的发展,原先的 PC 网络业务逐渐被迁移到移动互联网的终端上,如邮件、即时信息、浏览器、多媒体播放器等,而原来一些通过网页访问的网站也都为移动设备提供了相应的应用(APP)。

(2)娱乐化仍在加强。移动互联网的主要终端是智能手机,用户平均年龄虽然正在向高

龄迁移,但依然以青年人为主,出于缓解压力的需求,移动互联网依然体现出较强的娱乐性,音乐、视频、手机游戏类应用的广泛应用可见一斑。

(3)区域化分化明显。由于我国省份较多,区域化明显,各地区发展不均衡,因此不同区域内人们对移动互联网的需求各不相同,受到当地经济、技术、学术发展水平的影响较大。这吸引了一部分移动互联网服务的提供者着眼于区域,提供有特色的区域性服务。

(4)功能化尚需加强。移动互联网虽然备受瞩目,但现有的基于移动互联网的服务依然趋于同质化,也有一些行业依然采取旧的思路在运营,互联网只是噱头而没有成为它们真正的工具。

4)移动互联网的发展趋势

(1)亚洲会成为移动领域的领先者。谷歌前董事长艾瑞克·施密特认为在某种程度上,亚洲发明了移动互联网,而基于亚洲网络基础设施建设良好及极速扩大的市场规模,使得亚洲将成为移动领域的领先者。亚洲的确有这样的资本,且不说亚洲有超过世界一半的人口规模,亚洲总体创新能力和水平的提升也使得亚洲的移动互联网产品与理念得到了世界的认可,如韩国的 kakao、中国的支付宝和微信等,都证明了亚洲有和先进国家比拼的能力。

(2)市场规模及用户体量继续高速增长。有“互联网女皇”之称的华尔街证券分析师玛丽·米克尔(Mary Meeker)在对移动互联网的发展态势进行分析时指出,全球 52 亿移动用户中目前仅有 30% 的智能手机使用率,剩余 70% 的广阔市场有待挖掘;中国移动互联网用户数目前已达到了中国互联网用户数的 80%,中国无疑将主导移动商务的革命。工信部公布的相关数据也证明了这一点,截至 2017 年年底,中国的移动互联网用户数已经达到 10.85 亿,这也充分说明了移动互联网的发展潜力。

(3)移动互联网技术——开源与多样化。技术是支撑移动互联网持续发展的根基,同时移动互联网的迅猛发展也推进其中所涉及的网络接入、应用开发、操作系统等方面做出相应的改进与更新。在使用移动互联网时,大多数用户都更青睐通过 Wi-Fi 来接入网络,所以新的 Wi-Fi 标准及更多的基础设施的需求将持续增长。在操作系统方面,安卓的成功意味着开源这种方式的成功,并且在很大程度上促进了移动互联网的技术多样化,以及适应各种需求的健康生态模式的建立。在应用的开发方面,适用于多平台多架构、有很好的稳定性、生产效率高的开发工具将成为主流。

(4)移动搜索市场稳步增长。2017 年上半年度,具有搜索功能的 APP 并未能撼动浏览器搜索的地位,通过浏览器进行搜索的占比为 88.5%,是用户最主要使用的搜索入口;使用搜索引擎类 APP 占比为 79.4%,排名第二。

(5)移动互联网带来在线教育行业爆点。如今,几乎每个人的智能手机里都会装上一两个学习型的 APP,无论是背单词的还是学习专业知识的。通过移动设备下载课程来进行学习的人也越来越多,在线教育的平台类型也在逐渐丰富。根据沪江网提供的数据,我们可以看到移动互联网将在线教育行业发展带入高速时期,沪江网 Web 端用户积累从 0 到 1 000 万用了 10 年的时间,而在移动端实现这个数字仅用了 1 年的时间。移动互联网的发展给教育这一古老且决定未来的领域提供了新的发展契机。

(6)社交进入视频时代。Facebook、Twitter 等在 2013 年就推出了视频分享功能,微信在 2013 年 9 月底推出了包含小视频分享功能的新版本,从分享文字、图片再到文件、视频、

声音,通过移动互联网可以让你的朋友、家人更直观地了解你身边所发生的一切,正是因为移动互联网的存在,社交在某种程度上越来越真实化了。

(7)移动电子商务随时随地。说到电子商务就必须提到“双十一”这一现象级节日。2017年,天猫“双十一”的交易总额达到1 682亿元,同比增长39.35%,其中移动端的交易额占总体交易额的90%。天猫作为具有中国电商行业代表性的电子商务平台,这一数据十分具有说服力。

(8)移动游戏市场产业链优化。可以肯定的是,在未来一段时间内移动游戏市场仍将保持爆发式的增长态势。据不完全统计,移动游戏市场持续增长,2016年,移动游戏市场销售收入为819.2亿元,同比增长59.2%,超越客户端游戏市场成为主流。

(9)移动设备新兴产品层出不穷。作为移动互联网的重要载体——智能手机、平板电脑等移动设备的销量猛增,市场研究机构IDC发布的报告显示,全球Q3智能手机的出货量同比增长25.2%,达到3.276亿部。虽然有人预测智能手机的市场增长将有所放缓,但是庞大的待开发的用户群体将是市场规模持续增长的有力支撑。

(10)移动广告——原生、跨屏、LBS。根据Forrester的预测,截至2020年,将会有20%的销售来自可穿戴设备收集的数据。营销永远都是要做全渠道准备,但面对用户的消费行为逐渐往移动互联网迁徙,战略的转变与制定也迫在眉睫。

在移动互联网时代,用户的需求更加个性化,选择也更多。大数据无疑可以通过对这些需求数据的分析来帮助营销更为精准,在与用户互动中达到传递信息的目的,无疑是最被用户接受的方式。所以,在移动互联时代,原生广告将成为移动广告中的主力。

(11)网络安全伴随发展始终。在移动互联网的高速发展下,移动互联网用户数猛增,移动设备存储了大量的数据,安全隐患逐渐暴露出来。移动互联网的确实实在在地便利着我们的生活,但频发的安全事故造成的负面影响还是巨大的,所以不管是未来1年或者未来5年甚至更久,网络安全的保障都将会与互联网的发展如影随形。

1.2.2 数据通信技术基础

数据通信是网络技术发展的基础,学习数据通信可以帮助读者理解网络中数据传输的原理和实现方法。

1. 基本概念

数据通信是通信技术和计算机技术相结合而产生的一种新的通信方式。要在两地间传输信息就必须有传输信道,根据传输媒体的不同,有有线数据通信与无线数据通信之分。但它们都是通过传输信道将数据终端与计算机联结起来,而使不同地点的数据终端实现软、硬件和信息资源的共享。图1-18所示为简单的通信模型。



图 1-18 简单的通信模型

下面介绍一些通信术语。

(1)数据。数据可分为模拟数据和数字数据。模拟数据是在某区间内连续变化的值,数字数据是离散的值。

(2)信号。信号是指数据的电子或电磁编码。信号可分为模拟信号和数字信号。模拟信号是随时间连续变化的电流、电压或电磁波,如图 1-19 所示;数字信号则是一系列离散的电脉冲,可选择适当的参量来表示要传输的数据,如图 1-20 所示。

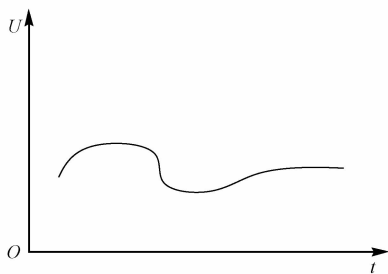


图 1-19 模拟信号

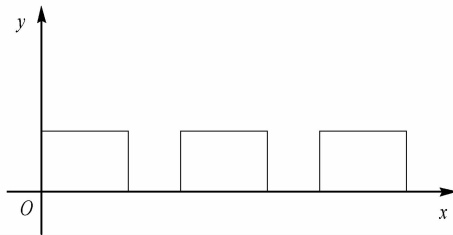


图 1-20 数字信号

(3)信息。信息是指数据的内容和解释,是客观事物属性和相互联系特征的表现,它反映了客观事物的存在形式和运动状态,是对客观事物存在形式的一种反映。

(4)信源。信源是指通信过程中产生和发送信息的设备或计算机。

(5)信宿。信宿是指通信过程中接收和处理信息的设备或计算机。

(6)信道。信道是指信源和信宿之间的通信线路。

(7)数据传输。数据传输是指用电信号把数据从发送端传送到接收端的过程。

2. 网络传输技术

在计算机网络通信中,常见的数据传输方式有并行传输和串行传输两种。

1) 并行传输

并行通信传输中有多个数据位,同时在两个设备之间传输。发送设备将这些数据位通过对应的数据线传送给接收设备,还可附加一位数据校验位。接收设备可同时接收到这些数据,不需要做任何变换就可直接使用。并行方式主要用于近距离通信。计算机内的总线结构就是并行通信的例子。这种方法的优点是传输速度快、处理简单。

2) 串行传输

串行数据传输时,数据是一位一位地在通信线上传输的,先由具有几位总线的计算机内的发送设备,将几位并行数据经并—串转换硬件转换成串行方式,再逐位经传输线到达接收站的设备中,并在接收端将数据从串行方式重新转换成并行方式,以供接收方使用。串行数据传输的速度要比并行传输慢得多,但对于覆盖面极其广阔的公用电话系统来说具有更大的现实意义。

假设有数据 1001010 需要传输,采用并行传输和串行传输两种不同的传输方式进行传输的过程分别如图 1-21 和图 1-22 所示。



动画
并行传输



动画
串行传输

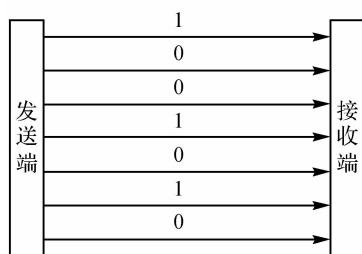


图 1-21 并行传输



图 1-22 串行传输

3. 调制与编码

数据在计算机中是以二进制形式表示的,而在传输时,数据编码的类型是由通信信道所支持的通信类型决定的。例如,模拟信号在模拟信道上传输,而数字信号则在数字通道上传输。如果模拟信号需要在数字通道上传输,又或者数字信号需要在模拟信道上传输,它们是如何实现的?这就需要了解调制和编码技术。

1) 调制技术

简单来说,调制技术就是数字数据和模拟数据的转换。所以在这里需要先了解两个基本的概念:调制和解调。

调制是在发送端将数据进行处理并放到载波上,使其变为可以在信道传输的信号的过程。它通过调制器完成。

解调是在接收端将被处理的数据还原成原来的数据的过程。它通过解调器完成。

调制技术在计算机网络中的应用过程如图 1-23 所示。

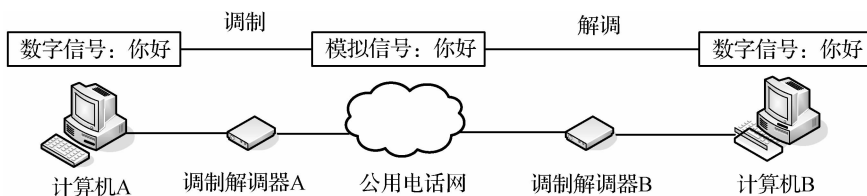


图 1-23 调制技术在计算机网络中的应用过程

在图 1-23 中,“计算机 A”如果要向“计算机 B”发送一个“你好”的数字信号,首先计算机 A 发出“你好”的数字信号,然后通过“调制解调器 A”把“你好”的数字信号调制为模拟信号并在“公用电话网”里面传输,再通过“调制解调器 B”把“你好”的模拟信号解调成数字信号并传输给“计算机 B”。

2) 编码技术

调制技术是为了解决数据传输的问题,而编码技术则是为了解决数据翻译的问题。编码技术就是将数据变换成数字信号。其中也有两个基本概念:编码和解码。

编码是将模拟数据或数字数据变成数字信号。

解码就是在接收端把数字信号变成原来的形式。

在数据传输系统中,主要采用了数字数据的数字信号编码、数字数据的模拟信号编码、模拟数据的数字信号编码 3 种数据编码技术。

(1)数字数据的数字信号编码。数字信号的传输一般采用两个电压平来表示两个二进制数字。例如,无电压用“0”表示,而恒定的正电压用“1”表示,称为不归零制(NRZ)编码,如图 1-24(a)所示。

不归零制编码虽然在传输中效率最高,但是存在发送方和接收方的同步问题。克服上述缺点的另外一个编码方案就是曼彻斯特编码,如图 1-24(b)所示,这种编码通常用于局部网络传输。在曼彻斯特编码中,每一位的中间有一个跳变。位之间的跳变既作为时钟,又作为数据;由高到低的跳变表示 1,由低到高的跳变表示为 0。还有一种编码称为差分曼彻斯特编码,如图 1-24(c)所示,为跳变仅提供时钟定时,用每位周期开始时有无跳变来表示 0 和 1 的编码,它的缺点是编码效率低。

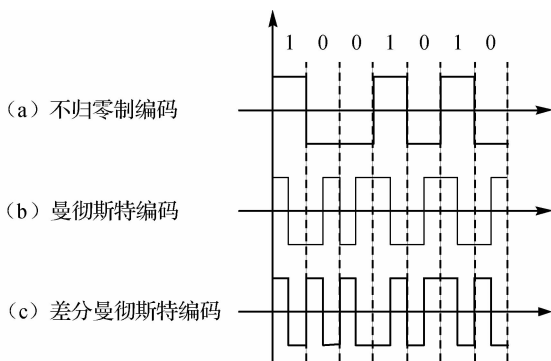


图 1-24 数字信号编码

(2)数字数据的模拟信号编码。要在模拟信道上传输数字数据,首先数字信号要对相应的模拟信号进行调制,即用模拟信号作为载波运载要传送的数字数据。

载波信号可以表示为正弦波形式 $f(t) = A\sin(\omega t + \varphi)$,其中幅度 A 、频率 ω 和相位 φ 的变化均影响信号波形。因此,通过改变这三个参数可实现对模拟信号的编码。相应的调制方式分别称为幅度调制 ASK、频率调制 FSK 和相位调制 PSK。结合 ASK、FSK 和 PSK 可以实现高速调制,常见的组合是 PSK 和 ASK 的结合。

①幅度调制。幅度调制简称调幅,也称为幅移键控(ASK amplitude-shift keying)。调制原理为用两个不同振幅的载波分别表示二进制值 0 和 1。

②频率调制。频率调制简称调频,也称为频移键控(FSK frequency-shift keying)。调制原理为用两个不同频率的载波分别表示二进制值 0 和 1。

③相位调制。

a. 绝对相移键控法。绝对相移键控用两个固定的不同相位表示数字“0”和“1”,用公式可表示为

$$\begin{aligned} U(t) &= U_m \sin(\omega t + \pi) && \text{数字“1”} \\ &= U_m \sin(\omega t + 0) && \text{数字“0”} \end{aligned}$$

b. 相对相移键控法。相对相移键控用载波在两位数字信号的交接处产生的相位偏移来表示载波所表示的数字信号。最简单的相对调相方法是:与前一个信号同相表示数字“0”,相位偏移 180° 表示“1”。这种方法具有较好的抗干扰性。

(3)模拟数据的数字信号编码。在数字化的电话交换和传输系统中,通常需要将模拟的



动画
模拟信号与
数字信号

话音数据编码成数字信号后再进行传输。这里常用的一种称为 PCM(pulse code modulation)的脉冲编码调制技术。

PCM 基于以下的采样定理:如果在规定的时间内,以有效信号 $f(t)$ 最高频率的两倍或两倍以上速率对该信号进行采样,则这些采样值包含了无混叠而又便于分离的全部原始信号信息。利用低通滤波器可不失真地从这些采样值中重新构造出 $f(t)$ 。PCM 原理如图 1-25 所示。

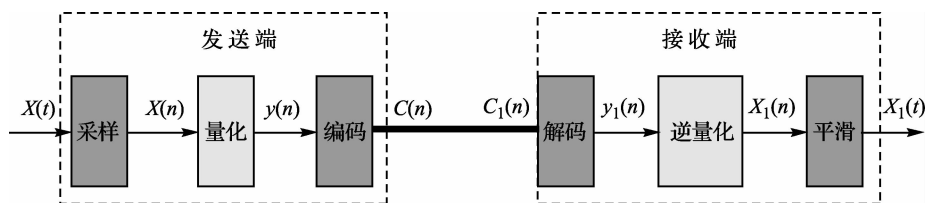


图 1-25 PCM 原理

信号数字化的转换过程可包括采样、量化和编码 3 个步骤。

①采样。每个固定的时间间隔,取出模拟数据的瞬时值,作为本次抽样到下次抽样之间该模拟数据的代表值。 $X(n)$ 就是采样处理后的脉冲调幅信号。

②量化。把抽样取得的电平幅值按照一定的分级标度转换成对应的数字值,并取整数。这样把连续的电平幅值转换成离散的数字 $y(n)$ 。

③编码。它是将量化后的整数表示为一定位数的二进制数 $C(n)$ 。

在发送端,经过信号数字化过程后,就可把模拟信号转换成二进制数码脉冲序列,然后经过信道进行传输。在接收端,将接收到的信号 $C_1(n)$ 解码成 $y_1(n)$,在通过逆量化获得信号 $X_1(n)$,最后平滑之后的信号 $X_1(t)$ 就是还原的模拟信号。 $X_1(t)$ 与 $X(t)$ 之差就是量化的误差。

根据原信号的频宽可以估算出采样的速度。如果声音数据限于 4 000 Hz 以下的频率,那么每秒钟 8 000 次的采样可以满足完整地表示声音信号的特征。如果使用七位二进制表示采样值,就允许有 128 个量化级,这就意味着,仅仅是声音信号就需要有每秒钟 8 000 次采样乘以每次采样 7 位等于 56 000 b/s 的数据传输速率。

以上是模拟数据,如声音经过 PCM 编码后成数字信号,就可以采用数字传输方式进行传输了。另外,计算机中的数字数据经过适当的编码后可直接采用数字传输方式传输。这样模拟数据和数字数据经过适当的编码后,可统一到相同的传输方式下进行传输。由于数字信号在传输过程中不引入噪声,传输可靠性高,因而应用相当广泛,如数字电话、数字传真、数字电视等。特别是目前多媒体技术的应用,要求将不同媒体的物理量(模拟量),如声音、图像、动画等,转换成数字信号后在计算机和网络系统内进行存储、处理和传输。这些都要用到模拟数据的数字传输技术。

4. 多路复用技术

多路复用技术就是把许多个单个信号在一个信道上同时传输的技术。频分多路复用 FDM 和时分多路复用 TDM 是两种最常用的多路复用技术。

1) 频分多路复用 FDM 技术原理

在物理信道的可用带宽超过单个原始信号所需带宽的情况下,可将该物理信道



动画
多路复用技术

的总带宽分割成若干个与传输单个信号带宽相同(略宽)的子信道,每个子信道传输一路信号,这就是频分多路复用。

多路原始信号在频分复用前,先要通过频谱搬移技术将各路信号的频谱搬移到物理信道频谱的不同段上,使各信号的带宽不相互重叠,然后用不同的频率调制每一个子信号,每个子信号可以并行传送一路信号。为了防止互相干扰,使用保护带来隔离每一个通道。

2)时分多路复用 TDM 技术原理

若媒体能达到的位传输速率超过传输数据所需的数据传输速率时,可采用时分多路复用 TDM 技术,即将一条物理信道按时间分成若干个时间片轮流地分配给多个信号使用。每一时间片由复用的一个信号占用,这样,利用每个信号在时间上的交叉就可以在一条物理信道上传输多个数字信号。

时分多路复用 TDM 技术不仅局限于传输数字信号,也可同时交叉传输模拟信号。

5. 数据交换技术

数据经编码后在通信线路上进行传输,按数据传送技术划分,交换技术又可分为电路交换、报文交换和分组交换。

1)电路交换

电路交换是指在电路交换网络中,在通信双方之间建立一条专用线路,最常见的有公用电话交换网(PSTN)。该方式的通信过程如下。

(1)建立电路。在传输任何数据之前,要先经过呼叫过程建立一条端到端的电路。如图 1-26 所示,若 A 站要与 D 站连接,典型的做法是:A 站先向与其相连的 1 节点提出请求,然后 1 节点在通向相邻节点的路径中找到下一个支路。例如,1 节点选择经 5 节点的电路,在此电路上分配一个未用的通道,并告诉 5 它还要连接 4 节点;5 再呼叫 4,建立电路 1-4 路径;最后,节点 4 完成到 D 站的连接。这样 1 与 4 之间就有一条专用电路 1-5-4,用于 A 站与 D 站之间的数据传输。

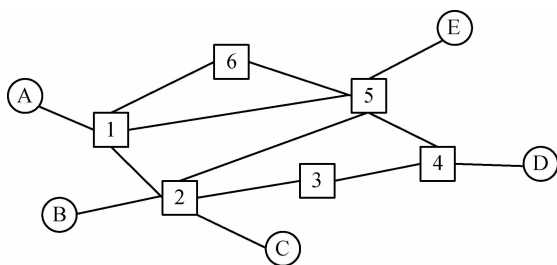


图 1-26 电路交换过程

(2)进行数据传输。建立电路 1-5-4 以后,数据就可以从 1 发送到 5,再由 5 交换到 4;4 也可以经 5 向 1 发送数据。在整个数据传输过程中,所建立的电路必须始终保持连接状态。

(3)电路拆除。数据传输结束后,由某一方(1 或 4)发出拆除请求,然后逐节拆除到对方节点。

电路交换技术的优点是数据传输可靠、迅速,数据不会丢失且保持原来的序列,但是它



动画
频分复用技术



动画
电路交换的
三个阶段

的缺点是在某些情况下,电路空闲时的信道容易被浪费。在短时间数据传输时电路建立和拆除所用的时间得不偿失。因此,它适用于系统间要求高质量的大量数据传输的情况。

2) 报文交换

这种方式不要求在两个通信节点之间建立专用通路。节点把要发送的信息组织成一个数据包,即报文。该报文中含有目标节点的地址,完整的报文在网络中一站一站地向前传送。每一个节点接收整个报文,检查目标节点地址,然后根据网络中的交通情况在适当时转发到下一个节点。经过多次的存储-转发,最后到达目标,因而这样的网络称为存储转发网络。其中的交换节点要有足够大的存储空间(一般是磁盘),用以缓冲收到的长报文。

(1) 报文交换的优点。

① 电路利用率高。由于许多报文可以分时共享两个节点之间的通道,因而对于同样的通信量来说,对电路的传输能力要求较低。

② 在电路交换网络上,当通信量变得很大时,就不能接收新的呼叫。在报文交换网络上,通信量大时仍然可以接收报文,不过传送延迟会增加。

③ 报文交换系统可以把一个报文发送到多个目的地,而电路交换网络很难做到这一点。

④ 报文交换网络可以进行速度和代码的转换。

(2) 报文交换的缺点。

① 不能满足实时或交互式的通信要求,报文经过网络的延迟时间长且不定。

② 有时节点收到过多的数据而无空间存储或不能及时转发时,就不得不丢弃报文,而且发出的报文不按顺序到达目的地。

3) 分组交换

分组交换是报文交换的一种改进,它将报文分成若干个分组,每个分组的长度有一个上限,有限长度的分组使得每个节点所需的存储能力降低了,分组可以存储到内存中,提高了交换速度。它适用于交互式通信,如终端与主机通信。分组交换有虚电路分组交换和数据报分组交换两种。它是计算机网络中使用最广泛的一种交换技术。

(1) 虚电路分组交换原理与特点。在虚电路分组交换中,为了进行数据传输,网络的源节点和目的节点之间要先建一条逻辑通路。每个分组除了包含数据之外,还包含一个虚电路标识符。在预先建好的路径上的每个节点都知道应该把这些分组引导到哪里去,不再需要路由选择判定。最后,由某一个站用清除请求分组来结束这次连接。它之所以是“虚”的,是因为这条电路不是专用的。

虚电路分组交换的主要特点是:在数据传送之前必须通过虚呼叫设置一条虚电路。但并不像电路交换那样有一条专用通路,分组在每个节点上仍然需要缓冲,并在线路上进行排队等待输出。

(2) 数据报分组交换原理与特点。在数据报分组交换中,每个分组的传送是被单独处理的。每个分组称为一个数据报,每个数据报自身携带足够的地址信息。一个节点收到一个数据报后,根据数据报中的地址信息和节点所储存的路由信息,找出一个合适的出路,把数据报原样地发送到下一节点。由于各数据报所走的路径不一定相同,因而不能保证各个数据报按顺序到达目的地,有的数据报甚至会中途丢失。整个过程中,没有虚电路建立,但要为每个数据报做路由选择。

1.3 实训

1.3.1 认识网络设备

1. 实训目的

认识实验室的网络设备的基本情况、简单使用网络设备。

2. 实训内容

认识设备、简单使用。

3. 实训方法

- (1)对设备的实务进行了了解。
- (2)对设备进行简单的连接操作。
- (3)画出设备的连接图。

1.3.2 使用 Visio 2010 绘制网络拓扑图

1. 实训目的

掌握用 Visio 2010 软件绘制网络拓扑图。

2. 实训内容

根据试验一的草图使用 Visio 2010 绘制网络拓扑图。

3. 实训方法

(1)运行 Visio 2010 软件,在打开的图 1-27 所示窗口的“模板类别”列表框中选择“网络”选项,然后在右边窗口中选择一个对应的选项,如“详细网络图”,打开图 1-28 所示的界面(在此仅以选择“详细网络图”选项为例)。



图 1-27 Visio 2010 主界面

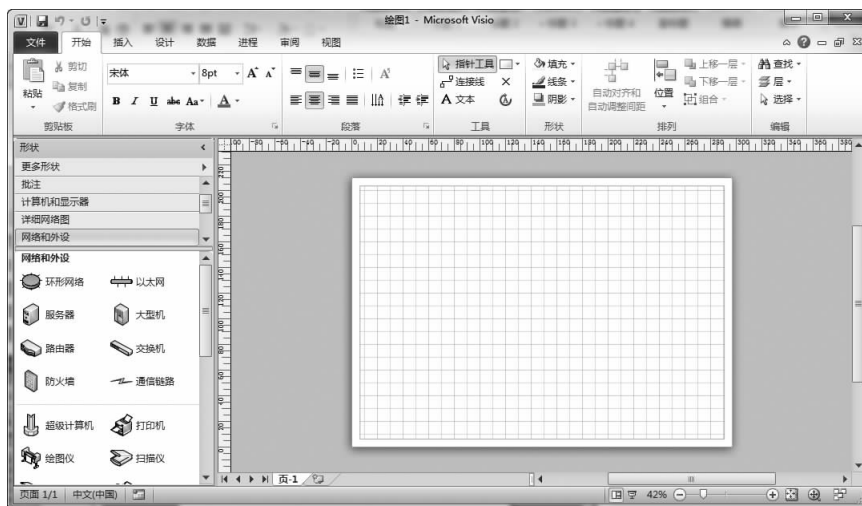


图 1-28 “详细网络图”拓扑结构绘制界面

(2)在左侧窗格中选择“网络和外设”选项,在其中的图像列表中选择“交换机”选项(因为交换机通常是网络的中心,首先确定好交换机的位置),按住鼠标左键不放并把“交换机”选项拖动到右边窗口中的相应位置,然后松开鼠标左键,得到一个交换机图像,如图 1-29 所示。用户还可以在按住鼠标左键的同时拖动四周的绿色方格来调整图像大小,通过按住鼠标左键的同时旋转图像顶部的绿色小圆圈,可以改变图像的摆放方向,再通过把鼠标放在图像上,然后在出现 4 个方向箭头时按住鼠标左键并拖动鼠标,可以调整图像的位置。图 1-30 所示为调整交换机图像大小、方向和位置后的图示。

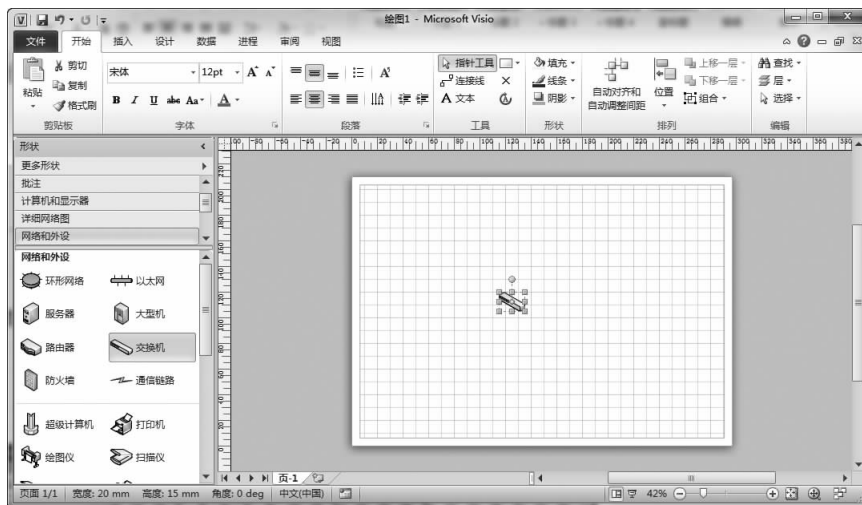


图 1-29 图像拖放到绘制平台后的图示

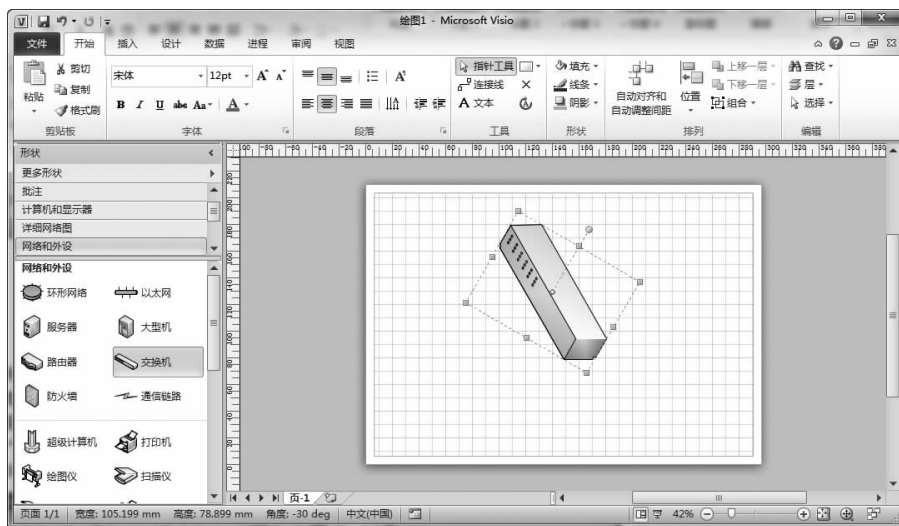


图 1-30 调整交换机图像大小、方向和位置后的图示

(3) 双击交换机图像,即可在图像下方显示一个小的文本框,此时用户可以在此输入交换机的型号或其他标注,如图 1-31 所示。输入完后在任意空白处单击即可,图像又恢复原来调整后的大小。标注文本的字体、字号和格式等都可以通过工具栏中的相关选项来调整,若要使调整适用于所有标注,则可在图像上右击,在弹出的快捷菜单中选择“格式”→“文本”菜单项,打开图 1-32 所示的对话框,在此可以进行详细的配置。标注的输入文本框位置也可通过按住鼠标左键移动。

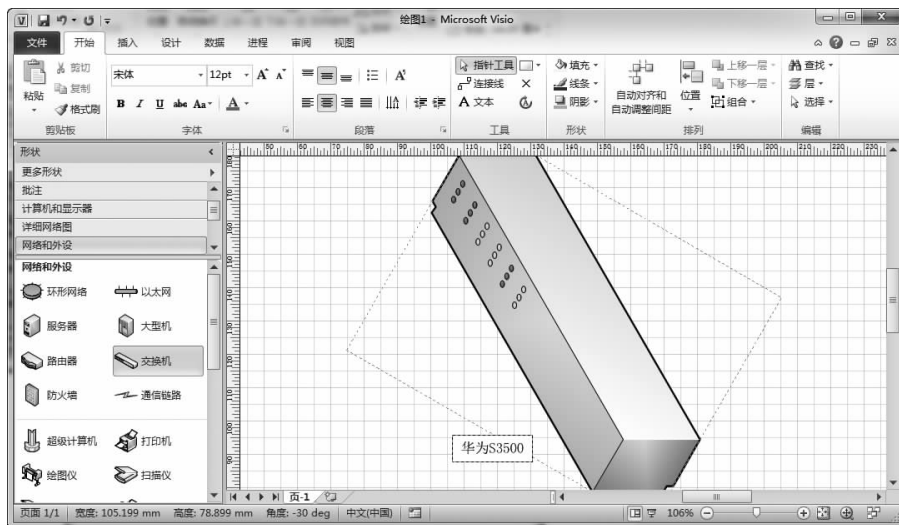


图 1-31 给图像输入标注



图 1-32 “文本”对话框

(4)以同样的方法添加一台服务器,并把它与交换机连接起来。服务器的添加方法与交换机一样,在此只介绍交换机与服务器的连接方法。在 Visio 2010 中介绍连接方法很复杂,其实可以不用管它,只需使用工具栏中的“连接线工具”进行连接即可。在选择该工具后,单击要连接的两个图像之一,此时会有一个红色的方框,移动鼠标选择相应的位置,当出现红色星状点时按住鼠标左键,把连接线拖到另一图像上,注意此时若出现一个大的红方框,则表示不宜选择此连接点,只有当出现小的红色星状点即可松开鼠标,连接成功,图 1-33 所示为图像之间的连接示例。

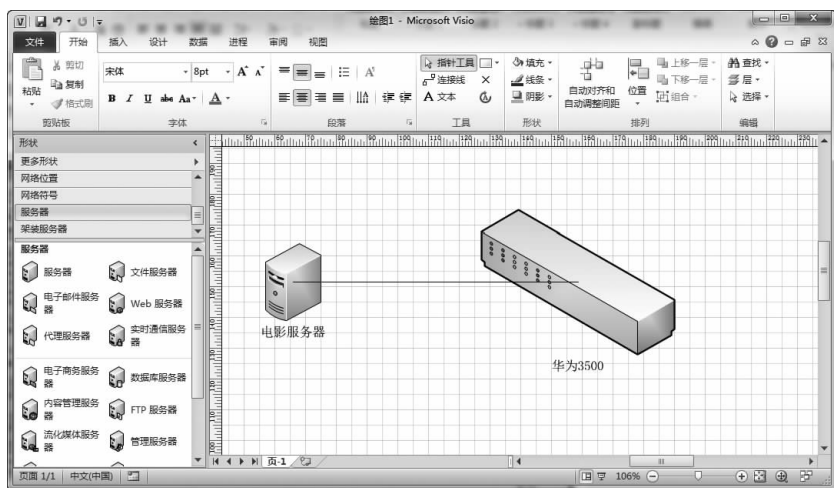


图 1-33 图像之间的连接示例

(5)把其他网络设备图像一一添加并与网络中的相应设备图像连接起来,当然这些设备图像可能会在左侧窗格的不同类别选项下面。如果左边已显示的类别中没有包括,则可通过选择“文件”→“新建”菜单项,在打开的子菜单中添加其他类别的图形,并显示在左侧窗格中。图 1-34 所示为用 Visio 2010 绘制的简单网络拓扑结构示意图。

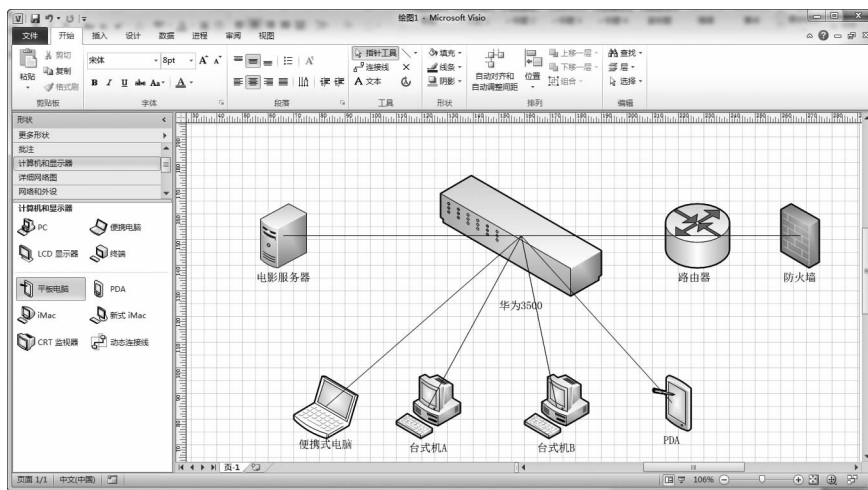


图 1-34 用 Visio 2010 绘制的简单网络拓扑结构示意图

以上只是介绍了 Visio 2010 的一部分网络拓扑结构绘制功能,因为它的使用方法比较简单,操作方法与 Word 类似,在此不再一一详细介绍。

1.4 思考与练习

一、名词解释

1. 数据。
2. 网络拓扑。
3. 服务器。

二、填空题

1. 计算机网络系统是由_____子网和_____子网组成的。
2. 计算机网络最突出的优点是_____。
3. 计算机网络按覆盖范围可分为_____、_____和_____。
4. 计算机网络技术是_____和_____结合的产物。

三、简答题

1. 什么是计算机网络?
2. 计算机网络的功能主要有哪些?
3. 简述计算机网络的分类。
4. 简述计算机发展过程。
5. 简述一个局域网的网络组成结构。

2

模块 2

计算机网络体系结构

▶ 本模块学习前的思考点

- 什么是协议?
- 为什么有参考模型?
- 参考模型的含义是什么?
- 如何使用命令进行日常网络维护?
- 子网划分的意义是什么?

2.1 必备知识

2.1.1 了解计算机网络协议与标准组织

前面对计算机网络的概念做了一些介绍,下面考虑计算机网络中另一个重要的术语——协议。什么是协议?协议是干什么的?如果遇见一个协议,如何识别它?

1. 人类活动的类比

要理解计算机网络协议的概念,最简单的方法就是和人类活动进行类比,因为人类无时无刻不在执行协议。例如,当你需要向某人询问时间时,将会怎么做?图 2-1 中显示了一种典型的交互过程。人类协议要求一方首先进行问候(如图 2-1 中的第一个“你好”),以此来开始与另外一个人的对话。对这个“你好”的典型响应是返回一个与“你好”相关的报文(如图 2-1 中第二个“你好”)。这表明能够继续向那人询问时间。如果对最初的“你好”回复的是不同的响应(例如,“不要烦我!”或“我不会说中文”等其他不合时宜的回答)可能表明:你可能不能继续和他通话。这种情况下,按照人类协议,发话者将不能够询问时间了。如果人们执行不同的协议(例如,一个人说中文,一个人说英文),该协议将不能够互动,因此不能完成本次通话。在网络中也一样,为了完成一次通信,就要求两个或两个以上通信实体执行相同的协议。

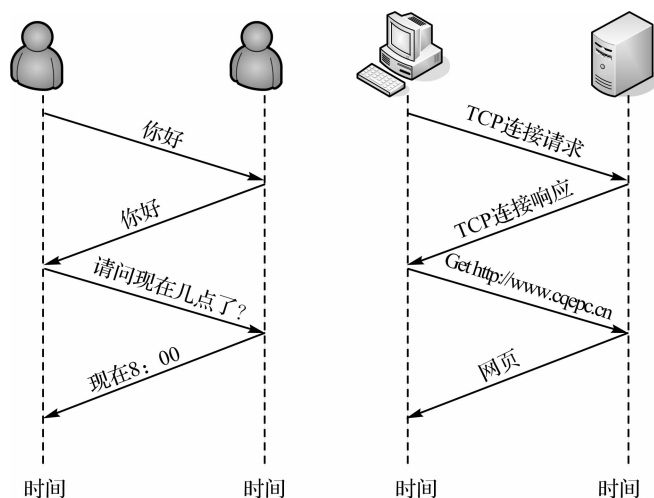


图 2-1 人类活动协议与计算机网络访问协议

2. 网络协议

网络协议类似于人类协议,只不过交谈的内容和交谈的对象是某些实体设备(计算机、PDA、移动电话、路由器等具有网络能力的设备)的硬件或软件组成。计算机网络中的所有活动,凡是涉及两个或者两个以上通信的实体都受到协议的制约。例如,在两台物理连接的计算机的网络接口卡中,硬件实现的协议控制了两块网络接口卡间的“线上”比特流;系统中的拥塞控制协议控制了发送方和接收方之间传输的分组的速率等。

以一个简单的计算机网络协议为例。例如,当计算机向一个 Web 服务器发出请求时,即使用者在 Web 浏览器中输入一个 Web 网站地址时,它是如何工作的?首先,计算机将向 Web 服务器发送一条“请求连接”的报文,并等待回答。Web 服务器最终将接收到该连接请求报文,并返回一条“连接这响应”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的名字。最后,Web 服务器向该计算机返回该 Web 网页(数据)。

从上述人类活动和网络例子中可见,报文的交换及当发送和接收这些报文时所采取的动作是一个协议的关键。

协议定义了在两个或者多个通信实体之间交换的报文格式和次序,以及在报文传输或接收等其他事件方面所采取的动作。

简单来说,协议就是通信双方必须遵守的规则与规定,也可以说协议是协议制定组织给出来的一个标准。

3. 计算机网络管理组织

网络的管理组织有以下几个。

1) Internet 协会

Internet 协会(Internet Society, ISOC)是一个专业性的会员组织,由来自 100 多个国家的 150 个组织以及 6 000 名个人成员组成,这些组织和个人影响着 Internet 现在和未来的技术。ISOC 由几个负责 Internet 结构标准的组织组成,包括 Internet 体系结构组(IAB)和

Internet 工程任务组(IETF)。ISOC 的主站点是 <http://www.ISOC.org/>。

2) Internet 体系结构组

Internet 体系结构组(Internet Architecture Board, IAB)以前称为 Internet 行动组,是 Internet 协会技术顾问,这个小组定期会晤、考查由 Internet 工程任务组和 Internet 工程指导组提出的新思想和建议,并给 IETF 带来一些新的想法和建议。IAB 的 Web 站点是 <http://www.IAB.org/>。

3) Internet 工程任务组

Internet 工程任务组(Internet Engineering Task Force, IETF)是由网络设计者、制造商和致力于网络发展的研究人员组成的一个开放性组织。IETF 一年会晤 3 次,主要的工作通过电子邮件组来完成, IETF 被分成多个工作组,每个组有特定的主题。IESG 工作组包括超文本传输协议(HTTP)工作组和 Internet 打印协议(IPP)工作组。

IETF 对任何人都是开放的,其站点是 <http://www.IETF.org>。

4) Internet 工程指导组

Internet 工程指导组(Internet Engineering Steering Group, IESG)负责 IETF 活动和 Internet 标准化过程的技术性管理, IESG 也保证 ISOC 的规定和规程能顺序进行。IESG 给出关于 Internet 标准规范采纳前的最后建议。通过访问 <http://www.IETF.org/iesg.html> 可获得更多关于 IESG 的信息。

5) Internet 编号管理局

Internet 编号管理局(Internet Assigned Numbers Authority, IANA)负责分配 IP 地址和管理域名空间, IANA 还控制 IP 协议端口号和其他参数, IANA 在 ICANN 下运作。IANA 的站点是 <http://www.IANA.org/>。

6) Internet 名字和编号分配组织(Internet Corporation for Assigned Names and Numbers, ICANN)

ICANN 是为国际化管理名字和编号而形成的组织。其目标是帮助 Internet 域名和 IP 地址管理从政府向民间机构转换。当前, ICANN 参与共享式注册系统(shared registry system, SRS),通过 SRS, Internet 域的注册过程是开放式公平竞争的。关于 ICANN 的更多信息可通过访问 <http://www.icann.org/> 获得。

7) Internet 网络信息中心和其他注册组织

Internet 网络信息中心(Internet network information center, InterNIC)从 1993 年起由 Network Solutions 公司运作,负责最高级域名的注册(.com,.org,.net,.edu), InterNIC 由美国国家电信和信息管理机构(NTIA)监督,这是商业部的一个分组。InterNIC 把一些责任委派给其他官方组织,如国防部 NIC 和亚太地区 NIC。最近有一些建议想把 InterNIC 分成更多的组,其中一个建议是已知共享式注册系统, SRS 在域注册过程中努力引入公平和开放的竞争。当前,有 60 多家公司进行注册管理。

8) RFC 编辑

RFC(Request For Comments)是关于 Internet 标准的一系列文档。RFC 编辑是 Internet RFC 文档的出版商,负责 RFC 文档的最后编辑检查。想获得关于 RFC 编辑的更多信息,可以访问 <http://www.RFC-editor.org/>。

9) Internet 服务提供商

20 世纪 90 年代, Internet 商业化之后, 大量的 Internet 服务提供商 (Internet Service Provider, ISP) 正“焦急”地等待着帮助成千上万个家庭和商业用户接入 Internet。ISP 是商业机构, 它们在办公室或计算机房内设有服务器, 这些服务器配置了调制解调器, 使用点到点协议 (PPP) 或串行线路接口协议 (SLIP)。这些协议允许远程用户使用拨号方式把个人计算机和 Internet 相连。为了获取费用, ISP 提供远程用户至 Internet 的接入支持。大多数 ISP 在服务器上也提供电子邮件账号, 甚至提供 UNIX Shell 账号。更大的 ISP 能够提供商业机构及其他 ISP 的 Internet 接入服务。这些 ISP 具有更快速的网络, 如 ISDN、分时 T-1 线路甚至更高。Internet.com 提供了 ISP 的数据库, 可通过电话区号搜索, 该 ISP 向导的站点网址为 <http://thelist.internet.com>。

2.1.2 熟悉 OSI 参考模型

在计算机网络产生之初, 每个计算机厂商都有一套自己的网络体系结构的概念, 它们之间互不相容。为此, 国际标准化组织 (International Organization for Standardization, ISO) 在 1979 年建立了一个分委员会来专门研究一种用于开放系统互连的体系结构 (open systems interconnection, OSI), “开放”这个词表示: 只要遵循 OSI 标准, 一个系统可以和位于世界上任何地方的、也遵循 OSI 标准的其他任何系统进行连接。这个分委员提出了开放系统互联, 即 OSI 参考模型, 它定义了连接不同计算机厂家的计算机的标准框架。

OSI 参考模型分为 7 层, 从下往上分别是物理层、数据链路层、网络层、传输层、会话层、表示层和应用层, 如图 2-2 所示。

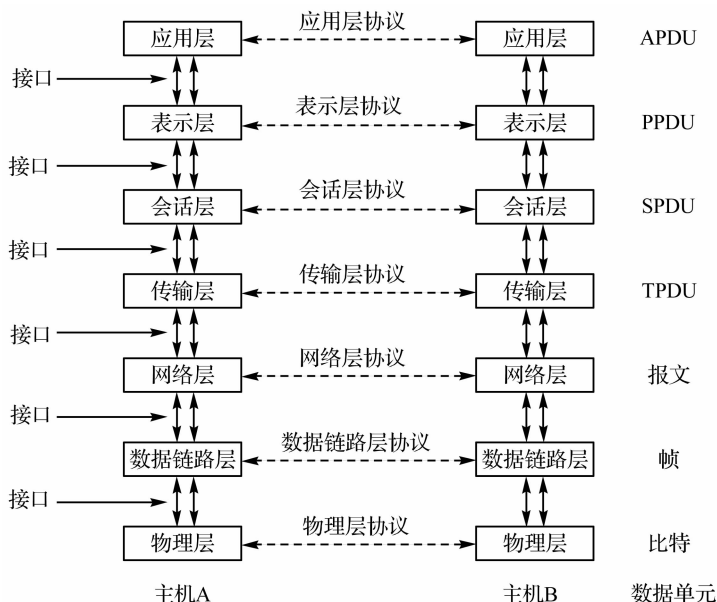


图 2-2 OSI 参考模型

1. 七个层次划分原则

ISO 为了使网络应用更为普及, 就推出了 OSI 参考模型。其目的就是使所有公司有统



动画
OSI 中信息
流动



动画
七层模型

一的标准,从而使得网络的开发与管理更加规范。这样所有公司都有相同的规范,就能互连了。提供各种网络服务功能的计算机网络系统是非常复杂的。根据分而治之的原则,ISO将整个通信功能划分为7个层次,划分原则如下。

- (1)网络中各节点都有相同的层次。
- (2)不同节点的同等层具有相同的功能。
- (3)同一节点内相邻层之间通过接口通信。
- (4)每一层使用下层提供的服务,并向其上层提供服务。
- (5)不同节点的同等层按照协议实现对等层之间的通信。

2. OSI 参考模型各层功能

1) 物理层

物理层(physical layer)是参考模型的最底层。它的主要功能是完成相邻节点之间原始比特流的传输。物理层协议关心的典型问题是使用什么样的物理信号来表示数据1和0;信号持续的时间多长;数据传输是否可同时在两个方向上进行;最初的连接如何建立和完成通信后连接如何终止;物理接口(插头和插座)有多少针及各针的用处。物理层的设计主要涉及物理层接口的机械、电气、功能和过程特性,以及物理层接口连接的传输介质等问题。物理层的设计还涉及通信工程领域内的一些问题。

2) 数据链路层

数据链路层(data link layer)的主要功能是如何在不可靠的物理线路上进行数据的可靠传输。数据链路层完成的是网络中相邻节点之间可靠的数据通信。为了保证数据的可靠传输,发送方把用户数据封装成帧(frame),并按顺序传送各帧。由于物理线路的不可靠,因而发送方发出的数据帧有可能在线路上发生出错或丢失(所谓丢失实际上是数据帧的帧头或帧尾出错),从而导致接收方不能正确地接收到数据帧。为了保证能让接收方对接收到的数据进行正确性判断,发送方为每个数据块计算出CRC(循环冗余检验)并加入帧中,这样接收方就可以通过重新计算CRC来判断数据接收的正确性。一旦接收方发现接收到的数据有错,发送方就必须重传这一帧数据。然而,相同帧的多次传送也可能使接收方收到重复帧。例如,接收方给发送方的确认帧被破坏后,发送方也会重传上一帧,此时接收方就可能接收到重复帧。数据链路层必须解决由帧的损坏、丢失和重复所带来的问题。数据链路层要解决的另一个问题是防止高速发送方的数据把低速接收方“淹没”。因此,需要某种信息流量控制机制使发送方得知接收方当前还有多少缓存空间。为了控制的方便,流量控制常常和差错处理一同实现。

3) 网络层

网络层(network layer)的主要功能是完成网络中主机间的报文传输,其关键问题之一是使用数据链路层的服务将每个报文从源端传输到目的端。在广域网中,这包括产生从源端到目的端的路由,并要求这条路径经过尽可能少的IMP。如果在子网中同时出现过多的报文,子网可能形成拥塞,必须加以避免,此类控制也属于网络层的内容。当报文不得不跨越两个或多个网络时,又会产生很多新问题。例如,第二个网络的寻址方法可能不同于第一个网络;第二个网络也可能因为第一个网络的报文太长而无法接收;两个网络使用的协议也可能不同等。网络层必须解决这些问题,使异构网络能够互连。在单个局域网中,网络层是冗余的,因为报文是直接从一个计算机传送到另一台计算机的,因此网络层所要做的工作很少。

4) 传输层

传输层(transport layer)的主要功能是完成网络中不同主机上的用户进程之间可靠的数据通信。传输层要决定对会话层用户(最终对网络用户)提供什么样的服务。最好的传输连接是一条无差错的、按顺序传送数据的管道,即传输层连接是真正端到端的。换言之,源端机上的某进程,利用报文头和控制报文与目标机上的对等进程进行对话。在传输层下面的各层中,协议是每台机器与它的直接相邻机器之间(主机-IMP、IMP-IMP)的协议,而不是最终的源端机和目标机之间(主机—主机)的协议。在它们中间,可能还隔着多个 IMP。即 1~3 层的协议是点到点的协议,而 4~7 层的协议是端到端的协议。由于绝大多数主机都支持多用户操作,因而机器上有多道程序,这意味着多条连接将进出于这些主机,因此需要以某种方式区别报文属于哪条连接。识别这些连接的信息可以放入传输层的报文头中。除了将几个报文流多路复用到一条通道上,传输层还必须管理跨网连接的建立和拆除。这就需要某种命名机制,使机器内的进程能够讲明它希望交谈的对象。另外,还需要有一种机制来调节信息流,使高速主机不会过快地向低速主机传送数据。尽管主机之间的流量控制与 IMP 之间的流量控制不尽相同,但是类似的原理对二者都适用。

5) 会话层

会话层(session layer)允许在不同机器上的用户之间建立会话关系。会话层允许进行类似传输层的普通数据的传送,在某些场合还提供了一些有用的增强型服务。允许用户利用一次会话在远端的分时系统上登录,或者在两台机器间传递文件。会话层提供的服务之一是管理对话控制。会话层允许信息同时双向传输,或任一时刻只能单向传输。若属于后者,则类似于物理信道上的半双工模式,会话层将记录此时该轮到哪一方。另一种与对话控制有关的服务是令牌管理(token management)。有些协议保证双方不能同时进行同样的操作,这一点很重要。为了管理这些活动,会话层提供了令牌,令牌可以在会话双方之间移动,只有持有令牌的一方可以执行某种关键性操作。另一种会话层服务是同步。如果在平均每小时出现一次大故障的网络上,两台机器间要进行一次两小时的文件传输,想想会出现什么样的问题?每一次传输中途失败后,都不得不重新传送这个文件。当网络再次出现大故障时,可能又会半途而废。为了解决这个问题,会话层提供了一种方法,即在数据中插入同步点。每次网络出现故障后,仅仅重传最后一个同步点以后的数据。

6) 表示层

表示层(presentation layer)完成某些特定的功能,对这些功能人们常常希望找到普遍的解决办法,而不必由每个用户自己来实现。值得一提的是,表示层以下各层只关心从源端机到目标机可靠地传送比特流,而表示层关心的是所传送的信息的语法和语义。表示层服务的一个典型例子是用一种大家一致选定的标准方法对数据进行编码。大多数用户程序之间并非交换随机的比特流,而是交换诸如人名、日期、货币数量和发票之类的信息。这些对象是用字符串、整型数、浮点数的形式,以及由几种简单类型组成的数据结构来表示。网络上计算机可能采用不同的数据表示,所以需要在数据传输时进行数据格式的转换。例如,在不同的机器上常用不同的代码来表示字符串(ASCII 和 EBCDIC)、整型数(二进制反码或补码)以及机器字的不同字节顺序等。为了让采用不同数据表示法的计算机之间能够相互通信并交换数据,我们在通信过程中使用抽象的数据结构(如抽象语法表示 ASN.1)来表示传送的数据,而在机器内部仍然采用各自的标准编码。管理这些抽象数据结构,并在发送方将

机器的内部编码转换为适合网上传输的传送语法以及在接收方做相反的转换等工作都是由表示层来完成的。另外,表示层还涉及数据压缩和解压、数据加密和解密等工作。

7)应用层

联网的目的在于支持运行于不同计算机的进程进行通信,而这些进程则是为用户完成不同任务而设计的。可能的应用是多方面的,不受网络结构的限制。应用层(application layer)包含大量人们普遍需要的协议。虽然对于需要通信的不同应用来说,应用层的协议都是必需的。例如,PC用户使用仿真终端软件通过网络仿真某个远程主机的终端并使用该远程主机的资源。这个仿真终端程序使用虚拟终端协议将键盘输入的数据传送到主机的操作系统,并接收显示于屏幕的数据。再如,当某个用户想要获得远程计算机上的一个文件拷贝时,他要向本机的文件传输软件发出请求,这个软件与远程计算机上的文件传输进程通过文件传输协议进行通信,这个协议主要处理文件名、用户许可状态和其他请求细节的通信。远程计算机上的文件传输进程使用其他特征来传输文件内容。由于每个应用有不同的要求,应用层的协议集在ISO/OSI模型中并没有定义,但是,有些确定的应用层协议,包括虚拟终端、文件传输和电子邮件等都可作为标准化的候选。

在这里,通过把七层比喻为真实世界收发信的两个老板的图(左为传输端,右为接收端)来更加形象地介绍七层之间的关系,如图2-3所示。

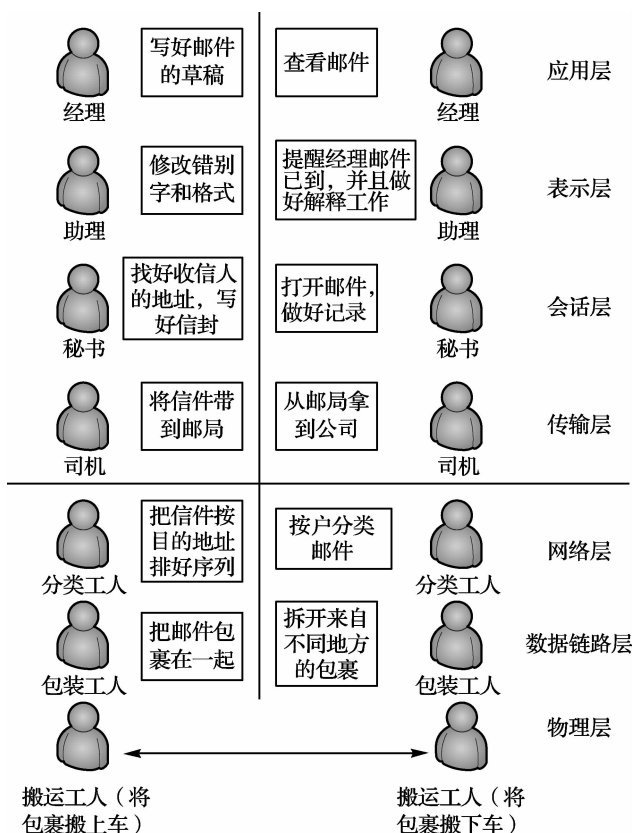


图 2-3 七层结构图与实际生活对比图

最低层物理层相当于邮局的搬运工人,负责搬运邮件。第二层数据链路层相当于邮局的装拆箱工人,负责包裹的拆与包。第三层网络层相当于邮局中的排序工人,按照地址把邮报排好。第四层传输层相当于公司中跑邮局的送信人员。第五层会话层相当于公司中的秘书,负责收寄信、写信封与拆信封。第六层表示层相当于公司助理,负责帮助老板写信收信等。最高层应用层就相当于是老板负责看信和发信件。其实从图 2-3 中可以看出,分界线下面三层协议更多基于设备,上面四层协议更多基于应用软件。

2.1.3 熟悉 TCP/IP 参考模型

OSI 参考模型的诞生为清晰地理解互联网络带来了极大的方便。但是 OSI 过于复杂,难以完全实现。OSI 各层功能具有一定的重复性,效率较低;再加上 OSI 参考模型提出的时候,TCP/IP 协议已经逐渐占据主导地位,因此 OSI 参考模型并没有流行开来,也从来没有存在一个完全遵守 OSI 参考模型的协议簇。

TCP/IP 起源于 20 世纪 60 年代末美国政府资助的一个分组交换网络项目,到 90 年代已发展成为计算机之间最常用的网络协议。它是一个真正的开放系统,因为协议簇的定义及其多种实现可以免费或花很少的钱获得。它已成为“全球互联网”或“因特网”(Internet)的基础协议簇。

与 OSI 参考模型一样,TCP/IP 也采用层次化结构,每一层负责不同的通信功能。但是 TCP/IP 协议简化了层次设计,只分为 4 层——应用层、传输层、网络层和网络接口层,如图 2-4 所示。



动画
四层模型工作
原理

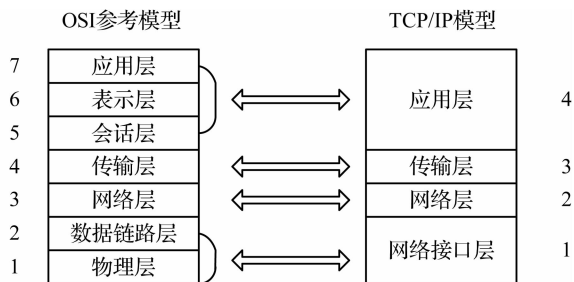


图 2-4 OSI 参考模型与 TCP/IP 模型对比图

通过对比可以清楚地看出 TCP/IP 模型的应用层综合了 OSI 参考模型中的应用层、表示层、会话层。传输层和网络层还是分别对应 OSI 参考模型的传输层和会话层,而网络接口层就是 OSI 参考模型中的数据链路层和物理层的集合。

1. 网络接口层

TCP/IP 本身对网络层之下并没有严格的描述,但是 TCP/IP 主机必须使用某种下层协议连接到网络,以便进行通信。而且,TCP/IP 必须运行在多种下层协议上,以便实现端到端的网络通信。TCP/IP 的网络接口层正是负责处理与传输介质相关的细节,为上层提供一致的网络接口。因此,TCP/IP 模型的网络接口层大体对应于 OSI 模型的数据链路层和物理层,通常包括计算机和网络设备的接口驱动程序和网络接口卡等。

TCP/IP 可以基于大部分局域网和广域网技术运行,这些协议便可以划分到网络接口层。典型的网络接口层技术包括常见的以太网、FDDI(fiber distributed data interface,光纤

分布式数据接口)和令牌环(token ring)等局域网技术,用于串行连接的 SLIP(serial line IP, 串行线路 IP)、HDLC(high-level data link control,高级数据链路控制)和 PPP(point-to-point protocol,点到点协议)等技术,以及常见的 X.25、帧中继(frame relay)和 ATM(asynchronous transfer mode,异步传输模式)等分组交换技术。

2. 网络层

网络层是 TCP/IP 体系的关键部分,它的主要功能是使主机能够将信息发往任何网络并传送到正确的目标。

基于这些要求,网络层定义了主要包格式及其协议——IP(Internet protocol,互联网协议)。网络层使用 IP 地址(IP address)标识网络节点;使用路由协议(routing protocol)生成路由信息,并且根据这些路由信息实现包的转发,使包能够准确地发送到目的地;使用 ICMP、IGMP 这样的协议管理网络。TCP/IP 网络层在功能上与 OSI 网络层极其相似。

3. 传输层

传输层主要为两台主机上的应用程序提供端到端的连接,使源、目的端主机上的对等实体可以进行回话。

在 TCP/IP 协议族的传输层协议主要包括 TCP(transmission control protocol)和 UDP(user datagram protocol)。其中 TCP 是面向连接的,可以保证通信两端的可靠传递,支持乱序恢复、差错重传和流量控制。而 UDP 是无连接的,它提供非可靠性数据传输,数据传输的可靠性由应用层保证。

4. 应用层

TCP/IP 模型没有单独的会话层和表示层,其功能融合在 TCP/IP 应用层中,应用层直接与用户和应用程序打交道,负责对软件提供接口以使程序能使用网络服务。这里的网络服务包括文件传输、文件管理、电子邮件的消息处理等。典型的应用层协议包括 Telnet、FTP、SMTP、SNMP 等。

Telnet(telecommunications network,远程登录)的名字具有双重含义,既指这种应用,也指协议自身。Telnet 给用户提供了一种通过连网的终端登录远程服务器的方式。

FTP(file transfer protocol,文件传输协议)是用于文件传输的 Internet 标准。FTP 支持文本文件(如 ASCII、二进制等)和面向字节流的文件结构。FTP 使用传输层协议 TCP 在支持 FTP 的终端系统间执行文件传输,因此,FTP 被认为提供了可靠的面向连接的文件传输能力,适合于远距离、可靠性较差的线路上的文件传输。

SMTP(simple mail transfer protocol,简单邮件传输协议)支持文本邮件的 Internet 传输。所有的操作系统具有使用 SMTP 收发电子邮件的客户端程序,绝大多数 Internet 服务提供者使用 SMTP 作为其输出邮件服务的协议。SMTP 被设计成在各种网络环境下进行电子邮件信息的传输。实际上,SMTP 真正关心的不是邮件如何被传送,而是关心邮件能够顺利到达目的地。SMTP 具有健壮的邮件处理特性,这种特性允许邮件依据一定标准自动路由。SMTP 具有当邮件地址不存在时立即通知用户的能力,并且具有把在一定时间内不可传输的邮件返回发送方的特点。

SNMP(simple network management protocol,简单网络管理协议)负责网络设备监控和维护,支持安全管理、性能管理等。

2.2 扩展知识

2.2.1 IP 地址规划与划分

网络上的每台主机都有一个唯一的 IP 地址。IP 协议就是使用这个地址在主机之间传递信息的,这是计算机网络能够运行的基础。IP 地址的长度为 32 位,分为 4 段,每段 8 位,用十进制数字表示,每段数字范围为 0~255,段与段之间用句点隔开。例如,159. 226. 1. 1。IP 地址由两部分组成,一部分为网络地址,另一部分为主机地址。IP 地址分为 A、B、C、D、E 共 5 类。

IP 地址就像是我们的家庭住址一样,如果你要写信给一个人,就需要知道他(她)的地址,这样邮递员才能把信送到。计算机发送信息是就好比是邮递员送信,它必须知道唯一的“家庭地址”才不至于把信送错。只不过我们的地址是用文字来表示的,而计算机的地址用十进制数字表示。

众所周知,在电话通信中,电话用户是靠电话号码来识别的。同样,在网络中为了区别不同的计算机,也需要给计算机指定一个联网专用号码,这个号码就是“IP 地址”。

最初设计互联网络时,为了便于寻址及层次化构造网络,每个 IP 地址包括两个标识码(ID),即网络 ID 和主机 ID。同一个物理网络上的所有主机都使用同一个网络 ID,网络上的一个主机(包括网络上的工作站、服务器和路由器等)有一个主机 ID 与其对应。Internet 委员会定义了 5 种 IP 地址类型以适应不同容量的网络,即 A 类~E 类。

其中 A、B、C 三类(见表 2-1)由 InternetNIC 在全球范围内统一分配,D、E 类为特殊地址。

表 2-1 A、B、C 三类 IP 地址的分配情况

网络类别	最大网络数	第一个可用的网络号	最后一个可用的网络号	每个网络中的最大主机数
A	126	1	126	16 777 214
B	16 382	128. 1	191. 255	65 534
C	2 097 150	192. 0. 1	223. 255. 255	254

(1)A 类 IP 地址。一个 A 类 IP 地址是指,在 IP 地址的四段号码中,第一段号码为网络号码,剩下的三段号码为本地计算机的号码。如果用二进制表示 IP 地址,那么 A 类 IP 地址就由 1 字节的网络地址和 3 字节主机地址组成,网络地址的最高位必须是 0。A 类 IP 地址中网络的标识长度为 8 位,主机标识的长度为 24 位,A 类网络地址数量较少,可以用于主机数达 1 600 多万台的大型网络。

A 类 IP 地址的地址范围为 1. 0. 0. 1~126. 255. 255. 255(二进制表示为:00000001 00000000 00000000 00000001~01111110 11111111 11111111 11111111)。

A 类 IP 地址的子网掩码为 255. 0. 0. 0,每个网络支持的最大主机数为 $256^3 - 2 = 16\,777\,214$ 台。

(2) B类 IP 地址。一个 B 类 IP 地址是指在 IP 地址的四段号码中,前两段号码为网络号码。如果用二进制表示 IP 地址,B 类 IP 地址就由 2 字节的网络地址和 2 字节主机地址组成,网络地址的最高位必须是 10。B 类 IP 地址中网络的标识长度为 16 位,主机标识的长度为 16 位,B 类网络地址适用于中等规模的网络,每个网络所能容纳的计算机数为 6 万多台。

B 类 IP 地址的地址范围为 128. 1. 0. 0~191. 255. 255. 255(二进制表示为:10000000 00000001 00000000 00000001~10111111 11111111 11111111 11111111)。

B 类 IP 地址的子网掩码为 255. 255. 0. 0,每个网络支持的最大主机数为 $256 \times 256 - 2 = 65\ 534$ 台。

(3) C 类 IP 地址。一个 C 类 IP 地址是指在 IP 地址的四段号码中,前三段号码为网络号码,剩下的一段号码为本地计算机的号码。如果用二进制表示 IP 地址,C 类 IP 地址就由 3 字节的网络地址和 1 字节主机地址组成,网络地址的最高位必须是 110。C 类 IP 地址中网络的标识长度为 24 位,主机标识的长度为 8 位,C 类网络地址数量较多,适用于小规模局域网,每个网络最多只能包含 254 台计算机。

C 类 IP 地址的地址范围为 192. 0. 1. 1~223. 255. 255. 255(二进制表示为:11000000 00000000 00000001 00000001~11011111 11111111 11111111 11111111)。

C 类 IP 地址的子网掩码为 255. 255. 255. 0,每个网络支持的最大主机数为 $256 - 2 = 254$ 台。

(4) 特殊的地址。1110 开始的地址都称多点广播地址。因此,任何第一个字节大于 223 小于 240 的 IP 地址(范围为 224. 0. 0. 1~239. 255. 255. 254)都是多点广播地址。

每一个字节都为 0 的地址(0. 0. 0. 0)对应于当前主机。

IP 地址中的每一个字节都为 1 的 IP 地址(255. 255. 255. 255)是当前子网的广播地址。

IP 地址中凡是以 11110 开头的 E 类 IP 地址都保留用于将来或实验使用。

IP 地址中不能以十进制 127 作为开头,该类地址中数字 127. 0. 0. 1~127. 1. 1. 1 用于回路测试,例如,127. 0. 0. 1 可以代表本机 IP 地址,用 `http://127. 0. 0. 1` 就可以测试本机中配置的 Web 服务器。

网络 ID 的第一个 6 位组也不能全置为 0,全 0 表示本地网络。

1. IP 子网划分的需求背景

早期的计算机网络是一个简单的二级网络结构。接入网络的结构由一个物理网络构成,该物理网络包括机构中需要接入网络的全部主机。采用自然分类法将 IP 地址划分为 A、B、C、D、E 类。每个 32 位的 IP 地址都被划分为由网络号和主机号构成的二级结构。为每个机构分配一个按照自然分类法得到的 Internet 网络地址,能够很好地适应当时的网络结构。随着时间的推移,网络计算逐渐成熟,网络的优势被许多大型组织认知,网络中出现了很多大型的接入机构。这些机构中需要接入的主机数量众多,单一物理网络容纳主机的数量有限,因此在同一机构内部需要划分多个物理网络。

早期解决这类大型机构接入网络的方法是为机构内的每一个物理网络划分一个逻辑网络,即对每一个物理网络都分配一个按照自然分类法得到的网络地址,称为“物理网络-自然分类 IP 网段”设计方案,如图 2-5 所示。

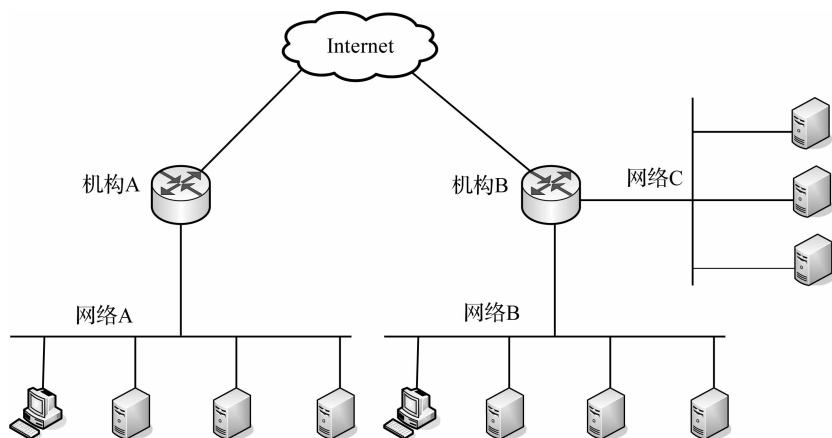


图 2-5 “物理网络-自然分类 IP 网段”设计方案

但是这种“物理网络-自然分类 IP 网段”分配方案存在严重的问题,主要有如下几个。

1) IP 地址资源严重浪费

举例来说,一个公司只有 1 个物理网络,其中需要 300 个 IP 地址。一个 C 类地址能提供 254 个主机地址,不满足需要,因此需要使用一个 B 类地址。一个 B 类地址能提供 65 534 个 IP 地址,网络中的地址得不到充分利用,大量的 IP 地址被浪费。

2) IP 网络数量不够用

举例来说,一个公司拥有 100 个物理网络,每个网络只需要 10 个 IP 地址。虽然需要的地址量仅有 1 000 个,但该公司仍然需要 100 个 C 类地址。很多机构都面临类似的问题,其结果是,在 IP 地址被大量浪费的同时,IP 网络数量却不能满足网络的发展需要。

3) 业务扩展缺乏灵活性

举例来说,一个公司拥有 1 个 C 类网络,其中只有 10 个地址被使用。该公司需要增加一个物理网络,就需要向 IANA 申请一个新的 C 类地址,在得到这个合法的网络地址前,它们就无法部署这个网络接入 Internet。这显然无法满足企业发展的灵活性需求。

综上所述,仅依靠自然分配的 IP 地址分配方案,对 IP 地址进行简单的两层划分,无法应对 Internet 爆炸式的增长需求。

2. 子网划分基础知识

20 世纪 80 年代中期,IETF 在 RFC950 和 RFC917 中针对简单的两层结构 IP 地址所带来的日趋严重的问题提出了解决方法,即子网划分(subnetting)。

子网划分技术又称子网寻径(subnet routing),英文简称为 subnetting,是最广泛使用的 IP 网络地址复用方式,目前已经标准化,并成为 IP 地址模式的一部分。

32 位的 IP 地址分为两部分,即网络号和主机号,分别把它们叫作 IP 地址的“网络部分”和“本地部分”。子网编址技术将“本地部分”进一步划分为“物理网络”和“主机”两部分,其中“物理网络”部分用于标识同一 IP 网络地址下的不同物理网络,常称为“掩码位”“子网掩码号”,或者“子网掩码 ID”,不同子网就是依据这个掩码 ID 来识别的。

按 IP 协议的子网标准规定,每一个使用子网的网点都选择一个 32 位的位模式,若位模式中的某位置为 1,则对应 IP 地址中的某位为网络地址(包括网络部分和子网掩码号)中的

一位;若位模式中的某位置为 0,则对应 IP 地址中的某位为主机地址中的一位。

例如,二进制位模式 11111111 11111111 11111111 00000000 中,前 3 个字节全为 1,代表对应 IP 地址中最高的 3 个字节为网络地址;后一个字节全为 0,代表对应 IP 地址中最后的一个字节为主机地址。为了使用的方便,常常使用“点分十进制数表示法”来表示一个 IP 地址和子网掩码。例如,C 类地址子网掩码(11111111 11111111 11111111 00000000)为 255.255.255.0。



图文
子网掩码

3. 子网掩码的划分

如果要将一个网络划分成多个子网,如何确定这些子网的子网掩码和 IP 地址中的网络号和主机号呢?子网划分的步骤如下。

第 1 步,将要划分的子网数目转换为 2 的 m 次方,例如,要分 8 个子网,则 $8=2^3$ 。如果不是刚好是 2 的 n (n 为任意整数)次方,则以取大为原则,要划分为 6 个,则同样要考虑 2^3 。

第 2 步,将上一步确定的幂 m 按高序占用主机地址 m 位后,转换为十进制,例如, m 为 3,则表示主机位中有 3 位被“网络标识号”占用,因网络标识号应全为 1,所以主机号对应的字节段为 11100000。转换成十进制后为 224,这就最终确定了子网掩码。若是 C 类网,则子网掩码为 255.255.255.224;如果是 B 类网,则子网掩码为 255.255.224.0;若是 A 类网,则子网掩码为 255.224.0.0。

在这里,子网个数与占用主机地址位数有如下关系式成立: $2^m \geq n$,其中 m 表示占用主机地址的位数; n 表示划分的子网个数。根据这些原则,将一个 C 类网络分成 4 个子网。

为了说明问题,现再举例。若使用的网络号为 192.9.200,则该 C 类网内的主机 IP 地址就是 192.9.200.1~192.9.200.254,现将网络划分为 4 个子网,按照以上步骤有 $4=2^2$,则表示要占用主机地址的 2 个高序位,即为 11000000,转换为十进制为 192。这样就可确定该子网掩码为:192.9.200.192。4 个子网的 IP 地址的划分是根据被网络号占住的两位排列进行的,这 4 个 IP 地址范围分别为:

第 1 个子网的 IP 地址是从 11000000 00001001 11001000 00000001 到 11000000 00001001 11001000 00111110,注意它们的最后 8 位中被网络号占住的两位都为 00,因为主机号不能全为 0 和 1,所以没有 11000000 00001001 11001000 00000000 和 11000000 00001001 11001000 00111111 这两个 IP 地址(下同)。

注意: 实际上此时的主机号只有最后面的 6 位。对应的十进制 IP 地址范围为 192.9.200.1~192.9.200.62,而这个子网的子网掩码(或网络地址)为 11000000 00001001 11001000 00000000,即 192.9.200.0。

第 2 个子网的 IP 地址是从 11000000 00001001 11001000 01000001 到 11000000 00001001 11001000 01111110,注意此时被网络号所占用的 2 位主机号为 01。对应的十进制 IP 地址范围为 192.9.200.65~192.9.200.126。对应这个子网的子网掩码(网络地址)为 11000000 00001001 11001000 01000000,即 192.9.200.64。

第 3 个子网的 IP 地址是从 11000000 00001001 11001000 10000001 到 11000000 00001001 11001000 10111110,注意此时被网络号所占用的 2 位主机号为 10。对应的十进制 IP 地址范围为 192.9.200.129~192.9.200.190。对应这个子网的子网掩码(网络地址)为 11000000 00001001 11001000 10000000,即 192.9.200.128。

第 4 个子网的 IP 地址是从 11000000 00001001 11001000 11000001 到 11000000 00001001 11001000 11111110, 注意此时被网络号所占住的 2 位主机号为 11。对应的十进制 IP 地址范围为 192. 9. 200. 193~192. 9. 200. 254。对应这个子网的子网掩码(网络地址)为 11000000 00001001 11001000 11000000, 即 192. 9. 200. 192。

4. 子网掩码的计算

1) 利用子网数来计算

在求子网掩码之前必须先弄清楚要划分的子网数目及每个子网内所需的主机数目。然后按以下基本步骤进行计算。

第 1 步, 将子网数目转化为二进制来表示。

第 2 步, 取得子网数二进制的位数(n)。

第 3 步, 取得该 IP 地址类的子网掩码, 然后将其主机地址部分的前 n 位置 1, 即得出该 IP 地址划分子网的子网掩码。

为了便于理解, 现举例说明如下: 假如要将一个 B 类 IP 地址 168. 195. 0. 0 划分成 27 个子网, 则它的子网掩码的划分方法如下(对应以上各基本步骤)。

第 1 步, 首先要将网络划分成 27 个子网, 27 的二进制为 11011。

第 2 步, 因为 $2^5 > 27$, 所以该子网数二进制为五位数, 即 $n=5$ 。

第 3 步, 将该 B 类地址的子网掩码 255. 255. 0. 0 的主机号前 5 位全部置 1, 即可得到 255. 255. 248. 0, 这就是划分成 27 个子网的 B 类 IP 地址 168. 195. 0. 0 的子网掩码。

2) 利用主机数来计算

利用主机数来计算子网掩码的方法与上类似, 基本步骤如下。

第 1 步, 将子网中需容纳的主机数转化为二进制。

第 2 步, 若主机数大于或等于 254(因为要去掉保留的两个 IP 地址), 则取得该主机的二进制位数为 n , 这里肯定有 $n > 8$, 也就是说主机地址将占据不止 8 位。

第 3 步, 将 255. 255. 255. 255 的主机地址位数全部置 1, 然后从后向前将 n 位全部置为 0, 即为子网掩码值。

举例如下。如要将一个 B 类 IP 地址为 168. 195. 0. 0 的网络划分成若干子网, 要求每个子网内有主机数为 700 台, 则该子网掩码的计算方法如下(也是对应以上各基本步骤)。

第 1 步, 首先将子网中要求容纳的主机数 700 转换成二进制, 得到 1010111100。

第 2 步, 计算出该二进制的位数为 10 位, 即 $n=10$ 。

第 3 步, 将 255. 255. 255. 255 从后向前的 10 位全部置 0, 得到的二进制数为 11111111. 11111111. 11111100. 00000000, 转换成十进制后即 255. 255. 252. 0, 这就是该要划分成主机数为 700 的 B 类 IP 地址 168. 195. 0. 0 的子网掩码。

3) 子网 ID 增量计算法

子网 ID 增量计算法的基本计算步骤如下。

第 1 步, 将所需的子网数转换为二进制, 如所需划分的子网数为 4, 则转换成二进制为 00000100。

第 2 步, 取子网数的二进制中有效位数, 即为向默认子网掩码中加入的位数(向主机 ID 中借用的位数), 如前面的 00000100, 有效位为 100, 为 3 位。

第 3 步, 决定子网掩码, 若 IP 地址为 B 类 1129. 20. 0. 0 网络, 则默认子网掩码为

255.255.0.0,借用主机 ID 的 3 位以后变为 255.255.224(11100000).0,即将所借的位全表示为 1,用作子网掩码。

第 4 步,将所借位的主机 ID 的起始位段最右边的 1 转换为十进制,即为每个子网 ID 之间的增量,例如,前面的借位的主机 ID 起始位段为 11100000,最右边的 1,转换成十进制后为 $2^5=32$ 。

第 5 步,产生的子网 ID 数为 2^m-2 (m 为向默认子网掩码中加入的位数),例如,本例向子网掩码中添加的位数为 3,则可用子网 ID 数为 $2^3-2=6$ 个。

第 6 步,将上面产生的子网 ID 增量附在原网络 ID 之后的第一个位段,便形成了第一个子网网络 ID:129.20.32.0。

第 7 步,重复上步操作,在原子网 ID 基础上加上一个子网 ID 增量,以此类推,直到子网 ID 中的最后位段为默认子网掩码位用主机 ID 位之后的最后一个位段值,这样就可得到所有的子网网络 ID,例如,默认子网掩码位用主机 ID 位之后的子网 ID 为 255.255.224.0,其中的 224 为借用主机 ID 后子网 ID 的最后一位段值,所以当子网 ID 通过以上增加增量的方法得到 129.20.224.0 时便终止,不必再添加。

2.2.2 IPv6 简介



图文
IPv4 与 IPv6
地址

IPv4(Internet protocol version 4,因特网协议版本 4)是因特网所使用的网络层协议。自 20 世纪 80 年代以来,IPv4 一直在因特网上良好、稳定地运行着。但是,IPv4 协议之初是为几百台计算机组成的小型网络设计的,随着因特网及其所提供的服务突飞猛进地发展,IPv4 已经暴露出一些不足之处。IPv6(Internet protocol version 6,因特网协议版本 6)也称为 IPng(IP next generation,下一代因特网协议),是 IETF(Internet Engineering Task Force,互联网工程任务组)设计的用于替代现行版本 IP 协议的下一代 IP 协议,是 IPv4 的升级版。IPv6 和 IPv4 最大的区别是,IP 的长度由现在的 32 位增加为 128 位。除此以外,IPv6 还在安全性、QoS 等方面进行了增强。

1. IPv6 的特点

实践证明,IPv4 是一个非常成功的协议,它本身也经受住了因特网从最初数目很少的计算机发展到目前上亿台计算机互联的考验。但是,IPv4 协议也不是十全十美的,随着因特网规模的快速扩张,逐渐地暴露出了一些问题。其中最严重的问题是 IPv4 可用地址的日益缺乏。

数据显示,目前的 IP 地址已于 2011 年 2 月 3 日分配完毕。其中北美占有 3/4,约 30 亿个,而人口最多的亚洲只有不到 4 亿个,截至 2010 年 6 月,中国拥有 IPv4 地址数量达到 2.5 亿,落后于 4.2 亿网民的需求。地址不足严重地制约了中国及其他国家互联网的应用和发展。一方面是地址资源数量的限制;另一方面是随着电子技术及网络技术的发展,计算机网络将进入人们的日常生活,可能身边的每一样东西都需要连入全球因特网。在这样的环境下,IPv6 应运而生。单从数量级上来说,IPv6 所拥有的地址容量是 IPv4 的约 8×10^{28} 倍,达到 2^{128} (算上全零的)个。这不但解决了网络地址资源数量的问题,同时也为除计算机外的设备连入互联网在数量限制上扫清了障碍。

IPv6 协议最大的特点是几乎无限的地址空间。IPv4 地址的位数是 32 位,但在 IPv6 中,地址的位数增长了 4 倍,达到 128 位。所以 IPv6 地址空间大得惊人。IPv4 中,理论上可编址的节点数是 2^{32} ,也就是 4 294 967 296,按照目前的全世界人口数,大约每 3 个人可以拥

有 2 个 IP 地址。而 IPv6 的 128 位长度的地址意味着 3.4×10^{38} 个地址。世界上每个人可以拥有 5.7×10^{28} 个 IPv6 地址。这个地址量是非常巨大的,有个夸张的说法,可以为全世界每一粒沙子都分配一个 IPv6 的地址。

IPv6 普及的一个重要的应用是网络实名制下的互联网身份证(VIeID)。目前基于 IPv4 的网络之所以难以实现网络实名制,一个重要原因就是 IP 资源的共用,因为 IP 资源不够,所以不同的人在不同的时间段共用一个 IP,IP 和上网用户无法实现一一对应。在 IPv4 下,根据 IP 查人也比较麻烦,电信局要保留一段时间的上网日志才行,通常因为数据量很大,运营商只保留 3 个月左右的上网日志,比如查前年某个 IP 发帖子的用户就不能实现。IPv6 的出现可以从技术上一劳永逸地解决实名制这个问题,因为到时 IP 资源将不再紧张,运营商有足够多的 IP 资源,运营商在受理入网申请的时候,可以直接给该用户分配一个固定 IP 地址,这样实际上就实现了实名制,也就是一个真实用户和一个 IP 地址的一一对应。

同时,IETF 在制定 IPv6 时,还考虑到了在 IPv6 中需要解决的其他一些 IPv4 协议中存在的问题,如配置不够简便、安全性差、QoS 功能弱等。从而使协议本身能够适应目前网络发展的需要。

2. IPv6 地址的表示方式

在 IPv4 中,地址是用 192.168.1.1 这种点分十进制方式来表达的。但在 IPv6 中,地址有 128 位,如果再用十进制表示的话就太长了。所以,IPv6 采用冒号十六进制表示法来表示地址。

IPv6 地址的 128 位被分成 8 段,每 16 位为一段,每段被转换为一个 4 位十六进制数,并用冒号隔开。例如,1234:5678:90AB:CDEF:ABCD:EF01:2345:6789。

为了尽量缩短地址的书写长度,IPv6 地址可以采用压缩方式来表示。在压缩时,有以下几个原则。

(1)每段中前导 0 可以去掉,但保证每段至少有一个数字,如 2012:0501:0000:0001:0000:0000:0001:E001 就可以压缩为 2012:0501:0:1:0:0:1:E001。但有效位 0 不能压缩。所以上述地址不能压缩为 212:51:0:1:0:0:1:E1。

(2)一个或多个连续的段内各位为 0 的时候可用双冒号(::)压缩表示,但一个 IPv6 中只允许有一个双冒号,如 2012:0501:0000:0001:0000:0000:0001:E001 就可以压缩为 2012:0501::1:0:0:1:E001 或 2012:0501:0:1::1:E001,但是不能表示为 2012:0501::1::1:E001。

IPv6 取消了 IPv4 的网络号、主机号和子网掩码的概念,代之以前缀、接口标识符、前缀长度;IPv6 也不再具有 IPv4 地址中的 A 类、B 类、C 类等地址分类的概念。

- ①前缀:前缀的作用与 IPv4 地址中的网络部分类似,用于标识该地址属于哪个网络。
- ②接口标识符:与 IPv4 地址中的主机部分类似,用于标识该地址在网络中的具体位置。
- ③前缀长度:作用类似于 IPv4 地址中的子网掩码,用于确定地址中哪一部分是前缀,哪一部分是接口标识符。

例如,地址 2012:0501:0000:0001:0000:0000:0001:E001/64,其中/64 表示此地址的前缀长度是 64 位,所以此地址的前缀就是 2012:0501:0000:0001,接口标识符就是 0000:0000:0001:E001。

3. IPv6 地址的分类

IPv4 地址包括单播、组播、广播等几种类型。与其类似,IPv6 地址也有不同类型,包括单播地址、组播地址和任播地址。IPv6 地址中没有广播地址,在 IPv4 协议中某些需要用到的广播地址的服务或功能,IPv6 协议都是用组播地址来完成。

1) 单播地址

用来唯一标识一个接口,类似于 IPv4 中的单播地址。单播地址只能分配给一个节点上的一个接口,发送到单播地址的数据报文将被传送给地址所标识的接口。IPv6 单播地址根据其作用范围的不同,又可分为链路本地地址、站点本地地址、全球单播地址等,还包括一些特殊地址,如未指定地址和环回地址。

(1)链路本地地址:用于链路本地节点之间的通信。在 IPv6 中,以路由器为边界的一个或多个局域网段称之为链路。使用链路本地地址作为目的地址的数据报文不会被转发到其他链路上。其前缀标识为 FE80::/10。

(2)站点本地地址:与 IPv4 中的私有地址类似。使用站点本地地址作为目的地址的数据报文不会被转发到本站点(相当于一个私有网络)外的其他站点。其前缀标识为 FEC0::/10。站点本地地址在实际应用中很少使用。

(3)全球单播地址:与 IPv4 中的公有地址类似,全球单播地址由 IANA 负责进行统一分配。全球单播地址的前缀表示是 2000::/3。

(4)未指定地址:地址“::”称为未指定地址,不能分配给任何节点。在节点获得有效的 IPv6 地址之前,可在发送的 IPv6 报文的源地址字段填入该地址,表示目前暂无有效地址,未指定地址不能作为 IPv6 报文中的目的地址。

(5)环回地址:单播地址“::1”称为环回地址,不能分配给任何物理接口。它的作用与 IPv4 中的环回地址 127.0.0.1 相同,节点可通过给自己发送 IPv6 报文而测试协议是否工作正常。

2) 组播地址

用来标识一组接口,类似于 IPv4 的组播地址。多个接口可配置相同的组播地址,发送到组播地址的数据报文被发送给此地址所标识的所有接口。IPv6 组播地址的范围是 FF00::/8。

3) 任播地址

任播地址是 IPv6 中特有的地址类型,也用来标识一组接口。当与组播地址不同的是,发送到任播地址的数据报文被传送给此地址所标识的一组接口中距离源节点最近的一个接口。例如,移动用户在使用 IPv6 协议接入因特网时,根据地理位置的不同,接入距离用户最近的一个接收站。

4. IPv6 的优势

与 IPV4 相比,IPV6 具有以下几个优势。

(1)IPv6 具有更大的地址空间。IPv4 中规定 IP 地址长度为 32,最大地址个数为 2^{32} ;而 IPv6 中 IP 地址的长度为 128,即最大地址个数为 2^{128} 。与 32 位地址空间相比,其地址空间增加了 2^{128} 个。

(2)IPv6 使用更小的路由表。IPv6 的地址分配一开始就遵循聚类(aggregation)的原

则,这使得路由器能在路由表中用一条记录(entry)表示一片子网,大大缩短了路由器中路由表的长度,提高了路由器转发数据包的速度。

(3)IPv6 增加了增强的组播(multicast)支持以及对流的支持(flow control),这使得网络上的多媒体应用有了长足发展的机会,为服务质量(quality of service, QoS)控制提供了良好的网络平台。

(4)IPv6 加入了对自动配置(auto configuration)的支持。这是对 DHCP 协议的改进和扩展,使得网络(尤其是局域网)的管理更加方便和快捷。

(5)IPv6 具有更高的安全性。在使用 IPv6 网络中用户可以对网络层的数据进行加密并对 IP 报文进行校验,在 IPv6 中的加密与鉴别选项提供了分组的保密性与完整性,极大地增强了网络的安全性。

(6)允许扩充。只有在新的技术或应用需要时,IPv6 才允许协议进行扩充。

(7)更好的头部格式。IPv6 使用新的头部格式,其选项与基本头部分开,如果需要,可将选项插入到基本头部与上层数据之间。这就简化和加速了路由选择过程,因为大多数的选项不需要由路由选择。

(8)新的选项。IPv6 有一些新的选项来实现附加的功能。

2.3 实训

2.3.1 常用网络命令的使用

1. 实训目的

掌握常见的几种 TCP/IP 网络命令的使用方法。

2. 实训内容

使用常见的几种 TCP/IP 网络命令。

3. 实训环境的搭建

- (1)PC 一台。
- (2)安装有 TCP/IP 协议的网络操作系统。

4. 实训操作实践与步骤

1) ping 命令

ping 命令是用于测试网络连接状况及信息包发送和接收状况的有用工具,是网络测试最常用的命令。ping 命令向目标主机(地址)发送一个回送请求数据包,要求目标主机收到请求后给予应答,从而判断网络的响应时间和本机是否与目标主机(地址)连通。若执行 ping 命令不成功,则可以预测故障出现在以下几个方面:网线故障、网络适配器配置不正确或 IP 地址不正确。如果执行 ping 命令成功而网络仍无法使用,那么问题很可能出在网络系统的软件配置方面,ping 成功只能保证本机与目标主机间存在一条连通的物理路径。

ping 命令的使用方式如下。

(1) 选择“开始”→“运行”菜单项，弹出“运行”对话框，如图 2-6 所示。



图 2-6 “运行”对话框

(2) 在“打开”的下拉列表框中输入 cmd 并按 Enter 键，打开命令提示符窗口，如图 2-7 所示。

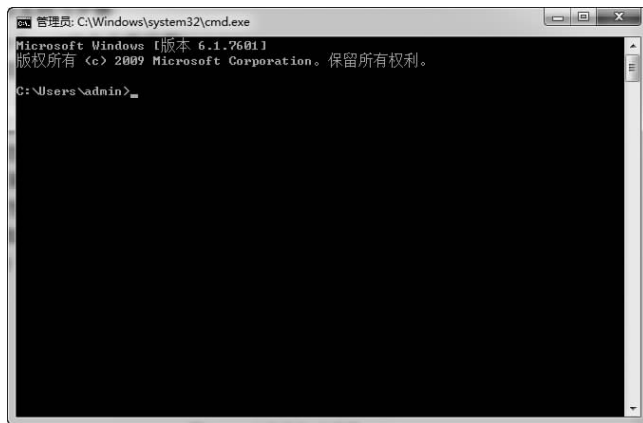


图 2-7 命令提示符窗口

(3) 在命令提示符窗口输入“ping www.163.com”，测试本机到网易服务器是否通行，如图 2-8 所示。

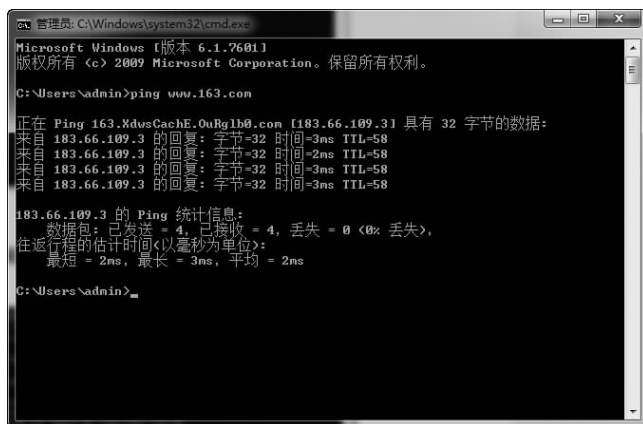


图 2-8 输入 ping 命令测试网络状况

2)tracert 命令

tracert 命令诊断实用程序通过向目标计算机发送具有不同生存时间的 ICMP 数据包,来确定至目标计算机的路由,也就是说,用来跟踪一个消息从一台计算机到另一台计算机所走的路径。该诊断实用程序将包含不同生存时间(TTL)值的 Internet 控制消息协议(ICMP)回显数据包发送。不过,有些路由器悄悄地下传包含过期 TTL 值的数据包,而 tracert 看不到。

tracert 命令的使用方式为:在命令提示符窗口输入“tracert www.163.com”并按 Enter 键,测试本机到网易服务器的路由,如图 2-9 所示。

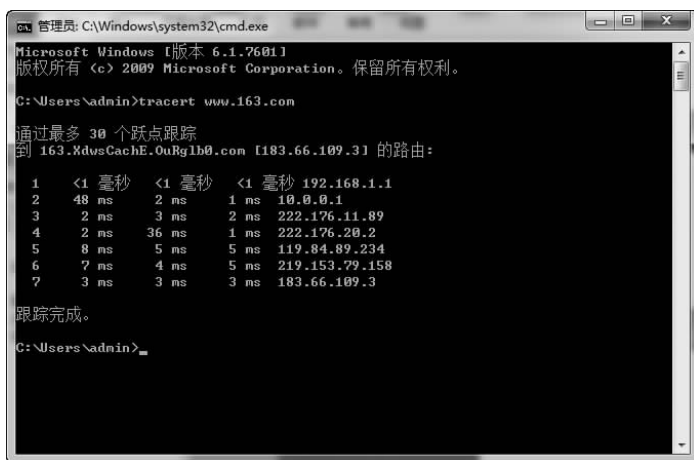


图 2-9 用 tracert 测试到网易服务器的路由

3)ipconfig 命令

ipconfig 命令用于显示系统当前的 IP 配置信息,它在动态地址分配(DHCP)环境中作用明显,显示内容包括当前 IP 地址、子网掩码和默认网关等。通过增加 all 参数,还可显示更多信息。ipconfig 命令的可选参数 release 和 renew 可以使管理员从 DHCP 服务器上动态地释放和重新搜索地址。

ipconfig 命令的使用方式为:在命令提示符窗口输入“ipconfig”并按 Enter 键,查看本地 IP 信息,如图 2-10 所示。

4)arp 命令

arp 命令用于显示和修改“地址解析协议(ARP)”缓存中的记录。ARP 缓存中包含一个或多个表,它们用于存储 IP 地址及其经过解析的网络物理地址。计算机上安装的每一个网络适配器都有自己单独的表。

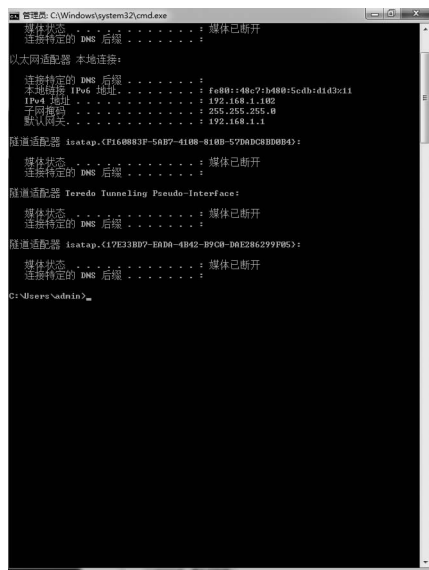


图 2-10 输入 ipconfig 查看本地信息

arp 命令的使用方式如下。

(1) 显示所有接口的 ARP 缓存表, 可使用可选参数“-a”, 如图 2-11 所示。

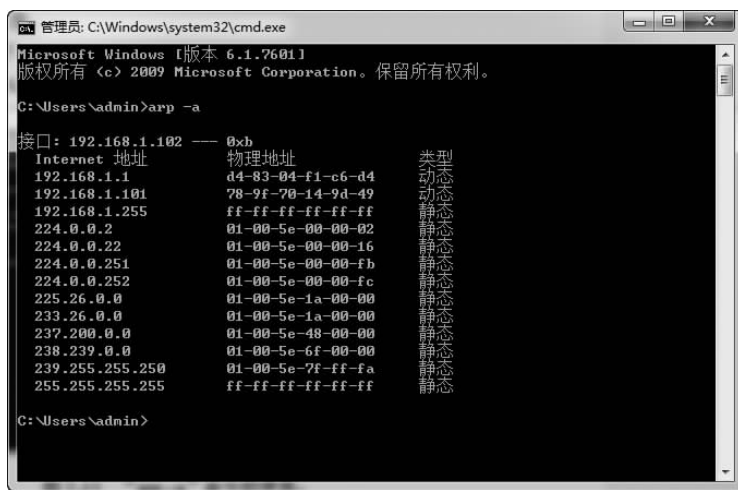


图 2-11 “arp-a”命令的使用

(2) 对于指派的 IP 地址为 192.168.1.101 的接口, 要显示其 ARP 缓存表, 可使用可选参数“-a”, 即输入命令“arp-a 192.168.1.101”并按 Enter 键, 如图 2-12 所示。

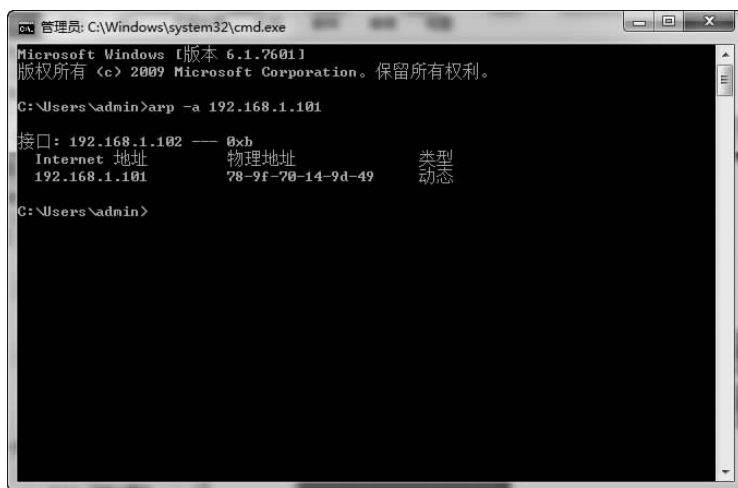


图 2-12 查看指派 IP 地址的 ARP 缓存表

(3) 静态 IP/MAC 绑定。

① 查看网口的 Idx 值。DOS 界面下, 输入命令“netsh i i show in”, 查看本地网卡对应的 Idx 值(后面会使用到该命令), 如图 2-13 所示。

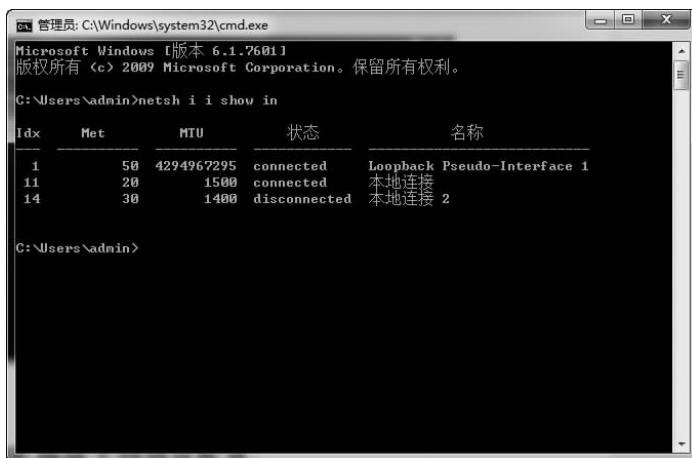


图 2-13 查看本地网卡对应的“Idx”值

②绑定 IP/MAC。DOS 界面下,输入命令“netsh-c "i i" add ne 11 192.168.1.250 00-aa-00-62-c6-09”,绑定 IP 与 MAC,默认是永久生效,即使重启系统还会保存,如图 2-14 所示。

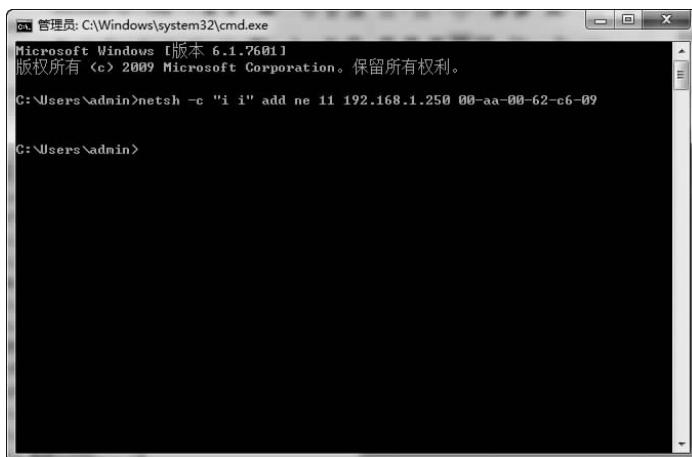


图 2-14 绑定 IP 地址

2.3.2 手动配置 IP 地址

1. 实训目的

掌握手动配置 IP 地址的方法。

2. 实训内容

通过手动配置 IP 地址来上网。

3. 实训环境的搭建

- (1) PC 一台。
- (2) 安装有 TCP/IP 协议的网络操作系统。

4. 实训操作步骤

(1) 在桌面上右击“网络”图标，在弹出的快捷菜单中选择“属性”选项，如图 2-15 所示。

(2) 单击“更改适配器设置”链接，如图 2-16 所示。



图 2-15 选择“属性”选项



图 2-16 单击“更改适配器设置”链接

(3) 打开“网络连接”窗口，如图 2-17 所示。



图 2-17 “网络连接”窗口

(4)在其中双击“本地连接 3”选项,在弹出的“本地连接 属性”对话框的“此连接使用下列项目”列表框中选择“Internet 协议(TCP/IP)”选项,如图 2-18 所示。

(5)单击“属性”按钮,在弹出的对话框中选中“使用下面的 IP 地址”单选按钮,在 IP 地址栏中输入的 IP 地址、子网掩码、默认网关及 DNS 服务器,如图 2-19 所示。

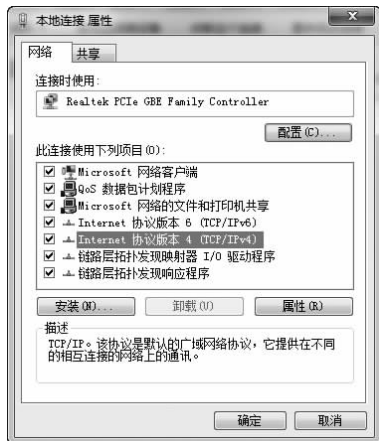


图 2-18 选择 Internet 协议(TCP/IP)

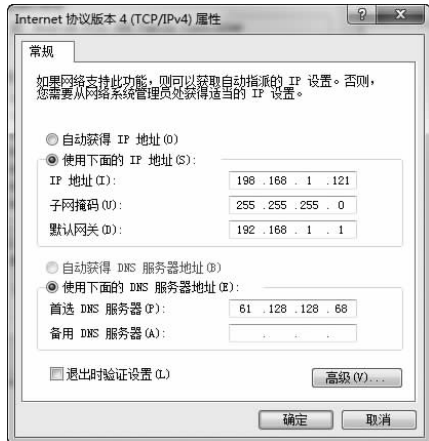


图 2-19 输入 IP 地址等信息

2.4 思考与练习

一、名词解释

1. OSI 参考模型。
2. 协议。
3. SNMP。

二、填空题

1. OSI 参考模型分为 7 层,从下往上分别是 _____、_____、_____、_____、_____、_____和 _____。
2. TCP/IP 参考模型分为 4 层,分别是 _____、_____、_____和 _____。
3. IPv4 的 IP 长度是 _____,IPv6 的 IP 长度是 _____。

三、简答题

1. 计算机网络管理组织有哪些?
2. OSI 七层划分原则是什么?
3. IPv4 地址的分类有哪些?
4. IPv6 的优势有哪些?

3

模块 3

局域网技术

▶ 本模块学习前的思考点

- 什么是局域网?
- 局域网有哪些常见的传输介质?
- 局域网常用的网络设备有哪些?
- 未来局域网的发展趋势是什么?
- 用户终端如何接入局域网?

3.1 必备知识

3.1.1 局域网概念

局域网(local area network, LAN)即计算机局部区域网,它是在一个局部的地理范围内,将各种计算机、外围设备、数据库等互相连接起来组成的计算机通信网。由此可以看出,它包括 3 个方面的含义。

(1)局域网所覆盖的地理范围小,通常是一个办公室、一幢大楼、一个机关、厂矿、公司、学校等,一般不超过几千米。

(2)局域网中所连接的数据通信设备的含义是广义的,它包含在传输介质上进行通信的各种设备,如计算机和各种外部设备等。

(3)局域网是为数据传输而构建的一种通信网络。

3.1.2 局域网的传输方式与介质

数据信号在局域网中传输,有不同的传输方式和不同的传输介质。选用不同的传输介质,对局域网性能的影响是不同的。

1. 局域网的传输方式

局域网中的传输方式有单工、半双工和全双工3种。这些概念源于电话系统的串行数据通信,表示通信双方信息交换的方式,在20世纪90年代初被引入到局域网中,传输方式主要是从时间和方向上区分。

1) 单工

单工即单向通信,如图3-1所示。它是指在一条线路上只能存在单个方向的通信,反向无法进行。单工通信多用于无线广播、电视广播等,在局域网中并不采用。

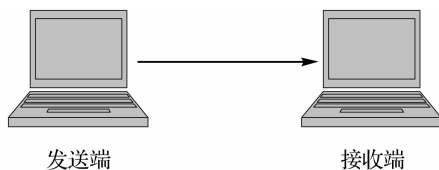


图 3-1 单工通信

2) 半双工

半双工即双向交替通信,如图3-2所示。这是指在同一时间内,通信双方中只能有一方发送信息,另一方只能接收,过一段时间再反过来,如对讲机。局域网中最早使用的传输方式就是这种。

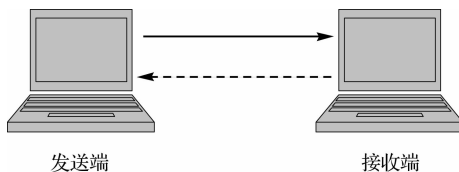


图 3-2 半双工通信

3) 全双工

全双工即双向同时通信,如图3-3所示。这是指在同一时间内,通信双方既可以发送信息,也可以接收信息。全双工传输方式在局域网中的应用,提高了网络的通信能力,现在的计算机网络大部分采用的是全双工方式。

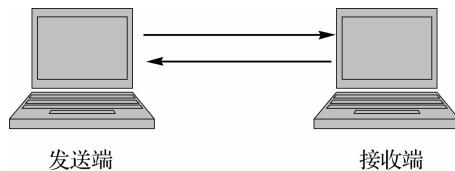


图 3-3 全双工通信

2. 局域网传输介质

传输介质是网络中传输信息的物理通道,它的性能对网络的通信、速度、距离、价格及网络中的节点数和可靠性都有很大影响。因此,必须根据网络的具体要求,选择适当的传输介质。常见的介质有很多种,可分成两大类:一类是有线传输介质,如双绞线、同轴电缆、光纤等;另一类是无线传输介质,如微波、卫星通信、红外线和激光通信等。

1) 双绞线

双绞线(twisted pair, TP)是最常用的一种传输介质,它由两条具有绝缘保护层的铜导线按一定密度互相绞合在一起,可降低信号干扰的程度。为了区分,每根铜线的颜色都不一样。

双绞线通常有屏蔽式和非屏蔽式两种。

(1)非屏蔽双绞线(unshielded twisted pair, UTP)电缆由多对双绞线和一个塑料外皮组成,如图 3-4 所示。

(2)屏蔽双绞线(shielded twisted pair, STP)电缆是在一对双绞线外面有金属筒缠绕,有的还在几对双绞线的外层用铜编织网包上,均用于屏蔽,最外层再包上一层具有保护性的聚乙烯塑料。与非屏蔽双绞线相比,其误码率明显下降,但价格相对较高,安装时也比非屏蔽双绞线电缆困难。屏蔽双绞线如图 3-5 所示。



图 3-4 非屏蔽双绞线

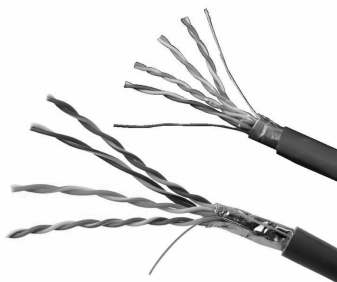


图 3-5 屏蔽双绞线

2) 同轴电缆

同轴电缆由内导体铜质芯线、绝缘层、网状编织的外导体屏蔽层及保护塑料外层组成,如图 3-6 所示。这种结构中的金属屏蔽网可防止中心导体向外辐射电磁场,也可用来防止外界电磁场干扰中心导体的信号,因而具有很好的抗干扰特性,被广泛用于较高速率的数据传输。

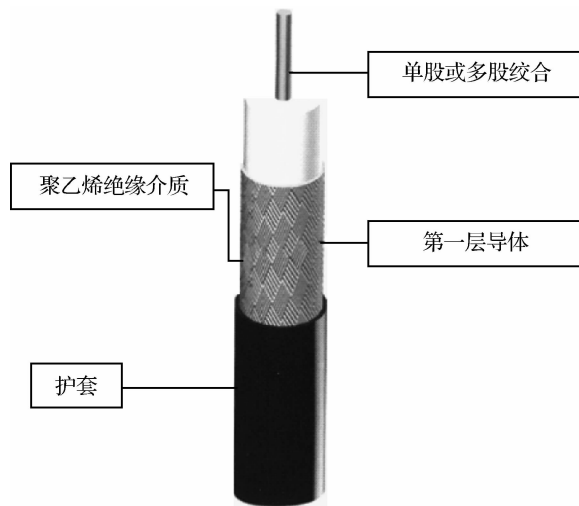


图 3-6 同轴电缆

根据同轴电缆的直径,同轴电缆可以分为粗缆和细缆两种。粗缆价格较贵,但可连接

较多的站点,支持较长距离的传输。细缆功耗损耗较大,一般只用于 500 m 内的数据传输。

通常按特性阻抗数值的不同,可将同轴电缆分为基带同轴电缆和宽带同轴电缆。基带同轴电缆一般只用来传输数据,不使用 Modem,因此较宽带同轴电缆成本较低,适合传输距离较短、速度要求较低的局域网。基带同轴电缆的外导体是用铜做成网状的,特征阻抗为 $50\ \Omega$ 。宽带同轴电缆传输速率较高,距离较远,但成本较高。它不仅能传输数据,还可以传输图像和语音信号。宽带同轴电缆的特性阻抗为 $75\ \Omega$ 。

3) 光纤

光导纤维(optical fiber)简称光纤,是目前发展最为迅速、应用最为广泛的传输介质。它是一种能够传输光束的、细而柔软的通信媒体。光纤通常是由石英玻璃拉成细丝,其结构一般是由双层的同心圆柱体组成,中心部分为纤芯,如图 3-7 所示。

与其他传输介质相比,光缆的电磁绝缘性能好,信号衰变小,频带较宽,传输距离较大。光缆主要是在要求传输距离较长,布线条件特殊的情况下用于主干网的连接。



图 3-7 光纤

4) 地面微波通信

地面微波通信常用于电缆或光缆铺设不便的特殊地理环境或作为地面传输系统的备份和补充。地面微波通信在数据通信中占有重要位置。

微波是一种频率很高的电磁波,其频率为 $300\ \text{MHz}\sim 300\ \text{GHz}$,地面微波通信主要使用的是 $2\sim 40\ \text{GHz}$ 的频率范围。地面微波一般沿直线传输。由于地球表面为曲面,所以,微波在地面的传输距离有限,一般为 $40\sim 60\ \text{km}$ 。但这个传输距离与微波的发射天线的高度有关,天线越高传输距离就越远。为了实现远距离传输,就需要在微波信道的两个端点之间建立若干个中继站,中继站把前一个站点送来的信号经过放大后再传输到下一站。经过这样的多个中继站点的“接力”,信息就被从发送端传输到接收端,如图 3-8 所示。

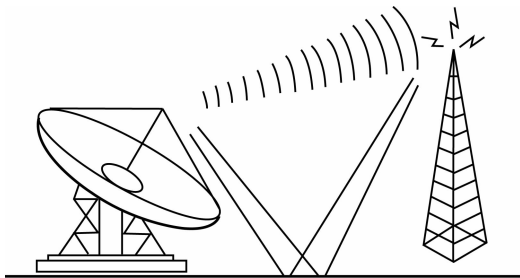


图 3-8 地面微波通信

微波通信具有频带宽、通信容量大、初建费用低、建设速度快、应用范围广等优点,其缺点是保密性能差、抗干扰性能差,两微波站天线间不能被建筑物遮挡。这种通信方式逐渐被很多计算机网络采用,有时在大型互联网中与有线介质混用。

5) 卫星通信

卫星通信实际上是使用人造地球卫星作为中继器来转发信号的,它使用的波段也是微波。卫星通信通常被定位在几万千米的高空,因此,卫星作为中继器可使信息的传输距离很远(几千至上万千米)。例如,每个同步卫星可覆盖地球的 $1/3$ 表面。卫星通信已被广泛应

用于远程计算机网络中。例如，国内很多证券公司显示的证券行情都是通过 VSAT 接收的卫星通信广播信息，而证券的交易信息则是通过延迟小的数字数据网 DDN 专线或分组交换网进行转发的。

卫星通信具有通信容量极大、传输距离远、可靠性高、一次性投资大、传输距离与成本无关等特点。

6) 红外线和激光通信

应用于计算机网络的无线通信除地面微波及卫星通信外，还有红外线和激光通信等。红外线和激光通信的收发设备必须处于视线范围内，具有很强的方向性，因此，防窃取能力强。但由于它们的频率太高，波长太短，不能穿透固体物质，且对环境因素（如天气）较为敏感，因而，只能在室内和近距离使用。

3.1.3 网络设备介绍

当要把两台以上的计算机联成局域网时，不仅需要为每台联网计算机安装网卡，还需要一条通信线路（介质）和介质连接器等附属部件，将计算机与电缆连接在一起。根据不同的联网技术，有时还需要使用集线器、交换机、路由器等网络连接设备，以实现局域网的物理连接。



图片
网卡

1. 网卡

网络互联的每台计算机上都必须安装网卡（network interface card, NIC），网卡是连网的计算机中所需的基本部件。它一方面连接计算机，另一方面连接局域网中的传输介质。典型的网卡结构如图 3-9 所示。

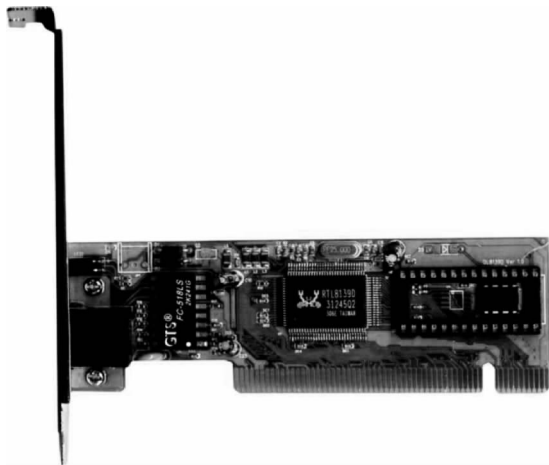


图 3-9 典型的网卡结构

网卡完成物理层和数据链路层的大部分功能，包括网卡与传输介质的物理连接、介质访问控制（CSMA/CD）、数据帧的拆装、帧的发送与接收、错误校验、数据信号的编/解码、数据的串/并行转换等功能。网卡是局域网通信接口的关键设备，它是决定计算机网络性能指标的重要因素之一。

1) 网卡的分类

根据网卡所支持的物理层标准与主机接口的不同，网卡可以分为不同的类型，如以太网卡和令牌环网卡。由于以太网是目前最主要的局域网类型，因而与之对应的网卡——以太

网卡也成了网卡的代名词,如在不特别指明的情况下,以下所描述的网卡分类实际上是指以太网卡。根据网卡与主板上总线的连接方式、网卡的传输速率和网卡与传输介质连接的接口的不同,网卡也可分为不同的类型。

(1)按照支持的计算机种类,网卡主要分为标准以太网卡和 PCMCIA 网卡两类。标准以太网卡用于台式计算机联网,而 PCMCIA 网卡用于笔记本电脑联网。PCMCIA 是个人计算机内存卡国际协会(Personal Computer Memory Card International Association)制定的一种便携式插卡标准,符合这种标准的网卡和信用卡大小相似,它仅适用于将便携机联入互联网。

(2)按照网卡支持的传输速率分类,目前主要分为四类:10 Mb/s 网卡、100 Mb/s 网卡、10/100 Mb/s 自适应网卡、1 000 Mb/s 网卡。10 Mb/s 和 100 Mb/s 网卡仅支持10 Mb/s和100 Mb/s 的传输速率,在使用非屏蔽双绞线(UTP)作为传输介质时,通常10 Mb/s网卡与3类 UTP 配合使用,而100 Mb/s 网卡与5类 UTP 相连接。10 或 100 Mb/s 自适应网卡是由网卡自动检测网络的传输速率,保证网络中两种不同传输速率的兼容性。随着局域网传输速率的不断提高,1 000 Mb/s 网卡大多应用于高速的服务器中。

(3)按照所支持的传输介质类型,网卡主要分为双绞线网卡、粗缆网卡、细缆网卡和光纤网卡4类。根据不同的传输介质,网卡提供了相应的接口。适用于粗缆的网卡应提供 AUI 接口,适用于细缆的网卡应提供 BNC 接口,适用于非屏蔽双绞线的网卡应提供 RJ-45 接口,适用于光纤的网卡应提供光纤的 F/O 接口。

(4)按照所支持的总线类型,网卡主要可以分为 ISA、EISA 和 PCI3 类网卡。ISA 总线的网卡又可分为8位和16位的两种。8位 ISA 总线接口的网卡由于 CPU 的占用率较高,因此数据传输速率较慢;16位 ISA 总线接口的网卡数据传输速率比8位总线的快,应用较多的是 ISA 接口的10 Mb/s 网卡,价格较便宜。但随着计算机技术的飞速发展,ISA 总线接口的网卡的使用越来越少。EISA 总线接口的网卡能够并行传输32位数据,数据传输速度快,但价格较贵。PCI 总线接口网卡的 CPU 占用率较低,常用的32位 PCI 网卡的理论传输速率为133 Mb/s,因此支持的数据传输速率可达100 Mb/s。

2)网卡的物理地址

在网卡的存储器中保存了一个全球唯一的网络节点地址,这个地址称为介质访问控制地址(media access control,MAC),又称为硬件地址或网卡物理地址。MAC 地址用12个十六进制数来表示,它的地址长度是48位(比特),前6个十六进制数(也就是24比特)代表网卡生产厂商的标识符信息,后6个十六进制数代表生产厂商分配的网卡序号。MAC 地址格式如图3-10所示。

生产厂商标识符	厂商分配的序号
6个十六进制数(24位二进制数)	6个十六进制数(24位二进制数)

图3-10 MAC地址格式

一个典型的用12个十六进制数表示的MAC地址的写法为00-0B-DB-A3-D4-B6。每块网卡都有一个唯一的MAC网络地址,这个地址是网卡生产厂商在生产时写入网卡上的ROM芯片中的。MAC地址的主要作用是在以太网传输数据时,在所传输的数据包中包含源节点和目标节点的MAC地址,网络中每台节点设备的网卡会检查所传输的数据中MAC地址是否与自己的MAC地址相匹配,如果不匹配,则网卡会丢弃该数据包。

在命令提示符状态下,输入“ipconfig /all”命令并按Enter键,可以查看到当前计算机的

网卡的 MAC 地址,如图 3-11 所示。

```

Connection-specific DNS Suffix . : 
Description . . . . . : WAN (PPP/SLIP) Interface
Physical Address. . . . . : 00-53-45-00-00-00
Dhcp Enabled. . . . . : No
IP Address. . . . . : 113.206.130.107
Subnet Mask . . . . . : 255.255.255.255
Default Gateway . . . . . : 113.206.130.107
DNS Servers . . . . . : 221.5.203.98
                        221.7.92.98
NetBIOS over Tcpip. . . . . : Disabled
    
```

图 3-11 在 Windows 中查看网卡的物理地址



图片
集线器

2. 集线器

集线器(hub)是局域网中计算机和服务器的连接设备,是局域网的星型连接点,每个工作站用双绞线连接到集线器上,由集线器对工作站进行集中管理。

集线器的主要功能是对接收到的信号进行再生整形放大,以增加网络的传输距离,同时把所有节点集中在以它为中心的节点上。

集线器工作于 OSI 参考模型的第一层,即“物理层”。

集线器是一种共享设备,它本身不能识别目的地址,当同一局域网内的 A 主机给 B 主机传输数据时,数据包在以集线器为架构的网络上是以广播方式传输的,由每一台终端通过验证数据包头的地址信息来确定是否接收。也就是说,在这种工作方式下,同一时刻网络上只能传输一组数据帧的通信,如果发生碰撞还得重试。这种方式就是共享网络带宽。

集线器技术发展至今,也经历了许多不同主流应用的历史发展时期,所以集线器产品也有许多不同类型。集线器的分类方法有按端口数量划分、按带宽划分、按配置的形式划分是否可进行网络管理来划分等。



图 3-12 集线器

按端口数量来分是最基本的分类标准之一。主流集线器主要有 8 口、16 口和 24 口等大类,但也有少数品牌提供非标准端口数,如 4 口和 12 口的,还有 5 口、9 口、18 口的集线器产品,这主要是想满足部分对端口数要求过严、资金投入比较谨慎的用户需求。此类集线器一般用作家庭或小型办公室等。图 3-12 所示为一款 16 口的集线器。

3. 交换机

交换机(switch)是集线器的升级换代产品,所以交换机和集线器在功能上一样,也是一种网络集中设备,它是用来集中连接其他网络的,其“交换”概念的提出是相对于集线器的共享工作模式的改进。交换机可以把网络“分段”,通过对照 MAC 地址表,交换机只允许必要的网络流量通过交换机。

通过交换机的过滤和转发,可以有效地隔离广播风暴,减少误包和错包的出现,避免共享冲突。

交换机在同一时刻可进行多个端口对之间的数据传输。每一端口都可视为独立的网段,连接在其上的网络设备独自享有全部的带宽,无须同其他设备竞争使用。当节点 A 向节点 D 发送数据时,节点 B 可同时向节点 C 发送数据,而且这两个传输都享有网络的全部带宽,都有着自己的虚拟连接。假设这里使用的是 10 Mb/s 的以太网交换机,那么该交换机这时的总流量就等于 $2 \times 10 \text{ Mb/s} = 20 \text{ Mb/s}$,而使用 10 Mb/s 的共享式集线器时,一个集线器的总流量也不会超过 10 Mb/s。

目前,在局域网中使用的交换机一般分为第二层交换机和第三层交换机,这两类交换机的最大区别是第二层交换机不带路由功能,第三层交换机带路由功能。图 3-13 所示为思科交换机。



图 3-13 思科交换机

第二层交换机工作在 OSI 参考模型的第二层,即数据链路层。第二层交换机依赖于链路层中的信息(如 MAC 地址)完成不同端口数据间的线速交换,主要功能包括物理编址、错误校验、帧序列及数据流控制。这是最原始的交换技术产品,目前桌面型交换机一般是属于这类型,因为桌面型的交换机一般来说所承担的工作复杂性不是很强,又处于网络的基层,所以也就只需要提供最基本的数据链接功能即可。目前第二层交换机应用最为普遍(主要是价格便宜,功能符合中、小企业实际应用需求),一般应用于小型企业或中型以上企业网络的桌面层次。

第三层交换机工作于 OSI 参考模型的网络层,具有路由功能,它是将 IP 地址信息提供给网络路径选择,并实现不同网段间数据的线速交换。当网络规模较大时,可以根据特殊应用需求划分为小面独立的 VLAN 网段,以减小广播所造成的影响。

要说明的是,所有的交换机从协议层次上来说都是向下兼容的,也就是说所有的交换机都能够工作在第二层。

4. 路由器

路由器(router)是将一个网络(一般为局域网)接入另一个网络,或实现网络之间互联的必选设备。其主要功能是实现不同类型网络之间的数据“翻译”,使一种类型的网络(如以太网)能读懂另一种类型的网络(如令牌环网)发送过来的数据。

路由器用于连接网络层、数据链路层、物理层执行不同协议的网络,协议的转换由路由器完成,从而消除了网络层协议之间的差别。

路由器在网络层实现网络互联,它主要完成网络层的功能。路由器负责将数据分组从源端主机经最佳路径送到目的端主机。路由器必须具备路由选择和数据转发两个基本功能。

1) 路由选择

所谓路由选择就是通过路由选择算法确定到达目的地址(目的端的网络地址)的最佳路径。路由选择实现的方法是:路由器通过路由选择算法,建立并维护一个路由表。在路由表中包含着目的地址和下一跳路由器地址等多种路由信息。路由表中的路由信息告诉每一台路由器应该把数据包转发给谁,它的下一跳路由器地址是什么。



图片
交换机和路由器



动画
路由与路由选择

路由器根据路由表提供的下一跳路由器地址,将数据包转发给下一跳路由器。通过一级一级地把包转发到下一跳路由器的方式,最终把数据包传送到目的地。

2) 数据转发

数据转发通常也称数据交换。路由器接收到来自源端主机发送的、带有目的主机网络地址的分组后,检查数据包的目的地址,再根据路由表来确定它是否知道怎样转发这个数据包,若它不知道下一跳路由器的地址,则将包丢弃。如果它知道怎么转发这个包,路由器将改变目的地址为下一跳路由器的地址,并且将包传给下一跳路由器。

为了简单地说明路由器的工作原理,假设有这样一个简单的网络。如图 3-14 所示,A、B、C、D 四个网络通过路由器连接在一起。

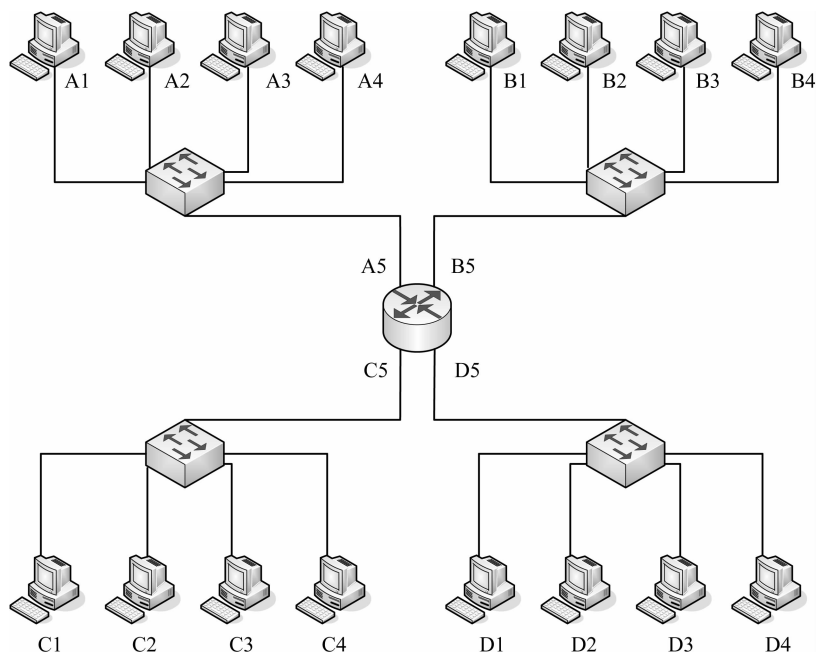


图 3-14 路由器的工作原理

路由器是如何发挥其路由、数据转发作用的? 现假设网络 A 中一个用户 A1 要向 C 网络中的 C3 用户发送一个请求信号,信号传递的步骤如下。

(1) 用户 A1 将目的用户 C3 的地址 C3,连同数据信息以数据帧的形式通过集线器或交换机以广播的形式发送给同一网络中的所有节点,当路由器 A5 端口侦听到这个地址后,分析得知所发目的节点不是本网段的,需要路由转发,就把数据帧接收下来。

(2) 路由器 A5 端口接收到用户 A1 的数据帧后,先从报头中取出目的用户 C3 的 IP 地址,并根据路由表计算出发往用户 C3 的最佳路径。因为从分析得知到 C3 的网络 ID 号与路由器的 C5 网络 ID 号相同,所以由路由器的 A5 端口直接发向路由器的 C5 端口应是信号传递的最佳路径。

(3) 路由器的 C5 端口再次取出目的用户 C3 的 IP 地址,找出 C3 的 IP 地址中的主机 ID 号,如果在网络中有交换机则可先发给交换机,由交换机根据 MAC 地址表找出具体的网络节点位置;如果没有交换机设备则根据其 IP 地址中的主机 ID 直接把数据帧发送给用户 C3,这样一个完整的数据通信转发过程也完成了。

从上面可以看出,不管网络有多复杂,路由器所做的工作就是这么几步,所以整个路由器的工作原理基本都差不多。当然在实际的网络中远比图 3-14 所示的要复杂,实际的步骤也不像上述那么简单,但总的过程是一样的。

目前,生产路由器的厂商,国外主要有 CISCO(思科)公司、北电网络等,国内厂商包括华为等。图 3-15 所示为 H3C 的路由器。



图 3-15 H3C 的路由器

3.2 扩展知识

3.2.1 局域网体系结构

为了使不同厂商生产的网络设备之间具有兼容性、互换性和互操作性,以便让用户更灵活地进行设备选型,国际标准化组织开展了局域网的标准化工作。美国电气与电子工程师协会(Institute of Electrical and Electronic Engineers, IEEE)于 1980 年 2 月成立了局域网络标准化委员会(简称 IEEE 802 委员会),专门进行局域网标准的制定。经过多年的努力,IEEE 802 委员会公布了一系列标准,称为 IEEE 802 标准。

1. 局域网体系结构

IEEE 802 参考模型与 OSI 参考模型的对应关系如图 3-16 所示。局域网参考模型只对应于 OSI 参考模型的数据链路层与物理层,它将数据链路层划分为两个子层:逻辑链路控制(logical link control, LLC)子层与介质访问控制(media access control, MAC)子层。

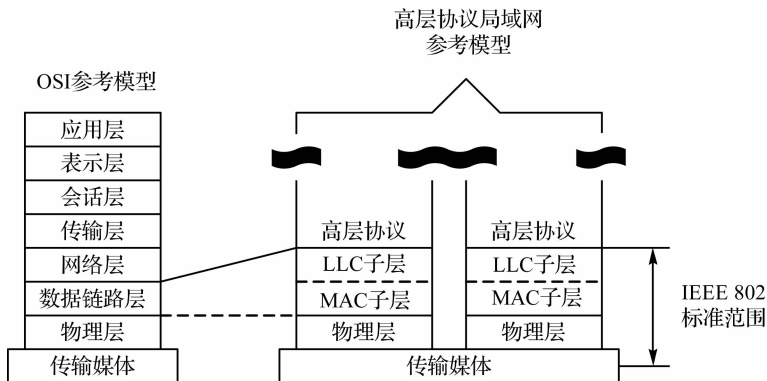


图 3-16 IEEE 802 参考模型与 OSI 参考模型的对应关系

1) 物理层

物理层涉及通信在信道上传输的原始比特流,它的主要作用是确保二进制位信号的正确传输,包括位流的正确传送与正确接收。也就是说物理层必须保证在双方通信时,一方发送二进制 1,另一方接收的也是 1,而不是 0。

2) MAC 子层

介质访问控制是数据链路层的一个功能子层。MAC 构成了数据链路层的下半部,它直接与物理层相邻。MAC 子层主要制定管理和分配信道的协议规范,换句话说,就是用来决定广播信道中信道分配的协议属于 MAC 子层。MAC 子层是与传输介质有关的一个数据链路层的功能子层,它的主要功能是进行合理的信道分配,解决信道的竞争问题。它支持在 LLC 子层中完成介质访问控制功能,为竞争的用户分配信道使用权,并具有管理多链路的功能。MAC 子层为不同的物理介质定义了介质访问控制标准。目前,IEEE 802 已制定的介质访问控制标准有著名的带冲突检测的 CSMA/CD、Token-Ring 和 Token-Bus 等。介质访问控制方法决定了局域网的主要性能,它对局域网的响应时间、吞吐量和网络利用率等都有十分重要的影响。

3) LLC 子层

逻辑链路控制也是数据链路层的一个功能子层。它构成了数据链路层的上半部,与网络层和 MAC 子层相邻。LLC 在 MAC 子层的支持下向网络层提供服务。可运行于所有 802 局域网和城域网协议之上的数据链路协议被称为逻辑链路控制 LLC。LLC 子层与传输介质无关,它独立于介质访问控制方法,隐藏了各种 802 网络之间的差别,向网络层提供一个统一的格式和接口。LLC 子层的作用是在 MAC 子层提供的介质访问控制和物理层提供的比特服务的基础上,将不可靠的信道处理为可靠地信道,确保数据帧的正确传输。LLC 子层的具体功能包括数据帧的组装与拆卸、帧的收发、差错控制、数据流控制和发送顺序控制等功能,并为网络层提供两种类型的服务,即面向连接服务和无连接服务。



随堂测试

2. IEEE 802 局域网标准

1980 年 2 月,IEEE 成立了专门负责制定局域网标准的 IEEE 802 委员会。该委员会开发了一系列局域网(LAN)和城域网(MAN)标准,最广泛使用的标准是以太网家族、令牌环、无线局域网和虚拟网等。IEEE 802 委员会于 1984 年公布了 5 项标准,即 IEEE 802.1~IEEE 802.5,随着局域网技术的迅速发展,新的局域网标准不断被推出,最新的吉位以太网技术目前也已标准化。下面是 IEEE 802 委员会颁布的标准。

- (1) IEEE 802.1: 局域网概述、体系结构、网络管理和网络互联。
- (2) IEEE 802.2: 逻辑链路控制。
- (3) IEEE 802.3: CSMA/CD 访问方法和物理层规范。
- (4) IEEE 802.4: 令牌总线。
- (5) IEEE 802.5: 令牌环访问方法和物理层规范。
- (6) IEEE 802.6: 城域网访问方法和物理层规范。
- (7) IEEE 802.7: 宽带技术咨询和物理层课题与建议实施。
- (8) IEEE 802.8: 光纤技术咨询和物理层课题。
- (9) IEEE 802.9: 综合语音/数据服务的访问方法和物理层规范。
- (10) IEEE 802.10: 安全与加密访问方法和物理层规范。
- (11) IEEE 802.11: 无线局域网访问方法和物理层规范,包括 IEEE 802.11a、IEEE 802.11b、

IEEE 802.11c 和 IEEE 802.11q 标准。

(12) IEEE 802.12:100VG-AnyLAN 快速局域网访问方法和物理层规范。

IEEE 802 标准仅包含 OSI 参考模型物理层和数据链路层协议,其他较高层次的协议目前还没有制定,一般会参考使用 OSI 和其他相应标准(如 TCP/IP)。IEEE 802 已增加到十几个分委员会,其各分委员的结构关系及其制定的局域网标准如图 3-17 所示。

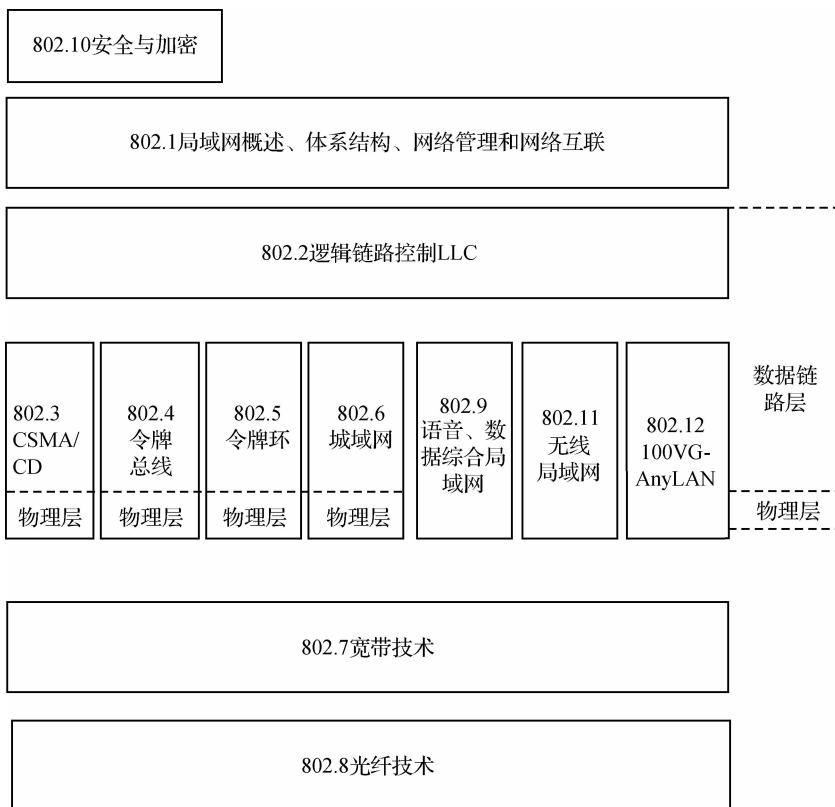


图 3-17 IEEE 802 各分委员会的结构关系及其制定的局域网标准

目前局域网协议的大部分功能由硬件实现,即通过网络适配器(网卡)来实现,再加上网卡驱动程序就可以实现 LAN 协议的整个功能。在 IEEE 802 标准中,目前应用最为广泛的是 IEEE 802.3,因为该标准是在 Ethernet 标准上制定的,所以现在人们也将 IEEE 802.3 局域网称为 Ethernet。

Token-Ring 是由美国 IBM 公司率先推出的环形基带网络,采用 4 Mb/s 和 16 Mb/s 两种传输速率,传输电缆为 IBM 专用电缆系统。IEEE 802.5 标准就是在 IBM Token-Ring 的基础上制定的,因此两者无太大的差别。Token-Ring 比较适合在传输距离远、负载重及实时性要求高的环境下应用,其网络性能优于以太网,而在轻负载时,性能反而不如以太网。此外,Token-Ring 网络的价格相对要贵一些。

3.2.2 局域网的发展趋势

计算机网络于 20 世纪 60 年代起源于美国,原本用于解决军事通信问题,随技术的不断发展成熟后转入民用,经过几十年不断的发展和完善,现已广泛应用于各个领域,并以高速

不断向前迈进。信息时代的到来不断为计算机网络开创着新时代的格局。作为计算机网络中一个不可或缺的部分,局域网的发展也是非常值得关注的一个问题。

1. 网络的协议发展

IPv4 是目前因特网所使用的网络层协议。自 20 世纪 80 年代初以来,IPv4 一直在因特网上良好稳定地运行着。但是,IPv4 协议设计之初没有考虑到因特网及其所提供的服务突飞猛进的发展,现在已经暴露出一些不足之处,未来必将被 IPv6 所取代。

网络协议从 IPv4 到 IPv6 的转变,不仅意味着增加 IP 地址数,还将给整个互联网产业带来重大影响。在万物互联的背景下,未来联网终端会越来越多,普及 IPv6 将为物联网、5G 等新技术发展铺平道路。

2. 局域网速度的发展及传输介质的变化

局域网传输的介质由最初的双绞线、同轴电缆,发展到现在越来越普及的光纤应用;相应地,局域网的传输速度由原来的 10 M 发展到现在的百兆、千兆、万兆,甚至某些厂家已经推出十万兆级核心设备,而且这个速度还会增长。无论是因为铜资源的稀缺还是为了承载更多的业务数据,未来的局域网中必然大量用到光纤作为传输介质,逐步取代双绞线等传统介质。

3. 有线与无线共建的混合网络持续

随着便携式联网终端设备的广泛应用,人们对移动无线网络的通信需求越来越高。传统有线局域网络在复杂环境和高灵活性要求的环境下已经越来越不能满足人们的需求,于是无线局域网络(wireless local area network, WLAN)应运而生,且发展迅速。

与有线局域网速度的不断提升相对应,无线局域网经历了 802.11b、g、n 等一系列进化,速度也是越来越快。虽然在稳定性及可靠性等方面落后于有线网络,但是已经能够满足大多数的接入需求。它还具备有线网络无法比拟的灵活性易扩展性等,被广泛运用于广场、校园等人员流动较大或布线不便的场合。

另外,无线局域网的弱势也很明显。容易受到天气、障碍物等客观环境影响,故障点查找不便,接入控制管理不便等注定它不能作为稳定的局域网骨干存在。所以,混合局域网作为有线网络的延伸和补充,能最大限度地发挥出 WLAN 的优势,当前混合型网络无论在中小型企业还是家庭局域网中应用都极为广泛。未来,有线和无线的充分混合网络在局域网的组建中仍将处于主导地位。

4. 更加安全稳定智能化网络

一直以来,针对网络层和应用层的攻击一直是网络安全的重大威胁。局域网由于网速快、设备相对简单、专业技术人员不足等因素成为网络安全问题的重灾区。传统的局域网防护多依赖于普通防火墙,对数据包的 IP 层和网络层进行检查,无法对网络和服务器区域形成有效的,特别是技术类攻击的防护。随着局域网应用越来越丰富,网络的攻击危害越发凸显,单纯依靠服务器技术和软件技术的防护更加捉襟见肘。对由很多中小型应用组成的大中型局域网服务来说,根本不可能通过应用层防护堵住所有漏洞,保障局域网服务的安全。

由网络互联带动的信息化,智能化的应用环境下,网络本身将承担起更多的安全责任。凭借承载数据的特性和强大的数据处理能力,新一代网络安全设备需要实现对数据包应用层的动作进行分析,并根据网络管理员的思想形成基于应用层的 ACL 规则,有效避免针对