第 3 章　计算机网络体系结构与 IP 协议

本章学习目标

- 了解计算机网络协议与网络体系结构。
- 了解 OSI/RM 模型。
- 了解 TCP/IP 模型。
- 了解 IP 协议。
- 掌握子网的划分。
- 了解 IPv6 的特点。

本章首先介绍计算机网络的协议及特点，然后介绍计算机网络参考模型，包括 OSI/RM 模型和 TCP/IP 模型，接着介绍 IPv4 协议和子网的划分，最后介绍 IPv6 协议的特点。

3.1　计算机网络协议与体系结构

扫一扫

3.1.1　计算机网络协议与体系结构概述

1. 计算机网络协议

网络上的计算机之间要进行通信，就必须遵循一定的规则或约定，这些为网络数据交换而制定的规则或约定称为协议。就像人们说话用某种语言一样，在网络上的各台计算机之间也有一种相互交流信息的规则，这就是网络协议，不同的计算机之间必须使用相同的网络协议才能进行通信。网络协议是网络上所有设备（网络服务器、计算机及交换机、路由器、防火墙等）之间的通信规则的集合，它规定了通信时信息必须采用的格式和这些格式的意义。

网络协议是由以下 3 个要素组成的。

（1）语义。解释控制信息每个部分的意义。它规定了通信双方“讲什么”，即需要发出何种控制信息以及完成的动作与做出的响应。

（2）语法。用户数据与控制信息的结构与格式以及数据出现的顺序。它规定了通信双方“如何讲”，即协议元素的格式、数据及控制信息的格式、编码和信号电平等。

（3）时序。时序也称为同步，是对事件发生顺序的详细说明。它规定了通信双方“做的顺序”，主要涉及传输速度匹配和顺序问题。

2. 计算机网络体系结构

不同系统之间的相互通信建立在各层实体之间能够相互通信的基础上，因此一个系统的通信协议是各层通信协议的集合。计算机网络由若干层来实现，每一层都有自己的协议。计算机网络的层次模型及其协议的集合称为网络的体系结构。

在网络体系结构中，每一层协议都实现了与另一层中对等实体之间的通信，所以称之为对等层协议。另外，每一层协议还要向相邻上层协议提供服务接口。网络体系结构的描述必须包含足够的信息，使实现者可以为每一层编写程序和设计硬件，并使之符合有关协议。

网络体系结构具有以下特点。

(1) 以功能作为划分层次的基础。

(2) 第 N 层的实体在实现自身定义的功能时,只能使用第 $N-1$ 层的功能。

(3) 第 N 层向第 $N+1$ 层提供服务时,此服务不仅包含第 N 层本身的功能,还包含由其下层提供的功能。

(4) 仅在相邻层之间有接口,且下层所提供的服务的具体细节对上一层完全屏蔽。

图 3-1 显示了计算机网络体系结构的层次模型。其中,实体是指在每一层中实现该层功能的活动元素,如终端、应用程序以及进程等。

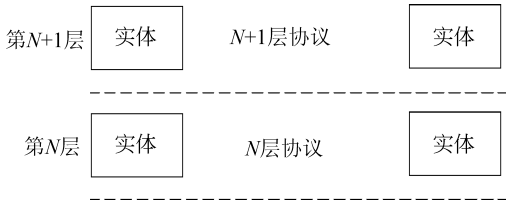


图 3-1 计算机网络体系结构的层次模型

3.1.2 计算机网络体系的产生与发展

在计算机网络的发展过程中,存在专用网络体系结构,如 IBM 公司的系统网络体系结构(systems network architecture,SNA);也存在开放网络体系结构,如 OSI/RM。

1. SNA

SNA 是 IBM 公司开发的网络体系结构,在 IBM 公司的主机环境中得到广泛的应用。SNA 主要是 IBM 公司的大型机(ES/9000、S/390 等)和中型机(AS/400)的主要联网协议,SNA 于 1974 年首次公布,是 IBM 公司为了连接 3270 系列产品而推出的方案。

SNA 设计在与 IBM 主机系统相连的大多数终端是不可编程终端的年代。SNA 在互连的主机之间提供了静态路由选择,所以用户在一个终端上可以访问其他任何互连的主机。在 SNA 出台之前,用户要访问主机,必须登录到一个单独的终端上。SNA 只是针对集中化的 IBM 主机计算环境设计的,所以它不适用于现在对等、客户/服务器、多供应商产品和多协议的环境。一般这些环境建立在部门级,每个管理者设计和建立自己的局域网。

为了提供程序间通信功能,IBM 公司引进了高级程序对程序通信(advanced program-to-program communication,APPC),并且为了对抗 TCP/IP 的威胁,还推出了高级对等自治网(advanced peer-to-peer networking,APPN)。APPN 在保持主机系统多样性的同时提供了一个企业范围内的非集中网络计算环境。在 APPN 上,大型和小型系统相互对等操作。IBM 公司的最新策略是,在包容工业标准协议(如 TCP/IP 和 OSI 协议)的同时继续支持 APPN。这个思想在联网方案中体现出来了,并且已有遵守该标准的产品。MPTN(多协议传输网)就是一个例子,它使应用程序从基层网络协议解脱开,允许应用程序使用其他协议。

SNA 最大的特性就是封闭性,它是 IBM 公司开发的专有协议。如果 SNA 要在其他主机系统中应用,需在网络的每一个节点增加支持 SNA 的软件和硬件。SNA 环境具有 COS(服务分类)与安全的能力,但 SNA 的安全仅仅是基于主机的 XID(交换 ID)。XID 类似于

用户的口令，在 SNA 会话建立过程中要交换 XID 来确认用户的合法性。XID 需要在主机中静态配置，网络逻辑单元需要与主机上的逻辑单元在建立会话时进行 XID 交换；在 SNA 环境中，除 XID 交换外，没有其他安全方面的考虑。由于 SNA 协议的封闭性，大多数人不熟悉它，使用 SNA 协议的网络也因此很少受到攻击；但对于熟悉它的人来说，SNA 的安全机制并不严密。基于 SNA 设计和开发的安全系统也不常见。单单依靠封闭性实际上是不可靠的。由于 SNA 协议的非开放性及开发的复杂性，在 SNA 环境下开发应用系统比较复杂，而且系统的迁移性比较弱，这不符合业界开放的潮流。维持一大批专家来支持 SNA 网络也日趋困难。随着因特网的发展和普及，越来越多的用户采用开放的 TCP/IP，IBM 公司也不例外。现在，在 IBM 公司传统的只支持 SNA 环境的计算机上也开始支持 TCP/IP，包括 ES/9000 和 AS/400。越来越多的 SNA 用户向 TCP/IP 环境迁移。

2. 开放系统互连参考模型

OSI 是 ISO 制定的网络体系结构模型，为开放式互联信息系统提供了一种功能结构的框架。它包括 7 层，从低到高分别是物理层、数据链路层、网络层、传输层、会话层、表示层和应用层。

OSI 参考模型还保持在模型阶段，它并不是一个已经被完全接受的国际标准。考虑到大量现存的事实上的标准，许多厂商只能支持许多在工业界使用的不同协议，而不是仅仅接受一个国际标准。

3.1.3　计算机网络体系的分层结构

扫一扫

由于计算机网络的复杂性，很难用一个单一的协议来为网络中的所有通信规定一套完整的规则。因此，把通信问题划分为许多小问题，然后给每个小问题设计一个单独的协议，从而使得每个协议的设计、分析、编码和测试都变得容易，这就是网络体系结构中使用的分层概念。

计算机网络体系采用分层结构主要是基于以下几点考虑。

(1) 各层之间相互独立。高层不需要知道低层的功能是采取何种硬件技术来实现的，它只需要知道通过与低层的接口就可以获得需要的服务。

(2) 灵活性好。各层都可以采用最适当的技术来实现。例如，某一层的实现技术发生了变化，用硬件代替了软件，只要这一层的功能与接口保持不变，实现技术的变化并不会对其他各层以及整个系统的工作产生影响。

(3) 易于实现和标准化。由于采用规范的层次结构去组织网络功能与协议，因此可以将计算机网络复杂的通信过程划分为有序的连续动作与有序的交互过程，有利于将网络复杂的通信工作过程分解为一系列可以控制和实现的功能模块，使得复杂的计算机网络系统变得易于设计、实现和标准化。如果想加上一些新的服务，只需要修改一层的功能即可。中间层可以使用低层提供的服务以及为高层提供服务。

下面以邮政通信系统为例说明分层的概念。一个邮政通信系统由用户(写信人和收信人)、邮局、运输部门和运输工具组成，可以将邮政通信系统按照功能分为 4 层，各层分工明确，功能独立，如图 3-2 所示。

分层之后，还需要在对等层之间约定一些通信规则，即对等协议。例如，通信双方写信时，必须约定用同样的语言，这样对方收到信后才能看懂信的内容。另外，一个邮局将用户的信件收集起来后，要进行分拣、打包等操作，而分拣和打包等规则必须在邮局之间事先协

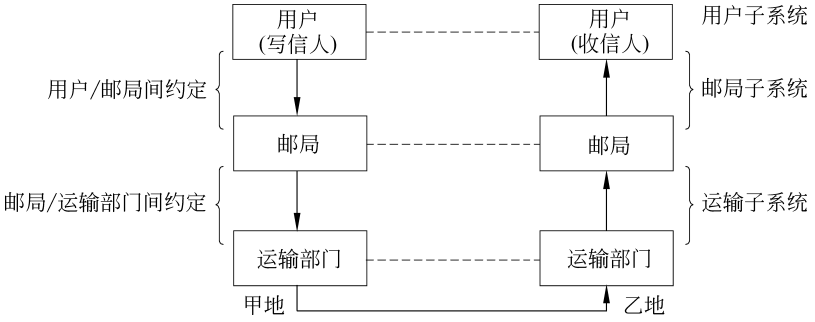


图 3-2　邮政通信系统的分层模型

商好,这就是邮局层次的协议。同样,在运输部门之间也应该有一致的协议。

当信写好之后,写信人把信装入信封中,信封上按照邮局规定顺序写上收信人的邮政编码、地址、姓名和发信人的地址、姓名和邮编,贴好邮票后把信封投入邮筒。这些就是写信人和邮局之间的约定,这些约定就是相邻层之间的接口。这封信如何传递到收信人的手中呢?写信人不需要去考虑,而是交给邮局去处理。

邮局将信件打包后交付运输部门进行运输,如民航、铁路或者公路运输部门等。这时,邮局和运输部门也存在着接口的问题,如到站地点、时间、包装形式等。信件运送到目的地后经历相反的过程,最终将信件送到收信人手中,收信人依照约定的格式才能读懂信件。

从一封信的运输过程可以看出,虽然两个用户、两个邮局、两个运输部门分处甲乙两地,但它们是两两对等的实体,即对等层实体。而处于同一地的不同实体则是上下层关系,存在着服务与被服务的关系。对等层实体间的约定称为协议,而上下层实体间的约定称为接口。

采用分层模型的网络系统结构主要有两个优点。

(1) 它将建造一个网络的问题分解为几层,每一层解决一部分问题。当所有层的问题都解决了,整个网络的问题也就解决了。

(2) 它提供了一种模块化的设计。如果想加上一些服务,只需要修改某一层的功能即可,其他层不需要修改。

3.2　OSI 参考模型

扫一扫

3.2.1　OSI 参考模型概述

OSI 一般都称 OSI 参考模型,是国际标准化组织在 1985 年研究的网络互联模型。OSI 参考模型定义了开放系统的层次结构、层次之间的相互关系及各层所包含的可能的服务。OSI 参考模型并不是一个标准,而是一个在制定标准时使用的概念性框架,其作为一个框架来协调和组织各层协议的制定。

OSI 参考模型制定过程中所采用的方法是将整个庞大而复杂的问题划分为若干个容易处理的小问题,这就是分层的体系结构方法。OSI 中采用了三级抽象,即体系结构、服务定义和协议规定说明。

OSI 参考模型的服务定义详细说明了各层所提供的服务。某一层的服务就是该层及其下各层的一种能力,它通过接口提供给更高一层。各层所提供的服务与这些服务是怎么实

现的无关。同时，各种服务定义还定义了层与层之间的接口和各层所使用的原语，但是不涉及接口是怎么实现的。

OSI 参考模型中的各种协议精确定义了应当发送什么样的控制信息，以及应当用什么样的过程来解释这个控制信息。协议的规程说明具有最严格的约束力。

OSI 参考模型并没有提供一个可以实现的方法，而只是描述了一些概念，用来协调进程间通信标准的制定。在 OSI 参考模型范围内，只有在各种协议是可以被实现的，而各种产品只有和 OSI 参考模型的协议相一致时才能互联。也就是说，OSI 参考模型并不是一个标准，只是一个在制定标准时所使用的概念性的框架。

3.2.2　OSI 参考模型各层功能

OSI 参考模型定义了网络互联的 7 层框架，包括物理层、数据链路层、网络层、传输层、会话层、表示层和应用层，如图 3-3 所示。

图 3-3　OSI 参考模型的 7 层框架

OSI 参考模型各层的功能如下。

(1) 物理层(physical layer)是 OSI 参考模型的最底层，它利用传输介质为数据链路层提供物理连接。为此，该层定义了与物理链路的建立、维护和拆除有关的机械、电气、功能和规程特性。

物理层包括信号线的功能、0 和 1 信号的电平表示、数据传输速率、物理连接器规格及其相关的属性等。物理层的作用是通过传输介质发送和接收二进制比特流。

(2) 数据链路层(data link layer)是为网络层提供服务的，解决两个相邻节点之间的通信问题，传送的协议数据单元称为数据帧。数据帧中包含物理地址(又称 MAC 地址)、控制码、数据及校验码等信息。该层的主要作用是通过校验、确认和反馈重发等手段，将不可靠的物理链路转换成对网络层来说无差错的数据链路。此外，数据链路层还要协调收发双方的数据传输速率，即进行流量控制，以防止接收方因来不及处理发送方发来的高速数据而导致缓冲器溢出及线路阻塞。

(3) 网络层(network layer)是为传输层提供服务的，传送的协议数据单元称为数据包或分组。该层的主要作用是解决如何使数据包通过各节点传送的问题，即通过路径选择算法(路由)将数据包送到目的地。另外，为避免通信子网中出现过多的数据包而造成网络阻塞，需要对流入的数据包数量进行控制(拥塞控制)。当数据包要跨越多个通信子网才能到达目的地时，还要解决网际互联的问题。

(4) 传输层(transport layer)的作用是为上层协议提供端到端的可靠、透明的数据传输服务，包括处理差错控制和流量控制等问题。该层向高层屏蔽了下层数据通信的细节，使高层用户看到的只是在两个传输实体间的一条主机到主机的、可由用户控制和设定的、可靠的数据通路。

(5) 会话层(session layer)的主要功能是管理和协调不同主机上各种进程之间的通信(会话)，即负责建立、管理和终止应用程序之间的会话。会话层得名的原因是它很类似于两个实体间的会话概念。例如，一个交互的用户会话以登录到计算机开始，以注销结束。

(6) 表示层(presentation layer)。处理流经节点的数据编码的表示形式问题，以保证一

个系统应用层发出的信息可被另一系统的应用层读出。如果必要,该层可提供一种标准表示形式,用于将计算机内部的多种数据表示形式转换成网络通信中采用的标准表示形式。数据压缩和加密也是表示层可提供的转换功能之一。

(7) 应用层(application layer)。是 OSI 参考模型的最高层,是用户与网络的接口。该层通过应用程序来完成网络用户的应用需求,如文件传输、收发电子邮件等。

因此,OSI 参考模型是一个定义良好的协议规范集,并有许多可选部分完成类似的任务。它定义了开放系统的分层结构、层之间的相互关系以及各层所包括的可能的任务,是作为一个框架来协调和组织各层所提供的服务的。OSI 参考模型并没有提供一个可以实现的方法,而是描述了一些概念,用来协调进程间通信标准的制定。

图 3-4 显示了 OSI 各层的主要功能,图 3-5 显示了 OSI 各层传送的协议数据单元。OSI 分层模型自顶向下分别是应用层、表示层、会话层、传输层、网络层、数据链路层、物理层。其中低三层称为通信子网,提供路由完成和数据传输、转发的功能。高三层为资源子网,负责全网的数据处理业务,并向网络用户提供网络资源共享服务,并且低层通过接口向高层提供服务。

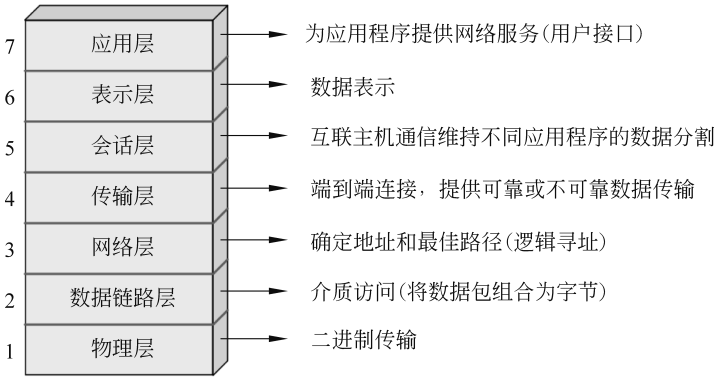


图 3-4 OSI 各层的主要功能

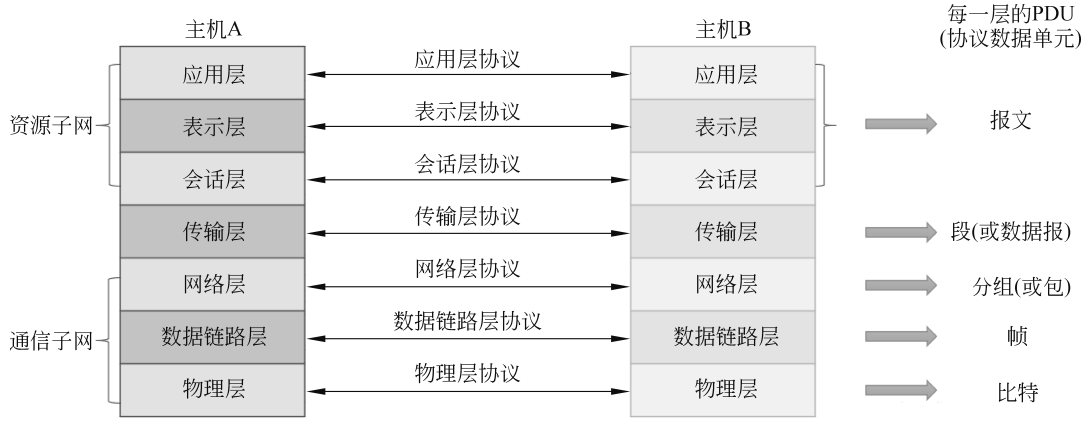


图 3-5 OSI 各层传送的协议数据单元

在 OSI 参考模型中,物理层的协议数据单元是比特(b),数据链路层的协议数据单元是

帧，网络层的协议数据单元是分组，传输层的协议数据单元是段或数据报，会话层、表示层和应用层的协议数据单元是报文。

3.2.3　OSI 参考模型通信过程

通过 OSI 参考模型，信息可以从一台计算机的应用程序传输到另一台计算机的应用程序上。例如，系统 A 上的应用程序要将信息发送到系统 B 的应用程序，则系统 A 中的应用程序需要将信息先发送到其应用层（第 7 层），然后此层将信息发送到表示层（第 6 层），表示层将数据转送到会话层（第 5 层），如此继续，直至物理层（第 1 层）。在物理层，数据被放置在物理传输信道中，并被发送至系统 B。系统 B 的物理层接收来自物理媒介的数据，然后将信息向上发送至数据链路层（第 2 层），数据链路层再转送给网络层，依次继续，直到信息到达系统 B 的应用层。最后，计算机 B 的应用层再将信息传送给应用程序接收端，从而完成通信过程。

此外，OSI 参考模型的 7 层模型还可以运用各种各样的控制信息，来和其他计算机系统的对应层通信。这些控制信息包含特殊的请求和说明，它们在对应的 OSI 层间进行交换。对于从上一层传送下来的数据，附加在前面的控制信息称为头，附加在后面的控制信息称为尾。

值得注意的是，OSI 参考模型的 1 个特定层通常是与另外 3 个 OSI 层联系：与之直接相邻的上一层和下一层还有目标联网计算机系统的对应层。例如，系统 A 的数据链路层应与其网络层、物理层以及对应系统 B 的数据链路层进行通信。

图 3-6 显示了 OSI 参考模型的通信过程。

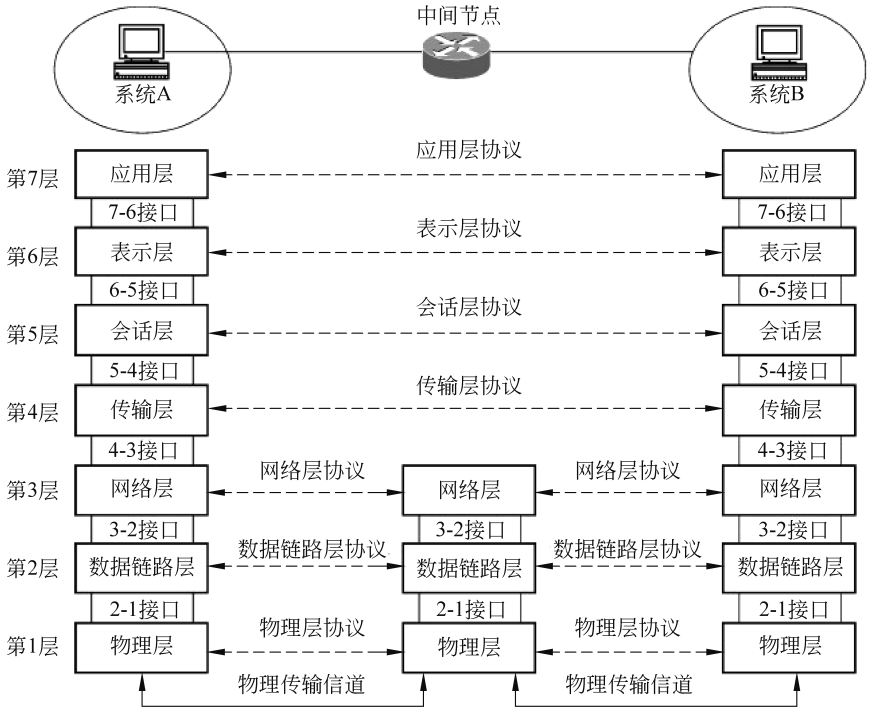


图 3-6　OSI 参考模型的通信过程

扫一扫

3.3 TCP/IP 模型

计算机网络体系结构采用分层结构。OSI 参考模型是严格遵循分层模型的典范,自推出之日起就成为网络体系结构的蓝本。但是在 OSI 参考模型推出之前,便捷、高效的 TCP/IP 体系结构就已经随着因特网的流行而成为事实上的国际标准。

3.3.1 TCP/IP 概述

传输控制协议/互联网协议(transmission control protocol/internet protocol,TCP/IP),又名网络通信协议,是因特网最基本的协议,也是 Internet 国际互联网络的基础。

1. TCP/IP 简介

TCP/IP 是用于计算机通信的一组协议,通常称为 TCP/IP 协议簇。它是 70 年代中期美国国防部为其 ARPANET 广域网开发的网络体系结构和协议标准,以它为基础组建的 Internet 是目前国际上规模最大的计算机网络,正因为 Internet 广泛使用,使得 TCP/IP 成了事实上的标准。之所以说 TCP/IP 是一个协议簇,是因为 TCP/IP 包括 TCP、IP、UDP、ICMP、RIP、TELNETFTP、SMTP、ARP、TFTP 等许多协议,这些协议一起称为 TCP/IP 协议簇,而 TCP/IP 主要由网络层的 IP 和传输层的 TCP 而得名。

TCP/IP 的目的不是推行一种标准,而是在承认已有不同标准的基础上解决这些不同导致的问题。因此,网络互联是 TCP/IP 的核心。TCP/IP 在设计时的侧重点不是具体的通信实现,也没有定义具体的网络接口协议,因此,TCP/IP 允许任何类型的通信子网参与通信。总体来看,TCP/IP 定义了电子设备如何连入因特网,以及数据在它们之间传输的标准。

2. TCP/IP 结构

TCP/IP 采用了 4 层的层级结构,分别是网络接口层、网际层、传输层和应用层,每一层都呼叫它的下一层所提供的协议来完成自己的需求。OSI/RM 参考模型与 TCP/IP 模型的对比如图 3-7 所示。

OSI/RM参考模型	TCP/IP模型
应用层	应用层
表示层	
会话层	
传输层	传输层
网络层	网际层
数据链路层	网络接口层
物理层	

图 3-7 OSI/RM 参考模型与 TCP/IP 模型的对比

3.3.2　TCP/IP 各层功能

TCP/IP 各层的主要功能如下。

1. 应用层

应用层对应 OSI 参考模型的高层，为用户提供所需要的各种服务，例如 FTP、Telnet、DNS、SMTP 等。应用层包含所有的高层协议，比如远程登录协议（telecommunications network，Telnet）；FTP；简单邮件传输协议（simple mail transfer protocol，SMTP）；域名服务（domain name service，DNS）；网络新闻传输协议（net news transfer protocol，NNTP）和 HTTP 等。Telnet 允许一台机器上的用户登录远程机器进行工作，FTP 提供将文件从一台机器上移到另一台机器上的有效方法，SMTP 用于电子邮件的收发，DNS 用于把主机名映射到网络地址，NNTP 用于新闻的发布、检索和获取，HTTP 用于在 WWW 上获取主页。

2. 传输层

传输层对应 OSI 参考模型的传输层，为应用层实体提供端到端的通信功能，保证了数据包的顺序传送及数据的完整性。该层定义了两个主要的协议：传输控制协议（TCP）和用户数据报协议（UDP）。

TCP 是面向连接的协议，它提供可靠的报文传输和对上层应用的连接服务。为此，除了基本的数据传输外，它还有可靠性保证、流量控制、多路复用、优先权和安全性控制等功能。UDP 是面向无连接的不可靠传输协议，主要用于不需要 TCP 的排序和流量控制等功能的应用程序。

TCP 和 UDP 最不同的地方是 TCP 提供了一种可靠的数据传输服务。TCP 是面向连接的，也就是说，利用 TCP 通信的两台主机首先要经历一个建立连接的过程，等到连接建立后才开始传输数据，而且传输过程中采用“带重传的肯定确认”技术来实现传输的可靠性。此外，TCP 还采用了一种称为“滑动窗口”的方式进行流量控制，发送完成后还会关闭连接。所以 TCP 要比 UDP 可靠得多。

具体实现中，TCP 采用三次握手策略，以确保数据的准确传输。发送端首先发送一个带有 SYN 标志的数据包给对方；接收端收到之后，回传一个带有 SYN/ACK 标志的数据包，以示传达确认信息；最后发送端再回传一个带 ACK 标志的数据包，代表握手结束。图 3-8 显示了 TCP 采用的三次握手策略。

表 3-1 显示了 TCP 中常见的标志位和序列号。

表 3-1　TCP 中常见的标志位和序列号

术　　语	说　　明
SYN	TCP 报文标志位，该位为 1 时，表示发起一个新连接
ACK	TCP 报文标志位，该位为 1 时，确认序号有效，确认接收到消息
seq	报文初始序列号，代表发送的第一个字节的序号
ack	报文确认序号，代表希望收到的下一个数据的第一个字节的序号

三次握手的目的是连接服务器指定端口，建立连接，并同步连接双方的序列号和确认号，交换窗口大小信息。图中 SYN＝1 表示这是一个连接请求或连接接受报文，当 SYN＝1

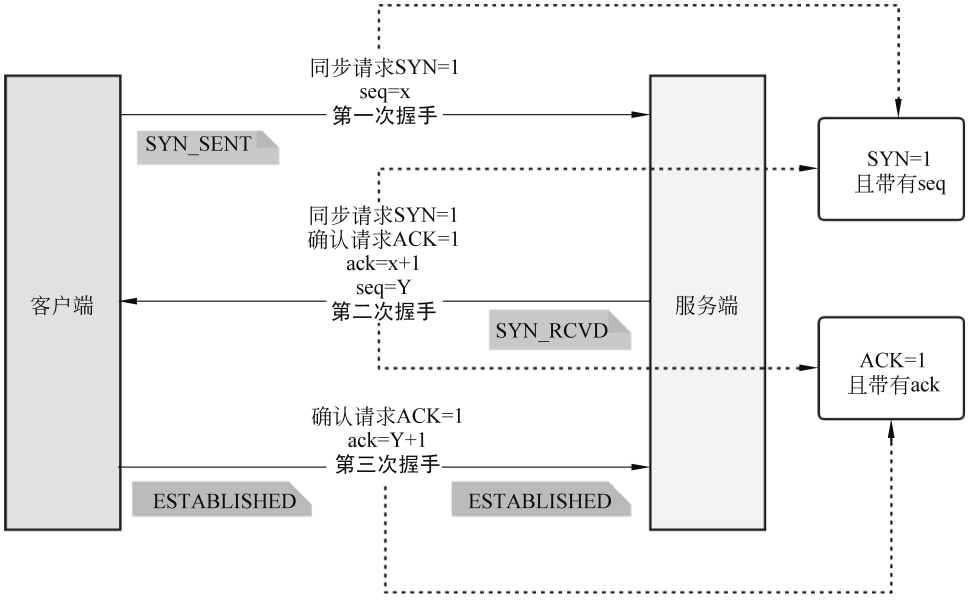


图 3-8　TCP 的三次握手

而 ACK＝0 时,表明这是一个连接请求报文段。对方若同意建立连接,则响应的报文段中使 SYN＝1、ACK＝1。

第一次握手时,客户端将标志位 SYN 置为 1,随机产生一个值 seq＝x,并将该数据包发送给服务器端,客户端进入 SYN_SENT 状态,等待服务器端确认。第二次握手时,服务器端收到数据包后,由标志位 SYN＝1 知道客户端请求建立连接,服务器端将标志位 SYN 和 ACK 都置为 1,ack＝x＋1,随机产生一个值 seq＝Y,并将该数据包发送给客户端以确认连接请求,服务器端进入 SYN_RCVD 状态。第三次握手时客户端收到确认后,检查 ack 是否为 Y＋1,ACK 是否为 1,如果正确,则将该数据包发送给服务器端,服务器端检查 ACK 是否为 K＋1,ACK 是否为 1,如果正确,则连接建立成功,客户端和服务器端进入 ESTABLISHED 状态,完成三次握手,随后客户端与服务器端之间即可以开始传输数据了。

图 3-9 显示了 TCP 的各种状态。

状态	说明
CLOSED	没有连接
LISTEN	收到了被动打开，等待SYN
SYN-SENT	已发送SYN；等待ACK
SYN-RCVD	已发送SYN+ACK；等待ACK
ESABLISHED	连接已建立；数据传送在进行
FIN-WAIT-1	第一个FIN已发送；等待ACK
FIN-WAIT-2	对第一个FIN的ACK已收到；等待第二个FIN
CLOSE-WAIT	收到第一个FIN，已发送ACK；等待应用程序关闭
TIME-WAIT	收到第二个FIN，已发送ACK；等待2MSL超时
LAST-ACK	已发送第二个FIN；等待ACK
CLOSING	双方都已决定关闭

图 3-9　TCP 的各种状态

3. 网际层

网际层也称为网际互联层，对应 OSI 参考模型的网络层，主要解决主机到主机的通信问题。它所包含的协议设计数据包在整个网络上的逻辑传输，赋予主机一个 IP 地址来完成对主机的寻址，还负责数据包在多种网络中的路由。该层有 3 个主要协议：网际协议、互联网组管理协议和互联网控制报文协议。

IP 协议是网际层最重要的协议，它提供的是一个不可靠、无连接的数据报传递服务。互联网组管理协议(internet group management protocol，IGMP)是 TCP/IP 协议簇中负责 IPv4 组播成员管理的协议。互联网控制报文协议(internet control message protocol，ICMP)是 TCP/IP 协议簇的一个子协议，用于在 IP 主机、路由器之间传递控制消息。

4. 网络接口层

网络接口层与 OSI 参考模型中的物理层和数据链路层相对应。它负责监视数据在主机和网络之间的交换。事实上，TCP/IP 本身并未定义该层的协议，而由参与互联的各网络使用自己的物理层和数据链路层协议，然后与 TCP/IP 的网络接口层进行连接。地址解析协议(ARP)工作在此层，即 OSI 参考模型的数据链路层。

3.4　IPv4

3.4.1　IPv4 简介

1. IP 的主要功能

IP 是英文 internet protocol 的缩写，意思是网络互联协议，是为计算机网络相互连接进行通信而设计的协议。在因特网中，它是能使连接到网上的所有计算机网络实现相互通信的一套规则，规定了计算机在因特网上进行通信时应当遵守的规则。任何厂家生产的计算机系统，只要遵守 IP 就可以与因特网互联互通。正是因为有了 IP，因特网才得以迅速发展，成为世界上最大的、开放的计算机通信网络。因此，IP 也可以叫作因特网协议。

IP 的基本任务是屏蔽下层各种物理网络的差异，向上层提供统一的 IP 数据报。由 IP 控制传输的协议单元称为 IP 数据报，各个 IP 数据报是相互独立的。IP 的基本功能是对数据报进行寻址和路由，并从一个网络转发到另一个网络。它屏蔽了形形色色的物理网络的差异，向上一层提供无连接的 IP 数据报服务。IP 在每个发送的数据报前加入一个控制信息，其中包含了源主机的 IP 地址和其他的一些信息。IP 的另一项工作是分割和重编在传输层被分割的数据报。由于数据报要从一个网络转发到另一个网络，当两个网络支持的数据报大小不同时，IP 就要在发送端将数据报分割，然后在分割的每一分片前再加入控制信息进行传输。当接收端接收到数据报的所有分片后，IP 将这些分片重新组合形成原始的数据报。

2. IP 的特征

IP 是一个无连接的协议。无连接是指主机之间不建立用于可靠通信的端到端的连接，源主机只是简单地将 IP 数据报发送出去，但是 IP 数据报在传输的过程中可能会丢失、重复、延迟或者失序(顺序混乱)等，因此要实现数据报的可靠传输，必须依靠高层的协议或者应用程序，如传输层的 TCP。

IP的重要特性是非连接和不可靠。非连接是指经过IP处理过的数据报是相互独立的,可以按照不同的路径传输到目的地,到达的顺序可以不一致。不可靠是指没有提供对数据流在传输时的可靠性控制,是尽力传送的数据报协议。它没有任何质量保证体系,对底层子网也没有提供任何纠错功能,用户数据报可能发生丢失、重复或者失序。IP无法保证数据报传输的结果,IP服务本身也不关心这些结果,也不将结果通知收发双方。

3. IP数据报的格式

TCP/IP定义了一个在因特网上传输的包,称为IP数据报(IP datagram)。这是一个与硬件无关的虚拟包,由首部和数据两部分组成。首部的前一部分是固定长度的20B,是所有IP数据报必须具有的。在首部的固定部分的后面是一些可选字段,其长度是可变的。首部中的源地址和目的地址都是IP协议地址。IP数据报格式如图3-10所示。

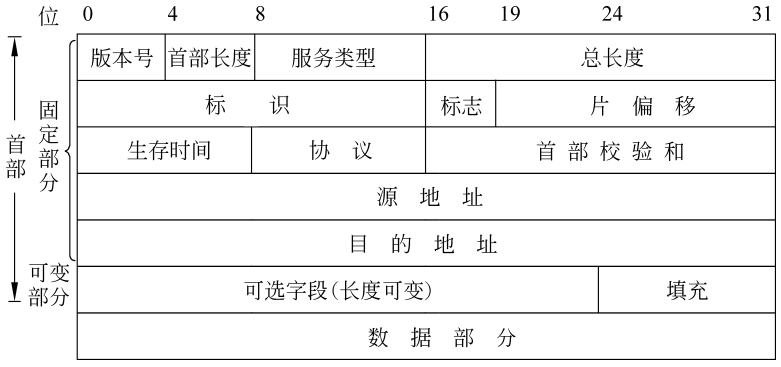


图3-10 IP数据报格式

IP数据报各字段功能如下。

(1) 版本号:占4个二进制位,表示该IP数据报使用的IP版本。目前因特网中使用的主要是TCP/IP协议族中版本号为6的IP协议。

(2) 首部长度:占4位,可表示的最大十进制数值是15。请注意,这个字段所表示数的单位是32位字(1个32位字长是4B),因此,当IP的首部长度为二进制的1111(即十进制的15)时,首部长度就达到60B。当IP数据报的首部长度不是4B的整数倍时,必须利用首部最后的填充字段加以填充。因此数据部分永远从4B的整数倍开始,这样在实现IP时较为方便。首部长度限制为60B的缺点是有时可能不够用。这样做的目的是希望用户尽量减少开销。最常用的首部长度就是20B(即首部长度为二进制的0101),这时不使用任何选项。

(3) 服务类型(type of service,TOS):占8位,用于规定本数据报的处理方式。服务类型字段的8位分成5个子域,如图3-11所示。

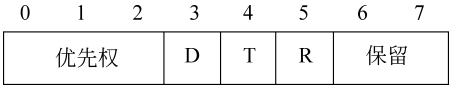


图3-11 服务类型

① 优先权的取值为0~7,数越大,表示该数据报优先权越高。在网络中,路由器可以使用优先权进行拥塞控制,例如,当网络发生拥塞时,可以根据数据报的优先权来决定数据报的取舍。

② 短延迟位D(delay):该位置1时,数据报请求以短延迟信道传输;0表示正常延迟。

③ 高吞吐量位T(throughput):该位置1时,数据报请求以高吞吐量信道传输;0表示普通吞吐量。

④ 高可靠位 R(reliability)：该位置 1 时，数据报请求以高可靠性信道传输；0 表示普通可靠性。

⑤ 保留位占两位。

(4) 总长度：占 16 位，总长度字段是指整个 IP 数据报的长度(首部＋数据区)，以字节为单位。利用首部长度字段和总长度字段就可以计算出 IP 数据报中数据部分的起始位置和长度。由于该字段占 16 位，因此理论上 IP 数据报最长可达 65 536B(事实上受物理网络的限制，要比这个数值小很多)。

(5) 标识(identification)：占 16 位。IP 在存储器中运行一个计数器，每产生一个数据报，计数器就加 1，并将此值赋给标识字段。但这个“标识”并不是序号，因为 IP 是无连接的服务，数据报不存在按序接收的问题。当数据报由于长度超过网络的 MTU(maximum transfer unit，最大传送单元)而必须分片时，标识字段的值就被复制到所有数据报的标识字段中。相同的标识字段的值使分片后的各数据报片最后能正确地重装成为原来的数据报。

(6) 标志 (flag)：占 3 位，但目前只有 2 位有意义。标志字段中的最低位记为 MF (more fragment)。MF＝1 表示后面还有分片的数据报。MF＝0 表示这已是若干数据报分片中的最后一个。标志字段中间的一位记为 DF(don't fragment)，意思是不能分片。只有当 DF＝0 时才允许分片。

(7) 片偏移：占 13 位。用于指出较长的数据报在分片后，某片在原数据报中的相对位置，也就是相对于用户数据部分的起点，该分片从何处开始。片偏移以 8B 为偏移单位。这就是说，每个分片的长度一定是 8B(64b)的整数倍。

(8) 生存时间(time to live，TTL)：占 8 位，它指定了数据报可以在网络中传输的最长时间。实际应用中把生存时间字段设置成数据报可以经过的最大路由器数。TTL 的初始值由源主机设置(通常为 32、64、128 或 256)，一旦经过一个处理它的路由器，它的值就减 1。当该字段为 0 时，数据报就丢弃，并发送 ICMP 报文通知源主机，因此可以防止数据报进入一个循环回路时无休止地传输下去。

(9) 协议：上层协议标识，占 8 位。IP 可以承载各种上层协议，目的端根据协议标识就可以把收到的 IP 数据报送到 TCP 或 UDP 等处理此报文的上层协议了。

表 3-2 显示了常用的网际协议。

表 3-2　常用的网际协议

十进制编号	协 议	说　　明	十进制编号	协 议	说　　明
0	无	保留	6	TCP	传输控制协议
1	ICMP	互联网控制报文协议	8	EGP	外部网关协议
2	IGMP	互联网组管理协议	9	IGP	内部网关协议
3	GGP	网关-网关协议	11	NVP	网络声音协议
4	无	保留	17	UDP	用户数据报协议
5	ST	流			

(10) 首部校验和：占 16 位，用于 IP 数据报首部有效性的校验，可以保证 IP 数据报首部在传输时的正确性和完整性。首部校验和字段是根据 IP 数据报首部计算出的校验和，它

不对首部后面的数据进行计算。

(11) 源地址：占32位,表示发送端IP地址。

(12) 目的地址：占32位,表示目的端IP地址。

(13) 可选字段：用来支持排错、测量以及安全等措施,内容很丰富。此字段的长度可变,为1～40B,取决于所选择的项目。某些选项只需要1B,只包括1B的选项代码,但还有些选项需要多个字节。这些选项一个个拼接起来,中间不需要分隔符,最后用全0的填充字段补齐为4B的整数倍。

(14) 填充：可选字段大小要求是4B的整数倍,否则要用填充字段来补齐,通常用0来填充。

3.4.2 IPv4地址

1. IP地址简介

在网络中,对于主机的识别要依靠地址。大多数局域网通过网卡的物理地址,来标识一个联网的计算机或其他设备。所谓物理地址,是指固化在网卡EPROM中的地址,这个地址应该保证在全网是唯一的。如果固化在网卡中的地址为112233445566,那么这块网卡插到主机A中,主机A的地址就是112233445566,不管主机A是连接在局域网1上还是在局域网2上,也不管这台主机移到什么位置,它的物理地址一直是112233445566,是不变的,而且不会和世界上任何一台计算机相同。使用物理地址,当主机A发送一帧时,网卡执行发送程序,直接将这个地址作为源地址写入该帧。当主机A接收一帧时,直接将这个地址与接收帧目的地址比较,以决定是否接收。

使用物理地址寻址有以下问题。

(1) 物理地址被固化在网络设备(网卡)中,通常不能被修改。

(2) 物理地址属于非层次化的地址,它只能标识出单个设备,不能标识出该设备连接的是哪一个网络。

针对物理地址存在的问题,采用逻辑地址的编址方案,使用网络层的IP地址。IP地址通过上层软件对各个物理地址进行统一编址,这种软件的方式没有改变物理地址,而是屏蔽了它们,建立了一种IP地址与物理地址的映射关系。这样,在网络层使用IP地址,到了底层,通过映射得到物理地址。

IP地址作为互联网的逻辑地址是有层次的,IP地址为32位长,理论上可以有2^{32}个地址。互联网给每一台上网的计算机分配一个32位长的二进制数字编号,这个编号就是IP地址。

IP地址由地址类别、网络号和主机号3部分组成,其结构如图3-12所示。其中,地址类别用来标识网络类型,网络号用来标识一个逻辑网络,主机号用来标识网络中的一台主机。互联网中的主机至少有一个IP地址,而且这个IP地址是全网唯一的。

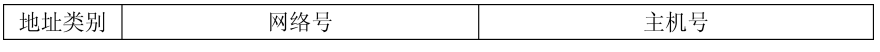


图3-12 IP地址的结构

用二进制数表示的IP地址不方便记忆和阅读,所以采用点分十进制的方法来表示IP地址。即,将32个二进制位每8位单独转成一个十进制数,共4个十进制数,中间用点隔

开。例如，某主机 IP 地址是 11001010 00101001 00001000 00010000，按上面的规则写成 202.41.8.16。

2. IP 地址的分类

IP 协议把 IP 地址分为 5 类，即 A 类、B 类、C 类、D 类和 E 类。5 类 IP 地址的格式如图 3-13 所示。

图 3-13　A～E 类 IP 地址的格式

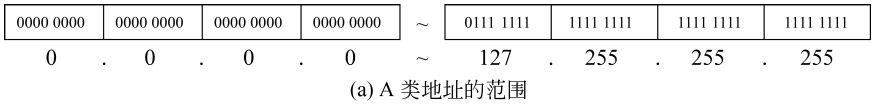
A 类地址适用于大型网络，B 类地址适用于中型网络，C 类地址适用于小型网络，D 类地址用于多播，E 类地址预留。地址的类别可以从 IP 地址的最高 8 位判断，如表 3-3 所示。

表 3-3　IP 地址分类

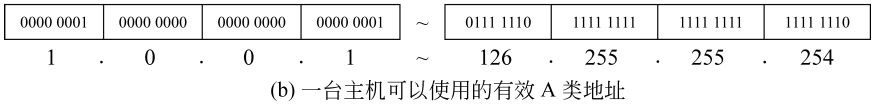
IP 地址分类	高 8 位数值范围	IP 地址分类	高 8 位数值范围
A 类	0～127	D 类	224～239
B 类	128～191	E 类	240～255
C 类	192～223		

1）A 类地址

A 类地址用最高位为 0 表示网络类别，接下来的 7 位表示网络号，低 24 位表示主机号。通过网络号和主机号的位数可以知道，A 类地址的网络数为 $2^7=128$ 个，每个网络中包含的主机数为 $2^{24}=16\ 777\ 216$ 个，A 类网络地址的范围是 0.0.0.0～127.255.255.255，如图 3-14(a)所示。



(a) A 类地址的范围



(b) 一台主机可以使用的有效 A 类地址

图 3-14　A 类地址

由于网络号全 0 和全 1 保留用于特殊目的，所以 A 类地址的有效网络数为 126 个，每个网络中包含的主机数是 $2^{24}-2=16\ 777\ 214$。因此，一台主机能使用的 A 类地址有效范

围为1.0.0.1～126.255.255.254,如图3-14(b)所示。

2) B类地址

B类地址用最高两位为10表示网络类别,接下来的14位表示网络号,低16位表示主机号。因此,B类地址的有效网络数为$2^{14}-2=16\ 382$个,每个网络中包含的主机数为$2^{16}-2=65\ 534$个。B类地址的范围为128.0.0.0～191.254.254.255,一台主机可以使用的B类地址的有效范围是128.1.0.1～191.255.255.254,如图3-15所示。

1000 0000	0000 0000	0000 0000	0000 0000	～	1011 1111	1111 1111	1111 1111	1111 1111
128	0	0	0	～	191	255	255	255

(a) B类地址的范围

1000 0000	0000 0001	0000 0000	0000 0001	～	0111 1110	1111 1111	1111 1111	1111 1110
128	1	0	1	～	191	255	255	254

(b) 一台主机可以使用的有效B类地址

图3-15　B类地址

3) C类地址

C类地址用最高3位为110表示网络类别,接下来的21位表示网络号,低8位表示主机号。因此,C类地址网络个数为2^{21}个,每个网络中包含主机256个。C类地址的范围是192.0.0.0～223.255.255.255,同理,一台主机能够使用的有效范围是192.0.1.1～223.255.254.254,如图3-16所示。

1100 0000	0000 0000	0000 0000	0000 0000	～	1101 1111	1111 1111	1111 1111	1111 1111
192	0	0	0	～	223	255	255	255

(a) C类地址的范围

1100 0000	0000 0000	0000 0001	0000 0001	～	1101 1111	1111 1111	1111 1110	1111 1110
192	0	1	1	～	223	255	254	254

(b) 一台主机可以使用的有效C类地址

图3-16　C类地址

4) D类地址

D类地址用于多播。多播就是同时把数据发送给一组主机,只有那些已经登记可以接收多播地址的主机才能接收多播数据。

5) E类地址

E类地址是预留的,同时也可以用于实验目的,但它们不能分配给主机。

3. 特殊IP地址

在IP中,规定了一些特殊的IP地址,这些IP地址具有特别的作用,不能分配给主机。特殊IP地址如表3-4所示。

(1) 网络地址。网络地址又称为网段地址,是网络号不空而主机号全0的IP地址,即网络本身。例如地址222.220.32.0为网络地址,表示整个网络。

(2) 直接广播地址。网络号不空而主机号全1的IP地址,表示这一网段下的所有用户。例如,222.220.32.255就是直接广播地址,表示222.220.32网段下的所有用户。若其他

表 3-4　特殊 IP 地址

地址类型	网络号	主机号	说　　明
本机地址	全 0	全 0	启动时使用
网络地址	有网络号	全 0	标识一个网络
直接广播地址	有网络号	全 1	在网络外向网络内的所有主机进行广播
内网广播地址	全 1	全 1	在本地网进行广播
回环测试地址	127	任意	回环测试

网络中的主机要向 222.220.32 网段下的所有用户进行广播，则向地址 222.220.32.255 进行广播即可。

(3) 内网广播地址。也叫有限广播地址，是网络号和主机号都是 1 的 IP 地址。向 255.255.255.255 这个地址发送的数据会在本地网络进行广播。

(4) 回环测试地址。网络号为 127 而主机号任意的 IP 地址。最常用的回环测试地址是 127.0.0.1。

4. 私有地址

私有地址专门用于企业内部网络中，不能作为互联网上的地址。内部网络由于不与外部互联，因而可以使用任意 IP 地址。使用私有地址的内部网络在接入因特网时，要使用地址翻译(NAT)将私有地址翻译成合法 IP 地址。在因特网上，这类地址是不能出现的。A、B、C 类地址都有私有地址段。

(1) A 类地址可以使用的私有地址段为 10.0.0.0～10.255.255.255。

(2) B 类地址可以使用的私有地址段为 172.16.0.0～172.131.255.255。

(3) C 类地址可以使用的私有地址段为 192.168.0.0～192.168.255.255。

3.4.3　子网技术

出于对管理、性能和安全方面的考虑，可以把单一网络划分为多个物理网络，并使用路由器将它们连接起来，子网划分技术能够使单个网络地址包含几个物理网络。

1. 划分子网的原因

划分子网的原因主要有以下几方面。

(1) 充分使用地址。由于 A 类网络和 B 类网络的地址空间太大，根本无法在单一网络中用完全部地址。因此，为了能够有效地利用地址空间，必须把可用地址分配给多个较小的网络。

(2) 划分管理职责。当一个网络被划分为多个子网后，每个子网的管理可以由子网管理人员负责，使网络变得更加容易控制。

(3) 提高网络性能。在一个网络中，随着网络用户的增长和主机的增加，网络通信也将变得非常繁忙。如果将一个大型网络划分为若干个子网，并通过路由器将其连接起来，就可以减少网络拥塞。

2. 划分子网的方法

IP 地址共 32 位，分为网络号和主机号两部分。为了划分子网，可以将单个网络的主机

号分为两个部分,一部分用于子网号的编址,另一部分用于主机号编址,如图 3-17 所示。

图 3-17　子网的划分

子网号的位数取决于具体的需要。子网号占用的位越多,则可以分配给主机的位数就越少,也就是在一个子网中包含的主机越少。假设一个 B 类网络地址为 172.16.0.0,将主机号分为两部分,其中 8 位用于子网号,8 位用于主机号,那么这个 B 类地址就被分为 254 个子网,每个子网可以容纳 254 台主机。

3. 子网掩码

将网络地址分为子网络号和主机号,那么网络中各个节点如何知道哪几位是网络号,哪几位是主机号呢? 方法是通过子网掩码来识别网络号和主机号。一个网络有一个子网掩码。子网掩码也是一个用点分十进制表示的 32 位二进制数,其取值的设置规则为:对应 IP 地址中网络号和子网号的位置设置为 1,对应主机号的位置设置为 0。通过判断是否为 1 即可知道哪几位是网络号。标准的 A 类、B 类、C 类地址都有一个默认的子网掩码,如表 3-5 所示。

表 3-5　A 类、B 类、C 类地址默认的子网掩码

地址类型	十进制表示	子网掩码的二进制位
A	255.0.0.0	1111 1111　0000 0000　0000 0000　0000 0000
B	255.255.0.0	1111 1111　1111 1111　0000 0000　0000 0000
C	255.255.255.0	1111 1111　1111 1111　1111 1111　0000 0000

为了识别网络地址,TCP/IP 对子网掩码进行按位与操作。按位与就是两个二进制数的对应位进行与运算。若两个值均为 1,则结果为 1;若其中一个值为 0,则结果为 0。例如,对于 B 类地址 172.16.31.55,若使用 B 类地址的默认子网掩码 255.255.0.0,则其网络号为 172.16.0.0;若使用划分了子网的子网掩码 255.255.255.0,则其网络号为 172.16.31.0。具体运算过程如图 3-18 所示。

IP 地址:172.16.31.55	1010 1100	0001 0000	0001 1111	0011 0111
子网掩码:255.255.0.0	1111 1111	1111 1111	0000 0000	0000 0000
网络号:172.16.0.0	1010 1100	0001 0000	0000 0000	0000 0000

(a) 使用默认子网掩码时的网络号

IP 地址:172.16.31.55	1010 1100	0001 0000	0001 1111	0011 0111
子网掩码:255.255.255.0	1111 1111	1111 1111	1111 1111	0000 0000
网络号:172.16.31.0	1010 1100	0001 0000	0001 1111	0000 0000

(b) 使用划分了子网的子网掩码时的网络号

图 3-18　使用不同子网掩码时网络号不同

在图 3-18 中，子网掩码使用一个字节用于划分子网，这种子网掩码叫边界子网掩码，其特点是子网掩码的取值不是 0 就是 255，非常整齐。但在实际的子网划分中，还会使用一个字节的某几位作为子网号，这就是非边界子网掩码。非边界子网掩码的取值除了 255 外，还有其他值。

4. 划分子网的规则

在 RFC 950 文档中规定了子网划分的规范，其中对网络地址中的子网号做了如下规定。

(1) 由于网络号全为 0 代表的是本网络，所以网络地址中的子网号不能全为 0。

(2) 由于网络号全为 1 表示广播地址，所以网络地址中的子网号也不能全为 1。

例如，划分子网的时候使用 3 位二进制数作为子网号，可以划分为 8 个子网(000、001、010、011、100、101、110、111)，但根据上述规则，全 0 和全 1 的子网不能使用，所以只有 6 个子网可以使用。

RFC950 禁止使用全 0 和全 1 的子网号，全 0 子网会给早期的路由器选择协议带来麻烦，全 1 子网与所有子网的直接广播地址冲突。随着网络技术的发展，很多供应商主机的产品可以支持全 0 和全 1 的子网。所以，当用户要使用全 0 或者全 1 的子网时，需要确认网络中的主机或者路由器是否支持。另外，对于可变长子网划分 VLSM 和无类域间路由 CIDR，由于它们属于现代网络技术，已经不再按照传统的 A 类、B 类、C 类地址的方式工作，因而不存在全 0 子网和全 1 子网的问题，全 0 和全 1 的子网都可以使用。

5. 子网划分实例

为了将网络划分为不同的子网，必须为每一个子网分配一个网络号。在划分子网前，需要确定需要的子网数和每个子网的最大主机数，有了这些信息后，就可以定义每个子网的子网掩码、网络地址的范围和主机号范围。划分子网的步骤如下。

(1) 确定需要多少个子网号来标识网络中的每一个子网。

(2) 确定需要多少个主机号来标识每个子网中的每一台主机。

(3) 定义一个符合网络要求的子网掩码。

(4) 确定标识每一个子网的网络地址。

(5) 确定每一个子网所使用的主机地址范围。

接下来用一个实例说明子网划分的过程。

假设某企业内部有 5 个部门：市场部、行政部、营销部、财务部和后勤部。该企业申请到一个 C 类网络地址 192.99.20.0/24，要对该网络进行子网划分，以确保不同部门的主机处于不同的子网中，提升网络的性能和确保通信的安全，请进行子网的规划。

该企业有 5 个部门，至少需要 5 个子网，则需要从第 4 个字节中划分出 3 位二进制数作为子网号。假设全 0 和全 1 的子网号不可用，则 3 位二进制数可以划分成 6 个子网，取其中的 5 个子网即可。

第 4 个字节划分出 3 位二进制数作为子网号，则剩下 5 位二进制数作为主机号，一共可以容纳的主机数目为 $2^5-2=30$，符合要求。

该划分子网的方案可行，接下来确定子网掩码。子网掩码为 255.255.255.224，如图 3-19 所示。

确定了子网掩码后，就可以确定可用的网络地址了。子网号为 3 位二进制数可以有 8 种取值组合，去掉全 0 和全 1 的两种组合后，还有 6 种组合，即 001、010、011、100、101、110，

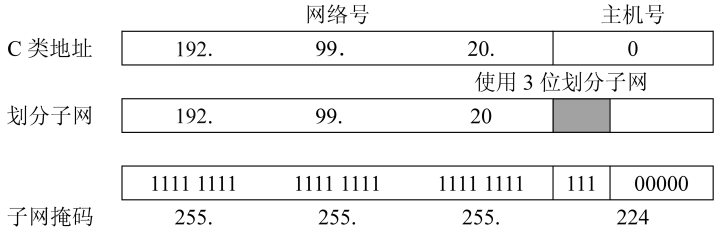


图 3-19　子网掩码划分

把前5种组合分配给公司的5个部门,网络地址分配如图3-20所示。

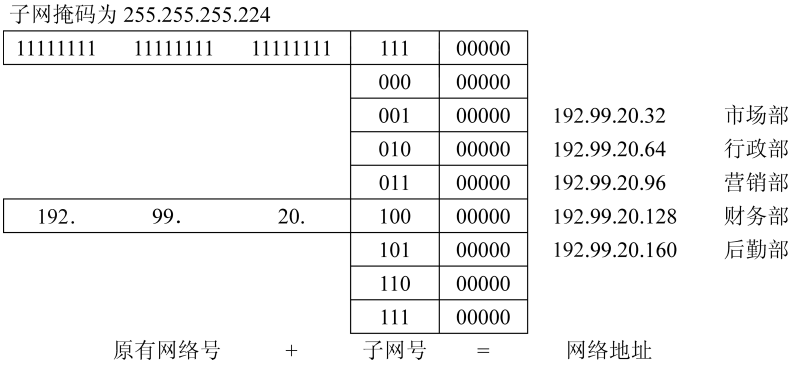


图 3-20　网络地址分配

根据每个子网的网络地址就可以确定每个子网的主机地址范围,如表3-6所示。

表 3-6　子网划分及主机地址范围

部　门	子网网络地址	子网号	子网主机号范围	子网主机地址范围
市场部	192.99.20.32	001	00001～11110	192.99.20.33～192.99.20.62
行政部	192.99.20.64	010	00001～11110	192.99.20.65～192.99.20.94
营销部	192.99.20.96	011	00001～11110	192.99.20.97～192.99.20.126
财务部	192.99.20.128	100	00001～11110	192.99.20.129～192.99.20.158
后勤部	192.99.20.160	101	00001～11110	192.99.20.161～192.99.20.190

6. 可变长子网划分

除了用点分十进制法表示外,子网掩码还可以用网络前缀标记法来表达。前面使用的C类地址默认子网掩码255.255.255.0就是点分十进制表示法。网络前缀法是一种表示子网掩码中网络地址长度的方法。由于网络号是从IP地址高字节以连续的方式从左到右选取若干位作为网络号,因此可以用“IP地址/网络号位数”这种方式来表示。

一个子网掩码为255.255.0.0的B类网络地址172.16.0.0用网络前缀标记法可以表示为172.16.0.0/16。若对这个B类网络进行子网划分,使用主机号的前8位用于子网号,网络号和子网号共计24位,因此,该网络地址的子网掩码为255.255.255.0,使用网络前缀法表示时,子网172.16.20可以表示为172.16.20.0/24。