

第3章

数据链路

本章主要介绍计算机网络最基本的内容——数据链路层。如果没有数据链路层，基于TCP/IP的通信也就无从谈起。因此，本章将着重介绍TCP/IP的具体数据链路，如以太网、无线局域网、PPP等。

7 应用层	<div><应用层> TELNET, SSH, HTTP, SMTP, POP, SSL/TLS, FTP, MIME, HTML, SNMP, MIB, SIP, RTP ...</div> <div><传输层> TCP, UDP, UDP-Lite, SCTP, DCCP</div> <div><网络层> ARP, IPv4, IPv6, ICMP, IPsec</div> <div>以太网、无线LAN、PPP…… (双绞线电缆、无线、光纤……)</div>
6 表示层	
5 会话层	
4 传输层	
3 网络层	
2 数据链路层	
1 物理层	

3.1

数据链路的作用

数据链路，指 OSI 参考模型中的数据链路层，有时也指以太网、无线局域网等通信手段。

TCP/IP 中对于 OSI 参考模型的数据链路层及以下部分（物理层）未作定义。因为 TCP/IP 以这两层的功能是透明的为前提。然而，数据链路的知识对于深入理解 TCP/IP 与网络起着至关重要的作用。

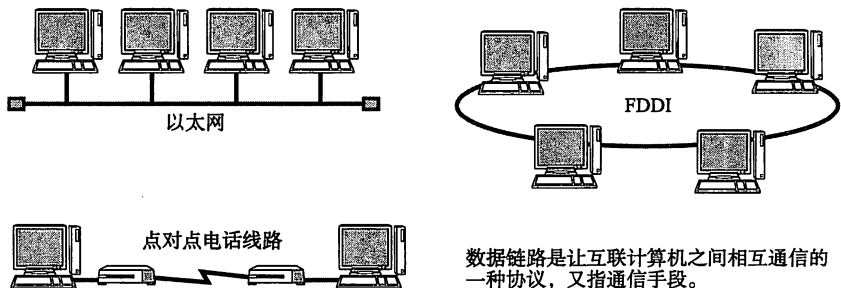
数据链路层的协议定义了通过通信媒介互连的设备之间传输的规范。通信媒介包括双绞线电缆、同轴电缆、光纤、电波以及红外线等介质。此外，各个设备之间有时也会通过交换机、网桥、中继器等中转数据。

实际上，各个设备之间在数据传输时，数据链路层和物理层都是必不可少的。众所周知，计算机以二进制 0、1 来表示信息，然而实际的通信媒介之间处理的却是电压的高低、光的闪灭以及电波的强弱等信号。把这些信号与二进制的 0、1 进行转换正是物理层（参考附录 3）的责任。数据链路层处理的数据也不是单纯的 0、1 序列，该层把它们集合为一个叫做“帧”的块，然后再进行传输。

本章旨在介绍 OSI 参考模型中数据链路层的相关技术，包括 MAC 寻址（物理寻址）、介质共享、非公有网络、分组交换、环路检测、VLAN（Virtual Local Area Network，虚拟局域网）等。本章也会涉及作为传输方式的数据链路，如以太网、WLAN（Wireless Local Area Network，无线局域网）、PPP（Point to Point Protocol，点对点协议）等概念。数据链路也可以被视为网络传输中的最小单位。其实，仔细观察连通全世界的互联网就可以发现，它也不外乎是由众多这样的数据链路组成的，因此又可以称互联网为“数据链路的集合”。

在以太网与 FDDI（Fiber Distributed Data Interface，光纤分布式数据接口）的规范中，不仅包含 OSI 参考模型的第 2 层数据链路层，也规定了第 1 层物理层的规格。而在 ATM（Asynchronous Transfer Mode，异步传输方式）的规范中，还包含了第 3 层网络层的一部分功能。

图 3.1
数据链路



■ 数据链路的段

数据链路的段是指一个被分割的网络。然而根据使用者不同，其含义也不尽相同。例如，引入中继器将两条网线相连组成一个网络。

这种情况下有两条数据链路：

- 从网络层的概念看，它是一个网络（逻辑上）→ 即，从网络层的立场出发，这两条网线组成一个段。
- 从物理层的概念看，两条网线分别是两个物体（物理上）→ 即，从物理层的观点出发，一条网线是一个段。

从网络层上看，是一个段

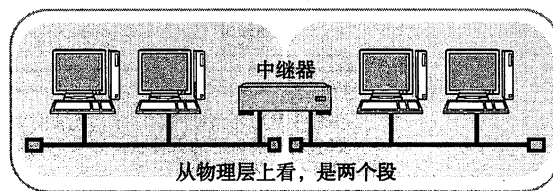


图 3.2

段的范围

■ 网络拓扑

网络的连接和构成的形态称为网络拓扑（Topology）。网络拓扑包括总线型、环型、星型、网状型等。拓扑一词不仅用于直观可见的配线方式上，也用于逻辑上网络的组成结构。两者有时可能会不一致。图 3.3 展示了配线上的拓扑结构。而目前实际的网络都是由这些简单的拓扑结构错综复杂地组合而成的。

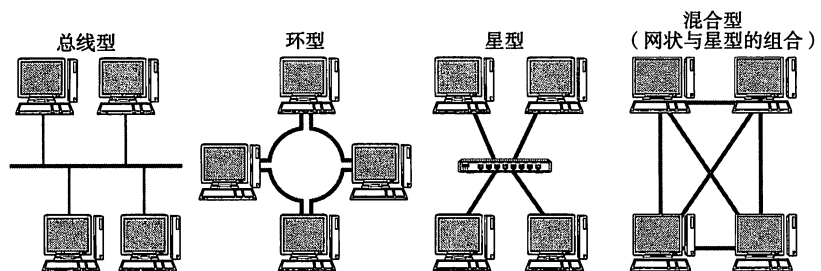


图 3.3

总线型、环型、星型、网状型

3.2

数据链路相关技术

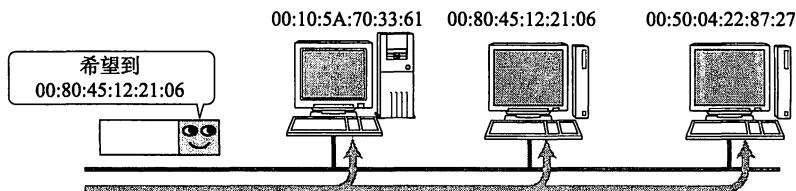
3.2.1 MAC 地址

▼ IEEE 指的是美国电气和电子工程师协会，也叫“1 triple E”。IEEE802 是制定局域网标准化相关规范的组织。其中 IEEE802.3 是关于以太网 (CSMA/CD) 的国际规范。

图 3.4

通过 MAC 地址判断目标地址

MAC 地址用于识别数据链路中互连的节点 (如图 3.4)。以太网或 FDDI 中, 根据 IEEE802.3 的规范使用 MAC 地址。其他诸如无线 LAN (IEEE802.11a/b/g/n 等)、蓝牙等设备中也是用相同规格的 MAC 地址。



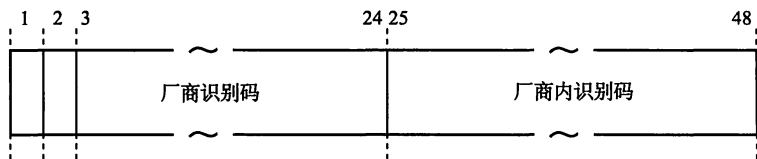
在总线型与环路型的网络中, 先暂时获取所有目标站的帧, 然后再通过 MAC 寻址。如果是发给自己的就接收, 如果不是就丢弃 (在令牌环的情况下, 依次转发给下一个站)。

MAC 地址长 48 比特, 结构如图 3.5 所示。在使用网卡 (NIC) 的情况下, MAC 地址一般会被烧入到 ROM 中。因此, 任何一个网卡的 MAC 地址都是唯一的, 在全世界都不会有重复▼。

▼也有例外, 具体请参考网页注解。

图 3.5

IEEE802.3 规范的 MAC 地址格式



第1位: 单播地址 (0) / 多播地址 (1)

第2位: 全局地址 (0) / 本地地址 (1)

第3~24位: 由IEEE管理并保证各厂家之间不重复

第25~48位: 由厂商管理并保证产品之间不重复

* 该图表示比特流在网络中的流动顺序。

MAC地址一般用十六进制数表示。注意, 如果以十六进制表示, 此图中已按照每8比特转换了对应的值, 并替换了前后顺序。

例如,

用十六进制多播MAC地址 (上图中第1比特为1) 表示……

01:00:XX:XX:XX:XX

用二进制比特表示

0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 …

1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 …

← 流过网络的比特流顺序

MAC 地址中 3~24 位（比特位）表示厂商识别码，每个 NIC 厂商都有特定唯一的识别数字。25~48 位是厂商内部为识别每个网卡而用。因此，可以保证全世界不会有相同 MAC 地址的网卡。

IEEE802.3 制定 MAC 地址规范时没有限定数据链路的类型，即不论哪种数据链路的网络（以太网、FDDI、ATM、无线 LAN、蓝牙等），都不会有相同的 MAC 地址出现。

■ 例外情况——MAC 地址不一定是唯一的

在全世界，MAC 地址也并不总是唯一的。实际上，即使 MAC 地址相同，只要不是同属一个数据链路就不会出现问题。

例如，人们可以在微机板上自由设置自己的 MAC 地址。再例如，一台主机上如果启动多个虚拟机，由于没有硬件的网卡只能由虚拟软件自己设定 MAC 地址给多个虚拟网卡，这时就很难保证所生成的 MAC 地址是独一无二的了。

但是，无论哪个协议成员通信设备，设计前提都是 MAC 地址的唯一性。这也可以说是网络世界的基本准则。

■ 厂商识别码

有一种设备叫网络分析器。它可以分析出局域网中的包是由哪个厂商的网卡发出的。它通过读取数据帧当中发送 MAC 地址里的厂商识别码进行识别。由于能够迅速定位是否有未知厂商识别码的网卡发送异常的包，这一功能在由多个厂商的设备构成的网络环境中，对于分析问题极为有效。

厂商识别码官方的叫法是 OUI（Organizationally Unique Identifier）。OUI 信息一般都会公开在以下网站上▼：

<http://standards.ieee.org/develop/regauth/oui/public.html>

此外，MAC 地址的分配，通过以下站点申请（收费）：

<http://standards.ieee.org/develop/regauth/oui/index.html>

▼由于最近网络设备厂商的收购与合并，OUI 的数据库和实际厂商名字也出现了不一致的情况。

3.2.2 共享介质型网络

从通信介质（通信，介质）的使用方法上看，网络可分为共享介质型和非共享介质型。

共享介质型网络指由多个设备共享一个通信介质的一种网络。最早的以太网和 FDDI 就是介质共享型网络。在这种方式下，设备之间使用同一个载波信道进行发送和接收。为此，基本上采用半双工通信（参考 3.2.3 节后面的详解）方式，并有必要对介质进行访问控制。

共享介质型网络中有两种介质访问控制方式：一种是争用方式，另一种是令牌传递方式。

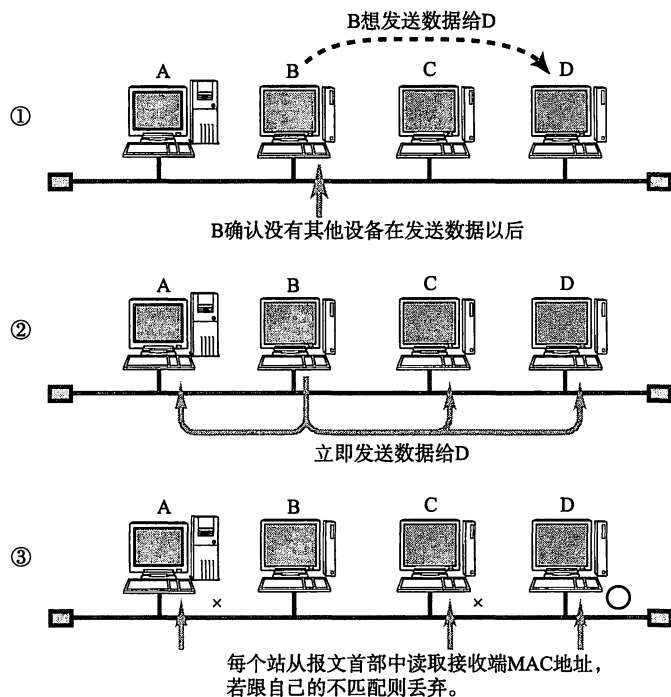
■ 争用方式

争用方式（Contention）是指争夺获取数据传输的权力，也叫 CSMA（载波监听多路访问）。这种方法通常令网络中的各个站采用先到先得的方式占用信道发送数据，如果多个站同时发送帧，则会产生冲突现象。也因此会导致网络拥堵与性能下降。

▼数据链路中很多情况下称节点为“站”。

图 3.6

争用方式



▼ Carrier Sense Multiple Access with Collision Detection

在一部分以太网当中，采用了改良 CSMA 的另一种方式——CSMA/CD[▼] 方式。CSMA/CD 要求每个站提前检查冲突，一旦发生冲突，则尽早释放信道。其具体工作原理如下：

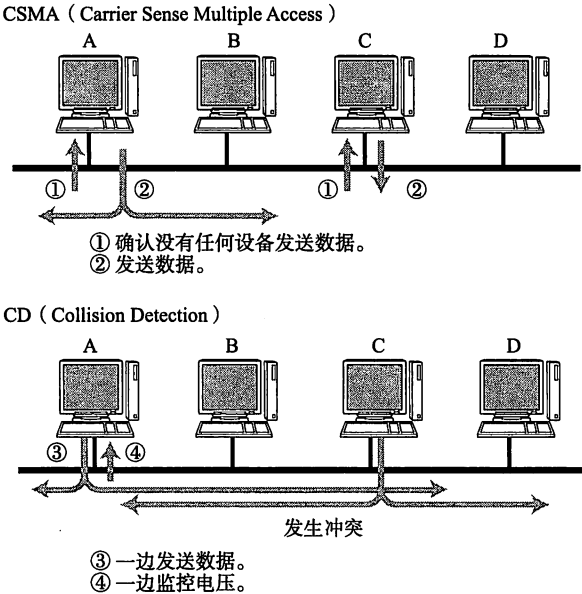
- 如果载波信道上没有数据流动，则任何站都可以发送数据。
- 检查是否会发生冲突。一旦发生冲突时，放弃发送数据[▼]，同时立即释放载波信道。
- 放弃发送以后，随机延时一段时间，再重新争用介质，重新发送帧。

CSMA/CD 具体工作原理请参考图 3.7。

▼实际上会发送一个 32 位特别的信号，在阻塞报文以后再停止发送。接收端通过发生冲突时帧的 FCS（参考 3.3.4 节），判断出该帧不正确从而丢弃帧。

图 3.7

CSMA/CD 方式



- 直到发送完数据，如果电压一直处于规定范围内，就认为数据已正常发送。
- 发送途中，如果电压一旦超出规定范围，就认为是发生了冲突。
- 发生冲突时先发送一个阻塞报文后，放弃发送数据帧，在随机延时一段时间后进行重发。

* 这种通过电压检查冲突的硬件属于同轴电缆。

■ 令牌传递方式

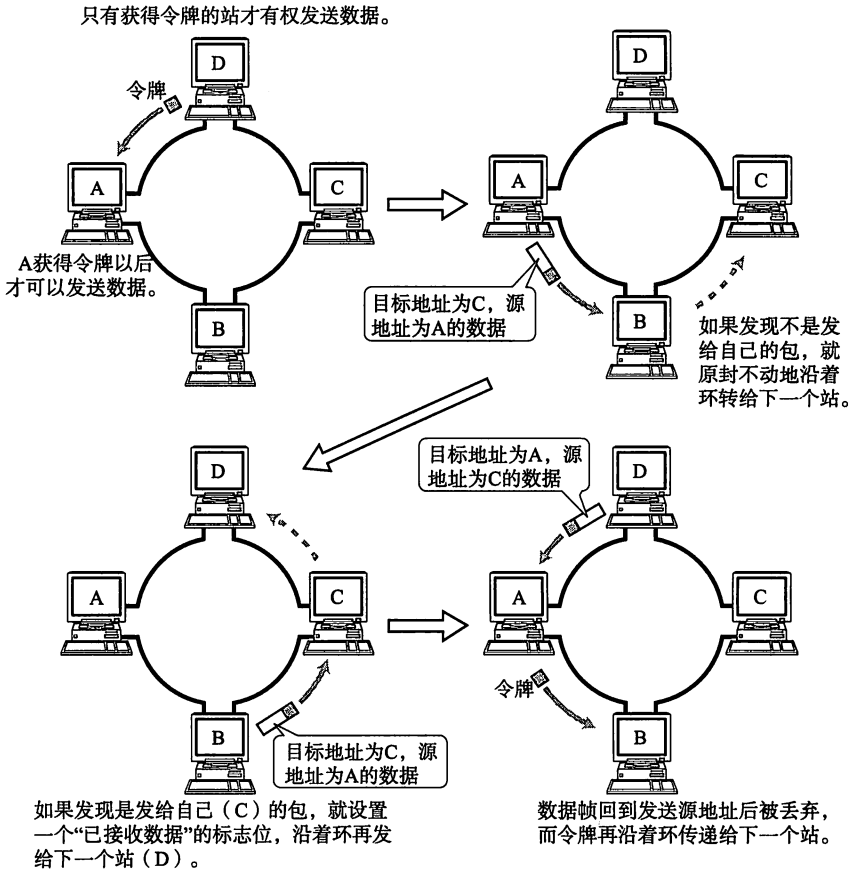
令牌传递方式是沿着令牌环发送一种叫做“令牌”的特殊报文，是控制传输的一种方式。只有获得令牌的站才能发送数据。这种方式有两个特点：一是不会有冲突，二是每个站都有通过平等循环获得令牌的机会。因此，即使网络拥堵也不会导致性能下降。

当然，这种方式中，一个站在没有收到令牌前不能发送数据帧，因此在网络不太拥堵的情况下数据链路的利用率也就达不到 100%。为此，衍生了多种令牌传递的技术。例如，早期令牌释放、令牌追加等方式以及多个令牌同时循环等方式。这些方式的目的是为了尽可能地提高网络性能。

▼ 不等待接收方的数据到达确认就将令牌发送给下一个站。

图 3.8

令牌传递方式



3.2.3 非共享介质网络

非共享介质网络是指不共享介质，是对介质采取专用的一种传输控制方式。在这种方式下，网络中的每个站直连交换机，由交换机负责转发数据帧。此方式下，发送端与接收端并不共享通信介质，因此很多情况下采用全双工通信方式（具体请参考本节最后的详解）。

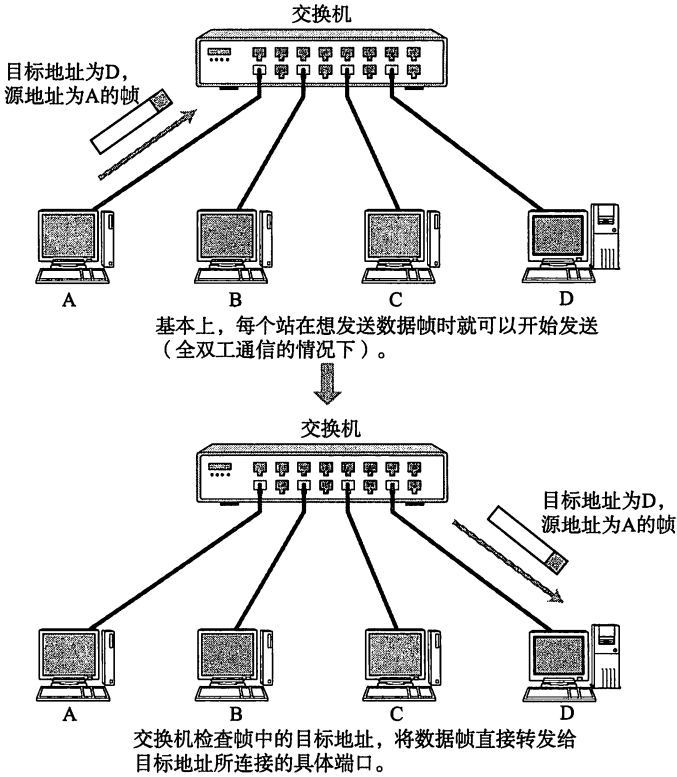
不仅 ATM 采用这种传输控制方式，最近它也成了以太网的主流方式。通过以太网交换机构建网络，从而使计算机与交换机端口之间形成一对一的连接，即可实现全双工通信。在这种一对一连接全双工通信的方式下不会发生冲突，因此不需要 CSMA/CD 的机制就可以实现更高效的通信。

该方式还可以根据交换机的高级特性构建虚拟局域网（VLAN，Virtual LAN）[▼]、进行流量控制等。当然，这种方式也有一个致命的弱点，那就是一旦交换机发生故障，与之相连的所有计算机之间都将无法通信。

▼关于 VLAN 的更多细节请参考 3.2.6 节。

图 3.9

非共享介质型网络



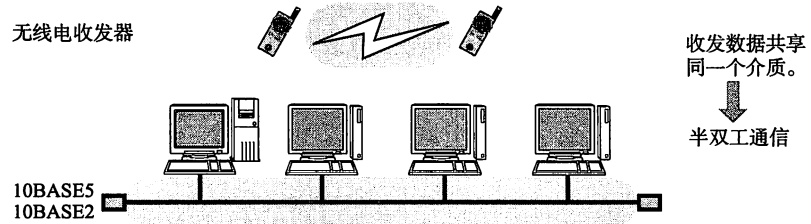
■ 半双工与全双工通信

半双工是指, 只发送或只接收的通信方式。它类似于无线电收发器, 若两端同时说话, 是听不见对方说的话的。而全双工不同, 它允许在同一时间既可以发送数据也可以接收数据。类似于电话, 接打双方可以同时说话。

采用 CSMA/CD 方式的以太网, 如图 3.7 所示, 首先要判断是否可以通信, 如果可以就独占通信介质发送数据。因此, 它像无线电收发器一样, 不能同时接收和发送数据。

图 3.10

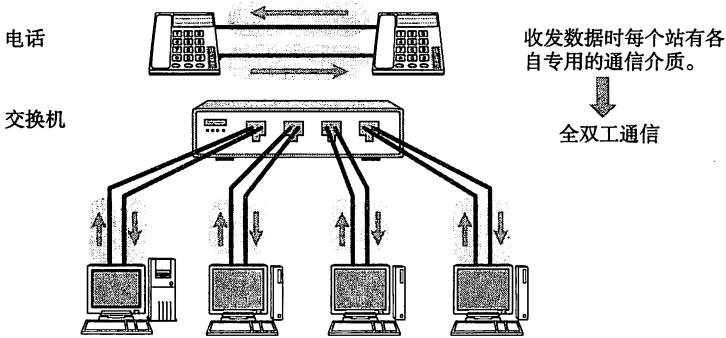
半双工通信



▼一般一根双绞线包着8个(4对)芯线。

图 3-11
全双工通信

同样是以太网，在使用交换机与双绞线电缆（亦或光纤电缆）的情况下，既可以通过交换机的端口与计算机之间进行一对一的连接，也可以通过相连电缆内部的收发线路分别进行接收和发送数据。因此，交换机的端口与计算机之间可以实现同时收发的全双工通信。



3.2.4 根据 MAC 地址转发

在使用同轴电缆的以太网（10BASE5、10BASE2）等介质共享网络中，同一时间只能有一台主机发送数据。当连网的主机数量增加时，通信性能会明显下降。若将集线器或集中器等设备以星型连接，就出现了一款新的网络设备——交换集线器，这是一种将非介质共享型网络中所使用的交换机用在以太网中的技术，交换集线器也叫做以太网交换机。

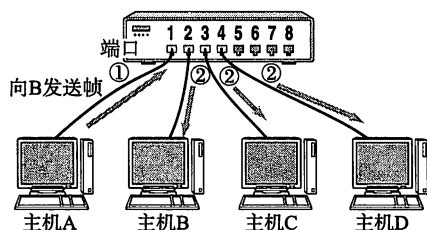
以太网交换机就是持有多个端口▼的网桥。它们根据数据链路层中每个帧的目标 MAC 地址，决定从哪个网络接口发送数据。这时所参考的、用以记录发送接口的表就叫做转发表（Forwarding Table）。

这种转发表的内容不需要使用者在每个终端或交换机上手工设置，而是可以自动生成。数据链路层的每个通过点在接到包时，会从中将源 MAC 地址以及曾经接收该地址发送的数据包的接口作为对应关系记录到转发表中。以某个 MAC 地址作为源地址的包由某一接口接收，实质上可以理解为该 MAC 地址就是该接口的目标。因此也可以说，以该 MAC 地址作为目标地址的包，经由该接口送出即可。这一过程也叫自学过程。

▼计算机设备的外部接口都称做端口。必须注意 TCP 或 UDP 等传输层协议中的“端口”另有其他含义。

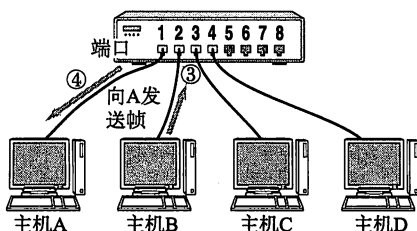
图 3.12

交换机的自学原理



① 从源MAC地址可以获知主机A与端口1相连接。

② 拷贝那些以“未知”MAC地址为目标的帧给所有的端口。



③ 从源MAC地址可以获知主机B与端口2相连接。

④ 由于已经知道主机A与端口1相连接，那么发给主机A的帧只拷贝给端口1。

以后，主机A与主机B的通信就只在它们各自所连接的端口之间进行。

▼关于地址的层次性，请参考1.8.2节。

由于MAC地址没有层次性[▼]，转发表中的入口个数与整个数据链路中所有网络设备的数量有关。当设备数量增加时，转发表也会随之变大，检索转发表所用的时间也就越来越长。当连接多个终端时，有必要将网络分成多个数据链路，采用类似于网络层的IP地址一样对地址进行分层管理。

▼关于FCS的更多细节请参考3.3.4节。

■ 交换机转发方式

交换机转发方式有两种，一种叫存储转发，另一种叫直通转发。

存储转发方式检查以太网数据帧末尾的FCS[▼]位后再进行转发。因此，可以避免发送由于冲突而被破坏的帧或噪声导致的错误帧。

直通转发方式中不需要将整个帧全部接收下来以后再进行转发。只需要得知目标地址即可开始转发。因此，它具有延迟较短的优势。但同时也不可避免地有发送错误帧的可能性。

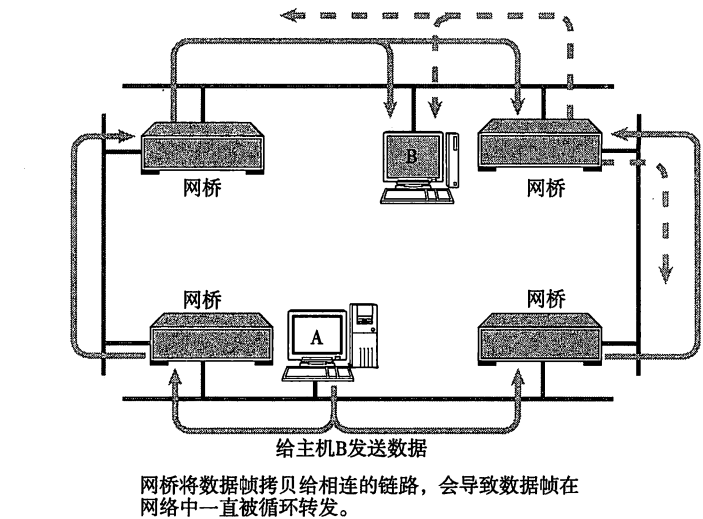
3.2.5 环路检测技术

通过网桥连接网络时，一旦出现环路该如何处理？这与网络的拓扑结构和所使用的网桥种类有直接关系。最坏的情况下，数据帧会在环路中被一而再再而三地持续转发。而一旦这种数据帧越积越多将会导致网络瘫痪。[▼]

为此，有必要解决网络中的环路问题。具体有生成树与源路由两种方式。如果使用具有这些功能的网桥，那么即使构建了一个带有环路的网络，也不会造成那么严重的问题。只要搭建合适的环路，就能分散网络流量，在发生某一处路由故障时选择绕行，可以提高容灾能力。

▼是指由于异常的数据帧遍布网络，造成无法正常通信的状态。很多情况下只有关掉网络设备的电源或断开网络才能恢复。

图 3.13
网桥搭建带有环路的网络



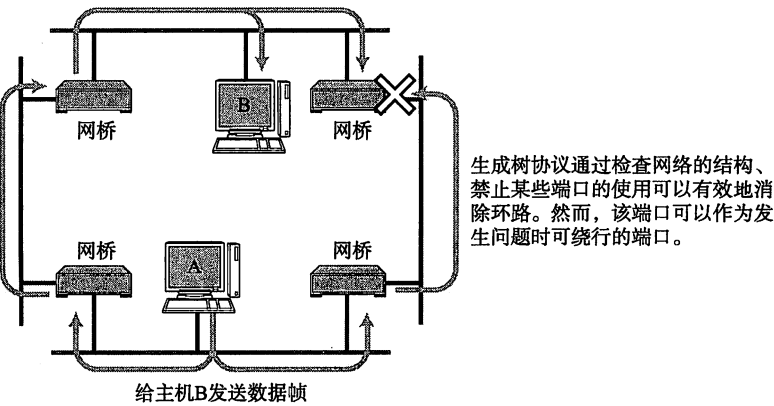
■ 生成树方式

该方法由 IEEE802.1D 定义。每个网桥必须在每 1~10 秒内相互交换 BPDU (Bridge Protocol Data Unit) 包，从而判断哪些端口使用哪些不使用，以便消除环路。一旦发生故障，则自动切换通信线路，利用那些没有被使用的端口继续进行传输。

例如，以某一个网桥为构造树的根 (Root)，并对每个端口设置权重。这一权重可以由网络管理员适当地设置，指定优先使用哪些端口以及发生问题时该使用哪些端口。

生成树法其实与计算机和路由器的功能没有关系，但是只要有生成树的功能就足以消除环路。

图 3.14
生成树法



IEEE802.1D 中所定义的生成树方法有一个弊端，就是在发生故障切换网络时需要几十秒的时间。为了解决这个用时过长的问题，在 IEEE802.1W 中定义了一个叫 RSTP (Rapid Spanning Tree Protocol) 的方法。该方法能将发生问题时的恢复时间缩短到几秒以内。

源路由法

▼关于 Token Ring 的更多细节，请参考 3.6.4 节。

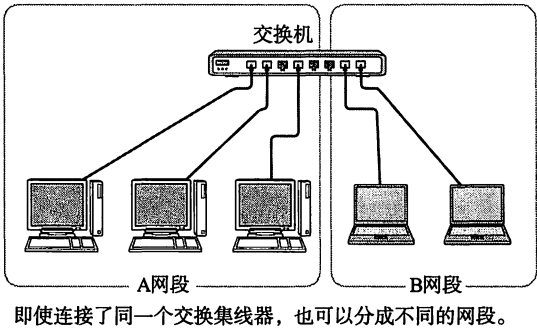
源路由法最早由 IBM 提出，以解决令牌环[▼]网络的问题。该方式可以判断发送数据的源地址是通过哪个网桥实现传输的，并将帧写入 RIF（Routing Information Field）。网桥则根据这个 RIF 信息发送帧给目标地址。因此，即使网桥中出现了环路，数据帧也不会被反复转发，可成功地发送到目标地址。在这种机制中发送端本身必须具备源路由的功能。

3.2.6 VLAN

进行网络管理的时候，时常会遇到分散网络负载、变换部署网络设备的位置等情况。而有时管理员在做这些操作时，不得不修改网络的拓扑结构，这也就意味着必须进行硬件线路的改造。然而，如果采用带有 VLAN 技术的网桥，就不用实际修改网络布线，只需修改网络的结构即可。VLAN 技术附加到网桥/2 层交换机（曾在 1.9.4 节做过介绍）上，就可以切断所有 VLAN 之间的所有通信。因此，相比一般的网桥/2 层交换机，VLAN 可以过滤多余的包，提高网络的承载效率。

那么 VLAN 究竟是什么？如图 3.15 所示，该交换机按照其端口区分了多个网段，从而区分了广播数据传播的范围、减少了网络负载并提高了网络的安全性。然而异构的两个网段之间，就需要利用具有路由功能的交换机（如 3 层交换机），或在各段中间通过路由器的连接才能实现通信。

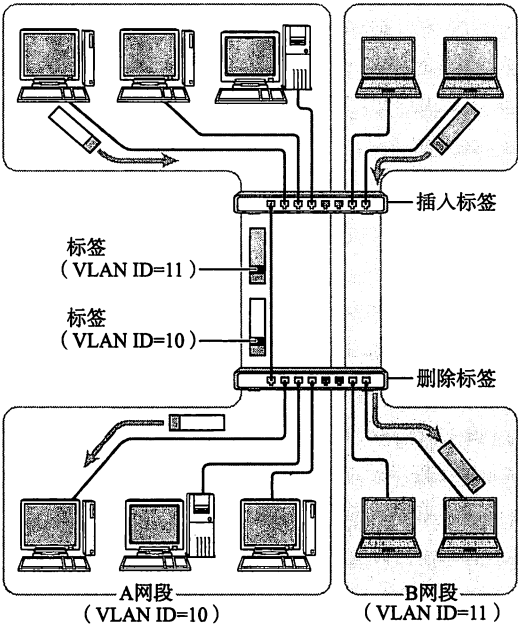
图 3.15
简单的 VLAN



对这种 VLAN 进行了扩展，又定义了 IEEE802.1Q 的标准（也叫 TAG VLAN），该标准允许包含跨越异构交换机的网段。TAG VLAN 中对每个网段都用一个 VLAN ID 的标签进行唯一标识。在交换机中传输帧时，在以太网首部加入这个 VID 标签，根据这个值决定将数据帧发送给哪个网段。各个交换机之间流动的数据帧的格式请参考图 3.21 中的帧格式。

随着 VLAN 技术的应用，不必再重新修改布线，只要修改网段即可。当然，有时物理网络结构与逻辑网络结构也可能会出现不一致的情况，导致不易管理。为此，应该加强对网段构成及网络运行等的管理。

图 3-16
跨交换机的 VLAN



3.3

以太网

▼以太网 (Ethernet) 一词源于 Ether (以太), 意为介质。在爱因斯坦提出量子力学之前, 人们普遍认为宇宙空间充满以太, 并以波的形式传递着光。

▼反之, 一般的以太网则有时被叫做 DIX 以太网。DIX 由 DEC、Intel 和 Xerox 等公司名称的首字母组成。

▼关于共享介质型的更多细节请参考 3.2.2 节。

在众多数据链路中最为著名、使用最为广泛的莫过于以太网 (Ethernet)▼。它的规范简单, 易于 NIC (网卡) 及驱动程序实现。因此, 在 LAN 普及初期, 以太网网卡相对于其他网卡, 价格也比较低廉。这也同时促进了以太网自身的普及。从最初的 10Mbps、1Gbps、10Gbps 到后来的 40Gbps/100Gbps 以太网已能够支持高速网络。现在, 以太网已成为最具兼容性与未来发展性的一种数据链路。

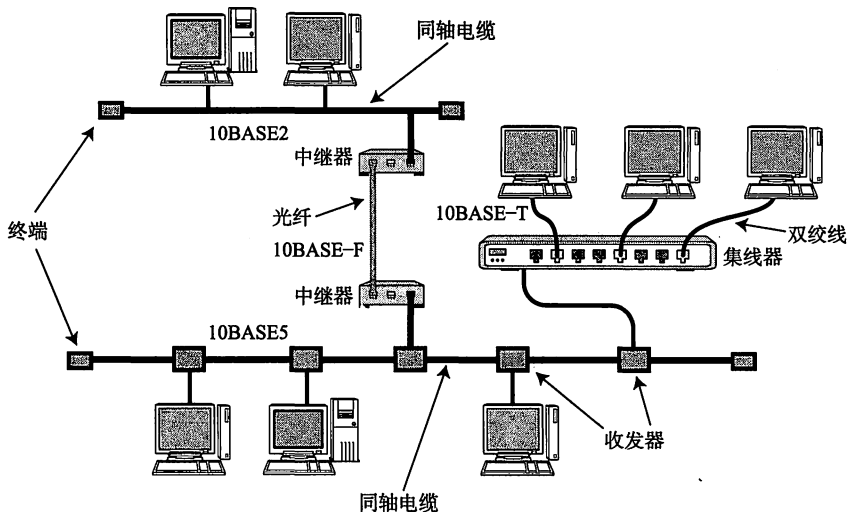
以太网最早是由美国的 Xerox 公司与前 DEC 公司设计的一种通信方式, 当时命名为 Ethernet。之后由 IEEE802.3 委员会将其规范化。但是这两者之间对以太网网帧的格式定义还是有所不同的。因此, IEEE802.3 所规范的以太网有时又被称为 802.3 以太网▼。

3.3.1 以太网连接形式

在以太网普及之初, 一般采用多台终端使用同一根同轴电缆的共享介质型▼连接方式。

图 3-17

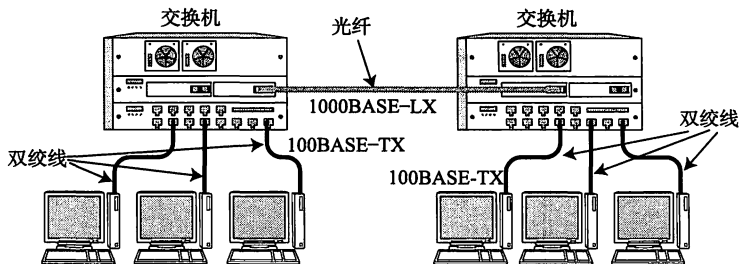
初期以太网结构举例



而现在, 随着互连设备的处理能力以及传输速度的提高, 一般都采用终端与交换机之间独占电缆的方式实现以太网通信, 如图 3.18。

图 3-18

现代以太网结构举例



3.3.2 以太网的分类

以太网因通信电缆的不同及通信速度的差异，衍生出了众多不同的以太网类型。

10BASE 中的“10”、100BASE 中的“100”、1000BASE 中的“1000”以及 10GBASE 中的“10G”分别指 10Mbps、100Mbps、1Gbps 以及 10Gbps 的传输速度。而追加于后面的“5”、“2”、“T”、“F”等字符表示的是传输介质。在传输速度相同而传输所用电缆不同的情况下，可以连接那些允许更换传输介质的中继器或集线器。而在传输速度不同的情况下，则必须采用那些允许变更速度的设备如网桥、交换集线器或路由器。

表 3-1

以太网主要分类及其特点

▼ Unshielded Twisted Pair Cable，非屏蔽双绞线。

▼ Category 的简写。TIA/EIA (Telecommunication Industries Association / Electronic Industries Alliance，美国电信工业协会/美国电子工业协议) 制定的双绞线规格。CAT 值越大，表明传输速度越高。

▼ Multi Mode Fiber，多模光纤。

▼ Shielded Twisted Pair Cable，屏蔽双绞线。

▼ Single Mode Fiber，单模光纤。

▼ Foil Twisted -Pair，铝箔总屏蔽双绞线。

以太网种类	电缆最大长度	电缆种类
10BASE2	185m (最大节点数为 30)	同轴电缆
10BASE5	500m (最大节点数为 100)	同轴电缆
10BASE-T	100m	双绞线 (UTP▼-CAT▼ 3-5)
10BASE-F	1000m	多模光纤 (MMF▼)
100BASE-TX	100m	双绞线 (UTP-CAT5/STP▼)
100BASE-FX	412m	多模光纤 (MMF)
100BASE-T4	100m	双绞线 (UTP-CAT3-5)
1000BASE-CX	25m	屏蔽铜线
1000BASE-SX	220m/550m	多模光纤 (MMF)
1000BASE-LX	550m/5000m	多模/单模光纤 (MMF/SMF▼)
1000BASE-T	100m	双绞线 (UTP-CAT5/5e)
10GBASE-SR	26m~300m	多模光纤 (MMF)
10GBASE-LR	1000m~2500m	单模光纤 (SMF)
10GBASE-ER	3000m/4000m	单模光纤 (SMF)
10GBASE-T	100m	双绞线 (UTP/FTP▼ CAT6a)

■ 传输速度与计算机内部的表现值

计算机内部采用二进制，因此以 2^{10} 表示最接近于 1000 的值。于是有如下等式。

- $1K=1024$
- $1M=1024K$
- $1G=1024M$

而以太网中以时钟频率决定传输速度。以下等式请不要与上面混淆。

- $1K=1000$
- $1M=1000K$
- $1G=1000M$

3.3.3 以太网的历史

最早被规范化的以太网采用同轴电缆的总线型 10BASE5 网络。之后，出现了使用细同轴电缆的 10BASE2（thin 以太网）、双绞线 10BASE-T（双绞线以太网）、高速 100BASE-TX（高速以太网）、100BASE-T（千兆以太网）以及 10G 以太网等众多以太网规范。

起初以太网的访问控制一般以半双工通信为前提采用 CSMA/CD 方式。CSMA/CD 前身与以太网同步使用，主要用来解决冲突检查的问题。然而，这时的 CSMA/CD 同时也成为了以太网高速化的一个主要瓶颈。即使出现了 100Mbps 的 FDDI，以太网仍然滞留在 10Mbps 的速度上，以至于人们一度认为要想获取更高速的网络，只能放弃以太网另寻他路。

而这种状况并没有持续太久，随着 ATM 交换技术的进步和 CAT5 UTP 电缆的普及很快就被打破。以太网的结构也发生了变化，逐渐采用像共享介质网络那样直接与交换机连接的方式。于是，冲突检查不再是必要内容，网络变得更加高速。此外，不使用交换机的半双工通信以及使用同轴电缆的总线型通信已经逐渐退出历史舞台。另外，在此需要指出的是没有交换机的半双工通信方式以及使用同轴电缆的总线型连接方式已渐渐退出舞台，使用范围在逐渐减少。

从此，由于不会产生冲突，早先人们所认为的那些在网络拥堵的情况下性能下降得都不如 FDDI 的观点也逐渐淡化。而且在同等性能的情况下，以太网简单的结构与低廉的成本是 FDDI 所不能比及的。难怪有人认为，随着以太网的迅速发展（从 100Mbps，1Gbps 到 10Gbps），可以说已经“没必要再研究其他有线局域网技术”了。

前面提及了多种以太网类型。不论哪种类型的以太网，它们都有一个共性：由 IEEE802.3 的分会（Ethernet Working Group）进行标准化。

▼ CSMA 或 CSMA/CD 相关的更多细节请参考 3.2.2 节的争用方式。

▼ ATM 中将固定长度的信元通过交换机快速传送。具体请参考 3.6.1 节。

▼ 100BASE-TX 在满足快速通信的同时，采用价格低廉的 CAT5 非屏蔽双绞线（UTP）。

IEEE802

IEEE（The Institute of Electrical and Electronics Engineers，美国电子和电气工程师协会）委员会中，依据不同的工作小组制定了各种局域网技术标准。以下是 IEEE802 委员会的构成。因于 1980 年 2 月启动局域网国际化项目，所以命名为 802。

- IEEE802.1 Higher Layer LAN Protocols Working Group
- IEEE802.2 Logical Link Control Working Group
- IEEE802.3 Ethernet Working Group（CSMA/CD）
10BASE5/10BASE2/10BASE-T/10Broad36
100BASE-TX/1000BASE-T/10Gb/s Ethernet
- IEEE802.4 Token Bus Working Group（MAP/TOP）
- IEEE802.5 Token Ring Working Group（4Mbps / 16Mbps）
- IEEE802.6 Metropolitan Area Network Working Group（MAN）
- IEEE802.7 Broadband TAG
- IEEE802.8 Fiber Optic TAG

IEEE802. 9	Isochronous LAN Working Group
IEEE802. 10	Security Working Group
IEEE802. 11	Wireless LAN Working Group
IEEE802. 12	Demand Priority Working Group (100VG-AnyLAN)
IEEE802. 14	Cable Modem Working Group
IEEE802. 15	Wireless Personal Area Network (WPAN) Working Group
IEEE802. 16	Broadband Wireless Access Working Group
IEEE802. 17	Resilient Packet Ring Working Group
IEEE802. 18	Radio Regulatory TAG
IEEE802. 19	Coexistence TAG
IEEE802. 20	Mobile Broadband Wireless Access
IEEE802. 21	Media Independent Handoff
IEEE802. 22	Wireless Regional Area Networks

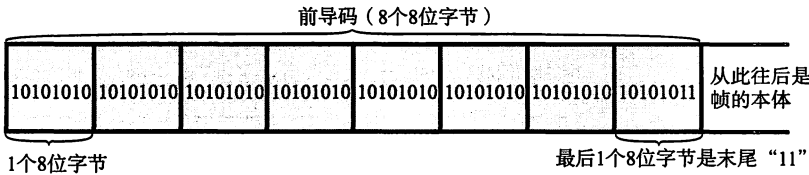
3.3.4 以太网帧格式

以太网帧前端有一个叫做前导码 (Preamble) 的部分, 它由 0、1 数字交替组合而成, 表示一个以太网帧的开始, 也是对端网卡能够确保与其同步的标志。如图 3.19 所示。前导码末尾是一个叫做 SFD (Start Frame Delimiter) 的域, 它的值是 “11”。在这个域之后就是以太网帧的帧本体 (图 3.20)。前导码与 SFD 合起来占 8 个字节▼。

▼8 位字节 (octet) 指包含 8 比特的 1 个字节。与人们平常说的字节 (Byte) 类似。关于它们的更多细节请参考后面的内容。

图 3.19

以太网帧的前导码



• 以太网中将最后2比特称为SDF, 而IEEE802.3中将最后8比特称为SDF。

以太网帧本体的前端是以太网的首部, 它总共占 14 个字节。分别是 6 个字节的目标 MAC 地址、6 个字节的源 MAC 地址以及 2 个字节的上层协议类型。

■ 比特 (位)、字节、8 位字节

- 比特 (位)
二进制中最小的单位。每个比特 (位) 的值要么是 0 要么是 1。
- 字节
通常 8 个比特构成一个字节。本书就以 8 个比特作为 1 个字节处理。然而在某些特殊的计算机中, 1 个字节有时包含 6 个比特、7 个比特或 9 个比特。
- 8 位字节
8 个比特也被称为 8 位字节。只有为了强调 1 个字节中包含 8 个比特时才会使用。

图 3-20

以太网帧格式

以太网帧格式

目标MAC地址 (6字节)	源MAC地址 (6字节)	类型 (2字节)	数据 (46~1500字节)	FCS (4字节)
------------------	-----------------	-------------	-------------------	--------------

IEEE802.3以太网帧格式

目标MAC地址 (6字节)	源MAC地址 (6字节)	帧长度 (2字节)	LLC (3字节)	SNAP (5字节)	数据 (38~1492字节)	FCS (4字节)
------------------	-----------------	--------------	--------------	---------------	-------------------	--------------

紧随帧头后面的是数据。一个数据帧所能容纳的最大数据范围是 46 ~ 1500 个字节。帧尾是一个叫做 FCS（Frame Check Sequence，帧检验序列）的 4 个字节。

在目标 MAC 地址中存放了目标工作站的物理地址。源 MAC 地址中则存放构造以太网帧的发送端工作站的物理地址。

类型通常跟数据一起传送，它包含用以标识协议类型的编号，即表明以太网的再上一层网络协议的类型。在这个字段的后面，则是该类型所标识的协议首部及其数据。关于主要的协议类型请参考表 3. 2。

表 3-2

以太网主要协议类型及其作用

类型编号（16 进制）	协 议
0000-05DC	IEEE802. 3 Length Field (01500)
0101-01FF	实验用
0800	Internet IP（IPv4）
0806	Address Resolution Protocol（ARP）
8035	Reverse Address Resolution Protocol（RARP）
8037	IPX（Novell NetWare）
805B	VMTP（Versatile Message Transaction Protocol）
809B	AppleTalk（EtherTalk）
80F3	AppleTalk Address Resolution Protocol（AARP）
8100	IEEE802. 1Q Customer VLAN
814C	SNMP over Ethernet
8191	NetBIOS/NetBEUI
817D	XTP
86DD	IP version 6（IPv6）
8847-8848	MPLS（Multi-protocol Label Switching）
8863	PPPoE Discovery Stage
8864	PPPoE Session Stage
9000	Loopback（Configuration Test Protocol）

本书中所涉及的协议类型有 IP 0800、ARP 0806、RARP 8035 以及 IPv6

86DD。关于协议类型对照表的更多细节，请参考以下两个网站。

```
http: //standards. ieee. org/regauth/ethertype/eth. txt
http: //www. iana. org/assignments/ethernet-numbers
```

▼ Frame Check Sequence

帧尾最后出现的是 FCS▼。用它可以检查帧是否有所损坏。在通信传输过程中如果出现电子噪声的干扰，可能会影响发送数据导致乱码位的出现。因此，通过检查这个 FCS 字段的值可以将那些受到噪声干扰的错误帧丢弃。

▼只是这时计算余数时，除了减法还会使用异或运算。

FCS 中保存着整个帧除以生成多项式的余数▼。在接收端也用同样的方式计算，如果得到 FCS 的值相同，就判定所接收的帧没有差错▼。

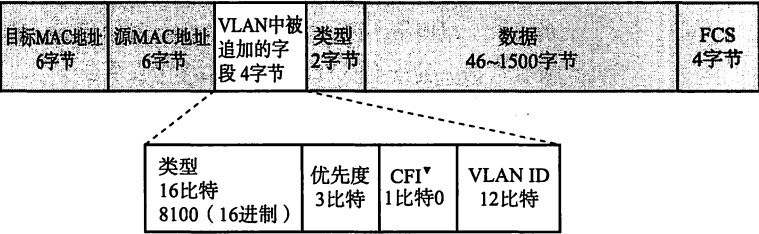
▼ FCS 具有较强的检错能力，能够检测出大量突发错误。

IEEE802.3 Ethernet 与一般的以太网在帧的首部上稍有区别。一般以太网帧中表示类型的字段，在 IEEE802.3 以太网中却表示帧的长度。此外，数据部分的前端还有 LLC 和 SNAP 等字段。而标识上一层协议类型的字段就出现在这个 SNAP 中。不过 SNAP 中指定的协议类型与一般以太网协议类型的意思基本相同。

在 3.2.6 节中将要介绍的 VLAN 中，帧的格式又会有所变化（图 3.21）。

图 3.21 VLAN 中以太网帧的格式

带有 VLAN 标记的交换机之间流动的以太网帧格式



▼ Canonical Format Indicator，标准格式指示位。当进行源路由时值为 1。

数据链路层分为两层

如果再进一步细分，还可以将数据链路层分为介质访问控制层▼和逻辑链路控制层▼。

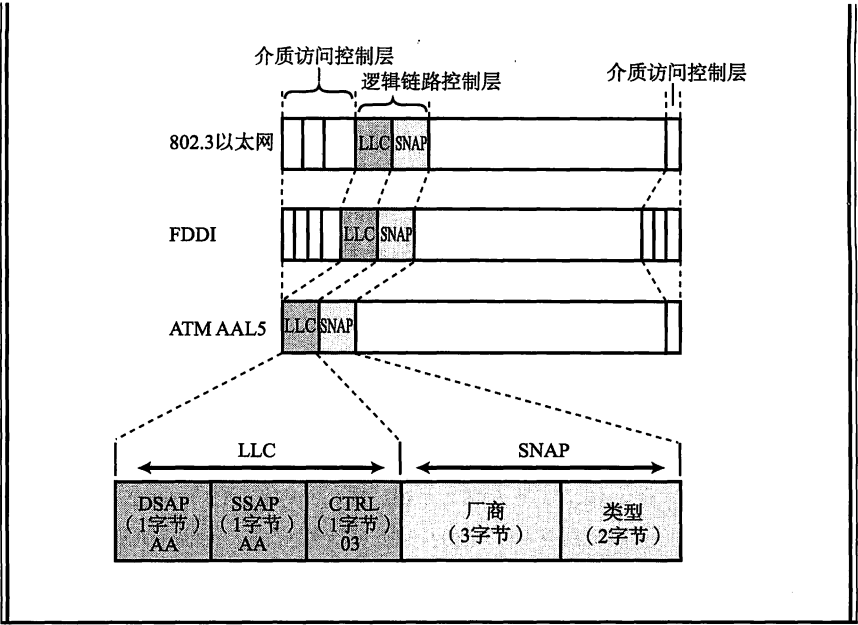
介质访问控制层根据以太网或 FDDI 等不同数据链路所特有的首部信息进行控制。与之相比，逻辑链路层则根据以太网或 FDDI 等不同数据链路所共有的帧头信息进行控制。

IEEE802.3 Ethernet 的帧格式中附加的 LLC 和 SNAP（由 IEEE802.2 制定）就是由逻辑链路控制的首部信息。从表 3.2 可以看出，当类型字段的值为 01500（05DC）时，表示 IEEE802.3 Ethernet 的长度。此时，即使参考类型对照表也无法确定上层协议的类型。在 IEEE802.3 Ethernet 中紧随其以太网首部的 LLC/SNAP 字段中包含了上层协议类型信息。因此只有查找到 SNAP 以后才能继而判断上层协议的类型。

▼介质访问控制层简称 MAC (Media Access Control)

▼逻辑链路控制层简称 LLC (Logical Link Control)

图 3.22
LLC/SNAP 格式



3.4 无线通信

无线通信通常使用电磁波、红外线、激光等方式进行传播数据。一般在办公室的局域网范围内组成的较高速的连接称为无线局域网。

无线通信不需要网线或其他可见电缆。因此，早期无线通信主要用于轻量级的移动设备。然而随着无线通信速度的不断提升，以及无线通信本身能够降低配线成本的优势，它很快在办公室、家庭、店铺以及车站和机场等环境中被广泛使用。

3.4.1 无线通信的种类

无线通信，依据通信距离可分为如表 3.3 所列出的类型。IEEE802 委员会制定了无线 PAN[▼]（802.15）、无线 LAN[▼]（802.11）、无线 MAN[▼]（802.16）以及无线 RAN[▼]（802.22）等无线标准。无线 WAN[▼] 的最典型代表就是手机通信。手机通过基站能够实现长距离通信。

- ▼ Personal Area Network
- ▼ Local Area Network
- ▼ Metropolitan Area Network
- ▼ Regional Area Network
- ▼ Wide Area Network

表 3.3
无线通信分类及其性质

分 类	通信距离	标准化组织	相关其他组织及技术
短距离无线	数米	个别组织	RF-ID
无线 PAN	10 米左右	IEEE802.15	蓝牙
无线 LAN	100 米左右	IEEE802.11	Wi-Fi
无线 MAN	数千米~100 千米	IEEE802.16、 IEEE802.20	WiMAX
无线 RAN	200 千米~700 千米	IEEE802.22	—
无线 WAN	—	GSM、CDMA2000、 W-CDMA	3G、LTE、4G、下 一代移动通信网络

* 通信距离因设备有所不同。

3.4.2 IEEE802.11

IEEE802.11 定义了无线 LAN 协议中物理层与数据链路层的一部分（MAC 层）。IEEE802.11 这个编号有时指众多标准的统称，有时也指无线 LAN 的一种通信方式。

IEEE802.11 是所有 IEEE802.11 相关标准的基础。其中定义的数据链路层的一部分（MAC 层）适用于所有 IEEE802.11 的其他标准。MAC 层中物理地址与以太网相同，都使用 MAC 地址，而介质访问控制上则使用 CSAM/CD 相似的 CSAM/CA[▼] 方式。通常采用无线基站并通过高基站实现通信。现在，各家厂商已经开始开发并销售一种具有网桥功能的（能够连接以太网与 IEEE802.11）基站设备。

作为一种通信方式，IEEE802.11 在物理层上使用电磁波或红外线，通信速度为 1Mbps 或 2Mbps。然而，这些通信速度在后续制定的 IEEE802.11b/g/a/n 等标准中逐渐被打破，以至于现在基本不被人们所使用。

- ▼ CSMA/CA Carrier Sense Multiple Access with Sollision Avoidance

表 3. 4

IEEE802. 11

标准名称	概 要
802. 11	IEEE Standard for Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications
802. 11a	Higher Speed PHY Extension in the 5 GHz Band
802. 11b	Higher Speed PHY Extension in the 2. 4 GHz Band
802. 11c	Media Access Control (MAC) Bridges – Supplement for Support by IEEE 802. 11
802. 11d	Operation in Additional Regulatory Domains
802. 11e	MAC Enhancements for Quality of Service
802. 11f	Inter – Access Point Protocol Across Distribution Systems Supporting IEEE 802. 11 Operation
802. 11g	Further Higher Data Rate Extension in the 2. 4 GHz Band
802. 11h	Spectrum and Transmit Power Management Extensions in the 5 GHz Band in Europe
802. 11i	MAC Security Enhancements
802. 11j	4. 9 GHz–5 GHz Operation in Japan
802. 11k	Radio Resource Measurement of Wireless LANs
802. 11m	802. 11 Standard Maintenance
802. 11n	High Throughput
802. 11p	Wireless Access in the Vehicular Environment
802. 11r	Fast Roaming Fast Handoff
802. 11s	Mesh Networking
802. 11t	Wireless Performance Prediction
802. 11u	Wireless Interworking With External Networks
802. 11v	Wireless Network Management
802. 11w	Protected Management Frame

出处: http://grouper.ieee.org/groups/802/11/Reports/802.11_Timelines.htm

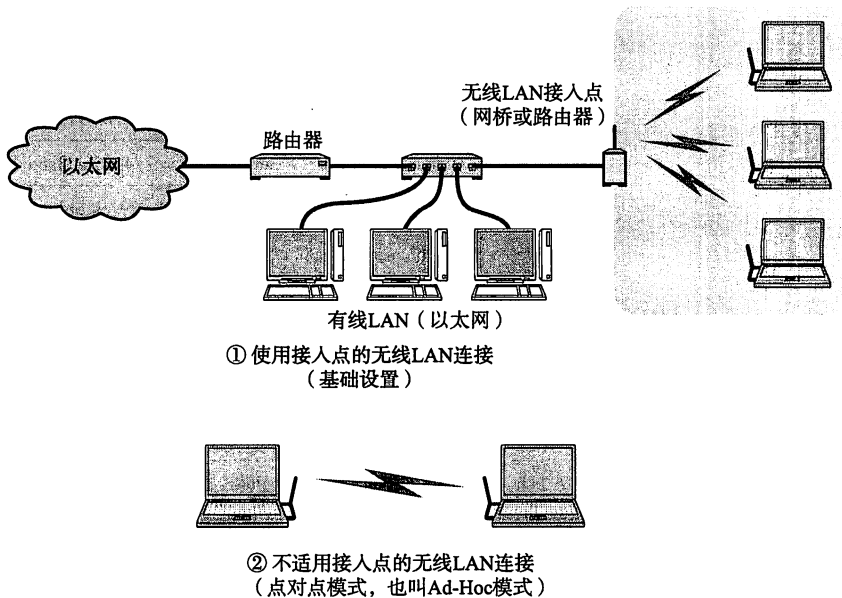
表 3. 5

IEEE802. 11 比较

传输层		TCP/UCP 等				
网络层		IP 等				
数据链路层	LLE 层	802. 2 逻辑链路控制				
	MAC 层	802. 11 MAC CSMA/CA				
物理层	方式	802. 11	802. 11a	802. 11b	802. 11g	802. 11n
	最大速度	2Mbps	54Mbps	11Mbps	54Mbps	150Mbps
	频率	2. 4GHz	5GHz	2. 4GHz	2. 4GHz	2. 4GHz/5GHz

图3.23

无线 LAN 的连接



3.4.3 IEEE802.11b 和 IEEE802.11g

▼ 2400~2497MHz

IEEE802.11b 和 IEEE802.11g 是 2.4GHz 频段▼中的无线局域网标准。它们的最大传输速度分别可达到 11Mbps (IEEE802.11b) 和 54Mbps (IEEE802.11g), 通信距离可以达到 30~50 米左右。它们与 IEEE802.11 相似, 在介质访问控制层使用 CSMA/CA 方式, 以基站作为中介进行通信。

3.4.4 IEEE802.11a

▼ 5150~5250MHz

在物理层利用 5GHz 频段▼, 最大传输速度可达到 54Mbps 的一种无线通信标准。虽然它与 IEEE802.11b/g 存在一定的兼容性问题, 但是市面上已经有支持这两方面的基站产品。再加上它不使用 2.4GHz 频段 (微波炉使用的频段), 因此也不易受干扰。

3.4.5 IEEE802.11n

▼ Multiple - Input Multiple - Output, 多入多出技术

IEEE802.11n 是在 IEEE802.11g 和 IEEE802.11a 的基础上, 采用同步多条天线的 MIMO▼技术, 实现高速无线通信的一种标准。在物理层使用 2.4GHz 或 5GHz 频段。

在使用 5GHz 频段的情况下, 若能不受其他 2.4GHz 频段系统 (802.11b/g 或蓝牙等) 的干扰, IEEE802.11n 可以达到 IEEE802.11a/b/g 的几倍带宽 (40MHz), 最大传输速度甚至可以达到 150Mbps。

■ Wi-Fi

Wi-Fi 是 WECA (Wireless Ethernet Compatability Alliance, 无线以太网兼容性联盟) 为普及 IEEE802. 11 的各种标准而打造的一个品牌名称。

WECA 从 2002 年 10 月开始已更名为 Wi-Fi Appliance。该组织向 Wi-Fi 设备厂商提供 IEEE802. 11 产品的互操作性测试, 并对合格的产品颁发 Wi-Fi Certified 认证。因此, 带有 Wi-Fi 标志的无线 LAN 设备意味着该产品已经过互操作测试并通过认证。

与音响中 Hi-Fi (High Fidelity: 高保真、高重现) 这个词类似, Wi-Fi (Wireless Fidelity) 指高质量的无线 LAN。

▣ 3.4.6 使用无线 LAN 时的注意事项

无线 LAN 允许使用者可以自由地移动位置、自由地放置设备, 通过无线电波实现较广范围的通信。这也意味着, 在其通信范围内, 任何人都可以使用该无线 LAN, 因此会有被监听或篡改的危险。

在无线 LAN 的标准中, 为防止监听或篡改, 已定义可以对传输数据进行加密。然而, 对于某些规范标准来说, 互联网上到处散布着解码的工具, 导致其弱点暴露无遗。对于即将普及的 IEEE802. 11i, 人们正在考虑使用增强型的加密技术。除了数据的加密, 应该对使用无线 LAN 的设备进行访问控制, 这样有利于构建更安全的网络环境。

此外, 无线 LAN 可以无需牌照使用特定频段。因此无线 LAN 的无线电波可能会收到其他通信设备的干扰, 导致信号不稳定。例如在一台微波炉附近使用一个 2.4GHz 带宽的 802. 11b/g 设备就得需要注意。微波炉启动后的放射出来的无线电波与设备频率相近, 产生的干扰可能会显著地降低设备的传输能力。

▣ 3.4.7 蓝牙

蓝牙与 IEEE802. 11b/g 类似, 是使用 2.4GHz 频率无线电波的一种标准[▼]。数据传输速率在 V2 中能达到 3Mbps (实际最大吞吐量为 2.1Mbps)。通信距离根据无线电波的信号的强弱, 有 1m、10m、100m 三种类型。通信终端最多允许 8 台设备[▼]。

如果说 IEEE802. 11 是针对笔记本电脑这样较大的计算机设备的标准, 那么蓝牙则是为手机或智能手机、键盘、鼠标等较小设备而设计的标准。

IEEE 在其 IEEE802. 15 规范中对 WPAN (Wireless Personal Area Network) 进行标准化。

▣ 3.4.8 WiMAX

WiMAX (Worldwide Interoperability for Microwave Access) 是使用微波在企业或家庭实现无线通信的一种方式。它如 DSL 或 FTTH 一样, 是实现无线网络关键步骤[▼]的一种方式。

▼因此, 当 IEEE802. b/g 等设备与蓝牙设备一起使用时, 无线电波信号削减有可能导致通信性能的下降。

▼其中一台为主节点, 其他 1~7 台为受管节点。这种网络也叫做 piconet, 微微网。

▼也常被形容为“最后一公里”。表示家庭或企业接入互联网时连接运营商网络的最后一段。

WiMAX 属于无线 MAN (Metropolitan Area Network), 支持城域网范围内的无线通信。由 IEEE802.16 标准化。此外, 移动终端由 IEEE802.16e (Mobile WiMAX) 标准化。

WiMAX 由 WiMAX Forum (WiMAX 论坛) 命名。该论坛还对厂商设备之间的兼容性及服务连通性进行检查。

3.4.9 ZigBee

ZigBee 主要应用于家电的远程控制^①, 是一种短距离、低功耗的无线通信技术。它最多允许 65536 个终端之间互连通信。ZigBee 的传输速度随着所使用的频率有所变化。但在日本, 使用 2.4GHz 频率的设备最高可达 250kpbs^②。

① 实际上, 工业控制、商业、公共场所以及农业控制、医疗等领域的远程控制也在广泛使用 ZigBee。——译者注
② 在我国同样最高可达 250kpbs。——译者注

3.5 PPP

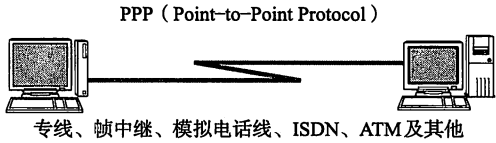
3.5.1 PPP 定义

PPP (Point-to-Point Protocol) 是指点对点, 即 1 对 1 连接计算机的协议。PPP 相当于位于 OSI 参考模型第 2 层的数据链路层。

PPP 不像以太网和 FDDI。后两者不仅与 OSI 参考模型的数据链路层有关, 还与第 1 层的物理层有关。具体来讲, 以太网使用同轴电缆或双绞线电缆, 它可以决定其中的 0、1 该被解释为何种电子信号。与之相比, PPP 属于纯粹的数据链路层, 与物理层没有任何关系。换句话说, 仅有 PPP 无法实现通信, 还需要有物理层的支持。

图 3.24

PPP



PPP 可以使用电话线或 ISDN、专线、ATM 线路。此外, 近些年人们更多是在用 ADSL 或有线电视通过 PPPoE (PPP over Ethernet) 实现互联网接入。PPPoE 是在以太网的数据中加入 PPP 帧进行传输的一种方式。

3.5.2 LCP 与 NCP

▼在使用电话线的情况下, 首先要保证电话线物理层面的连接以后才能在它之上建立 PPP 连接。

在进行数据传输前, 要先建立一个 PPP 级的连接▼。当这个连接建立以后就可以进行身份认证、压缩与加密。

在 PPP 的主要功能中包括两个协议: 一个是不依赖上层的 LCP 协议 (Link Control Protocol), 另一个是依赖上层的 NCP 协议 (Network Control Protocol)。如果上层为 IP, 此时的 NCP 也叫做 IPCP (IP Control Protocol)。

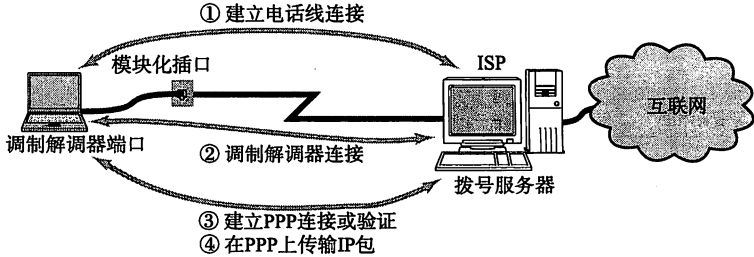
LCP 主要负责建立和断开连接、设置最大接收单元 (MRU, Maximum Receive Unit)、设置验证协议 (PAP 或 CHAP) 以及设置是否进行通信质量的监控。

而 IPCP 则负责 IP 地址设置以及是否进行 TCP/IP 首部压缩等设备▼。

▼设备之间的这种交互也叫协商 (Negotiation)。

图 3.25

PPP 中建立连接



通过 PPP 连接时, 通常需要进行用户名密码的验证, 并且对通信两端进行双向的验证▼。其验证协议有两种, 分别为 PAP (Password Authentication Protocol) 和 CHAP (Challenge Handshake Authentication Protocol)。

▼通过 ISP 接入互联网时, 一般对 ISP 端不验证。

PAP 是 PPP 连接建立时，通过两次握手进行用户名和密码验证。其中密码以明文方式传输。因此一般用于安全要求并不很高的环境，否则会有窃听或盗用连接的危险。

CHAP 则使用一次性密码 OTP（One Time Password），可以有效防止窃听。此外，在建立连接后还可以进行定期的密码交换，用来检验对端是否中途被替换。

3.5.3 PPP 的帧格式

PPP 的数据帧格式如图 3.26 所示。其中标志码用来区分每个帧。这一点与 HDLC 协议非常相似，因为 PPP 本身就是基于 HDLC 制定出来的一种协议。

HDLC 就是在每个帧的前后加上一个 8 位字节“01111110”用来区分帧。这一个 8 位字节叫做标志码。在两个标志码中间不允许出现连续 6 个以上的“1”。因此，在发送帧的时候，当出现连续 5 个“1”时后面必须插入一个 0。而当接收端在接收帧时，如果收到连续的 5 个“1”且后面跟着的是 0，就必须删除。由于最多只会出现 5 个连续的“1”，就可以比较容易地通过标志码区分帧的起始与终止。而 PPP 标准帧格式与此完全相同。

▼ HDLC High Level Data Link Control Procedure，高级数据链路控制。

图 3.26 PPP 数据帧格式

PPP数据帧格式（按照标准设定）

标志1 字节 (01111110)	地址1 字节 (11111111)	控制1 字节 (00000011)	类型 2字节	数据0~1500字节	FCS 4字节	标志1 字节 (01111110)
-------------------------	-------------------------	-------------------------	-----------	------------	------------	-------------------------

另外，在通过电脑进行拨号时，PPP 已在软件中实现。因此，那些插入或删除“0”的操作或 FCS 计算都交由电脑的 CPU 去处理。这也是为什么人们常说 PPP 这种方式会给计算机带来大量负荷的原因所在。

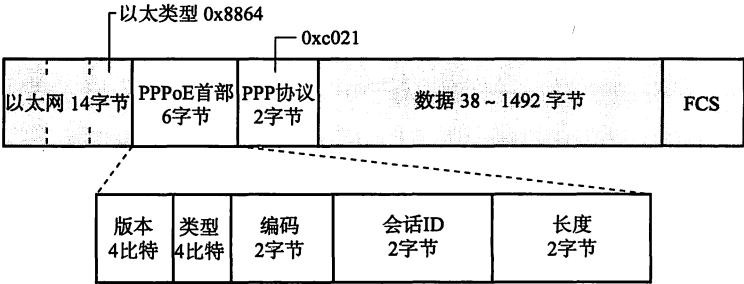
3.5.4 PPPoE

有些互联网接入服务商在以太网上利用 PPPoE（PPP over Ethernet）提供 PPP 功能。

在这种互联网接入服务中，通信线路由以太网模拟。由于以太网越来越普及，在加上它的网络设备与相应的 NIC 价格比较便宜，因而 ISP 能够提供一个单价更低的互联网接入服务。

单纯的以太网没有验证功能，也没有建立和断开连接的处理，因此无法按时计费。而如果采用 PPPoE 管理以太网连接，就可以利用 PPP 的验证等功能使各家 ISP 可以有效地管理终端用户的使用。

图 3.27 PPPoE 数据帧格式



3.6

其他数据链路

▼其中很多类型可能已经不再使用。

到此为止，我们已经介绍过以太网、无线通信以及 PPP 等数据链路。除此之外，还很多其他类型的数据链路▼。本节将对它们做一个简单介绍。

3.6.1 ATM

ATM (Asynchronous Transfer Mode) 是以一个叫做信元 (5 字节首部加 48 字节数据) 的单位进行传输的数据链路，由于其线路占用时间短和能够高效传输大容量数据等特点主要用于广域网的连接。ITU▼ 和 ATM 论坛负责对 ATM 进行标准化。

ATM 的特点

ATM 是面向连接的一种数据链路。因此在进行通信传输之前一定要设置通信线路。这一点与传统电话很相似。使用传统电话进行通话时，需要事先向交换机发出一个信令要求，建立交换机与通话对端的连接▼。而 ATM 又与传统电话不同，它允许同时与多个对端建立通信连接。

ATM 中没有类似以太网和 FDDI 那种发送权限的限制。它允许在任何时候发送任何数据。因此，当大量计算机同时发送大量数据时容易引发网络拥堵甚至使网络进入收敛状态▼。为了防止这一现象的出现，ATM 中也增加了限制带宽的细分功能。

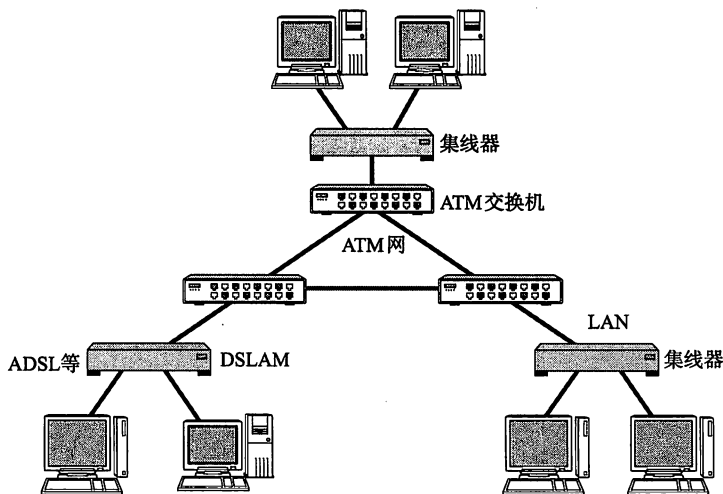
▼ International Telecommunication Union, 国际电信联盟。

▼ ATM 中把它叫做 SVC (Switched Virtual Circuit, 交换式虚电路)。另外也有使用固定线路的方式，叫做 PVC (Permanent Virtual Circuit, 永久虚电路)。

▼收敛状态指当网络非常拥堵时，路由器或交换机无法完成包的处理，从而丢弃这些包的一种状态。

图 3.28

ATM 网络



▼时分复用设备。

▼实际上它采用 TDM 方式的 SONET (Synchronous Optical Network) 或 SDH (Synchronous Digital Hierarchy) 的线路。

▼在 VPI 所标识的通信线路中, 用 VCI 识别多个通信。

▼网络开销是指在通信传输中, 除了发送实际想要发送的数据, 还需要附加的一些控制信息所耗的带宽开销以及处理这些信息所耗的时间开销。

图 3-29

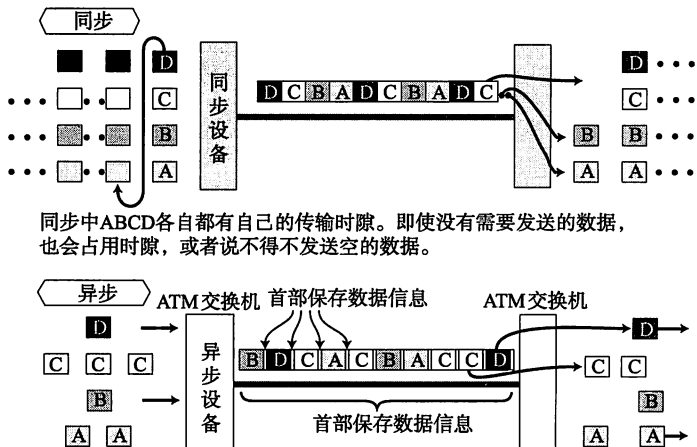
同步与异步

同步与异步

以多个通信设备通过一条电缆相连的情况为例。首先, 这样连接的设备叫做 TDM[▼]。TDM 通常在两端 TDM 设备之间同步的同时, 按照特定的时间将每个帧分成若干个时隙, 按照顺序发送给目标地址。这一过程与装配零件的车间作业非常相似。例如在汽车零件装配工厂, 传送带上传送着各种颜色的汽车。工人们或自动化设备可以根据汽车的颜色将特定的零件附加到相应的车身上。在这里每个颜色的汽车叫做插槽, 就相当于 TDM 中的时隙。即使某个汽车的车身缺少某些零件, 如果颜色不同就无法将零件安装上去。在 TDM 中也是如此, 不论是否还有想要发送的数据, 时隙会一直被占有, 从而可能会出现很多空闲的时隙。因此, 这种方式的线路利用率比较低。

ATM 扩展了 TDM, 能够有效地提高线路的利用率[▼]。ATM 在 TDM 的时隙中放入数据时, 并非按照线路的顺序而是按照数据到达的顺序放入。然而, 按照这样的顺序存放的数据在接收端并不易辨认真正的内容。为此, 发送端还需要附加一个 5 字节的包首部, 包含 VPI (Virtual Path Identifier)、VCI (Virtual Channel Identifier) 等识别码[▼]用来标识具体的通信类型。这种 VPI 与 VCI 的值只在直连通信的两个 ATM 交换机之间设置。在其他交换机之间意思则完全不同。

ATM 中信元传输所占用的时隙不固定, 一个帧所占用的时隙数也不固定, 而且时隙之间并不要求连续。这些特点可以有效减少空闲时隙, 从而提高线路的利用率。只不过需要额外附加 5 个字节的首部, 增加了网络的开销[▼], 因此也在一定程度上降低了通信速度。也就是说, 在一个 155Mbps 的线路上由于 TDM 和 ATM 的网络开销, 实际的网络吞吐也只能到 135Mbps。



同步中ABCD各自都有自己的传输时隙。即使没有需要发送的数据, 也会占用时隙, 或者说不发送空的数据。

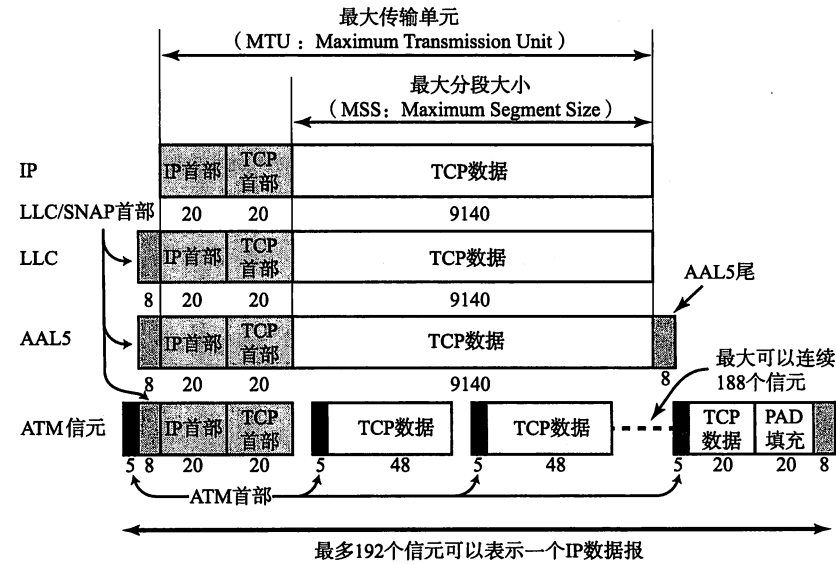
异步中在包首部位明确指明了目标地址, 因此只有在有必要发送时发送数据。

■ ATM 与上层协议

在以太网中一个帧最大可传输 1500 个字节，FDDI 可以最大传输 4352 字节。而 ATM 的一个信元却只能发送固定的 48 字节数据。这 48 字节的数据部分中若包含 IP 首部和 TCP 首部，则基本无法存放上层的数据。为此，一般不会单独使用 ATM，而是使用上层的 AAL (ATM Adapter Layer)▼。在上层为 IP 的情况下，则叫做 AAL5。如图 3.30 所示，每个 IP 包被附加各层的协议首部以后，最多可以被分为 192 个信元发送出去。

▼从 ATM 的角度是上一层，但对 IP 来说属于下一层。

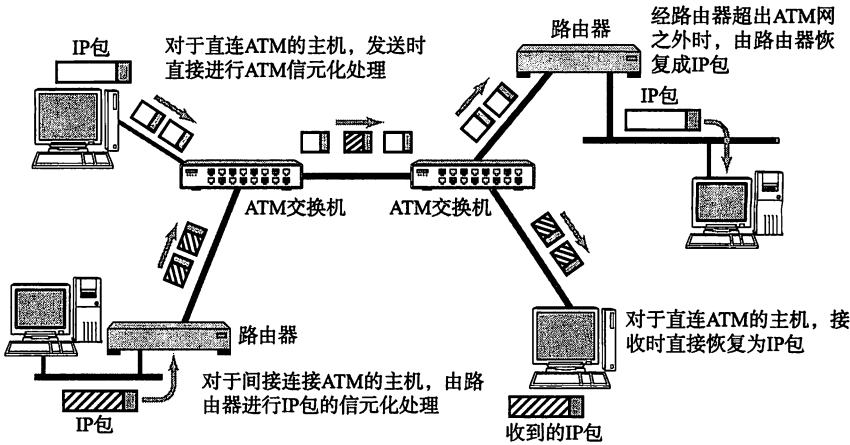
图 3.30 数据包的 ATM 信元封装



从这个图中还可以看出，在整个 192 个信元中只要有一个丢失，那么整个 IP 包就相当于被损坏。此时，AAL5 的帧检查位报错，导致接收端不得不丢弃所有的信元。前面曾提到 TCP/IP 在包发生异常的时候可以实现重发，因此在 ATM 网中即使只是一个信元丢失，也要重新发送最多 192 个信元。这也是 ATM 到目前为止的最大弊端。一旦在网络拥堵的情况下，只要丢掉哪怕 1% 的信元也会导致整个数据都无法接收。特别是由于 ATM 没有发送权限上的控制，很容易导致网络收敛。为此，在构建 ATM 网络的时候，必须保证终端的带宽合计小于主干网的带宽，还要尽量保证信元不易丢失。目前人们已经开始研究在发生网络收敛时，动态调整 ATM 网络带宽的技术。

图 3.31

ATM 中 IP 包的发送



3.6.2 POS

▼ Synchronous Digital Hierarchy, 同步数字体系。

▼ Synchronous Optical Network, 同步光纤网络。

POS (Packet over SDH/SONET) 是一种在 SDH (SONET) 上进行包通信的一种协议。SDH (SONET) 是在光纤上传输数字信号的物理层规范。

SDH 作为利用电话线或专线等可靠性较高的方式进行光传输的网络，正被广泛应用。SDH 的传输速率以 51.84Mbps 为基准，一般为它的数倍。目前，已经有针对 40Gbps SDH 的 OC 768 产品^①。

3.6.3 FDDI

FDDI (Fiber Distributed Data Interface) 叫做分布式光线数据接口。曾几何时，人们为了用光纤和双绞线实现 100Mbps 的传输速率，在主干网或计算机之间的高速连接上广泛使用了 FDDI。但是由于后来高速 LAN 提供了 Gbps 级的传输速率，FDDI 也就逐渐淡出了应用领域。

FDDI 采用令牌 (追加令牌) 环的访问方式。令牌环访问方式在网络拥堵的情况下极容易导致网络收敛。

FDDI 中的每个站通过光纤连接形成环状，如图 3.32 所示。FDDI 为了防止在环在某处断开时导致整个通信的中断，采用双环的结构。双环中站叫做 DAS，单环中的站叫做 SAS。

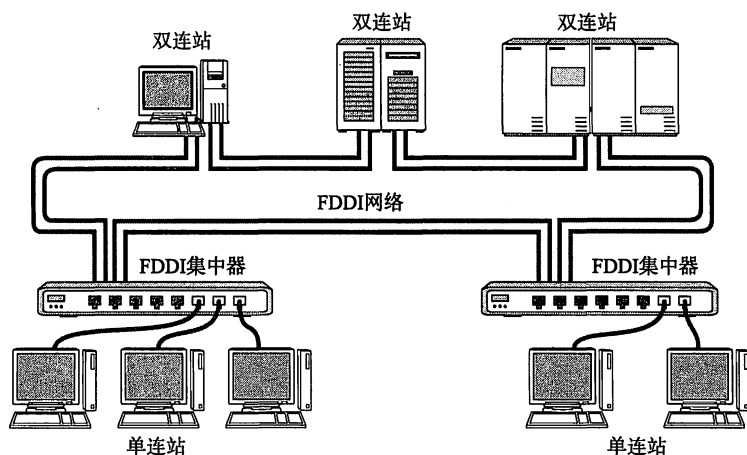
▼ Dual Attachment Station, 双连站。

▼ Single Attachment Station, 单连站。

① OC (光学载波) 是 SONET 光纤网络中的一组信号带宽，通常表示为 OC-n，其中，n 是一个倍数因子，表示是基本速率 51.84Mbit/s 的倍数。——译者注

图 3.32

FDDI 网络



3.6.4 Token Ring

令牌环网 (Token Ring) 源自 IBM 开发的令牌环 LAN 技术, 可以实现 4Mbps 或 16Mbps 传输速率。前面提到的 FDDI 实际上是扩展了 Token Ring 的一个产物。

令牌环由于其价格一直居高不下以及所支持的提供商逐渐较少等原因, 除了在 IBM 的环境以外始终未能得到普及, 而且随着以太网的广泛使用, 人们已经不再采用令牌环技术。

3.6.5 100VG-AnyLAN

100VG-AnyLAN 是 IEEE802.12 规范定义的一种网络协议。VG 为 Voice Grade 的缩写, 指语音级。它以语音级的 3 类 UTP 电缆实现 100Mbps 的传输速率。它的数据帧格式既能应对以太网又能应对令牌环网。在传输方式上, 它采用扩展了令牌传递方式的需求优先[▼]访问方式。在这种方式中, 交换机负责控制发送权。鉴于 100Mbps 以太网 (100BASE-TX) 的普及, 100VG-AnyLAN 也几乎不再被使用。

▼ Demand Priority。在数据帧里附加了一个优先级的信息, 使得包可以按照优先级发送给对端。

3.6.6 光纤通道

光纤通道 (Fiber Channel) 是实现高速数据通信的一种数据链路。与其说它是一种网络, 不如说它更像是 SCSI 那样类似于连接计算机周边设备的总线一样的规范。数据传输速率为 133Mbps~4Gbps。近些年被广泛用于搭建 SAN[▼], 成为其主要数据链路。

▼ Storage Area Network, 存储域网络。服务器与多台存储设备 (硬盘、磁带备份) 之间高速传输数据的网络系统。一般在企业当中用于保存超大容量数据。

3.6.7 HIPPI

HIPPI 用于连接超大型计算机传输速率为 800Mbps 或 1.6Gbps。铜缆的实际传输距离在 25 米以内, 但是如果使用光纤作为传输介质时, 可以延长到数公里。

3.6.8 IEEE1394

也叫 FireWire 或 i.Link, 是面向家庭的局域网, 主要用于连接 AV 等计算机外围设备。数据传输速率为 100~800Mbps 以上。

3.6.9 HDMI

HDMI 是 High-Definition Multimedia Interface 的缩写, 意为高清晰度多媒体接口。它可以通过一根缆线实现图像和声音等数字信号的高品质传输。曾主要用于 DVD/蓝光播放器、录像机、AV 功放等设备与电视机、投影仪的连接, 现在也逐渐开始用于计算机或平板电脑、数码相机与显示器的连接。从 2009 年发布的 1.4 版开始它可以传输以太网帧, 使得采用 HDMI 介质实现 TCP/IP 通信变为可能。关于它今后的发展, 让我们拭目以待。

3.6.10 iSCSI

▼ RFC3720、RFC3783

它是将个人电脑连接硬盘的 SCSI 标准应用于 TCP/IP 网络上的一种标准▼。它把 SCSI 的命令和数据包含进 IP 包内, 进行数据传输。由此, 人们就可以像使用个人电脑内嵌的 SCSI 硬盘一样使用网络上直连的大规模硬盘了。

3.6.11 InfiniBand

▼ 如 4 链接或 12 链接。

InfiniBand 是针对高端服务器的一种超高速传输接口技术。它最大的特点是高速、高可靠性以及低延迟。它支持多并发链接, 将多个线缆▼合并为一个线缆。可以实现从 2Gbps 至数百 Gbps 的传输速率。以后甚至还计划提供数千 Gbps 的高速传输速率。

3.6.12 DOCSIS

▼ Multidmedia Cable Network
System Patners Limited

DOCSIS 是有线电视 (CATV) 传输数据的行业标准, 由 MCNS▼制定。该标准定义了有线电视的同轴电缆与 Cable Modem (电缆调制解调器) 的连接及其与以太网进行转换的具体规范。此外, 有一个叫做 CableLabs (有线电视业界的研究开发机构) 的组织对 Cable Modem 进行认证。

3.6.13 高速 PLC

▼ Power line Communication,
高速电力线通信。

高速 PLC▼是指在家里或办公室内利用电力线上数 MHz~数十 MHz 频带范围, 实现数十 Mbps~200Mbps 传输速率的一种通信方式。使用电力线不用重新布线, 也能进行日常生活以及家电设备或办公设备的控制。然而, 本不是为通信目的而设计的电力线在传输高频信号时, 很容易收到电波干扰, 一般仅限于室内 (家里、办公室内) 使用。

表 3-6

主要数据链路类型及其特点

数据链路名称	介质传输速率	用 途
以太网	10Mbps ~ 1000Gbps	LAN、MAN
802.11	5.5 Mbps ~ 150 Mbps	LAN
Bluetooth	上限 2.1 Mbps, 下限 177.1 kbps	LAN
ATM	25 Mbps、155 Mbps、622 Mbps、2.4GHz	LAN ~ WAN
POS	51.84 Mbps ~ 约 40Gbps	WAN
FDDI	100 Mbps	LAN、MAN
Token Ring	4 Mbps、16 Mbps	LAN
100VG-AnyLAN	100 Mbps	LAN
光纤通道	133 Mbps ~ 4Gbps	SAN
HIPPI	800 Mbps、1.6Gbps	两台计算机之间的连接
IEEE1394	100 Mbps ~ 800 Mbps	面向家庭

3.7

公共网络

前面介绍了很多局域网连接相关的知识。本小节旨在介绍连接公共通信服务相关的细节。所谓的公共通信服务类似于电信运营商（如 NTT、KDDI 或软银等）提供的电话网络。人们通过与这些运营商签约、付费不仅可以实现联网还可以与距离遥远的机构组织进行通信。

这里将分别介绍模拟电话线路、移动通信、ADSL、FTTH、有线电视、专线、VPN 以及公共无线 LAN 等内容。

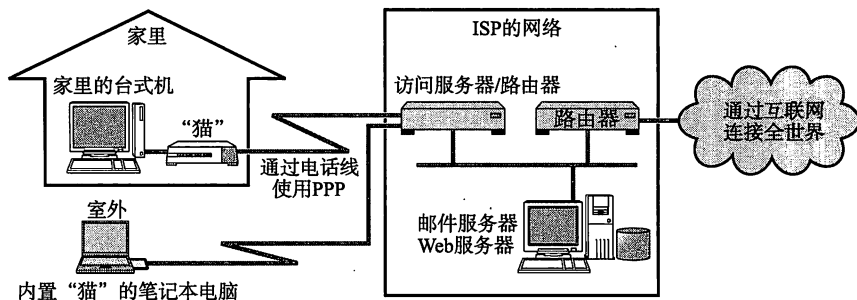
3.7.1 模拟电话线路

模拟电话线路其实就是利用固定电话线路进行通信。电话线中的音频带宽用于拨号上网。该方法不需要特殊的通信线路，完全使用已普及的电话网。

让计算机与电话线相连需要有一个将数字信号转换为模拟信号的调制解调器（俗称“猫”）。“猫”的传输速率一般只在 56kbps 左右，所以现在已逐渐被淘汰。

图 3.33

拨号连接



3.7.2 移动通信服务

在日本，移动通信服务包括手机和 PHS^① 服务。它们的特点是：只要在服务区范围内，就可以连接到运营商的网络。

以前手机通信的传输速率相对较低。而现在随着 Mobile WiMAX（参考 3.4.8 节）和 LTE（参考 1.10.3 节）等技术的发展，手机的传输速率可以达到数 Mbps 甚至几十 Mbps 不等。

PHS 的数字通信方式有以电路交换为基础的 PIAFS[▼]（最大 64kbps）和分包通信（最大 800kbps）两种方式。此外，还有更多实现高速通信的全新方式也在被不断提出。

3.7.3 ADSL

ADSL[▼] 是对已有的模拟电话线路进行扩展的一种服务。模拟电话线路虽然也能传输高频数字通信，但是它与电信局的交换机之间只有发送音频信号时才能显

▼ Internet Access Forum Standard 的缩写。

▼ Asymmetric Digital Subscriber Line，非对称数字用户环路。

① Personal Handy-phone System，类似于我国的小灵通。——译者注

示极好的传输效率，并会对其他多余频率的信号进行丢弃。尤其是在近几年，随着电话网逐渐数字化，通过电话线路的信号再经过电信的交换机时会变成 64kbps 左右的数字信号。因此，从理论上就无法传输 64kbps 更快的数字信号。然而，每个话机到电信局交换机之前的这段线路，是可以实现高速传输的。

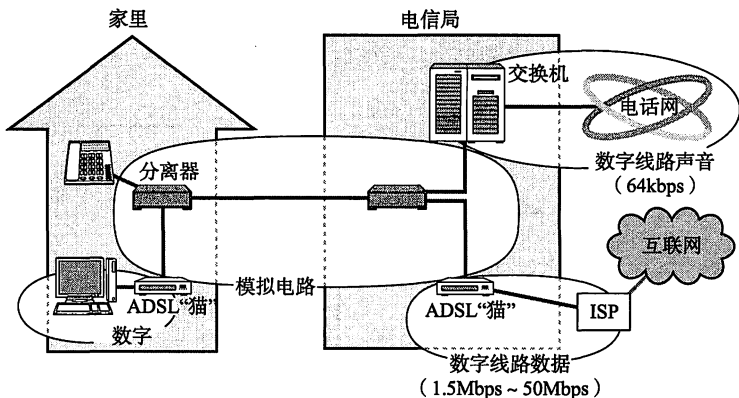
ADSL 正是利用话机到电信局交换机之间这段线路，附加一个叫做分离器的装置，将音频信号（低频信号）和数字信号（高频信号）隔离以免产生噪声干扰。

类似这种类型的通信方式除了 ADSL 还有其他诸如 VDSL、HDSL、SDSL 等。它们被统称为 xDSL。ADSL 是其中最为普及的一种方式。

ADSL 中的线路速度根据通信方式或线路的质量以及距离电信局的远近有所不同。从 ISP 到家里/办公室的速率在 1.5Mbps~50Mbps 左右，而从家里/办公室到 ISP 端的速率一般在 512kbps~2Mbps 左右。

图 3.34

ADSL 连接



3.7.4 FTTH

FTTH（Fiber To The Home）顾名思义就是一根高速光纤直接连到用户家里或公司建筑物处的方法。它通过一个叫做 ONU[▼] 的装置将计算机与之关联。该装置负责在光信号与电子信号之间的转换。使用 FTTH 可以实现稳定的高速通信。不过它的线路传输速率与具体的服务内容仍受个别运营商限制。

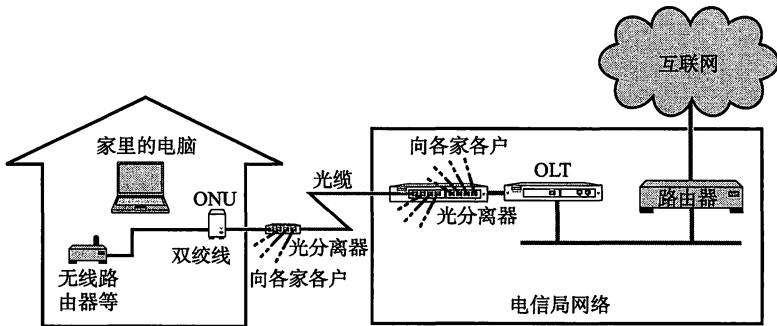
以上属于光纤到户。还有一种方式叫光纤到楼。它是指一个高速光纤直接连到某个大厦、公司或宾馆的大楼，随后在整个大楼内部再通过布线实现联网。简称 FTTB（Fiber To The Building）。甚至还有一种方式是将光纤接入到某个家庭以后，再通过布线实现周围几户住家共同联网。这种方式简称为 FTTC（Fiber To The Curb[▼]）。

▼ Optical Network Unit，光网络单元。其局端光线路终端叫做 OLT（Optical Line Terminal）。

▼ Curb 意指住宅周边的绿石小路。

图 3.35

FTTH 连接



▼有关光纤电缆与 WDM 的更多细节请参考附录 4.3 节。

另外，光缆通常由一条用来发送数据和另一条用来接收数据的线对组成。然而在 FTTH 中使用的是 WDM[▼]，即发送和接收双方都使用同一根线缆。接入每家每户的这些光纤电缆又通过 ONU 与 OLT 之间的光分离器相互隔离。

3.7.5 有线电视

电视最初使用无线电波发送信号。后来发展为使用线缆的有线电视。使用无线电波的时候，电视信号经常会受天线的设置状况以及周围其他建筑物的干扰。而有有线电视则很少受这种干扰，因此传送画质也明显好于传统电视。

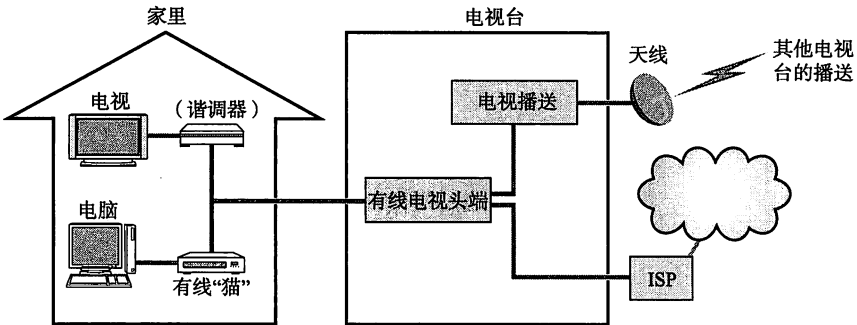
近几年通过有线电视接入互联网的服务又得到推广。这种方式通过利用空闲的频道传输数据实现通信。

其中从电视台到用户住宅使用与电视播送相同的频率带宽[▼]，而从住家住宅到电视台则使用播送当中未使用的低频带宽[▼]。因此这种方式有一个特点就是数据传输的上行速度低于下行速度。

▼称为下行 (DownStream)。

▼称为上行 (UpStream)。

图 3.36
通过有线电视连接互联网



通过有线电视连接互联网时，首先需要到有线电视台申请该项服务。购置用来进行通信的有线调制解调器（有线“猫”）以后就可以与局端的有线电视头端相连。端头负责将数字播送或部分模拟播送与数字信息之间通过一根线缆进行收发转换。

连网时，用户发送的信息由有线“猫”进行转换，经由有线电视网以后再接入具体的 ISP。在有线电视网中使用者一种叫做 DOCSIS[▼] 的标准，最大可实现 160Mbps 的传输速率。

▼具体请参考 3.6.12 节。

3.7.6 专线

随着互联网用户的急剧上升，专线服务向着价格更低、带宽更广以及多样化的方向发展。现在市面上已经出现了各种各样的“专线服务”。以 NTT Group 的服务为例，有 Mega Data Nets（用 ATM 接口提供 3Mbps~42Mbps 的专线接入）、ATM Mega-Link、Giga Stream（用以太网或 SONET/SDH 接口提供 0.5Mbps~135Mbps 的专线接入）等众多专线接入服务。

专线的连接一定是一对一的连接。虽然 ATM 的设计初衷允许有多个目标地，但对于提供专线服务的 ATM Mega-Link 中也只能指定一个目的地。因此不可能像 ISDN 或帧中继那样引进一条线缆就能连接众多目的地。

3.7.7 VPN

虚拟专用网络（VPN）用于连接距离较远的地域。这种服务包括 IP-VPN 和广域以太网。

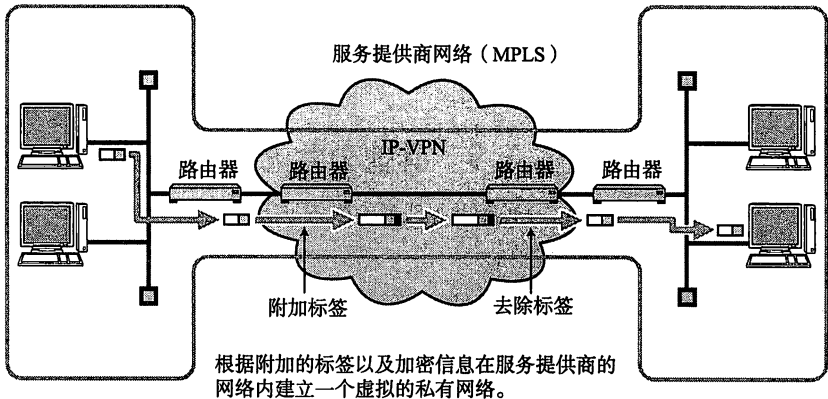
IP-VPN

意指在 IP 网络（互联网）上建立 VPN。

网络服务商提供一种在 IP 网络上使用 MPLS 技术构建 VPN 的服务。其中 MPLS（Multiprotocol Label Switching，多协议标签交换）在 IP 包中附加一个叫做标签（Label）的信息进行传输控制。每个用户的标签信息不同，因此在通过 MPLS 网时，可以轻松判断出目标地址。这样一来就可以将多个不同用户的 VPN 信息通过 MPLS 网加以区分，形成封闭的私有网络。此外，还能进行用户级的带宽控制。

有时也叫 tag。

图 3-37 IP-VPN (MPLS)



为了与 IP-VPN 相区别，这种方式的 VPN 也叫做企业互联网 VPN。
关于 IPsec 的更多细节请参考 9.4.1 节。

除了使用服务提供商的 IP-VPN 服务之外，有时企业还可以在互联网上建立自己的 VPN，一般采用的是 IPsec 技术。该方法对 VPN 通信中的 IP 包进行验证和加密，在互联网上构造一个封闭的私有网络。虽然这种方式可以利用价格低廉的互联网通信线路，并且还可以根据自己的情况对数据进行不同级别的加密，但有时会受到网络拥堵的影响。

广域以太网

服务提供商所提供的用于连接相距较远的地域的一种服务。IP-VPN 是在 IP 层面的连接，广域以太网则是在作为数据链路层的以太网上利用 VLAN（虚拟局域网）实现 VPN 的技术。该技术还可以使用 TCP/IP 中的其他协议。

广域以太网以企业专门使用服务提供商构建的 VLAN 网络为主要形式。只要指定同一个 VLAN，无论从哪里都能接入到同一个网络。由于广域以太网利用的是数据链路层技术，因此为了避免一些不必要的信息传输，使用者应谨慎操作。

3.7.8 公共无线 LAN

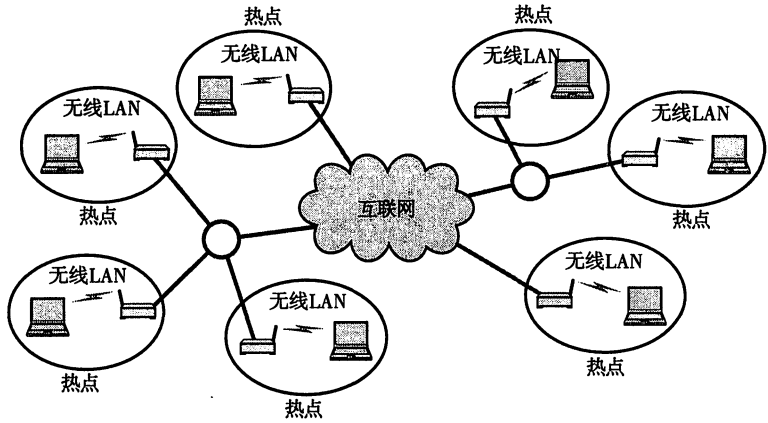
公共无线 LAN 是指公开的可以使用 Wi-Fi（IEEE802.11b 等）的服务。服务提供者可以在车站或餐饮店等人员相对比较集中的地方架设的一个叫做热点（HotSpot）的无线电波接收器。使用者到达这些区域就可以使用带有无线 LAN 网

卡的笔记本电脑或智能手机连接上网。

上网时使用者首先要通过这些热点建立互联网连接。连接以后，还可以通过那些利用 IPsec 技术实现的 VPN 连接到自己公司的内网。这种接入服务有时免费（如商场、车站等场所），有时也可能是收费的。

图 3.38

无线 LAN



3.7.9 其他公共无线通信服务

其他公共无线通信服务包括 X.25、帧中继和 ISDN。

X.25

X.25 网是电话网的改良版。它允许一个端点连接多个站点，传输速率为 9.6kbps 或 64kbps。由于现在已出现其他多种网络服务，X.25 已经不再使用。

帧中继

帧中继是对 X.25 进行精简并高速化的网络。与 X.25 相似，它允许 1 对 N 的通信，一般提供 64kbps~1.5Mbps 的传输速率。目前由于以太网和 IP-VPN 的广泛应用，帧中继的用户也在逐渐减少。

ISDN

ISDN 是 Integrated Services Digital Network（综合业务数字网）的缩写。它是一种集合了电话、FAX、数据通信等多种类型的综合公共网络。目前它的使用者也在日趋减少。