第 4 章　数据链路层

本章基本要求：掌握数据链路层的功能与作用，掌握差错控制的作用和原理，掌握数据链路层的设备与组件；理解常用的成帧方式，理解流量控制的作用和原理，理解 HDLC 协议的主要内容。

4.1　数据链路层功能

数据链路层是 OSI 参考模型中的第 2 层，在物理层提供服务的基础上向网络层提供服务。数据链路层为物理链路上提供可靠的数据传输。数据链路层的主要功能包括帧同步、差错控制、流量控制、链路管理、寻址等。

4.1.1　相邻结点

所谓相邻结点是指由同一物理链路连接的所有结点。相邻结点的最主要特征是结点之间的数据通信不需要经过其他交换设备的转发。

为实现相邻结点之间的可靠传输，数据链路层必须解决以下问题：在相邻的结点之间确定一个接收目标，即实现物理寻址；提供一种机制使得接收方能识别数据流的开始与结束；提供相应的差错检测与控制机制以使有差错的物理链路对网络层表现为一条无差错的数据链路；提供流量控制机制以保证源和目标之间不会因发送和接收速率不匹配而引起数据丢失。

4.1.2　帧同步

数据链路层采用了被称为帧（Frame）的协议数据单元作为数据链路层的数据传输逻辑单元。不同的数据链路层协议的核心任务就是根据所要实现的数据链路层功能来规定帧的格式。

1．帧的基本格式

尽管不同的数据链路层协议给出的帧格式都存在一定的差异，但它们的基本格式还是大同小异的。图 4.1 给出了帧的基本格式，组成帧的那些具有特定意义的部分被称为域或字段（Field）。

图 4.1　帧的基本格式

其中，帧开始字段和帧结束字段分别用以指示帧或数据流的开始和结束。地址字段给出结点的物理地址信息，物理地址可以是局域网网卡地址，也可以是广域网中的数据

链路标识，地址字段用于设备或机器的物理寻址。第 3 个字段则提供有关帧的长度或类型的信息，也可能是其他一些控制信息。数据字段承载的是来自高层即网络层的数据分组（Packet）。帧检验序列（FCS，Frame Check Sequence）字段提供与差错检测有关的信息。通常数据字段之前的所有字段被统称为帧头部分，而数据字段之后的所有字段被称为帧尾部分。

2．成帧与拆帧

引入帧机制不仅可以实现相邻结点之间的可靠传输，还有助于提高数据传输的效率。例如，若发现接收到的某一个（或几个）比特出错时，可以只对相应的帧进行特殊处理（如请求重发等），而不需要对其他未出错的帧进行这种处理；如果发现某一帧被丢失，也只要请求发送方重传所丢失的帧，从而大大提高了数据处理和传输的效率。但是，引入帧机制后，发送方的数据链路层必须提供将从网络层接收的分组（Packet）封装成帧的功能，即为来自上层的分组加上必要的帧头和帧尾部分，通常称此为成帧（Framing）；而接收方数据链路层则必须提供将帧重新拆装成分组的拆帧功能，即去掉发送端数据链路层所加的帧头和帧尾部分，从中分离出网络层所需的分组。在成帧过程中，如果上层的分组大小超出下层帧的大小限制，则上层的分组还要被划分成若干个帧才能被传输。

发送端和接收端数据链路层所发生的帧发送和接收过程大致如下：发送端的数据链路层接收到网络层的发送请求之后，便从网络层与数据链路层之间的接口处取下待发送的分组，并封装成帧，然后经过其下层物理层送入传输信道；这样不断地将帧送入传输信道就形成了连续的比特流；接收端的数据链路层从来自其物理层的比特流中识别出一个一个的独立帧，然后利用帧中的 FCS 字段对每一个帧进行校验，判断是否有错误。如果有错误，就采取收发双方约定的差错控制方法进行处理；如果没有错误，就对帧实施拆封，并将其中的数据部分即分组通过数据链路层与网络层之间的接口上交给网络层，从而完成了相邻结点的数据链路层关于该帧的传输任务。

3．帧的定界

帧定界就是标识帧的开始与结束。有 4 种常见的定界方法，即字符计数法、带字符填充的首尾界符法、带位填充的首尾标志法和物理层编码违例法。

（1）字符计数法

字符计数法是在帧头部中使用一个字符计数字段来标明帧内字符数。接收端根据这个计数值来确定该帧的结束位置和下一帧的开始位置。

（2）带字符填充的首尾界符法

带字符填充的首尾界符法是在每一帧的开头用 ASCII 字符 DLE STX，在帧末尾用 ASCII 字符 DLE ETX。但是，如果在帧的数据部分也出现了 DLE STX 或 DLE ETX，那么接收端就会错误判断帧边界。为了不影响接收方对帧边界的正确判断，采用了填充字符 DLE 的方法。即如果发送方在帧的数据部分遇到 DLE，就在其前面再插入一个 DLE。这样数据部分的 DLE 就会成对出现。在接收方，若遇到两个连续的 DLE，则认为是数据部分，并删除一个 DLE。

（3）带位填充的首尾标志法

带位填充的首尾标志法一次只填充一个比特“ 0 ”而不是一个字符“ DLE ”。另外，带位填充的首尾标志法用一个特殊的位模式“ 01111110 ”作为帧的开始和结束标志，而不是分别用“ DLE STX ”和“ DLE ETX ”作为帧的首标志和帧的尾标志。

（4）物理层编码违例法

物理层编码违例法就是利用物理层信息编码中未用的电信号来作为帧的边界。

4.1.3 差错控制

所谓差错是指接收端收到的数据与发送端实际发出的数据出现不一致的现象。产生差错主要是因为在通信线路上噪声干扰的结果。根据噪声类型不同，可将差错分为随机错和突发错。热噪声所产生的差错称为随机错，冲击噪声（如电磁干扰、无线电干扰等）所产生的错误称为突发错。

差错的严重程度由误码率来衡量，误码率 P_e 等于错误接收的码元数与所接收的码元总数之比。显然，误码率越低，信道的传输质量越高，但是由于信道中的噪声是客观存在的，所以不管信道质量多高，都要进行差错控制。

1．差错控制的作用与机制

为了提高传输的准确性，采用了专门的校验错误方法，用来发现所产生的错误，并给出出现错误的信号或者校正错误。差错控制是采用可靠、有效的编码以减少或消除计算机通信系统中传输差错的方法，其目的在于提高传输质量。

为了有效地提高传输质量，一种方法是改善通信系统的物理性能，使误码的概率降低到满足要求的程度，但这种方法受经济和技术上的限制。另一种方法是差错控制，它是利用编码的手段将传输中产生的错码检测出来，并加以纠正。差错控制是数据通信中常用的方法。

差错控制的主要作用是通过发现数据传输中的错误，采取相应的措施减少数据传输错误。差错控制的核心是对传输的数据信息加上与其满足一定关系的冗余码，形成一个加强的、符合一定规律的发送序列。所加入的冗余码称为校验码（Frame Check Sequence，简称 FCS）。

校验码按功能的不同被分为纠错码和检错码。纠错码不仅能发现传输中的错误，还能利用纠错码中的信息自动纠正错误，其对应的差错控制措施为自动前向纠错。汉明码（Hamming code）为典型的纠错码，具有很高的纠错能力。检错码只能用来发现传输中的错误，但不能自动纠正所发现的错误，需要通过反馈重发来纠错。常见的检错码有奇偶校验码和循环冗余校验码。目前计算机网络通信中大多采用检错码方案。

2．常见检错码

（1）奇／偶校验码

奇／偶校验的规则是在原数据位后附加一个校验位，将其值置为“ 0 ”或“ 1 ”，使附加该位后的整个数据码中“ 1 ”的个数成为奇数或偶数。使用奇数个“ 1 ”进行校验的方案被称为奇校验；对应于偶数个“ 1 ”的校验方案被称为偶校验。奇／偶校验有 3 种使用方式，即水平奇／偶校验、垂直奇／偶校验和水平垂直奇／偶校验。下面以奇校验为例

进行介绍。

水平奇校验码是指在面向字符的数据传输中，在每个字符的 7 位信息码后附加一个校验位“ 0 ”或“ 1 ”，使整个字符中二进制位“ 1 ”的个数为奇数。

例如，设待传输字符的比特序列为“ 1100001 ”，则采用奇校验码后的比特序列形式为“ 11000010 ”。接收方在收到所传输的比特序列后，通过检查序列中的“ 1 ”的个数是否仍为奇数来判断传输是否发生了错误。若比特在传输过程中发生错误，就可能会出现“ 1 ”的个数不为奇数的情况。水平奇校验只能发现字符传输中的奇数位错，而不能发现偶数位错。例如上述发送序列“ 11000010 ”，若接收端收到“ 11001010 ”，则可以校验出错误，因为有一位“ 0 ”变成了“ 1 ”；但是若收到“ 11011010 ”，则不能识别出错误，因为有两位“ 0 ”变成了“ 1 ”。不难理解，水平偶校验也存在同样的问题。

为了提高奇／偶校验码的检错能力，引入了水平垂直奇／偶校验，即由水平奇／偶校验和垂直奇／偶校验综合构成。

垂直奇／偶校验也称为组校验，是将所发送的若干个字符组成字符组或字符块，形式上相当于是一个矩阵，如图 4.2 所示，每行为一个字符，每列为所有字符对应的相同位。在这一组字符的末尾即最后一行附加一个校验字符，该校验字符中的第 i 位分别对应组中所有字符第 i 位的校验位。显然，如果单独采用垂直奇／偶校验，则只能检出字符块中某一列中的一位或奇数位错。

但是，如果同时采用了水平奇／偶校验和垂直奇／偶校验，既对每个字符作水平校验，同时也对整个字符块作垂直校验，则奇／偶校验码的检错能力可以明显提高。这种方式的奇／偶校验被称为水平垂直奇／偶校验，图 4.3 给出了一个水平垂直奇／偶校验的例子。但是从总体上讲，奇／偶校验方法的检错能力仍较差，虽然其实现方法简单。故这种校验一般只用于通信质量要求较低的环境。

（2）循环冗余校验码（CRC）

字母	前 7 行为对应字母的 ASCII 码，最后一行是垂直奇校验编码（粗体）
a	1 1 0 0 0 0 1
B	1 1 0 0 0 1 0
c	1 1 0 0 0 1 1
d	1 1 0 0 1 0 0
e	1 1 0 0 1 0 1
f	1 1 0 0 1 1 0
g	1 1 0 0 1 1 1
校验位	**0 0 1 1 1 1 1**

图 4.2　垂直奇校验

字母	最后一行是垂直奇校验码，最后一列是水平奇校验编码（粗体）
a	1 1 0 0 0 0 1 **0**
B	1 1 0 0 0 1 0 **0**
c	1 1 0 0 0 1 1 **1**
d	1 1 0 0 1 0 0 **0**
e	1 1 0 0 1 0 1 **1**
f	1 1 0 0 1 1 0 **1**
g	1 1 0 0 1 1 1 **0**
校验位	**0 0 1 1 1 1 1 0**

图 4.3　水平垂直奇校验

循环冗余校验码（Cycle Redundancy Check，简称 CRC）是一种被广泛采用的多项式编码。CRC 码由两部分组成，前一部分是 k+1 个比特的待发送信息，后一部分是 r 个比特的冗余码。由于前一部分是实际要传输的内容，因此是固定不变的，CRC 码的产生关键在于后一部分冗余码的计算。

计算中主要用到两个多项式：$f(x)$和 $G(x)$。其中，$f(x)$是一个 k 阶多项式，其系数是待发送的 k+1 个比特序列；$G(x)$是一个 r 阶的生成多项式，由发收双方预先约定。

例如，设实际要发送的信息序列是 1010001101（10 个比特，k=9），则以它们作为 $f(x)$

的系数，得到对应的 9 阶多项式为

$$f(x)=1\times x^9+0\times x^8+1\times x^7+0\times x^6+0\times x^5+0\times x^4+1\times x^3+1\times x^2+0\times x+1$$
$$=x^9+x^7+x^3+x^2+1。$$

再假设发收双方预先约定了一个 5 阶（r=5）的生成多项式 $G(x)=x^5+x^4+x^2+1=1\times x^5+1\times x^4+0\times x^3+1\times x^2+0\times x+1$，则其系数序列为 110101。

CRC 码的产生方法如下。

① 生成 r 个比特的冗余码：用模 2 除法进行 $x^r f(x)/G(x)$运算，得余式 $R(x)$，其系数即是冗余码。

例如，$x^5 f(x)=x^{14}+x^{12}+x^8+x^7+x^5$，对应的二进制序列为 101000110100000，也就是 $f(x)$信息序列向左移动 r=5 位，低位补 0。

$x^5 f(x)/G(x)=(101000110100000)/(110101)$，得余数为 01110，也就是冗余码，对应的余式 $R(x)=0\times x^4+x^3+x^2+x+0\times x^0$（注意：若 $G(x)$为 r 阶，则 $R(x)$对应的比特序列长度为 r）。

注意，模 2 除法在做减法时不借位，相当于在进行异或运算。

② 得到带 CRC 校验的发送序列：用模 2 减法进行 $x^5 f(x)-R(x)$运算得到带 CRC 校验的发送序列，即 $x^5 f(x)-R(x)=101000110101110$。从形式上看，也就是简单地在原信息序列后面附加上冗余码。

在接收方，用同样的生成多项式 $G(x)$除所收到的序列。若余数为 0，则表示传输无差错，否则说明传输过程出现差错。例如，若收到的序列是 101000110101110，则用它除以同样的生成多项式 $G(x)=x^5+x^4+x^2+1$（即 110101）。因为所得余数为 0，所以收到的序列无差错。

CRC 校验方法是由多个数学公式、定理和推论得出的，尤其是 CRC 中的生成多项式对于 CRC 的检错能力会产生很大的影响。生成多项式 $G(x)$的结构及检错效果是在经过严格的数学分析和实验后才确定的，有其国际标准。常见的标准生成多项式如下。

$$\text{CRC-12：}G(x)=x^{12}+x^{11}+x^3+x^2+1$$
$$\text{CRC-16：}G(x)=x^{16}+x^{15}+x^2+1$$
$$\text{CRC-32：}G(x)=x^{32}+x^{26}+x^{23}+x^{22}+x^{16}+x^{12}+x^{11}+x^{10}+x^8+x^7+x^5+x^4+x^2+x+1$$

可以看出，只要选择足够的冗余位，就可以使得漏检率减少到任意小的程度。由于 CRC 码的检错能力强，且容易实现，因此是目前应用最广泛的检错码编码方法之一。CRC 码的生成和校验过程可以用软件或硬件方法来实现，如可以用移位寄存器和半加法器方便地实现。

3．反馈重发机制

由于检错码本身不提供自动的错误纠正能力，所以需要提供一种与之相配套的错误纠正机制，即反馈重发。通常当接收方检出错误的帧时，首先将该帧丢弃，然后给发送方反馈信息请求发送方重发相应的帧。反馈重发又被称为自动请求重传（ARQ，Automatic Repeat request）。反馈重发有两种常见的实现方法，即停止等待方式和连续 ARQ 方式。

4.1.4 流量控制

由于系统性能的不同，如硬件能力（包括 CPU、存储器等）和软件功能的差异，会

导致发送方与接收方处理数据的速度有所不同。若一个发送能力较强的发送方给一个接收能力相对较弱的接收方发送数据，则接收方会因无能力处理所有收到的帧而不得不丢弃一些帧。如果发送方持续高速地发送，则接收方最终还会被“淹没”。也就是说，在数据链路层只有差错控制机制还是不够的，它不能解决因发送方和接收方速率不匹配所造成的帧丢失问题。

为此，在数据链路层引入了流量控制机制。流量控制的作用就是使发送方所发出的数据流量不要超过接收方所能接收的速率。流量控制的关键是需要有一种信息反馈机制，使发送方能了解接收方是否具备足够的接收及处理能力。

虽然有各种不同的流量控制机制，但大部分已知流量控制方案的基本原理都是相同的。协议中包括了一些定义完整的规则，这些规则描述了发送方在什么时候发送下一帧，在未获得接收方直接或间接允许之前，禁止发送帧。例如，当一个连接建好后，接收方可以说：现在你可以给我发 n 个帧，但是此后，直到我告诉你继续时，你才能再发。如简单停止等协议就可以实现流量控制功能，但其实现效率太低。滑动窗口协议可以将确认机制与流量控制机制巧妙地结合在一起。

滑动窗口协议是指一种采用滑动窗口机制进行流量控制的方法。通过限制已经发送但还未得到确认的数据帧的数量，滑动窗口协议可以调整发送方的发送速度。许多使用位填充技术的数据链路层协议（如 HDLC 协议）都使用滑动窗口协议进行流量控制。

滑动窗口协议在提供流量控制机制的同时，还可以同时实现帧的确认和差错控制。正是滑动窗口协议这种集帧确认、差错控制、流量控制融为一体的良好特性才使得该协议不仅被广泛地应用于数据链路层中，还被作为传输层实现相应功能的重要机制。

4.1.5 链路管理

链路管理功能主要用于面向连接的服务。在链路两端的结点要进行通信前，必须首先确认对方已处于就绪状态，并交换一些必要的信息以对帧序号初始化，然后才能建立连接。在传输过程中则要维持该连接。如果出现差错，需要重新初始化，重新自动建立连接。传输完毕后则要释放连接。数据链路层连接的建立、维持和释放就称做链路管理。

4.2 数据链路层所提供的基本服务

通常，数据链路层有 3 种基本服务可供选择，即无确认的无连接服务（Unacknowledged Connectionless Service）、有确认的无连接服务（Acknowledged Connectionless Service）、有确认的面向连接服务（Acknowledged Connection-oriented Service）。

① 在无确认的无连接服务方式下，两个相邻机器之间在发送数据帧之前，事先不建立连接，事后也不存在释放连接。源机器向目标机器发送独立的数据帧，而目标机器不对收到的帧作确认。由于线路上的噪声而造成的帧丢失，数据链路层将不作努力去恢复，而是将该工作留给上层（通常为传输层）去完成。这类服务通常适用于误码率很低的信道，如大多数局域网都使用这种无确认的无连接服务方式。

② 在有确认的无连接服务方式下，仍然不需要建立连接，源机器向目标机器发送独立的数据帧，但是接收站点要对收到的每一帧作确认，即在收到数据帧之后回送一个确认帧，而发送站点在收到确认帧之后才会发送下一帧。当在一个确定的时间段内没有收

到确认帧时，发送方就认为所发送的数据帧丢失并自动重发此帧。自动重发可能会产生接收站点收到重复的数据帧的问题。有确认的无连接服务方式适用于像无线网之类的不可靠信道。

③ 在有确认的面向连接服务方式下，发送数据之前，需要首先建立连接，然后才会启动帧的传输。在发送数据阶段，为所传输的每一帧都要编上号，数据链路层提供相应的确认和流量控制机制来保证每一帧都只被正确接收一次，并保证所有帧都按正确的顺序被接收。当数据传输完成之后，还需要拆除或释放所建立的连接。也就是说，面向连接的服务方式分 3 个阶段：链路建立阶段、数据传输阶段和链路拆除阶段。可以这么说，只有有确认的面向连接的服务方式才真正为网络层提供了可靠的无差错传输服务。这类服务实现复杂度及代价很高，通常被用于误码率较高的不可靠信道，如某些广域网链路。

4.3 高级数据链路控制协议（HDLC）

高级数据链路控制（规程）（HDLC，High Level Data Link Control）是一个在同步网上传输数据、面向位的数据链路层协议，它是由国际标准化组织（ISO）制定的。HDLC 是 IBM 的同步数据链路控制规程（SDLC）的一个超集。

HDLC 是面向比特的协议，支持全双工通信，采用位填充的成帧技术，以滑动窗口协议进行流量控制。

4.3.1 HDLC 的帧格式

HDLC 的功能集中体现在 HDLC 帧格式中，HDLC 的帧格式如图 4.4 所示。

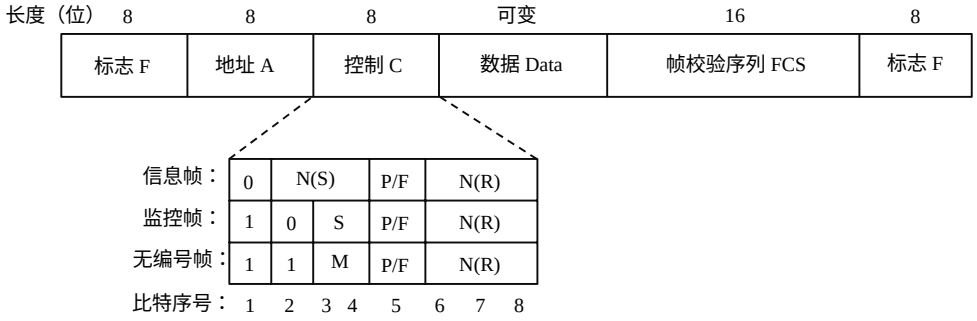

图 4.4 HDLC 帧格式及控制字段的结构

- 帧头和帧尾的位模式串“ 01111110 ”为帧的开始和结束标记（Flag）。可以看出，HDLC 协议在帧定界上采用的是带位填充技术首尾界符法。
- A 是地址字段（Address），由 8 位组成。对于命令帧，存放接收站的地址；对于响应帧，存放发送响应帧的站点地址。
- C 是控制字段（Control），由 8 位组成，该字段是 HDLC 协议的关键部分。它标志了 HDLC 的 3 种类型帧：信息（Information）帧、监控（Supervisory）帧和无序号（Unnumbered）帧。如图 4.4 中关于控制字段结构所示，若帧的第 1 比特为“ 0 ”，则代表这是一个用于发送数据的信息帧，相应地，其第 2 至第 4 比特代表当前发送的信息帧的序号，而第 6 至第 8 比特则代表接收序号即期望收到的帧的发送序

号。若帧的第 1 和第 2 比特为“ 10 ”，则代表这是一个用于协调双方通信状态的监控帧，相应地，其第 3 和第 4 比特用以代表 4 种不同类型的监控帧。“ 00 ”表示接收准备就绪；“ 01 ”表示传输出错，并要求采用拉回方式重发；“ 10 ”表示接收准备尚未就绪，要求发送方暂停发送；“ 11 ”则表示传输出错并要求采用选择重发。监控帧中不包含 Data（数据）部分，若帧的第 1 和第 2 比特为“ 11 ”，则代表用于数据链路控制的无序号帧，其第 3、4、6、7 和 8 比特用 M（Modifier）表示，M 的取值不同表示不同功能的无序号帧。无序号帧可用于建立连接和拆除连接。在所有 3 种情况下，第 5 比特是轮询／终止（Poll/Final）比特，简称 P/F，用于询问对方是否有数据要发送或告诉对方数据传输结束。

- Data 是数据字段，可以包含任意信息且可以是任意长的，但实际上受多种条件的制约，如帧校验效率就会随着数据长度的增加而下降。
- FCS 是校验序列字段，采用 16 位的 CRC 校验，其生成多项式为 CRC–16：$G(x)=x^{16}+x^{12}+x^5+1$，校验的内容包括 A 字段、C 字段和 Data 字段。

4.3.2 HDLC 用于实现面向连接的可靠传输

图 4.5 给出了将 HDLC 用于实现有确认的面向连接数据传输服务的例子。图 4.5 为正常传输，其中将无序号帧用于链路连接的建立、维护与拆除，而信息帧用于发送数据并实现捎带的帧确认。图 4.6 则表示出现差错后的处理过程，但省略了关于连接建立的过程。由于 B 方没有数据帧要发送给 A 方，所以不能利用信息帧的捎带来反馈帧出错信息，只有专门发送一个监控帧用于告诉 A 方数据帧传输出错并同时给出建议的差错控制方式，显然在该例子中差错控制采用了选择重发方式。

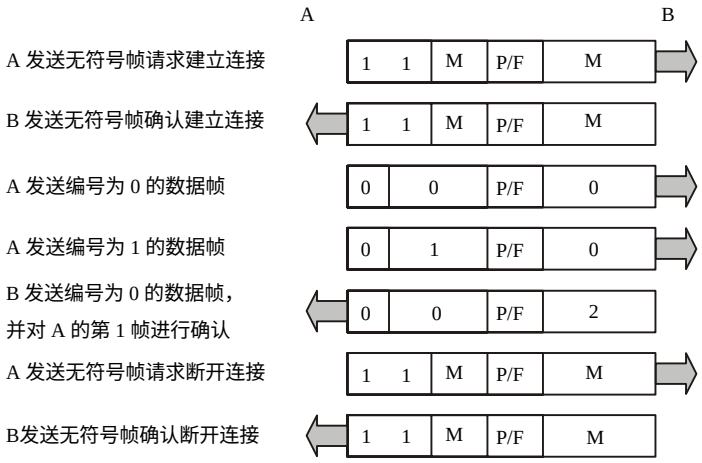

图 4.5　有确认的面向连接 HDLC 的连接建立、数据传输和连接拆除

4.4　点对点协议（PPP）

PPP 是点对点协议（Point-to-Point Protocol）的简称，它是一个工作于数据链路层的广域网协议。PPP 由 IETF（Internet Engineering Task Force）开发，目前已被广泛使用并成为国际标准。无论是同步电路还是异步电路，PPP 协议都能够建立路由器之间或者主

机到网络之间的连接，如图 4.7 所示。例如利用 Modem 进行拨号上网（163、169、165 等）就是使用 PPP 实现主机到网络连接的典型例子。

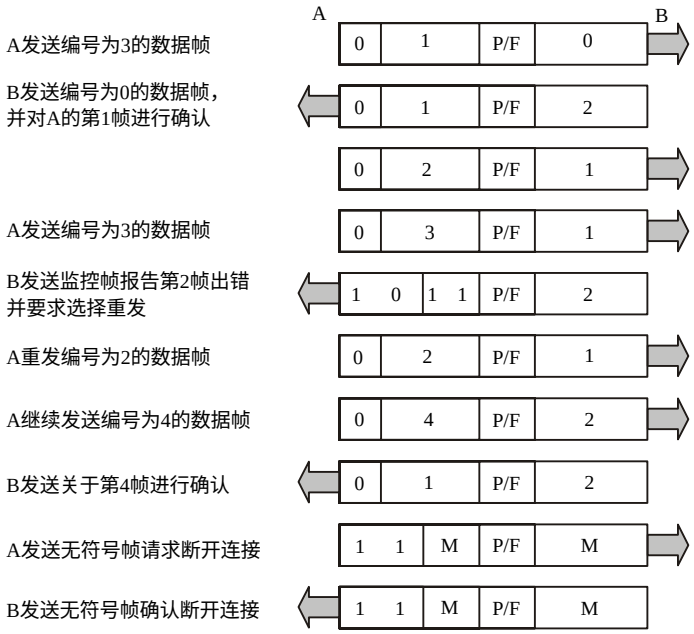

图 4.6　有确认的面向连接 HDLC 差错控制的实现连接建立、数据传输和连接拆除

图 4.7　PPP 提供多种连接

4.4.1　PPP 的特性

PPP 协议是目前使用最广泛的广域网协议，这是因为它具有以下特性：

- 能够控制数据链路的建立；
- 能够对 IP 地址进行分配和使用；
- 允许同时采用多种网络层协议；
- 能够配置和测试数据链路；
- 能够进行错误检测；
- 有协商选项，能够对网络层的地址和数据压缩等进行协商。

PPP 是现在主流的一种国际标准 WAN 封装协议，可支持如下连接类型：

- 同步串行连接；
- 异步串行连接；

- ISDN 连接；
- HSSI 连接。

4.4.2 PPP 的组成

PPP 作为第 2 层的协议，在物理上可使用各种不同的传输介质，包括双绞线、光纤及无线传输介质，在数据链路层提供了一套解决链路建立、维护、拆除和上层协议协商、认证等问题的方案；在帧的封装格式上，PPP 采用的是一种 HDLC 的变化形式；其对网络层协议的支持则包括了多种不同的主流协议，如 IP 和 IPX 等。图 4.8 给出了 PPP 的体系结构，其中，链路控制协议（LCP，Link Control Protocol）用于数据链路连接的建立、配置与测试，NCP（Network Control Protocols）则是一组用来建立和配置不同数据链路的网络层协议。

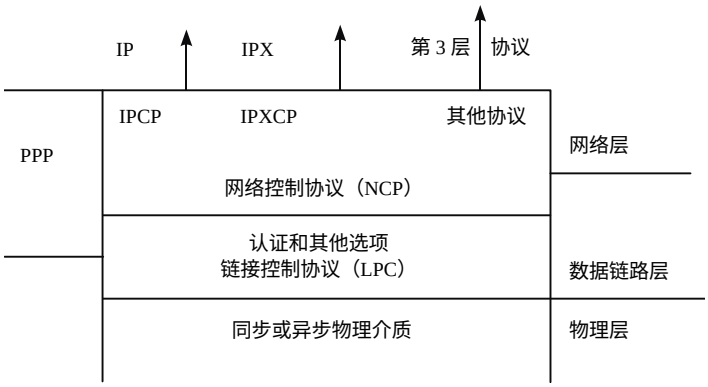

图 4.8　PPP 协议结构

4.4.3 PPP 会话建立的过程

PPP 提供了建立、配置、维护和终止点到点连接的方法。PPP 经过以下 4 个阶段在一个点到点的链路上建立通信连接。

① 链路的建立和配置协调：通信的发起方发送 LCP 帧来配置和检测数据链路。LCP 帧有链路建立帧、链路终止帧和链路维护帧 3 种。在链路建立阶段主要是通过发送 LCP 的帧来对链路进行相关的配置，包括数据的最大传输单元、是否采用 PPP 的压缩、PPP 的认证方式等。

② 链路质量检测：在链路建立、协调之后。这一阶段是可选的。主要用于对链路质量进行测试，以确定其能否为上层所选定的网络协议提供足够的支持，另外，若连接的双方已经要求采用安全认证，则在该阶段还要按所选定的认证方式进行相应的身份认证。

③ 网络层协议配置协调：通信的发起方发送 NCP 帧以选择并配置网络层协议。配置完成后，通信双方可以发送各自的网络层协议数据报。通过发送 NCP 包来选择网络层协议并进行相应的配置，不同的网络层协议要分别进行配置。此时，一条完整的 PPP 链路就建立起来了，可在所建立的 PPP 链路上进行数据传输。

④ 关闭链路：通信链路将一直保持到 LCP 或 NCP 帧关闭链路或者是发生一些外部事件（如空闲时间超长或用户干预）。

需要说明的是，尽管 PPP 的验证是一个可选项，但一旦采用身份验证，则必须在网络

层协议阶段之前进行。有两种类型的 PPP 验证，即 PAP（Password Authentication Protocol）与 CHAP（Challenge Handshake Authentication Protocol）方式。PAP 采用的是两次握手方式，远程结点提供用户名与密码，由本地结点提供身份验证的确认或拒绝。

用户名与密码对由远程网络结点不断地在链路上发送，直到验证被确认或被终结。密码在传输过程中采用的是明文方式，而且发送登录请求的时间和频率完全由远程结点控制，所以这种验证方式虽然实现简单但易受到攻击。CHAP 所使用的是三次握手的验证方式，本地结点提供一个用于身份验证的挑战值，由远程结点根据所收到的挑战值计算出一个回应值发送回本地结点，若该值与本地结点的计算结果一致，则远程结点被验证通过。显然一个没有获得挑战值的远程结点是不可能尝试登录并建立连接的，也就是说 CHAP 由本地来控制登录的时间与频率，并且由于每次所发送的挑战值都是一个不可预测的随机变量，所以 CHAP 较之 PAP 更加安全有效，因此在通常情况下，更多采用的是 CHAP 验证方式。

4.5 数据链路层的设备与组件

数据链路层的设备与组件是指那些同时具有物理层和数据链路层功能的设备或组件。数据链路层的主要设备与组件有网卡、网桥和交换机，下面分别给予介绍。

4.5.1 网卡

网卡是局域网中提供各种网络设备与网络通信介质相连的接口，全名是网络接口卡（NIC，Network Interface Card），也叫网络适配器，其品种和质量的好坏，直接影响网络的性能和网上所运行软件的效果。网卡作为一种 I/O 接口卡插在主机板的扩展槽上，其基本结构包括接口控制电路、数据缓冲器、数据链路控制器、编码解码电路、内收发器、介质接口装置等 6 大部分。网卡主要实现数据的发送与接收、帧的封装与拆封、编码与解码、介质访问控制和数据缓存等功能。因为网卡的功能涵盖了 OSI 参考模型的物理层与数据链路层，所以通常将其归于数据链路层的组件。

每一网卡在出厂时都被分配了一个全球惟一的地址标识，该标识被称为网卡地址或 MAC 地址，由于该地址是固化在网卡上的，所以又被称为物理地址或硬件地址。网卡地址由 48 位长度的二进制数组成。其中，前 24 位表示生产厂商（由 IEEE802.3 委员会分配给各网卡生产厂家），后 24 位为生产厂商所分配的产品序列号。若采用 12 位的十六进制数表示，则前 6 个十六进制数表示厂商，后 6 个十六进制数表示该厂商网卡产品的序列号。如网卡地址 00-90-27-99-11-cc，其中前 6 个十六进制数表示该网卡由 Intel 公司生产，相应的网卡序列号为 99-11-cc。网卡地址主要用于设备的物理寻址，与 IP 地址所具有的逻辑寻址作用有着截然不同的区别。

网卡的分类方法有多种，例如按照传输速率、总线类型、所支持的传输介质、用途或网络技术等来进行分类。

按照网络技术的不同可分为以太网卡、令牌环网卡、FDDI 网卡等。目前以太网网卡最常见。

按照传输速率，单单以太网卡就提供了 10 Mbps、100 Mbps、1000 Mbps 和 10 Gbps 等多种速率。数据传输速率是网卡的一个重要指标。

按照总线类型分类，网卡可分为 ISA 总线网卡、EISA 总线网卡、PCI 总线网卡及其他总线网卡等。16 位 ISA 总线网卡的带宽一般为 10 Mbps，没有 100 Mbps 以上带宽的 ISA 网卡。目前 PCI 网卡最常用，PCI 总线网卡常用的为 32 位，其带宽从 10 Mbps 到 1000 Mbps。

按照所支持的传输介质，网卡可分为双绞线网卡、粗缆网卡、细缆网卡、光纤网卡和无线网卡。连接双绞线的网卡带有 RJ-45 接口，连接粗缆的网卡带有 AUI 接口，连接细缆的网卡带有 BNC 接口，连接光纤的网卡则带有光纤接口。当然有些网卡同时带有多种接口，如同时具备 RJ-45 口和光纤接口。目前，市场上还有带 USB 接口的网卡，这种网卡可以用于具备 USB 接口的各类计算机网络。

另外，按照用途，网卡还可分为工作站网卡、服务器网卡和笔记本电脑网卡等。

4.5.2 网桥

网桥提供了一种最简单的将局域网网段连接成可维护、高可靠性的扩展网络的方法。网桥工作在 OSI 模型中数据链路层的 MAC 子层。

网桥可以将局域网分成两个或更多的网段，它通过隔离每个网段内部的数据流量，从而增加了每个结点所能使用的有效带宽。

网桥的重要功能是不受介质访问子层中冲突域的限制而扩展网长，对于众多的共享 LAN 可以隔离 LAN 段，为每一个 LAN 段提供相同的带宽，这就等于扩大了总带宽，可使各个 LAN 段内部信息包、冲突包都不会广播到另一个 LAN 段，明显地提高了利用效率。同时，网桥又具有存储、转发、过滤功能，使应当发送到另一个 LAN 的信息正确转发。

最基本的网桥用来连接两个或更多的局域网网段。网桥和每个局域网网段之间的接口称为端口。连接到每个端口的局域网被称为一个网段。

所有网桥都在数据链路层提供连接服务，一种常用的分类方法是将网桥分为本地网桥和远程网桥。本地网桥在同一区域中为多个局域网网段提供一个直接的连接，而远程网桥则通过电信线路，将分布在不同区域的局域网网段互连起来，如图 4.9 所示。

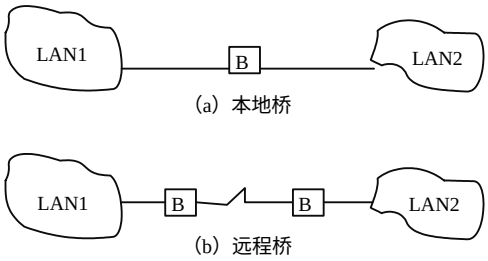

图 4.9　本地网桥和远程网桥

网桥主要具有如下功能。

(1) 在物理上扩展网络

一个网桥可以连接多个网络，同时一个网络又可以使用多个网桥与其他网络互连。所以通过网桥，可以在物理上将多个不同的网段互连在一起，从而扩大了网络的地址覆盖范围和主机规模。从这一点上看，网桥具备和中继器、集线口类似的在物理上扩展网络的功能。

(2) 数据过滤功能

在网桥中，要维持一个交换表，该表给出关于网桥不同接口所连主机的 MAC 地址信

息，网桥根据数据帧中的目的地址判断是否转发该帧。也就是说，网桥从某一接口收到数据帧时，将首先获取目的 MAC 地址，然后查看交换表，若发送结点与目的结点在同一个网段内时，则网桥就不转发该帧，只有源结点与目的结点不在同一个网段时，网桥才转发该帧。也就是说，网桥具有基于第 2 层地址进行帧过滤的功能。

（3）逻辑划分网络的功能

通过对帧的过滤，网桥实现了物理网络内部通信的相互隔离，源和目标在同一物理网段中的数据帧由于网桥的数据过滤作用是不会被转发或渗透到其他网段中的，尽管从物理上看，这些网段通过网桥和源与目标主机所在的网段是互连在一起的。我们将网桥所具备的这种隔离功能称为逻辑划分网络的功能，这项功能也是网桥与物理网络互连设备中继器及集线器之间的最大区别，物理层设备只能转发原始比特流，从而不能根据某种地址信息实现数据过滤功能。

（4）数据推进功能

网桥根据数据过滤的结果实现数据帧的转发。在网桥中可以设置缓冲区以缓存输出端口无法立即传输的数据，从而可以使网桥输出帧的速率与接收 LAN 的速率相同。

（5）帧格式转换功能

当数据帧通过网桥到达另一个执行不同局域网协议的 LAN 时，网桥还能够对帧格式进行转换处理。也就是将一种帧格式转换为另一种帧格式，其中包括位组的重新排列、帧长度的限制以及重新生成校验序列等。

4.5.3 交换机

随着局域网对容量和性能方面需求的增高，1993 年，局域网交换设备出现。1994 年，国内掀起了交换网络技术的热潮。

1．交换技术的基本原理

局域网交换技术是 OSI 参考模型中的第 2 层——数据链路层（Data-Link Layer）上的技术，所谓“ 交换 ”实际上就是指转发数据帧（Frame）。在数据通信中，所有的交换设备（即交换机）执行两个基本的操作：

- 交换数据帧，将从输入介质上收到的数据帧转发至相应的输出介质；
- 维护交换操作，构造和维护交换地址表。

（1）交换数据帧

交换机根据数据帧的 MAC（Media Access Control）地址（即物理地址）进行数据帧的转发操作。交换机转发数据帧时，遵循以下规则：

- 如果数据帧的目的 MAC 地址是广播地址或者组播地址，则向交换机所有端口转发（除数据帧来的端口）。
- 如果数据帧的目的地址是单播地址，但是这个地址并不在交换机的地址表中，那么也会向所有的端口转发（除数据帧来的端口）。
- 如果数据帧的目的地址在交换机的地址表中，那么就根据地址表转发到相应的端口。
- 如果数据帧的目的地址与数据帧的源地址在一个网段上，它就会丢弃这个数据帧，交换也就不会发生。

下面，以图 4.10 为例来看看具体的数据帧交换过程。

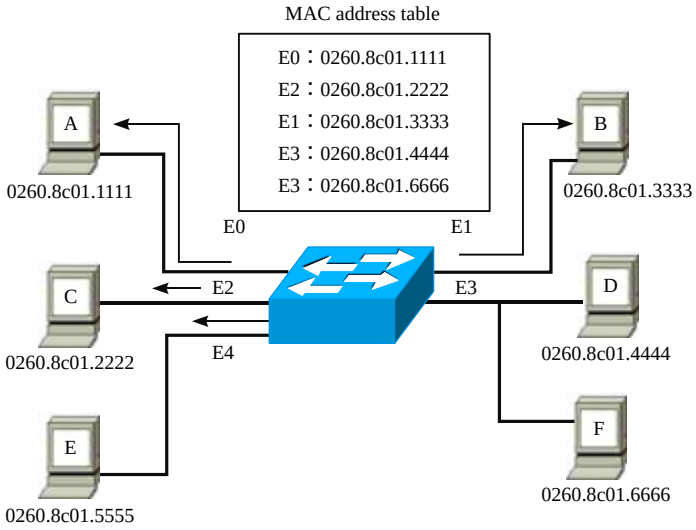

图 4.10　数据帧交换过程

① 当主机 D 发送广播帧时，交换机从 E3 端口接收到目的地址为 ffff.ffff.ffff 的数据帧，则向 E0、E1、E2 和 E4 端口转发该数据帧。

② 当主机 D 与 E 主机通信时，交换机从 E3 端口接收到目的地址为 0260.8c01.5555 的数据帧，查找地址表后发现 0260.8c01.5555 并不在表中，因此交换机仍然向 E0、E1、E2 和 E4 端口转发该数据帧。

③ 当主机 D 与主机 F 通信时，交换机从 E3 端口接收到目的地址为 0260.8c01.6666 的数据帧，查找地址表后发现 0260.8c01.6666 也位于 E3 端口，即与源地址处于同一个网段，所以交换机不会转发该数据帧，而是直接丢弃。

④ 当主机 D 与主机 A 通信时，交换机从 E3 端口接收到目的地址为 0260.8c01.1111 的数据帧，查找地址表后发现 0260.8c01.1111 位于 E0 端口，所以交换机将数据帧转发至 E0 端口，这样主机 A 即可收到该数据帧。

⑤ 如果在主机 D 与主机 A 通信的同时，主机 B 也正在向主机 C 发送数据，交换机同样会把主机 B 发送的数据帧转发到连接主机 C 的 E2 端口。这时 E1 和 E2 之间，以及 E3 和 E0 之间，通过交换机内部的硬件交换电路，建立了两条链路，这两条链路上的数据通信互不影响，因此网络亦不会产生冲突。所以，主机 D 和主机 A 之间的通信独享一条链路，主机 C 和主机 B 之间也独享一条链路。而这样的链路仅在通信双方有需求时才会建立，一旦数据传输完毕，相应的链路也随之拆除。这就是交换机主要的特点。

从以上的交换操作过程中，可以看到数据帧的转发都是基于交换机内的 MAC 地址表，但是这个地址表是如何建立和维护的呢？下面我们就来介绍这个问题。

（2）构造维护交换地址表

交换机的交换地址表中，一条表项主要由一个主机 MAC 地址和该地址所位于的交换机端口号组成。整张地址表的生成采用动态自学习的方法，即当交换机收到一个数据帧以后，将数据帧的源地址和输入端口记录在交换地址表中。思科的交换机中，交换地址表放置在内容可寻址存储器（CAM，Content-Addressable Memory）中，因此也被称为 CAM 表。

当然，在存放交换地址表项之前，交换机首先应该查找地址表中是否已经存在该源地址的匹配表项，仅当匹配表项不存在时才能存储该表项。每一条地址表项都有一个时间标记，用来指示该表项存储的时间周期。地址表项每次被使用或者被查找时，表项的时间标记就会被更新。如果在一定的时间范围内地址表项仍然没有被引用，它就会从地址表中被移走。因此，交换地址表中所维护的一直是最有效和最精确的地址——端口信息。

2．交换机 3 种交换技术

（1）端口交换

端口交换技术最早出现在插槽式的集线器中，这类集线器的背板通常划分有多条以太网段（每条网段为一个广播域），不用网桥或路由连接，网络之间是互不相通的。以太网主模块插入后通常被分配到某个背板的网段上，端口交换用于将以太网模块的端口在背板的多个网段之间进行分配、平衡。根据支持的程度，端口交换还可细分为以下几种。

- 模块交换：将整个模块进行网段迁移。
- 端口组交换：通常模块上的端口被划分为若干组，每组端口允许进行网段迁移。
- 端口级交换：支持每个端口在不同网段之间进行迁移。这种交换技术是基于 OSI 第 1 层完成的，具有灵活性和负载平衡能力等优点。如果配置得当，还可以在一定程度进行容错，但没有改变共享传输介质的特点，因而不能称之为真正的交换。

（2）帧交换

帧交换是目前应用最广的局域网交换技术，它通过对传统传输媒介进行微分段，提供并行传输的机制，以减小冲突域，获得高的带宽。一般来讲每个公司的产品实现技术均会有差异，但对网络帧的处理方式一般有以下几种。

- 直通交换：提供线速处理能力，交换机只读出网络帧的前 14 个字节，便将网络帧传输到相应的端口上。
- 存储转发：通过对网络帧的读取进行验错和控制。

前一种方法的交换速度非常快，但缺乏对网络帧进行更高级的控制，缺乏智能性和安全性，同时也无法支持具有不同速率的端口的交换。因此，各厂商把后一种技术作为重点。

有的厂商甚至对网络帧进行分解，将帧分解成固定大小的信元，该信元处理极易用硬件实现，处理速度快，同时能够完成高级控制功能（如美国 MADGE 公司的 LET 集线器的优先级控制）。

（3）信元交换

ATM 技术代表了网络和通信技术发展的未来方向，也是解决目前网络通信中众多难题的一剂“良药”。ATM 采用固定长度 53 个字节的信元交换，由于长度固定，因而便于用硬件实现。ATM 采用专用的非差别连接，并行运行，可以通过一个交换机同时建立多个结点，但并不会影响每个结点之间的通信能力。ATM 还容许在源结点和目标结点建立多个虚拟链接，以保障足够的带宽和容错能力。ATM 采用了统计时分电路进行复用，因而能大大提高通道的利用率。ATM 的带宽可以达到 25 Mbps、155 Mbps、622 Mbps 甚至数吉比特每秒的传输能力。

3．局域网交换机的种类

① 从广义上来看，交换机分为两种：广域网交换机和局域网交换机。广域网交换机主要应用于电信领域，提供通信基础平台。而局域网交换机则应用于局域网络，用于连接终端设备，如 PC 及网络打印机等。

② 按照现在复杂的网络构成方式，网络交换机被划分为接入层交换机、汇聚层交换机和核心层交换机。其中，核心层交换机全部采用机箱式模块化设计，已经基本上都设计了与之相配备的 1000Base-T 模块。接入层支持 1000Base-T 的以太网交换机基本上是固定端口式交换机，以 10/100M 端口为主，并且以固定端口或扩展槽方式提供 1000Base-T 的上联端口。汇聚层 1000Base-T 交换机同时存在机箱式和固定端口式两种设计，可以提供多个 1000Base-T 端口，一般也可以提供 1000Base-X 等其他形式的端口。接入层和汇聚层交换机共同构成完整的中小型局域网解决方案。

③ 从传输介质和传输速度上看，局域网交换机可以分为以太网交换机、快速以太网交换机、千兆位以太网交换机、FDDI 交换机、ATM 交换机和令牌环交换机等多种，这些交换机分别适用于以太网、快速以太网、FDDI、ATM 和令牌环网等环境。

④ 从规模应用上又有企业级交换机、部门级交换机和工作组交换机等。一般来讲，企业级交换机都是机架式，部门级交换机可以是机架式，也可以是固定配置式，而工作组级交换机则一般为固定配置式，功能较为简单。另一方面，从应用的规模来看，作为骨干交换机时，支持 500 个信息点以上大型企业应用的交换机为企业级交换机，支持 300 个信息点以下中型企业的交换机为部门级交换机，而支持 100 个信息点以内的交换机为工作组级交换机。

⑤ 按照 OSI 的参考网络模型，交换机又可以分为第 2 层交换机、第 3 层交换机、第 4 层交换机等，一直到第 7 层交换机。基于 MAC 地址工作的第 2 层交换机最为普遍，用于网络接入层和汇聚层。基于 IP 地址和协议进行交换的第 3 层交换机普遍应用于网络的核心层，也少量应用于汇聚层。部分第 3 层交换机也同时具有第 4 层交换功能，可以根据数据帧的协议端口信息进行目标端口判断。第 4 层以上的交换机称之为内容型交换机，主要用于互联网数据中心。

⑥ 按照交换机的可管理性，又可把交换机分为可管理型交换机和不可管理型交换机，它们的区别在于对 SNMP、RMON 等网管协议的支持。可管理型交换机便于网络监控、流量分析，但成本也相对较高。大中型网络在汇聚层应该选择可管理型交换机，在接入层视应用需要而定，核心层交换机则全部是可管理型交换机。

⑦ 按照交换机是否可堆叠，交换机又可分为可堆叠型交换机和不可堆叠型交换机两种。设计堆叠技术的一个主要目的是为了增加端口密度。

⑧ 按照最广泛的普通分类方法，局域网交换机可以分为桌面型交换机（Desktop Switch）、工作组型交换机（Workgroup Switch）和校园网交换机（Campus Switch）3 类。桌面型交换机是最常见的一种交换机，使用最广泛，尤其是在一般办公室、小型机房和业务受理较为集中的业务部门、多媒体制作中心、网站管理中心等部门。在传输速度上，现代桌面型交换机大都提供多个具有 10/100 Mbps 自适应能力的端口。工作组型交换机常用来作为扩充设备，在桌面型交换机不能满足需求时，大多直接考虑工作组型交换机。虽然工作组型交换机只有较少的端口数量，但却支持较多的 MAC 地址，并具有良好的扩充能力，端口的传输速度基本上为 100 Mbps。校园网交换机的应用相对较少，仅应用于

大型网络，且一般作为网络的骨干交换机，并具有快速数据交换能力和全双工能力，可提供容错等智能特性，还支持扩充选项及第 3 层交换中的虚拟局域网（VLAN）等多种功能。

⑨ 根据交换技术的不同，有人又把交换机分为端口交换机、帧交换机和信元交换机 3 种。

⑩ 从应用的角度划分，交换机又可分为电话交换机（PBX）和数据交换机（Switch）。当然，目前在数据上的语音传输 VoIP 又有人称之为“ 软交换机 ”。

4．交换机之间的连接

交换机之间最简单的一种连接方法就是采用一根交叉的双绞线（1、2 和 3、6 对调）并将它们连接起来。如果下级交换机有 Uplink 口，也可以接到 Uplink 口上，用直连线连接。总的来讲，交换机之间的连接有以下几种。

（1）级联

交换机可以通过上联端口实现与骨干交换机的连接。

（2）冗余连接

在以太网环境下是不允许出现环路的，生成树（Spanning Tree）则可以在交换机之间实现冗余连接又避免出现环路。当然，这要求交换机支持 Spanning Tree。

不过，Spanning Tree 冗余连接的工作方式是 Stand By，也就是说，除了一条链路工作外，其余链路实际上是处于待机（Stand By）状态，这显然影响传输的效率。一些最新的技术，例如 FEC（Fast Ethernet Channel）、ALB（Advanced Load Balancing）和 Port Trunking 技术，则可以允许每条冗余连接链路实现负载分担。其中 FEC 和 ALB 技术用来实现交换机与服务器之间的连接（Server to Switch），而 Port Trunking 技术则实现交换机之间的连接（Switch to Switch）。通过 Port Trunking 的冗余连接，交换机之间可以实现几倍于线速带宽的连接。

（3）堆叠

提供堆叠接口的交换机之间可以通过专用的堆叠线连接起来。通常，堆叠的带宽是交换机端口速率的几十倍，例如，一台 100 Mbps 交换机，堆叠后两台交换机之间的带宽可以达到几百兆位甚至上千兆位。

多台交换机的堆叠是靠一个提供背板总线带宽的多口堆叠母模块与单口的堆叠子模块相联实现的，并插入不同的交换机实现交换机的堆叠。上联交换机可以通过上联端口实现与骨干交换机的连接。例如，一台具有 24 个 10 Mbps 和 1 个 100 Mbps 端口的交换机，就可以通过 100 Mbps 端口与 100 Mbps 主干交换机实现 100 Mbps 速率的连接。

交换机作为多端口网桥，确实具备了网桥所拥有的全部功能，如物理上扩展网络、逻辑上划分网络等。但是作为对网桥的改进设备，首先，交换机可以提供高密度的连接端口；其次，交换机由于采用的基于交换背板的虚电路连接方式，从而可为每个交换机端口提供更高的专用带宽，而集中网桥在数据流量大时易形成瓶颈效应。另外，交换机的数据转发是基于硬件实现的，所以较网桥采用软件实现数据的存储转发也具有更高的交换性能。正因为如此，在交换机问世后，网桥已逐渐退出了第 2 层网络互连设备的市场。

4.6　技能训练：交换机和集线器的级联

交换机和集线器的连接都可以采用级联方式，两者级联的方法基本相同，只是集线器的级联应严格遵守集线器级联配置规则（5-4-3-2-1 规则）。下面以交换机为例介绍级联的方法。

如果交换机具有级联端口，那么可以通过一条 UTP 直连线将一条交换机的级联端口连入另一台交换机的普通端口。

如果交换机没有级联端口，那么可以通过一条 UTP 交叉线将两台交换机的普通端口进行连接。

习　　题

1．数据链路层的主要任务是什么？

2．数据链路层的常用成帧方法有哪些？

3．何为差错？引起差错的原因是什么？

4．试计算传输信息 1011001 的 CRC 编码，假设其生成多项式 $G(x)=x^4+x^3+1$。

5．简述滑动窗口的原理。

6．简述 HDLC 帧各字段的意义。HDLC 帧可分为哪几个大类？简述各类帧的作用。

7．网卡的主要功能是什么？

8．试对网桥和交换机的异同之处进行比较。