哈爾濱工業大學
HARBIN INSTITUTE OF TECHNOLOGY

立足航天，服务国防，面向国民经济主战场

《计算机网络》

第5章 数据链路层与局域网

主讲人：李全龙

数据链路层与局域网

- 数据链路层服务
- 差错编码
- 多路访问控制(MAC)协议
  - 信道划分MAC协议
  - 随机访问MAC协议
  - 轮转访问MAC协议
- ARP协议
- 以太网
  - 交换机
  - 虚拟局域网（VLAN）
- PPP协议
- 802.11无线局域网

数据链路层与局域网

数据链路层服务

差错编码

多路访问控制(MAC)协议

- 信道划分MAC协议
- 随机访问MAC协议
- 轮转访问MAC协议

ARP协议

以太网

- 交换机
- 虚拟局域网（VLAN）

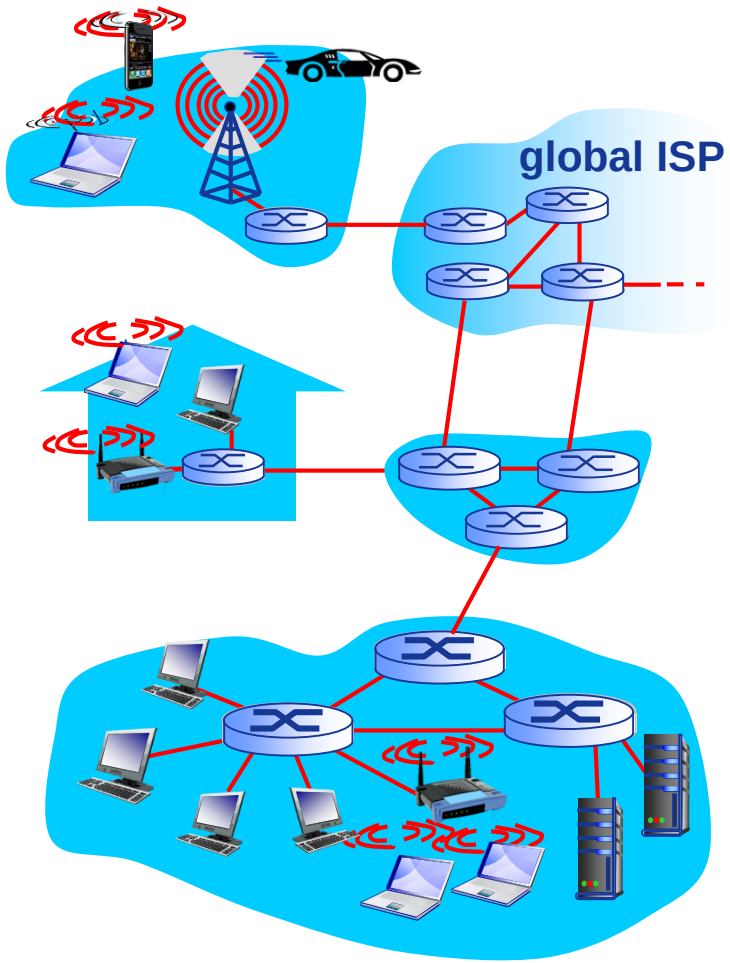
PPP协议

802.11无线局域网

概述

术语:

- 主机和路由器：结点(nodes)
- 连接相邻结点的通信信道：链路(links)
  - 有线链路(wired links)
  - 无线链路(wireless links)
  - 局域网(LANs)
- 链路层(第2层)数据分组：帧(frame)，封装网络层数据报



数据链路层负责通过一条链路从一个节点向另一个物理链路直接相连的相邻结点传送数据报。

链路层服务

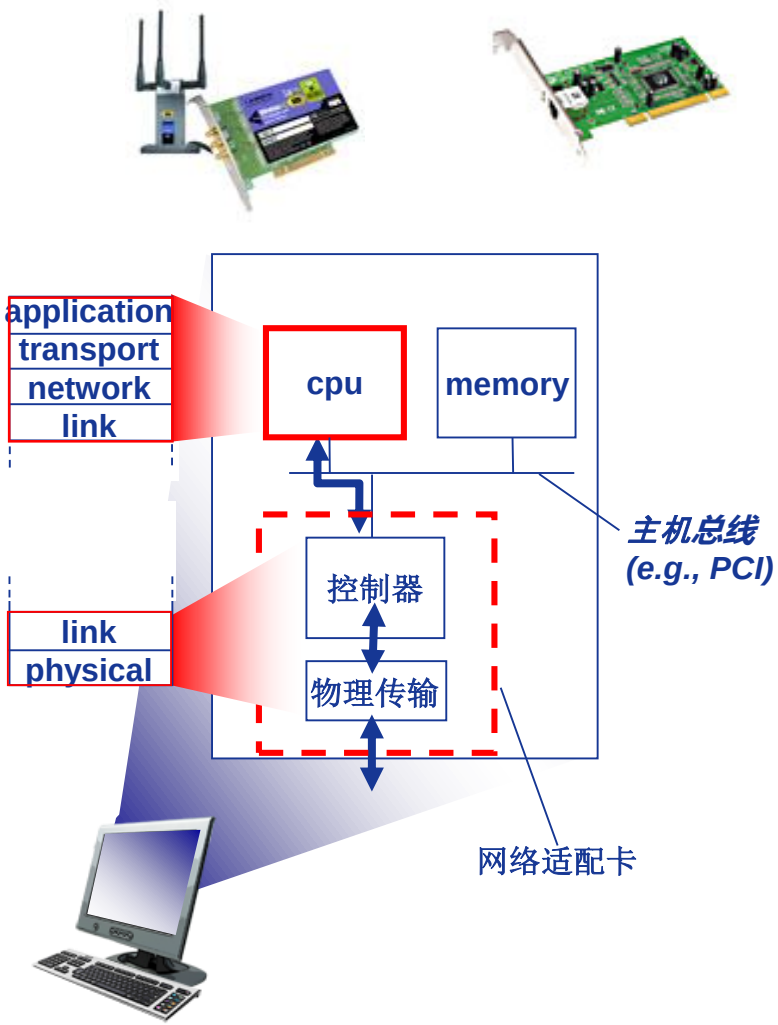
- 组帧(framing)
  - 封装数据报构成数据帧，加首部和尾部
  - 帧同步
- 链路接入(link access)
  - 如果是共享介质，需要解决信道接入(channel access)
  - 帧首部中的“MAC”地址，用于标识帧的源和目的
    - 不同于IP地址！
- 相邻结点间可靠交付
  - 在低误码率的有线链路上很少采用 (如光纤，某些双绞线等)
  - 无线链路：误码率高，需要可靠交付

链路层服务

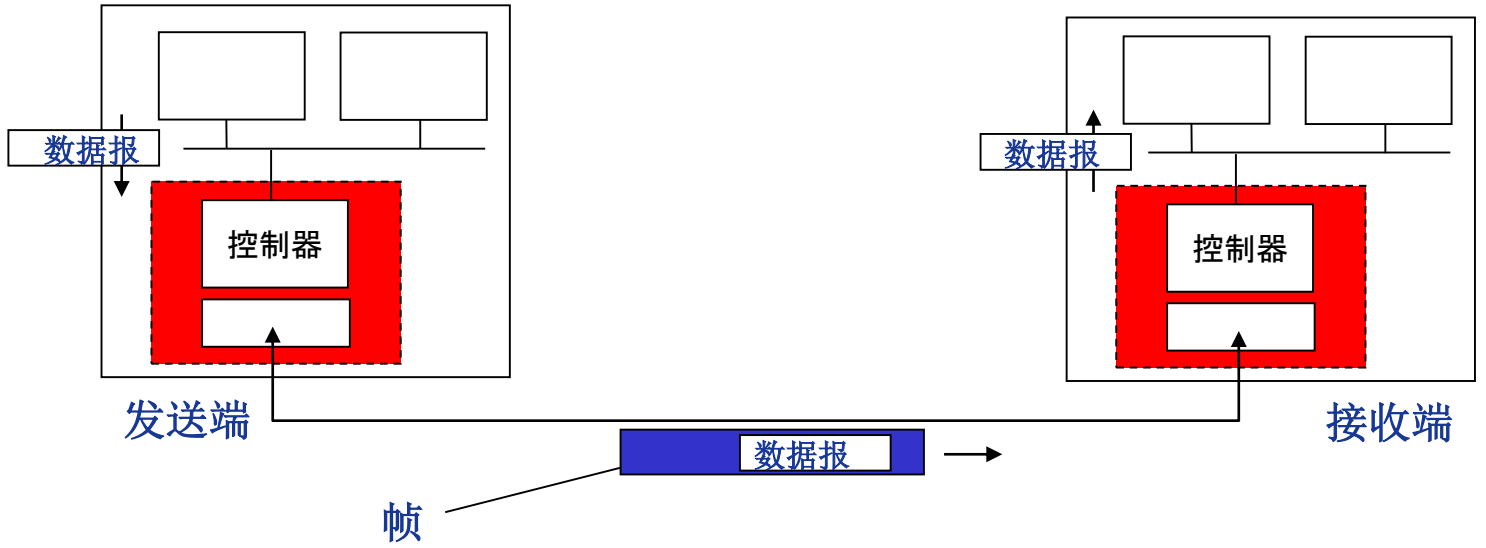
- 流量控制(flow control)
  - 协调(pacing)相邻的发送结点和接收
- 差错检测(error detection)
  - 信号衰减和噪声会引起差错.
  - 接收端检测到差错:
    - 通知发送端重传或者直接丢弃帧
- 差错纠正(error correction)
  - 接收端直接纠正比特差错
- 全双工和半双工通信控制
  - 全双工：链路两端结点同时双向传输
  - 半双工：链路两端结点交替双向传输

链路层的具体实现？

- 每个主机或路由器接口
- 链路层在“适配器”(即网络接口卡-NIC)中实现或者在一个芯片上实现
  - 以太网网卡，802.11网卡；以太网芯片组
  - 实现链路层和物理层
- 链接主机的系统总线
- 由硬件、软件与固件组成



网卡间通信



- 发送端：
  - 将数据报封装成帧
  - 增加差错检测比特，实现可靠数据传输和流量控制等.
- 接收端：
  - 检测差错，实现可靠数据传输和流量控制等
  - 提取数据报，交付上层协议实体

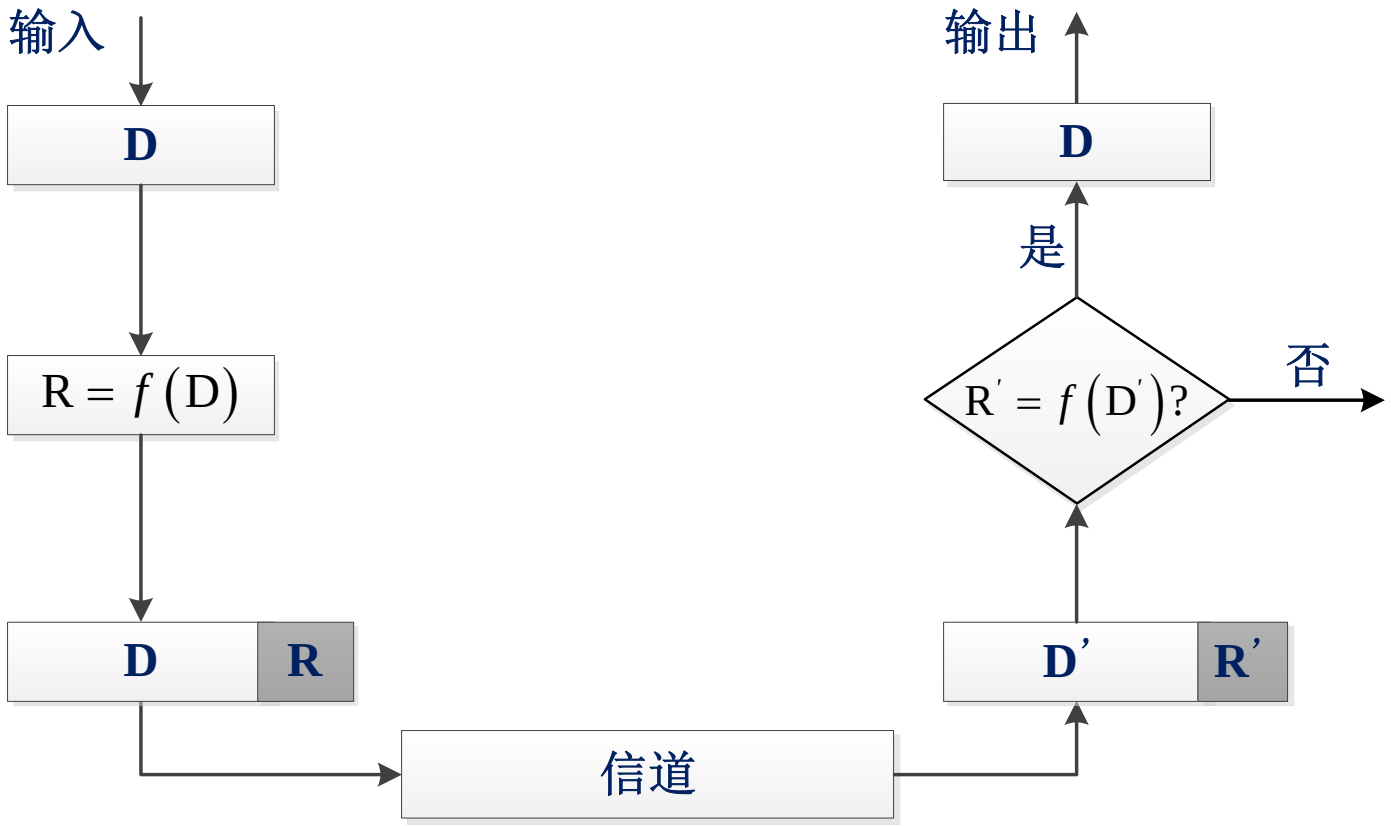
数据链路层与局域网

- 数据链路层服务
- 差错编码
- 多路访问控制(MAC)协议
  - 信道划分MAC协议
  - 随机访问MAC协议
  - 轮转访问MAC协议
- ARP协议
- 以太网
  - 交换机
  - 虚拟局域网（VLAN）
- PPP协议
- 802.11无线局域网

差错检测：差错编码

差错编码基本原理：

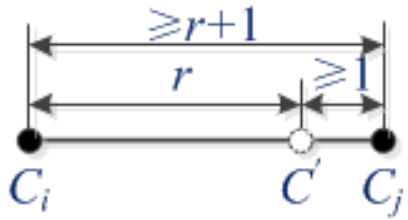
D→DR，其中R为差错检测与纠正比特（冗余比特）



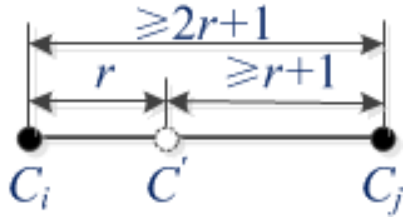
差错编码不能保证100%可靠！

差错编码的检错能力

- 差错编码可分为检错码与纠错码
- 对于检错码，如果编码集的汉明距离$d_s=r+1$，则该差错编码可以检测r位差错



  - 例如，编码集 {0000,0101,1010,1111} 的汉明距离$d_s=2$，可以100%检测1比特差错
- 对于纠错码，如果编码集的汉明距离$d_s=2r+1$，则该差错编码可以纠正r位差错

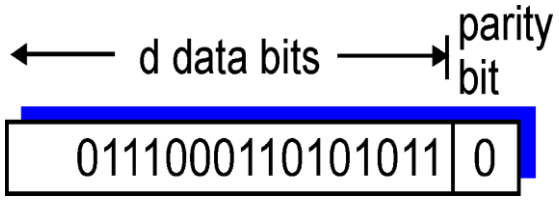


  - 例如，编码集 {0000000,0100101,1001010,1111111} 的汉明距离$d_s=3$，可以纠正1比特差错，如100010纠正为101010。

奇偶校验码

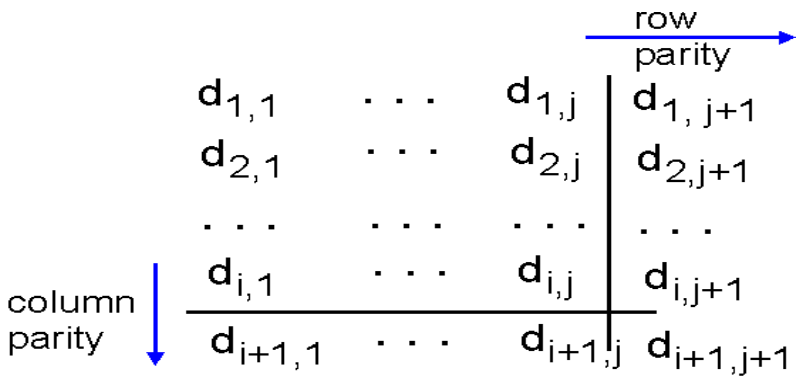
1比特校验位:

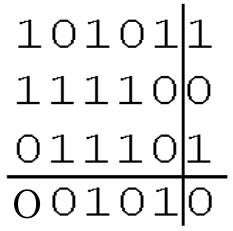
- 检测奇数位差错

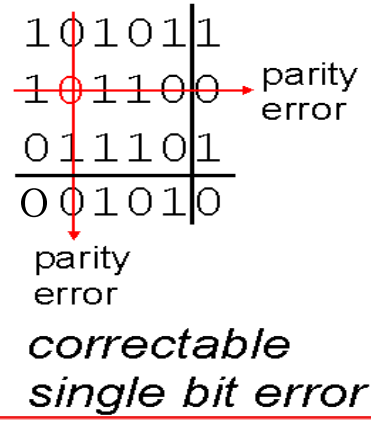


二维奇偶校验:

- 检测奇数位差错、部分偶数位差错
- 纠正同一行/列的奇数位错







Internet校验和(Checksum)

发送端:

- 将“数据”(校验内容)划分为16位的二进制“整数”序列
- 求和(sum)：补码求和 (最高位进位的“1”，返回最低位继续加）
- 校验和(Checksum)：sum的反码
- 放入分组(UDP、TCP、IP)的校验和字段

接收端:

- 与发送端相同算法计算
- 计算得到的"checksum"：
  - 为16位全0（或sum为16位全1）：无错
  - 否则：有错

循环冗余校验码(CRC)

- 检错能力更强大的差错编码
- 将数据比特，D，视为一个二进制数
- 选择一个r+1位的比特模式 (生成比特模式)，G
- 目标：选择r位的CRC比特，R，满足
  - <D,R>刚好可以被G整除(模2)
  - 接收端检错：利用G除<D,R>，余式全0，无错；否则，有错！
  - 可以检测所有突发长度小于r+1位差错。
- 广泛应用于实际网络 (以太网，802.11 WiFi，ATM)

CRC举例

期望:

$D \cdot 2^r$ XOR R = nG

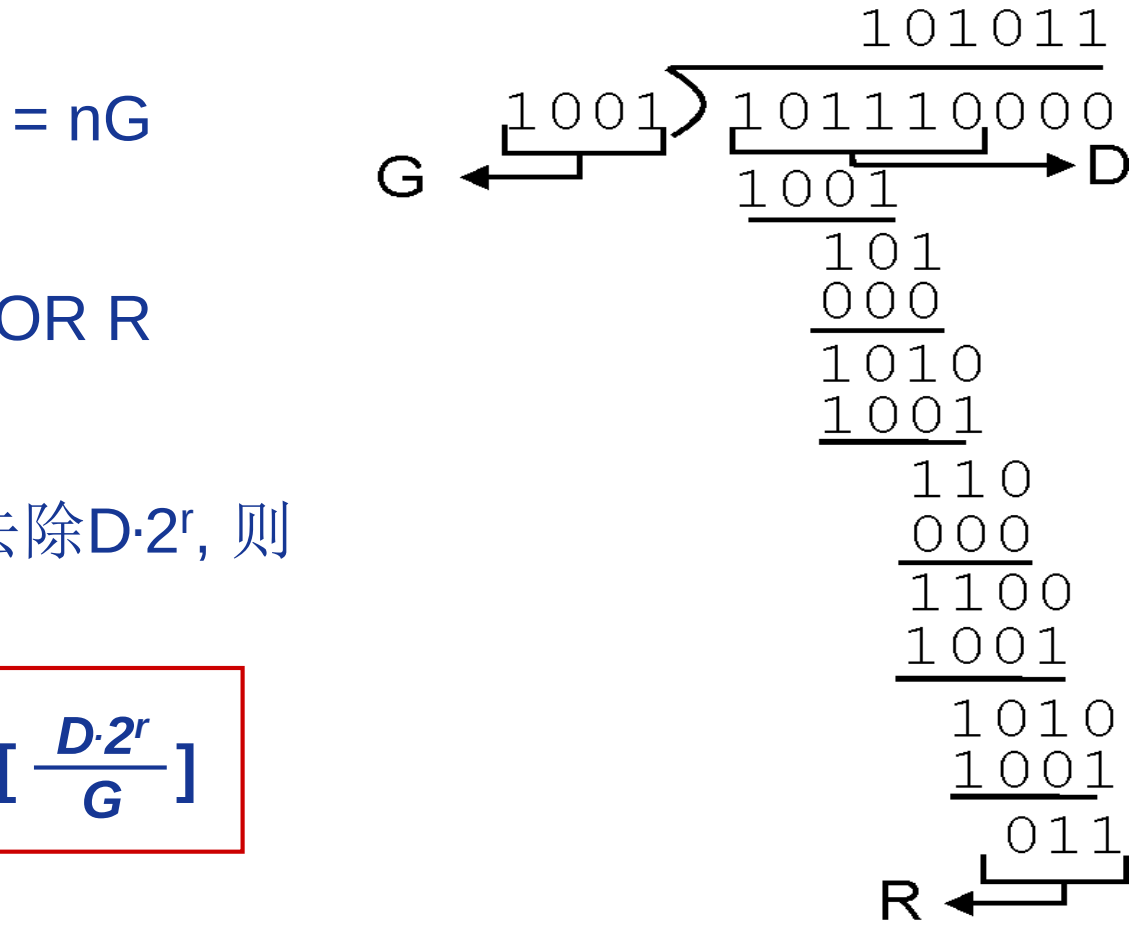
相当于:

$D \cdot 2^r$ = nG XOR R

相当于:

如果利用G去除$D \cdot 2^r$, 则余式即为R:

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



数据链路层与局域网

- 数据链路层服务
- 差错编码
- 多路访问控制(MAC)协议
  - 信道划分MAC协议
  - 随机访问MAC协议
  - 轮转访问MAC协议
- ARP协议
- 以太网
  - 交换机
  - 虚拟局域网（VLAN）
- PPP协议
- 802.11无线局域网

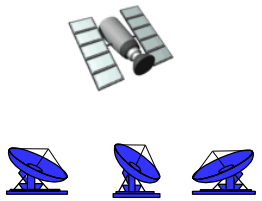
多路访问控制(MAC)协议

两类“链路”：

- 点对点链路
  - 拨号接入的PPP
  - 以太网交换机与主机间的点对点链路
- 广播链路 (共享介质)
  - 早期的总线以太网
  - HFC的上行链路
  - 802.11无线局域网

共享线路
(e.g., 总线以太网)

共享RF
(e.g., 802.11 WiFi)



共享RF
(e.g., 卫星网络)

共享空气、声频
(e.g., 鸡尾酒会)

多路访问控制(MAC)协议

- 单一共享广播信道
- 两个或者两个以上结点同时传输：干扰(interference)
  - 冲突(collision)
    - 结点同时接收到两个或者多个信号→接收失败！

多路访问控制协议(multiple access control protocol)

- 采用分布式算法决定结点如何共享信道，即决策结点何时可以传输数据
- 必须基于信道本身，通信信道共享协调信息！
  - 无带外信道用于协调

理想MAC协议

给定：速率为R bps的广播信道

期望：

1. 当只有一个结点希望传输数据时，它可以以速率 R发送.
2. 当有M个结点期望发送数据时，每个节点平均发送数据的平均速率是R/M
3. 完全分散控制:
   - 无需特定结点协调
   - 无需时钟、时隙同步
4. 简单

MAC协议分类

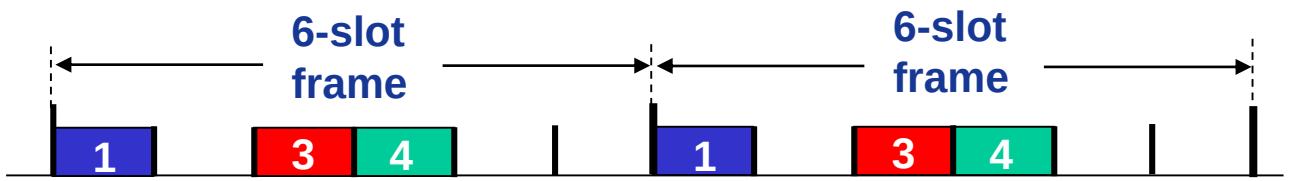
三大类：

- 信道划分(channel partitioning)MAC协议
  - 多路复用技术
  - TDMA、FDMA、CDMA、WDMA等
- 随机访问(random access)MAC协议
  - 信道不划分，允许冲突
  - 采用冲突“恢复”机制
- 轮转(“taking turns”)MAC协议
  - 结点轮流使用信道

信道划分MAC协议：TDMA

TDMA: time division multiple access

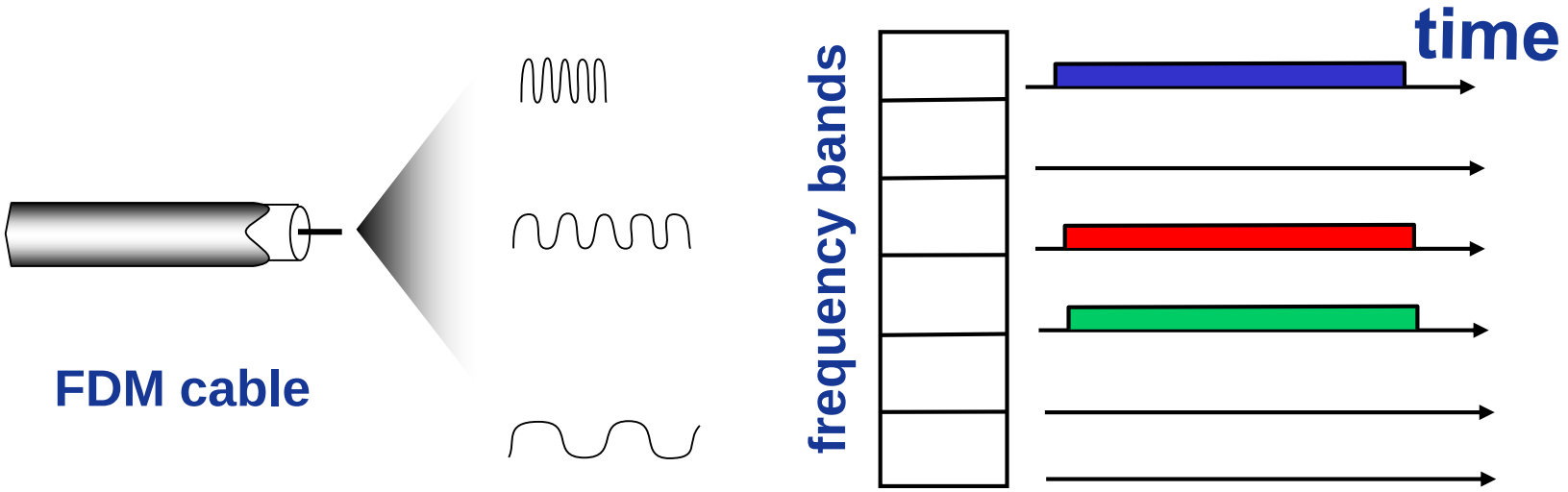
- “周期性”接入信道
- 每个站点在每个周期，占用固定长度的时隙(e.g.长度=分组传输时间)
- 未用时隙空闲(idle)
- 例如：6-站点LAN，1,3,4传输分组，2,5,6空闲



信道划分MAC协议：FDMA

FDMA: frequency division multiple access

- 信道频谱划分为若干频带(frequency bands)
- 每个站点分配一个固定的频带
- 无传输频带空闲
- 例如: 6站点LAN, 1,3,4频带传输数据, 2,5,6频带空闲。



随机访问MAC协议

- 当结点要发送分组时：
  - 利用信道全部数据速率R发送分组
  - 没有事先的结点间协调
- 两个或多个结点同时传输：➜ “冲突”
- 随机访问MAC协议需要定义:
  - 如何检测冲突
  - 如何从冲突中恢复 (e.g., 通过延迟重传)
- 典型的随机访问MAC协议：
  - 时隙(sloted)ALOHA
  - ALOHA
  - CSMA、CSMA/CD、CSMA/CA

时隙ALOHA协议

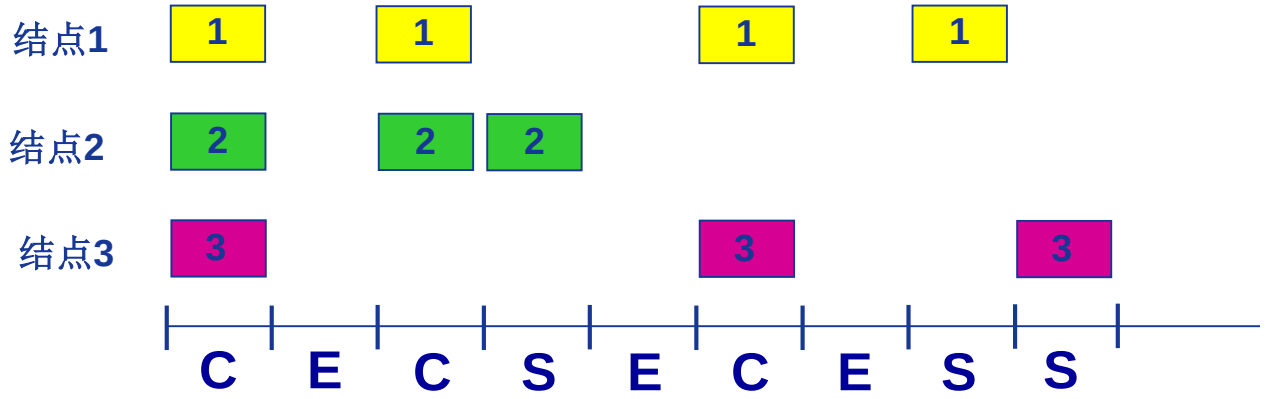
假定：

- 所有帧大小相同
- 时间被划分为等长的时隙(每个时隙可以传输1个帧)
- 结点只能在时隙开始时刻发送帧
- 结点间时钟同步
- 如果2个或2个以上结点在同一时隙发送帧，结点即检测到冲突

运行:

- 当结点有新的帧时，在下一个时隙(slot)发送
  - 如果无冲突：该结点可以在下一个时隙继续发送新的帧
  - 如果冲突：该结点在下一个时隙以概率p重传该帧，直至成功

时隙ALOHA协议



优点:

- 单个结点活动时，可以连续以信道全部速率传输数据
- 高度分散化：只需同步时隙
- 简单

缺点:

- 冲突，浪费时隙
- 空闲时隙
- 结点也许能以远小于分组传输时间检测到冲突
- 时钟同步

时隙ALOHA协议

效率(efficiency): 长期运行时，成功发送帧的时隙所占比例(很多结点，有很多帧待发送)

- 假设: N个结点有很多帧待传输，每个结点在每个时隙均以概率p发送数据
- 对于给定的一个结点，在一个时隙将帧发送成功的概率= $p(1-p)^{N-1}$
- 对于任意结点成功发送帧的概率= $Np(1-p)^{N-1}$

- 最大效率: 求得使$Np(1-p)^{N-1}$最大的p^*
- 对于很多结点，求$Np^*(1-p^*)^{N-1}$当N趋近无穷时的极限，可得:

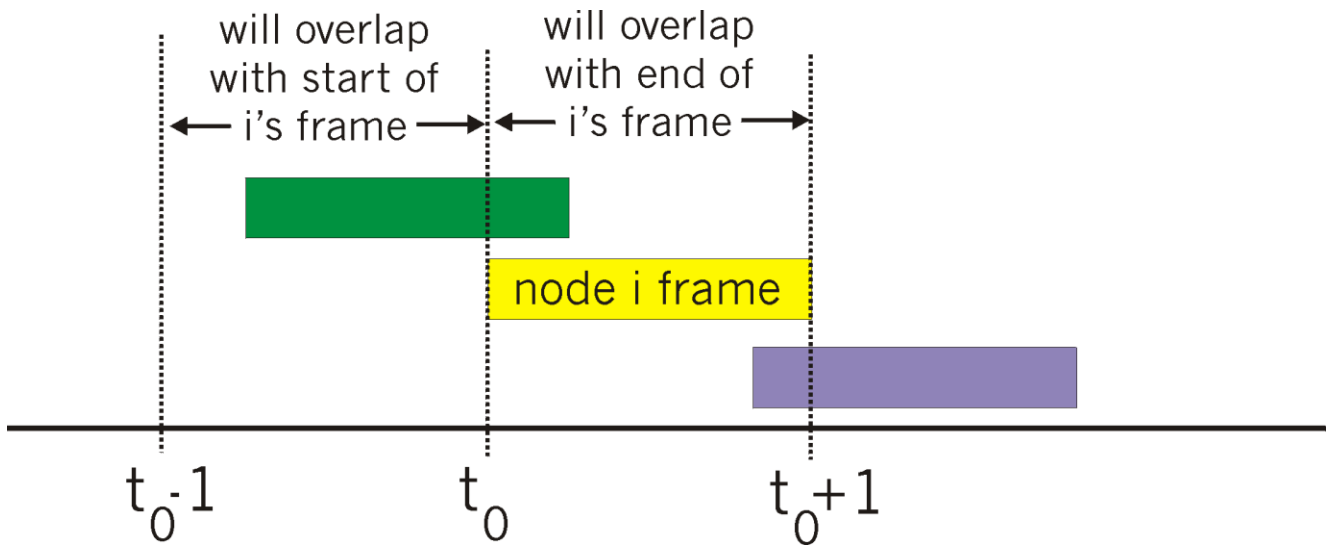
最大效率= 1/e = 0.37

最好情况: 信道被成功利用的时间仅占37%!

!

ALOHA协议

- 非时隙(纯)Aloha：更加简单，无需同步
- 当有新的帧生成时
  - 立即发送
- 冲突可能性增大:
  - 在t_0时刻发送帧，会与在$[t_0-1, t_0+1]$期间其他结点发送的帧冲突



ALOHA协议

P(给定结点成功发送帧) = P(该结点发送) ·
P(无其他结点在$[t_0-1, t_0]$期间发送帧) ·
P(无其他结点在$[t_0, t_0+1]$期间发送帧)

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

$= p \cdot (1-p)^{2(N-1)}$

… 选取最优的p，并令$n \longrightarrow \infty$

$= 1/(2e) = 0.18$

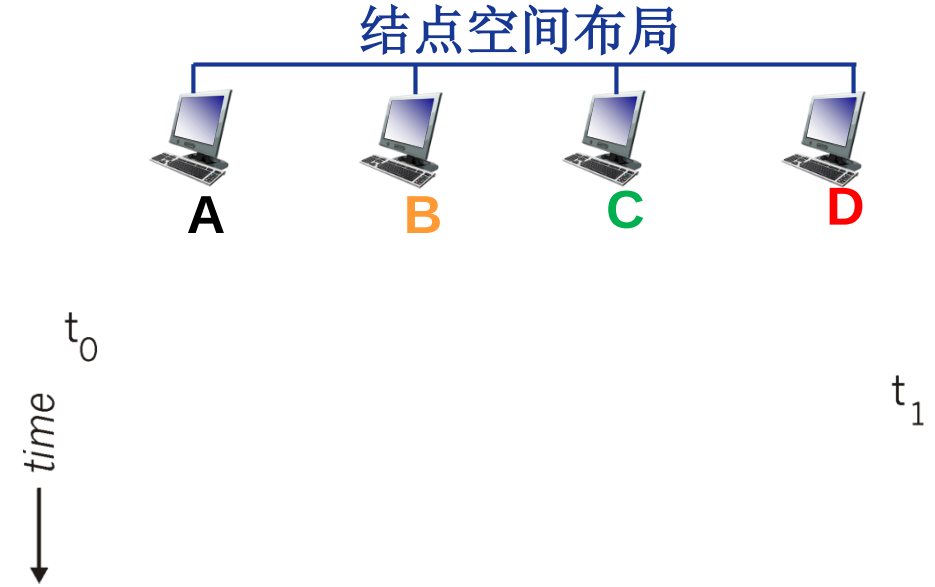
比时隙ALOHA协议更差!

CSMA协议

- 载波监听多路访问协议
  CSMA (carrier sense multiple access)
- 发送帧之前，监听信道(载波)：
  - 信道空闲：发送完整帧
  - 信道忙：推迟发送
    - 1-坚持CSMA
    - 非坚持CSMA
    - P-坚持CSMA
- 冲突可能仍然发生：
  信号传播延迟

CSMA协议

- 载波监听多路访问协议 CSMA (carrier sense multiple access)
- 发送帧之前，监听信道(载波)：
  - 信道空闲：发送完整帧
  - 信道忙：推迟发送
    - 1-坚持CSMA
    - 非坚持CSMA
    - P-坚持CSMA
- 冲突可能仍然发生：信号传播延迟
- 继续发送冲突帧：浪费信道资源



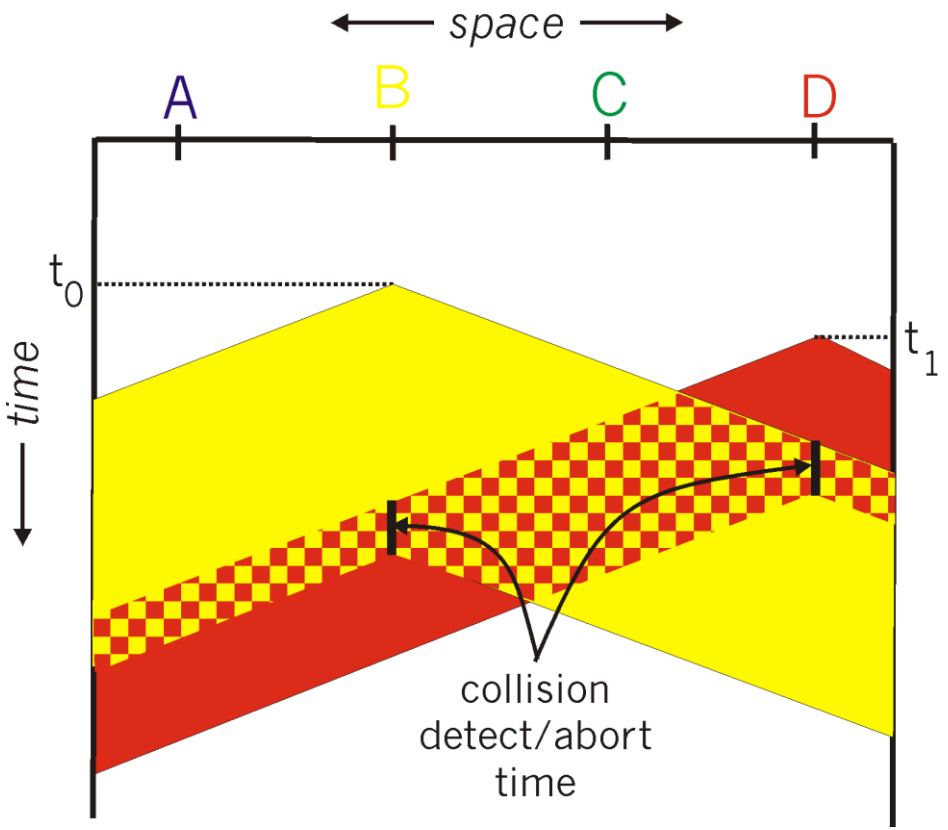
CSMA/CD协议

CSMA/CD: CSMA with Collision Detection

- 短时间内可以检测到冲突
- 冲突后传输中止，减少信道浪费

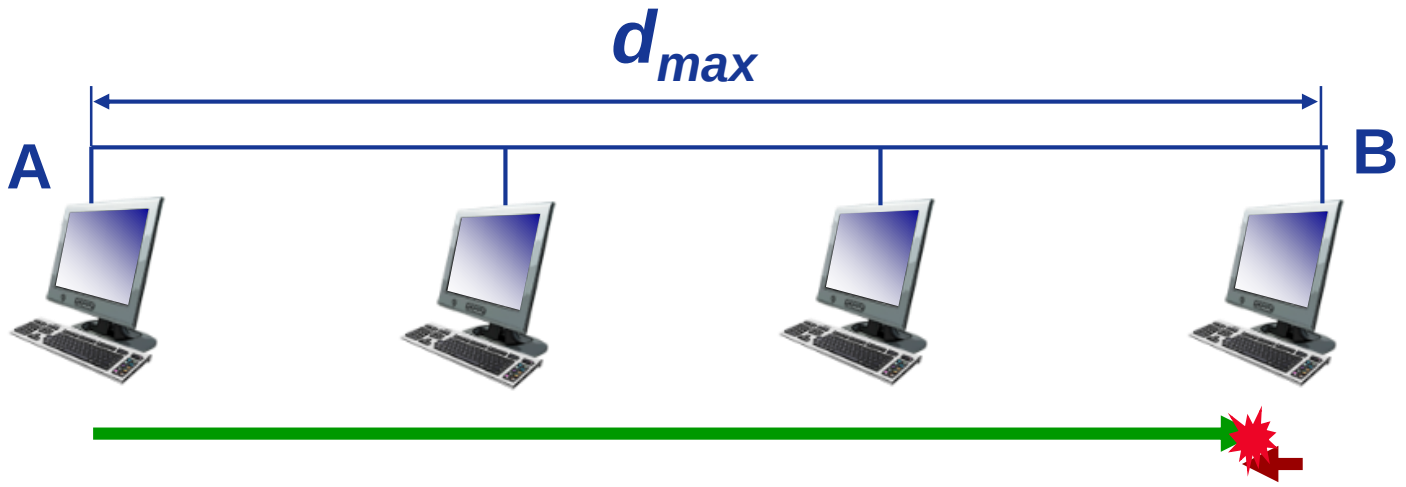
❖ 冲突检测:

- 有线局域网易于实现：测量信号强度，比较发射信号与接收信号
- 无线局域网很难实现：接收信号强度淹没在本地发射信号强度下



“边发边听，不发不听”

CSMA/CD协议



网络带宽：R bps

数据帧最小长度：L_{min} (bits)

信号传播速度：V (m/s)

$$L / R \geq 2d_{max} / V$$

$$L_{min} / R = 2d_{max} / V$$

$$L_{min} / R = RTT_{max}$$

例题

在一个采用CSMA/CD协议的网络中，传输介质是一根完整的电缆，传输速率为1 Gbps，电缆中的信号传播速度是200 000 km/s。若最小数据帧长度减少800比特，则最远的两个站点之间的距离至少需要

A.增加160 m　　B.增加80 m

C.减少160 m　　D.减少80 m

解：根据CSMA/CD协议工作原理，有

$L_{min}/R=2*d_{max}/V$,则$d_{max}=(V/2R)*L_{min}$，于是

$\Delta d_{max}=(V/2R)*\Delta L_{min}$

将V=200 000 km/s, R=1 Gbps, $\Delta L_{min}=-800$bit,代入得：

$\Delta d_{max}=(200000*10^3/(2*10^9))*(-800)=-80$ m

答案：D

CSMA/CD效率

- T_{prop} = LAN中2个结点间的最大传播延迟
- t_{trans} = 最长帧传输延迟

$$效率 = \frac{1}{1+5t_{prop}/t_{trans}}$$

- t_{prop} 趋近于0或者t_{trans} 趋近于∞时，效率趋近于1
- 远优于ALOHA，并且简单、分散！

轮转访问MAC协议

信道划分MAC协议：

- 网络负载重时，共享信道效率高，且公平
- 网络负载轻时，共享信道效率低！

随机访问MAC协议：

- 网络负载轻时，共享信道效率高，单个结点可以利用信道的全部带宽
- 网络负载重时，产生冲突开销

轮转访问MAC协议：

综合两者的优点！

轮转访问MAC协议

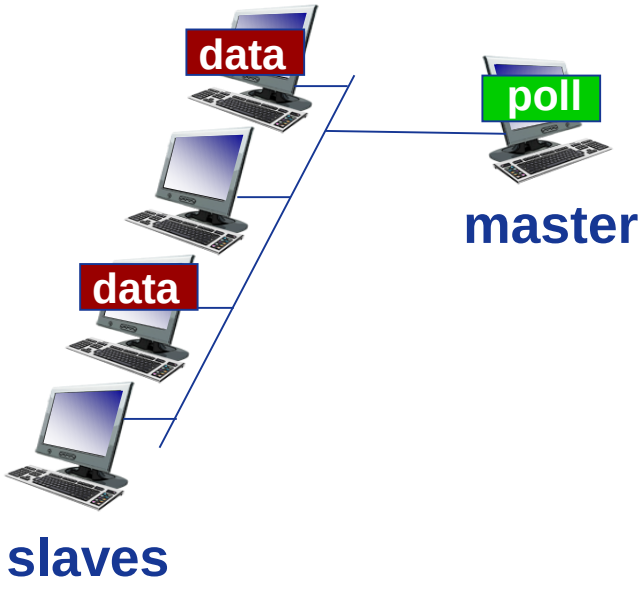
轮询(polling):

- 主结点轮流“邀请”从属结点发送数据
- 典型应用：“哑(dumb)”从属设备

轮转访问MAC协议

轮询(polling):

- 主结点轮流“邀请”从属结点发送数据
- 典型应用：“哑(dumb)”从属设备
- 问题：
  - 轮询开销
  - 等待延迟
  - 单点故障



轮转访问MAC协议

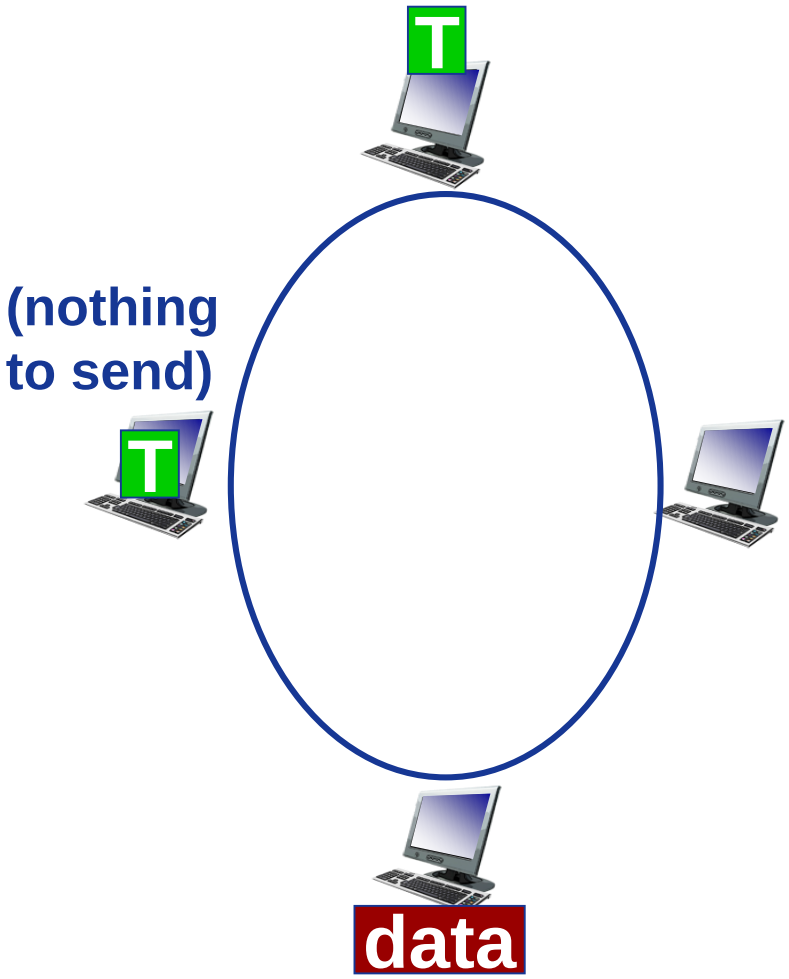
令牌传递(token passing):

- 控制令牌依次从一个结点传递到下一个结点.
- 令牌：特殊帧

轮转访问MAC协议

令牌传递(token passing):

- 控制令牌依次从一个结点传递到下一个结点.
- 令牌：特殊帧
- 问题:
  - 令牌开销
  - 等待延迟
  - 单点故障



MAC协议总结

- 信道划分MAC协议：时间、频带、码片划分
  - TDMA、FDMA、CDMA
- 随机访问MAC协议：
  - ALOHA, S-ALOHA, CSMA, CSMA/CD
  - CSMA/CD应用于以太网
  - CSMA/CA应用802.11无线局域网
- 轮转访问MAC协议：
  - 主结点轮询；令牌传递
  - 蓝牙、FDDI、令牌环网

数据链路层与局域网

- 数据链路层服务
- 差错编码
- 多路访问控制(MAC)协议
  - 信道划分MAC协议
  - 随机访问MAC协议
  - 轮转访问MAC协议
- ARP协议
- 以太网
  - 交换机
  - 虚拟局域网（VLAN）
- PPP协议
- 802.11无线局域网

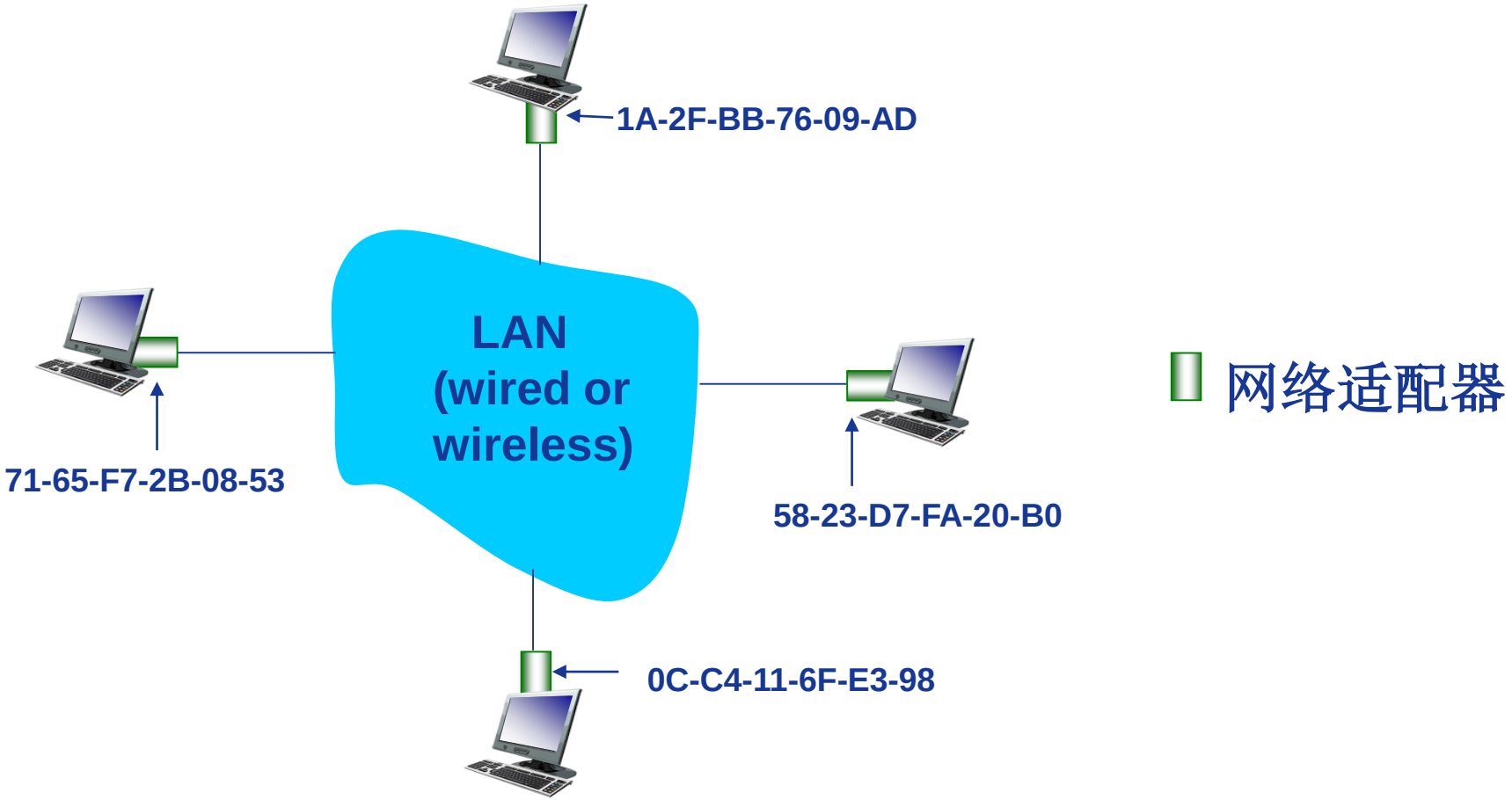
MAC地址

- 32位IP地址:
  - 接口的网络层地址
  - 用于标识网络层(第3层)分组，支持分组转发
- MAC地址(或称LAN地址,物理地址,以太网地址)：
  - 作用：用于局域网内标识一个帧从哪个接口发出，到达哪个物理相连的其他接口
  - 48位MAC地址(用于大部分LANs)，固化在网卡的ROM中，有时也可以软件设置
  - e.g.: 1A-2F-BB-76-09-AD

16进制表示

MAC地址

局域网中的每块网卡都有一个唯一的*MAC地址*

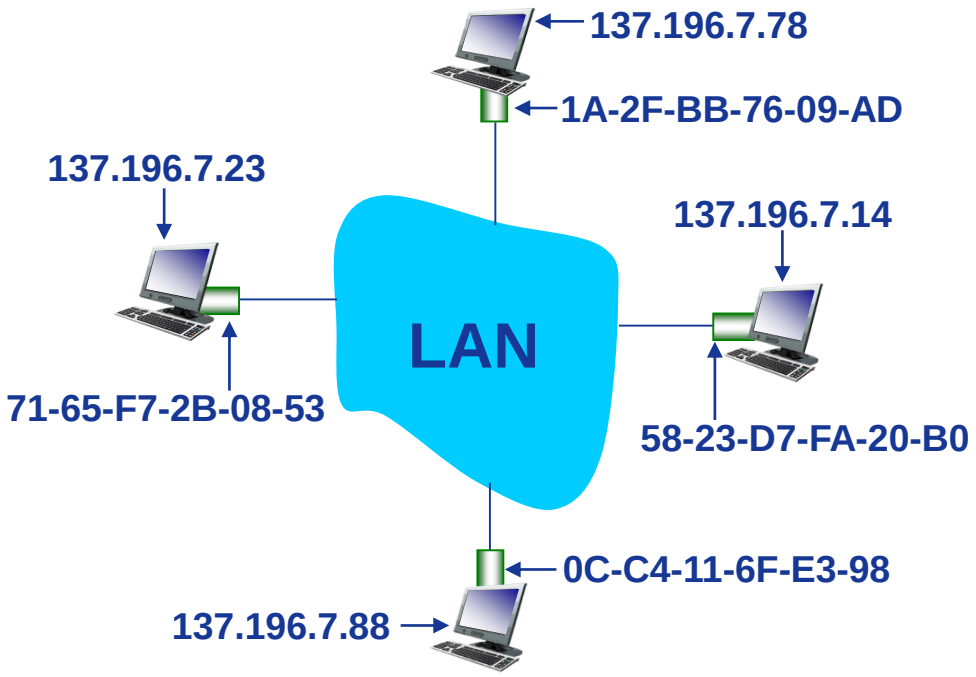


MAC地址

- MAC地址由IEEE统一管理与分配
- 网卡生产商购买MAC地址空间(前24比特)
- 类比：
  - MAC地址：身份证号
  - IP地址：邮政地址
- MAC地址是“平面”地址： ➜ 可“携带”
  - 可以从一个LAN移到另一个LAN
- IP地址是层次地址： ➜ 不可“携带”
  - IP地址依赖于结点连接到哪个子网

ARP: 地址解析协议

问题:（在同一个LAN内）如何在已知目的接口的IP地址前提下确定其MAC地址?



ARP表: LAN中的每个IP结点(主机、路由器)维护一个表

- 存储某些LAN结点的IP/MAC地址映射关系:
  < IP地址; MAC地址; TTL>
- TTL (Time To Live)：经过这个时间以后该映射关系会被遗弃(典型值为20min)

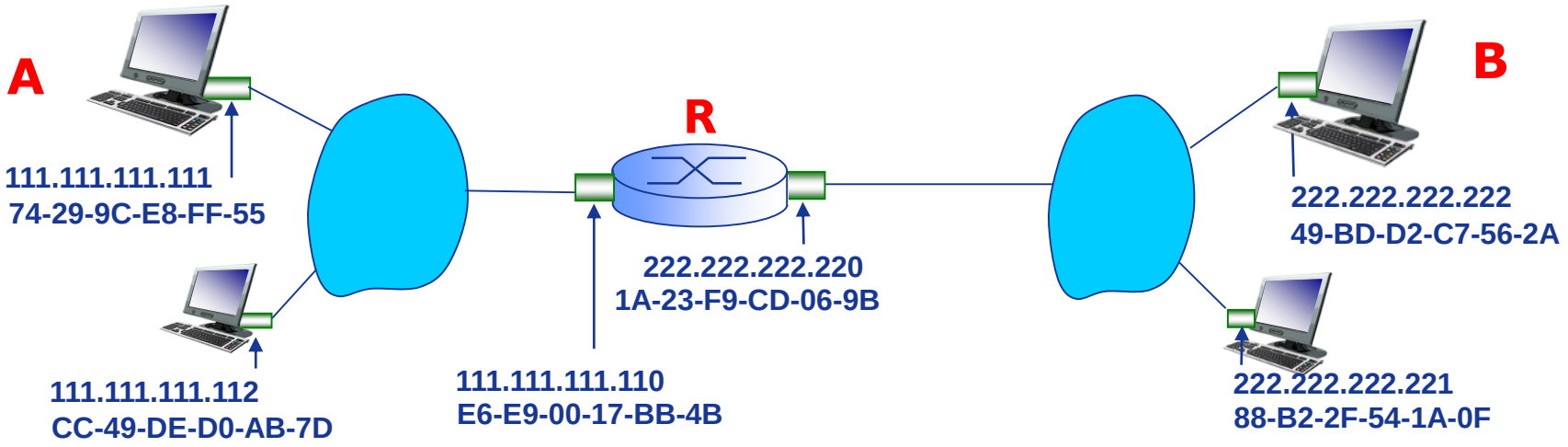
ARP协议: 同一局域网内

- A想要给同一局域网内的B发送数据报
  - B的MAC地址不在 A的ARP 表中.
- A广播ARP查询分组，其中包含B的IP地址
  - 目的MAC地址 = FF-FF-FF-FF-FF-FF
  - LAN中所有结点都会接收ARP查询
- B接收ARP查询分组，IP地址匹配成功，向A应答B的MAC 地址
  - 利用单播帧向A发送应答
- A在其ARP表中，缓存B的IP-MAC地址对，直至超时
  - 超时后，再次刷新
- ARP是“即插即用”协议:
  - 结点自主创建ARP表，无需干预

寻址: 从一个LAN路由至另一个LAN

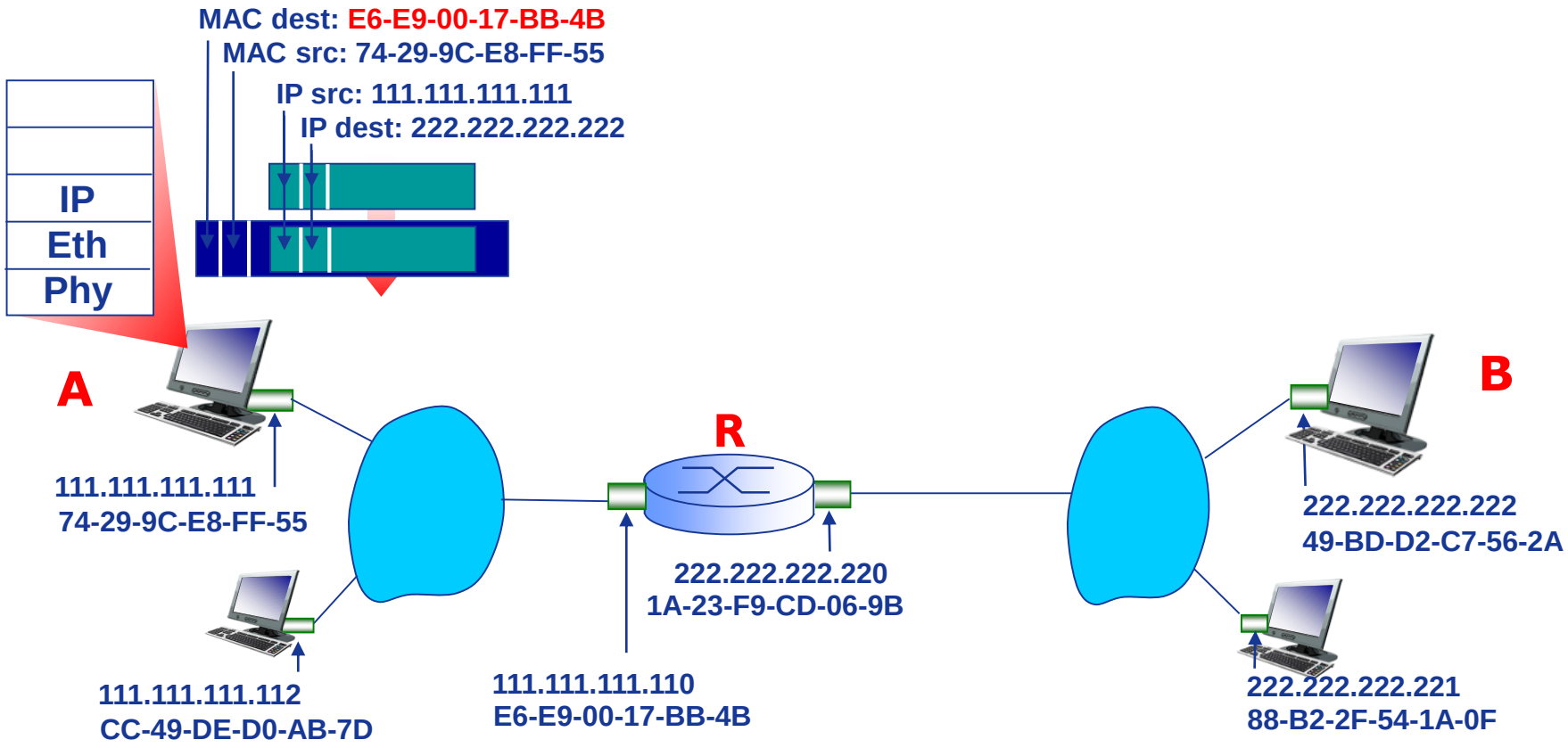
通信过程: A通过路由器R向B发送数据报

- 关注寻址：IP地址(数据报中)和MAC地址(帧中)
- 假设A知道B的IP地址(怎么知道的?)
- 假设A知道第一跳路由器R (左)接口IP地址 (怎么知道的?)
- 假设A知道第一跳路由器R (左)接口MAC地址 (怎么知道的?)



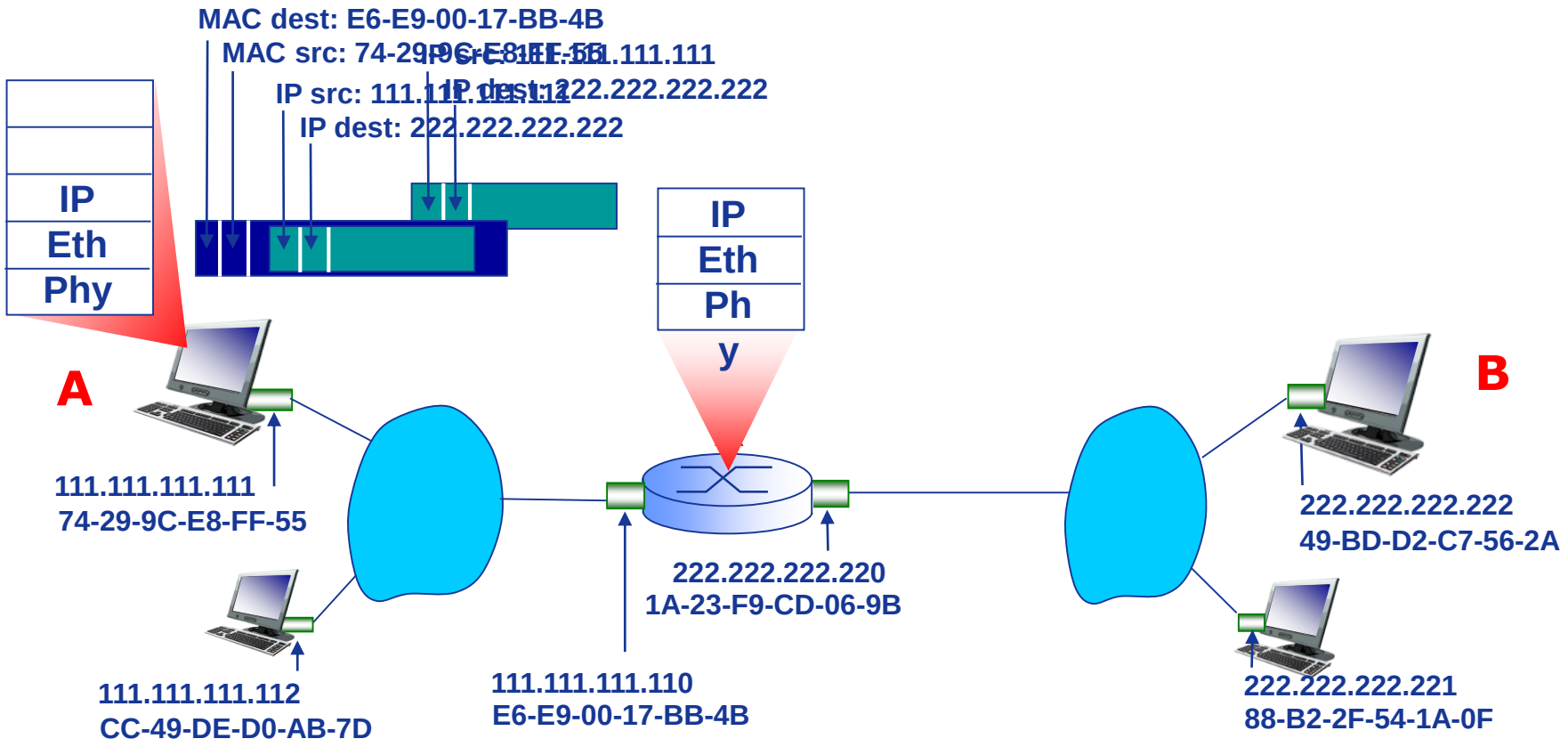
寻址：从一个LAN路由至另一个LAN

- A构造IP数据报，其中源IP地址是A的IP地址，目的IP地址是B的IP地址
- A构造链路层帧，其中源MAC地址是A的MAC地址，目的MAC地址是R(左)接口的MAC地址，封装A到B的IP数据报。



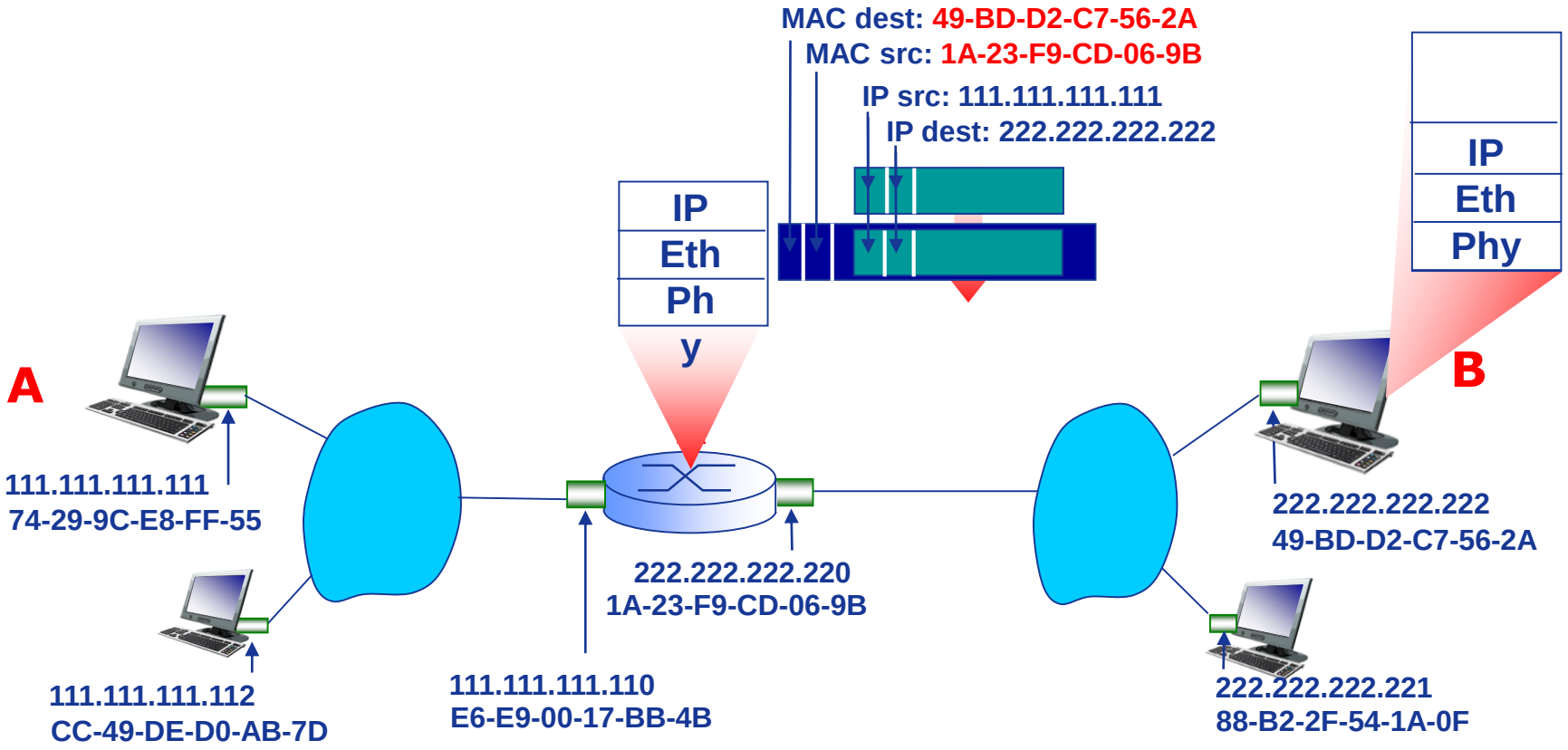
寻址：从一个LAN路由至另一个LAN

- 帧从A发送至R
- R接收帧，提取IP数据报，传递给上层IP协议



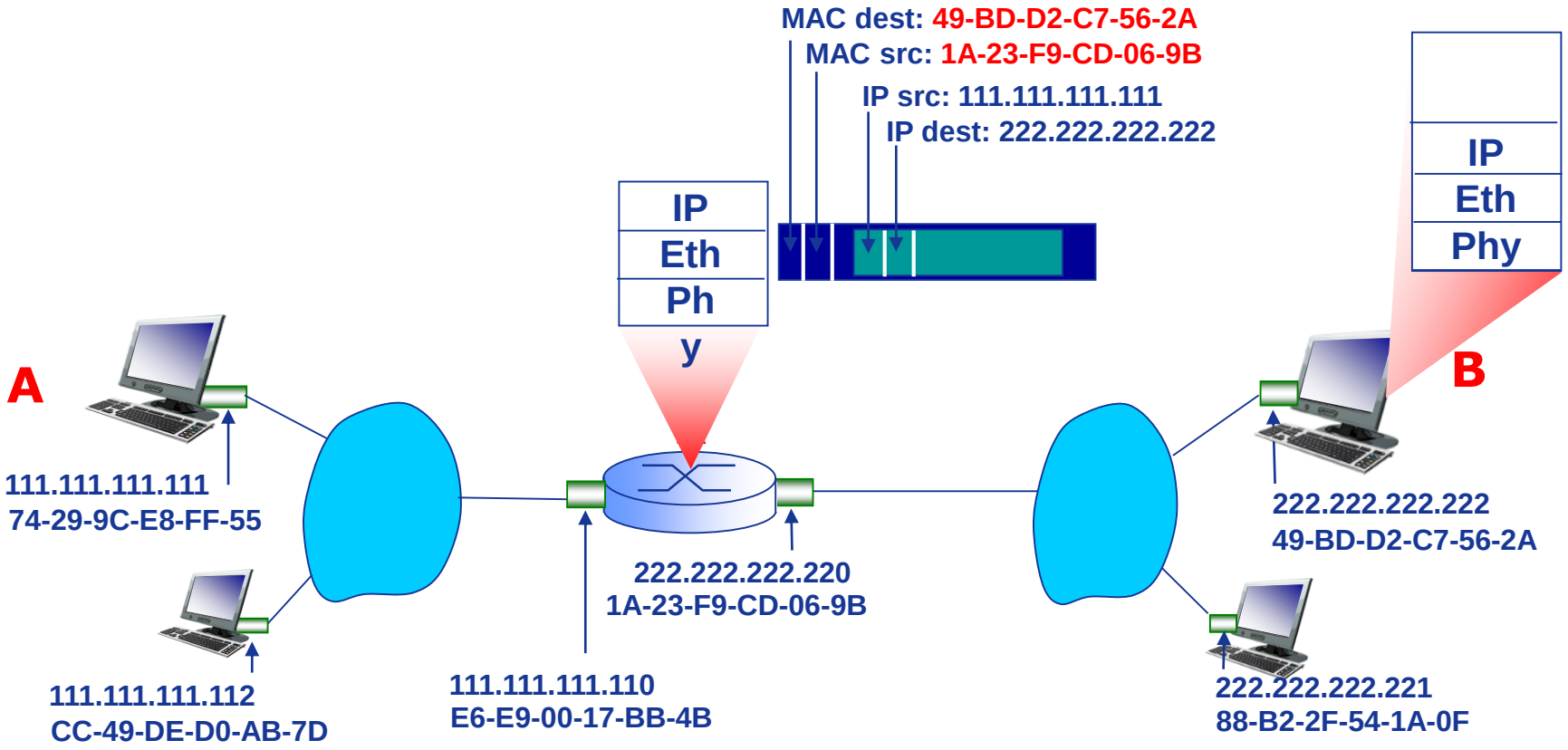
寻址：从一个LAN路由至另一个LAN

- R转发IP数据报（源和目的IP地址不变！）
- R创建链路层帧，其中源MAC地址是R(右)接口的MAC地址，目的MAC地址是B的MAC地址，封装A到B的IP数据报。



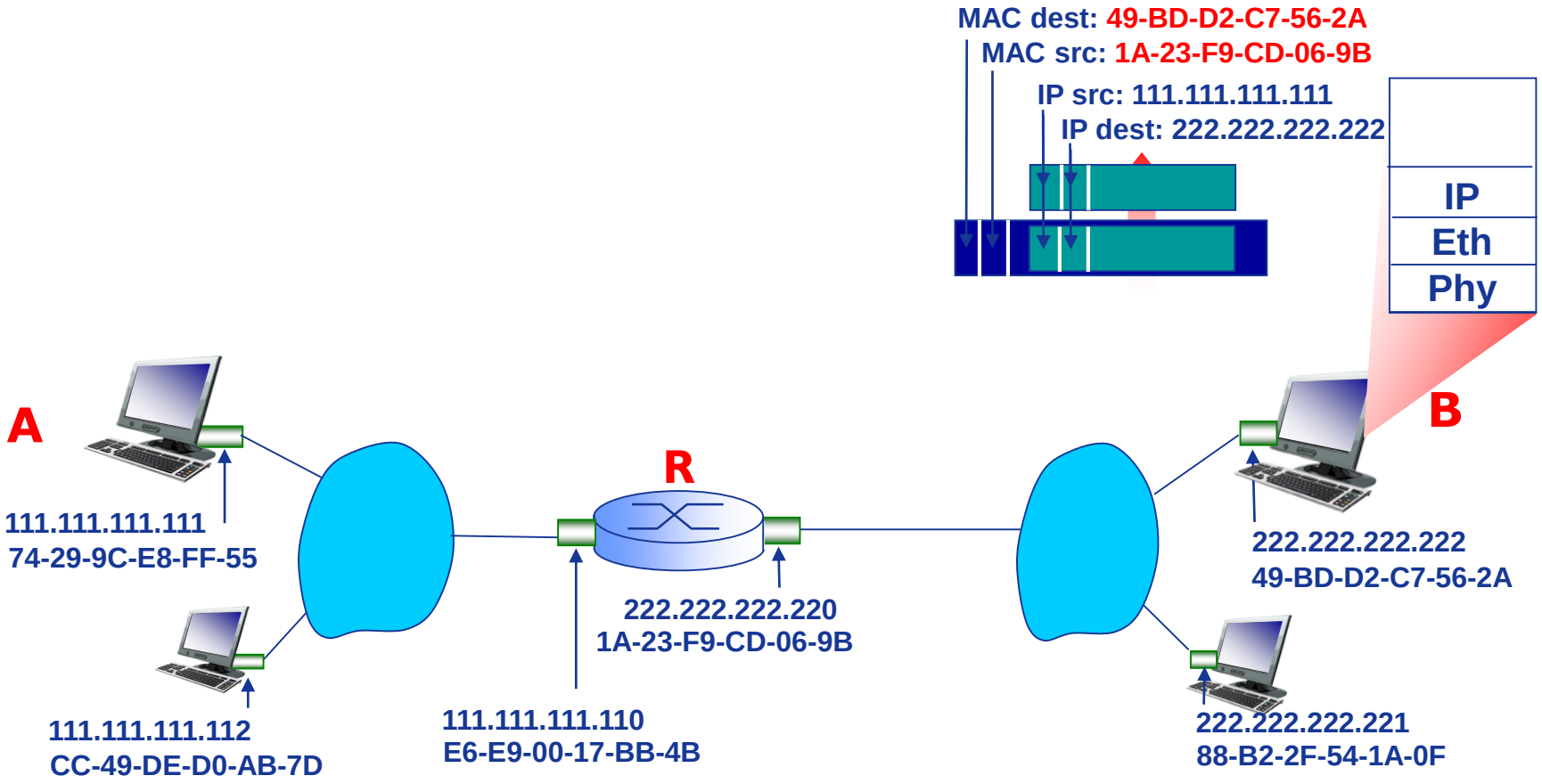
寻址：从一个LAN路由至另一个LAN

- R转发IP数据报（源和目的IP地址不变！）
- R创建链路层帧，其中源MAC地址是R (右)接口的MAC地址，目的MAC地址是B的MAC地址，封装A到B的IP数据报。



寻址：从一个LAN路由至另一个LAN

- R转发IP数据报（源和目的IP地址不变！）
- R创建链路层帧，其中源MAC地址是R (右)接口的MAC地址，目的MAC地址是B的MAC地址，封装A到B的IP数据报。



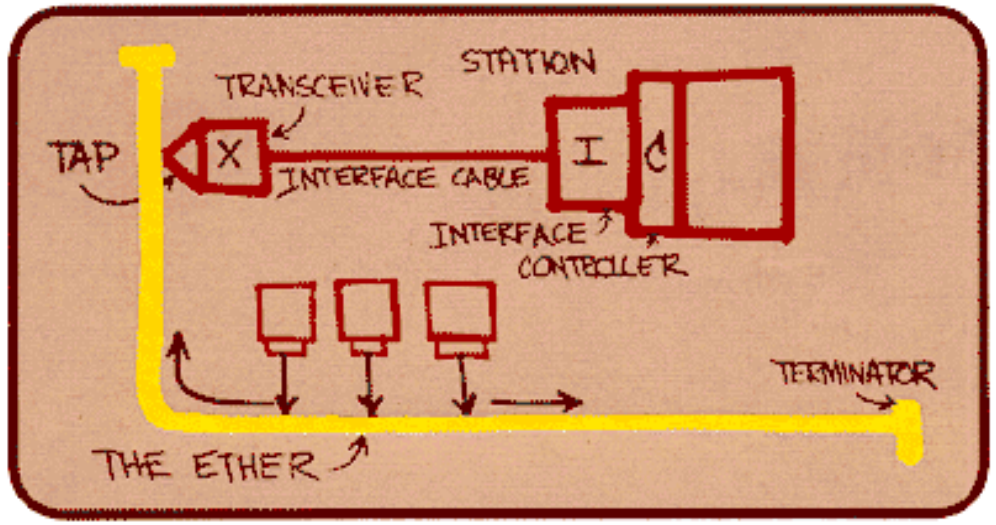
数据链路层与局域网

- 数据链路层服务
- 差错编码
- 多路访问控制(MAC)协议
  - 信道划分MAC协议
  - 随机访问MAC协议
  - 轮转访问MAC协议
- ARP协议
- 以太网
  - 交换机
  - 虚拟局域网（VLAN）
- PPP协议
- 802.11无线局域网

以太网(ETHERNET)

“统治地位”的有线LAN技术:

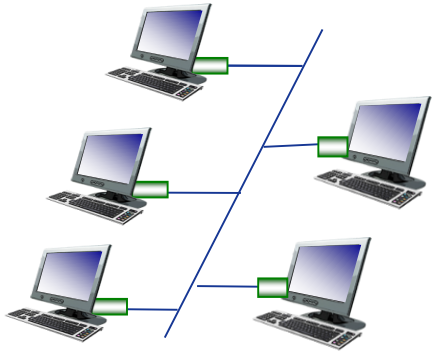
- 造价低廉(NIC不足￥100.00)
- 应用最广泛的LAN技术
- 比令牌局域网和ATM等，简单、便宜
- 满足网络速率需求：10 Mbps – 10 Gbps



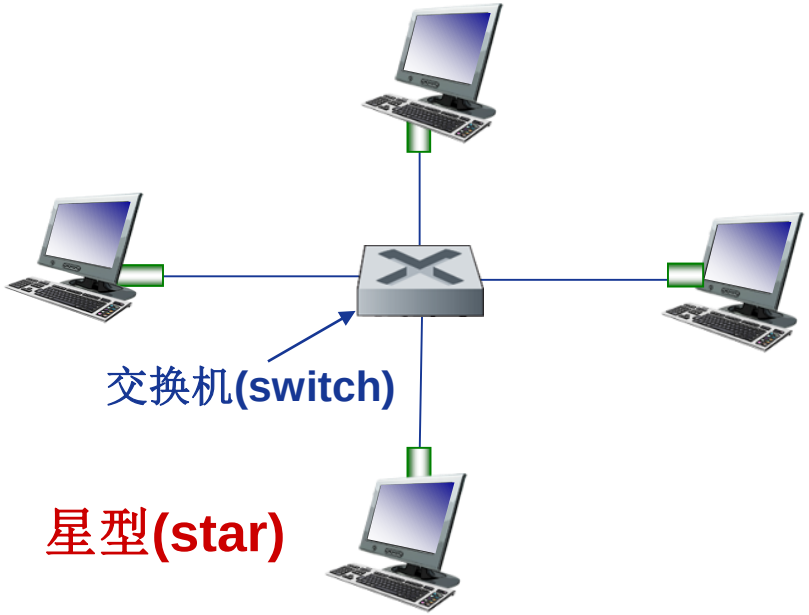
Metcalfe的以太网草图

以太网：物理拓扑

- 总线(bus): 上世纪90年代中期前流行
  - 所有结点在同一冲突域(collision domain) (可能彼此冲突)
- 星型(star): 目前主流网络拓扑
  - 中心交换机(switch)
  - 每个结点一个单独冲突域(结点间彼此不冲突)



总线(bus): 同轴电缆



星型(star)

以太网：不可靠、无连接服务

- 无连接(connectionless): 发送帧的网卡与接收帧的网卡间没有“握手”过程
- 不可靠(unreliable): 接收网卡不向发送网卡进行确认
  - 差错帧直接丢弃，丢弃帧中的数据恢复依靠高层协议 (e.g., TCP)，否则，发生数据丢失
- 以太网的MAC协议: 采用二进制指数退避算法的CSMA/CD

以太网CSMA/CD算法

1. NIC从网络层接收数据报，创建数据帧。
2. 监听信道：

   如果NIC监听到信道空闲，则开始发送帧；

   如果NIC监听到信道忙，则一直等待到信道空闲，然后发送帧。
3. NIC发送完整个帧，而没有检测到其他结点的数据发送，则NIC确认帧发送成功！
4. 如果NIC检测到其他结点传输数据，则中止发送，并发送堵塞信号(jam signal)
5. 中止发送后，NIC进入二进制指数退避：
   - 第m次连续冲突后：
     - 取$n = \mathrm{Min}(m, 10)$
     - NIC 从$\{0,1,2, \ldots, 2^n-1\}$中随机选择一个数K
     - NIC等待$K \cdot 512$比特的传输延迟时间，再返回第2步
   - 连续冲突次数越多，平均等待时间越长。

以太网帧结构

发送端网卡将IP数据报(或其他网络层协议分组)封装到以太网帧中：

前导码(Preamble)(8B):

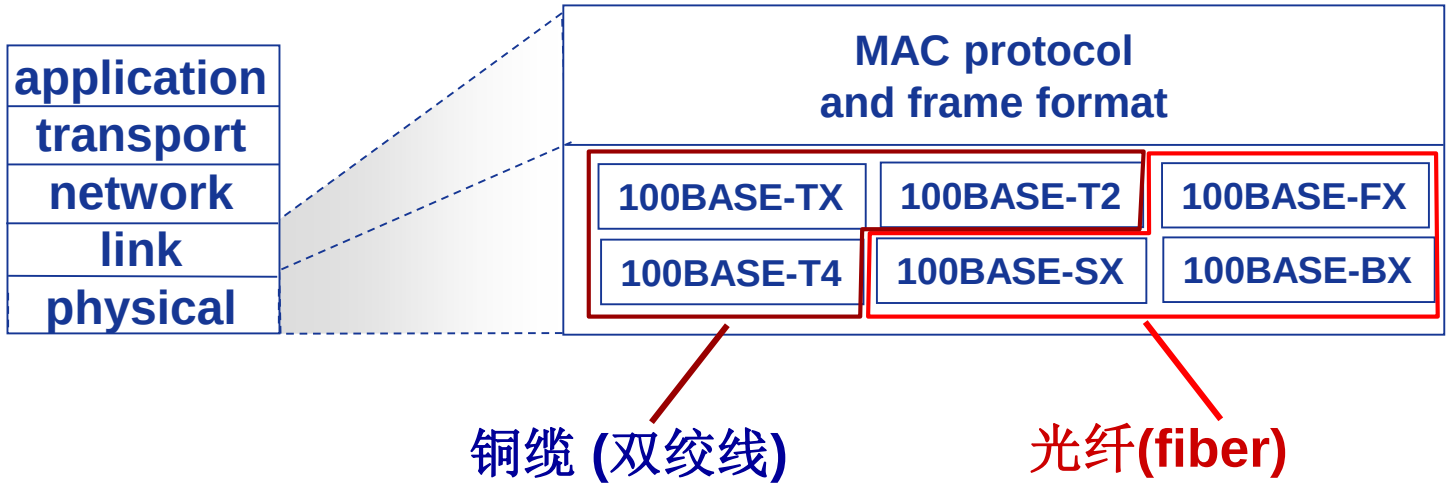
- 7个字节的10101010，第8字节为10101011
- 用于发送端与接收端的时钟同步

以太网帧结构

- 目的MAC地址、源MAC地址(各6B):
  - 如果网卡的MAC地址与收到的帧的目的MAC地址匹配，或者帧的目的MAC地址为广播地址(FF-FF-FF-FF-FF-FF)，则网卡接收该帧，并将其封装的网络层分组交给相应的网络层协议。
  - 否则，网卡丢弃(不接收)该帧。
- 类型(Type)(2B): 指示帧中封装的是哪种高层协议的分组(如，IP数据报、Novell IPX数据报、AppleTalk数据报等)
- 数据(Data)(46-1500B): 指上层协议载荷。
  - R=10Mbps，RTT_{max}=512μs，$L_{min} / R = RTT_{max}$
  - L_{min}=512bits=64B，$Data_{min}=L_{min}$-18=46B
- CRC(4B): 循环冗余校验码
  - 丢弃差错帧

802.3以太网标准: 链路与物理层

- 许多不同的以太网标准
  - 相同的MAC协议和帧格式
  - 不同速率: 2 Mbps, 10 Mbps, 100 Mbps, 1Gbps, 10G bps
  - 不同物理介质: 光纤, 线缆

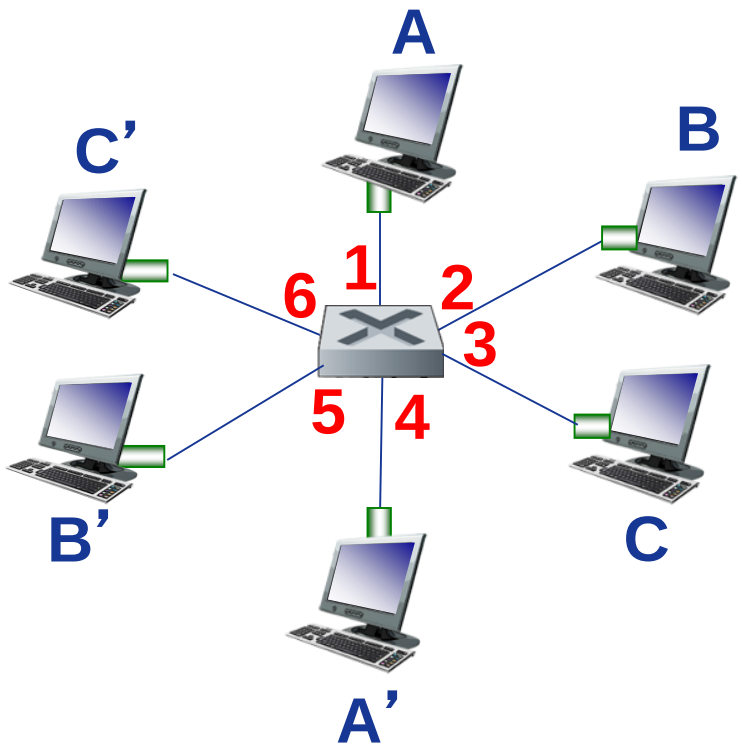


以太网交换机(switch)

- 链路层设备
  - 存储-转发以太网帧
  - 检验到达帧的目的MAC地址，选择性(selectively) 向一个或多个输出链路转发帧
  - 利用CSMA/CD访问链路，发送帧
- 透明(transparent)
  - 主机感知不到交换机的存在
- 即插即用(plug-and-play)
- 自学习(self-learning)
  - 交换机无需配置

交换机: 多端口间同时传输

- 主机利用独享(dedicated)链路直接连接交换机
- 交换机缓存帧
- 交换机在每段链路上利用CSMA/CD收发帧，但无冲突，且可以全双工
  - 每段链路一个独立的冲突域
- 交换(switching): A-A’与B-B’的传输可以同时进行，没有冲突

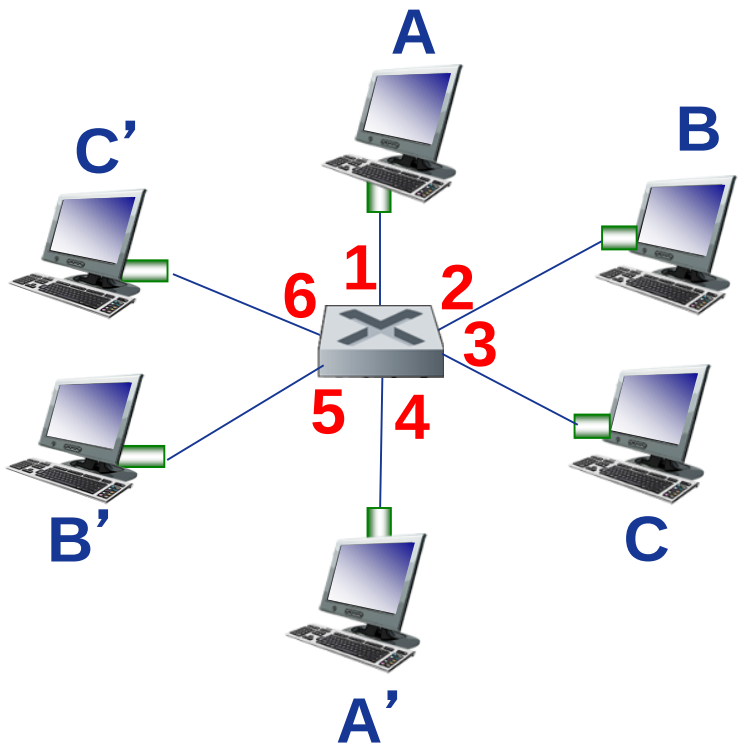


6个接口交换机
(1,2,3,4,5,6)

交换机转发表：交换表

Q: 交换机怎么知道A’可以通过接口4到达，而B’可以通过接口5到达?

- A: 每个交换机有一个交换表(switch table), 每个入口(entry):
  - (主机的MAC地址, 到达主机的接口, 时间戳)
  - 看起来很像路由表！
- Q: 交换表入口信息如何创建和维护的那？
  - 类似于路由协议？



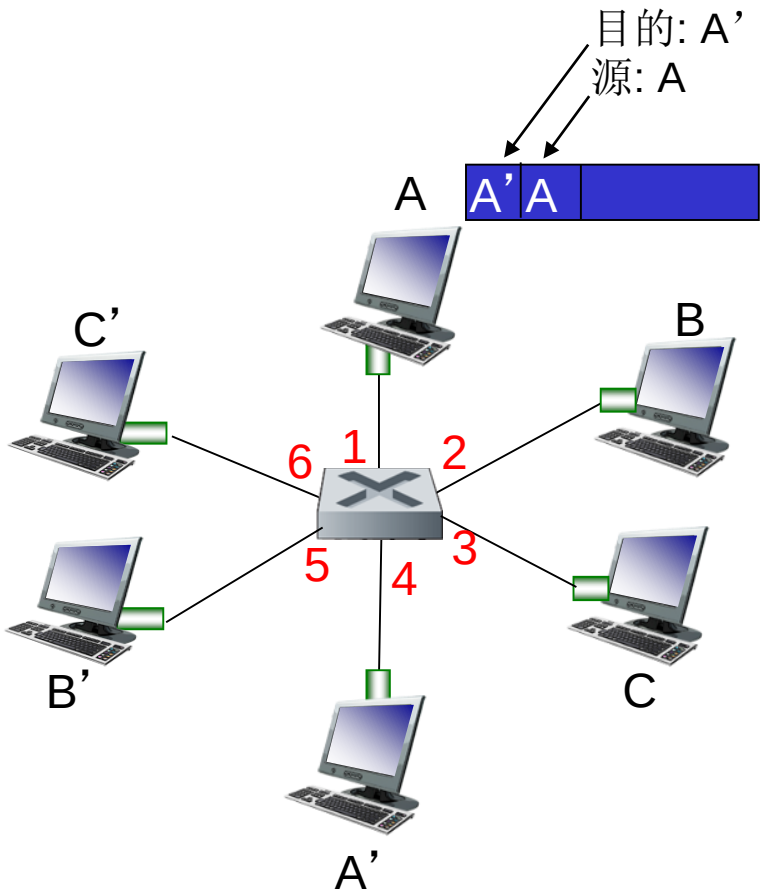
6个接口交换机 (1,2,3,4,5,6)

交换机: 自学习

- 交换机通过自学习，获知到达主机的接口信息
  - 当收到帧时，交换机“学习”到发送帧的主机（通过帧的源MAC地址），位于收到该帧的接口所连接的LAN网段
  - 将发送主机MAC地址/接口信息记录到交换表中

交换表
(初始为空)

MAC地址	接口	TTL
A	***1***	***60***



交换机: 帧过滤/转发

当交换机收到帧:

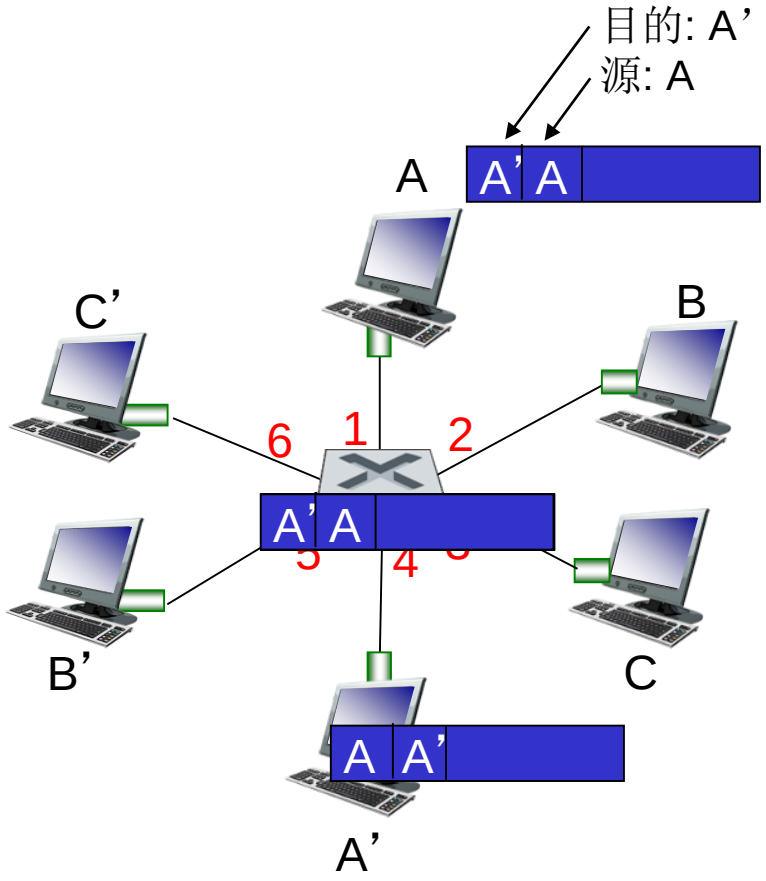
1. 记录帧的源MAC地址与输入链路接口
2. 利用目的MAC地址检索交换表
3. if 在交换表中检索到与目的MAC地址匹配的入口(entry)
   then {
   if 目的主机位于收到帧的网段
   then 丢弃帧
   else 将帧转发到该入口指向的接口
   }
   else 泛洪(flood) /* 向除收到该帧的接口之外的所有接口转发 */

自学习与转发过程举例

- 目的MAC地址A’，位置未知：
  泛洪
- 目的MAC地址A，位置已知：
  选择性转发

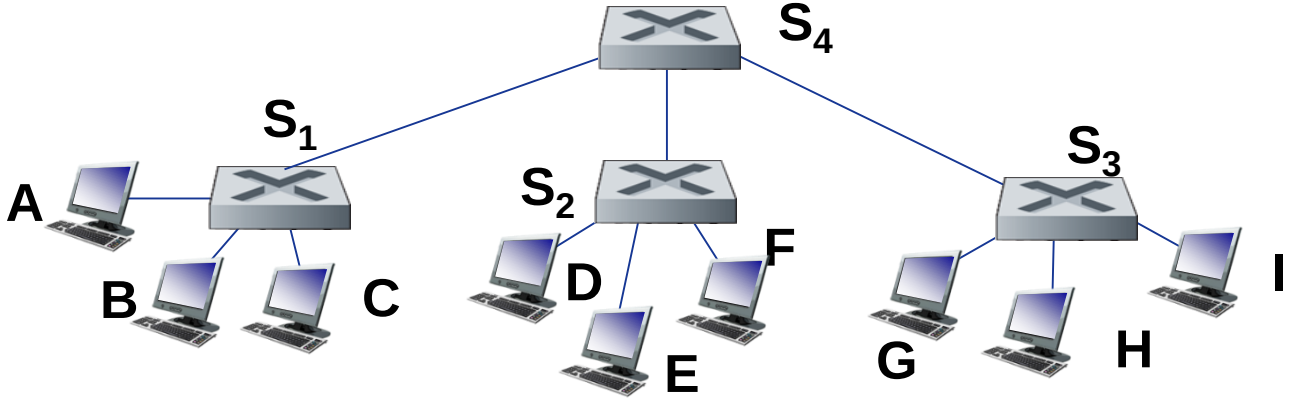
交换表
(初始为空)

MAC地址	接口	TTL
A *A’*	*1* *4*	*60* *60*



交换机互联

- 交换机可以互联

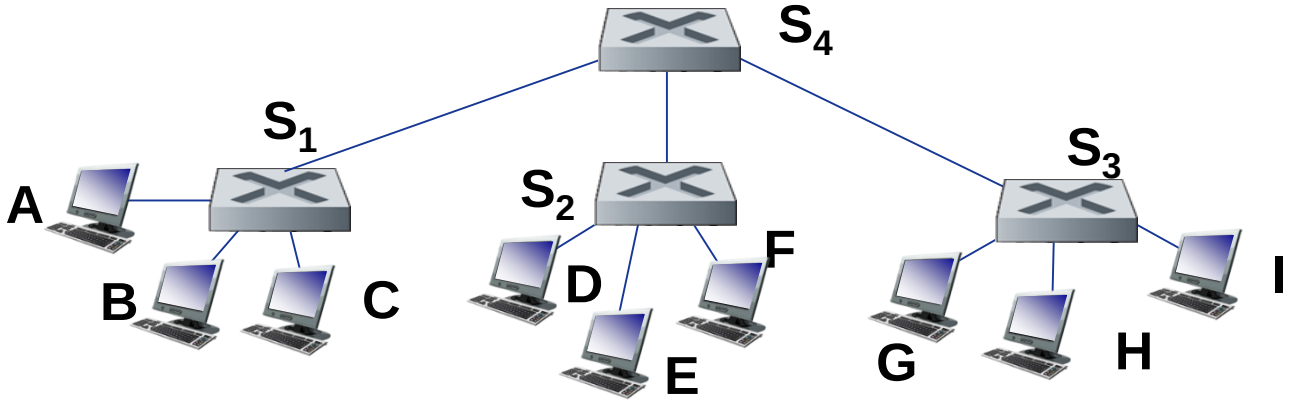


Q: 当A向G发送帧时，S_1怎么知道通过S_4转发？S_4又怎么知道通过S_3转发?

- **A:** 自学习！(工作过程与单一交换机情形相同！)

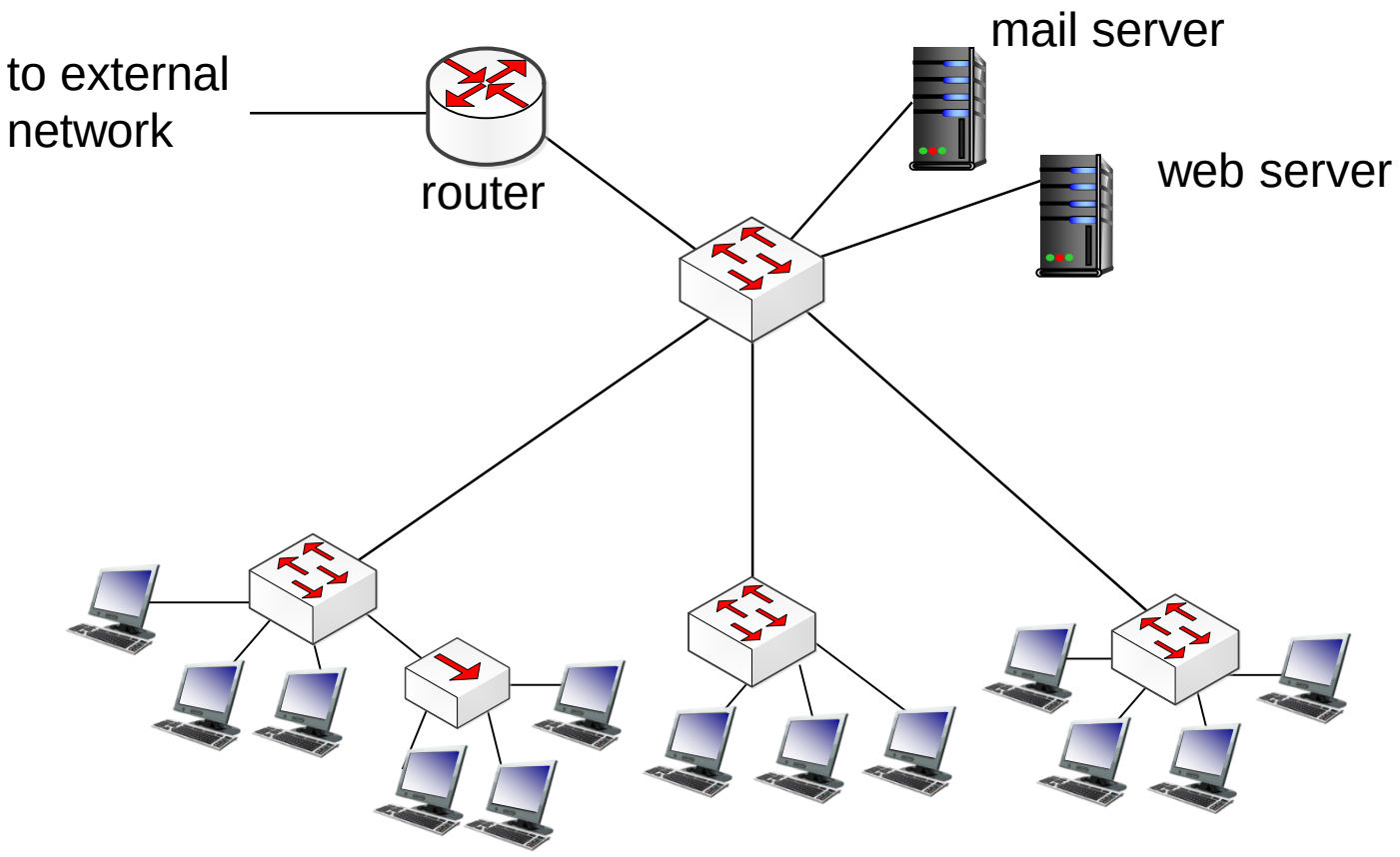
多交换机自学习举例

假设C向H发送帧，H向C发送应答帧

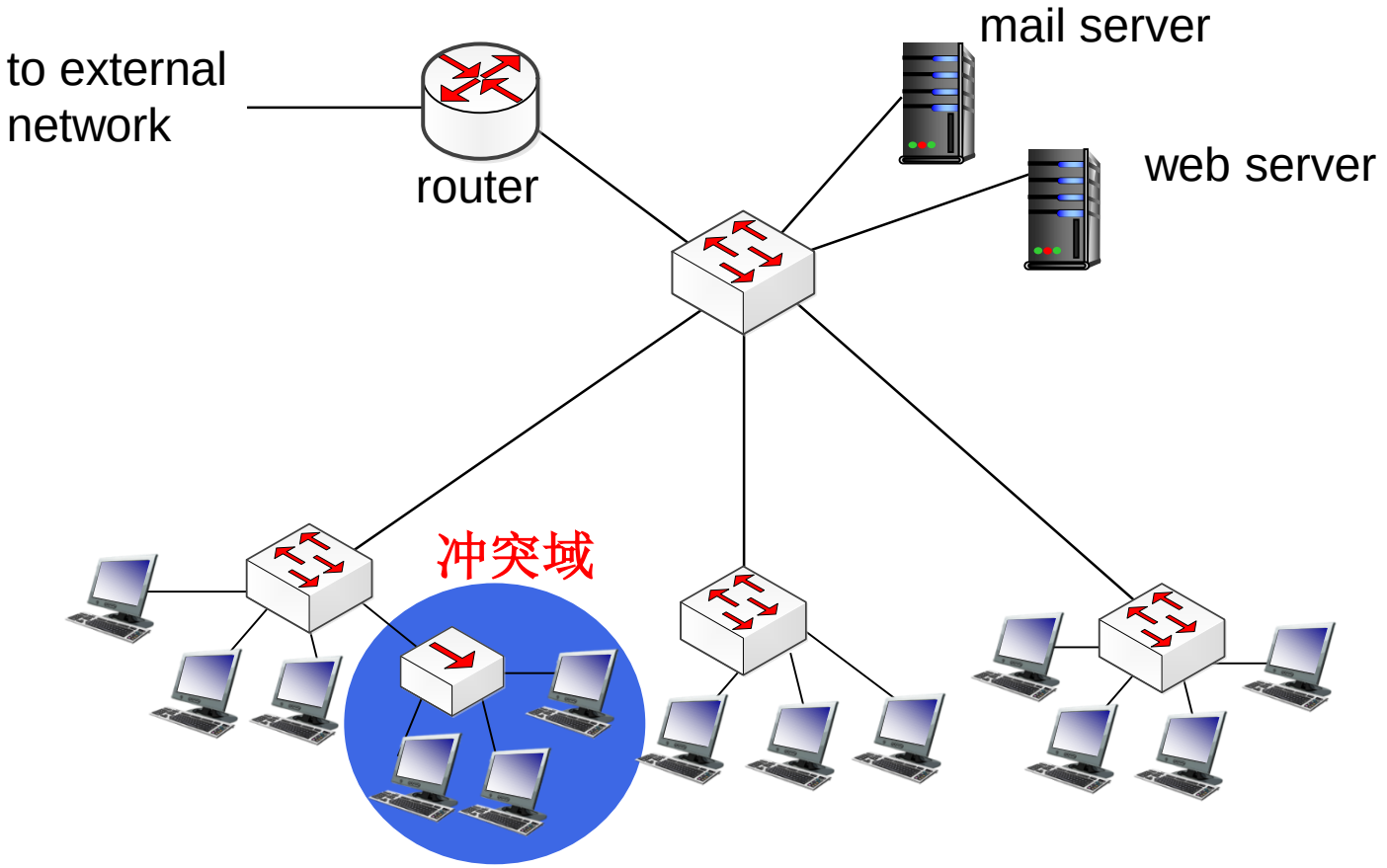


- **Q:** 请给出S_1, S_2, S_3, S_4的交换表，并说明帧的转发过程?

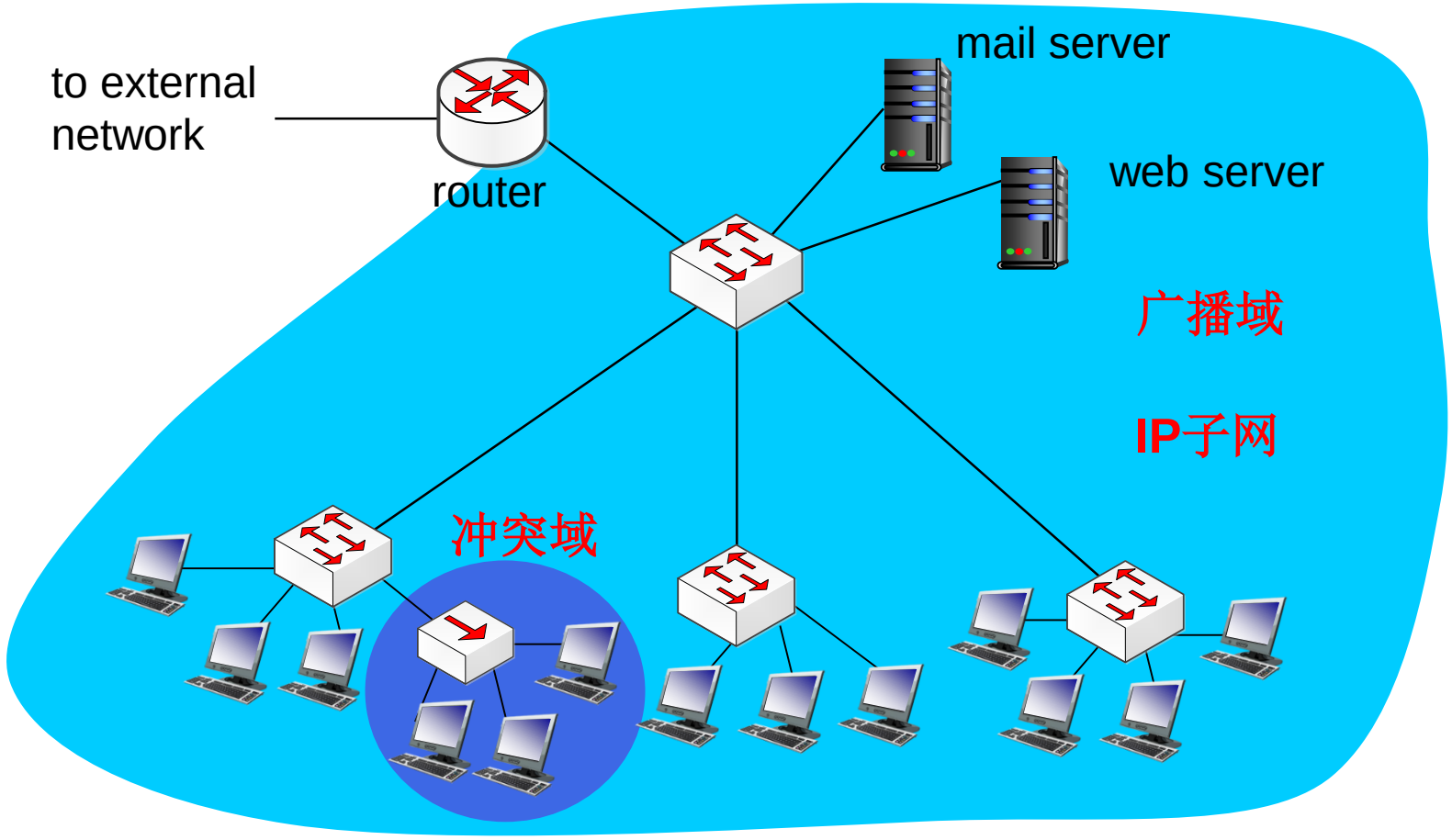
组织机构(Institutional)网络



冲突域？广播域？IP子网？



冲突域？广播域？IP子网？


to external network
router
mail server
web server
广播域
IP子网
冲突域

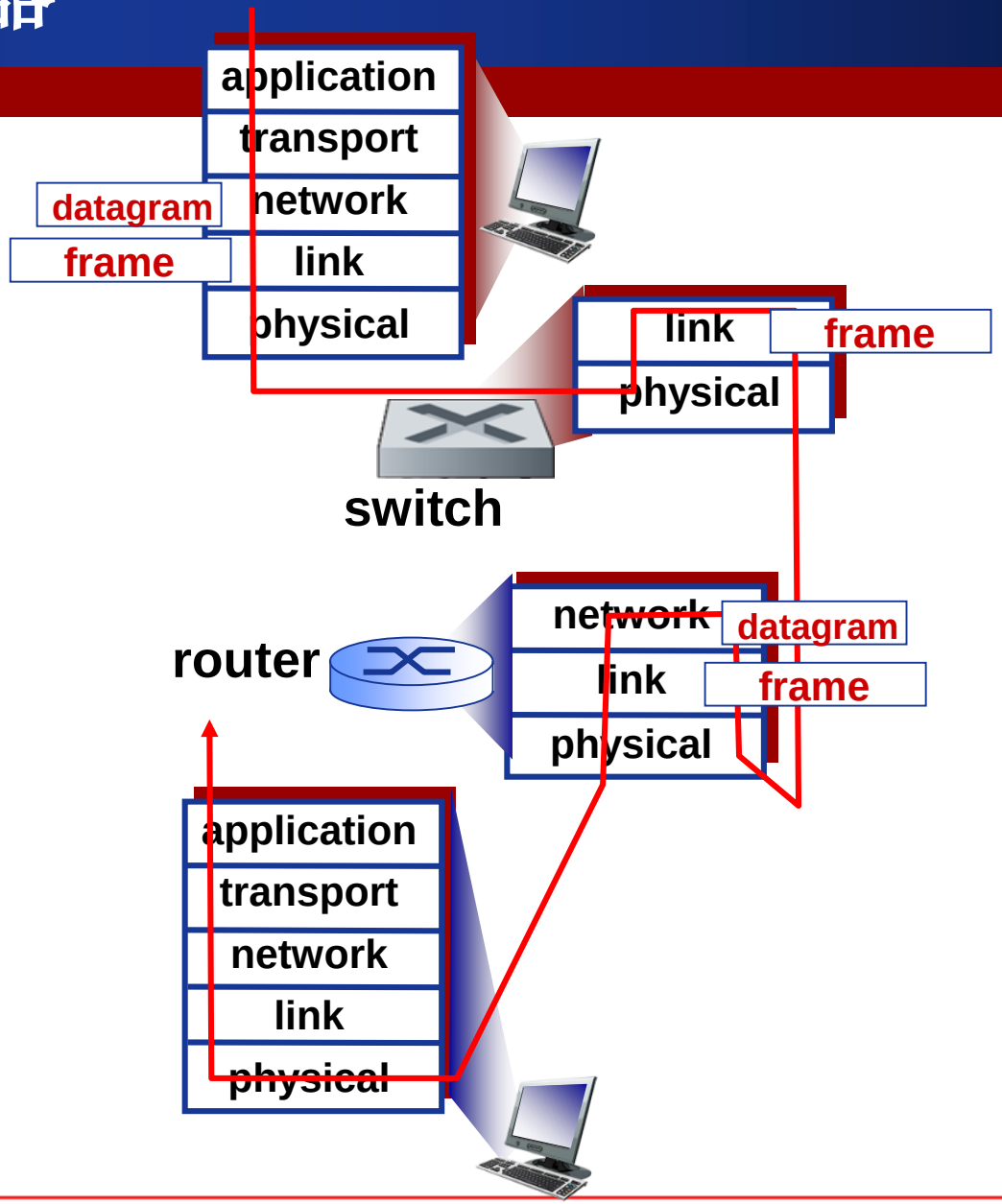
交换机 vs. 路由器

两者均为存储-转发设备:

- 路由器: 网络层设备 (检测网络层分组首部)
- 交换机: 链路层设备 (检测链路层帧的首部)

二者均使用转发表:

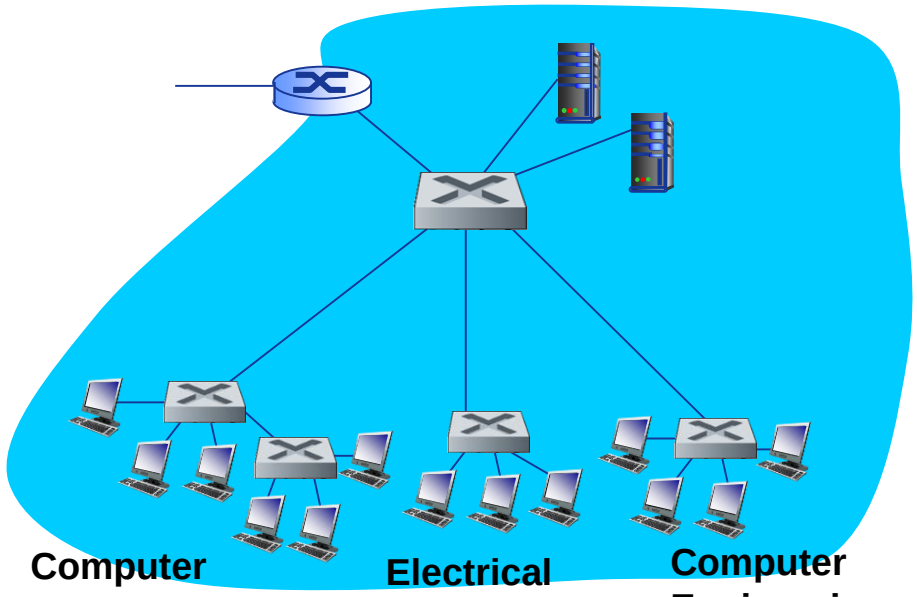
- 路由器: 利用路由算法(路由协议)计算(设置), 依据IP地址
- 交换机: 利用自学习、泛洪构建转发表, 依据MAC地址



网络设备对比

	集线器 (hub)	交换机 (switche)	网桥 (bridge)	路由器 (router)
层次	1	2	2	3
流量(冲突域)隔离	no	yes	yes	yes
广播域隔离	no	no	no	yes
即插即用	yes	yes	yes	no
优化路由	no	no	no	yes
直通传输 (Cut through)	yes	yes	yes	no

VLANs: 动机



考虑一下情形:

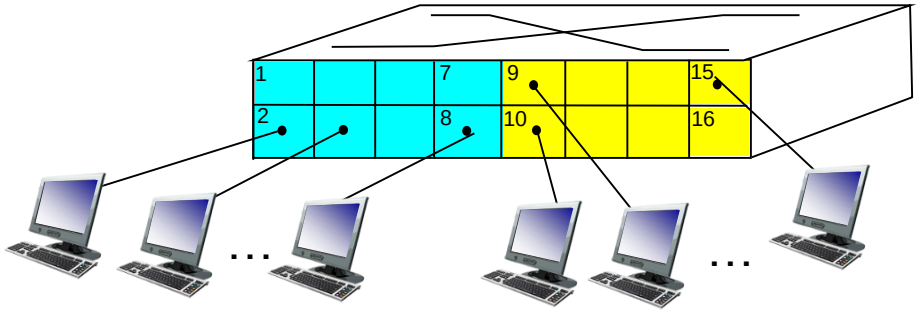
- CS用户迁移到EE，但是希望连接至CS交换机，怎么办?
- 单一广播域:
  - 所有第2层广播流量(ARP, DHCP, 未知目的MAC地址位置)必须穿越整个LAN
  - 安全/隐私、效率问题

VLANs

虚拟局域网(Virtual Local Area Network)

支持VLAN划分的交换机，可以在一个物理LAN架构上配置、定义多个VLAN

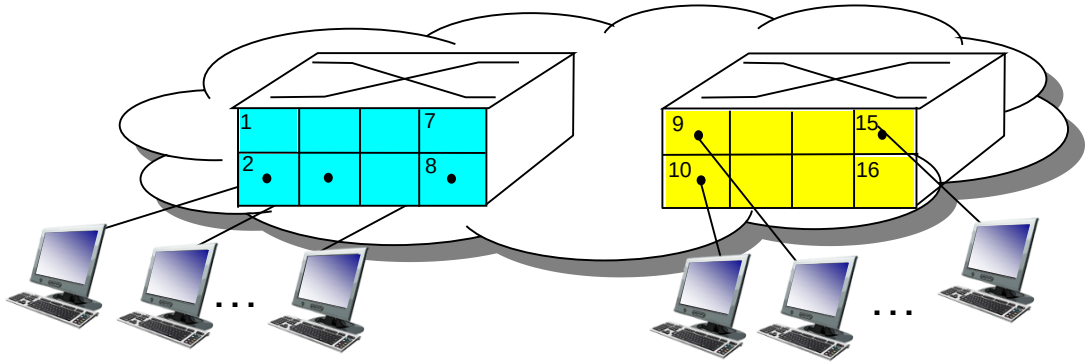
基于端口的VLAN: 分组交换机端口 (通过交换机管理软件)，于是， 单一的物理交换机 ……



Electrical Engineering
(VLAN ports 1-8)

Computer Science
(VLAN ports 9-15)

…就像多个虚拟交换机一样运行

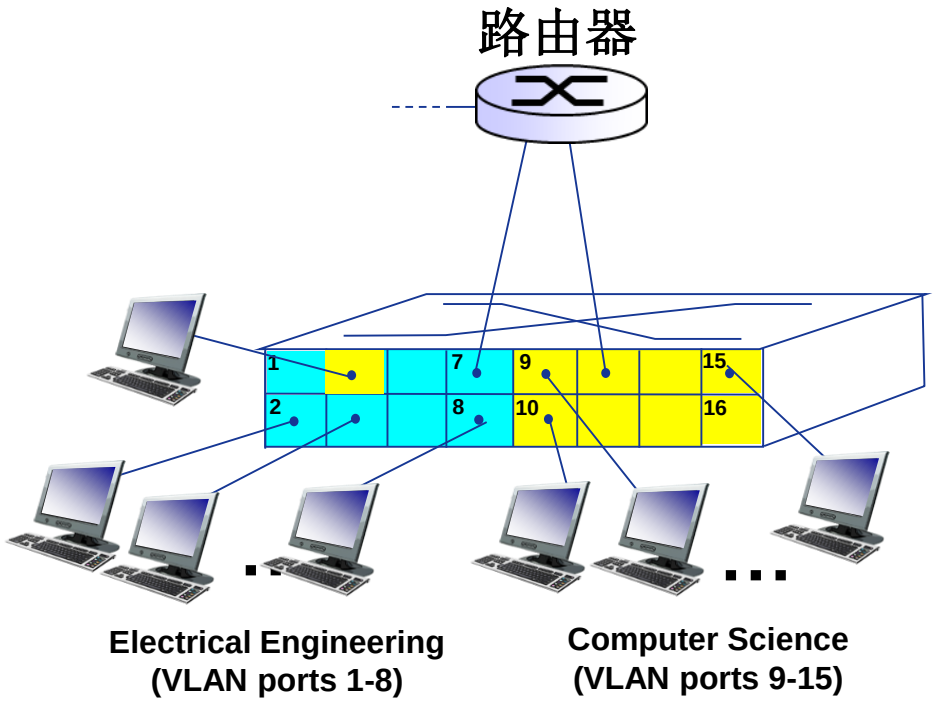


Electrical Engineering
(VLAN ports 1-8)

Computer Science
(VLAN ports 9-16)

基于端口的VLAN

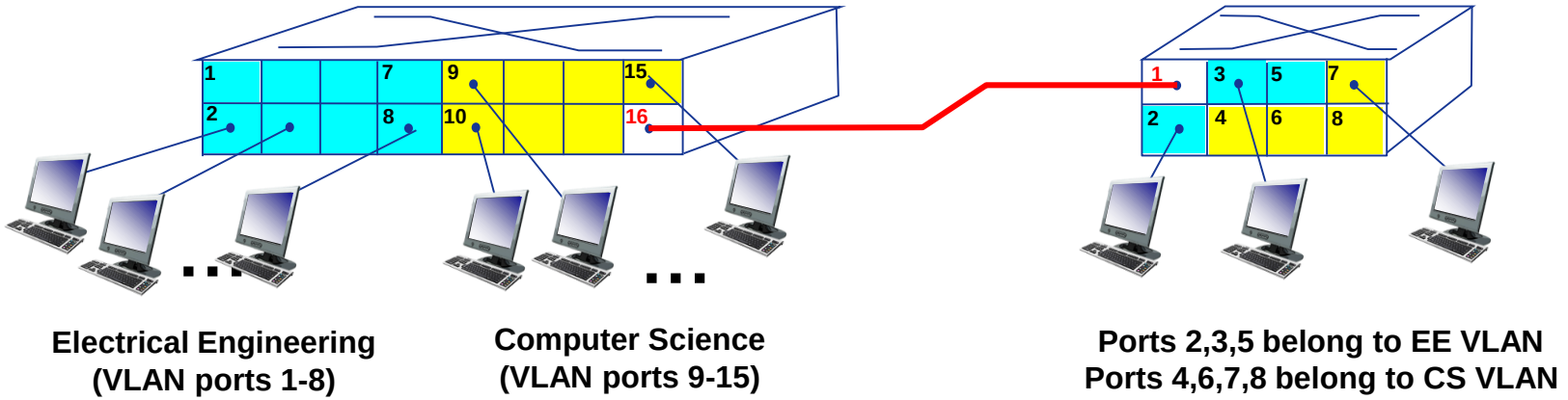
- 流量隔离(traffic isolation): 去往/来自端口1-8的帧只到达端口1-8
  - 也可以基于MAC地址定义VLAN, 而不是交换端口
- 动态成员: 端口可以动态分配给不同VLAN



- 在VLAN间转发: 通过路由(就像在独立的交换机之间)

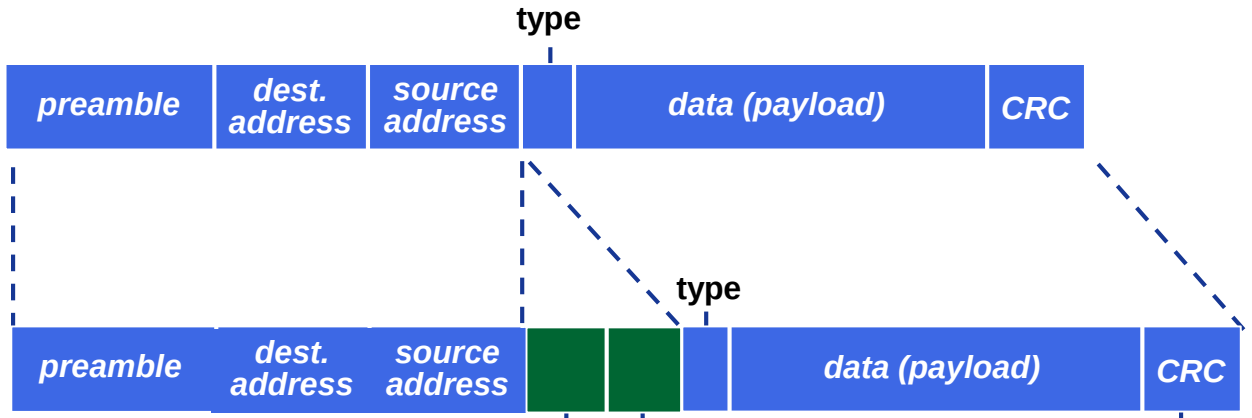
  实践中，厂家会将交换机与路由器集成在一起

跨越多交换机的VLAN



- 多线缆连接
  - 每个线缆连接一个VLAN
- 中继端口(trunk port): 在跨越多个物理交换机定义的VLAN承载帧
  - 为多VLAN转发802.1帧容易产生歧义 (必须携带VLAN ID信息)
  - 802.1q协议为经过中继端口转发的帧增加/去除额外的首部域

802.1Q VLAN帧格式



2字节的标记协议ID
(Tag Protocol Identifier)
(固定值: 0x8100)

重新计算的CRC

2字节的标记控制信息(Tag Control Information)
(12bit VLAN ID字段；3bit优先级字段，类似于IP的TOS)

数据链路层与局域网

- 数据链路层服务
- 差错编码
- 多路访问控制(MAC)协议
  - 信道划分MAC协议
  - 随机访问MAC协议
  - 轮转访问MAC协议
- ARP协议
- 以太网
  - 交换机
  - 虚拟局域网（VLAN）
- PPP协议
- 802.11无线局域网

点对点数据链路控制

- 一个发送端，一个接收端，一条链路：比广播链路容易
  - 无需介质访问控制(Media Access Control)
  - 无需明确的MAC寻址
  - e.g., 拨号链路, ISDN链路
- 常见的点对点数据链路控制协议：
  - HDLC: High Level Data Link Control
  - PPP (Point-to-Point Protocol)

PPP设计需求[RFC 1557]

- 组帧：将网络层数据报封装到数据链路层帧中
  - 可以同时承载任何网络层协议分组(不仅IP数据报)
  - 可以向上层实现分用（多路分解）
- 比特透明传输：数据域必须支持承载任何比特模式
- 差错检测：(无纠正)
- 连接活性(connection liveness)检测：检测、并向网络层通知链路失效
- 网络层地址协商：端结点可以学习/配置彼此网络地址

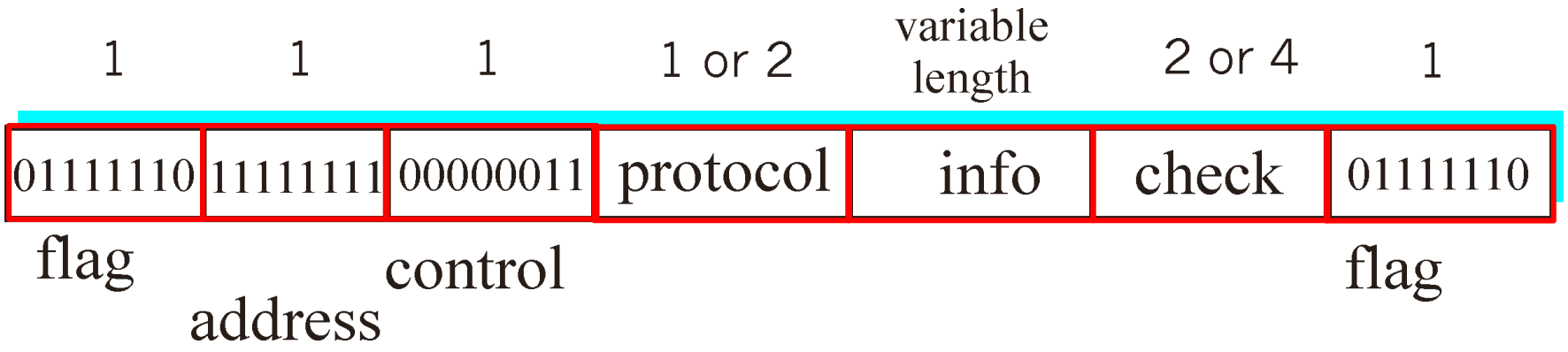
PPP无需支持的功能

- 无需差错纠正/恢复
- 无需流量控制
- 不存在乱序交付
- 无需支持多点链路

差错恢复、流量控制等由高层协议处理！

PPP数据帧

- 标志(Flag): 定界符(delimiter)
- 地址(Address): 无效(仅仅是一个选项)
- 控制(Control): 无效；未来可能的多种控制域
- 协议(Protocol): 上层协议 (eg, PPP-LCP, IP, IPCP, etc)
- 信息(info): 上层协议分组数据
- 校验(check): CRC校验，用于差错检测



字节填充(Byte Stuffing)

- “数据透明传输”需求: 数据域必须允许包含标志模式<01111110>
  - Q: 如何判断该作为数据接收，还是作为标志处理？

- 发送端: 在数据中的<01111110>和<01111101>字节前添加额外的字节<01111101> (“填充(stuffs)”)
- 接收端:
  - 单个字节<01111101>表示一个填充字节；
  - 连续两个字节<01111101>：丢弃第1个，第2个作为数据接收
  - 单个字节<01111110>: 标志字节

字节填充(Byte Stuffing)

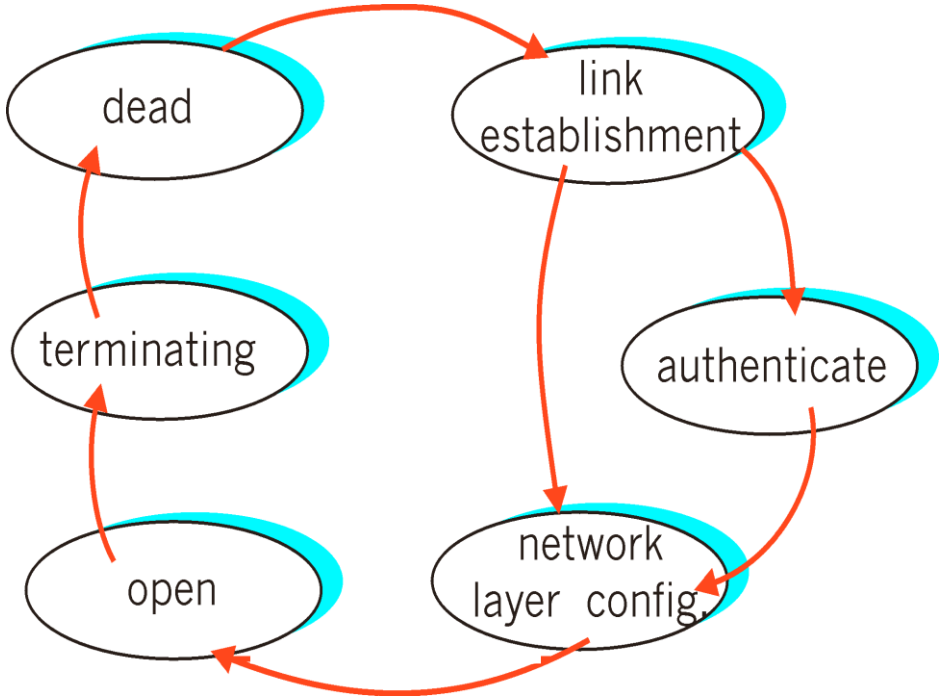
数据中包含
标志(flag)
字节

数据中的标志(flag)字节前，插入填充字节

PPP数据控制协议

在交换网络层数据之前，PPP数据链路两端必须：

- 配置PPP链路
  - 最大帧长
  - 身份认证(authentication)
  - etc.
- 学习/配置网络层信息
  - 对于IP协议: 通过交换IPCP协议 (IP Control Protocol) 报文 (IP分组首部的“上层协议”字段取值: 8021)，完成IP地址等相关信息配置



数据链路层与局域网

- 数据链路层服务
- 差错编码
- 多路访问控制(MAC)协议
  - 信道划分MAC协议
  - 随机访问MAC协议
  - 轮转访问MAC协议
- ARP协议
- 以太网
  - 交换机
  - 虚拟局域网（VLAN）
- PPP协议
- 802.11无线局域网

IEEE 802.11无线局域网

802.11b

- 2.4-2.5GHz免费频段(unlicensed spectrum)
- 最高速率：11 Mbps
- 物理层采用直接序列扩频(DSSS)技术
  - 所有主机使用相同的码片序列

802.11a

- 5-6 GHz频段
- 最高速率：54 Mbps

802.11g

- 2.4-2.5 GHz频段
- 最高速率：54 Mbps

802.11n: 多天线(MIMO)

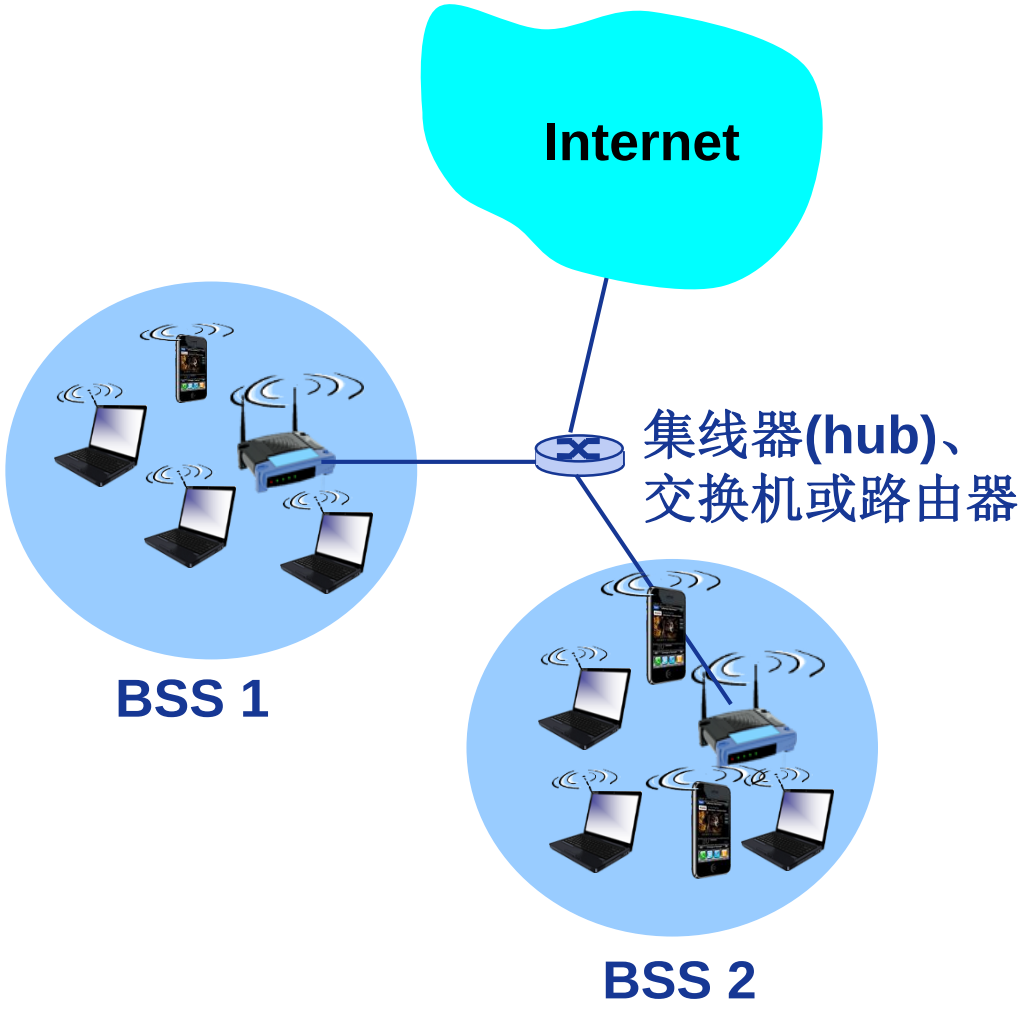
- 2.4-2.5 GHz频段
- 最高速率：600 Mbps

- **均使用CSMA/CA多路访问控制协议**
- **均有基础设施(基站)网络模式和特定网(自组网)网络模式**

IEEE 802.11无线局域网

标准	频段	数据速率	物理层	优缺点
802.11b	2.4 GHz	最高11 Mb/s	扩频	最高数据率较低，价格最低， 信号传播距离最远，且不易受阻碍
802.11a	5 GHz	最高54 Mb/s	OFDM	最高数据率较高，支持更多用户同时上网， 价格最高，信号传播距离较短，且易受阻
802.11g	2.4 GHz	最高54 Mb/s	OFDM	最高数据率较高，支持更多用户同时上网， 信号传播距离最远，且不易受阻， 价格比802.11b贵
802.11n	2.4 GHz 5 GHz	最高600 Mb/s	MIMO OFDM	使用多个发射和接收天线以允许更高的数据传输率，当使用双倍带宽(40 MHz)时速率可达600 Mb/s

IEEE 802.11体系结构

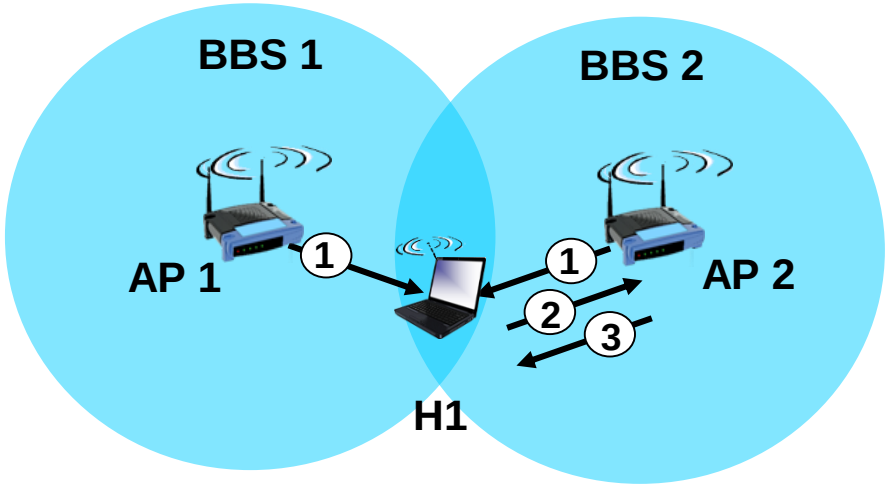


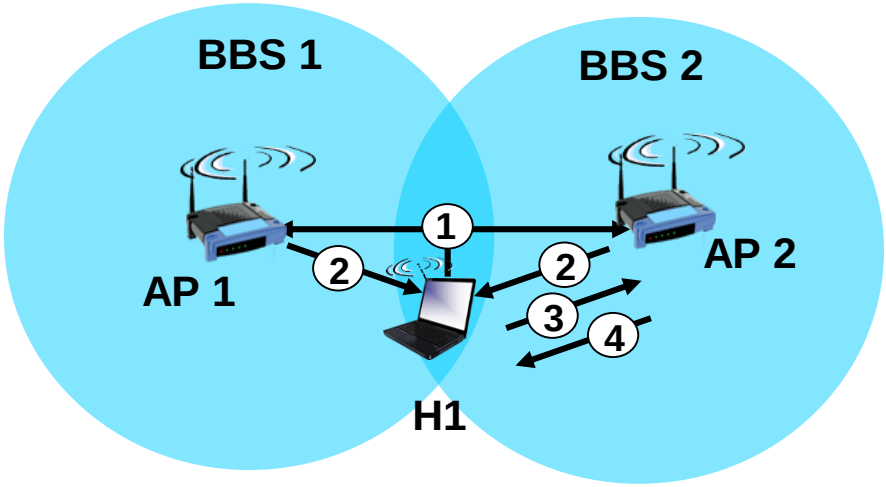
- 无线主机与基站通信
  - 基站(base station) = 访问点(access point-AP)
- 基本服务集BSS(Basic Service Set)，也称为单元(cell)
  - 基础设施网络模式:
    - 无线主机
    - AP: 基站
  - 自组网(ad hoc)模式:
    - 只有主机

802.11：信道与AP关联

- 802.11b: 2.4GHz-2.485GHz频谱划分为11个不同频率的信道
  - 每个AP选择一个频率(信道)
  - 存在干扰可能: 相邻的AP可能选择相同的信道！
- 主机: 必须与某个AP关联(associate)
  - 扫描信道，监听包含AP名称(服务集标识符-SSID)和MAC地址的信标(beacon)帧
  - 选择一个AP进行关联
  - 可能需要进行身份认证
  - 典型情形：运行DHCP获取IP地址等信息

802.11AP关联：被动扫描与主动扫描





被动扫描(scanning):

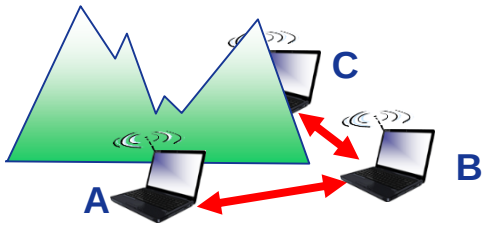
- 各AP发送信标帧
- 主机(H1)向选择的AP发送关联请求帧
- AP向主机(H1)发送关联响应帧

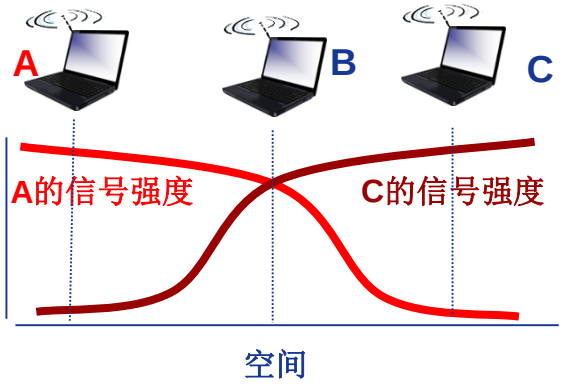
主动扫描:

- 主机(H1)主动广播探测请求帧(Probe Request Frame)
- AP发送探测响应帧(Probe Response Frame)
- 主机(H1)向选择的AP发送关联请求帧
- AP向主机(H1)发送关联响应帧

802.11：多路访问控制

- 避免冲突: 2^{+}结点同时传输
- 802.11: CSMA – 发送数据前监听信道
  - 避免与正在进行传输的其他结点冲突
- 802.11: 不能像CSMA/CD那样，边发送、边检测冲突！
  - 无线信道很难实现
  - 无法侦听到所有可能的冲突：隐藏站、信号衰落
  - 目标: 避免冲突(avoid collisions)-CSMA/C(ollision)A(voidance)





IEEE 802.11 MAC协议: CSMA/CA

802.11 sender

1 if 监听到信道空闲了**DIFS时间** then

发送整个帧(无同时检测冲突，即CD)

2 if 监听到信道忙 then

开始随机退避计时

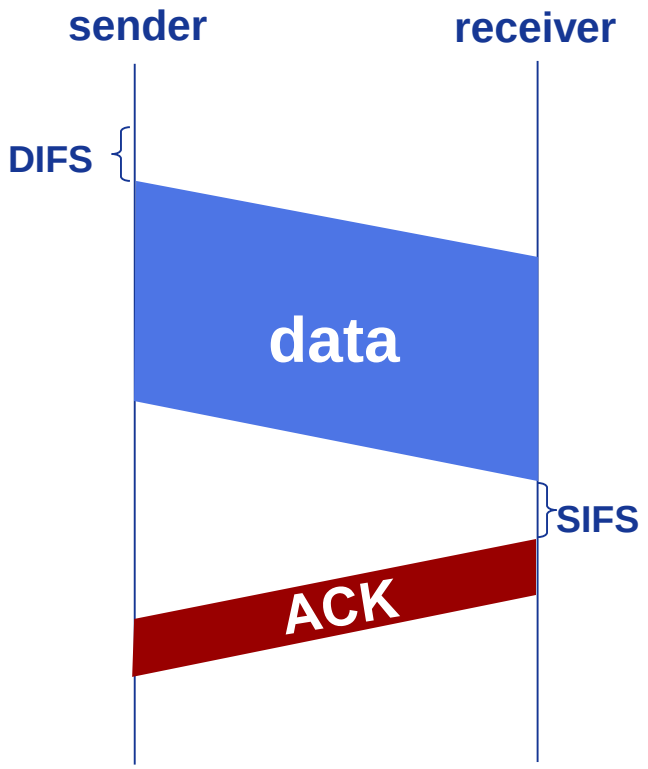
当信道空闲时，计时器倒计时

当计时器超时时，发送帧

if 没有收到ACK then

增加随机退避间隔时间

重复第2步



802.11 receiver

- if 正确接收帧

延迟SIFS时间后，向发送端发送ACK

(由于存在隐藏站问题)

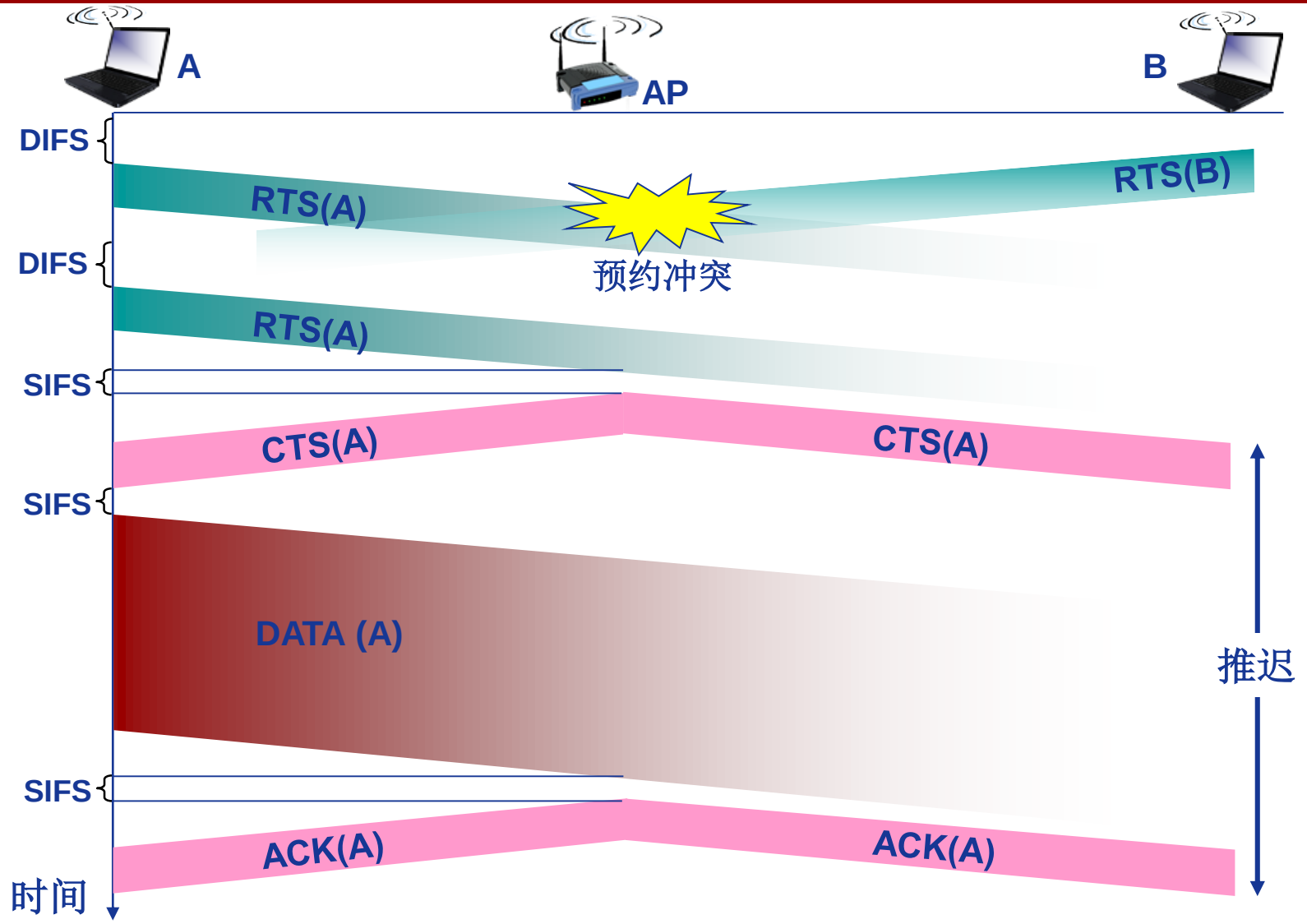
IEEE 802.11 MAC协议: CSMA/CA

基本思想：允许发送端“预约”(reserve)信道，而不是随机发送数据帧，从而避免长数据帧的冲突

- 发送端首先利用CSMA向BS发送一个很短的RTS (request-to-send)帧
  - RTS帧仍然可能彼此冲突 (但RTS帧很短)
- BS广播一个CTS(clear-to-send)帧作为对RTS的响应
- CTS帧可以被所有结点接收
  - 消除隐藏站影响
  - 发送端可以发送数据帧
  - 其他结点推迟发送

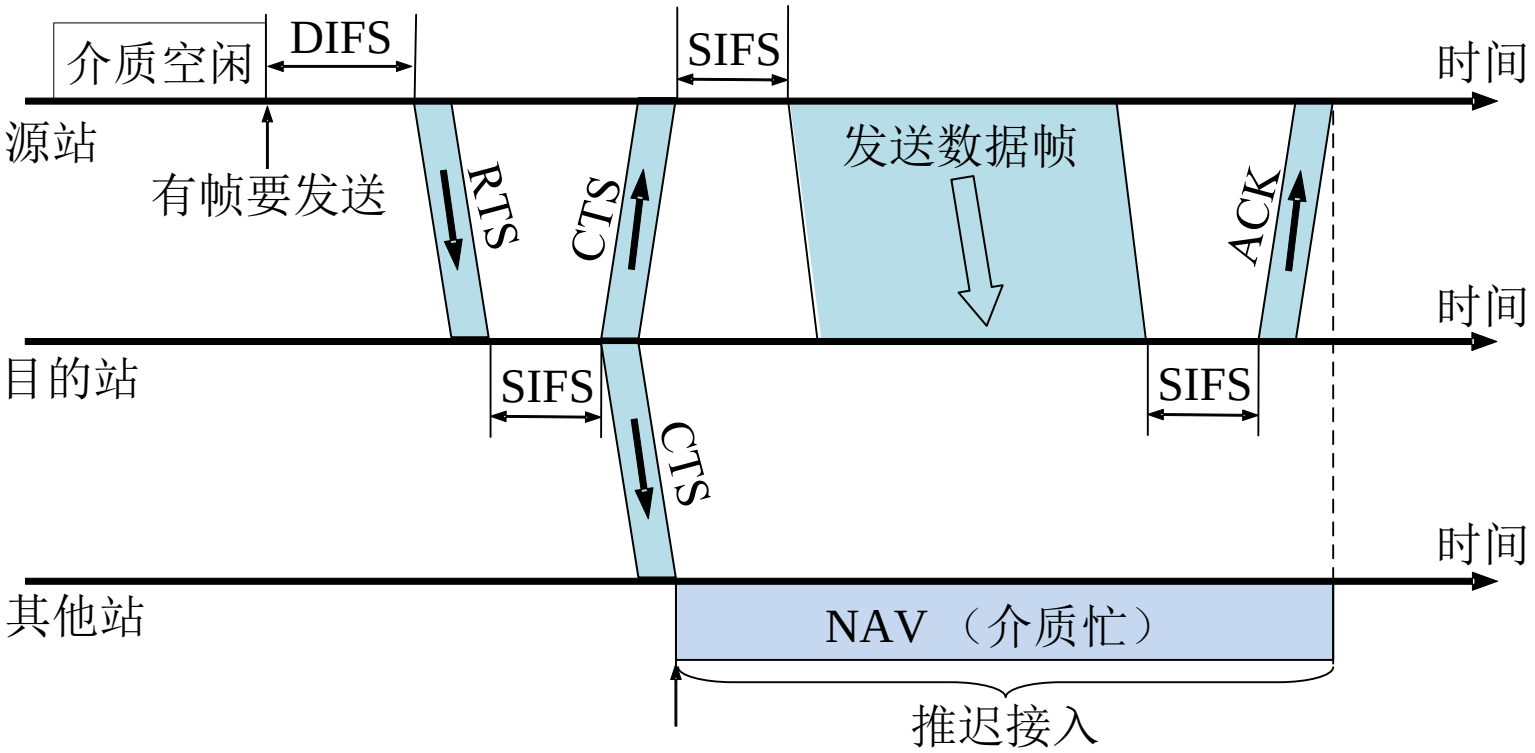
利用很小的预约帧彻底避免了数据帧冲突！

冲突避免(CA): RTS-CTS交换

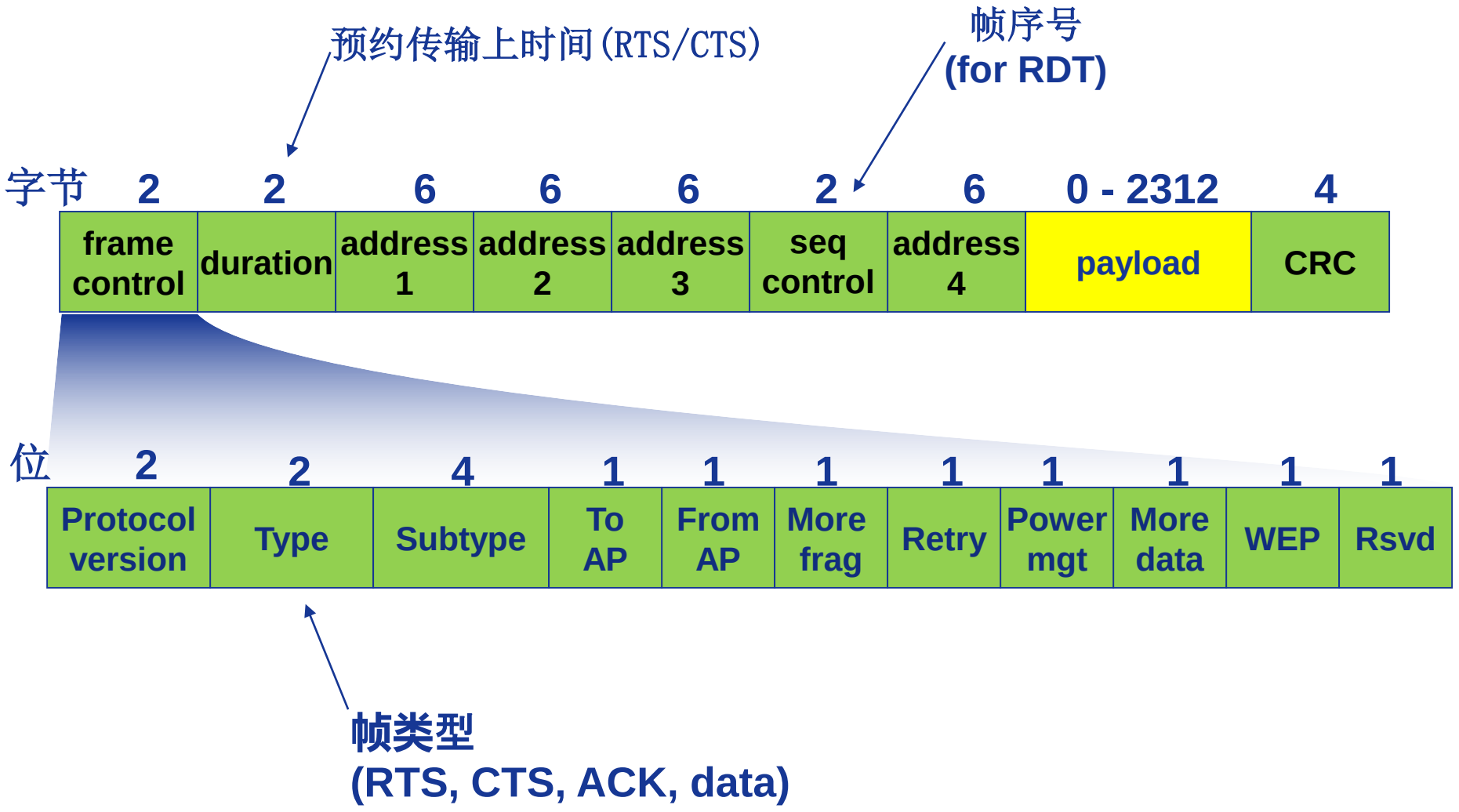


例

- CSMA/CA协议如何实现信道预约？一个主机从期望发送数据开始，到确认数据被接收方正确接收最快需要多少时间？（注：忽略传播时延、处理时延和传输时延）
  - DIFS+3SIFS



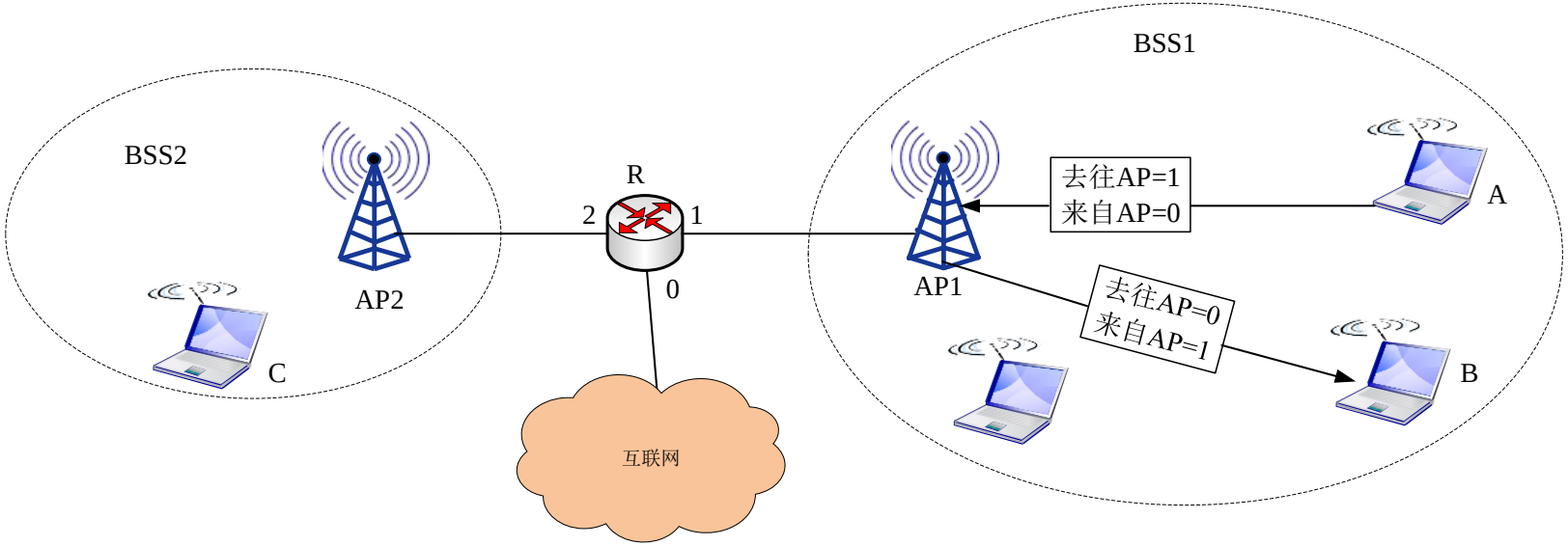
IEEE 802.11 MAC帧



IEEE 802.11数据帧地址

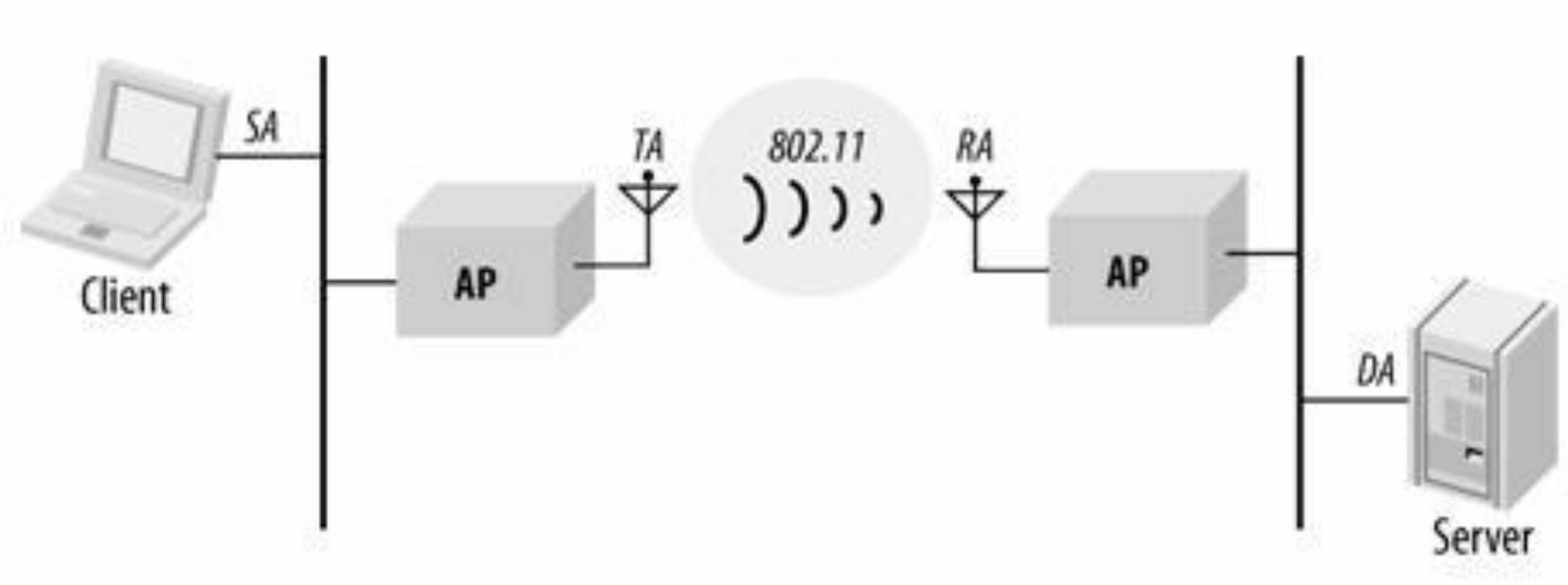
- 802.11数据帧有4个地址字段
- 地址 4 用于采用无线桥接（无线DS）等情形
- IEEE802.11帧中的几个地址分别如何取值？

去往AP	来自AP	地址1	地址2	地址3	地址4
0	1	目的地址	AP地址	源地址	——
1	0	AP地址	源地址	目的地址	——



IEEE 802.11数据帧地址

无线分布系统WDS

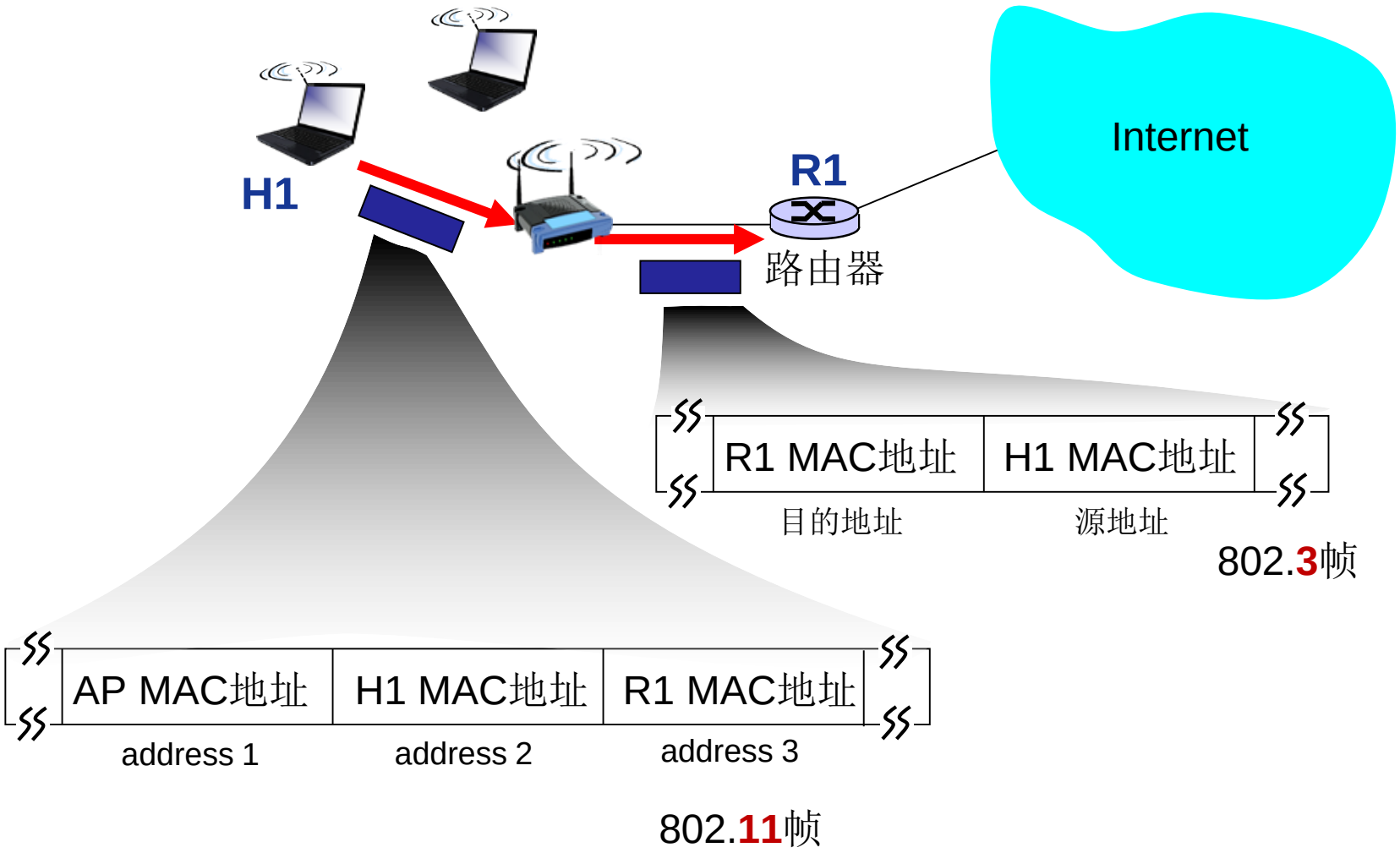


IEEE 802.11数据帧地址

- 802.11 数据帧四个地址字段取值

功能	去往AP	来自AP	地址 1	地址 2	地址 3	地址 4
IBSS	0	0	DA (目的地址)	SA (源地址)	BSSID	——
架构网络	0	1	DA	AP 地址 (BSSID)	SA	——
	1	0	AP 地址 (BSSID)	SA	DA	——
WDS	1	1	RA	TA	DA	SA

IEEE 802.11数据帧地址



哈爾濱工業大學
HARBIN INSTITUTE OF TECHNOLOGY

立足航天，服务国防，面向国民经济主战场

谢 谢！