计网笔记 - 第五章

第五章 链路层: 链路, 接入网和局域网

网络层数据报如何被封装进链路层帧，不同的链路能够采用不同的链路层协议吗，如何解决传输碰撞，链路层存在编址吗.

5.1 链路层概述

- 链路：沿着通信路径连接相邻结点的通信信道.

  通过特定链路时，传输结点将数据报封装在链路层帧中，并将该帧传送到链路中.

5.1.1 链路层提供的服务

- 成帧 (framing)：网络层数据报经链路传输之前 都要用链路层帧封装起来. 一个帧由一个数据字段和若干个首部字段组成.
- 链路接入：媒体访问控制 (Medium Access Control MAC)协议规定了帧在链路上传输的规则. MAC协议用于协调多个结点的帧传输.
- 可靠交付：保证无差错地移动每个网络层数据报，通常是通过确认和重传取得的.

- 差错检测和纠正：(运输层和网络层提供了因特网检验和：有限形式的差错检测)．链路层差错检测通常更复杂并且用硬件实现．

5.1.2 链路层在何处实现

链路层的主体部分：在网络适配器(network adapter)(或称 网络接口卡 NIC)，位于网络适配器核心的是链路层控制器．一个实现了许多服务的芯片．

5.2 差错检测和纠正技术

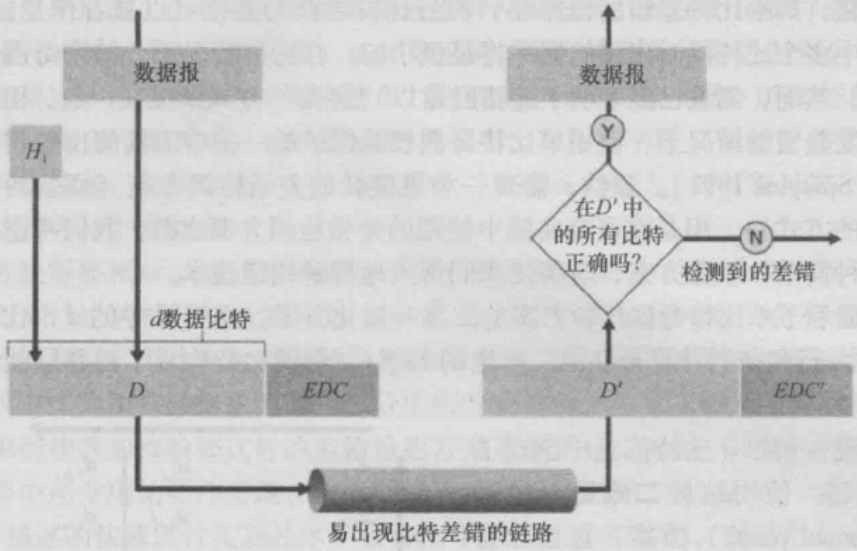

图 5-3　差错检测与纠正的场景

- 比特级差错检测和纠正：

  差错检测和纠正比特 EDC：增强数据 D．

  需要在只收到 D' 和 EDC' 的情况下，确定 D' 是否和初始的 D 相同．

  目的是让 未检出比特差错 发生概率越小．

5.2.1 奇偶校验

只能检测出奇数个比特差错．

- 使用单比特奇偶校验 未检测出差错的概率：50 %．
- 二维奇偶校验：

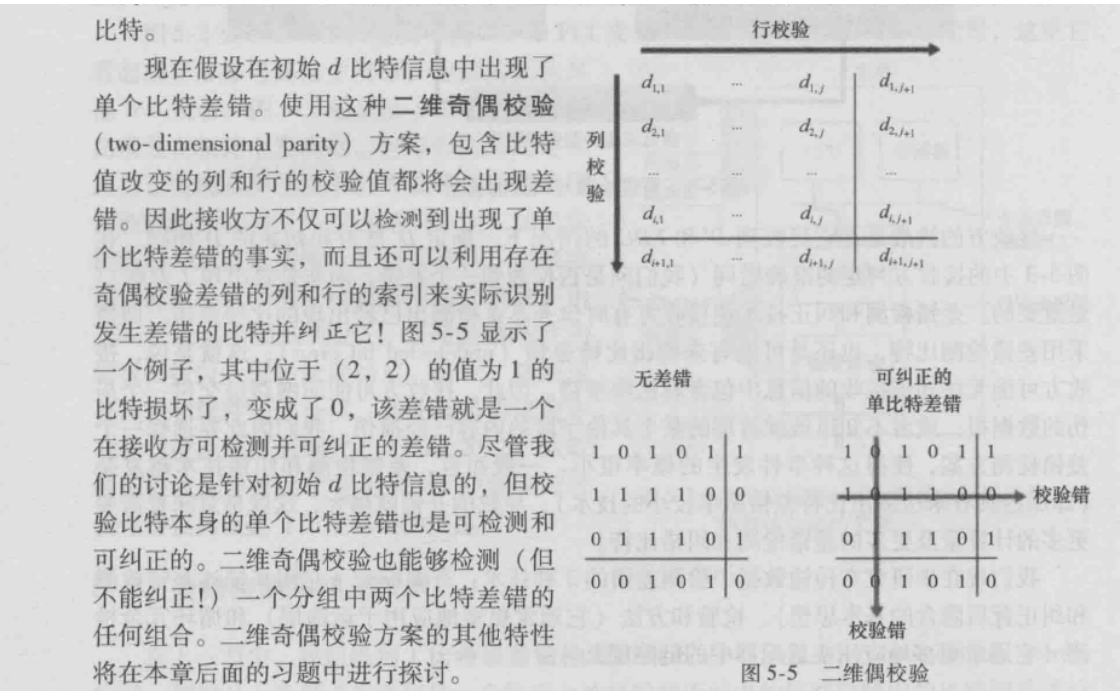

图 5-5 二维偶校验

不仅可以检测到差错，还可以纠错.

- 前向纠错：接收方检测和纠正差错的能力．这避免了不得不等待的往返时延.

5.2.2 检验和方法

因特网检验和，去看习题.

5.2.3 循环冗余检测

循环冗余检测编码 (CRC)也称为多项式编码．该编码将发送的比特串看成 系数是 0/1 的多项式．去看习题和王道.

- 生成多项式 G 的最高位有效比特(最左边)是1.
- 每个CRC标准都可以检测小于 $r+1$ 的比特突发差错．意味着所有连续的 r 比特或者更小的差错都可以检测到.

  长度大于 $r+1$ 比特的突发差错以概率 $1-0.5^r$ 被检测到.

  每个CRC标准也都能检测到任何奇数个比特的差错.

5.3 多路访问链路和协议

如何协调多个发送和接受结点对一个共享广播信道的访问，多路访问问题.

两种链路：1. 点对点链路 2. 广播链路.

- 点对点链路：由链路一端的单个发送方和链路另一端的单个接收方组成.
- 广播链路：多个发送方和接受方连接到相同的，单一的，共享的广播信道上.

发送碰撞：传输的帧在所有接收方处发送碰撞，没有一个接受结点能够有效地获得任何传输的帧.

我们所希望的是(对于速率 R bps 的信道：

1. 当仅有一个结点有数据发送时，该结点具有 R bps 的吞吐量.
2. 当有 M 个结点要发送数据时，每个节点(在适当的时间间隔内)吞吐量为 R/M bps，可以是平均传输速率.
3. 协议是分散的，不会因为某主结点故障导致整个系统崩溃.
4. 协议是简单的，实现不昂贵.

5.3.1 信道划分协议

时分多路复用(TDM)，频分多路复用(FDM)：用于在所有共享信道结点之间划分广播信道带宽的技术.

TDM

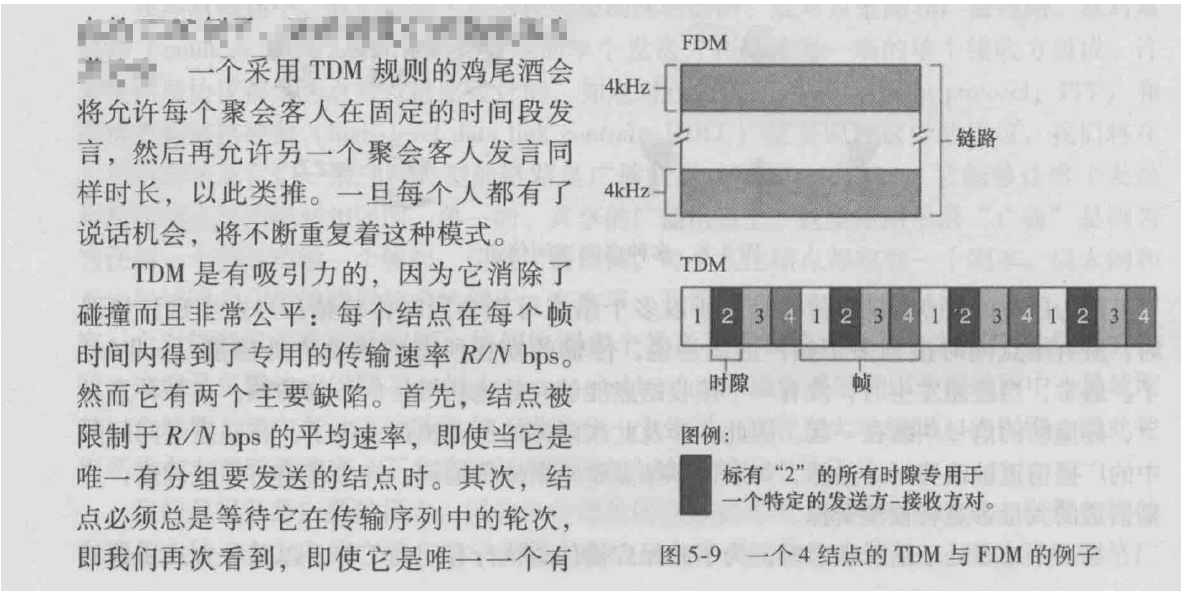

图 5-9　一个 4 结点的 TDM 与 FDM 的例子

TDM消除了碰撞并且非常公平，请注意TDM每个节点在每个帧时间内得到了专用的 R/N bps，因为本来是全部时间可以传的，现在只有 $1/N$ 的时间可以传了.

缺陷：

- 结点被限制于 R/N bps 的平均速率，即使只有一个结点要发送分组.
- 结点必须总是等待它在传输序列中的轮次，轮到它了才可以发.

FDM

与TDM一样有相同的优缺点. FDM将 R bps信道划分为不同的频段. 每个频段具有 R/N 带宽.

码分多址 CDMA

CDMA 对每个节点分配一种不同的编码. 这会被用来对其发送的数据进行编码.

不同的结点能够同时传输，并且它们各自相应的接收方仍能正常接收发送方编码的数据比特，而不在乎其他结点的干扰传输.

5.3.2 随机接入协议

一个传输结点总是以信道的全部速率（即 R bps）进行发送. 有碰撞时涉及碰撞的每个结点反复重发它的帧. 直到无碰撞地通过.

1 时隙 ALOHA

我们做如下假设：

- 所有帧由 L 比特组成.
- 时间被划分成长度为 L/R 秒的时隙，一个时隙等于传输一帧的时间.
- 结点只在时隙起点开始传输帧.
- 结点是同步的，每个结点都知道时隙何时开始.
- 如果在一个时隙中有两个或者更多个帧碰撞，则所有节点在该时隙结束之前 检测到该碰撞事件.

具体操作：

- 在下一个时隙开始时传输整个帧.
- 如果没有碰撞，则传输成功.
- 如果有碰撞，以概率 p 在后续的每个时隙中重传，直到成功.

请注意如果当前时隙没重传($1-p$ 概率)，并不代表下一个时隙就一定会重传，仍然是概率 p.

- 优点：

    - 仅有一个结点时，时隙ALOHA是全速 R bps连续传输的.
    - 高度分散，每个结点独立决定什么时候重传.
    - 简单的协议.
- ALOHA最大效率的推导:
    - N 个结点，一个传，其他不传，所有加起来:

      $$N \cdot p(1-p)^{N-1}$$

      极限推导不难 $\frac{1}{e}$，看习题.

2 纯ALOHA

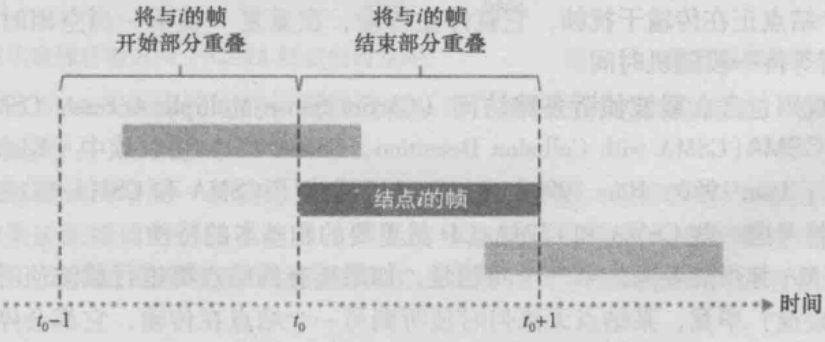

图 5-11　纯 ALOHA 中的干扰传输

如果碰撞，立即(在完全传输完碰撞帧之后)以概率p重传. $(1-p)$ 概率等待一个帧运输时间.

帧要在 t_0 成功传输，需要前面和后面都没有帧来碰撞:

- $[t_0-1, t_0]$ 不能有帧，$(1-p)^{N-1}$.
- t_0 之后开始传了也不能，$(1-p)^{N-1}$.
- $p(1-p)^{2(N-1)}$，极限效率为 $\frac{1}{2e}$

时隙ALOHA和纯ALOHA的传输都是决定独立于连接到这个广播信道的其他结点的活动.

3 载波侦听多路访问

- 说话之前必听:

  载波侦听，直到检测到一小段时间没有传输.
- 如果与他人同时开始说话，则停止.

  碰撞检测，传输时一直侦听，如果检测到干扰就停止.

类型:

- CSMA: 载波侦听多路访问.
- CSMA/CD: 具有碰撞检测的CSMA. (CSMA with Collision Detection).

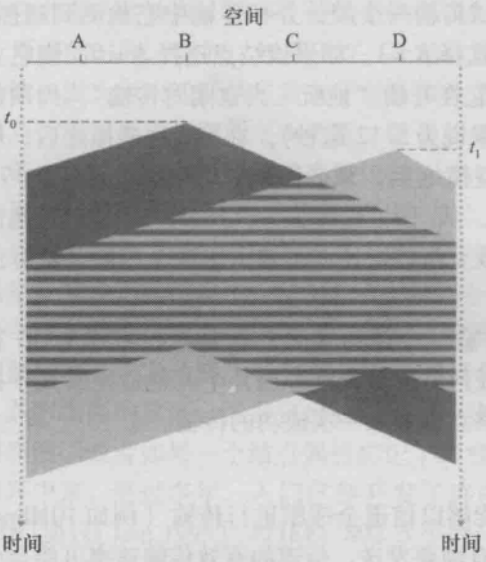

图 5-12　发生碰撞传输的两个 CSMA 结点的时空图

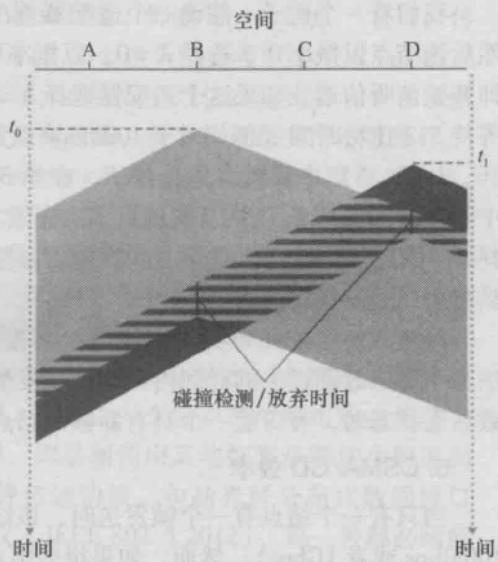

图 5-13　具有碰撞检测的 CSMA

如上图：

- 左边：CSMA

  因为 端到端的信道传播时延(信号从一个结点到另一个结点的时间)，B，D结点都听到了信道空闲，此时发送并产生了碰撞.

- 右边：CSMA/CD

  这两个结点在检测到碰撞后很短时间内放弃了传输.

CSMA/CD 运行：

1. 适配器从网络层获取数据报，封装成链路层帧.
2. 适配器如果侦听到信道空闲，则开始传输. 否则等待.
3. 传输过程中监视来自其他使用该广播信道的适配器的信号能量存在.
4. 如果监视到其他信号能量，则停止传输.
5. 停止后等待随机时间量，转到2. 转到2还要侦听是否空闲啊.

碰撞后等待的随机时间量：二进制指数退避法，经历 n 次碰撞则随机地从 $\{0,1,\cdots,2^n-1\}$ 中选择一个 K. 等待的实际时间量是 $K\cdot 512$ 比特时间. 例子：

- 第一次碰撞：K 从 $\{0,1\}$ 中随机.
- 第二次碰撞：K 从 $\{0,1,2,3\}$ 中随机.
- 第三次碰撞：K 从 $\{0,1,2,3,4,5,6,7\}$ 中随机.

5 CSMA/CD效率

$$\frac{1}{1+5d_{prop}/d_{trans}}$$

其中 d_{prop} 表示信号能量在任意两个适配器之间传播所需的最大时间. d_{trans} 表示传输一个最大长度的以太网帧的时间.

请注意：ALOHA和CDMA协议可以做到只有一个结点是全速传输，但是当有 M 个结点活跃时，每个活跃结点的吞吐量接近 R/M 是不具备的.

5.3.3 轮流协议

- 轮询协议：

  要求结点之一要被指定为主结点. 主结点以循环的方式轮询每个结点. 告诉其能传输的最大数量.

  消除了困扰随机接入协议的 碰撞和空时隙.

  - 缺陷：

引入了轮询时延.

主结点故障则整个信道都完蛋.

- 令牌传递协议:

  一个称为令牌的 小的特殊帧 在结点之间以某种固定的次序进行交换.

  - 有帧要发送: 发送最大数目的帧数.
  - 没有则立即向下一个结点转发该令牌.

5.4.1 链路层寻址和APR

1 MAC 地址

适配器(网络接口)才具有链路层地址.

具有多个网络接口的主机或路由器 将具有与之相关联的多个链路层地址.

链路层地址又叫 MAC地址(LAN地址, 物理地址).

MAC地址长度 6字节, 所以有 2^{48} 个可能的MAC地址. IEEE管理着MAC地址空间. 没有两块适配器具有相同的地址.

具有扁平结构, 无论在哪都不会变化.

MAC 广播地址: 要让局域网上所有其他适配器来接受并处理它打算发送的帧. 十六进制表示法: `FF-FF-FF-FF-FF-FF`

2 地址解析协议 ARP

IP地址: 点分十进制表示法, MAC地址: 十六进制表示法.

每台主机和路由器有一个单一的IP地址和MAC地址表示. 问题在于如何 在已知目的主机IP地址情况下确定 MAC地址?

(很多方面上和DNS类似: DNS将主机名解析为IP地址. 但是:

- DNS为在因特网的任何地方主机解析主机名.
- ARP只为同一个子网上的主机和路由器接口解析IP地址.

如何工作 (可以看lab):

- 每台主机或路由器在内存中有一个 ARP 表, 表项:
  - IP地址, 对应的MAC地址, 相应的TTL.
- 假设现在需要IP寻址到本子网上的另一台主机或路由器.
  - 如果发送主机的ARP表中没有目的结点的表项:
    1. 构造ARP查询分组. 包括发送和接受方 IP地址 MAC地址. 目的是询问子网上其他主机和路由器.
    2. 适配器以MAC广播地址 `FF-FF-FF-FF-FF-FF` 发送查询分组.
    3. 所有其他的适配器ARP模块检查IP地址是否与ARP分组中的目的IP地址匹配, 如果有则返回一个相应ARP分组.
    4. 发送主机更新ARP表, 并发送IP数据报.
    5. 数据报封装在一个链路层帧中.
- ARP查询报文是在广播帧中发送的, 响应报文是在标准帧中发送的.
- ARP是即插即用的.
- ARP分组封装在链路层帧中. 但是ARP也包含网络层地址, 所以ARP是跨越链路层和网络层的协议.

3 发送数据报到子网以外

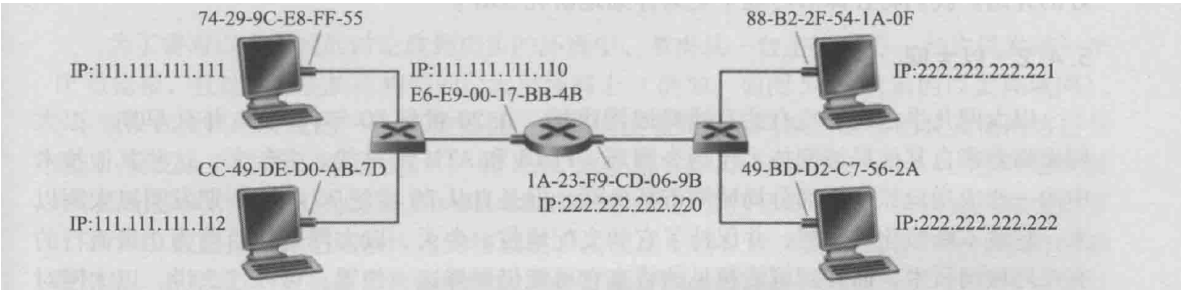

图 5-19　由一台路由器互联的两个子网

- 子网一：`111.111.111.111/24`，子网二：`222.222.222/24`

假设主机 `111.111.111.111` 要向主机 `222.222.222.222` 发送一个IP数据报.

- 请注意此时目的MAC地址不是 子网二主机 `222.222.222.222` 的MAC地址，因为这样数据报在子网一发出后不会传递到网络层.

主要过程:

1. 所以目的MAC，应该填子网一出口的接口的MAC地址：`111.111.111.110`.
2. 给到子网一出口接口后，由第四章中 路由器中转发表，通过路由器接口 `222.222.222.220` 转发.
3. 然后接口 `222.222.222.220` 的适配器中 把数据报封装到一个新的帧中，这时就是目的主机的MAC地址了.

5.4.2 以太网

最流行的有线局域网技术. 便宜，高数据速率.

- 集线器 (hub):

  作用于各个比特而不是帧，只是将比特的能量强度放大并从其他所有接口传输. 如果两个接口同时接收到帧，则发生碰撞，将重新传输.

以太网帧结构

6个字段:

- 数据字段 (46~1500字节): 承载了IP数据报，以太网的最大传输单元 (MTU) 是1500字节. 超过则需要分片. 数据字段的最小长度是46字节，不足则需填充.
- 目的地址 (6字节): 目的适配器的MAC地址.
- 源地址 (6字节): 传输该帧到局域网上的适配器的MAC地址.
- 类型字段 (2字节): 类型字段允许以太网复用多种网络层协议. 为了把一层中的某协议与上一层的某协议结合起来.
- CRC (4字节): 循环冗余检测，目的是使接受适配器检测帧中是否引入了差错. 接收端适配器进行CRC校验不通过，不会发送否定确认帧，只是简单丢弃. 同理通过了也不发送确认帧. 所以是不可靠服务，但是这使以太网简单和便宜.
- 前同步码 (8字节): Preamble. 用于 发送端适配器和接收端适配器的时钟同步. 因为发送端适配器可能不会以精确的速率传输帧. 相对于额定速率有所偏移. 接收端适配器需要通过锁定前同步码的前7字节比特，来锁定发送端适配器的时钟.

所有的以太网技术都向网络层提供无连接服务，不可靠服务.

如果由于丢弃了以太网帧而存在间隙，主机 B 上的应用也会看见这个间隙吗？如我们在第 3 章中学习的那样，这只取决于该应用是使用 UDP 还是使用 TCP。如果应用使用的是 UDP，则主机 B 中的应用的确会看到数据中的间隙。另一方面，如果应用使用的是 TCP，则主机 B 中的 TCP 将不会确认包含在丢弃帧中的数据，从而引起主机 A 的 TCP 重传。注意到当 TCP 重传数据时，数据最终将回到曾经丢弃它的以太网适配器。因此，从这种意义上来说，以太网的确重传了数据，尽管以太网并不知道它是正在传输一个具有全新数据的全新数据报，还是一个包含已经被传输过至少一次的数据的数据报。

2 以太网技术

吉比特以太网：IEEE 802.3z

基于交换机的以太局域网中，一台交换机和一个结点能够在同时向对方发送帧而不干扰，就是说任何时候决不会向相同的接口转发超过一个帧.

5.4.3 链路层交换机

交换机的转发功能.

交换机对子网中的主机和路由器是透明的：主机/路由器向其他传输一个帧，并不知道交换机会接收并转发.

1 交换机转发和过滤

- 过滤：决定一个帧应该转发还是丢弃.
- 转发：决定一个帧应该被导向哪个接口.
- 过滤和转发借助于交换机表完成，交换机表表项：

  一个MAC地址，通向该MAC地址的交换机接口，表项放置在表中的时间.
- 交换机转发分组基于MAC地址而不是IP地址.

交换机过滤和转发的工作流程，假定目的地址 `DD-DD-DD-DD-DD-DD` 的帧从接口 x 到达，交换机在表中查找，可能有下列三种情况：

- 表中没有对于 `DD-DD-DD-DD-DD-DD` 的表项，此时交换机向除了接口 x 外的所有接口前面的输出缓存转发 该帧的副本. 即没有该目的地址的表项，就广播该帧.
- 表中有一个表项将 `DD-DD-DD-DD-DD-DD` 与接口 x 关联. 此时说明帧从包含目的地址为 `DD-...` 的适配器来，所以这里 过滤，直接丢弃即可.
- 表中有一个表项将 `DD-DD-DD-DD-DD-DD` 与接口 $y \neq x$ 关联. 此时转发，转发到与接口 y 相连的局域网网段.

交换机表如果是完整，准确的，则无需任何广播就向着目的地转发帧. 如何配置交换机表?

2 自学习

交换机表是自动，动态和自治地建立的，交换机是自学习的.

1. 交换机表初始为空.
2. 对于每个接口接收到的每个入帧，需要存储：

   1. 帧的源 MAC地址.
   2. 帧到达的接口.
   3. 当前时间.

   用这种方式在它的表中记录了发送结点所在的局域网网段.
3. 一段时间后 (老化期)，交换机没有接受到以改地址为源地址的帧，就在表中删除这个地址.

交换机是即插即用设备，因为他们不需要网络管理员或用户做其他事.

交换机也是双工的，意味着任何交换机接口能够同时发送和接收.

3 链路层交换机的性质

- 消除碰撞：所以交换机的最大聚合带宽是该交换机所有接口速率之和.
- 异质的链路：交换机将链路彼此隔离，因此局域网中的不同链路能够以不同的速率运行，并且能够在不同的媒体上运行.
- 管理：易于进行网络管理.

交换机毒化

攻击行为：它向交换机发送大量的具有不同 伪造源MAC地址的分组．用伪造表项填满了交换机表．

嗅探：如果交换机转发，仅向主机B发送帧时，不能嗅探到，但是广播时候可以嗅探到那些不是明确寻址的帧．

4 交换机和路由器比较

路由器是使用网络层地址转发分组的 存储转发分组交换机．

交换机与路由器是根本不同的，因为它用MAC地址转发分组．

交换机是第二层的分组交换机，路由器是第三层的分组交换机．

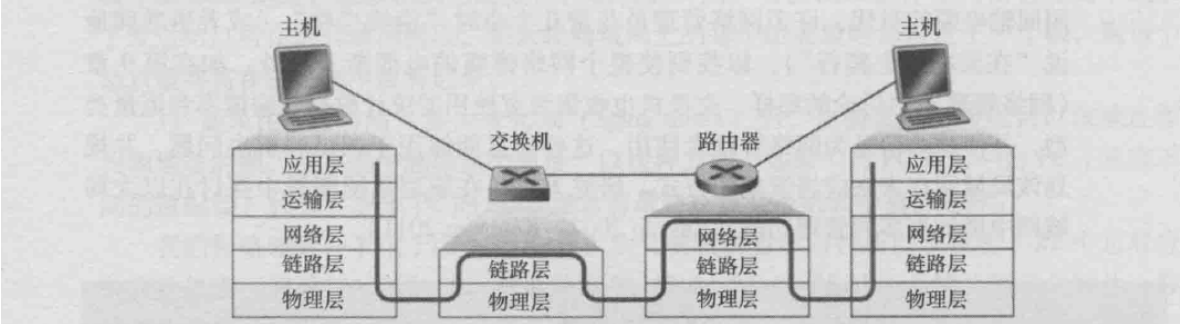

图 5-24 在交换机、路由器和主机中分组的处理

- 交换机的优点：
  - 即插即用．
  - 能具有相对高的分组过滤和转发速率．
  - 如图5-24，交换机必须处理高至第二层的帧．
- 交换机的缺点：
  - 为了防止广播帧循环，交换网络的活跃拓扑限制为一棵生成树．
  - 交换机对于广播风暴并不提供保护措施，如果某主机出现故障并输出无限的广播帧流，交换机将转发所有并使以太网崩溃．
- 路由器的优点：
  - 网络寻址是分层的，不向MAC寻址那样是扁平的．
  - 由上拓扑结构不会被限制为生成树，路由器表被正确配置时，分组不会通过路由器循环，并可以使用源和目的地之间的最佳路径．
  - 在规模更大的网络中，路由器提供了更健壮的流量隔离方式和对广播风暴的控制，并在主机之间选择更"智能"的路由．
- 路由器的缺点：
  - 不是即插即用的，路由器连接到它们的主机都需要人为地配置IP地址．
  - 路由器对每个分组的处理时间通常比交换机更长．因为需要处理高达第三层的字段．

几百台主机组成的小网络通常有几个局域网网段．此时交换机就足够了．但是在由几千台主机组成的更大网络中，通常还包含路由器．

5.4.4 虚拟局域网

现代机构的局域网通常配置为 等级结构．每个组有交换局域网，经过交换机和其他组互联．但是这样的结构会带来一些问题：

- 缺乏流量隔离．广播流量(ARP，DHCP，交换机广播帧)仍然必须横跨整个网络．
- 交换机在组人数少的时候组间交换机可能是多余的．
- 管理用户不方便．如果用户在组间移动，需要改变物理布线．

支持虚拟局域网 (VLAN) 的交换机可以解决这些问题的．

支持VLAN的交换机允许经一个单一的物理局域网基础设施定义多个虚拟局域网．在一个VLAN内的主机彼此通信，仿佛它们与交换机连接．交换机的接口由管理员划分为组，每个组构成一个VLAN．实现方面就是交换机中位于一张端口到VLAN的映射表．

两个组的流量不是完全隔离的，VLAN交换机的一个端口与一台外部的路由器相连，并将该端口配置为属于这两个组的，所以就是相当于两个组具有分离的 经路由器连接的 交换机．数据报先到一个组的交换机，再到路由器，再到另一个组的交换机．