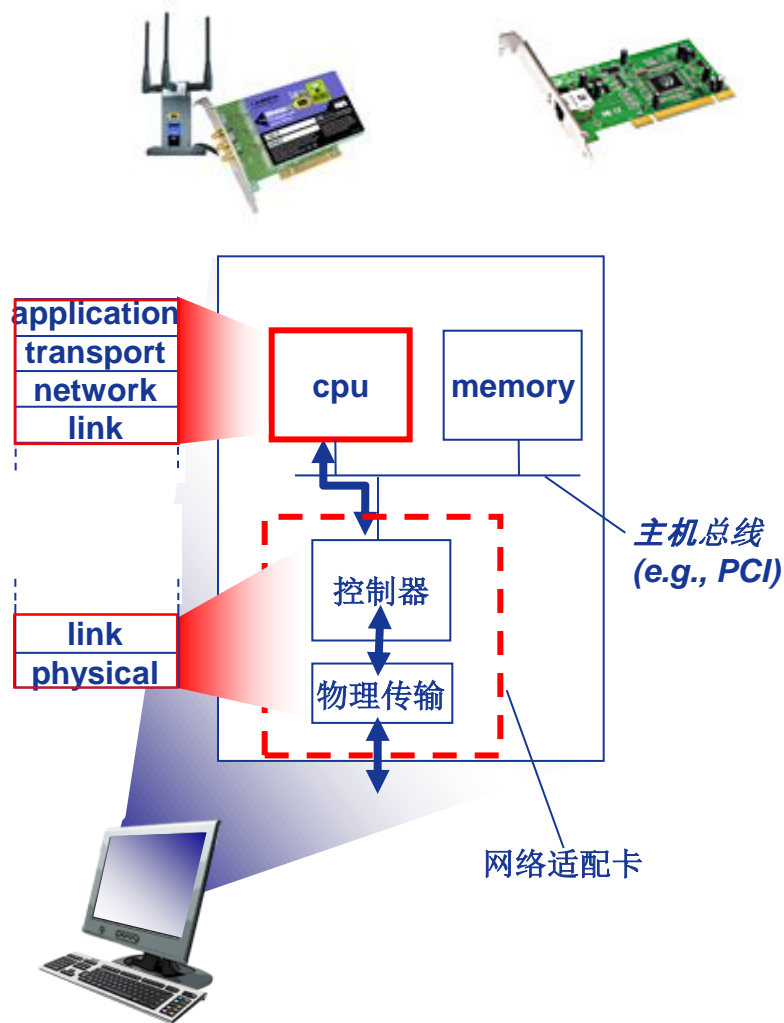


本讲主题

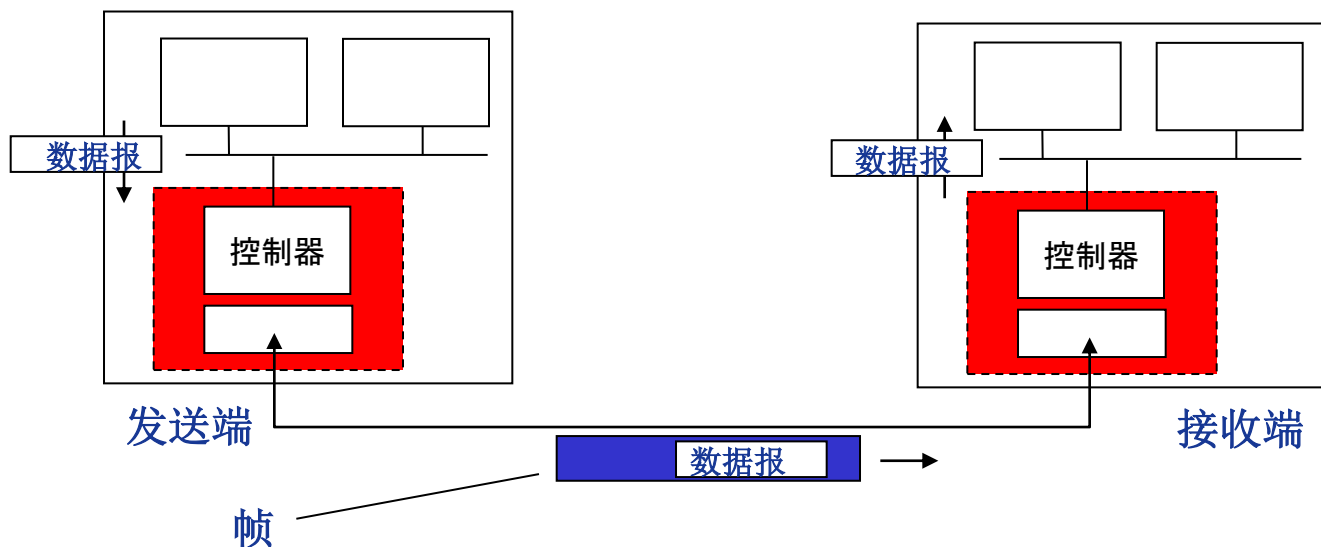
数据链路层服务

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

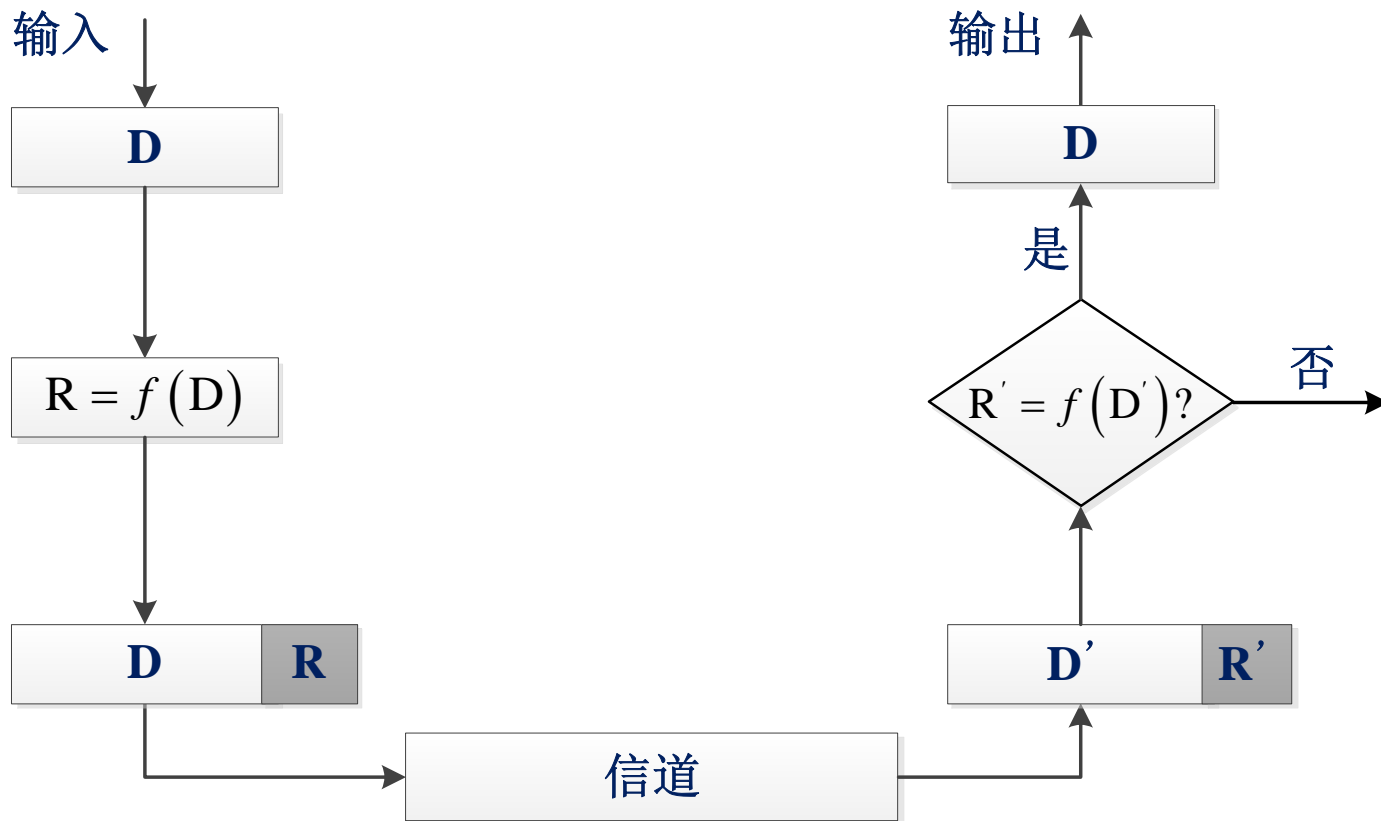
本讲主题

差错编码

差错检测： 差错编码

差错编码基本原理：

$D \rightarrow DR$ ，其中R为差错检测与纠正比特（冗余比特）

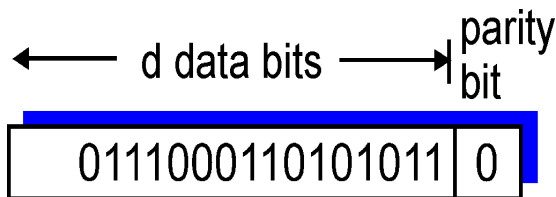


差错编码不能保证100%可靠！

奇偶校验码

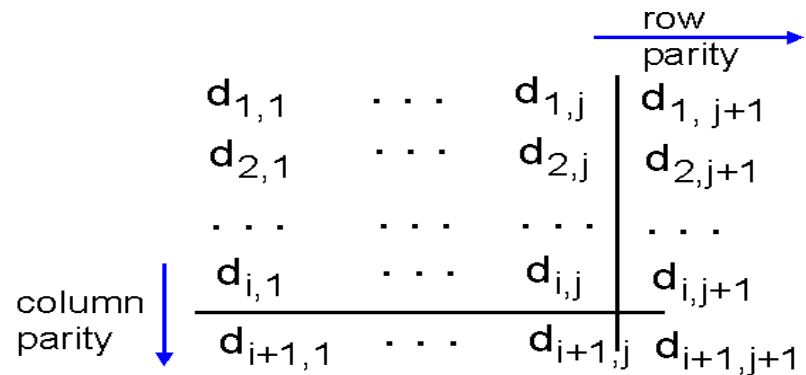
1 比特校验位:

- ❖ 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



1	0	1	0	1	1
1	1	1	1	0	0
0	1	1	1	0	1
0	0	1	0	1	0

no errors

1	0	1	0	1	1
1	0	1	1	0	0
0	1	1	1	0	1
0	0	1	0	1	0

parity error

correctable
single bit error

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
 - $\langle D, R \rangle$ 刚好可以被 G 整除(模2)
 - 接收端检错: 利用 G 除 $\langle D, R \rangle$, 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:

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

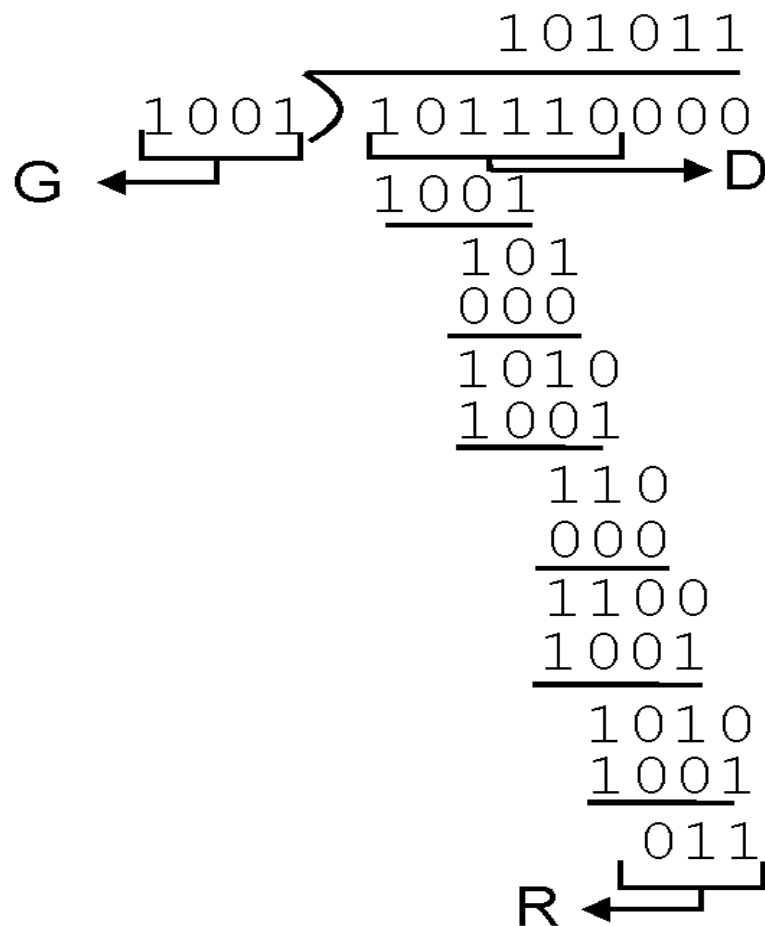
相当于:

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

相当于:

如果利用G去除 $D \cdot 2^r$, 则
余式即为R:

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



本讲主题

多路访问控制(MAC)协议

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

MAC协议分类

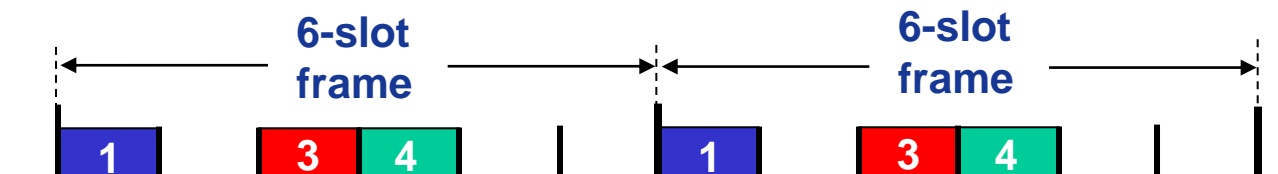
三大类:

- ❖ 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

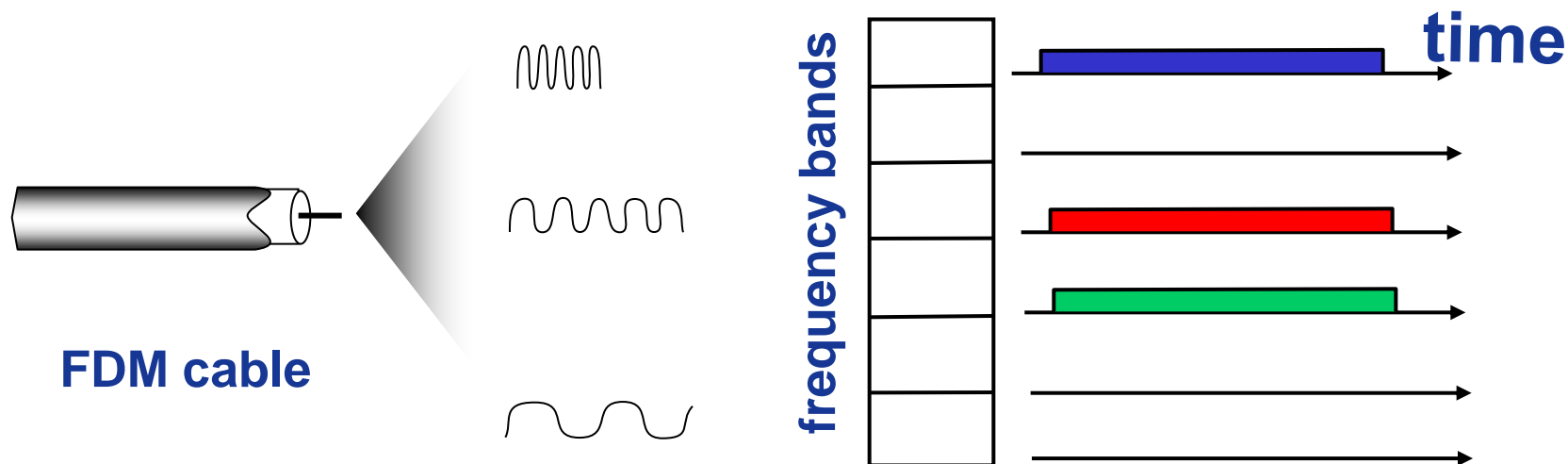
- ❖ “周期性” 接入信道
- ❖ 每个站点在每个周期， 占用固定长度的时隙(e.g.长度=分组传输时间)
- ❖ 未用时隙空闲(idle)
- ❖ 例如： 6-站点LAN， 1,3,4传输分组， 2,5,6空闲



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。



本讲主题

随机访问**MAC**协议（1）

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(slotted)ALOHA
 - ALOHA
 - CSMA、CSMA/CD、CSMA/CA

时隙ALOHA协议

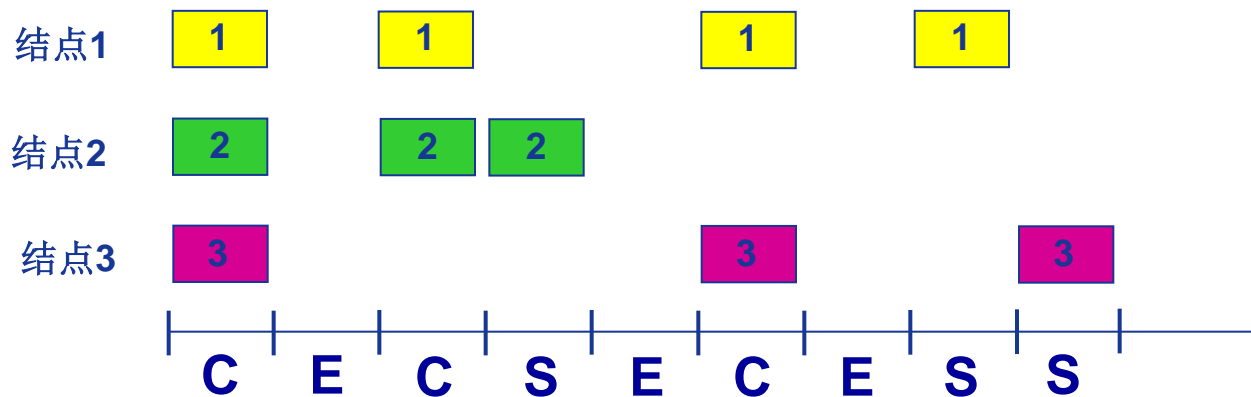
假定:

- ❖ 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:

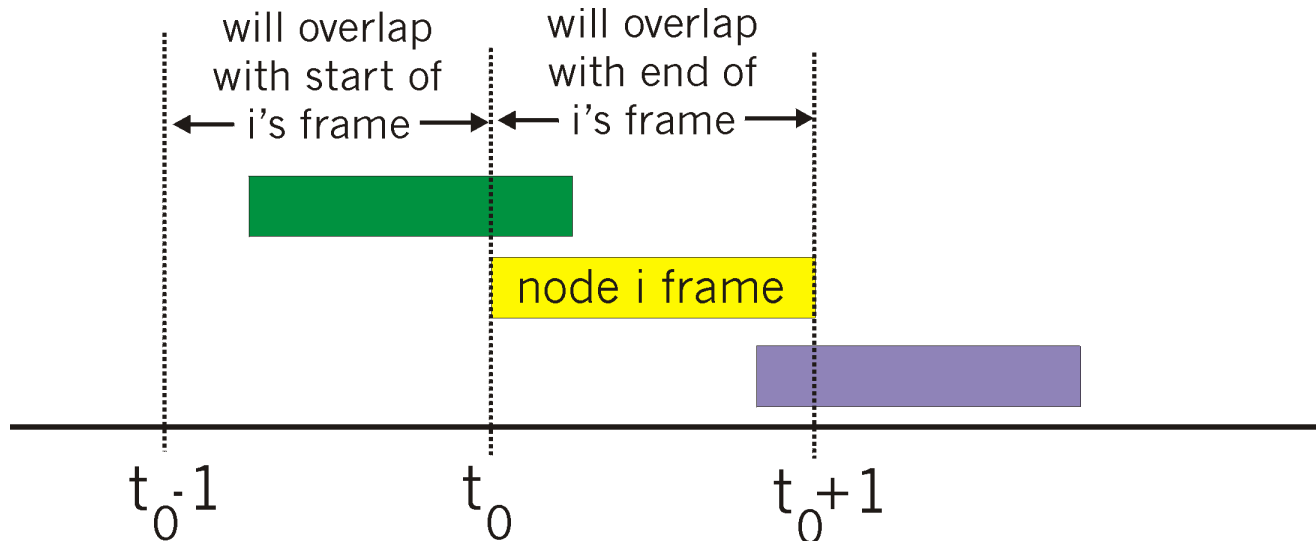
- ❖ 单个结点活动时，可以连续以信道全部速率传输数据
- ❖ 高度分散化：只需同步时隙
- ❖ 简单

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

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



本讲主题

随机访问**MAC**协议（2）

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

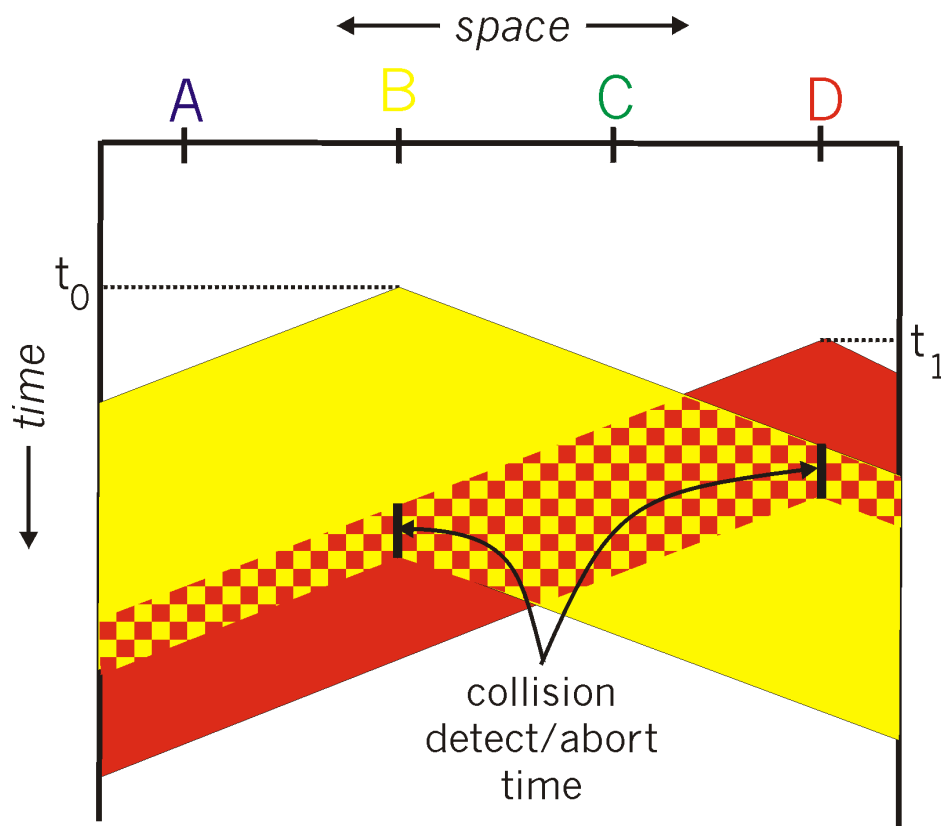
CSMA/CD协议

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

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

❖ 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！

本讲主题

轮转访问**MAC**协议

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

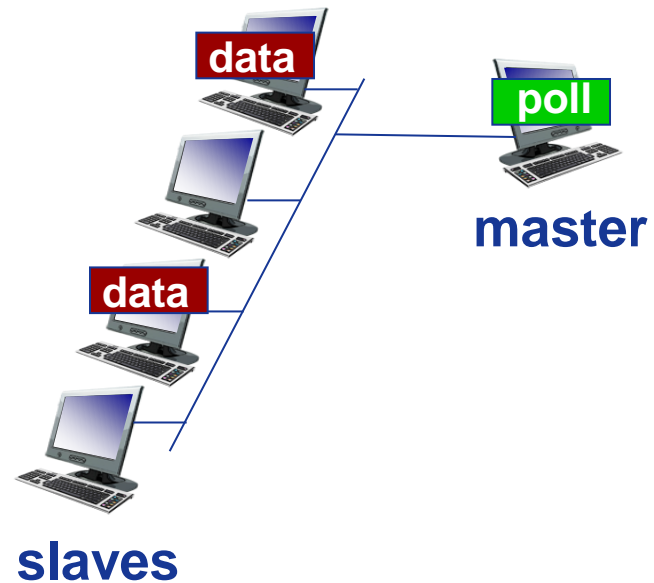
轮询(polling):

- ❖ 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

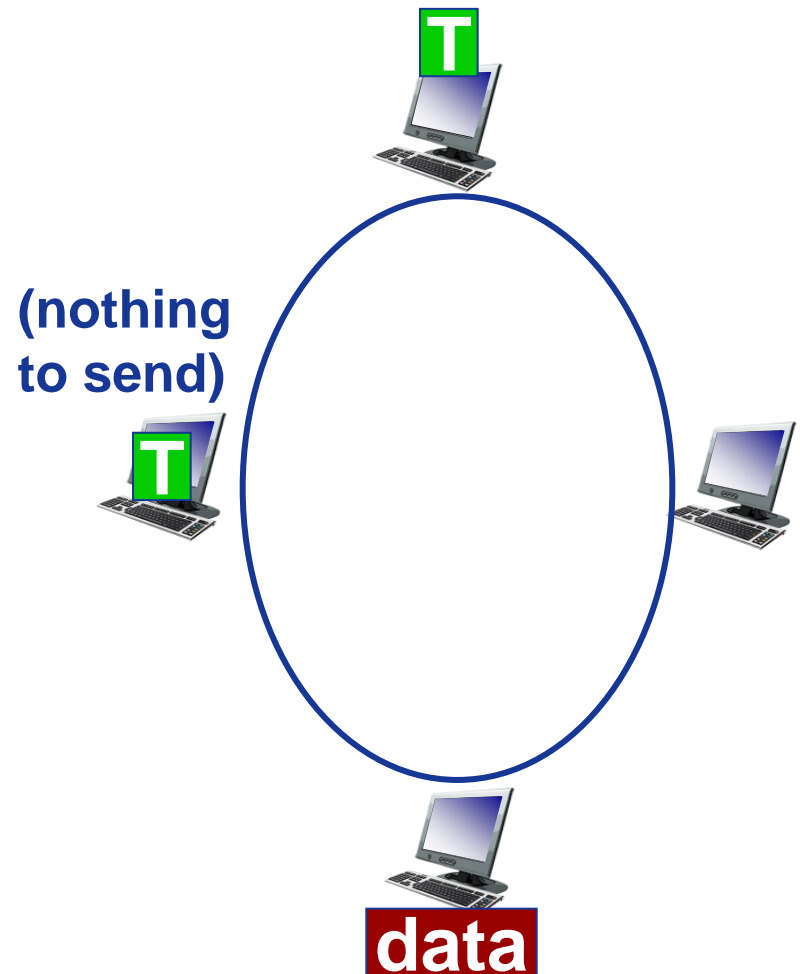
令牌传递(token passing):

- ❖ 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

本讲主题

ARP协议（1）

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

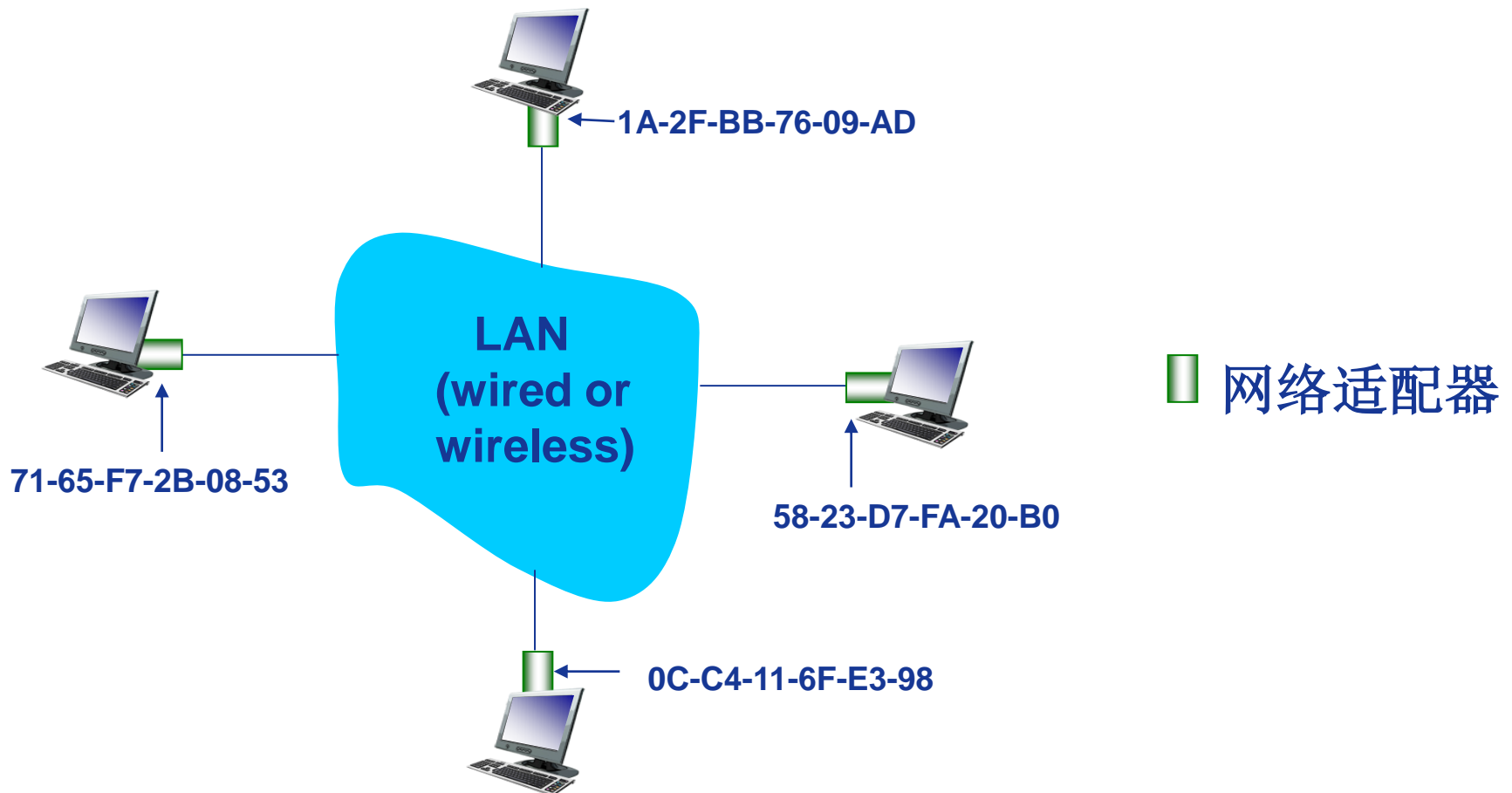
❖ MAC地址(或称LAN地址,物理地址,以太网地址):

- 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**

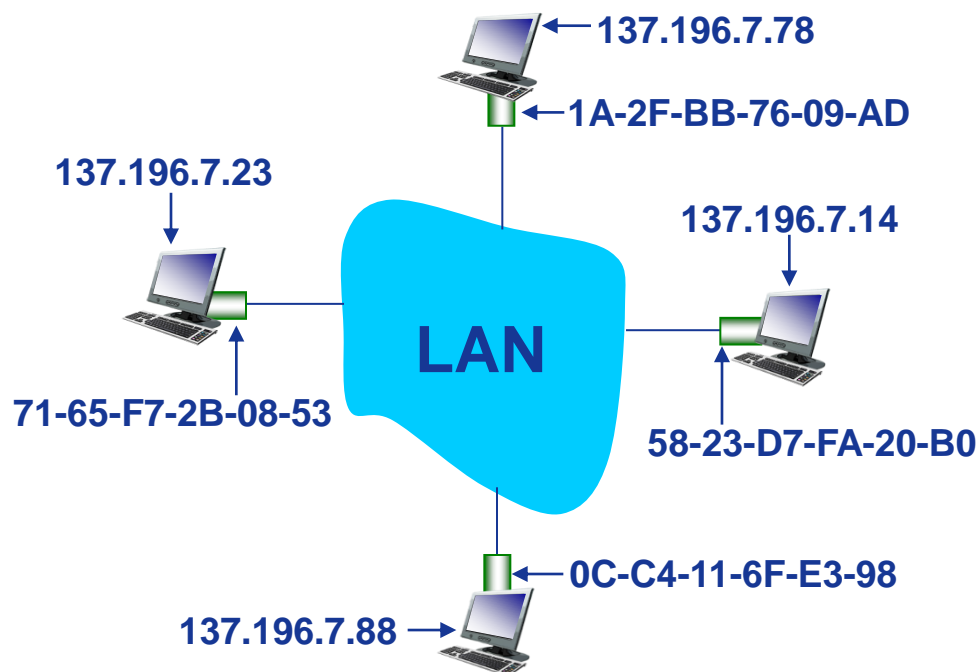


MAC地址

- ❖ 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

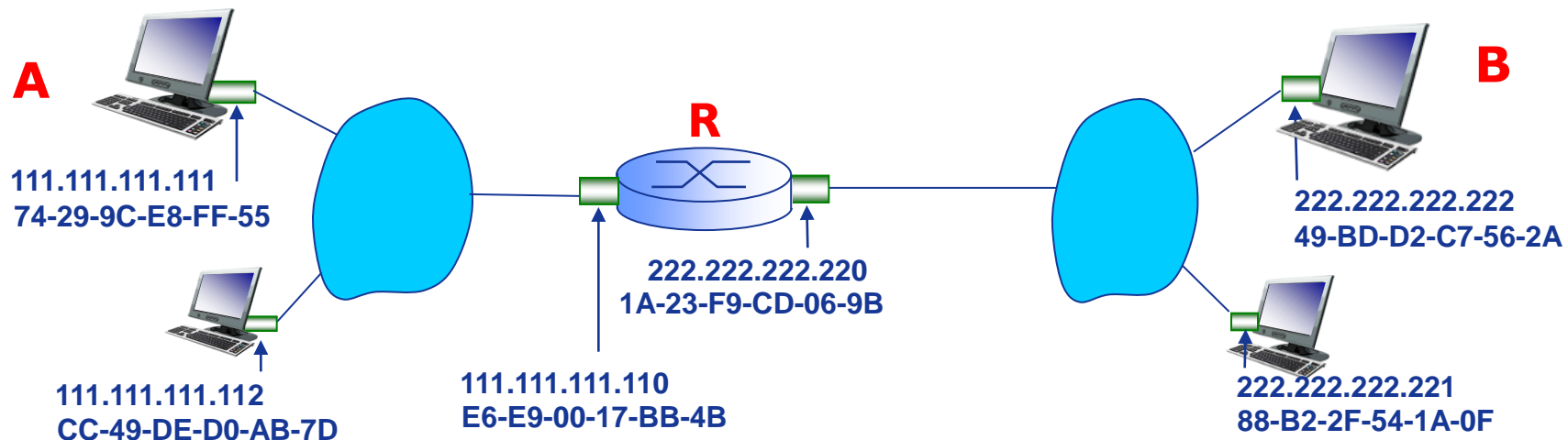
本讲主题

ARP协议（2）

寻址: 从一个LAN路由至另一个LAN

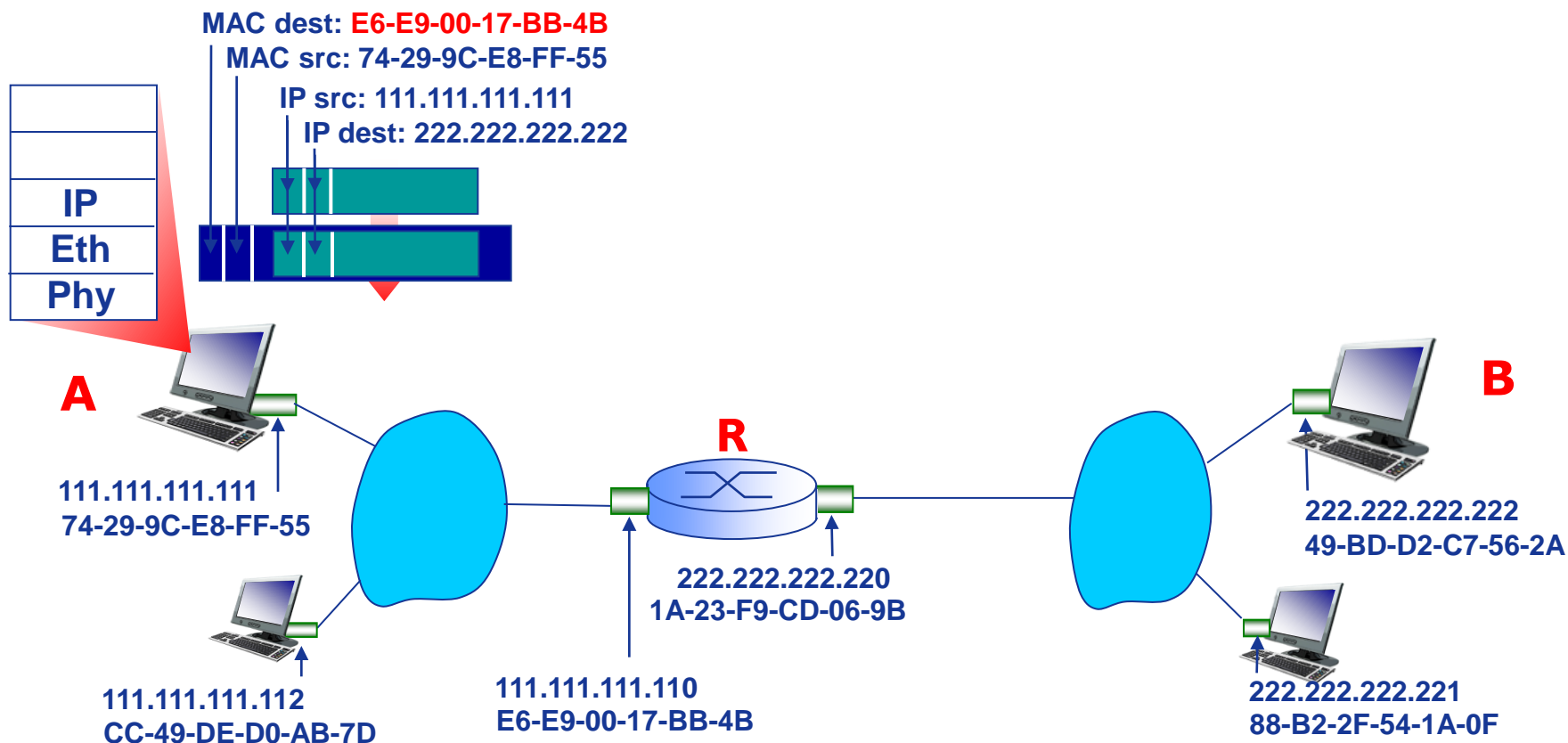
通信过程: **A**通过路由器**R**向**B**发送数据报

- 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?)



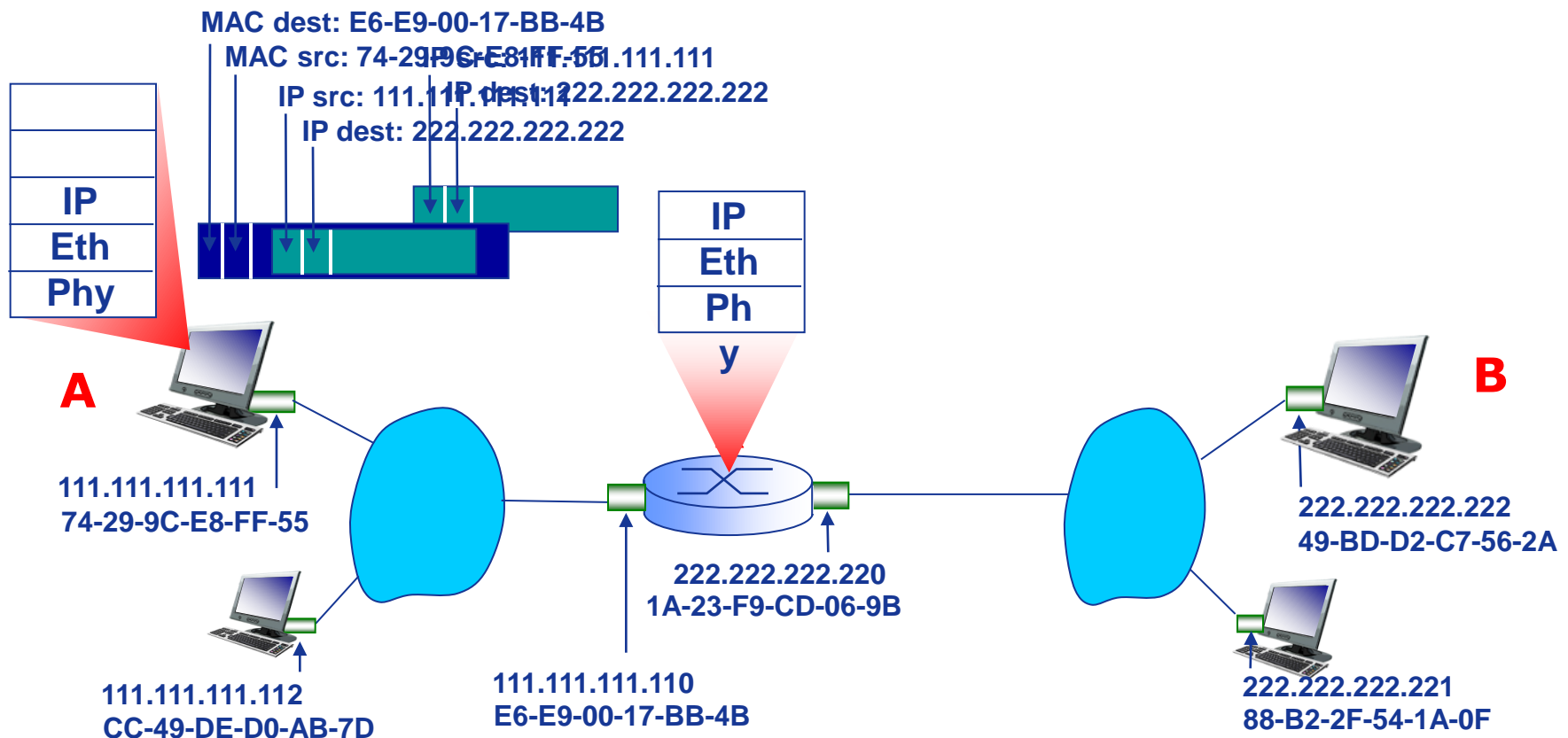
寻址：从一个LAN路由至另一个LAN

- ❖ 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。



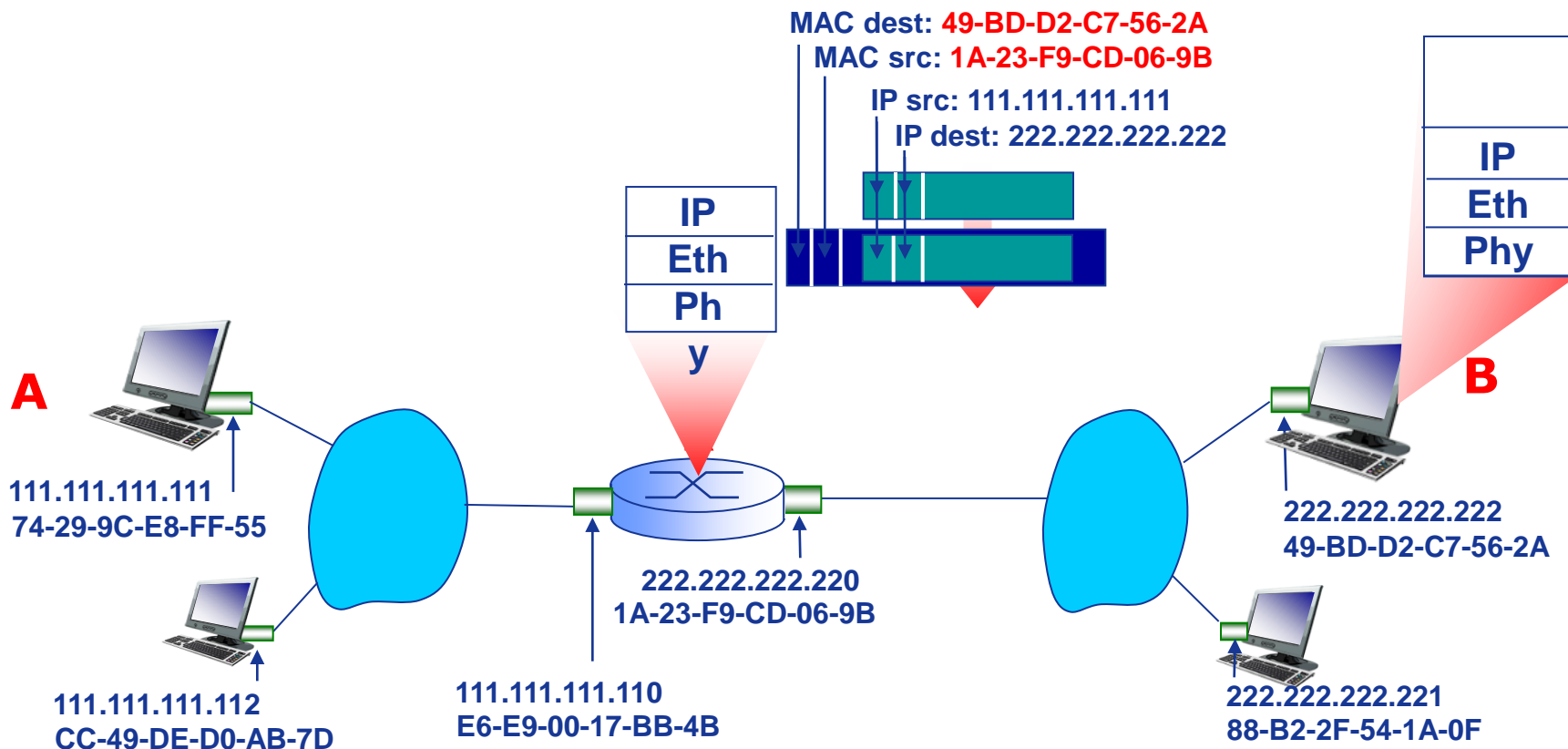
寻址：从一个LAN路由至另一个LAN

- ❖ 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



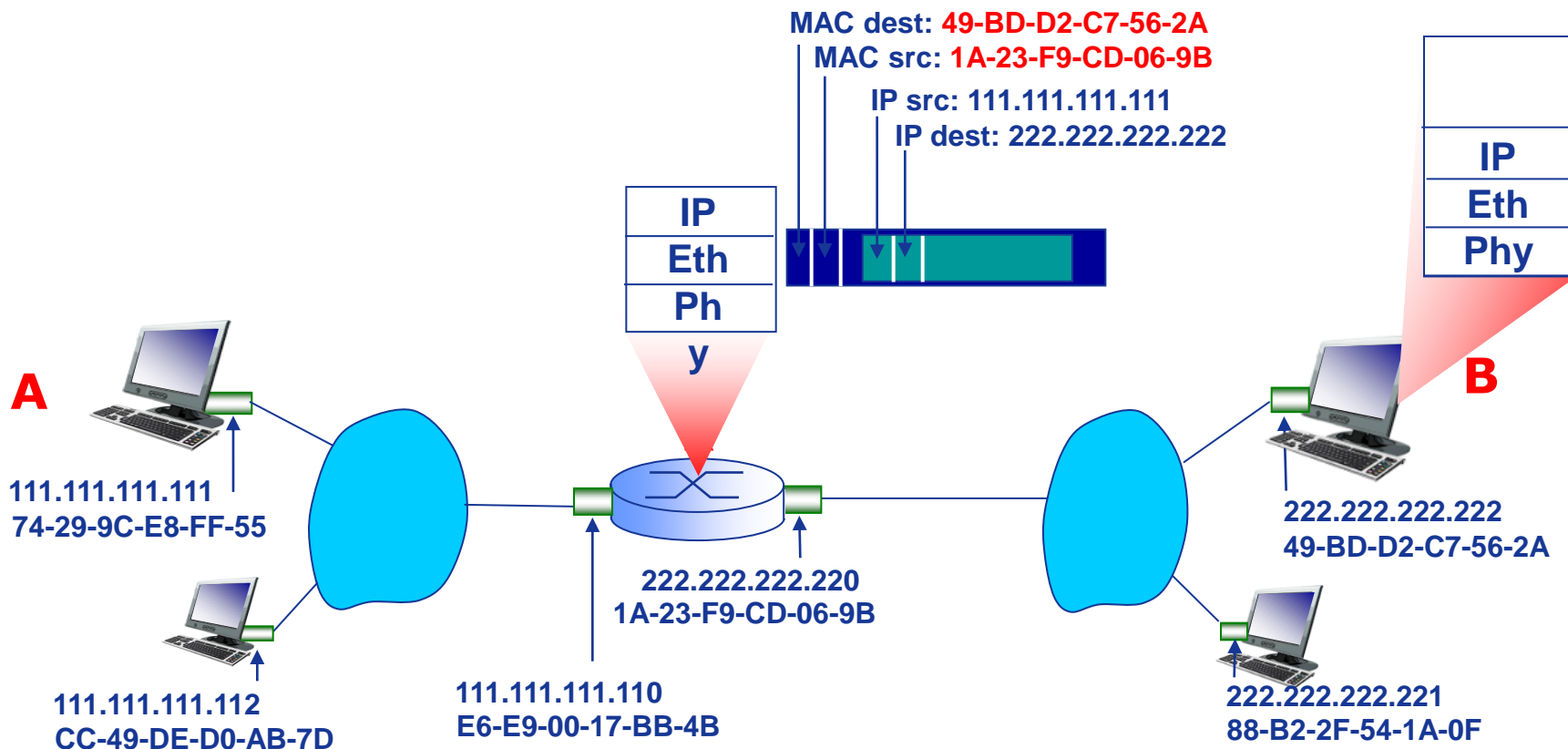
寻址：从一个LAN路由至另一个LAN

- ❖ 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。



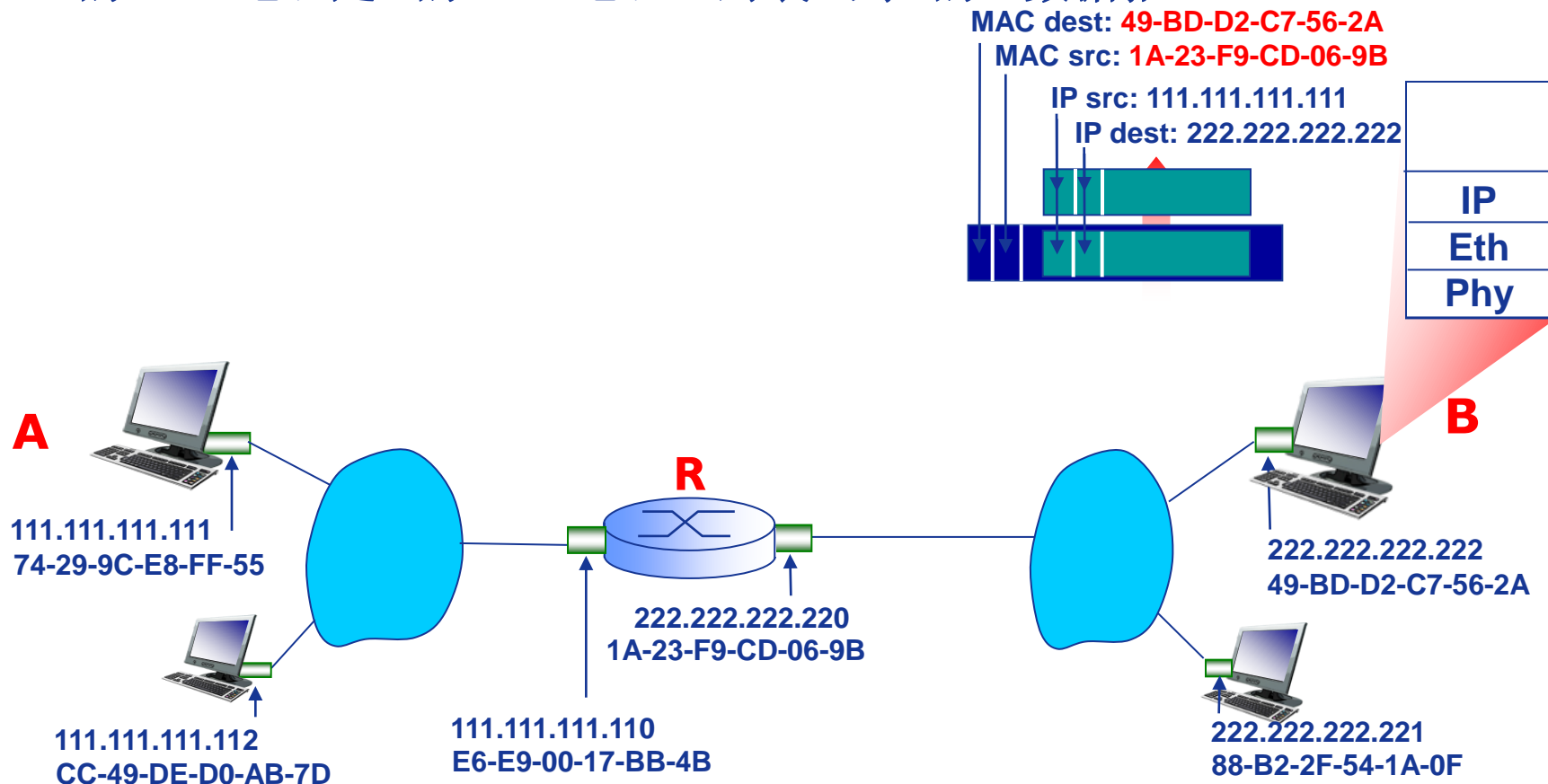
寻址：从一个LAN路由至另一个LAN

- ❖ 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。



本讲主题

以太网

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 $n = \text{Min}(m, 10)$
 - NIC 从 $\{0, 1, 2, \dots, 2^n - 1\}$ 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。

- ❖ $R=10\text{Mbps}$, $RTT_{\max}=512\mu\text{s}$, $L_{\min} / R = RTT_{\max}$
- ❖ $L_{\min}=512\text{bits}=64\text{B}$, $\text{Data}_{\min}=L_{\min}-18=46\text{B}$

❖ CRC(4B): 循环冗余校验码

- 丢弃差错帧



本讲主题

交换机（1）

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

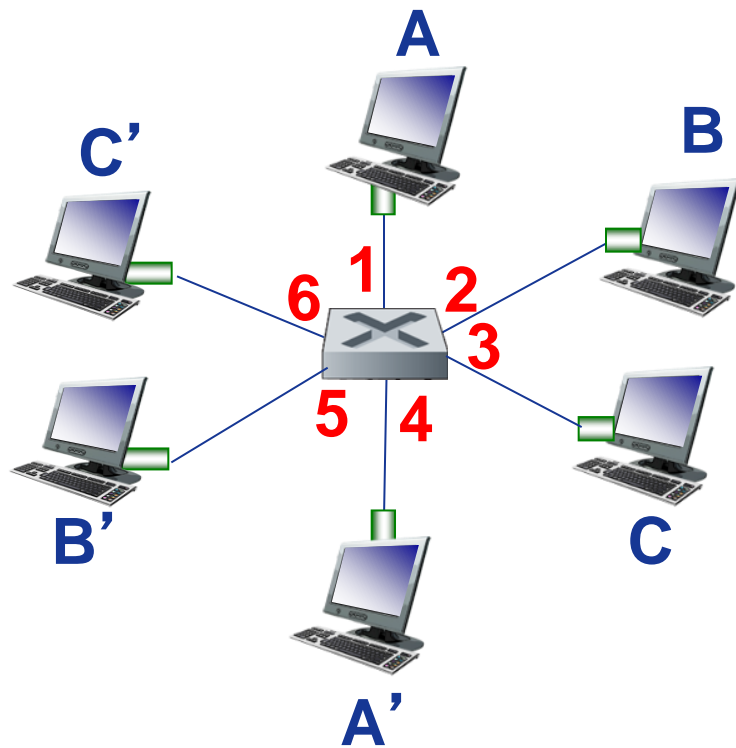
Q: 交换机怎么知道A' 可以通过接口4到达，而B' 可以通过接口5到达？

❖ **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)

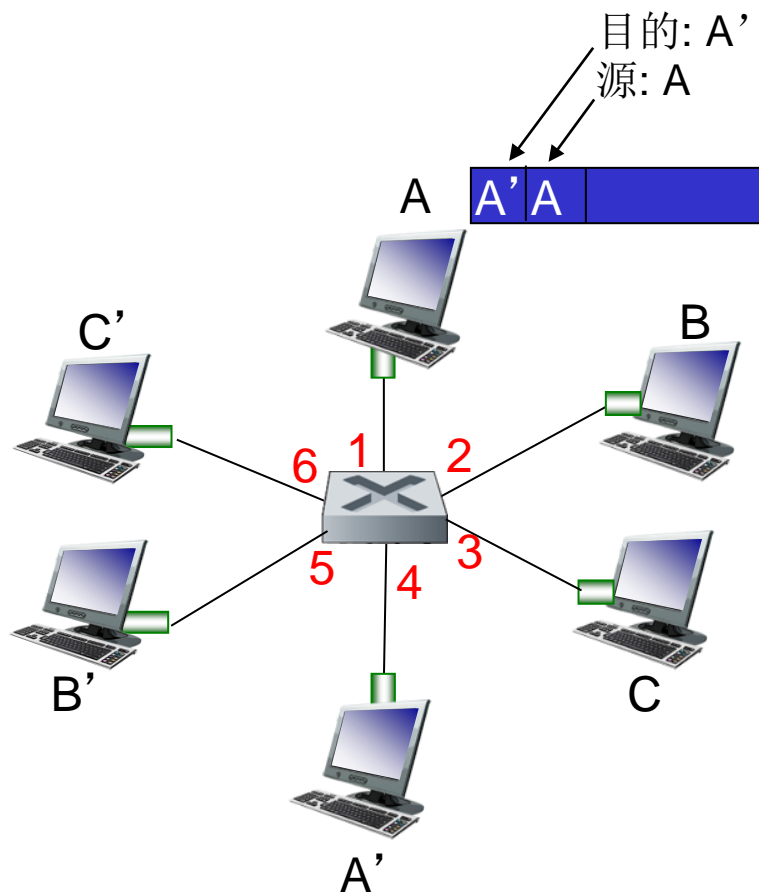
交换机：自学习

❖ 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 */

本讲主题

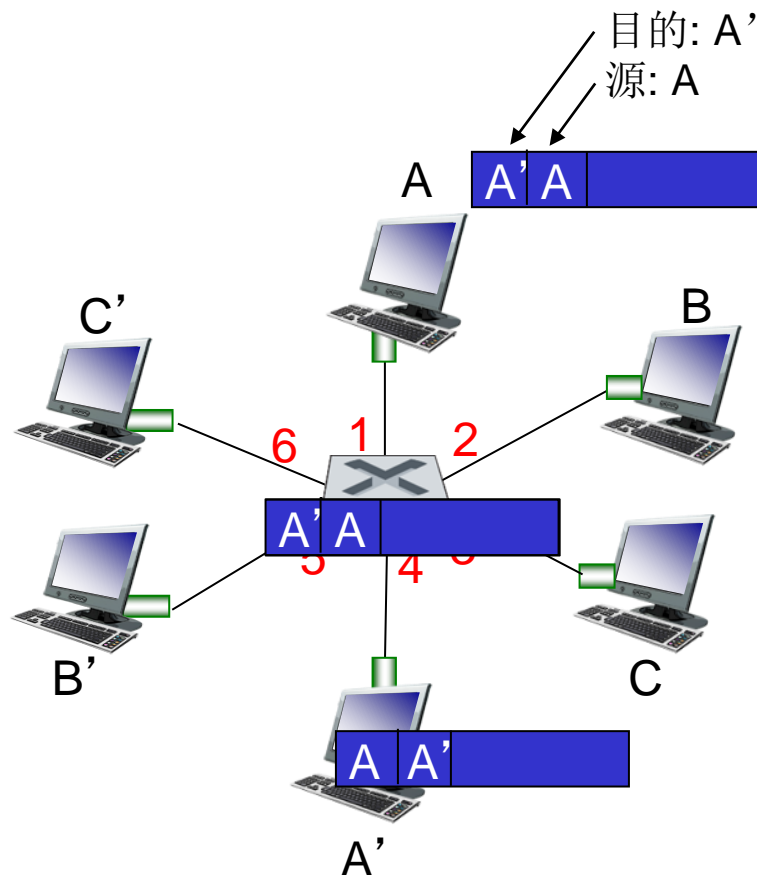
交换机（2）

自学习与转发过程举例

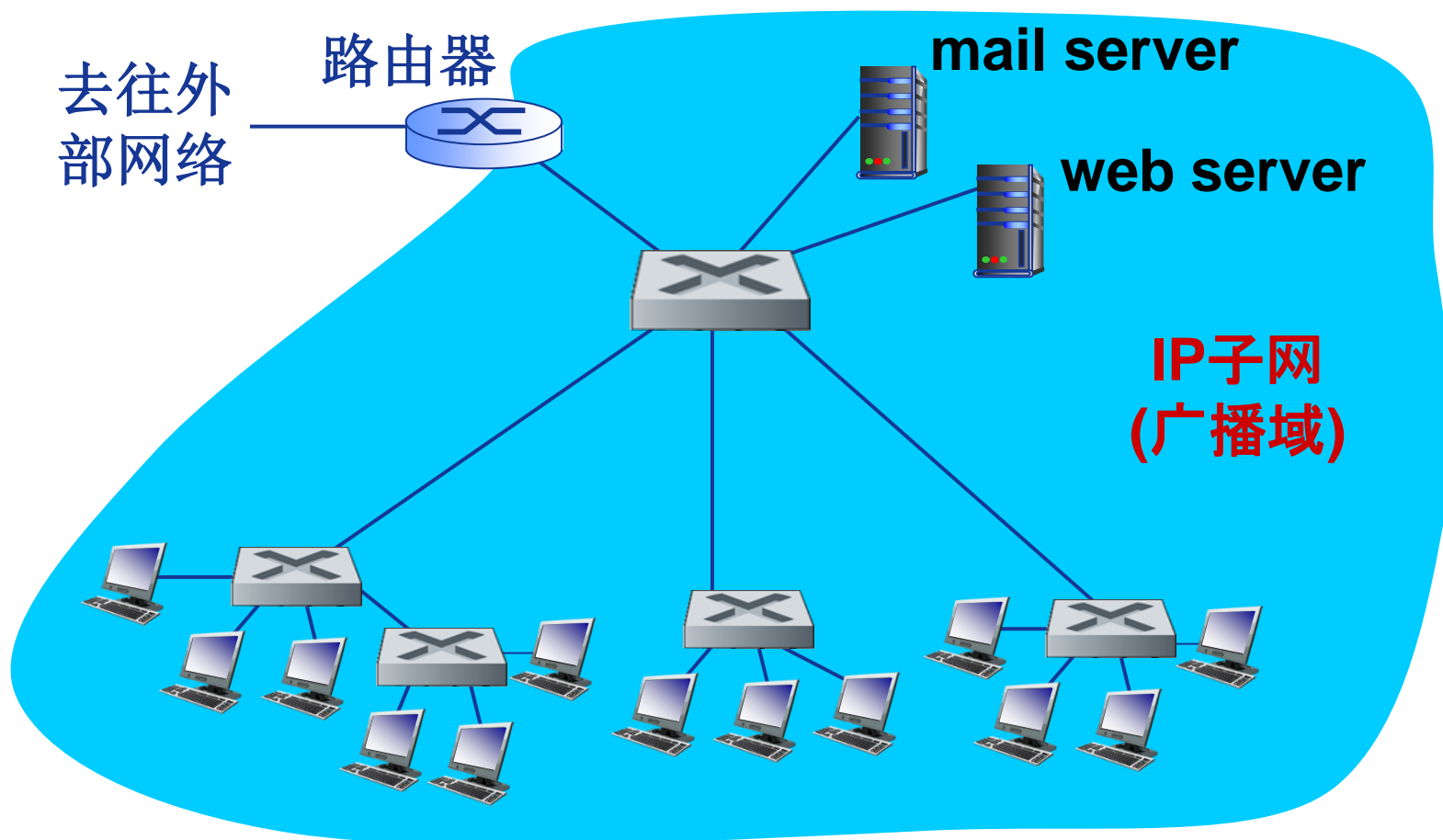
- ❖ 目的MAC地址A'，位置未知：
泛洪
- ❖ 目的MAC地址A，位置已知：
选择性转发

交换表
(初始为空)

MAC地址	接口	TTL
A	1	60
A'	4	60



组织机构(Institutional)网络



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

