

第 5 章 互联网与 IP

互联网是将物理网络相互连接形成的计算机网络。那么,为什么要进行网络互联? 网络互联的解决方案都有哪些? 本章将对网络互联的这些基本问题进行讨论。对目前应用最广泛的 IP 互联问题,将在第 6~9 章详细介绍。

5.1 互联网

有线以太网和无线局域网是两种使用比较广泛的局域网络,它们能够在一定的范围内提供高速的数据通信服务。除了以太网和无线局域网外,世界上存在着各种各样的物理网络,每种物理网络都有其与众不同的技术特点。这些网络有的提供短距离高速服务(如以太网、无线局域网等),有的提供长距离大容量服务(如 DDN 网、ATM 网等)。到目前为止,没有一种物理网络能够满足所有应用的需求。因为在介质访问控制方法、寻址机制、分组最大长度、差错恢复技术、状态报告和用户接入等方面存在很大差异,所以不同种类的物理网络之间不能直接相连,形成了相互隔离的物理网络孤岛,如图 5-1 所示。

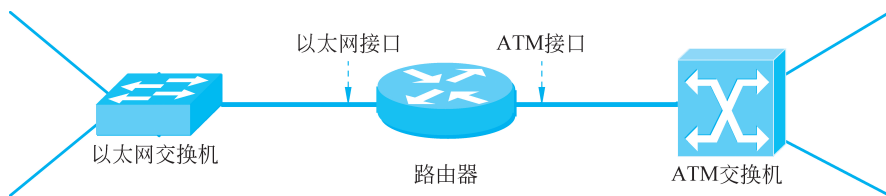


图 5-1 相互隔离的物理网络

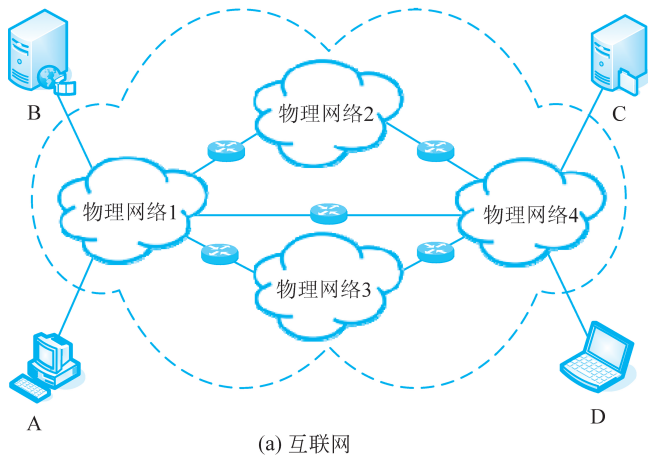
随着网络应用的深入和发展,用户越来越不满足物理网络相互隔离的现状。不但一个物理网络上的用户有与另一个物理网络上的用户通信的需要(如图 5-1 中的用户 A 与用户 B 需要直接进行在线通信或相互发送电子邮件),而且一个物理网络上的用户也有共享另一个网络上资源的需求(如图 5-1 中的用户 A 和用户 C 需要共享服务器 B 中的文件)。在用户强大需求的推动下,互联网络诞生了。

连接物理网络之间的设备称为路由器。路由器通常具有两个或多个物理网络接口,用于连接两个或多个物理网络。图 5-2 显示了一个连接两个物理网络的路由器,该路由器具有两个物理网络接口,一个接口与以太网交换机相连,另一个接口与 ATM 交换机相连。该路由器接收以太网交换机转发的以太帧,并将其转换成 ATM 帧后发送到 ATM 交换机。同时,它也可以从 ATM 交换机接收 ATM 帧,并将其转换成以太帧后发送给以太网交换机。

互联网屏蔽了各个物理网络的差别(如寻址机制的差别、分组最大长度的差别、差错恢复的差别等),隐藏了各个物理网络实现细节,为用户提供通用服务(universal service)。因此,用



户常常把互联网看成一个虚拟网络(virtual network)系统,如图 5-3 所示。这个虚拟网络系统是对互联网结构的抽象,它提供通用的通信服务,能够将所有主机互联起来,实现全方位的通信。



5.2 网络互联解决方案

网络互联是 ISO/OSI 参考模型的网络层或 TCP/IP 体系结构的互联层需要解决的问题。广义上讲,网络互联可以采用面向连接的解决方案和面向非连接的解决方案。

5.2.1 面向连接的解决方案

面向连接的解决方案要求两个结点在通信时建立一条逻辑通道,所有信息单元都沿着这

条逻辑通道传送。路由器将一个网络中的逻辑通道连接到另一个网络中的逻辑通道,最终形成一条从源结点至目的结点的完整通道。

在图 5-4 中,结点 A 和结点 B 通信时形成一条逻辑通道。该通道经过物理网络 1、物理网络 2 和物理网络 4,并利用路由器 i 和路由器 m 连接起来。一旦该逻辑通道建立起来,结点 A 和结点 B 之间的后续信息传输就会沿着该通道进行。

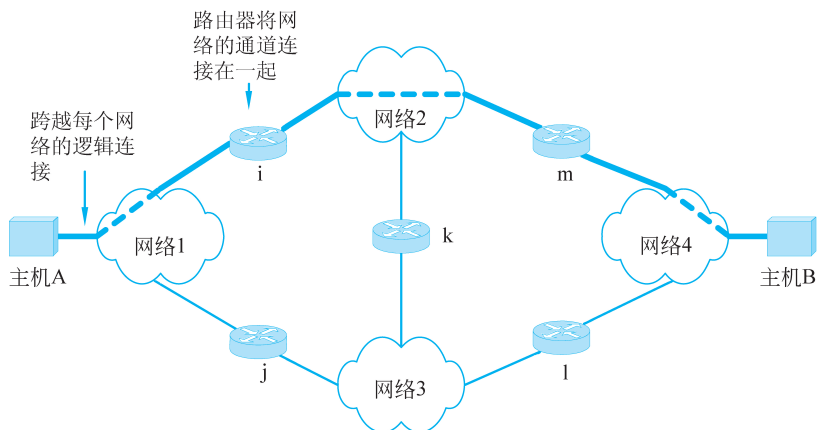


图 5-4 面向连接的解决方案

在面向连接的解决方案中,通信开始时建立逻辑通道的过程就是路由选择。由于后续的数据传输沿逻辑通道顺序进行,因此,源结点只需通知物理网络这些后续数据发往的逻辑通道号,物理网络就会将它们正确传输到目的地址。逻辑通道上的各结点不需要对后续数据进行路由选择,后续数据也不需要携带源地址、目的地址等信息。但是,面向连接的解决方案可以看成一种有状态的解决方案,它对互联网中间结点要求较高。每个中间结点不但在通信开始时要进行路由选择(建立逻辑通道),而且在通信过程中要保持和维护逻辑通道状态。由于面向连接的解决方案的实现较为复杂,因此,尽管很多学者在这方面做了很大的努力,但是始终没有被业界接受。

5.2.2 面向非连接的解决方案

与互联网面向连接的解决方案不同,面向非连接的解决方案并不需要建立逻辑通道。网络中的信息单元都被独立对待,这些信息单元经过一系列的网络和路由器,最终到达目的结点。

图 5-5 显示了一个面向非连接的解决方案示意图。当主机 A 需要发送一个数据单元 P1 到主机 B 时,主机 A 首先进行路由选择,判断 P1 到达主机 B 的最佳路径。如果它认为 P1 经过路由器 i 到达主机 B 是一条最佳路径,那么主机 A 就将 P1 投递给路由器 i。路由器 i 收到主机 A 发送的数据单元 P1 后,根据自己掌握的路由信息为 P1 选择一条到达主机 B 的最佳路径,从而决定将 P1 传递给路由器 k,还是传递给路由器 m。这样,P1 经过多个路由器的中继和转发,最终将到达目的主机 B。

如果主机 A 需要发送另一个数据单元 P2 到主机 B,那么主机 A 同样需要对 P2 进行路由选择。在面向非连接的解决方案中,由于设备对每一数据单元的路由选择独立进行,因此,数据单元 P2 到达目的主机 B 可能经过一条与 P1 完全不同的路径。

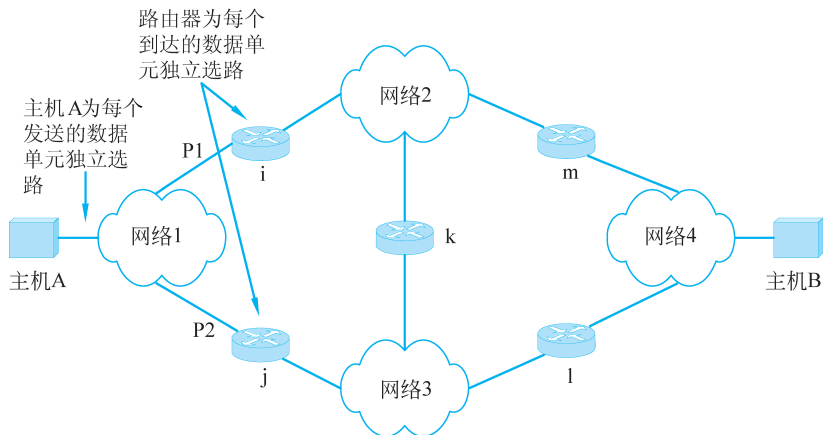


图 5-5 面向非连接的解决方案

在面向非连接的互联网解决方案中,由于中间结点独立地对待每个数据单元,因此,每个数据单元都需要携带完整的源地址、目的地址等信息,以便中间结点为它们选路。另外,面向非连接的解决方案并不能保证发送结点发送的数据单元按顺序到达目的结点。但是,面向非连接的解决方案是一种无状态的解决方案,中间结点不需要维护前后传输数据单元的因果关系,因此,实现起来比较简单。事实上,目前流行的互联网都采用了这种方案。

IP(Internet Protocol)和IPX(Internet work Packet eXchange)协议是两种比较重要的互联网协议,它们都采用了面向非连接的互联网解决方案。其中,IPX 协议主要用于 Novell 网络,有些路由器(如 Cisco 路由器、华为路由器等)也可对其进行支持。但是,随着 Novell 网络市场份额的逐渐萎缩,IPX 协议已风光不再,逐渐被兴起的 IP 取代。

IP 是面向非连接的互联解决方案中设计最成功、应用最广泛的互联协议。尽管 IP 不是国际标准,但由于它效率高、互操作性好、实现简单、比较适合于异构网络,因此,被众多著名的网络供应商(如 IBM、Microsoft、Novell、Cisco 等公司)采用,成为事实上的标准。支持 IP 的路由器称为 IP 路由器,IP 处理的数据单元称为 IP 数据包(有时又称为 IP 数据分组)。

实际上,世界上最具影响力的 Internet(因特网,又称国际互联网)就是一种计算机互联网。它是由分布在世界各地的、数以万计的、各种规模的计算机网络借助网络互联设备——路由器,相互连接形成的全球性互联网。这个正以惊人速度发展的 Internet 采用的互联协议就是 IP。高效、可靠的 IP 为 Internet 的发展起到了不可低估的作用。

5.3 IP 与 IP 层服务

如果说 IP 数据包是 IP 互联网中行驶的车辆,那么 IP 就是 IP 互联网中的交通规则,连入互联网的每台计算机以及处于十字路口的路由器都必须熟知和遵守该交通规则。发送数据的主机需要按 IP 装载数据,路由器需要按 IP 指挥交通,接收数据的主机需要按 IP 拆卸数据。满载着数据的 IP 数据包从源主机出发,在沿途各个路由器的指挥下,就可以顺利地到达目的主机。

IP 精确定义了 IP 数据包格式,并且对数据包寻址和路由、数据包分片和重组、差错控制和处理等做出了具体规定。

5.3.1 IP 互联网的工作机理

图 5-6 给出了一个 IP 互联网工作机理示意图,它包含两个以太网和一个广域网,其中主机 A 与以太网 1 相连,主机 B 与以太网 2 相连,两台路由器除了分别连接两个以太网外,还与广域网相连。从图 5-6 中可以看到,主机 A、主机 B、路由器 X 和路由器 Y 都加有 IP 层并运行 IP。由于 IP 层具有将数据单元从一个网转发至另一个网的功能,因此,互联网上的数据可以进行跨网传输。

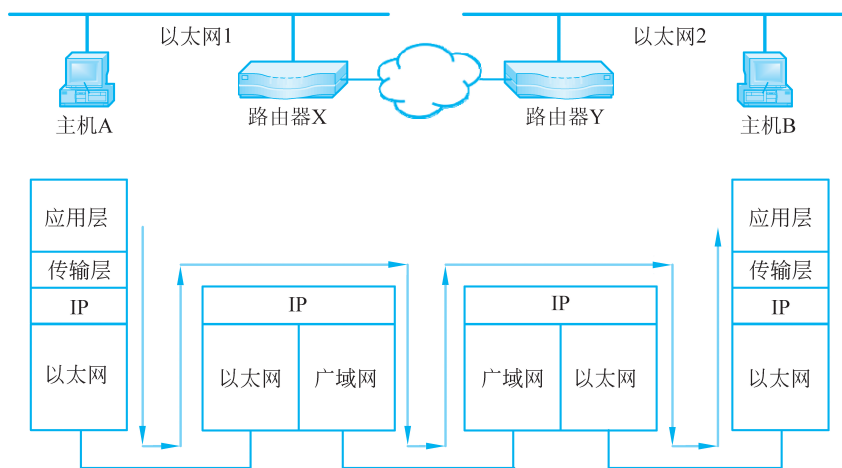


图 5-6 IP 互联网工作机理示意图

如果主机 A 发送数据至主机 B,IP 互联网封装、处理和投递该信息的过程如下:

- (1) 主机 A 的应用层形成要发送的数据,并将该数据经传输层送到 IP 层处理。
- (2) 主机 A 的 IP 层将该数据封装成 IP 数据包,并对该数据包进行路由选择,最终决定将它投递到路由器 X。
- (3) 主机 A 把 IP 数据包送交给它的以太网控制程序,以太网控制程序将其封装成以太网帧后,传递到路由器 X。
- (4) 路由器 X 的以太网控制程序收到主机 A 发送的以太网帧后,进行拆封和处理,将得到的 IP 数据包送到它的 IP 层处理。
- (5) 路由器 X 的 IP 层对该 IP 数据包进行处理。经过路由选择得知该数据必须穿越广域网,才能到达目的地。
- (6) 路由器 X 将 IP 数据包送到它的广域网控制程序,封装成广域网帧。
- (7) 广域网控制程序负责将封装后的广域网帧从路由器 X 传递到路由器 Y。
- (8) 路由器 Y 的广域网控制程序将收到的广域网帧拆封后,提交给它的 IP 层处理。
- (9) 与路由器 X 相同,路由器 Y 对收到的 IP 数据包进行处理。通过路由选择得知,路由器 Y 与目的主机 B 处于同一以太网,可直接投递到达。
- (10) 路由器 Y 再次将该 IP 数据包转交给自己的以太网控制程序,并封装成以太网帧。
- (11) 以太网控制程序负责把封装后的以太网帧由路由器 Y 传送到主机 B。
- (12) 主机 B 的以太网控制程序将收到的以太网帧拆封后,提交给它的 IP 层处理。
- (13) 主机 B 的 IP 层处理该 IP 数据包,在确定数据目的地为本机后,将 IP 数据包的数据

部分经传输层提交给应用层。

5.3.2 IP 层服务

互联网应该屏蔽低层网络的差异,为用户提供通用的服务。具体地讲,运行 IP 的互联层可以为其高层用户提供的服务具有如下 3 个特点:

(1) 不可靠的数据投递服务。这意味着 IP 不能保证数据包的可靠投递,IP 本身没有能力证实发送的 IP 数据包是否被正确接收。数据包可能在线路延迟、路由错误、数据包分片和重组等过程中受到损坏,但 IP 不检测这些错误。在错误发生时,IP 也没有可靠的机制通知发送方或接收方。

(2) 面向无连接的传输服务。它不管数据包沿途经过哪些结点,甚至也不管数据包起始于哪台计算机、终止于哪台计算机。从源结点到目的结点的每个数据包都可能经过不同的传输路径,而且在传输过程中数据包有可能丢失,有可能正确到达。

(3) 尽最大努力投递服务。尽管 IP 层提供的是面向非连接的不可靠服务,但是 IP 并不随意丢弃数据包。只有在系统的资源用尽、接收数据错误或网络故障等状态下,IP 才被迫丢弃报文。

5.3.3 IP 互联网的特点

IP 互联网是一种面向非连接的互联网络,它对各个物理网络进行高度抽象,形成一个大的虚拟网络。总体来说,IP 互联网具有如下特点:

(1) IP 互联网隐藏了低层物理网络细节,为上层用户提供通用的、一致的网络服务。因此,尽管从网络设计者角度看,IP 互联网由不同网络借助 IP 路由器互联而成,但从用户观点看,IP 互联网是一个单一的虚拟网络。

(2) IP 互联网不指定网络互联的拓扑结构,也不要求网络之间全互联。因此,IP 数据包从源主机至目的主机可能经过若干中间网络。一个网络只要通过路由器与 IP 互联网中的任意一个网络相连,它就具有访问整个互联网的能力,如图 5-7 所示。

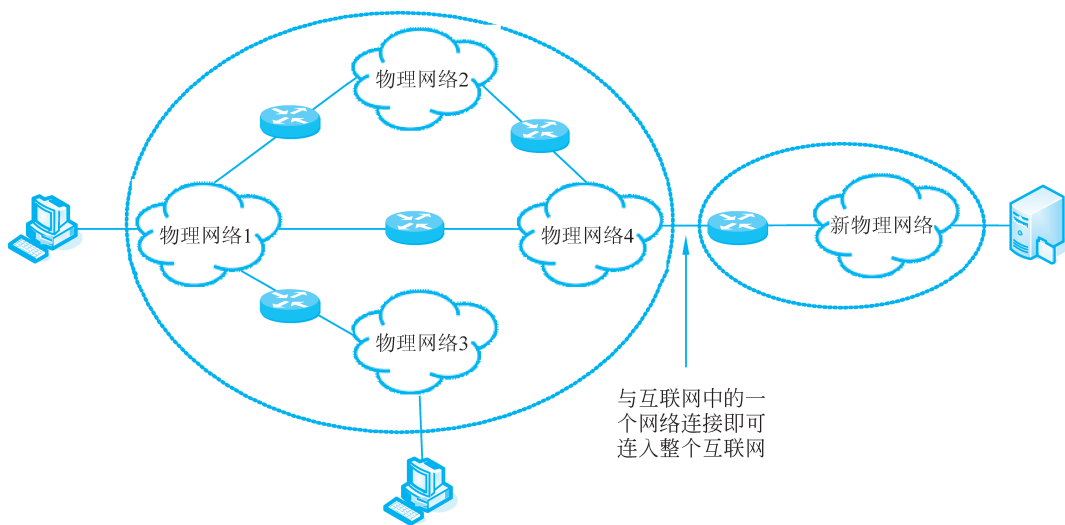


图 5-7 IP 互联网不要求网络之间全互联

(3) IP 互联网能在物理网络之间转发数据,信息可以跨网传输。

(4) IP 互联网中的所有计算机使用统一的、全局的地址描述法。

(5) IP 互联网平等地对待互联网中的每个网络,不管这个网络规模是大还是小,也不管这个网络的速度是快还是慢。实际上,在 IP 互联网中,任何一个能传输数据单元的通信系统均被看作网络(无论该通信系统的特性如何)。因此,大到广域网、小到局域网甚至两台机器间的点到点链接都被当作网络,IP 互联网平等地对待它们。

练习与思考

一、填空题

(1) 网络互联的解决方案有两种:一种是_____,另一种是_____。其中,_____是目前主要使用的解决方案。

(2) IP 提供服务的特点为_____、_____和_____。

二、单项选择题

(1) Internet 使用的互联协议是()。

- a) IPX b) IP c) AppleTalk d) NetBEUI

(2) 关于 IP 层提供的功能,错误的是()。

- a) 屏蔽各个物理网络的差异 b) 代替各个物理网络的数据链路层工作
c) 为用户提供通用的服务 d) 隐藏各个物理网络的实现细节

三、简答题

简述 IP 互联网的主要作用和特点。

第 6 章 IP 数据包

在 IP 层,需要传输的数据首先需要加上 IP 头信息,封装成 IP 数据包。IP 数据包(IP packet)是 IP 层使用的数据单元,互联层数据信息和控制信息的传递都需要通过 IP 数据包进行。

6.1 IP 数据包的格式

IP 数据包的格式可以分为包头区和数据区两大部分,其中数据区包括高层需要传输的数据,而包头区是为了正确传输高层数据而增加的控制信息。图 6-1 给出了 IP 数据包的具体格式。

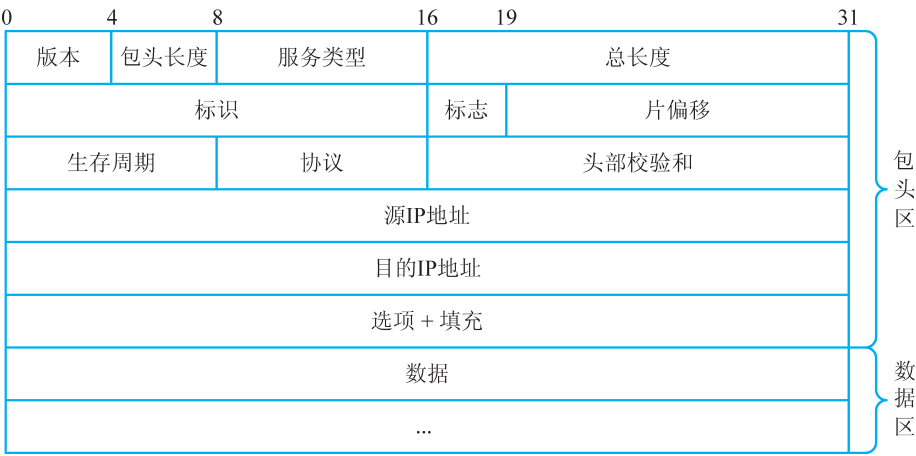


图 6-1 IP 数据包的具体格式

包头区包含了源 IP 地址、目的 IP 地址等控制信息,下面分别介绍各主要字段的功能。

1. 版本与协议

在 IP 包头中,版本字段表示该数据包对应的 IP 版本号,不同 IP 版本规定的数据包格式稍有不同,目前最常使用的 IP 版本号为 4。为了避免错误解释报文格式和内容,所有 IP 软件在处理数据包之前都必须检查版本号,以确保版本正确。

协议字段表示该数据包数据区数据使用的协议类型(如 TCP),用于指明数据区数据的格式。

2. 长度

包头中有两个表示长度的字段:一个是包头长度,另一个是总长度。

包头长度以 32 位双字为单位,指出该包头区的长度。在没有选项和填充的情况下,该值为 5。一个含有选项的包头长度取决于选项域的长度。但是,包头长度应当是 32 位的整数倍,否则需在填充域加 0 凑齐。

总长度以 8 位(1 字节)为单位,表示整个 IP 数据包的长度,其中包含头部长度和数据区长度。

3. 服务类型

服务类型字段规定对本数据包的处理方式。利用该字段,发送端可以为 IP 数据包分配一个转发优先级,并可以要求中途转发路由器尽量使用低时延、高吞吐率或高可靠性的线路投

递。但是,中途路由器能否按照 IP 数据包要求的服务类型进行处理,则依赖于路由器的实现方法和底层物理网络技术。

4. 生存周期

IP 数据包的路由选择具有独立性,因此从源主机到目的主机的传输延迟也具有随机性。如果路由表发生错误,数据包有可能进入一条循环路径,无休止地在网络中流动。利用 IP 包头中的生存周期字段,可以有效地控制这一情况发生。在网络中,生存周期域随时间而递减,该域为 0 时,报文将被删除,避免死循环发生。

5. 头部校验和

头部校验和用于保证 IP 数据包包头的完整性。注意,在 IP 数据包中只含有包头校验字段,而没有数据区校验字段。这样做的最大好处是可以节约路由器处理数据包的时间,并允许不同的上层协议选择自己的数据校验方法。

6. 地址

在 IP 数据包包头中,源 IP 地址和目的 IP 地址分别表示该 IP 数据包发送者和接收者地址。IP 地址采用 32 位的地址形式,其作用和使用方法将在第 7 章详细介绍。在整个数据包传输过程中,无论经过什么路由,无论如何分片,这两个字段都一直保持不变。

6.2 IP 封装、分片与重组

因为 IP 数据包可以在互联网上传输,所以它可能要跨越多个网络。作为一种高层网络数据,IP 数据包最终也需要封装成帧进行传输。图 6-2 显示了一个 IP 数据包从源主机至目的主机被多次封装和解封装的过程。

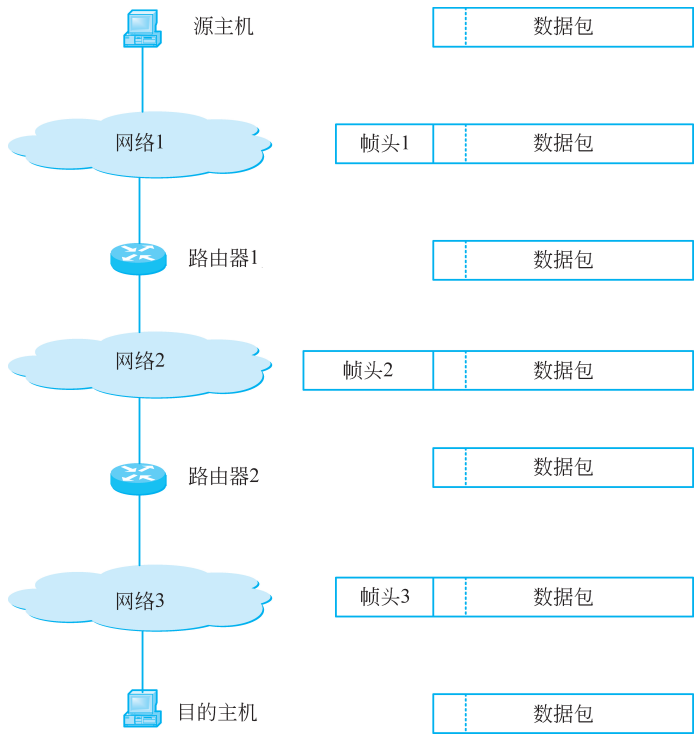


图 6-2 IP 数据包在各个网络中被重新封装

从图 6-2 中可以看出,主机和路由器只在内存中保留了整个 IP 数据包,而没有附加的帧头信息。只有当 IP 数据包通过一个物理网络时,才会被封装进一个合适的帧中。帧头的大小依赖于相应的网络技术。例如,如果网络 1 是一个以太网,则帧 1 有一个以太网头部;如果网络 2 是一个 FDDI 环网,则帧 2 有一个 FDDI 头部。注意,在数据包通过互联网的整个过程中,帧头并没有累积起来。当数据包到达它的最终目的地时,数据包的大小与其最初发送时是一样的。

6.2.1 MTU 与分片

根据网络使用的技术不同,每种网络都规定了一个帧最多能够携带的数据量,这一限制称为最大传输单元(Maximum Transmission Unit,MTU)。因此,一个 IP 数据包的长度只有小于或等于一个网络的 MTU,才能在这个网络中传输。

互联网可以包含各种各样的异构网络,一个路由器也可能连接具有不同 MTU 值的多个网络,能从一个网络上接收 IP 数据包并不意味着一定能在另一个网络上发送该数据包。在图 6-3 中,一个路由器连接了两个网络,其中一个网络的 MTU 为 1500B,另一个网络的 MTU 为 1000B。



图 6-3 路由器连接具有不同 MTU 的网络

主机 1 连接着 MTU 值为 1500B 的网络 1,因此,每次传送的 IP 数据包字节数不超过 1500B。而主机 2 连接着 MTU 值为 1000B 的网络 2,因此,主机 2 可以传送的 IP 数据包最大尺寸为 1000B。如果主机 1 需要将一个 1400B 的数据包发送给主机 2,路由器 R 尽管能够收到主机 1 发送的数据包,却不能在网络 2 上转发它。

为了解决这一问题,IP 互联网通常采用分片与重组技术。当一个数据包的尺寸大于将发往网络的 MTU 值时,路由器会将 IP 数据包分成若干较小的部分,这个过程称为分片,然后再将每片独立地进行发送。

与未分片的 IP 数据包相同,分片后的数据包也由包头区和数据区两部分构成,而且除一些分片控制域(如标志域、片偏移域)之外,分片的包头与原 IP 数据包的包头非常相似,如图 6-4 所示。

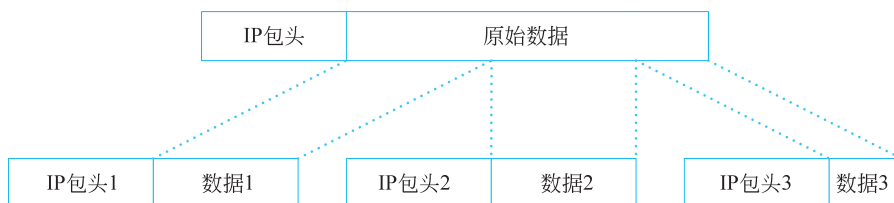


图 6-4 IP 数据包分片

一旦进行分片,每片都可以像正常的 IP 数据包一样经过独立的路由选择等处理过程,最终到达目的主机。

6.2.2 重组

在接收到所有分片的基础上,主机对分片进行重新组装的过程称为 IP 数据包重组。IP