

# 第5章 链路层

## 本章学习目标

- 理解支持数据链路层服务的原则:
  - 差错检测, 纠错
  - 共享广播信道: 多路访问
  - 链路层编址
  - 局域网:以太网, 虚拟局域网VLAN
- 各种链路层技术实例与实现

## 第5章 链路层

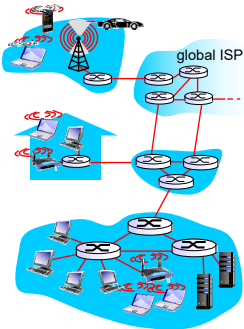
- 5.1 概述与服务
- 5.2 差错检测和纠错
- 5.3多路访问协议
- 5.4 局域网
  - 链路层寻址和ARP
  - 以太网
  - 链路层交换机
  - 虚拟局域网
- 5.5 PPP

## 链路层: 概述

### 某些术语:

- 主机和路由器: **节点**
- 沿通信路径的相邻节点的连接:
  - 通信链路**
    - 有线链路
    - 无线链路
    - 局域网
- 第二层的分组叫**帧**, 封装数据报

数据链路层能够: 经一条链路, 从一个节点传输数据到相邻节点



## 链路层: 相关内容

- 使用不同的链路协议, 经过不同的链路, 传输数据报:
    - 如第一段链路是以太网, 中间链路是帧中继, 最后链路是 802.11
  - 每个链路协议提供不同的服务
    - 例如: 可能或不能经链路提供可靠数据传输
- 运输类比**
- 从美国普林斯顿到瑞士洛桑的旅行
    - 轿车: 普林斯顿到JFK机场
    - 飞机: JFK机场到日内瓦机场
    - 火车: 日内瓦到洛桑
  - 旅行者 = **datagram**
  - 运输各段 = **通信链路**
  - 运输模式 = **链路层协议**
  - 旅行代理人 = **选路算法**

## 链路层服务

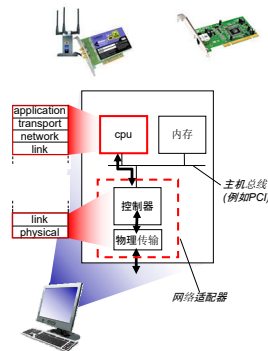
- 封装成帧:**
  - 将数据报封装进帧, 加上首部和尾部
  - 用特殊比特流实现帧同步 (涉及透明传输问题)
- 链路访问:**
  - 如果共享媒体: 多路访问, 协调多个节点的帧传输
  - 位于帧首部的 “MAC” 地址标识源、目的地
- 相邻节点间的可靠交付**
  - 我们已经知道如何做了 (传输层)!
  - 在比特差错低的链路很少使用 (光纤, 某些双绞线)
  - 无线链路: 高差错率
    - 同时使用链路级和端到端可靠性

## 链路层服务(续)

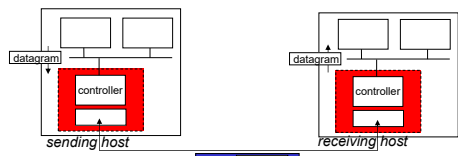
- 流量控制:
  - 发送节点和接收节点之间“步调”一致
- 差错检测:
  - 差错由信号衰减、噪声所致
  - 接收方检测差错的存在
    - 信号发送方负责重传或丢弃帧
- 纠错:
  - 接收方识别和纠正比特差错，而不采取重传
- 半双工和全双工通信控制:
  - 使用半双工，链路的两端节点能够传输数据，但只能交替进行
  - 使用全双工，链路的两端节点能够同时双向传输数据

## 链路层在何处实现？

- 每一个主机
- 链路层主体部分在“适配器”中实现的 (又称网络接口卡 NIC)
  - 核心是链路层控制器 (实现链路层功能：成帧、链路接入、差错检测等)
  - 以太网卡，PCMCIA卡，802.11卡
- 与主机系统总线连接
- 是硬件和软件的结合



## 适配器通信

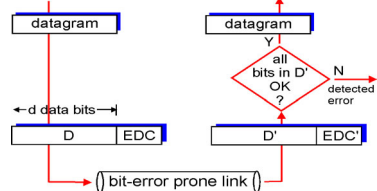


- 发送侧:
  - 将数据报封装在帧中
  - 增加差错检测比特，可靠数据传输，流量控制等
- 接收侧:
  - 检测错误，可靠数据传输，流量控制等
  - 提取数据报，传递到接收节点

## 第5章 链路层

- 5.1 概述与服务
- 5.2 差错检测和纠错
- 5.3 多路访问协议
- 5.4 局域网
  - 链路层寻址和ARP
  - 以太网
  - 链路层交换机
  - 虚拟局域网
- 5.5 PPP

## 差错检测的基本原理



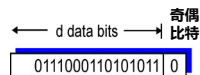
EDC= 差错检测和纠错 bits (冗余码)

D =需要进行差错检测的数据，可能包括首部字段

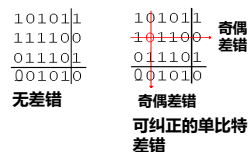
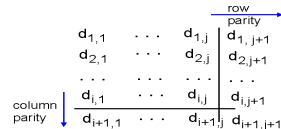
- 差错检测不是100%可靠!
  - 虽然发生的概率很低，但协议确实有可能漏掉某些差错
  - 较大的EDC字段能产生更好的检测和纠正

## 奇偶校验

单比特奇偶校验:  
检测单个比特差错



二维比特奇偶校验:  
检测和纠正单个比特差错



## 互联网传输层校验和回顾

### 发送方:

- 将待校验内容划分为16比特整数序列来处理
- 求和: 对16比特序列补码求和(最高位进位1回卷至最低位相加)
- 校验和: 求和结果取反
- 发送方将校验和的值放入UDP、TCP的校验和字段

### 接收方:

- 对接收到的数据计算校验和
- 检查计算的校验和是否等于校验字段的值:
  - 不等 – 检测到差错
  - 相等 – 没有检测到差错. 尽管如此, 还可能有错. 详情见后....

## 循环冗余码校验和

- 将数据比特D看作一个二进制数
- 选择r+1比特模式(生成式), G
- 目标: 选择r个CRC比特R, 使得
  - $\langle D, R \rangle$  能够被G整除 (以2为模)
  - 接收方知道G, 用 $\langle D, R \rangle$ 除以G, 余数全0: 没有差错; 否则有错!
  - 能够检测所有小于r+1比特的突发差错
- 广泛用于实践中 (以太网、802.11、WiFi、ATM, HDCL)



## 循环冗余校验码例子

### 希望:

$$D \cdot 2^r \text{ XOR } R = nG$$

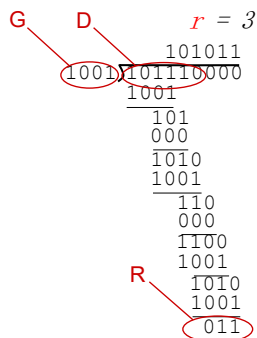
### 等价:

$$D \cdot 2^r = nG \text{ XOR } R$$

### 等价:

如果我们用 $D \cdot 2^r$ 除以G, 余数为 R

$$R = \text{余数} \left[ \frac{D \cdot 2^r}{G} \right]$$



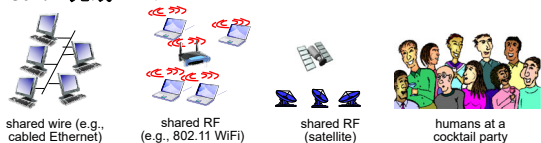
## 第5章 链路层

- 5.1 概述与服务
- 5.2 差错检测和纠错
- 5.3 多路访问协议
- 5.4 局域网
  - 链路层寻址和ARP
  - 以太网
  - 链路层交换机
  - 虚拟局域网
- 5.5 PPP

## 多路访问链路、协议

### 两类“链路”:

- 点对点链路
  - 用于拨号接入的PPP
  - 在以太网交换机和主机之间的点对点链路
- 广播链路(共享线路或媒体)
  - 传统的以太网
  - 向上游的HFC
  - 802.11无线LAN



## 多路访问控制 (MAC) 协议

- 单一共享广播信道
- 节点之间两个或更多的并行传输: 干扰
  - 碰撞 如果节点同时接收到两个或更多信号
  - 碰撞会导致接收失败。
- 多路访问控制协议
  - 采用分布式算法决定节点如何共享信道, 例如: 决定节点何时能够开始传输信号
  - 共享信道的通信必须通过信道本身传输!
    - 不用带外信道来协调

## 一个理想的多路访问控制协议

### 给定一个：速率为R bps的广播信道

1. 当只有一个节点传输数据时，它能够以速率R发送
2. 当M个节点要传输数据时，每个节点能以平均速率R/M发送
3. 完全分散控制：
  - 无特殊节点来协调传输
  - 无同步时钟、时隙
4. 简单

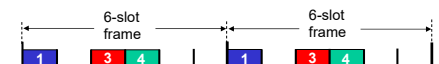
## MAC协议: 分类

- 按信道划分
  - 将信道划分为较小的“段”（时隙，频率，编码）
  - 为节点分配一部分专用
  - 如TDMA\FDMA\CDMA\WDM等多路复用技术
- 随机访问
  - 不划分信道，允许碰撞
  - 必须能从“碰撞”中恢复
- “轮流”访问
  - 节点轮流使用信道，但有更多数据要发送的节点能够获得更长的轮流时间。

## 信道划分MAC协议: TDMA

### TDMA: 时分多路访问

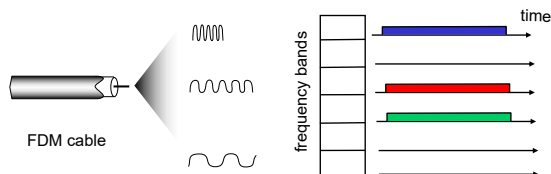
- “周期性”访问信道
- 每个站点在每个周期中获得固定长度的时隙(长度=分组传输时间)
- 不使用的时隙进入空闲状态
- 例子：6个站点的LAN, 时隙1、3、4有分组, 时隙2、5、6空闲



## 信道划分 MAC协议: FDMA

### FDMA: 频分多路访问

- 信道频谱划分为若干带
- 每个站点分配一个固定的频带
- 未有数据传输的频带为空闲状态
- 例子：6个站点的LAN, 频带1、3、4有分组, 频带2、5、6空闲



## 随机访问MAC协议

- 当站点有分组要发送
  - 以信道全部速率R传输
  - 节点间无优先权协调
- 两个或更多节点传输数据→“碰撞”，
- 随机访问MAC协议定义了：
  - 如何检测碰撞
  - 如何从碰撞中恢复(例如，一段延迟后重新传输)
- 典型的随机访问MAC协议：
  - 时隙ALOHA
  - 纯ALOHA
  - CSMA, CSMA/CD, CSMA/CA

## 时隙ALOHA

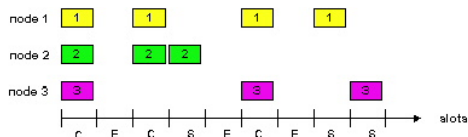
### 假定

- 所有帧有相同长度
- 时间划分为等长时隙（时隙长度足够传输1个帧）
- 节点仅在时隙开始时开始传输帧
- 节点必须时间同步
- 如果2个或多个节点在时隙中传输，所有节点将检测到碰撞

### 操作

- 当节点获得新帧，将在下一个时隙中传输
  - 如果无碰撞，节点能够在下一个时隙继续发送新帧
  - 如果碰撞，节点在每个后继时隙中以概率p重传帧，直到发送成功

## 时隙ALOHA



### 优点

- 单个活跃节点能够连续以信道的最高速率全速传输
- 高速分散：仅节点中的时隙需要同步
- 简单

### 缺点

- 碰撞，浪费时隙
- 存在空闲时隙
- 节点必须能够在小于分组传输时长的时间内检测到碰撞的发生
- 需要时钟同步

## 时隙Aloha的效率

**效率：**当多个节点都有许多帧要发送时，成功时隙与总时隙的长的比值

- 假定N个节点有很多帧要发送，每个节点在每个时隙以概率 $p$ 发送数据
- 对于一个给定节点，在一个时隙中成功发送的概率 =  $p(1-p)^{N-1}$
- 任何节点成功发送的概率 =  $Np(1-p)^{N-1}$

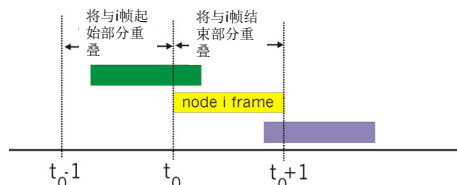
### 如何使效率最大化？

- 求 $p^*$  使得 $Np(1-p)^{N-1}$ 最大
- 对许多节点，当N趋近无穷大，取 $Np^*(1-p^*)^{N-1}$  极限，得到 $1/e = 0.37$

**最好情况：**信道用于有效数据传输的时间是37%!

## 纯ALOHA

- 非时隙Aloha: 更简单，无同步要求
- 当节点有新帧生成，立即传输
- 发生碰撞，节点以概率 $p$ 重传该帧，否则等待一个帧的传输时间；之后，节点以概率 $p$ 传输该帧，或者以概率 $1-p$ 在另一个帧时间等待。
- 显然碰撞的概率增加，效率较时隙ALOHA降低:  $1/(2e) \approx 18\%$ 
  - 在 $t_0$  发送的帧与在 $[t_0-1, t_0+1]$ 发送的其他帧碰撞



## CSMA (载波侦听多路访问)

**CSMA:** 在传输前侦听:

- 当一个站检测到的信号电压摆动值超过一定的门限值时，就认为总线上至少有两个站同时在发送数据，表明产生了碰撞。

## CSMA(载波侦听多路访问)

**CSMA:** 在传输前侦听:

- 侦听到信道空闲: 传输整个帧

- 侦听到信道忙, 推迟传输

- 碰撞还是会发生:

- 传播延迟意味着两个节点可能无法在准备发送数据时“听到”对方正在进行数据传输

- 碰撞发生后继续发送数据帧:

浪费信道资源



## CSMA/CD (载波侦听多路访问/带碰撞检测)

- **CSMA/CD:**

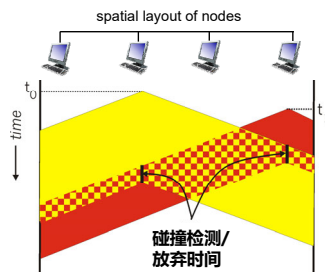
- 在短时间内检测到碰撞

- 发生碰撞的数据传输被放弃，减少信道浪费

- 碰撞检测:

- 有线局域网容易实现: 测量信号强度，比较传输的信号和接收的信号

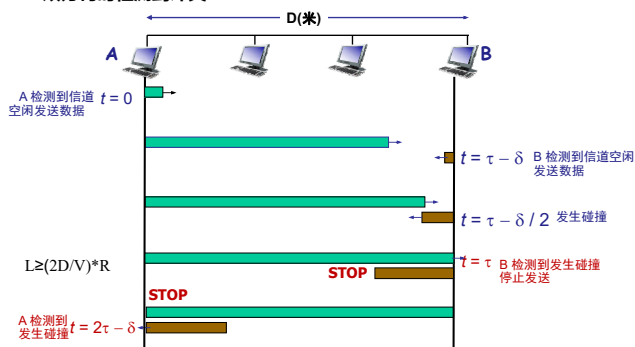
- 在无线局域网中困难: 由于信号衰减等因素，接收方接收到的信号强度不足



## CSMA/CD 协议的争用期

- 假设网络带宽: Rbps, 数据帧最小长度L(bits), 信号在信道上的传播速度: V(m/s), 通信双方距离为D(m)

- 双方何时检测到冲突?



## CSMA/CD 协议的争用期

- 端到端往返时延  $2\tau$  称为争用期, 或碰撞窗口

- 结论:

- ① 只有当A经过争用期这段时间还没有检测到碰撞, 才能肯定这次发送不会失败。
- ② A发送数据帧的最短长度L必须满足:  $L/R \geq (2D/V) \rightarrow L \geq (2D/V) * R$

- 举一个栗子

假定 1km 长的 CSMA/CD 网络的数据率为 1Gb/s。设信号在网络上的传播速率为 200000km/s。求能够使用此协议的最短帧长。

答: 对于 1km 电缆, 单程传播时间为  $1 \div 200000 = 5 \times 10^{-6}$ s, 即 5us, 来回路程传播时间为 10us。为了能够按照 CSMA/CD 工作, 最短帧的发射时间不能小于 10us。以 1Gb/s 速率工作, 10us 可以发送的比特数等于:

$$\frac{10 \times 10^{-6}}{1 \times 10^{-9}} = 10000$$

因此, 最短帧是 10000 位或 1250 字节长。

## “轮流”访问MAC协议

- 信道划分 MAC协议:

- 在高负载时高效、公平地共享信道
- 低负载时低效: 信道访问延时, 当仅有1个活跃节点传输数据时, 也仅分配了 1/N 带宽!

- 随机访问 MAC协议

- 低负载时效率高: 单个节点能够完全利用信道资源
- 高负载: 碰撞开销

- “轮流”访问MAC协议

- 兼有这方面的优点!

## “轮流”访问 MAC协议

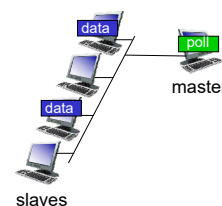
轮询:

- 主节点“邀请”从节点依次传输数据

- 通常使用“沉默”的从设备

- 存在问题:

- 轮询开销
- 等待时延
- 单点故障(主节点)



## “轮流”访问 MAC协议

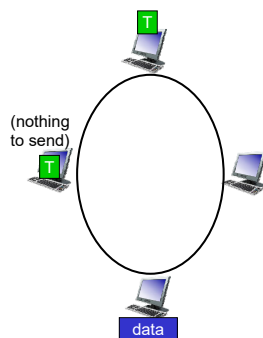
令牌传递:

- 控制令牌从一个节点顺序地传递到下一个。

- 令牌: 一种特殊的数据帧

- 存在问题:

- 令牌开销
- 等待时延
- 单点故障(令牌)



## MAC协议小结

- 信道划分: 通过时间、频率或编码
- 随机访问 (动态的)
  - ALOHA, S-ALOHA, CSMA, CSMA/CD
  - CSMA/CD 用在以太网中
  - CSMA/CA 用在 802.11 中
- 轮流
  - 主节点轮询; 令牌传递
  - 典型应用: 蓝牙、FDDI、令牌环

## 第5章 链路层

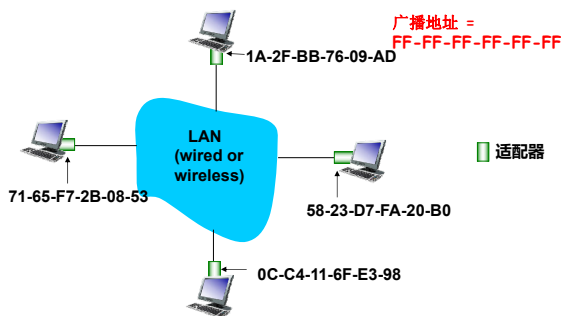
- 5.1 概述与服务
- 5.2 差错检测和纠错
- 5.3 多路访问协议
- 5.4 局域网
  - 链路层寻址和ARP
  - 以太网
  - 链路层交换机
  - 虚拟局域网
- 5.5 链路虚拟化：MPLS
- 5.6 数据中心网络
- 5.7 回顾：Web页面请求历程

## MAC地址和ARP

- 32-bit IP地址：
  - 网络层地址
  - 用于使数据报到达目的IP子网，支持分组转发
- MAC(或LAN 或物理或以太网)地址：
  - 用于使数据报从一个物理接口到达另一个物理连接的接口(同一个网络内)
  - 48 bit MAC地址(对多数LAN) 固化于适配器ROM中，有些可以软件设置
- 例如：b0-df-c1-20-49-b0

## LAN 编制和ARP

在LAN中的每块适配器具有独一无二的LAN地址

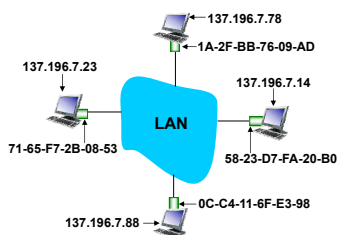


## LAN地址(续)

- MAC地址分配由IEEE管理
- 制造商购买部分MAC地址空间(前24比特，确保惟一性)
- 类比：
  - MAC地址：像居民身份证号
  - IP地址：像邮政地址
- MAC 扁平地址 → 可移动性
  - 能够将LAN卡从一个LAN移动到另一个LAN
- IP层次地址
  - 取决于节点所在的子网

## ARP: 地址解析协议

问题: 在同一个局域网内，已知节点的IP地址怎样获取节点的MAC地址?



LAN上的每个IP节点(主机、路由器)都有ARP表:

- IP/MAC地址映射 < IP地址; MAC地址; TTL >
- TTL(寿命): 地址映射过期的时间(通常20分钟)

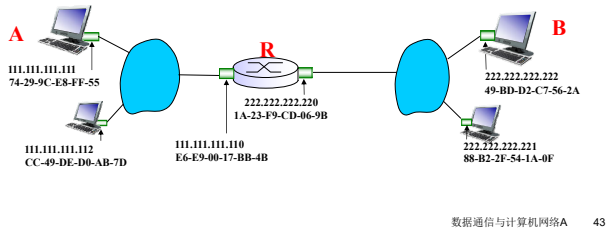
## ARP协议: 相同的LAN (网络)

- 主机A要向同一个局域网内的主机B发送数据报，并且B的MAC地址不在A的ARP表中：
  - A广播ARP 请求分组, 包含B的IP地址
    - 目的地MAC地址 = FF-FF-FF-FF-FF-FF
    - 在 LAN上的所有机器接收ARP请求
  - B接收ARP分组，用它的MAC地址回答 A
    - 应答帧发送到A的MAC地址 (单播)
- A在它的ARP表中缓存(保存) IP-MAC地址对，直到信息超时
  - 超时后被删除
- ARP是“即插即用”：
  - 节点创建它们的ARP表无需网络管理员干预



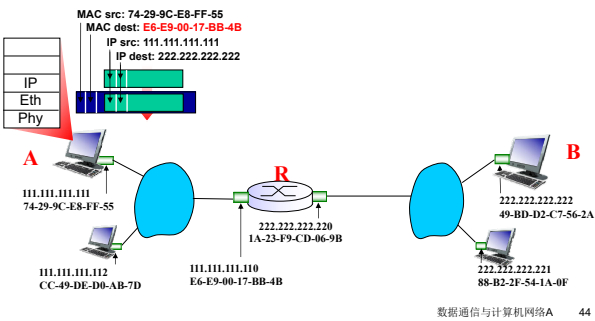
选路到另一个LAN

- 目的: 主机A经路由器R向主机B发送数据报
- 假设A知道B的IP地址
  - 假设A知道第一跳路由器R的IP地址
  - 假设A知道R的MAC地址



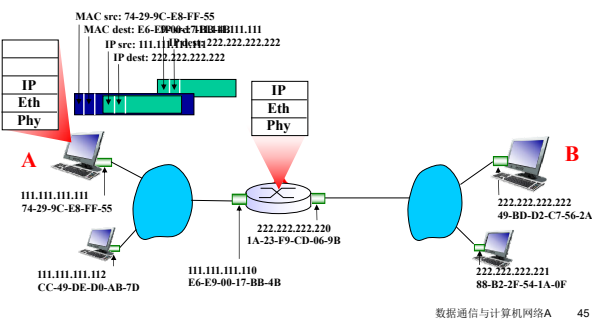
选路到另一个LAN

- A生成具有源IP地址A、目的IP地址B的数据报
- A使用ARP解析得到R的左侧接口 (111.111.111.110) 的MAC地址E6-E9-00-17-BB-4B，生成数据帧



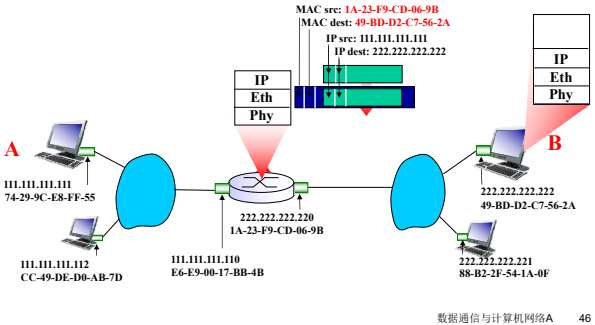
选路到另一个LAN

- A发送数据帧到R
- R接收数据帧，解封装数据报，发送到网络层



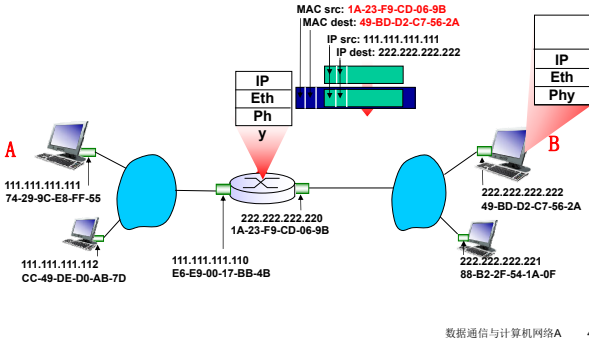
选路到另一个LAN

- R以源 IP地址A，目的地址B转发数据报
- R 使用B的MAC地址为目的地址，创建数据帧封装A-to-B IP数据报



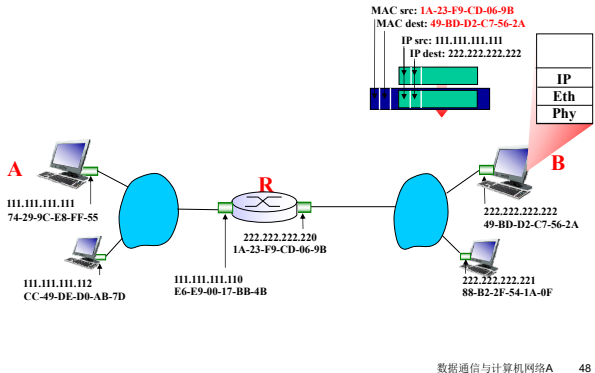
选路到另一个LAN

- R以源 IP地址A，目的地址B转发数据报
- R 使用B的MAC地址为目的地址，发送数据帧给B



选路到另一个LAN

- B接收数据帧，解封装数据报，发送到网络层





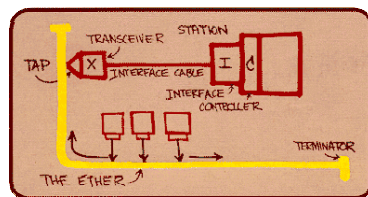
## 第5章 链路层

- 5.1 概述与服务
- 5.2 差错检测和纠错
- 5.3 多路访问协议
- 5.4 局域网
  - 链路层寻址和ARP
  - 以太网
  - 链路层交换机
  - 虚拟局域网
- 5.5 PPP

## 以太网

“占统治地位”的有线LAN技术:

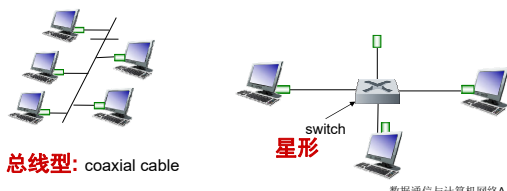
- 100Mbps网卡百元以内
- 最先广泛使用的LAN技术
- 比令牌LAN和ATM更便宜、配置更简单
- 满足用户对网络速率的要求: 10 Mbps – 10 Gbps



Metcalfe的以太网草图

## 以太网拓扑结构

- 总线拓扑: 流行于20世纪90年代, 是广播式局域网, 所有节点都在同一个冲突域
- 星型拓扑: 目前主流结构, 以集线器或交换机为中心
  - 基于集线器的星型结构以太网, 本质上还是广播式局域网
  - 基于交换机的星型结构以太网, 每个交换机接口都是一个单独的冲突域!

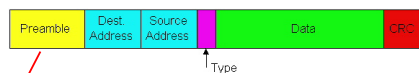


## 以太网提供的服务

- 以太网提供的服务是**不可靠交付**, 即**尽最大努力的交付**.
  - **无连接**: 在发送和接收适配器之间没有握手
  - **不可靠**: 接收适配器不向发送适配器发送确认或否定确认
- 当接收方收到有差错的数据帧时就丢弃此帧, 其他什么也不做。差错的纠正由高层来决定。
- 如果高层发现丢失了一些数据而进行重传, 但以太网不知道这是一个重传的帧, 而是当作一个新的数据帧发送。

## 以太网V2的帧结构

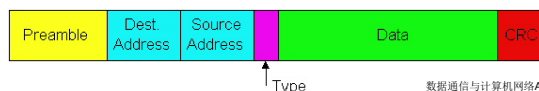
- 发送适配器在以太网帧(或其他网络层协议分组)中封装IP数据报



- **前同步码 (8字节)**:
  - 模式为前7个字节是10101010, 第8个字节为 10101011
  - 用于同步接收方, 发送方时钟速率

## 以太网V2的帧结构(续)

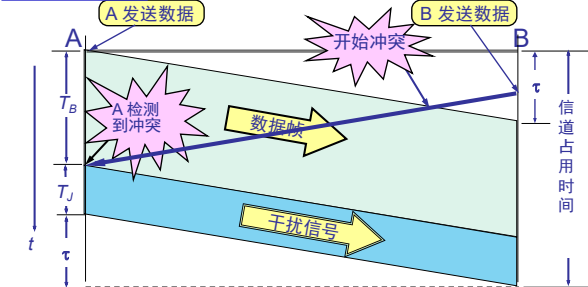
- **源/目的地址 (各6字节)**:
  - 如果适配器接收具有匹配的目的地地址或广播地址(如ARP分组)的帧, 它将帧中的数据提交给网络层协议
  - 否则, 适配器丢弃帧
- **类型 (2字节)**: 指示帧中封装的是哪个高层协议分组 (大多数为IP 但也可以支持其他类型如 Novell IPX和AppleTalk)
- **数据 (46~1500字节)**: 指示上层协议载荷, 有效的MAC帧长度64 字节~1518字节。
- **CRC (4字节)**: 在接收方核对, 如果检测到差错, 该帧就被丢弃



回顾：CSMA/CD 协议的争用期

- 端到端往返时延  $2\tau$  称为争用期，或碰撞窗口
- 结论：
  - ①只有当A经过争用期这段时间还没有检测到碰撞，才能肯定这次发送不会失败。
  - ②A发送数据帧的最短长度L必须满足： $L/R \geq (2D/V) \rightarrow L \geq (2D/V) * R$
- 10Mbps以太网争用期定义为51.2μs
  - 理想情况下，电磁波在电缆中的传输速率为  $2 * 10^8$  km/s
  - 早期的10Mbps以太网最多只能有5个网段，4个中继器，而其中只允许3个网段有设备，其他两个只是传输距离的延长，因此10Mbps以太网连接的极限距离是2500米。
  - 所以往返时延25μs+中继器的额外时延，预估时延取为45μs，再加上48bit的人为干扰信号（强化碰撞）耗时4.8μs，所以IEEE将10Mbps以太网的争用期定义为51.2μs。
  - $10\text{Mbps} * 51.2\mu\text{s} = 512\text{bit}$ ， $512/8 = 64$ 字节，即最小帧长度64字节

强化碰撞



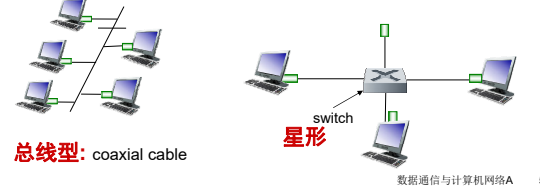
- 发送数据的站点检测到冲突时：
  - 立即停止发送数据；
  - 再继续发送若干比特（32或48bit）的人为干扰信号(jamming signal)，以便让所有用户都知道现在已经发生碰撞。

最小帧间间隔

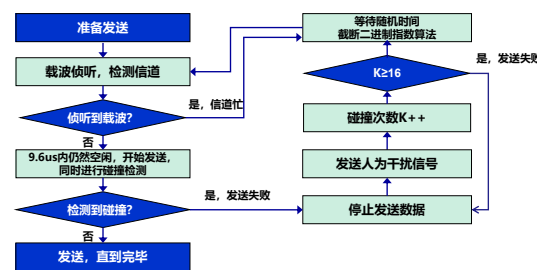
- 以太网还规定帧间最小间隔为9.6 μs（对于10Mbps以太网相当于96比特时间）。
- 站点在检测到总线开始空闲后，还要等待 9.6 μs 才能再次发送数据。
- 目的：为了使刚刚接收到数据帧的接收方的缓存来得及清理，做好接收下一帧的准备。

以太网的CSMA/CD协议

- 总线拓扑和使用集线器的星型拓扑以太网，显然多个节点同时发送时会出现帧碰撞，因此对于这种广播式以太网，使用CSMA/CD协议
- 对于基于交换机的星型结构以太网，采用的是存储转发分组交换，且现代交换机是全双工，不会再有碰撞发生，无需使用CSMA/CD协议



以太网的CSMA/CD协议



截断二进制指数退避算法

- 发生碰撞的站点在停止发送数据后，要推迟（退避）一个随机时间才能再发送数据。
  - 确定基本退避时间，一般为争用期  $2\tau$ 。
  - 从整数集合  $[0, 1, \dots, (2^k - 1)]$  中随机地取出一个数，记为  $r$ 。退避等待时间 =  $r$  倍的基本退避时间。
  - 参数  $k$  的计算公式： $k = \text{Min}[\text{重传次数}, 10]$
  - 当重发达 16 次仍不能成功时丢弃该帧，并向高层报告。

## 截断二进制指数退避算法

□ 例如：

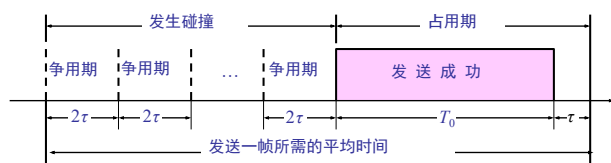
- 第1次碰撞发生后，每个站点等待0或1个争用期后重发；
- 第2次发生碰撞后，等待时间从0、1、2、3个争用期中随机选一个；
- 第3次发生碰撞后，等待时间从0~2<sup>3-1</sup>个争用期之间随机选一个；
- 检测 i 次碰撞后，等待时间从0~2<sup>i-1</sup>个争用期之间随机选一个；
- 检测10次碰撞后，等待时间就从0到1023个争用期之间随机选一个；
- 检测16次碰撞后，控制器不再动作，并宣告发送失败。

## 以太网的信道利用率

□ 以太网的信道被占用的情况：

- 争用期长度为  $2\tau$ ，即端到端传播时延的两倍。假设检测到碰撞后不发送干扰信号。
- 帧长为  $L$  (bit)，数据发送速率为  $C$  (b/s)，因而帧的发送时间为  $L/C = T_0$  (s)。

□ 一个帧从开始发送，经过可能发生的碰撞后，重传多次，直到发送成功且信道转为空闲(即再经过时间  $\tau$  使得信道上无信号在传播)时为止，是发送一帧所需的平均时间。



## 参数 a

□ 要提高以太网的信道利用率，就必须减小  $\tau$  与  $T_0$  之比。在以太网中定义参数  $a$ ，它是以太网单程端到端时延  $\tau$  与帧的发送时间  $T_0$  之比：

$$a = \frac{\tau}{T_0}$$

- $a \rightarrow 0$  表示一发生碰撞就立即可以检测出来，并立即停止发送，因而信道利用率很高。
- $a$  越大，表明争用期所占的比例增大，每发生一次碰撞就浪费许多信道资源，使得信道利用率明显降低。

## 极限信道利用率 Smax

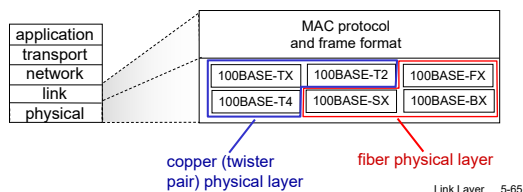
- 当数据率一定时，以太网的连线长度受到限制，否则  $\tau$  的数值会太大。
- 以太网的帧长不能太短，否则  $T_0$  的值会太小，使  $a$  值太大。
- 在理想情况下，以太网上的各站发送数据都不会产生碰撞，即总线一旦空闲就有某一个站立即发送数据。
- 发送一帧占用线路的时间是  $T_0 + \tau$ ，而帧本身的发送时间是  $T_0$ 。于是理想情况下极限信道利用率  $S_{\max}$  为：

$$S_{\max} = \frac{T_0}{T_0 + \tau} = \frac{1}{1 + a}$$

## 不同的以太网技术

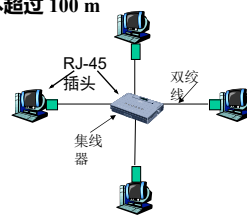
□ 许多不同的以太网技术

- 通用的MAC协议和帧格式
- 不同的速率: 2 Mbps, 10 Mbps, 100 Mbps, 1Gbps, 10G bps
- 不同的物理介质: 光纤, 电缆



## 不同的以太网技术

- 传统以太网最初使用粗同轴电缆，后来演进到使用较便宜细同轴电缆，最后发展为使用更便宜和灵活的双绞线。
- T 表示双绞线(Twisted Pair)，如10Base-T和100Base-T；10BASE-T的通信距离稍短，节点连接到一台集线器：“星型拓扑”；每个站到集线器的距离不超过 100 m
- 10/100 Mbps速率，后者被称为“快速以太网”
- 10BASE-T 无屏蔽双绞线星形以太网的出现，既降低成本，又提高可靠性为以太网在局域网中统治地位奠定牢固的基础。



## 第5章 链路层

- 5.1 概述与服务
- 5.2 差错检测和纠错
- 5.3 多路访问协议
- 5.4 局域网
  - 链路层寻址和ARP
  - 以太网
  - 链路层交换机
  - 虚拟局域网
- 5.5 PPP

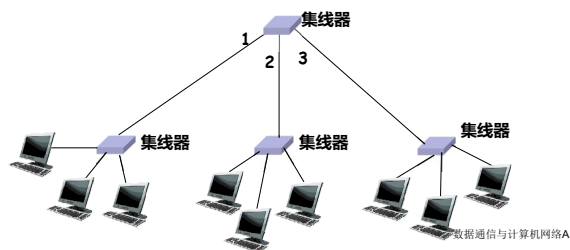
## 广播域和冲突域

- 广播域 (Broadcast Domain) 是一个逻辑上的计算机组，接收同样广播消息的节点集合。如：在该集合中的任何一个节点传输一个广播帧，则所有其他能收到这个帧的节点都被认为是该广播域的一部分。
- 所以类似Hub、二层交换机等节点都被认为是同一个广播域。而路由器，第三层交换机则连接不同的广播域。
- 冲突域 (Collision Domain) 是一种物理分段，是指连接在同一物理介质上的所有站点的集合，同一冲突域中的设备互联时，同一时刻只允许一个设备发送的数据从共享的物理通道中通过，其他设备发送的数据则要等到这个通道处于“空闲”时才可以通过，否则会出现冲突。
- 连接同一冲突域的设备有集线器，中继器或者其他进行简单复制信号的设备。

<http://www.kokojia.com/article/27609.html>

## 用集线器互联

- 主干集线器互联LAN网段
- 扩展局域网范围，但单独碰撞域会合并为一个大的碰撞域
- 不能互联使用不同以太网技术的局域网（如10BaseT和100BaseT）

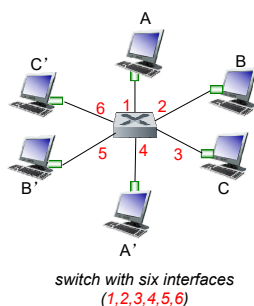


## 链路层交换机

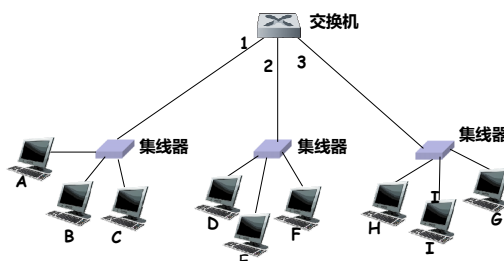
- 功能
  - 存储并转发以太网帧
  - 检查帧首部并基于MAC目的地址选择性地转发帧
- 优点：
  - 透明，主机不知道交换机的存在
  - 即插即用，自学习，交换机不必配置
- 缺点：
  - 主机和路由器ARP表较大，生成可观的ARP处理量和流量大
  - 交换机组成的网络是一个广播域，无法防范广播风暴（转发广播帧）

## 链路层交换机

- 主机直接连接到交换机接口
- 交换机输出接口设有缓存
- 全双工通信
- 交换：A-to-A'和B-to-B'可以同时发送，无碰撞（每一个交换机接口都是独立的冲突域），不再需要MAC协议



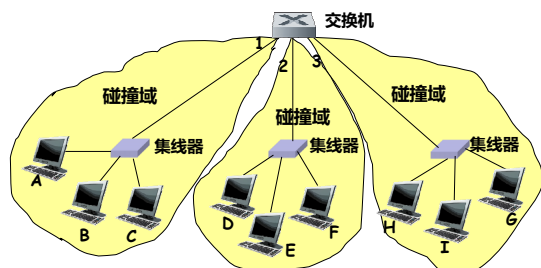
## 交换机连接LAN



- 当交换机接口直接连接了集线器，而集线器又连接了多台主机时
- 每个集线器连接的局域网需要使用MAC协议。

## 交换机连接LAN: 流量隔离

- 交换机将子网分割成LAN, LAN成为分离的碰撞域
- 交换机过滤分组:
  - LAN内的数据帧通常不在其他LAN上转发

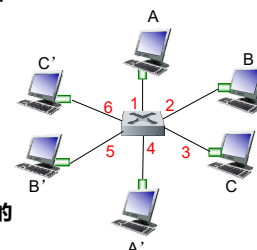


## 交换机: 转发表

- 交换机怎样决定向哪个接口转发数据帧呢?
- 一个交换机具有一个**转发表**:
  - (MAC地址, 接口, 生存期)
  - 超过生存期的表项被删除
  - 收到数据帧时, 交换机“得知”发送方在哪个接口
  - 交换机中记录了“发送方—接口”的映射关系

MAC addr	Interface	TTL
MAC-A	1	60

Switch table  
(initially empty)



## 交换机: 过滤/转发

**当交换机收到1个数据帧:**

使用目的地址MAC地址索引交换表

if 找到目的地址项

then{

if 数据帧的目的地址与此帧的来源地址相同

then 丢弃帧

else 在交换表指示的接口转发该帧

}

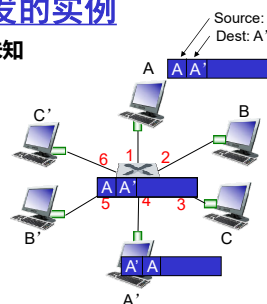
else 洪泛

向所有接口(除了该帧到达的)转发该帧

## 交换机自学习、转发的实例

- ❖ 帧的目的为A', 但接口位置未知  
使用洪泛

- ❖ 帧的目的地为A, 接口位置已知:  
有选择性发送到一个通信链路上

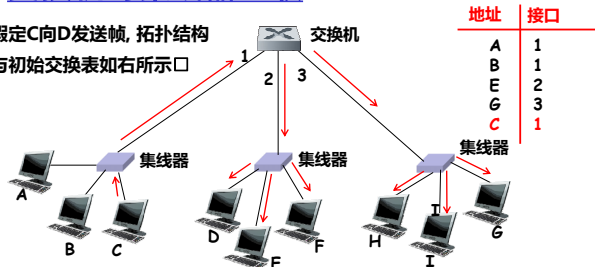


MAC addr	interface	TTL
A	1	60
A'	4	60

switch table  
(initially empty)

## 交换机与集线器互联

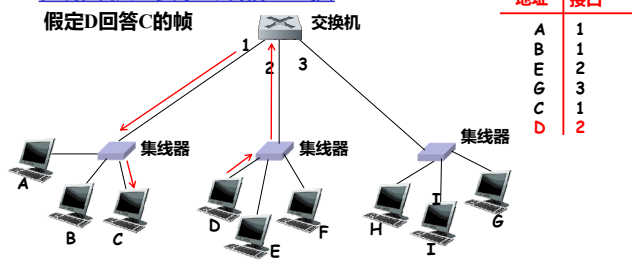
假定C向D发送帧, 拓扑结构与初始交换表如右所示



- 交换机从C接收帧
  - 注意到交换机表中C位于接口1(添加C的表项)
  - 因为D不在表中, 交换机将向接口2和3转发帧 (洪泛)
- D接收帧

## 交换机与集线器互联

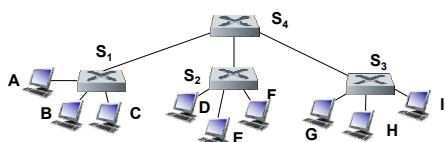
假定D回答C的帧



- 交换机从D接收帧
  - 注意到在交换机表中D位于接口2(添加表项D)
  - 因为C在表中, 交换机仅向接口1转发帧
- C接收帧

## 交换机与交换机互联

### 交换机互联的层级结构

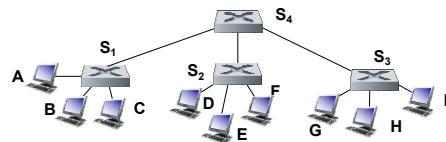


地址	接口
A	1
B	1
E	2
G	3

- 当A向G发送数据帧时，S1如何知道必须通过S4转发？而S4如何知道通过S3转发？
- 也是通过自学习生成交换表

## 交换机互联的自学习和转发

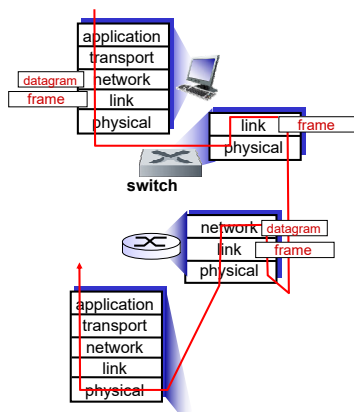
### 交换机互联的层级结构



- 练习：假设主机C发送数据帧给I，I应答C，初始转发表为空，请写出收发上述数据帧后，S1/S4的转发表

## 交换机 vs. 路由器

- 两者都是存储转发设备
  - 路由器：网络层设备(检查网络层首部)
  - 交换机：链路层设备(检查链路层首部)
- 两者都有转发表：
  - 路由器使用选路算法维护转发表，依靠IP地址
  - 交换机使用洪泛、自学习维护转发表，依靠MAC地址



## 对比小结

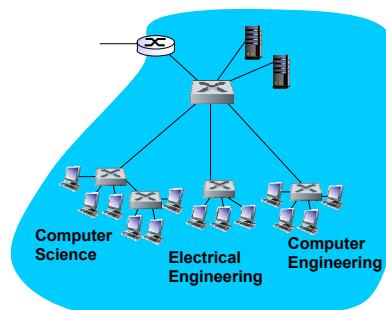
	集线器	路由器	交换机
流量隔离	no	yes	yes
即插即用	yes	no	yes
优化选路	no	yes	no
隔离广播	no	yes	no

**显然：**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，并且还在主机之间使用了更“智能”的路由，更适合于几千台甚至以上主机组成的更大规模网络

## 第5章 链路层

- 5.1 概述与服务
- 5.2 差错检测和纠错
- 5.3 多路访问协议
- 5.4 局域网
  - 链路层寻址和ARP
  - 以太网
  - 链路层交换机
  - 虚拟局域网
- 5.5 PPP

## VLAN：虚拟局域网

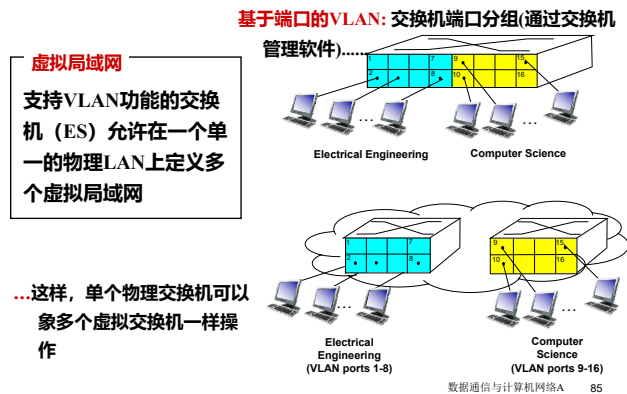


常见的机构网络配置

- 左图网络拓扑存在的问题：
  - ①管理用户：CS 用户移动到EE，但他可能希望继续连接CS的交换机？
  - ②可能造成交换机部署上的浪费（如机构内部存在多个小组，但每个小组规模很小）
  - ③单个广播域缺乏流量隔离：
    - 所有的二层广播流量（例如携带ARP和DHCP的报文、或那些目的MAC地址还未被交换机学习到的帧）仍然必须跨越整个机构网络
    - 带来安全性、效率问题

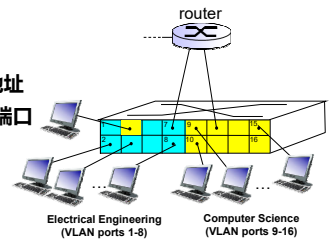


## VLAN：虚拟局域网

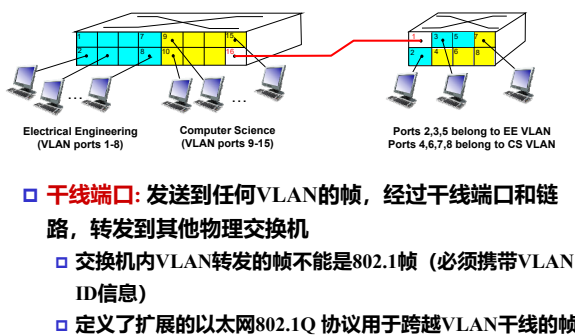


## 基于端口的VLAN

- 流量隔离：只能在属于相同VLAN的端口之间交付帧
  - 也可以根据端点的MAC地址定义VLAN，而不是交换端口
- 动态成员：端口可以动态的在VLAN之间分配
- VLAN之间转发：通过路由器完成。

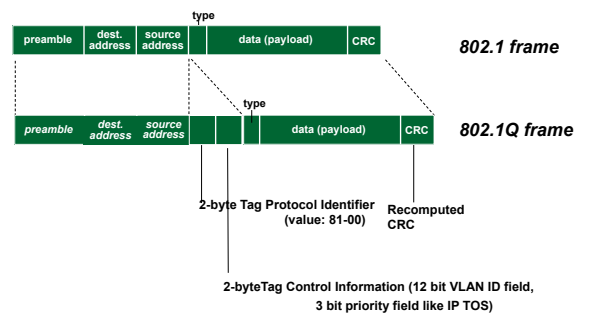


## 跨越多个多交换机的VLAN



## 802.1Q VLAN帧格式

- 标准的以太网帧+4字节的VLAN标签组成。



## 第5章 链路层

- 5.1 概述与服务
- 5.2 差错检测和纠错
- 5.3 多路访问协议
- 5.4 局域网
  - 链路层寻址和ARP
  - 以太网
  - 链路层交换机
  - 虚拟局域网
- 5.5 PPP

## 点对点链路层控制

- 一个发送方、一个接收方、一段链路：比广播链路容易处理：
  - 无需媒体访问控制
  - 不需要明确的MAC寻址
  - 如拨号链路、ISDN链路
- 流行的点对点DLC协议：
  - PPP (point-to-point协议)
  - HDLC：高级数据链路控制 (数据链路过去被认为位于协议栈的“高层”！)



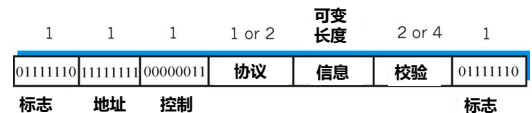
## PPP 设计要求[REC 1557]

- **分组成帧**: 在数据链路帧中封装网络层数据报
  - 同时承载任何网络层协议分组(不止是IP)
  - 向上层实现多路分解
- **比特透明传输**: 在数据字段必须承载任何比特模式
- **差错检测**: 不纠错, 丢弃差错帧
- **连接活跃性**: 检测、向网络层通告链路故障
- **支持多种类型链路**: 如PPPOE, 允许多个连接在以太网的用户共享一条到ISP的宽带链路。
- **网络层地址协商**: 网络层实体能够协商或配置彼此的网  
络地址

数据通信与计算机网络A 91

## PPP数据帧

- **标志**: 定界符(成帧)
- **地址**: 不起作用 (仅是一个选项)
- **控制**: 不起作用 ; 以后可能有多种控制字段
- **协议**: 该帧交付的高层协议 (如 PPP-LCP, IP, IPCP等)
- **信息**: 上层协议的分组数据
- **校验**: 实现差错检测的冗余循环校验



数据通信与计算机网络A 92

## PPP数据传输

- **考虑**: 发送方向接收方发送了若干数据帧, 接收方如何从比特流中区分出数据帧?
- **解决方案**: 使用**帧定界符**, PPP为**0x7E(01111110)**
- **考虑**: 如果数据字段中出现了和帧定界符01111110一样的比特流, 接收方会误判为上一个帧结束了, 如何解决?
- **解决方案**: 采用**透明传输**, 即按某种规则对数据进行“转义”

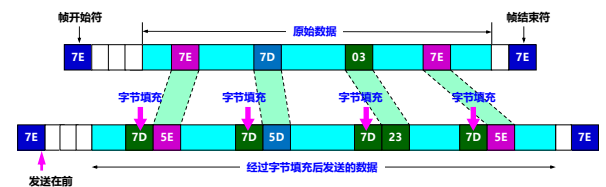
<https://tools.ietf.org/html/rfc1662#page-9>

数据通信与计算机网络A 93

## PPP异步传输的字符填充法

对于发送方的转义规则为:

- 信息字段中出现的每一个 0x7E 字节序列, 转变成成为 2 字节序列 (0x7D, 0x5E)。
- 信息字段中出现的每一个 0x7D 的字节序列, 转变成成为 2 字节序列 (0x7D, 0x5D)。
- 信息字段中出现 ASCII 码的控制字符 (即数值小于 0x20 的字符), 转变为2字节序列 ( 0x7D 字节, 该字符与0X20异或值) 。



## PPP同步传输的零比特填充

- PPP 协议用在 SONET/SDH 链路时, 使用同步传输 (一连串的比特连续传送)。这时 PPP 协议采用**零比特填充**方法来实现**透明传输**。
- 在发送端, 只要发现有 5 个连续 1, 则立即填入一个 0。
- 接收端对帧中的比特流进行扫描。每当发现 5 个连续1时, 就把这 5 个连续 1 后的一个 0 删除。



## PPP数据控制协议

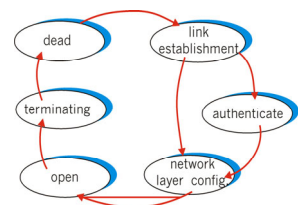
在交换网络层数据前, 数据链路

对方必须: LCP/NCP

- **配置PPP链路**(如最大的帧长, 身份鉴别协议等)

- **学习/配置网络层信息**

- 对IP: 使用IP控制协议 (IPCP)分组, 封装成PPP帧 (协议字段: 8021), 完成IP 地址等相关信息的配置



数据通信与计算机网络A 96