

第 3 章 广域网与 Internet

本章要点

- 理解广域网的基本概念；
- 理解 Internet 的基本概念；
- 掌握 IP 编址方案；
- 理解子网掩码的概念；
- 掌握子网划分方法；
- 熟悉 DNS 域名系统；
- 掌握 TCP/IP 协议族中的几个重要协议；
- 掌握几种常用接入 Internet 的方法；
- 掌握使用 Wireshark 抓包分析网络访问的方法。

3.1 广域网与 Internet 概述

3.1.1 广域网简介

前面章节介绍的局域网通常适用于较小物理范围、规划统一的环境,主要作用是将各主机互联起来。当主机之间的距离相隔较远时,局域网显然无法胜任。这时候就需要另一种类型的网络——广域网。

与局域网一样,广域网并没有严格的定义。通常广域网是指覆盖全国、全洲范围的计算机网络。广义来说,也可将 Internet 看成一个最大的广域网。

与较小物理范围的局域网相比,广域网具有以下特点:

(1) 覆盖范围更广,可达数百、甚至数千公里。

(2) 作用不同,需要解决的问题也不尽相同。局域网通常是将距离相隔较近的主机互联起来;广域网需要将距离相隔较远的众多主机和小型网络接入到网络中,更关心如何路由和拥塞控制等问题。

(3) 网络环境、拓扑结构更加复杂,很难做到统一规划和管理。局域网一般采用相对简单的星型、总线型或环型拓扑结构;但广域网一般采用更加复杂的树状结构和网状结构,网络中存在更多路径变化。

(4) 数据传输速率较低。

(5) 使用多种传输媒体,有线媒体有光纤、双绞线、同轴电缆等,无线媒体有微波、卫星、红外和激光等。

(6) 数据传输延时更大,由于物理路径较长,中转结点较多,相对传输所需时间也更多。如使用卫星通信的网络,延迟甚至可能达到几秒钟。

(7) 数据传输质量较差,一般来说,广域网中丢包率和误码率更高。

广域网由一些路由器和连接这些路由器的链路组成。路由器可以看作是一台特殊的计算机,完成分组存储转发的功能。路由器之间采用点到点线路连接,为了提高网络的可靠性,通常一个路由器往往与多个路由器相连,增加链路的冗余备份。

一组路由器相互连接构成广域网。一台路由器通常有多个输入输出端口,使它能形成多种不同的拓扑结构,连接多台计算机和网络。广域网示例如图 3-1-1 所示。

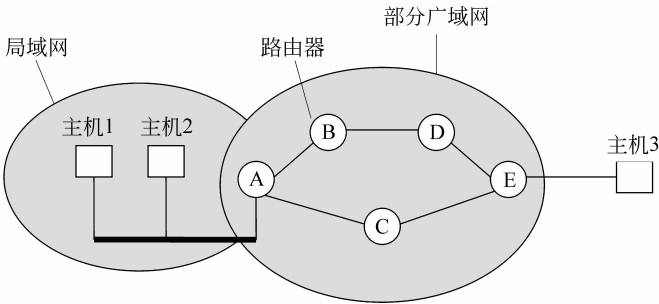


图 3-1-1 广域网示例

当数据在一条线路到达路由器时,路由器必须选择一条出线将数据转发出去。路由决策是在路由器本地进行的。在图 3-1-1 的网络环境中,当一个分组到达路由器 A 的时候,A 可以自行决定该分组应该被送到连接 B 的线路上,还是连接 C 的线路上。A 做出这个决定的过程称为路由算法。

路由算法非常多,可分为静态路由算法和动态路由算法。其中,距离矢量路由算法和链路状态路由算法是两个最常用的动态路由算法。

3.1.2 Internet 简介

Internet 也被称为因特网或国际互联网。它是当今世界上最大的国际性计算机广域网。

1. Internet 的产生及在我国的发展

在 20 世纪 60 年代末,美国建立了一个名为 ARPANET 的军用网络,利用这个网络可使数据通过几条可选路径从一台计算机传送到另一台计算机。

20 世纪 80 年代初,军事信息从 ARPANET 网上分离了出去,剩下的部分就被称为 Internet。Internet 的发展最初是由学术团体推动的,各大学之间建立了宽带连接,彼此利用 TCP/IP 交换数据。政府机构也把自己的网络与 Internet 相连接,此后,Internet 从学术界扩展至工商界,进而冲出美国,使其覆盖范围遍及全球 170 多个国家和地区。

1986 年,我国一些科研部门和高等院校开始研究 Internet 互联技术,这个阶段的网络应用仅限于为少数高等院校和研究机构提供电子邮件服务。1994 年 4 月,中关村地区教育与科研示范网络工程进入 Internet,开通了 Internet 全功能服务。从此我国被国际上正式承认为有互联网的国家。到今天为止,我国已建成中国教育科研计算机网(CERNET)、中国科技网(CSTNET)、中国公用计算机互联网(ChinaNET)、中国网通公

用互联网(CNCNET)、中国国际经济贸易互联网(CIETNET)、宽带中国(China169)等十大网络,并与 Internet 连通。

2. Internet 提供的服务

Internet 提供的服务主要是通过信息交流、信息获取和资源共享等功能实现的。常见的 Internet 服务包括 WWW 服务、信息检索、文件传输、电子邮件、远程管理、即时通信以及 BBS 等。

1) WWW 服务

万维网(World Wide Web, WWW)又称 3W,是 Internet 上的一个大型信息媒介,由分布全球的 Web 服务器组成的超文本资源大集合。WWW 采用客户机/服务器的工作方式,客户端软件是 WWW 浏览器(如 Internet Explorer),服务器端则运行 WWW 服务器程序,当用户使用浏览器向服务器发出查询请求时,服务器检索存储在 WWW 服务器上的所有资源,并将查询到的结果发送给客户端浏览器。上网时在地址栏中输入网址或单击超级链接,网页随后打开的过程就反映了这个工作方式。

2) 搜索引擎

搜索引擎是一种能够通过 Internet 接受用户的查询指令,并向用户提供符合查询要求的信息或相关链接的系统。搜索引擎既有供用户检索的界面,又有供检索服务的网站。所以,也可把搜索引擎称为 Internet 上具有检索功能的软件。

搜索引擎的使用技巧:

(1) 正确选择搜索引擎。

当用户尚未形成明确的检索概念,或仅对某一专题作泛泛浏览时,可先用目录搜索引擎的合适类目进行逐个浏览,直到发现相关的网址。若需进一步检索,则再从这些网址中寻找合适的检索关键词,利用全文搜索引擎或元搜索引擎进行检索。当用户已经明确了检索词,但对全文搜索引擎不够熟悉或想节省在多个全文搜索引擎之间进行转换的时间时,则可选用元搜索引擎作试探性的起始检索。

(2) 选择合适的关键词。

关键词是搜索引擎对网页进行分类的依据,选择合适的关键词可以用较短的时间检索到更多的信息,关键词的选择可以从以下几个方面考虑:

① 正确使用布尔运算符。各搜索引擎一般都支持使用布尔运算符进行检索。布尔运算符主要包括“与”、“或”、“非”,不同的搜索引擎有不同的表示方法。通常,用户在检索时需要用不同的布尔运算符把检索词与检索词连接起来,这样可以更为准确地表达检索要求,使检索结果更符合用户要求。

② 用双引号进行精确检索。当用户输入较长的检索词时,搜索引擎往往会反馈大量不需要的信息。如果要查找的是一个特定的词组、短语或一句确定的句子,最好的办法就是将它们加上西文双引号,这样得到的结果会完全符合双引号中的关键词,精确度更高。

③ 避免使用太常见的关键词。搜索引擎对常见的检索存在缺陷,因为这些词出现的频率太高,用它们进行检索往往无法找到有用的内容。当检索结果太多太乱时,应该尝试使用更多的关键词,或使用布尔运算符“非”来缩小检索范围。

④ 尝试使用近义词。如果检索返回的结果较少,可以适当扩大检索范围。除了去掉

一些修饰词外,还可以尝试使用近义词。比如用“计算机”代替“电脑”就可以找到不同的信息。

(3) 合理利用“网页快照”。

在利用 Internet 检索信息时,经常会遇到检索到的网页无法打开的情况,原因可能是网站已经搬走,转向地址未知,或是该网页已经从搜索引擎的数据库中删除了。如果检索时使用的是 Google、百度、Bing 等搜索引擎,可以在每个检索结果后发现类似“网页快照”、“缓存页”等超级链接,单击超级链接就可以观看该网页的快照内容,而且网页内的所有关键词都用不同的颜色进行了区分,比直接打开网页后慢慢查找要方便很多。

(4) 研究每种搜索引擎的帮助说明。

Google、百度、Bing 等许多网络信息检索工具都提供了高级检索功能,有些还附有详细的使用说明及检索技巧。花点时间了解一下相关内容,将有助于快速、高效地检索到所需的信息资源。

3) 电子邮件

电子邮件(E-mail)是 20 世纪 70 年代出现的一种新型通信方式,它的出现给人类的交流方式带来了一次新的革命。电子邮件有几个显著特点:

- 电子邮件可以与一个人交换信息,也可以同时与多个人交换信息。
- 电子邮件的内容无论有多长,传递的路径无论有多远,只要收信人地址正确,通过 Internet,片刻就可以到达目的地。
- 电子邮件使用简单、方便,用户无论走到哪里都能通过自己的电子邮箱收发信件。
- 电子邮件的内容除了文本外,还可以包含声音、图像、应用程序等各类计算机文件。

(1) 电子邮件地址。

电子邮件地址格式为:用户名@邮件服务器域名,如 admin@sina.com。其中,符号“@”读作 at,用以分隔用户名和邮件服务器域名。邮件服务器域名在 Internet 上是唯一的,用户名在该邮件服务器上也是唯一的,从而保证了 Internet 上用户电子邮件地址的唯一性。

(2) 电子邮件的收发。

在 Internet 上电子邮件的收发过程如下:

① 发送端客户使用电子邮件收发软件发送邮件,通过网络,提交给用户设置的邮件发送服务器。

② 邮件发送服务器将邮件转发至接收端客户所使用的邮件接收服务器。

③ 接收端客户使用电子邮件收发软件,通过网络,访问邮件接收服务器,接收邮件。

电子邮件的收发软件有很多,最为流行的是 Outlook Express、FoxMail 等,它们提供的功能主要有管理多个邮件和新闻账号、快速方便地浏览邮件、发送和阅读安全邮件、完善的邮件地址管理功能、查找并下载感兴趣的新闻组等。

4) 文件传输

Internet 上常有大量文件需要传输,文件传输协议(FTP)能使用户连接上一个远程计算机,将文件从远程计算机复制到本地计算机(此过程称“下载”),或将本地计算机的文

件传送到远程计算机上(此过程称“上传”)。使用 FTP 实现上传或下载的方法有三种: FTP 命令、浏览器和 FTP 应用软件,其中“浏览器”的使用方法是:在高版本的 IE 或其他浏览器的地址栏中输入 ftp://ftp. hhh. hhh. hhh,如输入 ftp://ftp. pku. edu. cn,就可以访问相应的 FTP 服务器,如图 3-1-2 所示。



图 3-1-2 FTP 网站内容

5) BBS

公告板服务(Bulletin Board Service,BBS)是一种电子信息服务系统,它向用户提供一块公共电子白板,每个注册用户都可以在上面发布信息或提出看法。现在大多数网站上都建立了自己的 BBS 系统,供网民通过网络结交更多的朋友,发表各自的看法。一些大型的 BBS 系统还按照不同的主题“分类”,访问者可以根据自己的喜好浏览感兴趣的主题或参与讨论,并发表自己的看法。图 3-1-3 所示是著名的“天涯论坛”。



图 3-1-3 天涯论坛

3.2 IP 地址及域名系统

3.2.1 IP 地址

IP 地址用来唯一地标识网络中的一个通信实体,如一台主机或路由器。

1. IP 地址及表示方法

目前 Internet 中使用的网际协议规范为 IP 版本 4,称为 IPv4。在 IPv4 中 IP 地址共 32 位,采用“点分十进制法”表示,即以点号划分成 4 段十进制数表示。每段数字的取值只能在 0~255 之间,如 202.127.18.49,128.56.156.18 等。由于计算机网络中实际使用的是二进制数字格式,因此,IP 地址的每段对应于 8 个二进制数,4 段实际上对应了 4 字节,即 32 位二进制数。

按照国际互联网组织的规定,IP 地址的 4 段分成主机所在的网络号部分和主机号部分。由于 IP 地址的总长度只有 4 字节(32 位),选择长的网络号可容纳大量网络,但限制了每个网络的大小;选择长的主机号意味着每个物理网络能包含大量的主机,但限制了网络的总数。

为了适应不同规模网络的需要,IP 地址的设计者将 IP 地址划分为 5 个不同的地址类别,其中 A、B、C 三类最为常用,它们用作主机地址,D 类地址用作组播,E 类地址保留未用。IP 编址方案中,地址的前 4 位决定了该地址所属的类别,并且确定了如何将地址的其余部分划分网络号和主机号。5 类 IP 地址的格式如图 3-2-1 所示。

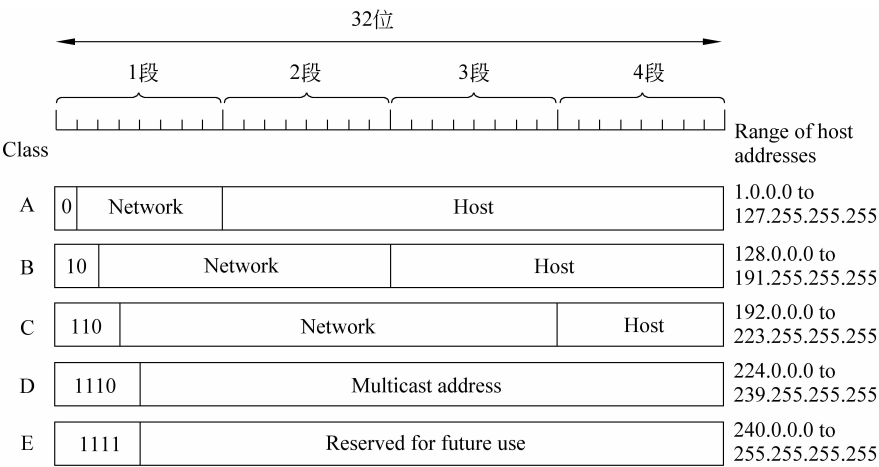


图 3-2-1 IP(IPv4)地址格式

(1) A 类 IP 地址的第 1 个字节取值范围为 1~127,用来表示网络号部分,其后 3 个字节一起表示主机号部分。A 类地址用于大型计算机网络。

(2) B 类 IP 地址的第 1 个字节取值范围为 128~191,地址中前两个字节合起来表示网络号,后两个字节合起来表示主机号。B 类地址用于中型计算机网络。

(3) C 类 IP 地址的第 1 个字节取值范围为 192~223,地址中第 1~第 3 个字节合在

一起表示网络号,第 4 段为主机号。C 类地址可用于计算机数量少的小型网络。

(4) D 类 IP 地址用于组播。一旦组播传输组建立,任何发送到组播地址的数据报将传送到该组中的每台主机。

(5) E 类 IP 地址暂时不用,用于实验和留作将来使用。

表 3-2-1 总结了 A、B、C 三类网络类别关于最大网络数、最大主机数和 IP 地址范围的比较。

表 3-2-1 各 IP 地址类别比较

地址类别	网络号位数	最大网络数	主机号位数	最大主机数	IP 地址范围
A	7	$2^7=128$	24	$2^{24}=16\,777\,216$	0.0.0.0~127.255.255.255
B	14	$2^{14}=16\,384$	16	$2^{16}=65\,536$	128.0.0.0~191.255.255.255
C	21	$2^{21}=2\,097\,152$	8	$2^8=256$	192.0.0.0~223.255.255.255

2. 保留 IP 地址

除了给每台计算机分配一个 IP 地址外,有时也使用 IP 地址表示整个网络或者一组计算机。这些特殊的地址,不用来分配给主机,具有特殊含义,称为保留 IP 地址。

(1) 本机地址,网络号和主机号部分全为 0,即地址 0.0.0.0,表示本计算机,仅用于计算机启动。

(2) 网络地址,主机号部分全为 0,用来表示一个网络。如网络地址 222.204.251.0 表示一个 C 类网络,网络号为 222.204.251。

(3) 本地网络广播地址,网络号和主机号部分全为 1,即地址 255.255.255.255,也称有限广播地址,用于在本地网络上进行广播。

(4) 定向广播地址,主机号部分全为 1,用于向指定网络发送广播数据。如定向广播地址 222.204.251.255,用来向 222.204.251.0 网络中的所有主机发送广播数据。

(5) 回送地址,网络号为 127,主机号任意,用来测试 TCP/IP 协议栈及本机应用程序间的通信。如回送地址 127.0.0.1。

观察这些保留地址的规定,可以发现一些规律:一般来说,可以把全 0 理解成“本”,全 1 理解成“所有”。表 3-2-2 所示为各保留 IP 地址。

表 3-2-2 保留 IP 地址

地 址 类 型	网络号	主机号	用 途
本机地址	全 0	全 0	启动时使用
网络地址	某网络号	全 0	表示一个网络
本地网络广播地址	全 1	全 1	在本地网络上广播
定向广播地址	某网络号	全 1	在指定网络上广播
回送地址	127	任意	测试 TCP/IP 协议栈及本机应用程序间的通信

3. IP 地址的分配和管理

为了保证 Internet 中每台主机的 IP 地址的唯一性,IP 地址不能随意定义,Internet 上主机的 IP 地址必须向 Internet 域名与地址管理机构(ICANN)提出申请才能得到。

一般来说,只有最大型的网络服务提供商才需要直接与 ICANN 联络,获得 IP 地址前缀。在实际使用中,如果要将网络接入 Internet,大多数单位不会直接向中央管理机构申请,而是由本地 ISP 分配一个有效 IP 地址。

4. 私网 IP 地址

对于大多数局域网而言,并不需要为每台主机申请一个唯一的 IP 地址。这些局域网主机可以通过 NAT(地址转换协议)路由器或 Proxy(代理)服务器等方式与 Internet 进行通信。为了便于管理,专门划分了一些地址段用于局域网主机。这些地址段不能用于 Internet 公网上,只能用于局域网内部。在不同的局域网内部,这些地址可以重复,但同一局域网内部,也必须保证地址的唯一性。这些地址段如下:

- 10.0.0.0~10.255.255.255
- 172.16.0.0~172.31.255.255
- 192.168.0.0~192.168.255.255

5. IPv6 简介

IPv4 为 TCP/IP 族和整个 Internet 提供了基本的通信机制。在 IPv4 中,IP 地址理论上可以标识的地址数为 $2^{32}=4\,294\,967\,296$,但实际中真正可以有效标识的地址数只有大约 1 亿个。随着各种智能设备、无线网络设备的出现,Internet 用户的数量呈爆炸性的增长,IP 地址已经出现短缺现象。除此之外,IPv4 还存在其他不足之处。因此,互联网工程任务组(Internet Engineering Task Force ,IETF)设计出下一代网际协议规范 IPv6 (Internet Protocol version 6)。IPv6 保留了 IPv4 所赖以成功的许多特点,并做了一定修改。

IPv6 的最大变化就是使用了更大的地址空间(IPv6 地址长度为 128 位,而 IPv4 仅 32 位)。这种方案足以使地球表面每平方米上的 IP 地址数量超过 10^{24} 个。如果以每微秒分配一百万个地址的速度进行分配,需要 10^{20} 年的时间才能将所有可能的地址分配完毕。

IPv6 使用 128 位地址长度,解决了地址不够的问题,但也为使用者带来了难以阅读、难以输入地址的问题。IPv6 的设计者们建议采用“冒分十六进制法”表示 IPv6 地址,即把地址分为 8 段,各段之间使用冒号分隔,每段包括 16 个二进制位,并以十六进制书写。例如,68E6: 8c64: FFFF: FFFF: 0: 1180: 96A: FFFF。IPv6 可以使用以下几种方法简化地址:

- (1) 允许零压缩。即一连串重复的零可以用一对冒号代替,如 700E: 0: 0: 0: 0: 0: AC21: B3 可以写成 700E: : AC21: B3。
- (2) “冒分十六进制表示法”可以结合“点分十进制表示法”的后缀,如 0: 0: 0: 0: 0: 0: 192.168.0.1。

与 IPv4 相比,IPv6 新特性包括以下 7 点:

- (1) 更大的地址。IPv6 把 IPv4 的 32 位地址增大到 128 位,在可预见的将来地址是不会耗尽的。
- (2) 扩展的地址层次结构。IPv6 使用更大的地址空间创建额外级别的地址层次,如可以定义 ISP 的地址层次。

- (3) 灵活的首部格式。与 IPv4 中固定格式的数据报首部不同,IPv6 使用了一组可选的首部。
- (4) 增强的选项。IPv6 还包含了新的选项,提供 IPv4 所不具备的新功能。
- (5) 对协议扩展的保障。IPv6 一个显著变化就是允许协议新增特性,这样可以摆脱要全面指定所有细节的状况。
- (6) 支持自动配置和重新编号。IPv6 允许孤立网络上的计算机给自己分配地址,不依赖于路由器或手工配置就可以进行通信。协议还包括允许管理员动态重新给网络编号的能力。
- (7) 支持资源分配。IPv6 提供了“数据流提取”和“分类服务说明”两种方法,实现对网络资源的预分配。
- 当然,IPv6 并非十全十美,不可能解决所有问题。IPv6 只能在发展中不断完善,也不可能在一夜之间发生,过渡需要时间和成本,但从长远看,IPv6 有利于互联网的持续和长久发展。目前,国际互联网组织已经成立了两个专门工作组,制定相应的国际标准。

3.2.2 划分子网

前面介绍的 IPv4 的地址编址方案中,使用了两层地址结构:网络号+主机号。这样做最大的优点是使路由器中的路由表不会太大。路由器不必为每个目的主机维持一个路由表项,而只需要为每个网络维持一个路由表项即可。但这种两层地址结构的设计存在着一个严重的缺陷:没有考虑到数量迅猛增加的小型网络。

现在越来越多的单位开始拥有多个小型网络,如果所有网络都分配不同的网络号,不仅会造成 IP 地址的大量浪费,而且会造成路由器中的路由表项过大。因此,需要开发某种机制来区别同一个地点内的多个逻辑网络,作为 Internet 的另一个层次。

目前应用最广泛的一种方法是采用子网编址技术,增加分层的方法组织这些逻辑子网,并且在它们之间进行路由。对于 Internet 来说,它们是一个网络(具有相同网络号),但对于网络内部的各子网来说,各子网具有不同的子网号,可以进行子网间路由。

1. 子网编址和子网掩码

子网编址技术,使用子网掩码将 IP 地址中主机号部分进一步划分为子网号和主机号。使用了子网编址技术后,IP 地址实际包含 3 个部分:网络号、子网号和主机号,其中子网号和主机号部分是由原先 IP 地址的主机号部分分割而成的,如图 3-2-2 所示。

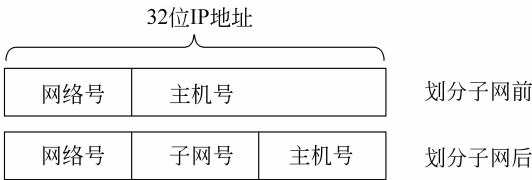


图 3-2-2 划分子网前后 IP 地址结构比较

子网掩码共 32 位,类似 IP 地址,使用“点分十进制法”表示。但子网掩码是由连续的“1”加上连续的“0”组成。将 IP 地址与子网掩码对应起来看,子网掩码中连续“1”的部分,

对应于 IP 地址中的网络号和子网号(根据 IP 地址的第一个字节,可以直接判断出该地址所属网络类型,从而得到网络号所占位数,剩余部分即为子网号),子网掩码中连续“0”的部分,对应于 IP 地址中的主机号。

例如,某公司有两个部门,部门 A 有 60 台主机,部门 B 有 30 台主机。如果直接给其分配两个 C 类网络地址(每个 C 类网络拥有最大主机数 256,减去主机号全 0 和全 1 的保留地址,实际可供使用的 IP 地址有 254 个),存在着巨大的地址浪费。

使用子网编址和子网掩码技术后,可采用子网掩码 255.255.255.192,将一个 C 类网络如 222.204.251.0 的 8 位主机号部分划分成 2 位子网号部分和 6 位主机号部分,这样为该公司两个部门分配两个子网即可。如图 3-2-3 所示。

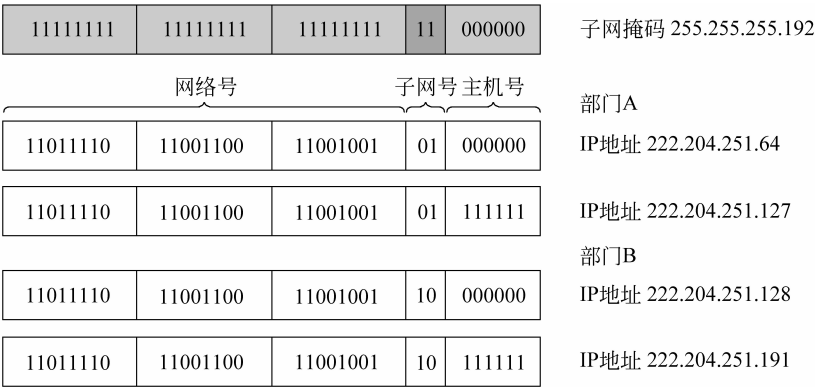


图 3-2-3 子网划分示例

注意：划分子网时,与网络号和主机号一样,子网号也不可为全 0 或全 1。

- 部门 A 所在子网 IP 地址段为 222.204.251.64~222.204.251.127,子网掩码为 255.255.255.192。地址段内所有 IP 地址具有相同网络号 222.204.251,相同子网号 01(二进制)。减去主机号全 0 和全 1 的地址,部门 A 实际可使用的 IP 地址为 62 个,多于所需的 60 个地址,满足要求。
- 部门 B 所在子网 IP 地址段为 222.204.251.128~222.204.251.191,子网掩码为 255.255.255.192。地址段内所有 IP 地址具有相同网络号 222.204.251,相同子网号 10(二进制)。减去主机号全 0 和全 1 的地址,部门 B 实际可使用的 IP 地址为 62 个,多于所需的 30 个地址,满足要求。

虽然划分子网有利于提高 IP 地址的使用效率,减少广播域的范围,改善网络性能,但每一次划分子网也会造成一定的 IP 地址浪费。对于上例而言,一个 C 类网络实际可供使用的 IP 地址数量为 254 个。但是划分子网后,对于子网号为全 0 和全 1 的所有地址将不能使用,即 IP 地址为 222.204.251.0~63(子网号全 0)和 222.204.251.192~222.204.251.255(子网号全 1)的地址不可使用。实际可供使用的 IP 地址,只有部门 A 所在子网的 62 个 IP 地址,和部门 B 所在子网的 62 个 IP 地址,造成了超过一半地址的浪费。