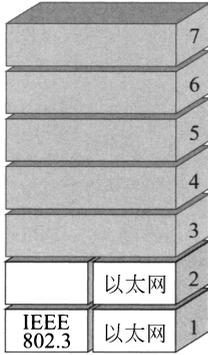


# 第 3 章 以太网和 IEEE 802.3 协议

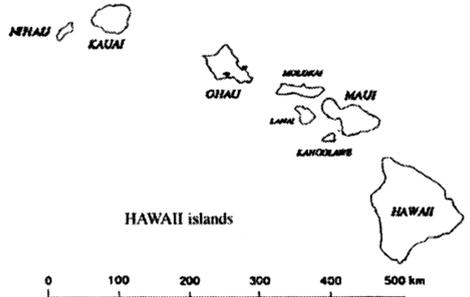


IEEE 802.3 和以太网协议属于互相竞争的协议族。尽管有不兼容性，这两个协议还是有许多共同点。

以太网实际上是最常用的协议。本章展示该协议的演化以及它如何能够适应技术进步来保持其竞争力并在大多数情况下得以加强其使用。

## 3.1 历史

以太网网络源于 20 世纪 70 年代初期的夏威夷岛。夏威夷大学分布于许多不同的岛屿上。校本部位于瓦胡岛，而其他 7 个校园位于夏威夷州的其他 4 个岛上。在 20 世纪 70 年代，为了能够与不同站点互连，开发了一套通过无线电发射来传输数据的系统。



每个校园的每座建筑都有一套无线电收发装置，以同一频率传送和接收数字数据。传送速度是 2400 b/s。每个装置都接收头部包含地址的数据块，这个地址标识了要读入该数据块的唯一接收方是谁。有时候数个站同时发送，造成数据块冲突。于是发射信号变得模糊，数据丢失。

夏威夷大学的 Norman Abramson 教授发表了一些关于这一系统的理论和应用的文档。这套系统名为 **Aloha**，在夏威夷语中意为问候语。

该系统激励了位于 Palo Alto 的施乐研究中心的 R. Metcalfe 和 D. Boggs 这两个人。在 1973 年 5 月 22 日，他们在一份报告中建立了以太网。以太网以 Aloha 网络为基础，其最初名称是 Alto Aloha 网络。这一网络用于链接个人计算机，并使用同轴电缆。该试验网络的速率是 3 Mb/s。

在科学的历史上，以太被定义为一种不可衡量的假定柔性流体，我们将其看成是光和电的传送代理。Michelson 和 Morley 在 1887 年证明了其不存在。他们展示了在地球和以太之间没有相对移动。在此基础上，爱因斯坦在 1905 年建立了相对论。这一术语再次被使用，这次用来对通过传输电磁波来传送分组广播的被动支撑体命名。

1976 年 7 月，R. Metcalfe 和 D. Boggs 在一篇标题为 *Ethernet: Distributed Packet*

*Switching for Local Computer Networks* 的文档中将以太网介绍给了科学界。在 1979 年 2 月, Digital Equipment 公司和 Xerox 公司开始协商开发这一标准。在当年 6 月, Intel 公司加入了这一小组, 因为需要将电子设备集成到以太网标准中。在 1980 年, 这 3 家公司宣布共同将以太网定义为工业标准。速率为 3 Mb/s 的以太网被认为是试验以太网, 而速率为 10 Mb/s 的以太网的部署是 Digital-Intel-Xerox 协作的第一个版本, 称为 DIX 以太网版本 1.0。该标准提交给了 IEEE 和欧洲计算机制造商协会 (ECMA)。

经过一些修改之后, DIX 的 2.0 版本完成了, 称为以太网 II。1982 年, ECMA 技术委员会正式接受这一版本。1983 年 6 月 23 日, IEEE 和美国国家标准学会 (ANSI) 采纳了以太网标准, 成为 ANSI/IEEE 802.3: “带有冲突侦测的载波监听多路访问 (CSMA/CD) 的访问方法以及物理层规范”。

IEEE 没有完全接受 DIX 版本 2.0, 它修改了一部分以便让其显得正式。这些不同带来标准的不兼容性, 不仅在访问支撑体上, 也在上层处理信息的方法上。

从创建以来, 该协议就考虑技术革新。1998 年的版本集成了 IEEE 802.3 标准的演化, 内容长达 1200 页。最重要的演化是从总线拓扑 (共享支撑) 变化到开始使用让所有设备通过点对点连接 (专用连接) 来链接的架构。速率从 10 Mb/s 改进到了 100 Mb/s。目前是 1000 Mb/s。有些研究提出 10 Gb/s 的速度。比如, 在以太网标准化之后帧格式没有修改以保证兼容性以及不同世代的设备之间的互连。

以太网成功的另一个因素是, 它易于使用和配置。无论选择的技术是什么, 都能操作网络 (第 2 层), 无须在终端或者互连设备上进行任何配置。自然地, 一旦网络构建完成, 就可能对功能做优化并添加一些修改默认配置的安全过程。

## 3.2 物 理 层

### 3.2.1 支撑体

IEEE 802.3 协议在数个物理支撑体上开发。每个都由一个代码来定义, 它表明:

- 以 Mb/s 标识的传输速率 (1、10、100 或者 1000);
- 编码方法 (基带 (base) 或宽带 (broad));
- 一个表示同轴电缆最大段长度的数字或者一个给出支撑体类型的字符 (T 表示双绞线, F 则表示光纤)。

表 3.1 给出了在 IEEE 802.3 网络中所用的不同支撑体类型。这个表也给出了能找到与这些支撑体有关的更多信息的参考处。

表 3.1 IEEE 802.3 支撑体

速度	支撑体	类型	模 式	注 释
1 Mb/s	PT	1BASE5	HD	在经典的电话线上的 1 Mb/s。设备和集线器之间的距离可以是 250 m。这是 IEEE 802.3 小组的最初规范。现在可以获得更高的速率了

续表

速度	支撑体	类型	模式		注 释	
10 Mb/s	同轴电缆	10BASE5 厚以太网	HD		50 $\Omega$ 同轴电缆上的 10 Mb/s, 最大长度 500 m。允许使用 4 个中继器, 从而达到 2.5 km 的最大覆盖范围。开始的时候广泛使用, 现在不再使用, 因为它相对难以部署 (10 mm 直径, 两个插头之间有 2.5 m)。每个段的最大节点数量为 100。 更多细节参见 3.5 节	
		10BASE2 瘦以太网	HD		50 $\Omega$ 同轴电缆上的 10 Mb/s, 但直径可以是 5 mm, 有利于安装。布线要比 10BASE5 容易得多。比如一个段可以只有 185 m, 节点间的距离 0.5 m, 每个段只有 30 个节点 更多细节参见 3.6 节	
	电视	10BROAD36	HD		在用于电视的 75 $\Omega$ 电缆上的 10 Mb/s。数据编码是为宽带准备的, 以便既传输以太网数据也传输其他信号。这种类型的支撑体在特殊情况下使用	
	F0		10BASE-F	1 0 B A S E - F B		在集线器和中继器之间同步传输 (见 3.8 节)
				1 0 B A S E - F L	HD FD	围绕着主动设备的异步传输 (集线器)
				1 0 B A S E - F P		围绕着被动设备异步传输 (光纤星型)
	PT	10BASE-T	HD FD		在双绞电话线上的 10 Mb/s 更多细节参见 3.7 节	

续表

速度	支撑体	类型	模 式		注 释
100 Mb/s	PT	100BASE-T	1 0 0 B A S E - T 2		两条双绞线，第3类，以脉冲码以频率和幅度对信息编码。这一解决方案价格昂贵，不怎么用
			1 0 0 B A S E - T 4	HD	4条无屏蔽双绞线，第3类、第4类和第5类
			1 0 0 B A S E - X	HD FD	两条无屏蔽双绞线，第5类，或者一条屏蔽双绞线
			1 0 0 B A S E - F X	HD FD	两条光纤
	FO				
1000 Mb/s	FO	1000BASE-X	1 0 0 B A S E - L X	HD FD	波长很大的光纤

续表

速度	支撑体	类型	模式	注 释
	PT		1 0 0 0 B A S E - S X	短波长光纤
			1 0 0 0 B A S E - C X	屏蔽双绞线（见 3.8.3 节）
		1000BASE-TX	HD FD	4 条双绞线，第 5 类

PT：双绞线；FO：光纤；HD：半双工；FD：全双工

开始的时候网络定义在同轴电缆之上（见图 3.1）。在这一模式中所有设备共享同一个物理支撑。在给定的瞬间必须使用互斥机制以便只有一个站发送数据。结果是，当一个站发送数据时，其他站就不能发送。这种功能模式称为半双工（half-duplex）。

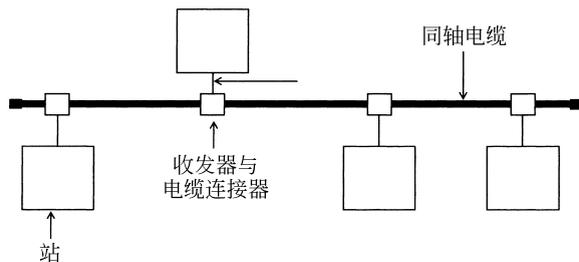


图 3.1 一个简单网络的布线

网络拓扑越来越朝向使用点对点连接的星形架构演化。在一个点对点链接中不可能有访问冲突。每个发射方向都有一条双绞线。第一个互连设备（集线器）只是将接收到的信号重复发送给其他端口。星形拓扑模拟了总线拓扑（见图 3.2）。访问冲突必须由终端来裁决。

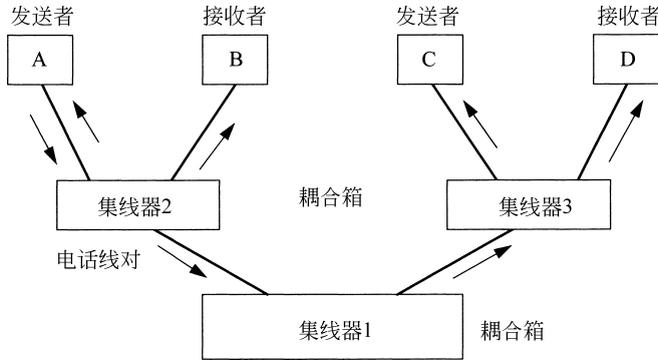


图 3.2 星形拓扑示例

星形拓扑是另一个演化的开始点。访问模式可以修改，以便让双方可以同时收发数据。这种模式称为**全双工 (full-duplex)**，而互连设备称为**交换机 (switch)**。

功能模式（半双工或者全双工）不仅与网络拓扑有关，也与终端所实现的功能以及互连设备有关。在标准中，所有的支撑体都可以半双工模式工作，但随着朝向 100 或者 1000 Mb/s 速率的演化，与访问模式有关的缺点意味着**全双工模式**是最常使用的模式。

### 3.2.2 接口和连接器

IEEE 802.3 标准定义了数个接口，如图 3.3 所示。这张图看起来有些复杂，而且接口的数量非常大。实际上，不是所有设备都具备全部这些接口。它们的角色是为了让传送和交换功能独立。

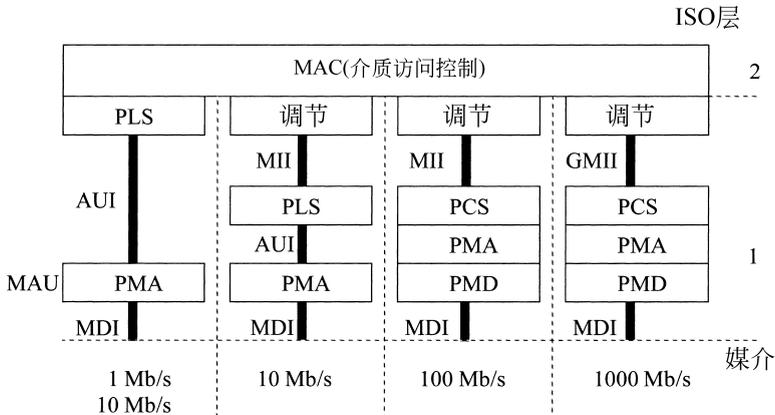


图 3.3 IEEE 802.3 协议栈中的接口

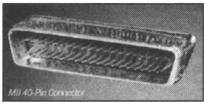
AUI (attachment unit interface, 连接单元接口) 是最著名、最古老的接口，其可视部分是一个 15 针的母连接器，我们可在一些网络板卡上见到。这个接口让物理层与所用的媒介（同轴电缆、电话线对、光纤等）无关。这一技术所用的收发器 (transceiver) 也称为 MAU (medium attachment unit, 媒介连接单元)，或者，所有解决方案所用的总称：PMA (physical media attachment, 物理介质连接)。收发器做信号转换，让它们适合于物理支撑网络。除了对



DB15公插头

传输支撑体的适应以外，收发器还控制站的发射时长。实际上，如果一个站在出现软件或硬件错误后开始在共享的支撑体上永久地向所有其他站发送数据，就会阻塞传输。**Jabber** 机制授权发射时长在 20~150ms 之间。MDI (media dependent device, 介质依赖设备) 接口是用于将收发器连接到发送支撑的插头的总称。它可以是 10BASE2 网络中的 BNC 插头或者双绞线网络的 RJ-45 插头，或者光纤连接器。

子层的定义随着不同的标准版本而演化。图 3.3 展示了其他协议栈。AUI 接口已经不再用在集线器和交换机上了。这是因为在机械上难以用这种插头连接设备。此外，设备和收发器之间的通信使用线对完成，每个发射方向都有，并使用曼彻斯特编码，这限制了向更高传送速率的演化。



MII公插头

为了定义 MII (media independent interface, 介质独立接口) 加入了调节层 (Reconciliation Layer)。这个接口主要用在交换机或集线器上。如同便携计算机上的 PCMCIA 卡一样，这个接口给互连设备带来一些可伸缩性。它增加了一个能适应新的、不在最初计划中的传输支撑体的收发器。MII 插头有 40 针，设备和收发器之间的数据传送是按 4 位来完成的，所以速度可提高到 100 Mb/s。

这一接口已经被修改，成为 1 Gb/s 速度的网络。GMII (gigabit media independent interface, 千兆位介质独立接口) 完全与 MII 兼容。不同的是数据在收发器和设备之间是按位传输。

收发器架构分为 3 个子层：

- PCS (physical coding sublayer, 物理编码子层) 对要通过支撑体传送的信息进行编码。实际定义的子层有 4 个：100BASE-X、100BASE-T4、100BASE-T2 和 1000BASE-X。
- PMA 用于给定的介质类型。其功能是：数据传送、冲突侦测、时钟同步等。
- PMD (physical medium dependent, 物理媒介依赖) 就在 MDI 接口之前，用于为 MDI 和传送支撑体提供接口。



RJ-45 插头是在以太网网络中常见的接口。其电缆的两端各有一个 RJ-45 插头，如图 3.4 所示。

有些解决方案 (比如 100BASE-T4 或者 1000BASE-TX) 使用 4 个线对。它们有自动协商功能 (自适应) 来避免混淆，并找出最优化的功能模式 (见 3.7.5 节)。

将集线器与站连接起来的电缆是单排扣的。连接两个集线器的电缆必须将线对交叉。如果不这样的话，发射线对和接受线对就成一对一的了 (见图 3.4)。



1 T+ 白-桔黄  
2 T- 桔黄  
3 R+ 白-绿  
4  
5  
6 R- 绿  
7  
8

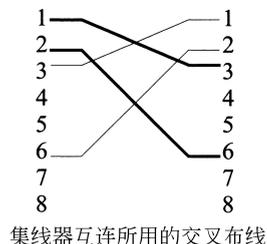


图 3.4 RJ-45 电缆插头

### 3.3 CSMA/CD 基础

在了解以太网设备的细节之前，理解 CSMA/CD 算法的工作原理是很有趣的，因为这一算法所基于的最大信号传播延迟会限制互连设备的数量并会造成需要相对严格的布线规则。

媒介访问协议依据两个基本原则：

- CSMA：这一方法减少冲突。在发送之前，站必须侦听信道。如果空闲，站可以发送其消息；否则，站延后发射。由于存在传播延迟，所以冲突的风险没有被压制，如图 3.5 所示。

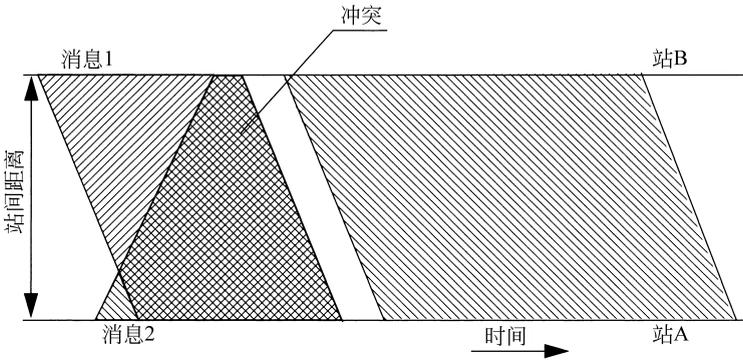


图 3.5 CSMA/CD 的原理

- CD (collision detect, 冲突侦测)：如果发送站意识到其消息卷入冲突中，它停止消息的发射。

站在想要发送消息的时候全局地发送消息。于是，当其他站不发送任何东西的时候，一个站可以处置整个信道。这种属性称为信道透明性 (transparence of channel)。鉴于这种信道随机访问算法，会有一些问题发生。冲突 (collision) 是指两个或更多站同时发送消息。冲突可由两个或多个设备同时发射而产生。

另一个问题是由于当站在等待信道空闲的时候产生的发射同步所引起的。图 3.6 展示了这一现象。站 A 和站 B 因为其他站传送另一个帧而造成延迟发射。当传送几乎完成时，传播延迟也几乎要结束时，两个站将同时开始发送，于是造成冲突。BEB 算法 (见 3.3.2 节) 可以将处于冲突中的设备隔开。

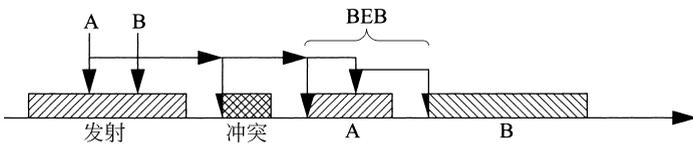


图 3.6 冲突

### 3.3.1 协议参数

时隙 (time slot) 是一个应用程序得知消息的正确传输所需的时间。这段时间的最小长度等于线路上消息传播的最大时长的两倍。这就证实了我们之前所见的布线约束。

考虑两个最远的站并且加上在信号传送一个来回时所能引入的所有延迟。计算的结果显示最大传播时长为  $44.99 \mu\text{s}$ 。对于一个经典的以太网而言, 其标准给出的值要稍微高一些。时隙的时长等于发射 512 个位的时长——以  $10 \text{ Mb/s}$  的话是  $51.2 \mu\text{s}$ 。

帧的发射时长总是大于或等于时隙。对于一个  $10 \text{ Mb/s}$  的网络而言,  $51.2 \mu\text{s}$  与 64 字节帧的发射时长相关。如果分组很小, 那么会引入一些填充 (padding) 位以便达到这一尺寸。这一最小时长的引入, 是为了让所有的站在传输之后处于相同的状态。

图 3.7 展示了发射时长比传播延迟小的协议。站 S1 发送消息 M1, 而站 S7 同时发送消息 M2。由于发射时长比传播时长小, 这些站在发射期间没有时间来发现冲突。它们将传送一个正确的传送给上一层。我们也可看到:

- 站 S2 已经正确接收帧 M1, 但没有正确接收帧 M2;
- 站 S6 已经正确接收帧 M2, 但没有正确接收帧 M1;
- 站 S3、S4 和 S5 没有正确接收到任何帧。

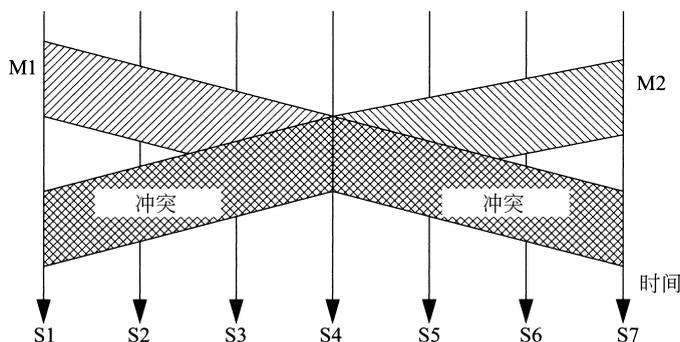


图 3.7 发射时长比时隙短

最大帧尺寸为 1518 字节 (1500 数据字节, 14 个头字节和 4 字节 CRC), 以避免一个站独占信道。这一尺寸的固定值是任意设定的, 不过更大的尺寸将增加冲突风险, 如图 3.6 所示。

当一个站侦测到冲突时, 它不会立即中断传输。而是发送干扰 (jamming) 数据来让其他站侦测到冲突。干扰数据的大小是 32 位。冲突帧的发射可持续不到一个时隙的时间。

如果不遵守布线规则, 冲突会在  $51.2 \mu\text{s}$  之后被侦测到。在网络中, 这种不正确的行为称为迟冲突 (late collision)。

### 3.3.2 BEB 算法

BEB (binary exponential backoff, 二进制指数退避<sup>1</sup>) 算法在冲突产生时限制网络负载。

<sup>1</sup> 在二进制指数之后重新传输。

在冲突期间，卷入冲突的站在一个时隙之后停止发射。站在冲突之后必须做什么没有定义。如果它们就在冲突之后重新开始发送数据，那么就会产生另一个冲突，如此这般。在支撑体上将不再有消息发射。有必要使用一种机制来分开各个站（不发送任何消息）。也有必要在站之间做一些计划来限制在拥塞时对物理支撑体的访问。冲突解决方案和新的分组也可造成新的冲突，这会限制可用带宽。这一问题可能重复，状况会被放大。

BEB 算法将抽出做出下一个传输尝试之前的等待时长。两个选择是 0 或 1。如果只有两个站参与冲突解决方案，那么会有 4 种可能性：

- 第一个站和第二个站抽出 0：两个站将在冲突之后重启传送，而冲突将再次发生。
- 第一个站抽出 0 而第二个站抽出 1：第一个站将在冲突之后开始传送数据而第二个站则等待一个时隙。于是当后者侦测到媒介上有活动时，它将等待来自第一个站的消息结束，然后再发送自己的消息。冲突解决。
- 第一个站抽出 1 而第二个站抽出 0。这种情况和上一个一样。避免了冲突。
- 两个站抽出 1：两个站都将等待一个时隙然后同时再次发送消息，造成另一个冲突。

这个示例展示，解决冲突有一半的机会。由于有没有解决方案的地方（两个站都抽出同一个数字或者有超过两个站卷入冲突），于是站将可抽空间扩展一倍。于是站可以获得 0、1、2、3 或者 4 个时隙。这将把两个站同时发送的几率减少到 1/4。站在发送消息之前将等待更长时间，这将减少网络的负载。

默认地，可抽空间在第 10 次尝试之后扩展一倍。如果在第 16 次尝试之后帧还没发送出，那么协议中止，上一层得到失败通知。

### 1. 发射算法

图 3.8 总结了在上述所有一切与帧发射有关的内容。

当数据被层的发送实体所考虑时，MAC 层加入目的地址和源地址字段而且计算 CRC。MAC 侦听支撑体，以确定当前是否有传送存在（CSMA 算法部分）。当支撑体上没有活动时，它开始在 9.6 μs 之后发送，这与帧间静默（inter-frame silence, IFS）相关。它继续侦听支撑体，侦测是否有冲突（冲突侦测算法部分）。如果没有侦测到冲突，MAC 层向上一层发送正面确认。

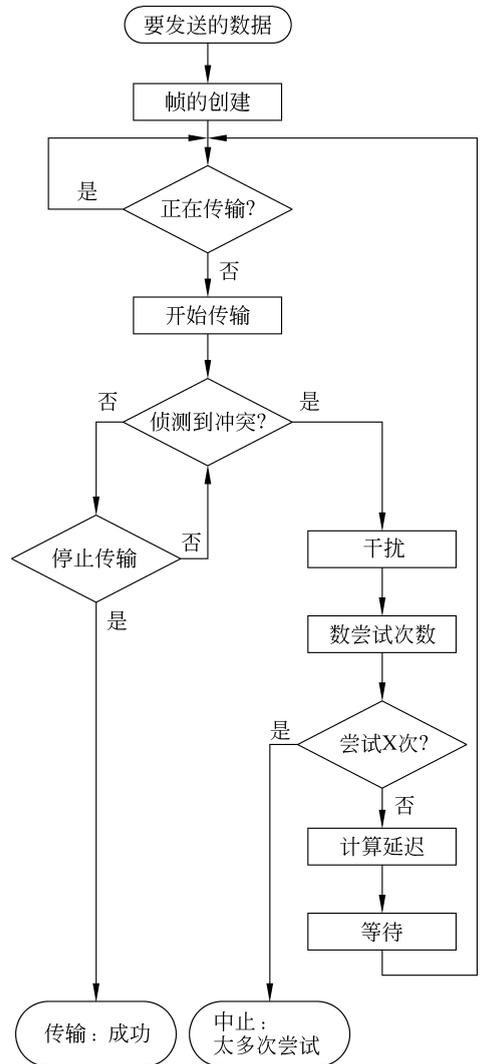


图 3.8 CSMA/CD + BEB 发射算法

如果侦测到冲突，那么 MAC 层在干扰期继续发送。它对试探性的尝试次数进行计数（BEB 算法）。如果尝试次数显著（16 次失败）则中止尝试并向上一层发送一个负面指示。如果不是，则根据尝试次数抽出每次间隔的时隙数，然后在此期间等待，并再次侦听支撑体，开始发送。

## 2. 接收算法

当一个站侦测到总线上的活动时，它就获取数据并分析帧。如果帧很小，则可能是冲突帧，将其丢弃（见图 3.9）。

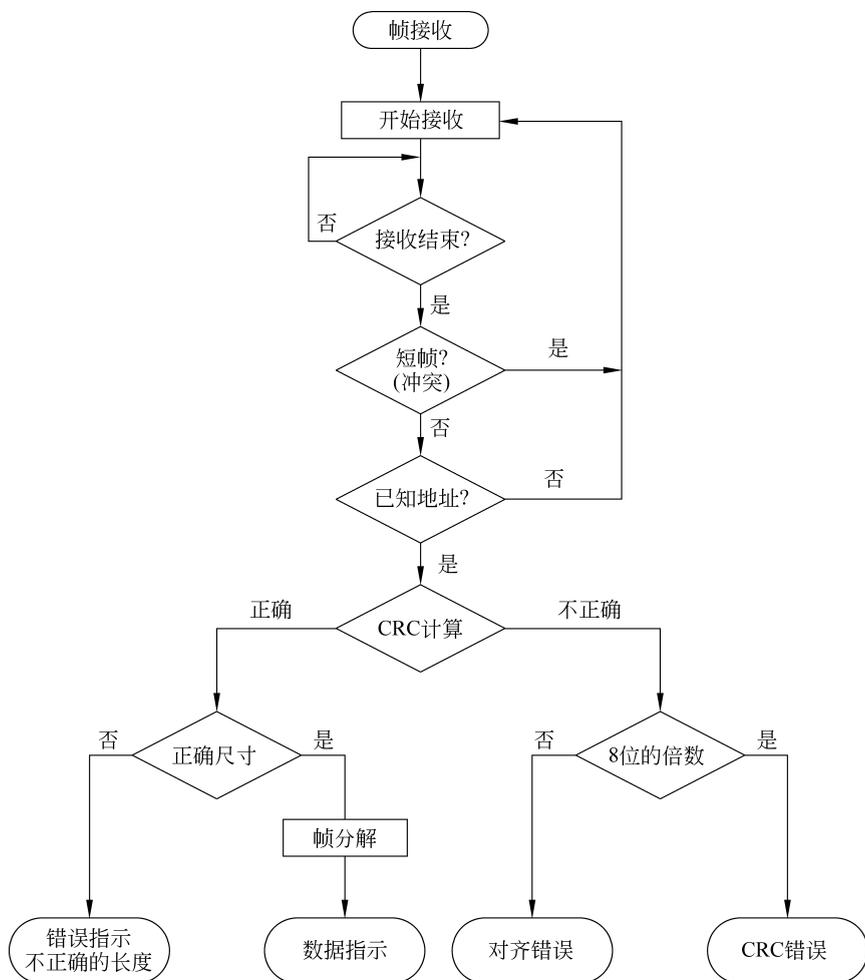


图 3.9 帧接收算法

这也是目的地址与存储地址（站地址、多播（multicast）或者广播（broadcast）地址）不相关时的情况。在地址正确的情况下，CRC 会被验证。如果正确，则把帧发送给上一层。如果校验和错误，那么会计算长度帧。如果它不是 8 位的倍数，那么它是个整数字节数，会给出对齐错误。如果不是，会将校验和错误发送给上一层。

在以太网协议里，帧的协议字段中的标识符指定协议实体。在 IEEE 802.3 中，数据会被发送给 LLC 层。

### 3.3.3 CSMA/CD 算法的限制

CSMA 访问方法在总线拓扑的局域网中工作得相对好一些。冲突侦测不适用于所有以广播扩散的局域网类型。于是：

- 对于无线网络（比如 IEEE 802.11），设备的发射强度会给接收造成干扰。要侦测冲突是不可能的。如果有障碍物或者发射范围不足，那么发送方将无法被其他消息所干扰，而接收方会被干扰。
- 对于电视网络（比如 IEEE 802.14），传播延迟太大。

### 3.3.4 中继器

中继器（见图 3.10）重新生成一个信号。在此之前，位编码会因为信号在支撑体中的传输而衰弱。中继器读出这些位，然后将其复制到另一个支撑体中。中继器也增强由于冲突而产生的信号。这可以增加网络的作用范围。中继器对其他层而言是不可见的；它不要求任何特别的管理。中继器也可将一个物理支撑体连通到另一个支撑体，比如 10BASE5 到 10BASE2 或者同轴电缆到光纤。

两个站之间最多可以有 4 个中继器。无须将中继器放置在段的边缘。一个段可支持数个中继器。对于双绞线或者光纤网络而言，集线器可当成中继器。

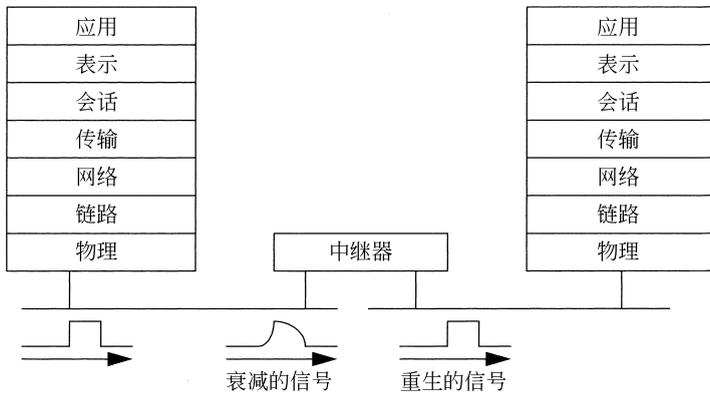


图 3.10 放置在参考模型中的中继器

## 3.4 帧格式

### 3.4.1 物理层

IEEE 802.3 的物理层以 7 个字节的导码 (preamble) 作为帧发射的开始 (参见图 3.11)。导码是值为 1010 1010 的二进制序列。这些字节的目的是在发送方和接收方之间同步时钟。由于网络站在没有东西可发送的时候网络上不会有通信量，所以此时没有理由让时钟保持

同步。当在帧传输过程中需要跨过数个中继器时，导码的第一位会被毁坏。中继器们无法恢复导码的头一位。于是，当需要进行数据重传时，它们必须重新构建完整的导码。

选择 1010 1010 的导码值是为了用曼彻斯特编码（10BASE5，10BASE2）给出方波，见 3.5.2 节。由于头几位可能丢失，所以传输必须使用经典的曼彻斯特编码，因为接收到的第一位不能作为参考来使用。

上一层的数据发送的开始位于 SDF（starting delimiter frame，启动定界符帧），值是 1010 1011。

接收方识别总线上连续的帧，因为它们由称为 IFS（inter frame spacing，帧间间隔）的 9.6 μs 静默来分隔。超市收银台的传送带是这一技术很好的阐释。每个顾客在他或她所购买的东西以及之前的顾客的东西之间留有一定距离，以便帮助收银员识别各个顾客的东西。IFS 让回声反射现象能够在新的传送开始之前在线缆边缘被吸收。这些回声反射现象会造成干扰，会被站解释成不适时的冲突<sup>1</sup>。

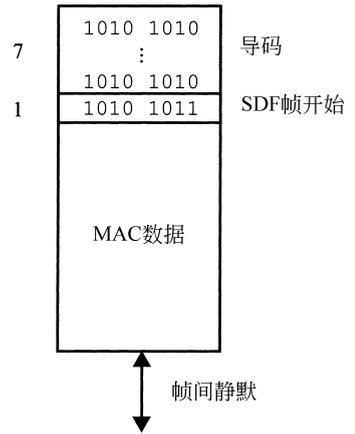


图 3.11 IEEE 802.3 物理层封装

## 3.4.2 MAC 层

### 1. IEEE 802.3 帧

IEEE 802.3 帧字段（见图 3.12）是：

- 目的地址符合第 2 章的 2.2 节所定义的格式。依据网络的部署，地址选择 2 个或者 6 个字节，但在实践中这一选择一直都是 6 个字节。

这是在传送开始、接收的时候第一个要识别的帧字段，用于确定该帧是否是给它们的。

- 源地址表示发送出帧的站。尺寸必须和目的地址的大小一样。
- 数据长度表示以字节计数的数据尺寸。MAC 层的一个帧的最小发射时间必须是 51.2 μs（原因之前已经解释了）。这个时长会根据想要发送的数据尺寸不同而改变。如果这一尺寸不符合最小时长要求，就需要加入填充来得到这个值。数据长度字段的重要性在于，接收方需要用它来区分数据和填充。
- 数据必须符合 LCC 标准定义的格式（见第 5 章）。
- CRC 按下列多项式求出：

$$x^{32} + x^{26} + x^{23} + x^{22} + x^{16} + x^{12} + x^{11} + x^{10} + x^8 + x^7 + x^5 + x^4 + x^2 + x + 1$$

### 2. 以太网帧

MAC 地址总是 6 个字节。以太网帧格式（见图 3.13）与 IEEE 802.3 帧在第三个字段上不同。这个帧具有的是上一层协议的标识符而不是数据长度。许多要点从这一改变推断出来：

<sup>1</sup> 我们将在 3.4.3 节中看到，以太网的演化使用了其他在超市中部署的“下一顾客”指示牌策略。其惯例是，这个指示牌在商店中不卖；也可使用不用来代表二进制数据的代码来指示帧的开始和结束。

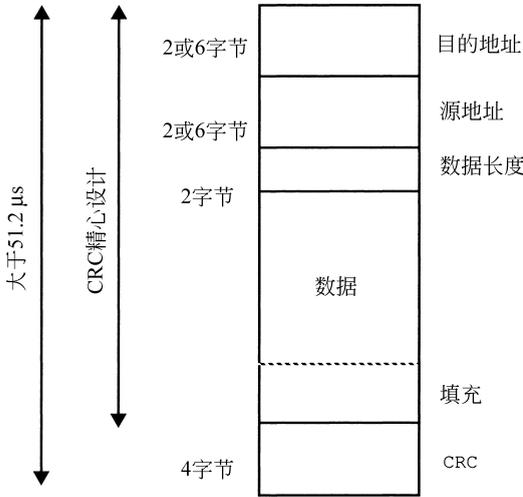


图 3.12 在 MAC 层的 IEEE 802.3 帧封装

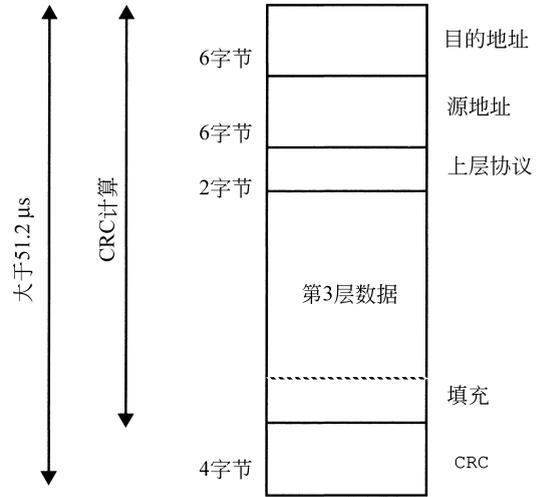


图 3.13 在 MAC 层的以太网帧封装

- LLC 层不存在。LLC 层的功能是实现对上层协议的切换。在这里 MAC 层直接进行切换。
- MAC 层不能省略由发送方在 MAC 层引入的填充位。这里与层架构的功能有一点违背，因为第 2 层数据将进入目的设备的上一层。
- 第 3 层协议必须具备去除这些填充位的方法，如果这些位在帧中存在的话（比如，使用分组的数据长度字段）。

IEEE 802.3 帧和以太网帧之间的区别位于头部中的第三字段。数据尺寸限制在 1500 字节。对于以太网帧，要使用更大的值来对协议编码。

在 Internet Assigned Numbers Authority<sup>1</sup> Web 服务器上，可以找到协议字段的一些值（表 3.2 给出了这些值中的一部分）。带有灰色底纹的值是可在 Internet 网络中找到的值（见第 7 章以及随后的内容）。

表 3.2 以太网帧中的协议字段值

协议字段（十进制）	协议字段（十六进制）	协议
2048	0x0800	DOD IP (Internet)
2049	0x0801	X.75 Internet
2050	0x0802	NBS Internet
2051	0x0803	ECMA Internet
2052	0x0804	Chaosnet
2054	0x0806	ARP
2055	0x0807	XNS 兼容性
32923	0x809B	Appletalk
32981	0x80D5	以太网上的 IBM SNA 服务
33011	0x80F3	AppleTalk AARP (Kinetics)

1 <http://www.iana.org/assignments/ethernet-numbers>。

续表

协议字段（十进制）	协议字段（十六进制）	协议
33024	0x8100	虚拟网络标签
34525	0x86dd	IPv6
34824	0x8808	Message PAUSE
34887	0x8847	MPLS 点对点
34888	0x8848	MPLS 多点
34915	0x8863	PPPoE 通信发现
34916	0x8864	PPPoE 信息运送
34958	0x888E	身份验证 (IEEE 802.1X)

## 3.5 10BASE5 网络

### 3.5.1 设备

以太网 10BASE5 或者“厚以太网”的同轴电缆的组成如下：

- 一条直径为 2.17 mm 的中心半导体。
- 一个（绝缘）介质，它可很容易地钻入，以方便安装收发器而无须修改电缆的特性并避免在屏蔽层和中心半导体之间短路的风险（参见图 3.14）。
- 编织体屏蔽层，保证数字信号不受外界电磁场的干扰。
- PVC 护套，符合实际的防火标准，它是一条颜色鲜艳的电缆（比如黄色）。每 2.5 m 要有一个可见的标志。这一距离与两个收发器之间的最小距离有关。

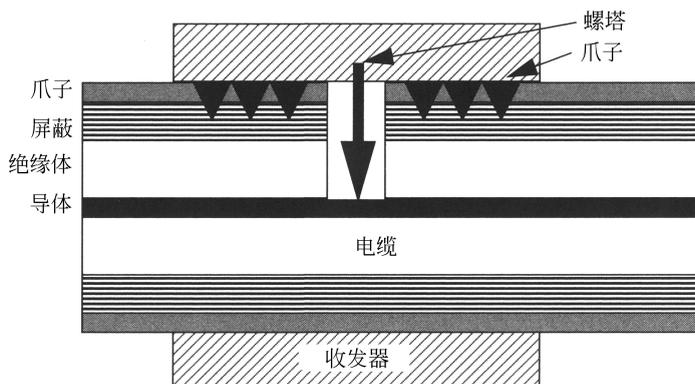
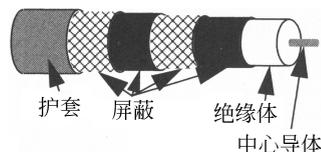


图 3.14 在 10BASE5 电缆上安装收发器

一个段（segment）是在两端都有阻抗器（用于限制回声反射现象）的同轴电缆。它的最长长度应为 500 m（见图 3.15）。

段发送和接收经编码的信号以及站和网络的电绝缘。收发器必须至少相隔 2.5 m。一个

段中不能有超过 100 个收发器。

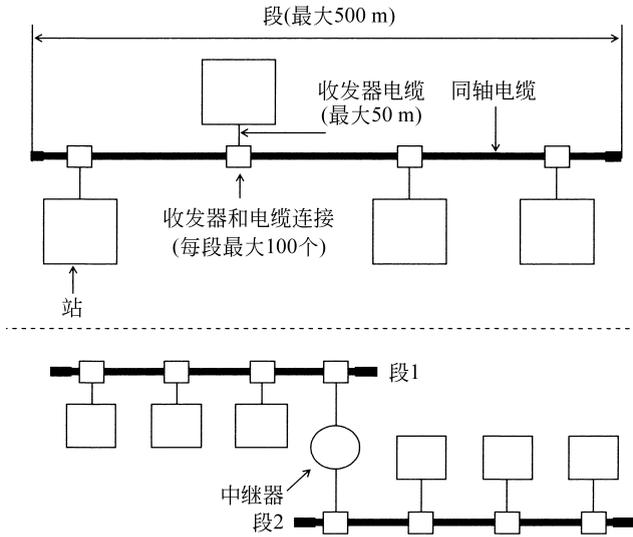
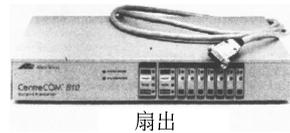


图 3.15 简单的网络布线



带有“吸血鬼式分接头”的收发器



扇出

为了安装收发器，要使用带有一个挡块的工具给电缆打孔。收发器盒子牢牢固定住电缆。盒子的爪会抓紧屏蔽层。螺塔旋入孔中，与中心导体连接（见图 3.14）。

一条电缆由各自带有金属屏蔽的一对线组成，外面可由抓牢收发器的屏蔽层包裹。从收发器到站的电缆最大长度为 50 m。

收发器之间最小距离 2.5 m 的要求会限制网络的伸缩性。设备可能位于同一间房间中，而布线则是在天花板上进行的。为了符合最小距离的要求，就必须在离站很远的地方做连接，或者要求使用很长的电缆。

扇出（fan out）或者访问复用器在一个收发器上通过 AUI 电缆连接许多用户（见图 3.16）。这里给出的模型可支持 8 个直接连接的用户，并且通过级联扇出，可在半径 150 m 内连接 64 个用户。

### 3.5.2 曼彻斯特编码

一旦流率变得越来越高，就值得频繁地在所发送的信号中引入转换。两阶段编码（或者曼彻斯特编码）使用这一原理强制在每个二进制周期的中间进行转换（跳变）。位中间的

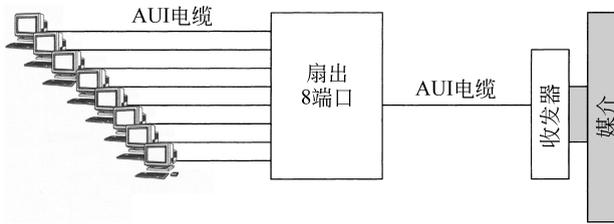


图 3.16 扇出连接

转换用于同步时钟。

在正常曼彻斯特编码中（见图 3.17），从高电平到低电平的转换代表位码为 0。从低电平到高电平的转换代表位码为 1。以太网使用这种类型的编码。

在令牌环网使用的差分曼彻斯特编码中（见图 3.18），如果要编码的位的值是 0，要在位时间开始的时候引入跳变（另外一个跳变在位的半个周期之后引入）。

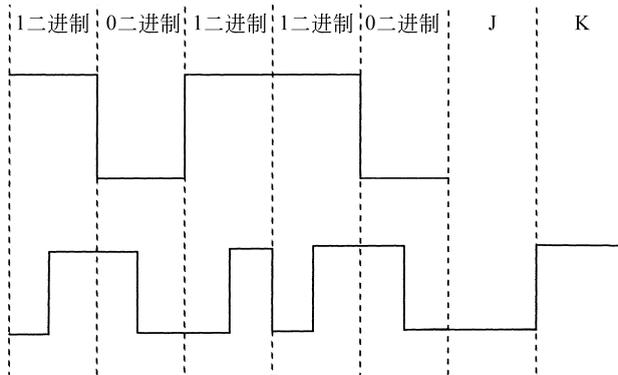


图 3.17 曼彻斯特编码

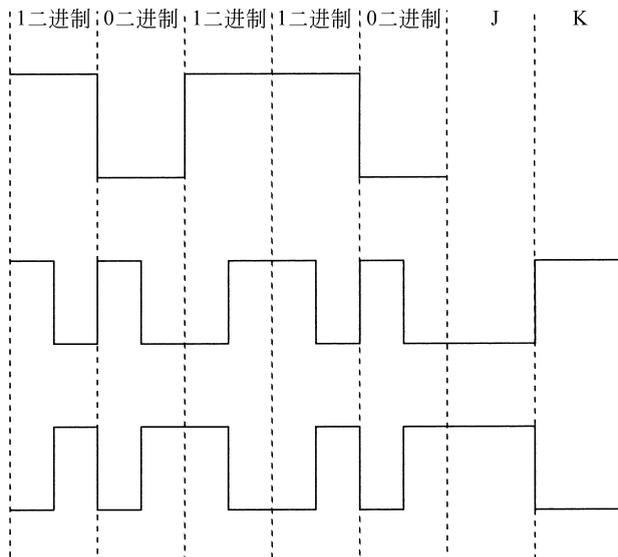


图 3.18 差分曼彻斯特编码

当位的值是 1 时，在位时间开始的时候没有跳变。有必要在传输开始的时候引入一个参考位来修正信号极性。令牌环网络使用差分曼彻斯特编码。

J 和 K 符号打破了上述规则，它们在两个半周期内保持相同状态。

曼彻斯特编码的优点如下：

- 为同步接收器和发送器的时钟而引入了转换；
- 传输电缆可用于为网络中的设备传输电力（以及在 IEEE 802.5 中使用的幻像电源）；
- 当出现全 1 或者全 0 的情况时，不会有连续的耗电组件让电子组件负荷过重；
- J 和 K 两个符号的存在授予了帧信号动作；
- 差分编码可忽略信号极性（如果使用转换器对站进行电隔离时需要如此）；
- 如果超过一个周期没有侦测到极性更改，很容易就能侦测出传送错误。

曼彻斯特编码的主要缺点是它占用更多带宽，因为编码的频率调制（以波特为单位）是传输率的两倍（以 b/s 为单位）。诸如 100 Mb/s 以太网这样的网络使用 4B/5B 编码类型，在支撑体上使用 5 个位对 4 位有用信息编码。这样，每 4 个或 5 个发送的位就至少引入一次转换。按照这种分配方式，曼彻斯特编码会是 1B/2B。

缺少布线灵活性加上其特异性使得这种类型的布线最不常用。

### 3.6 10BASE2 的设备

10BASE2 或者瘦以太网或廉价网（thin Ethernet or Cheapernet）的布线使用瘦同轴电缆，用两个 BNC 压接连接器作为两端。段的长度限制为 185 m，由数个通过 T 连接器互连的电缆组成。两个连接器之间的最小空间为 0.5 m，最大可在电缆上连接 30 个收发器。段的两端必须接上 50  $\Omega$  阻抗器（见图 3.19）。



图 3.19 以太网连接器

在支撑体所用的数据编码方式和 10BASE5 一样，是曼彻斯特编码。

所有的站必须使用 T 连接器连接，和边缘的站一样。常见的错误是直接将最后一个站连接到 BNC 连接器上：在插头上有 50  $\Omega$  的设备非常少见。

这种轻量级的连接是 10BASE2 的一个弱点。如果用户意外断开一个插头，则段会被分成两半，每一半都需要阻抗器。除非加上阻抗器，否则就不可能在两个段上进行任何通信。这种类型的断连接非常难以定位，虽然在使用分析器的时候不难看到：在这种类型的布线中，网络工程师将测试网络中的每个插头，每个插头放置一个盖子来定位连接断开位置。

10BASE2 网络越来越少使用双绞线布线，这是因为它糟糕的可靠性、布线规范及其发展的停滞不前。

## 3.7 双绞线设备

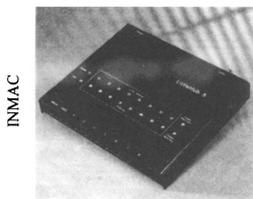
BASE-T (T 代表双绞线——Twisted pair) 并不构建在总线技术基础之上。它是个星形拓扑, 需要能模拟总线的活动设备 (集线器)。最初这种网络类型在标准电话线对上以 10 Mb/s 的速率工作。于是就能利用已经安装好的电话布线 (所以成本低)。在布线中使用了层次星团 (hierarchical starry cluster) (见图 3.2)。

这种类型的布线和第 2 章开始时给出的规则相去甚远, 这个现象很有趣。实际上, 网络功能是否正常将依赖于活动的中央元素是否正常, 虽然要挑战电子设备的可靠性要比偶然的总线断开连接更难。这一拓扑用得越来越多, 因为它利用了现有的布线 (电话、令牌环等)。高速网络协议将以如星星般分布的系统为基础 (100 Mb/s 以太网、ATM 等)。活动设备也可远程管理 (端口激活/禁用、配置、统计信息收集等)。

一般在布线的时候要使用 4 条双绞线。在将主机连接到集线器时只需要两条双绞线: 一条用于发射一条用于接收。另外两条不用。它们可用于更高速度的传输。这一最初用于模拟 10 Mb/s 总线的拓扑, 已经让标准演化出 3 个互补的方法:

- 增加速率, 提升到 100 Mb/s。增加的速率使用适合于这些速率的五类线 (100BASE-TX) 并且余地更多地对适应的方法进行编码 (100BASE-T2) 或者使用半双工的 4 条双绞线 (100BASE-T4)。
- 将集线器中的总线编码方法替换为能够允许同时传输数个帧的通信机制。
- 虚拟网络的创建, 真正隔开同一个基础设施上的通信量 (见第 13 章)。

### 3.7.1 集线器



集线器让从一条双绞线上接收来的信息经过集中后重传到其他双绞线上。当两个站同时传送的时候, 集线器生成冲突信号发送给所有设备。对于站而言, 访问支撑体的方法和访问总线的方法一样。可将集线器当成一个带有许多输入输出端口的中继器。

集线器使用曼彻斯特编码, 和 10BASE5 以及 10BASE2 一样。

### 3.7.2 交换机

交换机直接继承了电话标准中使用的一些技术以及一些对 ATM 上的高速网络的研究成果。可以将交换机看成是带有同样数量的输入输出端口的网桥。交换机读取来自输入端口的消息, 按照帧中的目的地址将其复制到输出端口。在交换机内, 电子器件和处理器快速建立交换矩阵, 以便在同一时间处理数个帧。

这一方法避免了冲突, 因为它不将消息在所有端口上发送。如果我们想保留站中的相同设备, 交换机也必须将数据复制给发送者, 以宣布有冲突产生。

理论上, 有一半的站可同时与其他另一半通信, 让网络的总带宽翻倍。每个站都能保

证有 10 Mb/s 的速率。图 3.20 代表了理想的情况：网络中的每个站都与另一个站关联。在这个示例中，网络上的总带宽是 50 Mb/s。

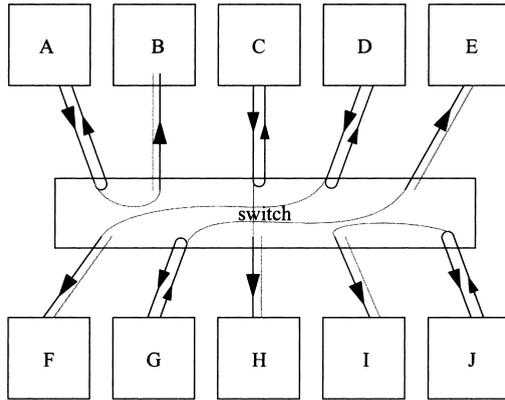


图 3.20 交换机的完美使用

在这个示例中，没有对站做修改，它们工作在半双工模式下。无论如何，通过允许同时发射和接收，有可能增加站的性能。在这种情况下，当站 A 给站 B 发送一个帧的同时，也可以从站 C 接收一个帧。

在实际中，通信量从来不会如此良好地在站之间传递。图 3.21 展示了一个以太网段上 MAC 层的通信量矩阵。在这张图中，站以圆圈形式放置。将每个站连接起来的线代表了对话的建立。对话主要发生在站 rsm（文件服务器）、nintendo & msg-rsm（连接其他段的路由器）中，两个站之间的对话非常不寻常。

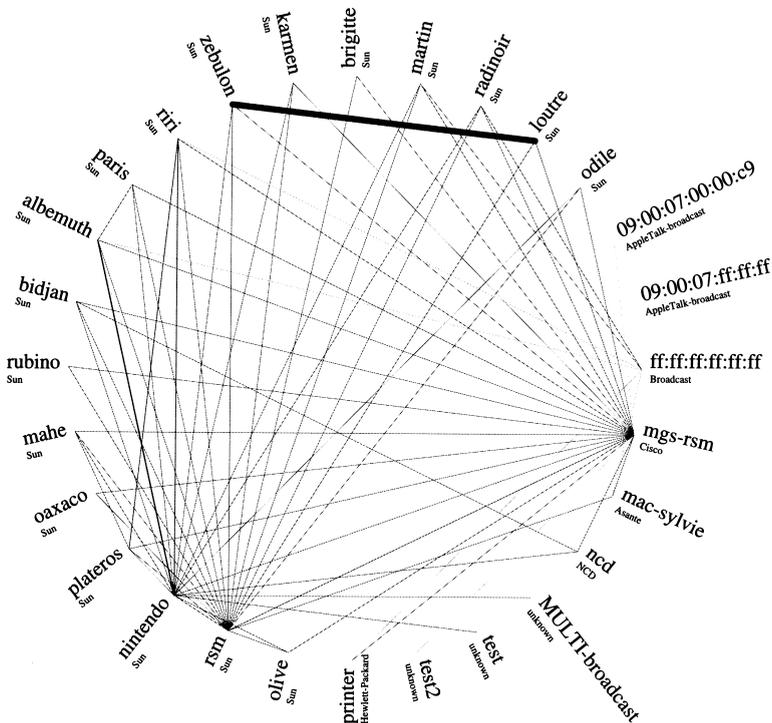


图 3.21 以太网段的流量矩阵