



# 第六章: 链路层

## 目标:

- 理解下列数据链路层服务原理:
  - 错误检测, 纠正
  - 共享广播信道: 多路访问
  - 链路层编址
  - LAN局域网技术: 以太网, VLANs
- 各类链路层技术的实现和实例

1

## 6.1 引论和服务

2

## 6.2 差错检测和纠正技术

3

## 6.3 多路访问链路和协议

4

## 6.4 链路层

1.寻址和ARP    2.以太网    3.交换机    4.虚拟局域网

5

## 6.5 链路虚拟化: MPLS

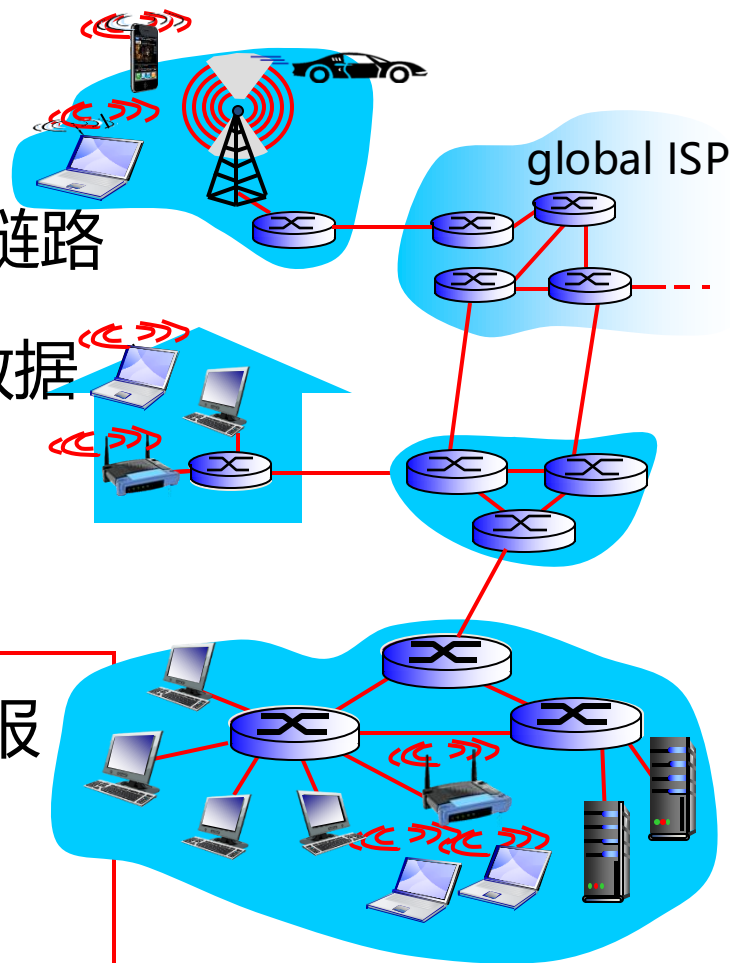
6

## 6.6 Web 请求生命周期中的一天

## 术语:

- 主机和路由器：结点
- 连接通信中相邻节点的通信信道：链路
- 第二层数据包：帧， 将网络层的数据包封装

数据链路层负责从一个节点传输数据报通过链路到物理上相邻的节点



- ❖ 数据报通过不同类型的链路层协议在不同的链路上传输：
  - 如以太网可能在第一段，帧中继在第二段，无线传输在最后的链路上
- ❖ 每一种链路层协议提供不同的服务
  - 如可能或不能在链路上提供可靠数据传输服务

## 交通类比:

- 旅行新校区 钟楼
  - 自行车：到新校区巴士站
  - 公交：到老校区巴士站
  - 出租车：到目的地
  - 你=数据报
- 传输段 = 通信链路
- 传输模式 = 链路层协议
- 行程代理 = 路由算法

## ■ 成帧，链路接入：

- 将数据报层成帧，添加帧头和帧尾
- 通过共享媒介实现多路访问机制
- 帧头包含标示数据来源和目的的MAC地址---不同于IP地址！

## ■ 相邻结点间可靠交付

- 传输层可靠交付已经学习，确认和重传机制也可用于链路层
- 但在低误码率的链路中（光纤，某些双绞线）很少使用
- 无线链路：误码率相当高

**Q：**为什么链路级和端到端传输中都要保证可靠性？

## ■ 流量控制:

- 保持发送结点和接收结点的同步

## ■ 差错检测:

- 差错可能由信号减弱, 噪声产生
- 接收结点检测差错时: 给发送端信号要求重发或丢弃出错帧

## ■ 差错纠正:

- 接收结点除了检测还要纠正比特位差错, 且无需比特重传

## ■ 半双工和全双工

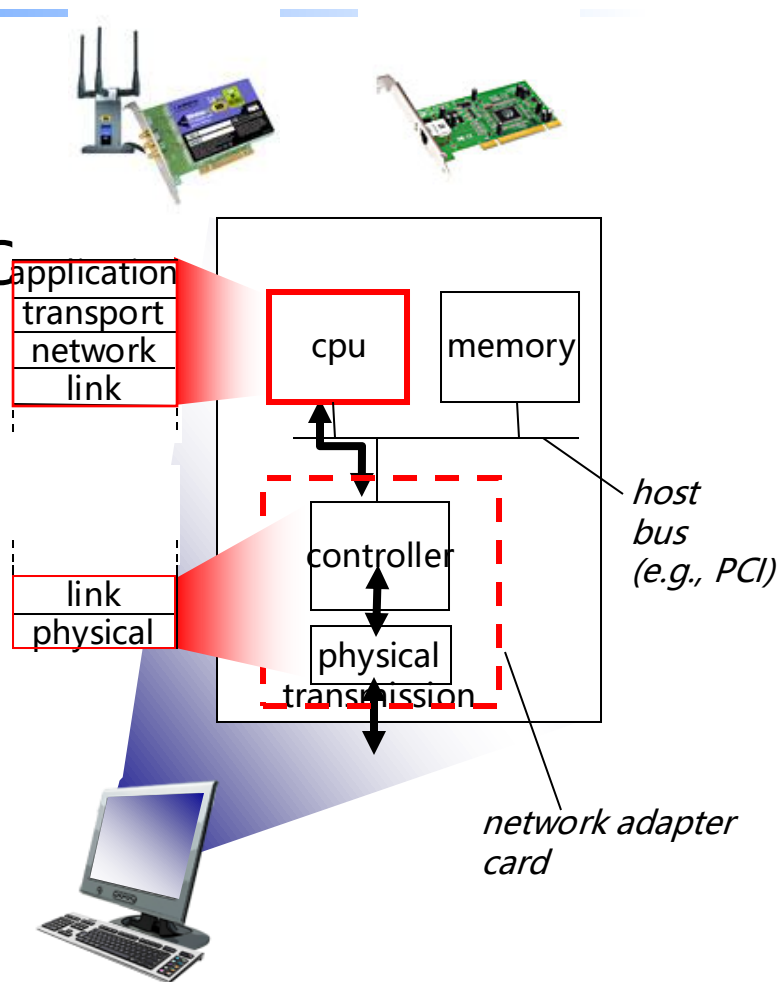
- 使用半双工, 链路两端的节点可以传输, 但不能同时传输

# 链路层在何处实现

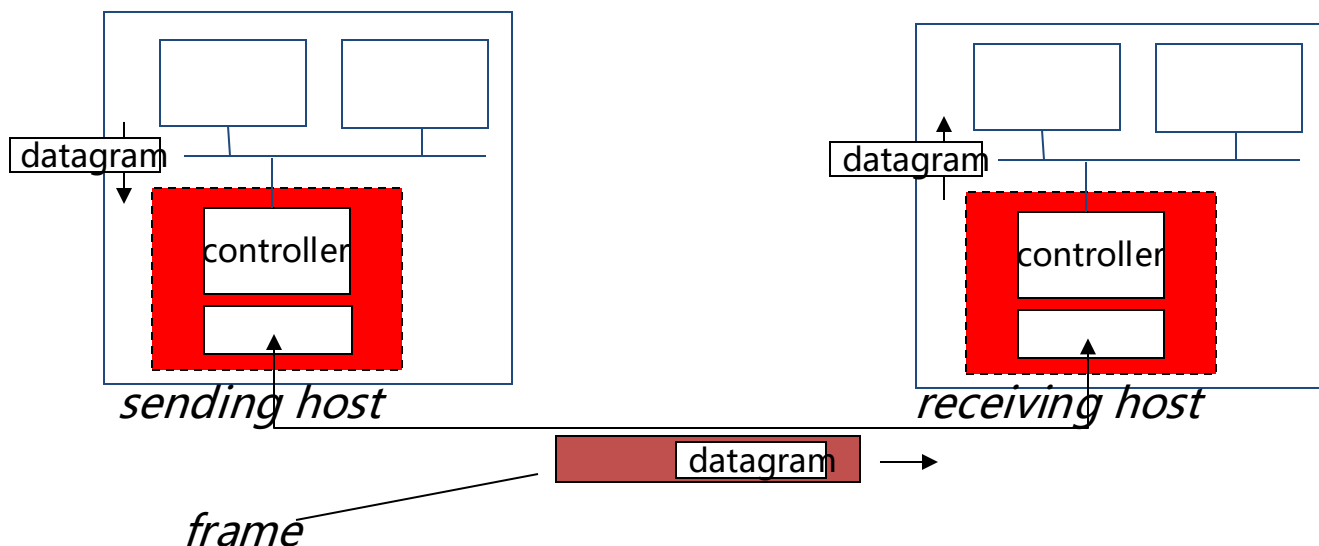
- 在所有主机
- 通过 “adapter (网卡或适配器)” 实现 (**Network interface card** , NIC)

- 以太网卡, 802.11无线网卡; 或者是集成至主板的以太网芯片
- 物理层上为链路接口

- 附属在主机的系统总线上
- 是硬件和固件的组合







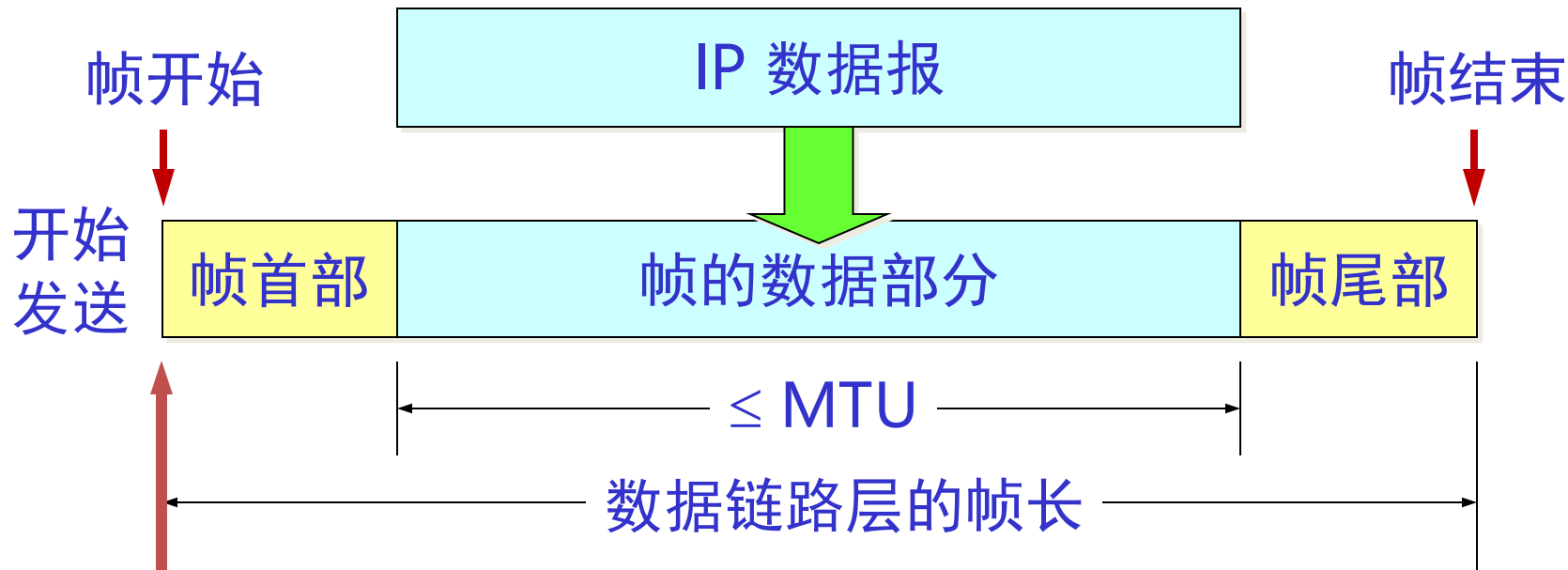
## ❖ 发送端

- 封装数据报至帧
- 增加差错检验位，可靠数据传输，流量控制

## ❖ 接收方

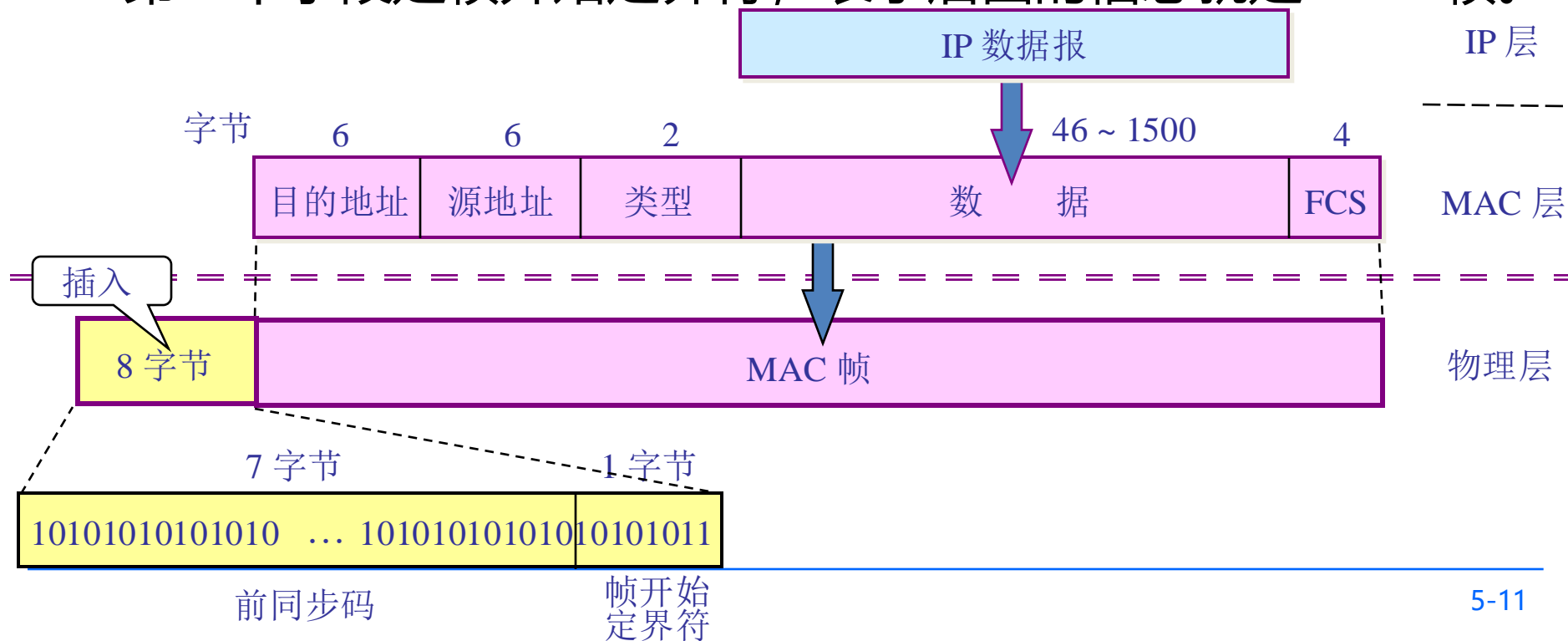
- 检测错误，可靠数据传输，流量控制
- 解封数据报，将其转至接收端高层处理

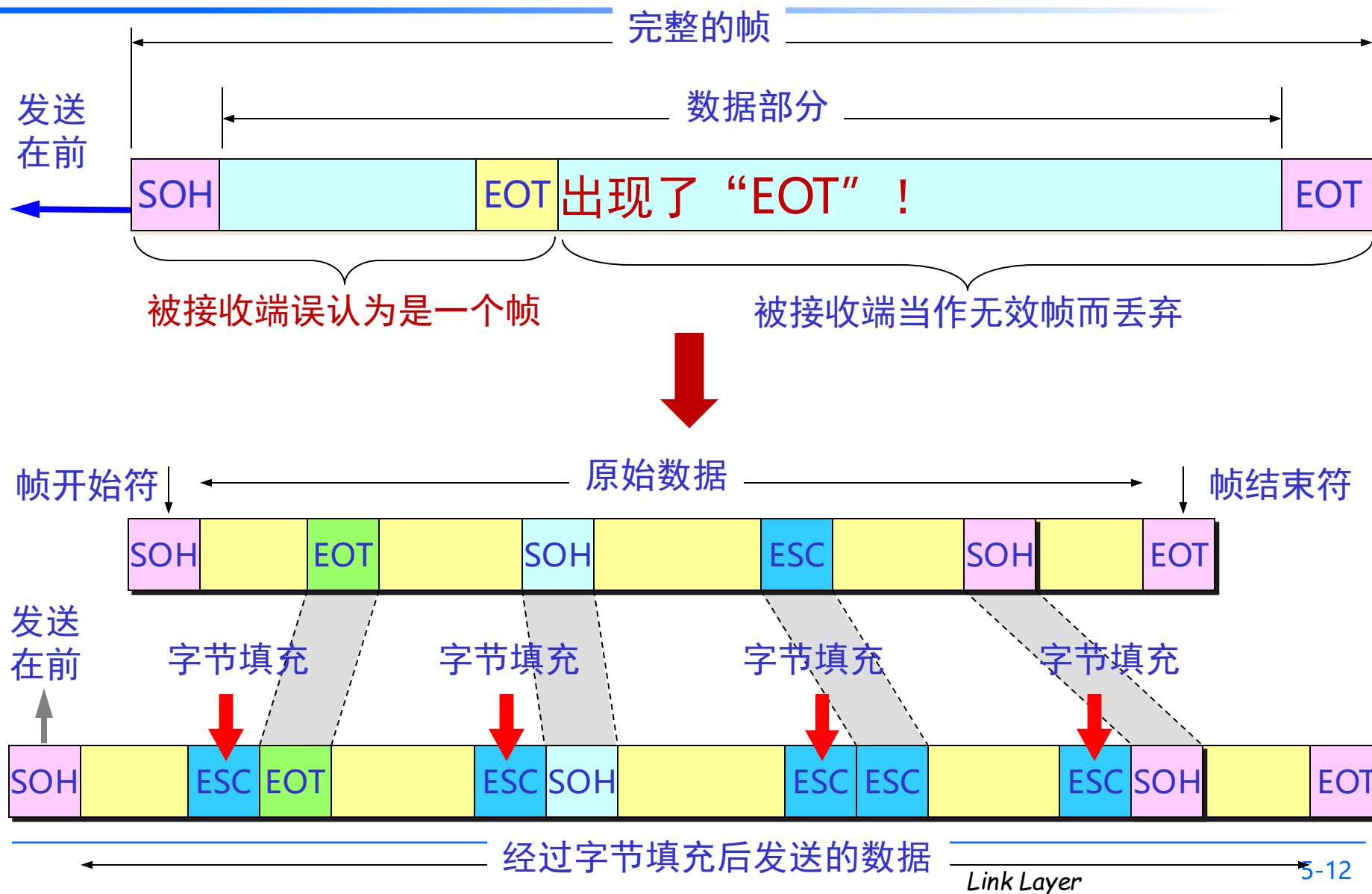
- 封装成帧(framing)就是在一段数据的前后分别添加首部和尾部，然后就构成了一个帧。确定帧的界限。
- 首部和尾部的一个重要作用就是进行**帧定界**。



■ 为了达到比特同步，在传输媒体上实际传送的要比 MAC 帧还多 8 个字节。

- 第一个字段共 7 个字节，是前同步码，用来迅速实现 MAC 帧的比特同步。
- 第二个字段是帧开始定界符，表示后面的信息就是 MAC 帧。





1

5.1引论和服务

2

5.2差错检测和纠正技术

3

5.3多路访问链路和协议

4

5.4链路层

1.寻址和ARP    2.以太网    3.交换机    4.虚拟局域网

5

5.5链路虚拟化: MPLS

6

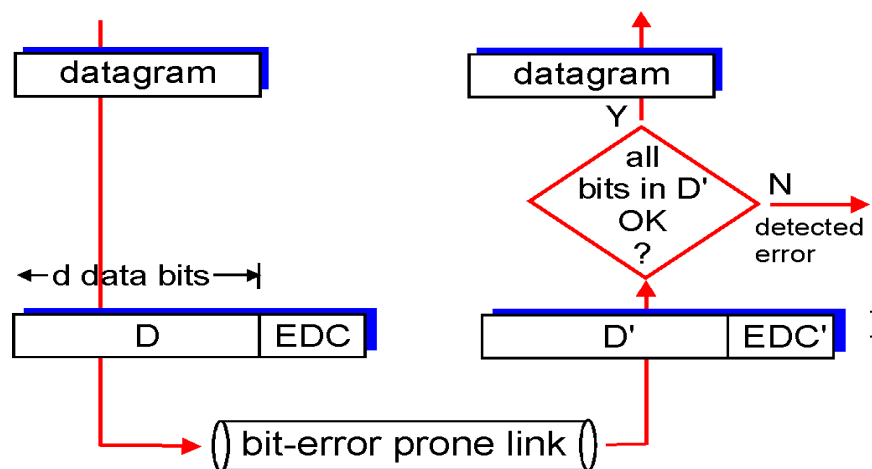
5.6 Web 请求生命周期中的一天

▶ 2003年

Schaerbeek

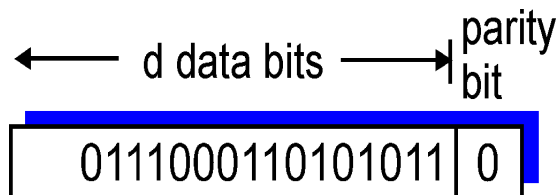
2003年比利时的斯哈尔贝克

- ❖ EDC = 错误检测校正 (Error Detection and Correction)
- ❖ D = 由检验位保护的数据, 包括首部字段
- 错误检测不可能达到 100% 可靠!
  - 协议算法可能会忽略了某些错误, 但比例极小
  - 较大的 EDC 字段可以产生较好的检错和纠错效果



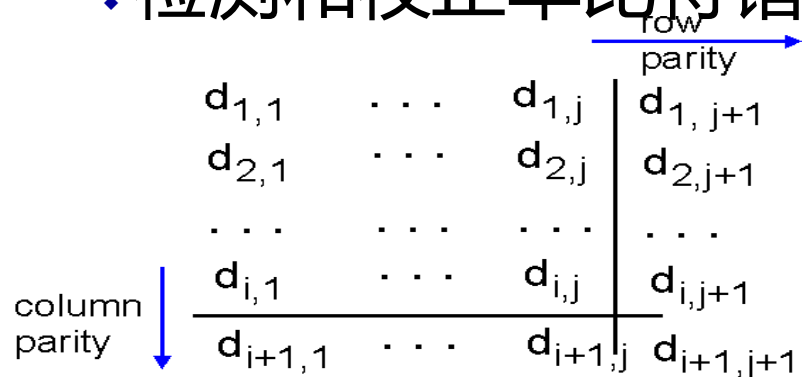
## 单比特校验::

### ❖ 检测一位错误



## 两维单比特校验:

### ❖ 检测和校正单比特错误



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

*no errors*

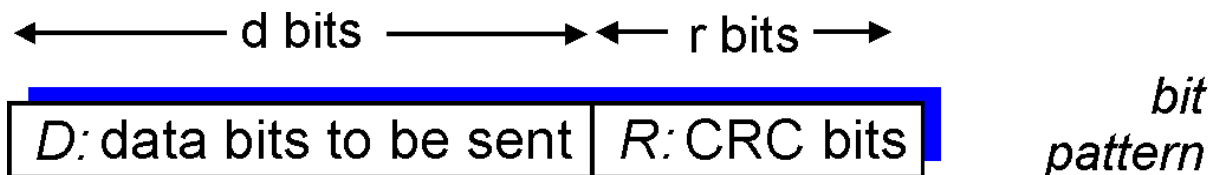
1	0	1	0	1	1
1	1	1	1	0	0
0	1	1	1	0	1
					0
0	0	1	0	1	0

parity error

*correctable  
single bit error*



- ❖ 更强的差错检测编码
- ❖ 把数据比特串,  $D$ , 看成二进制数据
- ❖ 选择一个  $r+1$  bit 型 (generator),  $G$
- ❖ 目的: 选择  $r$  个 CRC 位,  $R$ , 这样
  - #  $\langle D, R \rangle$  可为  $G$  整除(modulo 2)
  - #接收端已知  $G$ , 用 $G$ 来除  $\langle D, R \rangle$ . 如果余数不为0: 查出错误!
  - #可以查出所有长度不超过 $r+1$ 位的突发错误
- ❖ 在实践中被广泛应用 (ATM, HDCL)



$$D * 2^r \text{ XOR } R$$

mathematical  
formula

## CRC举例

如果要:

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

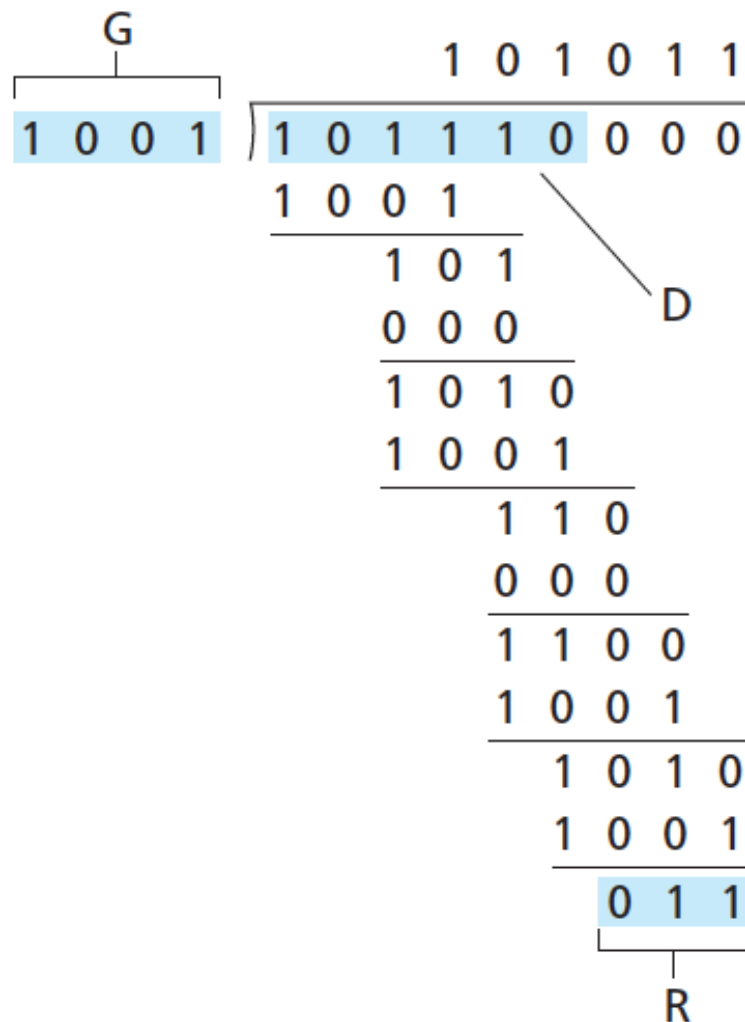
则等于:

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

也就等于:

如果要想  $D \cdot 2^r$  被  $G$  整除, 则  
需要余数

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



## 第六章: 链路层

1

5.1 引论和服务

2

5.2 差错检测和纠正技术

3

5.3 多路访问链路和协议

4

5.4 链路层

1. 寻址和ARP    2. 以太网    3. 交换机    4. 虚拟局域网

5

5.5 链路虚拟化: MPLS

6

5.6 Web 请求生命周期中的一天

有两种种类型的 “链路”：

点对点 ( e.g. PPP, HDLC)

- PPP,点对点协议
- HDLC,高级数据链路控制协议

广播式 (共享线路或介质; e.g, 共享式以太网, 无线网, etc.)即多路访问问题)

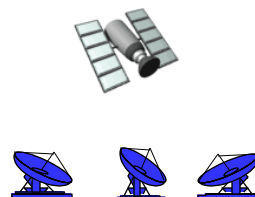
- 传统以太网
- 上行HFC
- 802.11 无线局域网



共享线路 (例如,  
有线以太网)



共享射频  
(e.g., 802.11 WiFi)



共享射频  
(卫星)



你和我在教室里  
(共享空气, 声学)

- 一条共享的通信信道
- 两个或多个结点可同时发送信号: 相互干扰
  - 在某一时刻只有一个结点可以成功地发送信号

## 介质访问控制协议 (MAC, Medium access control)

- 分布式的算法来决定如何共享信道, i.e., 决定结点何时可以发送
- 注意: 有关共享通道的通信 (协商) 也必须在该通道自身上解决!

假设：  $R$  bps带宽的广播信道

期望：

1. 当某一结点有数据发送时，该结点具有  $R$  bps的吞吐量；
2. 当有  $M$  个结点要发送数据时，每个结点吞吐量为  $R/M$  bps；
3. 协议是非集中式的：不会因某主结点故障而使整个系统崩溃

三大类:

## ■ 通道分割

- 将信道分割成较小的 “片” (时隙, 频率)
- 将小片分给各站点使用

## ■ 随机访问

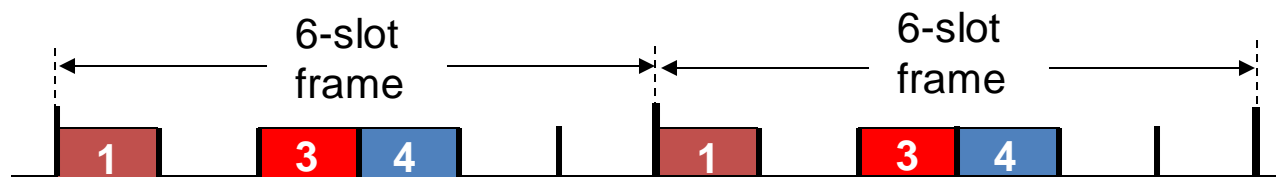
- 允许冲突
- 从冲突中 “恢复”

## ■ “排队”

- 严格协调访问来避免冲突

## TDMA: time division multiple access (时分多址)

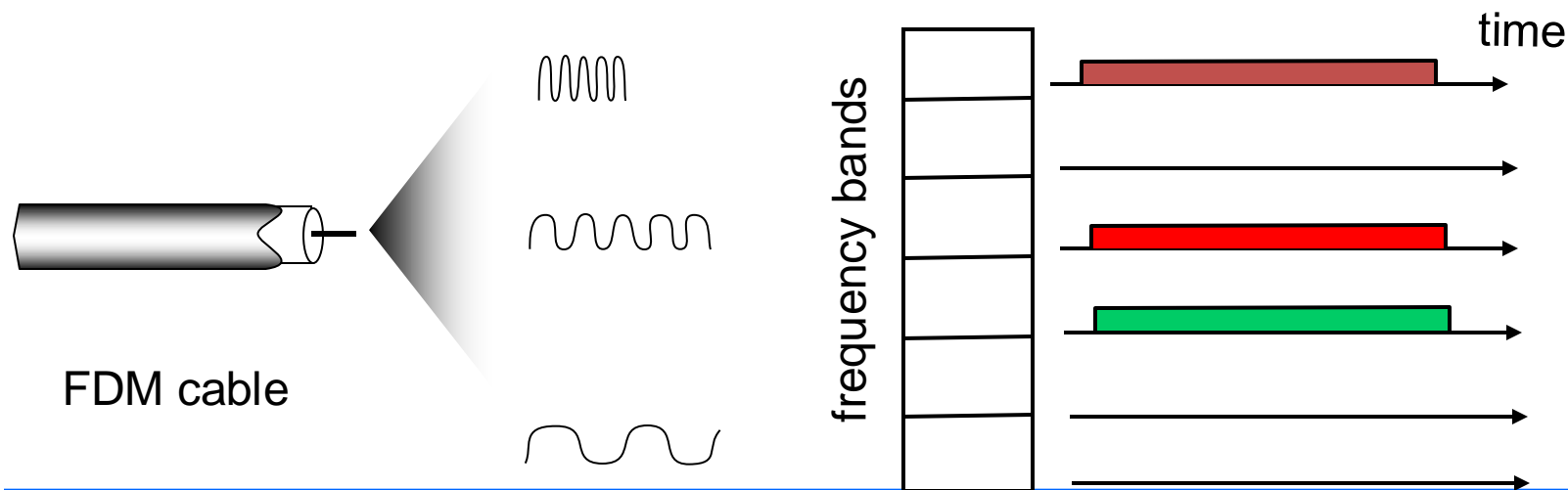
- “依次” 访问信道
- 每次每个站点分得固定长度的时隙  
(时长 = 分组的单位传输时间)
- 未用的时隙被闲置和浪费
- 例如: 6个站点的LAN, 1,3,4 有分组发送, 而 2,5,6的时隙则被闲置





## FDMA: frequency division multiple access (频分多址)

- 信道按频谱分成若干频段
- 每个站点分得固定的频段
- 在频段不用时该部分信道被闲置和浪费
- 例如: 6站点的 LAN, 1,3,4 发送分组, 而 2,5,6 的频段被闲置



# 信道划分协议

软件学院·计算机网络

电单 baba

▪

- 常用的名词是码分多址 CDMA (Code Division Multiple Access)。
- 各用户使用经过特殊挑选的不同码型，因此彼此不会造成干扰。
- 这种系统发送的信号有很强的抗干扰能力，其频谱类似于白噪声，不易被敌人发现。

- 每一个比特时间划分为  $m$  个短的间隔，称为码片(chip)。
- 每个站被指派一个唯一的  $m$  bit 码片序列。
  - 如发送比特 1，则发送自己的  $m$  bit 码片序列。
  - 如发送比特 0，则发送该码片序列的二进制反码。
- 例如，S 站的 8 bit 码片序列是 00011011。
  - 发送比特 1 时，就发送序列 00011011，
  - 发送比特 0 时，就发送序列 11100100。
- S 站的码片序列：(-1 -1 -1 +1 +1 -1 +1 +1)

- 假定S站要发送信息的数据率为  $b$  bit/s。由于每一个比特要转换成  $m$  个比特的码片，因此 S 站实际上发送的数据率提高到  $mb$  bit/s，同时 S 站所占用的频带宽度也提高到原来数值的  $m$  倍。
- 这种通信方式是扩频(spread spectrum)通信中的一种。
- 扩频通信通常有两大类：
  - 一种是直接序列扩频DSSS (Direct Sequence Spread Spectrum)，如上面讲的使用码片序列就是这一类。
  - 另一种是跳频扩频FHSS (Frequency Hopping Spread Spectrum)。

- 每个站分配的码片序列不仅**必须各不相同**，并且还**必须互相正交** (orthogonal)。
- 在实用的系统中是使用**伪随机码序列**。

- 令向量  $S$  表示站  $S$  的码片向量，令  $T$  表示其他任何站的码片向量。
- 两个不同站的码片序列正交，就是向量  $S$  和  $T$  的规格化内积 (inner product) 等于 0:

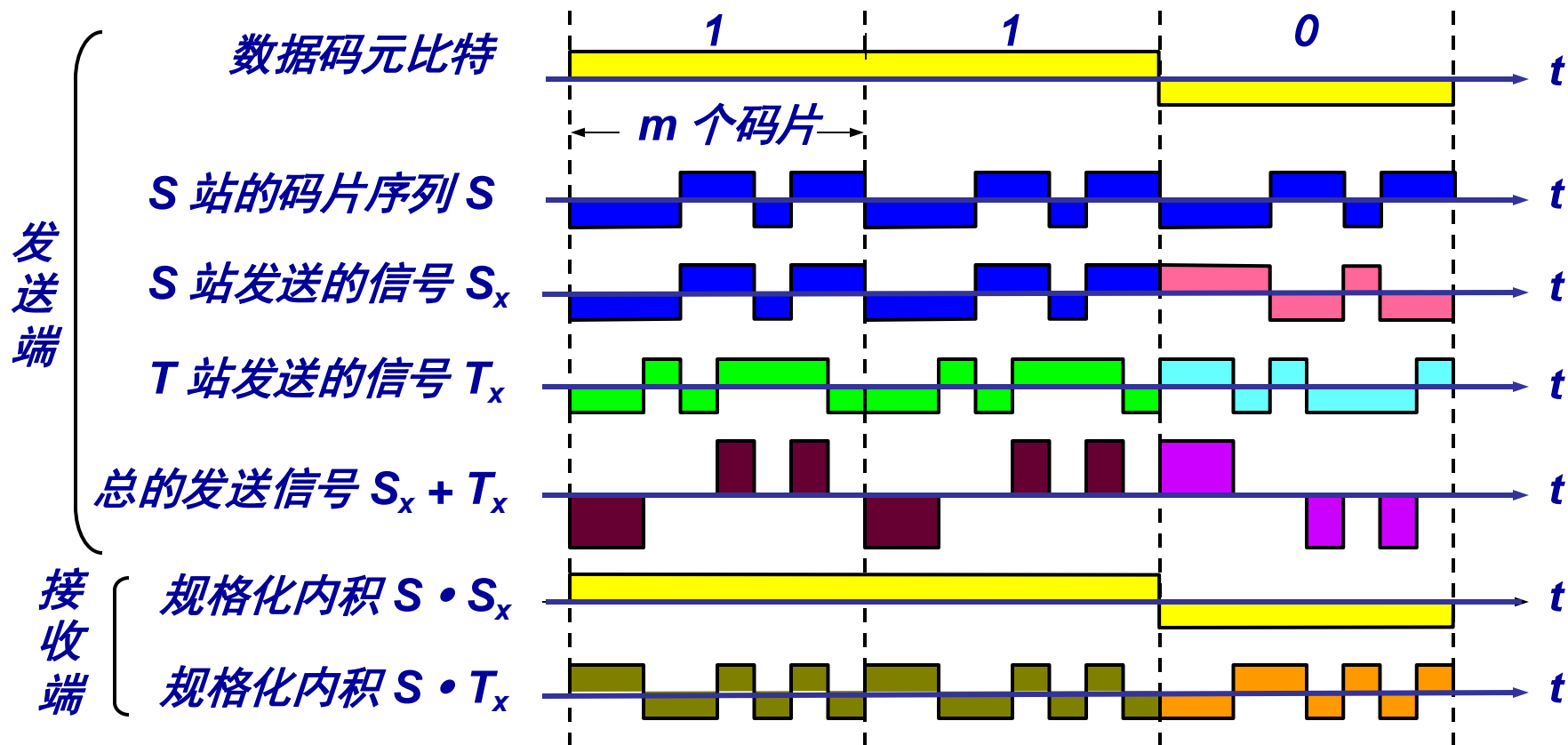
$$\mathbf{S} \bullet \mathbf{T} \equiv \frac{1}{m} \sum_{i=1}^m S_i T_i = 0$$

- 任何一个码片向量和该码片向量自己的规格化内积都是 1。

$$\mathbf{S} \bullet \mathbf{S} = \frac{1}{m} \sum_{i=1}^m S_i S_i = \frac{1}{m} \sum_{i=1}^m S_i^2 = \frac{1}{m} \sum_{i=1}^m (\pm 1)^2 = 1$$

- 一个码片向量和该码片反码的向量的规格化内积值是 -1。





# 信道划分协议

软件学院 · 计算机网络



[Learning Resources](#) [Blog](#) [Sponsor and Donate](#)

[At-home STEM Kits](#)

[Register for Cam](#)

[Programs](#) ▾

[Educators](#) ▾

[Museum](#) ▾

[Inductees](#) ▾

[Events](#) ▾

[Competitions](#) ▾

[About](#)



## Irwin Mark Jacobs

CDMA Technology

US Patent No. 4,901,307

Inducted in 2013

Born October 18, 1933

Qualcomm's co-founders Irwin Jacobs, Klein Gilhousen, and Andrew Viterbi made major contributions to code division multiple access (CDMA), the technology underlying all third generation cellular telephone networks. CDMA now supports over 1.6 billion subscribers in developing and developed countries with voice and high speed Internet access.

The first CDMA communications system was installed and commercialized in Hong Kong in 1995, followed closely by South Korea and then the United States.

Jacobs and Viterbi are also two of the three co-founders of Linkabit, initially providing consulting but then developing and manufacturing a number of innovative communication products. After Linkabit was sold, Jacobs, Viterbi, and five others founded Qualcomm in July 1985. Their first commercial product was OmniTRACS, a satellite communications and tracking system, still used by long haul trucking companies.



## UNITED STATES PATENT OFFICE

2,292,387

### SECRET COMMUNICATION SYSTEM

Hedy Kiesler Markey, Los Angeles, and George Antheil, Manhattan Beach, Calif.

Application June 10, 1941, Serial No. 397,412

6 Claims. (Cl. 250-2)

This invention relates broadly to secret communication systems involving the use of carrier waves of different frequencies, and is especially useful in the remote control of dirigible craft, such as torpedoes.

An object of the invention is to provide a method of secret communication which is relatively simple and reliable in operation, but at the same time is difficult to discover or decipher.

Briefly, our system as adapted for radio control of a remote craft, employs a pair of synchronous records, one at the transmitting station and one at the receiving station, which change the tuning of the transmitting and receiving apparatus from time to time, so that without knowledge of the records an enemy would be unable to determine at what frequency a controlling impulse would be sent. Furthermore, we contemplate employing records of the type used for many years in player pianos, and which consist of long rolls of paper having perforations variously positioned in a plurality of longitudinal rows along the records. In a conventional player piano record there may be 88 rows of perforations, and in our system such a record would permit the use of 88 different carrier frequencies, from one to another of which both the transmitting and receiving station would be changed at intervals. Furthermore, records of the type described can be made of substantial length and may be driven slow or fast. This makes it possible for a pair of records, one at the transmitting station and one at the receiving station, to run for a length of time ample for the remote control of a device such as a torpedo.

Fig. 2 is a schematic diagram of the apparatus at a receiving station;

Fig. 3 is a schematic diagram illustrating a starting circuit for starting the motors at the transmitting and receiving stations simultaneously;

Fig. 4 is a plan view of a section of a record strip that may be employed;

Fig. 5 is a detail cross section through a record-responsive switching mechanism employed in the invention;

Fig. 6 is a sectional view at right angles to the view of Fig. 5 and taken substantially in the plane VI-VI of Fig. 5, but showing the record strip in a different longitudinal position; and

Fig. 7 is a diagram in plan illustrating how the course of a torpedo may be changed in accordance with the invention.

Referring first to Fig. 7, there is disclosed a mother ship 10 which at the beginning of operations occupies the position 10a and at the end of the operations occupies the position 10b. This mother ship discharges a torpedo 11 that travels successively along different paths 12, 13, 14, 15 and 16 to strike an enemy ship 17, which initially occupies the position 17a but which has moved into the position 17b at the time it is struck by the torpedo 11. According to its original course, the enemy ship 17 would have reached the position 17c, but it changed its course following the firing of the torpedo, in an attempt to evade the torpedo.

In accordance with the present invention, the torpedo 11 can be steered from the mother ship 10a and its course changed from time to time as necessary to cause it to strike its target. In

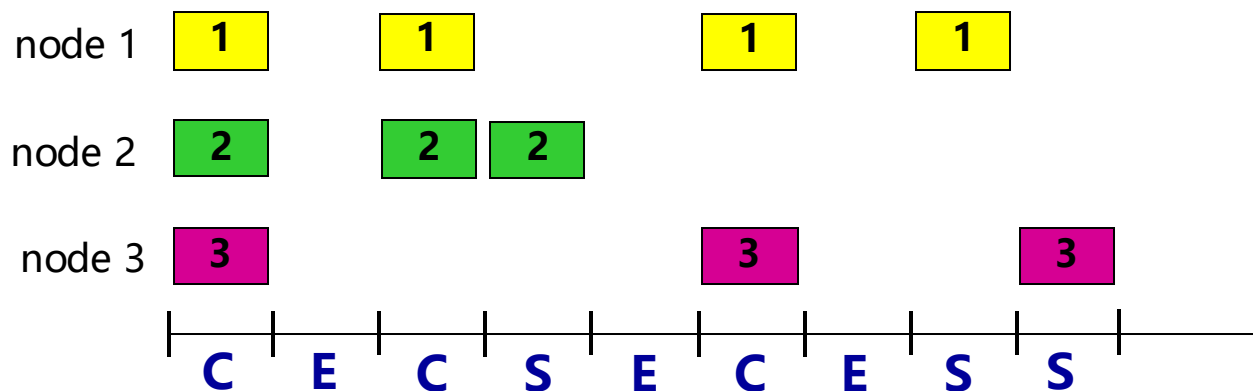
- 当结点有数据要发送时
  - 使用信道全部的传输速率  $R$ .
  - #在诸多结点中不存在“预先”协商的机制
- 可能发生两个以上结点同时传输 -> “冲突”
- 随机访问的 MAC协议定义了:
  - 如何检测冲突
  - #如何从冲突中恢复 (e.g., 通过延迟重发)
- 随机访问 MAC协议的实例:
  - 时隙ALOHA
  - ALOHA
  - CSMA and CSMA/CD

## 设定:

- ❖ 所有帧大小一致
- ❖ 时间划分成等长的时隙 (=分组单元传送的时长)
- ❖ 结点只在时隙起点开始传输帧
- ❖ 结点同步, 每个结点都知道时隙什么时候开始
- ❖ 同一个时隙中有两个或者更多帧碰撞, 则所有结点在该时隙结束之前检测到该碰撞事件。

## 操作:

- ❖ 当结点发送新在帧, 要等到下一个时隙开始并在该时隙传输整个帧
- ❖ 如果无碰撞, 成功传输该帧, 不需要考虑重传该帧 (同时准备新 帧)
- ❖ 如果有碰撞, 结点在时隙结束之前检测到这次碰撞。该结点以概率 $p$ 在后续的时隙中重传该帧, 直到无
- ❖ 碰撞成功传输



## 优点:

- 单个节点可以信道最高速率连续传输数据
- 高度非集中化：只有结点的时隙需要同步
- 简单

## 缺点:

- ❖ 碰撞发生，时隙被浪费
- ❖ 有空时隙
- ❖ 结点检测碰撞需要时间
- ❖ 需要同步时钟

**效率：**成功插槽的长期比例（许多节点，所有节点都有许多帧要发送）

- 假设有 N 站点有数据发送
- 每个站点以 p 的概率在时隙中传输数据传输成功的概率  $S := p(1-p)^{N-1}$

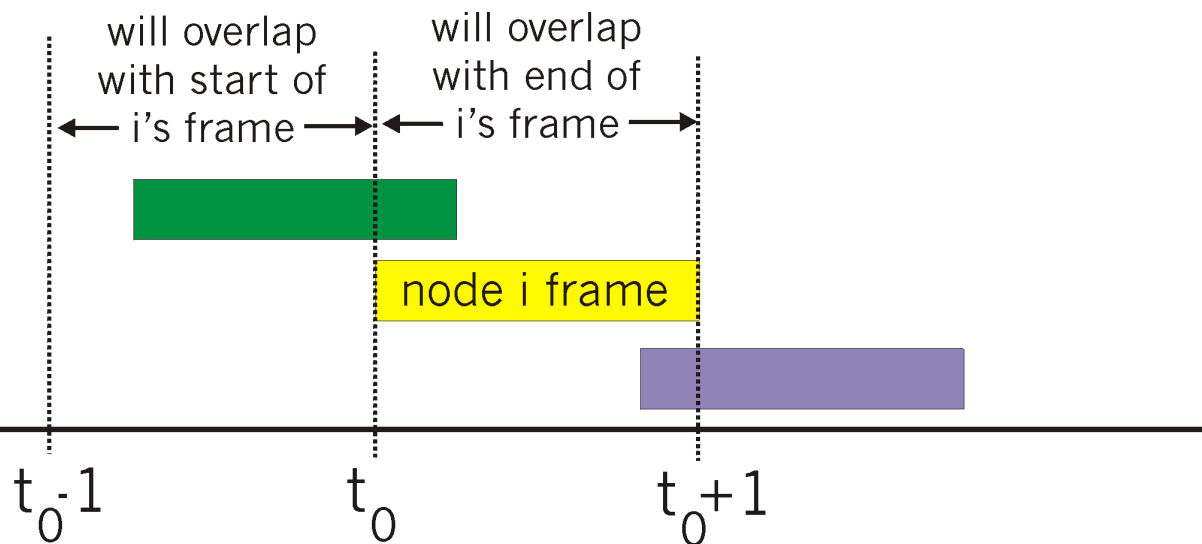
- 对 N 个节点中的任何一个：  $S = \text{概率}(\text{仅有一个站点传输的}) = N p (1-p)^{(N-1)}$
- ... 选择最佳的 p 当  $n \rightarrow \infty$  ...
- $S = 1/e = 0.37$

**充其量：**通道用于有用的传输 37% 时间!



# 无时隙ALOHA

- ❖ 无时隙 Aloha: 简单, 没有同步信号
- ❖ 需要传输分组时:
  - 立即发送而不必等待时隙的开始
- ❖ 冲突的概率增加了:
  - 在  $t_0$  时刻发送的分组与在  $[t_0-1, t_0+1]$  时刻发送的分组冲突





$P(\text{给定节点的成功率}) = P(\text{节点传输}) .$

$P(\text{无其他节点在 } [p_0-1, p_0] \text{ 时刻传输}) .$

$$P(\text{无其他节点在 } [p_0-1, p_0] \text{ 时刻传输}) \\ = p . (1-p) . (1-p)$$

$\rightarrow \infty$

$P(\text{在 } N \text{ 结点中任意结点的成功率}) = N p . (1-p) . (1-p)$

... 选择最佳的  $p$  当  $n \rightarrow \infty$  ...

$$= 1/(2e) = .18$$

比时隙ALOHA的效果还差，协议 限制了信道吞吐量的有效利用

**CSMA:** (载波侦听) 发送前侦听:

- 如果信道闲置: 发送整个分组
- 如果信道忙, 推迟发送

持续性 CSMA: 当信道闲置时, 以 $p$ 的概率立即重试  
(可能导致不稳定)

非持续性 CSMA: 在某个随机间隔以后再试

- 为人处事的规则之一: 不要打断别人的发言!

节点的空间布局



$t_c$   
time  
↓

$t_1$

- **冲突可能发生在:**由于传播延迟两个节点可能听不到对方的发送
- **冲突:**整个分组的传输时间被浪费
  - 这里的冲突概率是由距离和 传播延迟来决定的

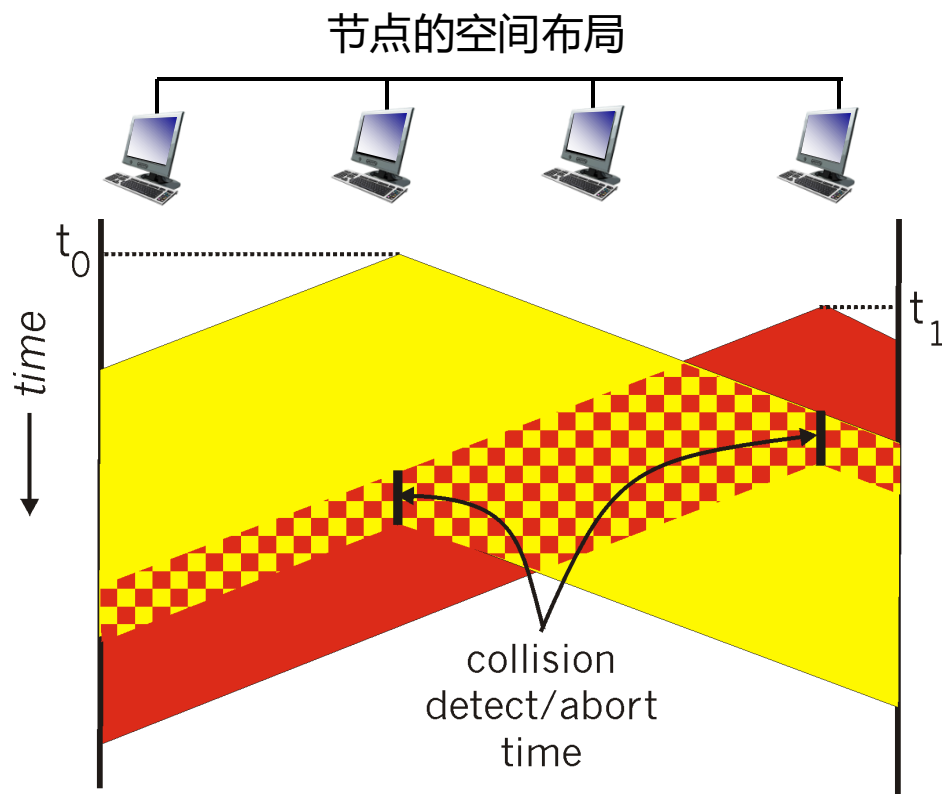
## CSMA/CD:

- 在冲突发生后, 短时间内可探测到
- 立即中断传输, 减少信道的时间浪费
- 坚持性或非坚持性重传

### ❖ 冲突检测:

- 在有线 LAN 中简便易行: 检测信号强度, 比较收、发的信号
- 在无线 LAN 比较困难: 接收的信号强度由发送强度决定

# CSMA/CD 带碰撞检测 的载波侦听多路访问



1. 网卡从网络层获得数据报，生成链路层帧
2. 如果网卡侦听到信道空闲，开始传输帧。如果网卡侦听到信道正忙，则等待，直到侦听到没有信号能量才开始自己的传输
3. 如果网卡传输整个帧没有检测到其他的传输（检测信号能量），网卡就完成了该帧的传输。
4. 如果网卡传输帧过程中检测到其他的传输（检测信号能量），网卡会中止当前帧传输
5. 中止传输后，网卡等待一个随机时间量，然后返回 2
  - 第  $m$  次碰撞后，NIC 从  $\{0, 1, 2, \dots, 2^m - 1\}$  中随机选择  $K$ 。网卡等待  $K \cdot 512$  比特时间，返回步骤 2
  - 退避间隔更长，冲突更多

- $t_{prop}$  = 局域网中 2 个节点之间的最大传播延迟

- $t_{trans}$  = 传输最大尺寸帧的时间

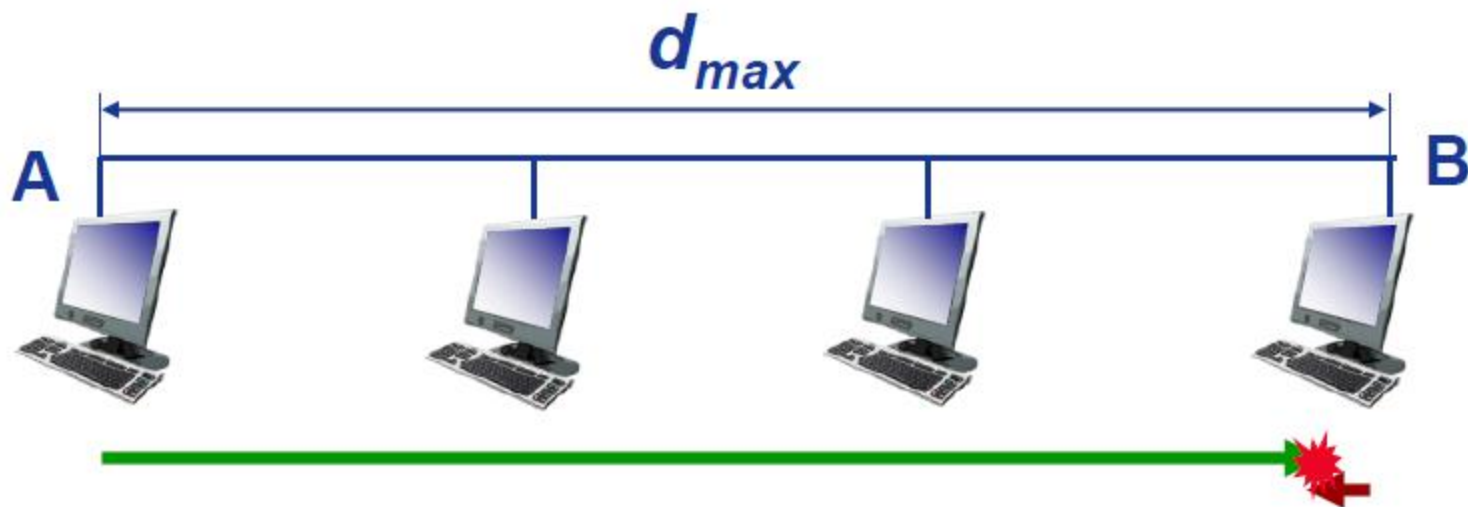
$$efficiency = \frac{1}{1 + 5t_{prop}/t_{trans}}$$

- 效率达到 1

- 当  $t_{prop}$  到达 0

- 当  $t_{trans}$  到达无限

- 比 ALOHA 更好的性能：简单、便宜、去中心化!



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)

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

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

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



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

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

答案：D

## 信道分割的 MAC 协议::

- 在重负荷时共享信道的效率高
- 在轻负荷时效率低: 延迟了信道访问, 当只有一个结点激活时, 分到的带宽只有信道带宽的 $1/N$ !

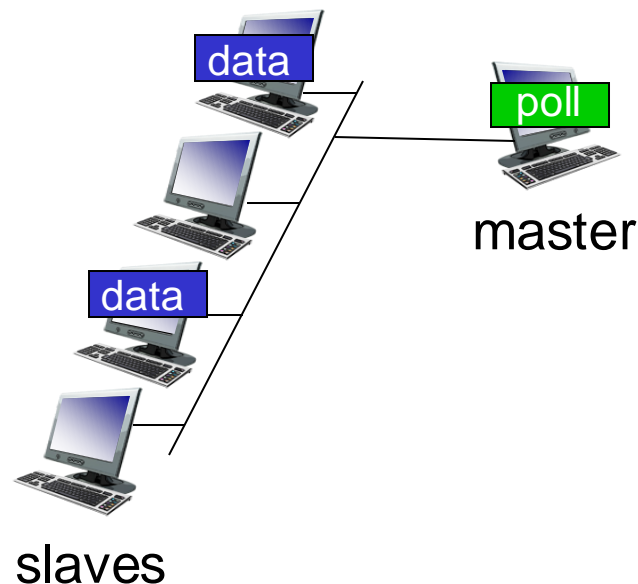
## 随机访问 MAC 协议

- 在轻负荷时有效率: 单个节点可以充分利用信道的 带宽
- #在重负荷下: 冲突的开销

“轮转式” 协议-----在两个方面寻求最佳的解决方案!

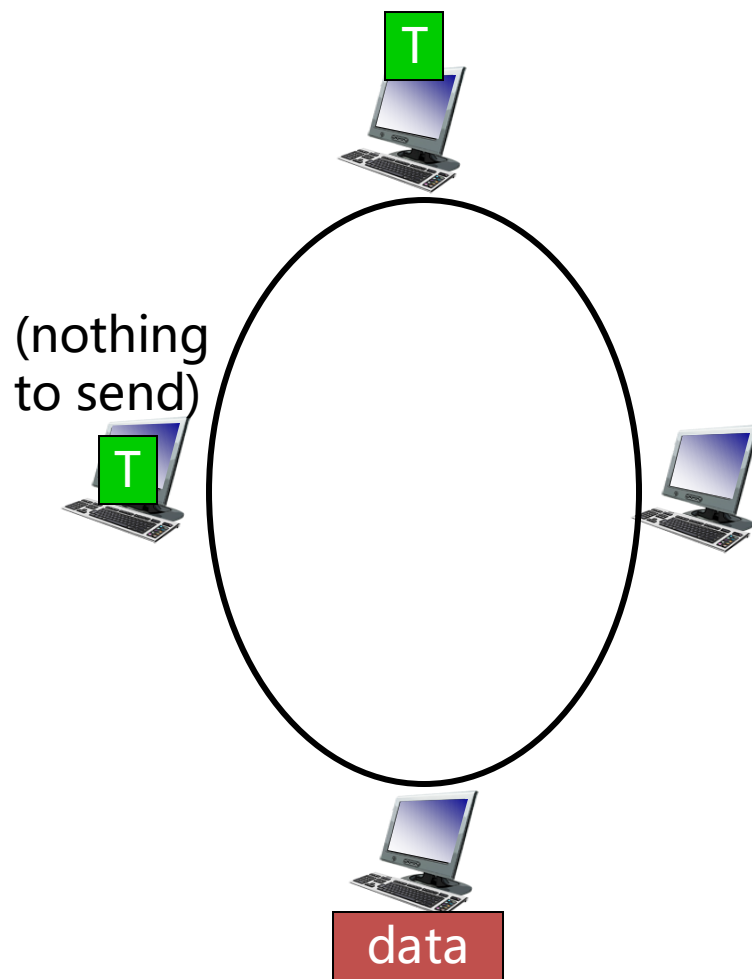
## 轮询（Polling）：

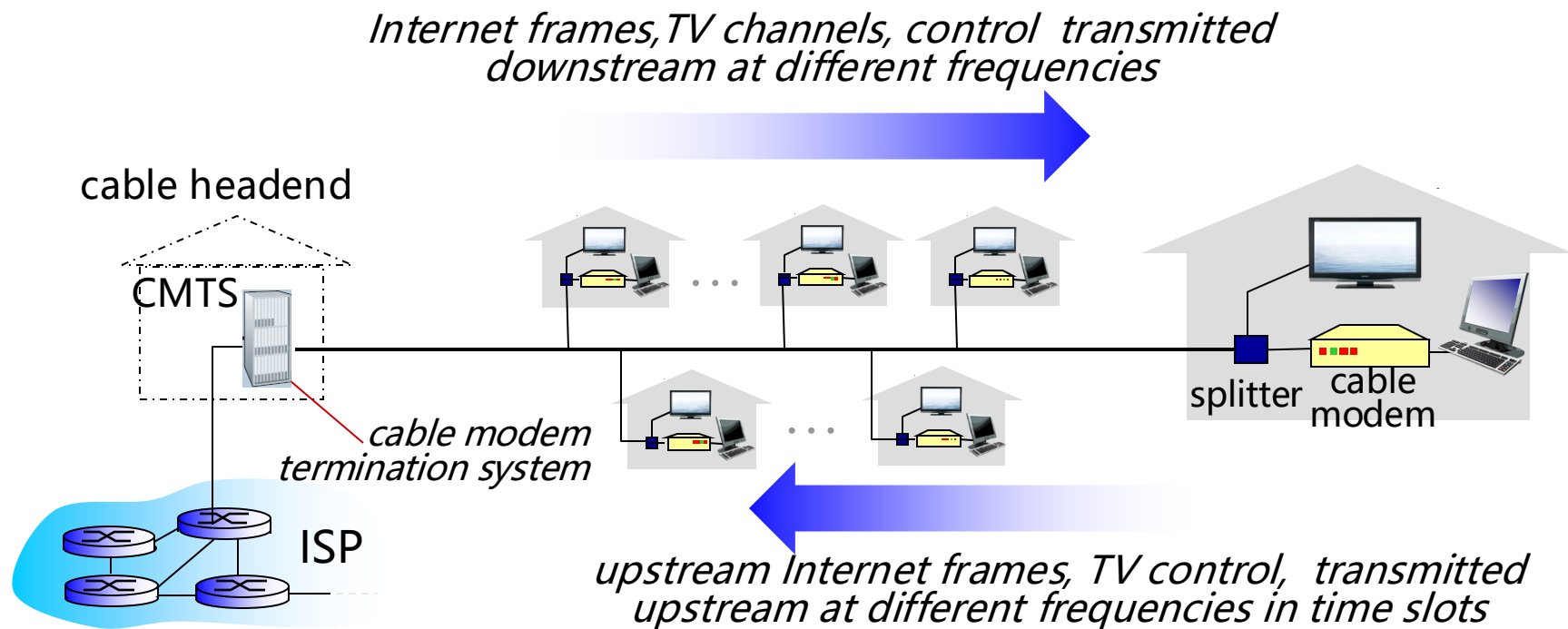
- 由主结点“邀请”从结点依次传输
- Request to Send, Clear to Send 报文
- 关注：
  - 轮询的开销
  - 等待时间
  - 单点失效（主结点）



## 令牌传递（Token passing 或 Token Ring）：

- ❖ 控制令牌依次通过各个结点
- ❖ 令牌报文
- ❖ 关注：
  - 令牌的开销
  - #等待时间
  - #单点失效(token)



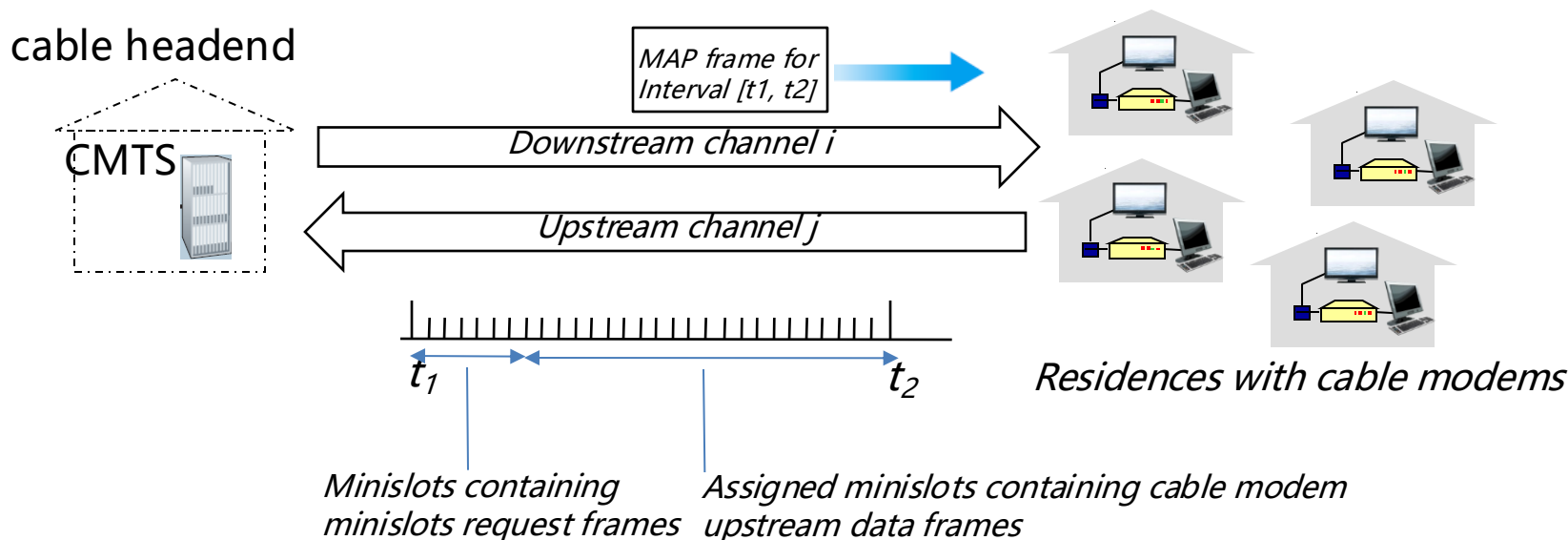


❖ 多个40Mbps下行(广播)频道

- 单个CMTS传输到通道中

❖ 多个30 Mbps上行通道

- 多址接入:所有用户竞争某些上行信道时隙(其他分配的)



## DOCSIS: 电缆数据服务接口规范

- ❖ 上行、下行频道上的FDM
- ❖ TDM上行: 一些时隙被分配, 一些有竞争
  - 下行映射帧: 分配上行时隙
  - 在所选时隙中请求上行时隙(和数据)传输随机接入(二进制退避)

- ❖ 信道分割, 按时间, 频率或编码
  - 时分、码分、频分
- ❖ 随机分割 (动态)
  - ALOHA, S-ALOHA, CSMA, CSMA/CD
  - #载波检测: 有线 “易行”、无线 “困难”
  - #CSMA/CD 被用在以太网中
- ❖ 轮转分割
  - 从主结点发出轮询, 令牌传递 (token passing)
  - # 蓝牙, FDDI, token ring

1

5.1引论和服务

2

5.2差错检测和纠正技术

3

5.3多路访问链路和协议

4

5.4链路层

1.寻址和ARP    2.以太网    3.交换机    4.虚拟局域网

5

5.5链路虚拟化: MPLS

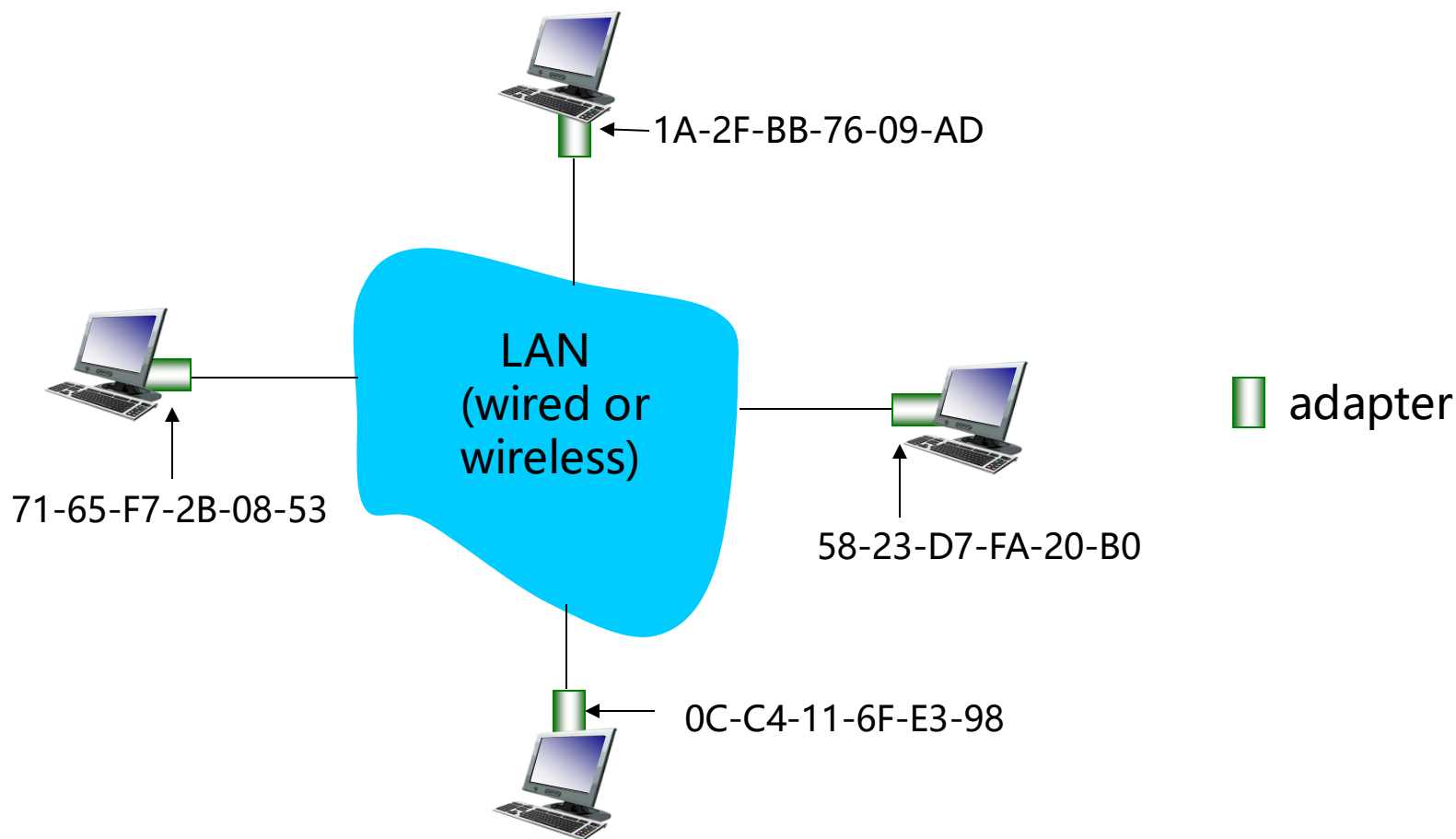
6

5.6 Web 请求生命周期中的一天



- 32位的IP地址:
  - 网络层 地址
  - 用于从目的网络获取分组 (参见 IP 地址定义, p171)
- LAN (或MAC 或物理) 地址:
  - 用来(在同一网络中)物理上互相连接的接口之间获取分组 (或帧)
  - 48 位MAC 地址 (绝大部分 LANs)
  - 烧制在适配器的 ROM中

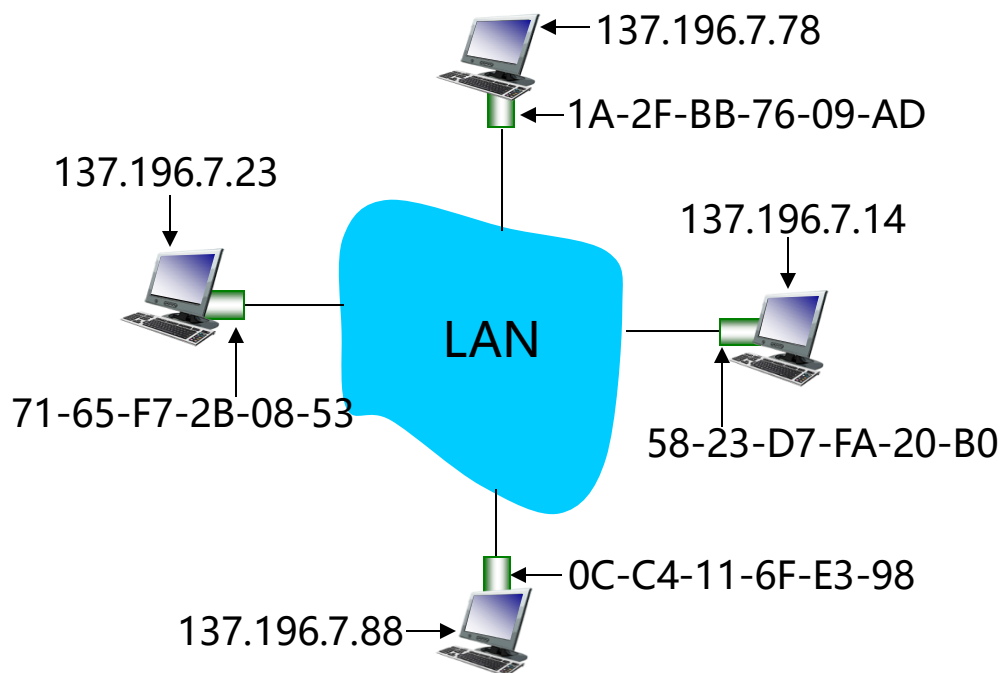
每个LAN的网卡都有唯一的LAN地址



## ■ MAC地址

- MAC 分配由 IEEE管理
- 制造商购买部分MAC地址空间 (以保证唯一性)
- 比方: (a) MAC地址: 美国人的社会保险号  
(b) IP地址: 类似邮政地址
- MAC 平面地址 => 可以迁移  
可以将 LAN卡从一个LAN换到另一个
- IP 层次性地址不可迁移  
取决于某个站点接入的网络

**Q:** 若已知B站点的IP地址, 如何确定其MAC地址?



- 每个LAN 上的IP 节点 (主机, 路由器) 都有 ARP 模块, 和表

- ARP 表: 是某些LAN 结点的IP/MAC 地址映射

< IP 地址; MAC 地址; TTL >

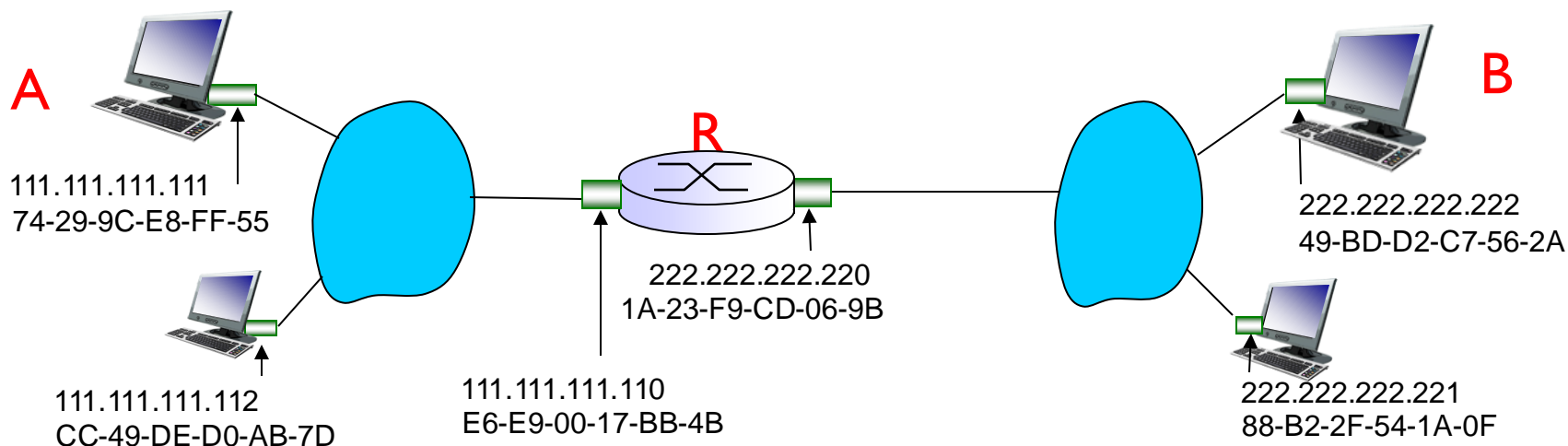
TTL (Time To Live): 超过TTL 的地址映射会被删除 (一般为 20 分钟)

- ❖ A想要发送数据报给B
  - b的MAC地址不在A的ARP表中
- ❖ A广播ARP查询数据包，其中包含B的IP地址
  - 目的MAC地址= FF-FF-FF-FF-FF-FF
  - 局域网上的所有节点都接收ARP查询
- ❖ B接收到 ARP帧, 将其物理地址返回给A
  - 发送到A的MAC地址的帧(单播)
- ❖ A 对收到的IP/MAC地址对进行缓存直到信息过期 (超时)
  - 软状态:除非定期刷新, 否则信息会超时(消失)
- ❖ ARP是 “即插即用” 的:
  - 节点创建自己的ARP表, 无需网络管理员的干预

# 寻址：发送数据报到子网之外

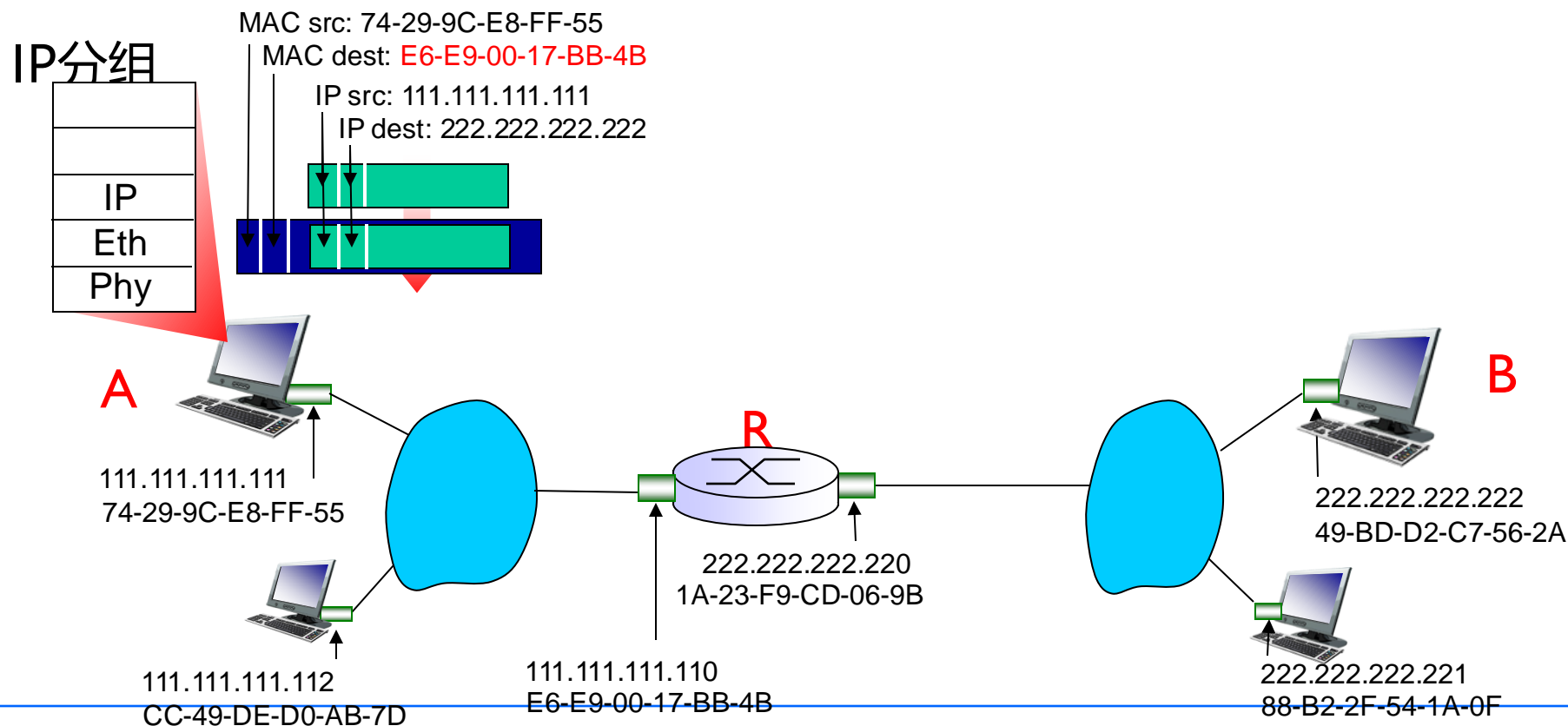
案例：通过 R 将数据报从 A 发送到 B

- 专注于寻址 – 在 IP（数据报）和 MAC 层（帧）
- 假设 A 直到 B 的 IP 地址
- 假设 A 知道第一跳路由器 R 的 IP 地址(怎么知道的？)
- 假设 A 知道 R 的 MAC 地址（怎么知道的？）



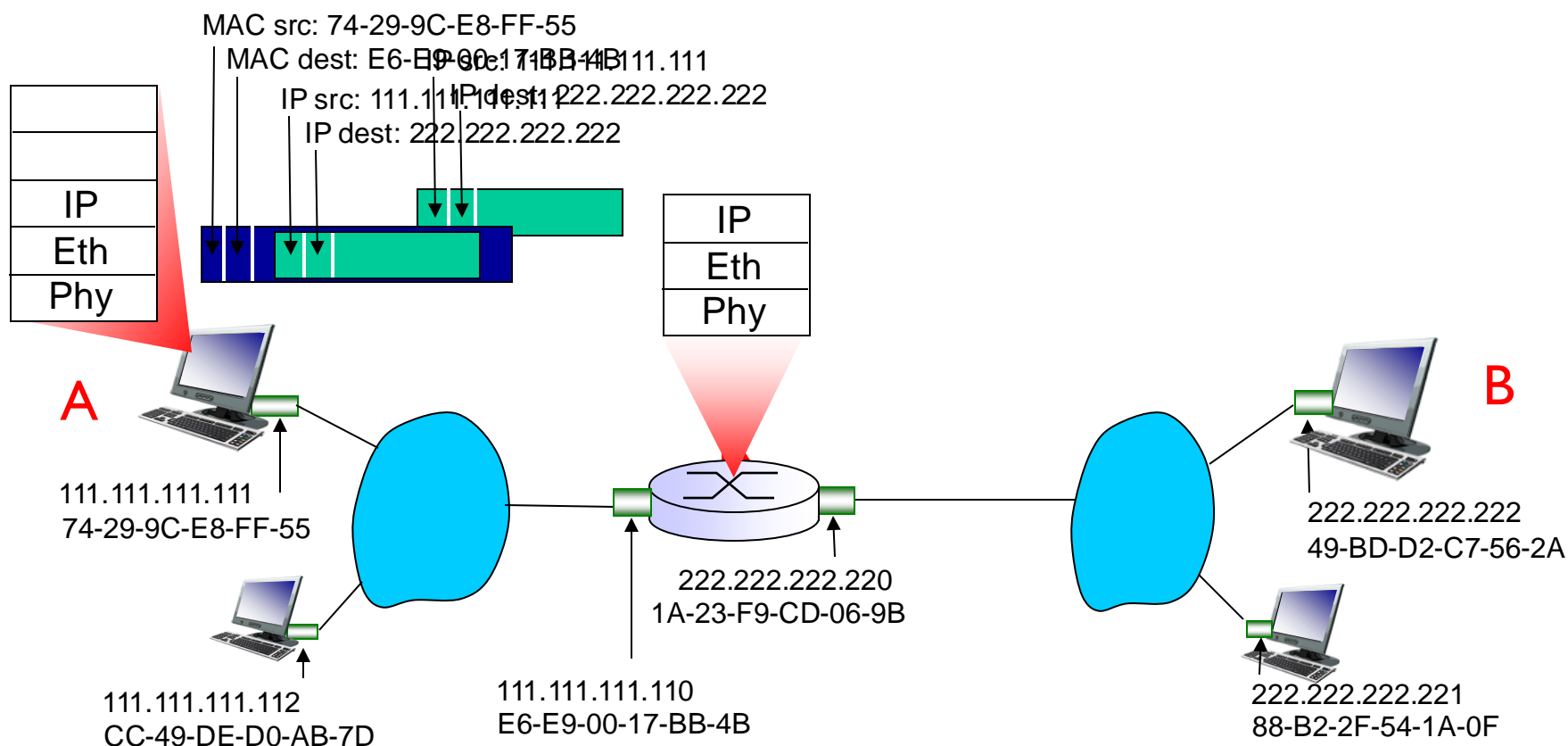
# 寻址：发送数据报到子网之外

- ❖ A 创建了 IP 分组，源地址为 A, 宿地址为 B
- ❖ A 使用 ARP 来获取 R 的与 111.111.111.110 对应的物理地址
- ❖ A 创建了以 R 的物理地址为宿地址的以太网帧，该帧包含的 A-to-B 的



# 寻址：发送数据报到子网之外

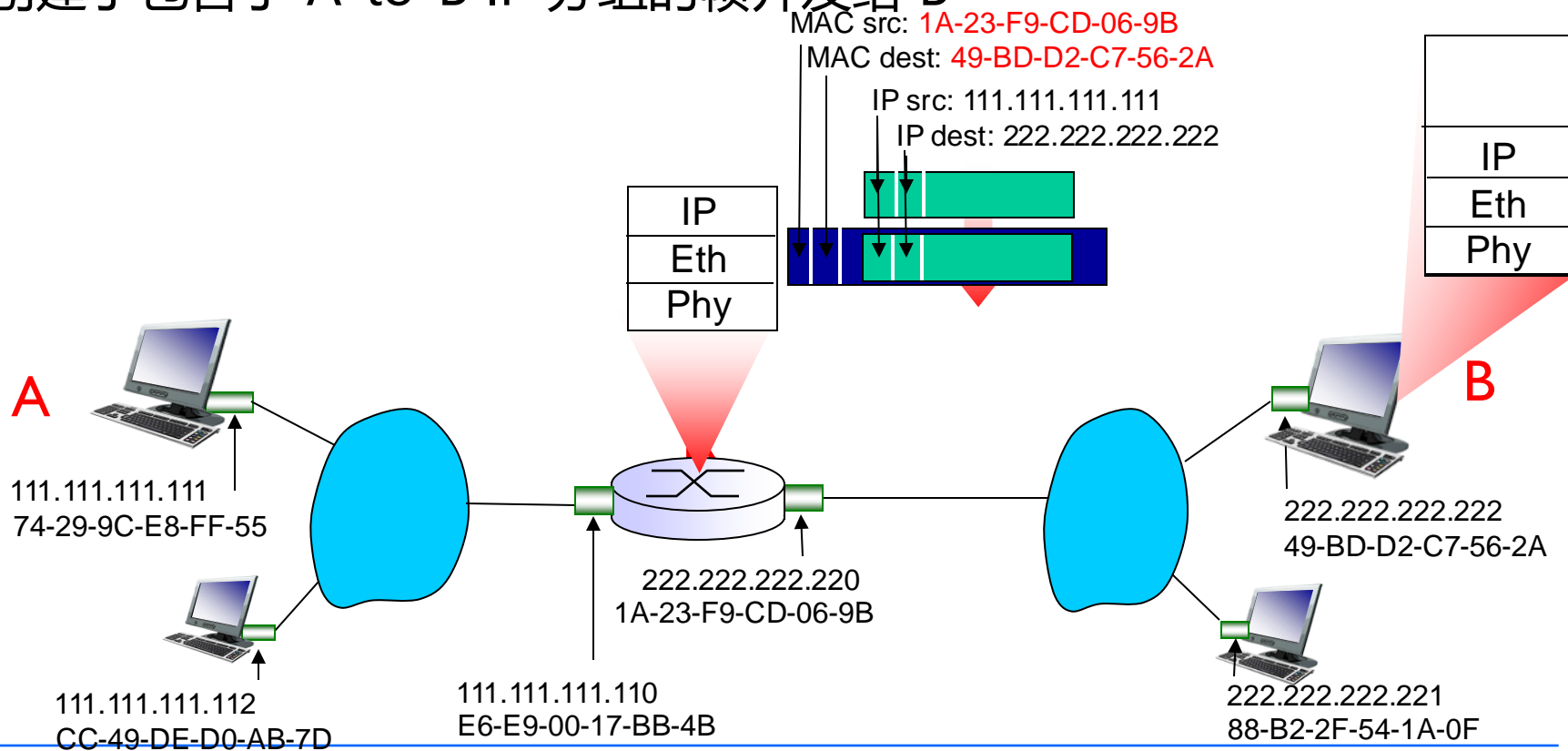
- ❖ A 的数据链路层发送以太网的帧
- ❖ R 的数据链路层接收到以太网的帧





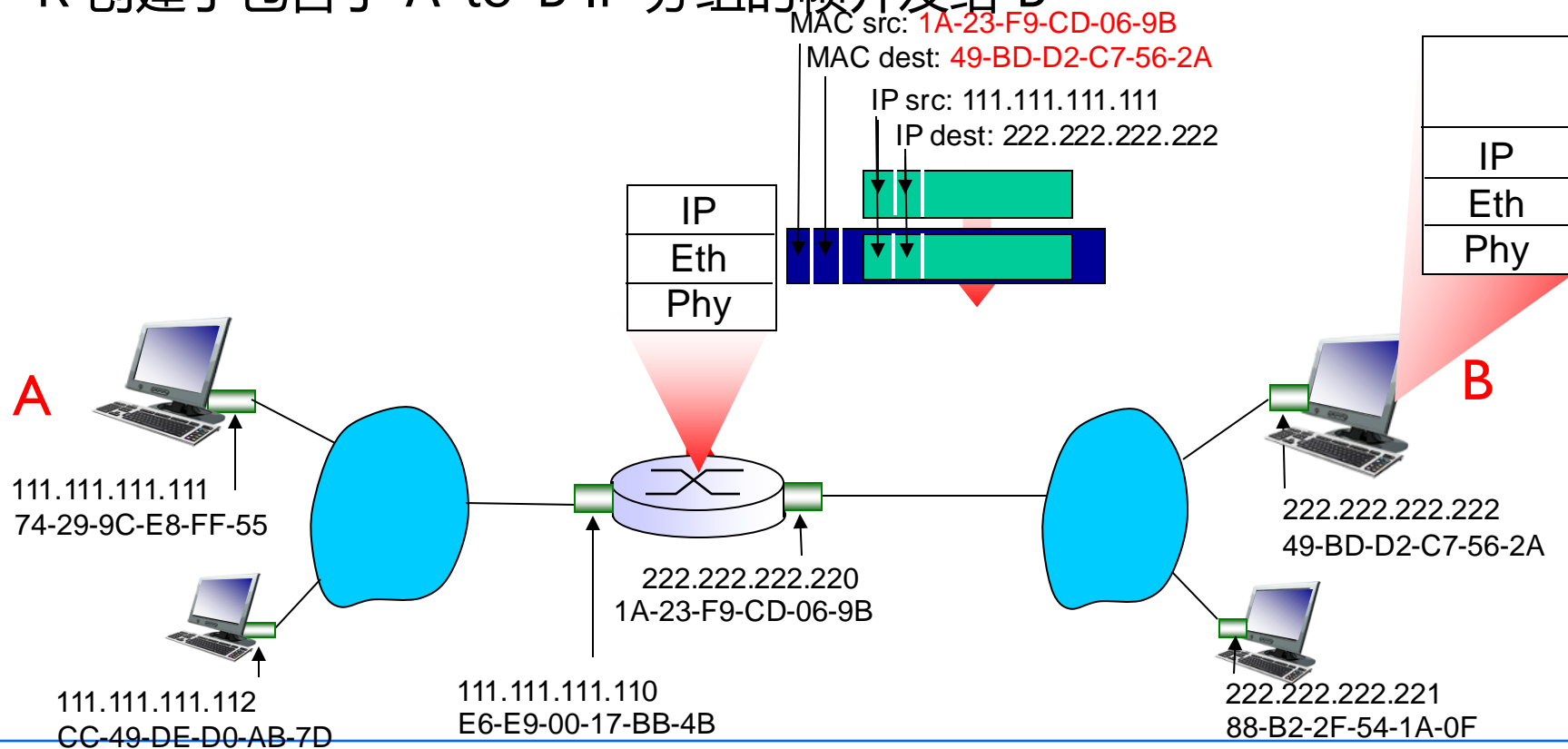
# 寻址：发送数据报到子网之外

- ❖ R 从以太网帧中取出 IP 分组,知道该分组的信宿为 B
- ❖ R 使用ARP 来取得 B的物理层地址
- ❖ R 创建了包含了 A-to-B IP 分组的帧并发给 B



# 寻址：发送数据报到子网外

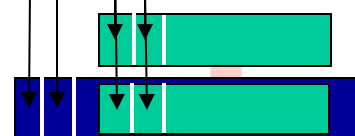
- ❖ R 从以太网帧中取出 IP 分组,知道该分组的信宿为 B
- ❖ R 使用ARP 来取得 B 的物理层地址
- ❖ R 创建了包含了 A-to-B IP 分组的帧并发给 B



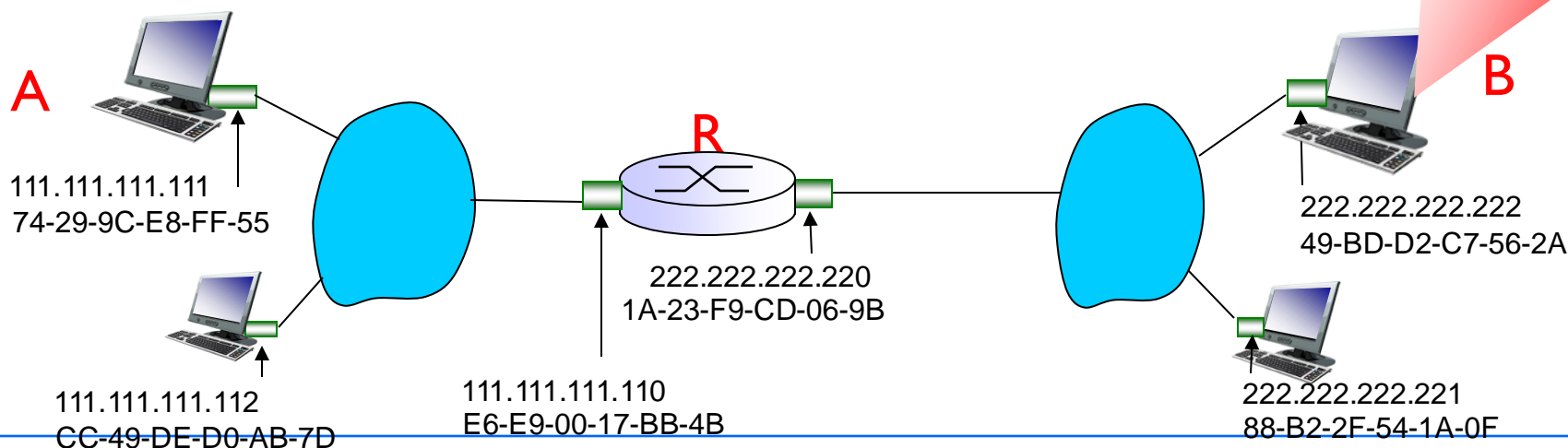
# 寻址：发送数据报到子网外

- ❖ R 从以太网帧中取出 IP 分组,知道该分组的信宿为 B
- ❖ R 使用ARP 来取得 B的物理层地址
- ❖ R 创建了包含了 A-to-B IP 分组的帧并发给 B

MAC src: 1A-23-F9-CD-06-9B  
MAC dest: 49-BD-D2-C7-56-2A  
IP src: 111.111.111.111  
IP dest: 222.222.222.222



IP
Eth
Phy



1

5.1 引论和服务

2

5.2 差错检测和纠正技术

3

5.3 多路访问链路和协议

4

5.4 链路层

1. 寻址和ARP    2. 以太网    3. 交换机    4. 虚拟局域网

5

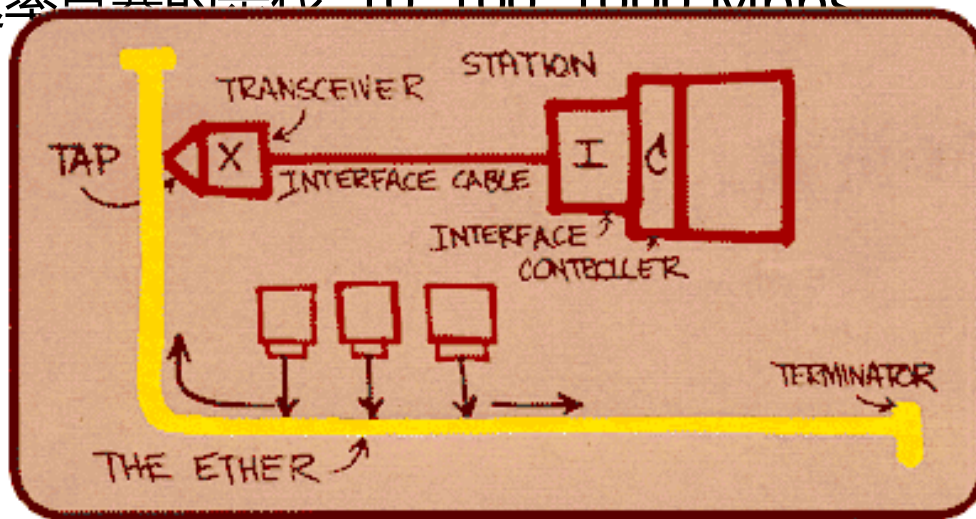
5.5 链路虚拟化: MPLS

6

5.6 Web 请求生命周期中的一天

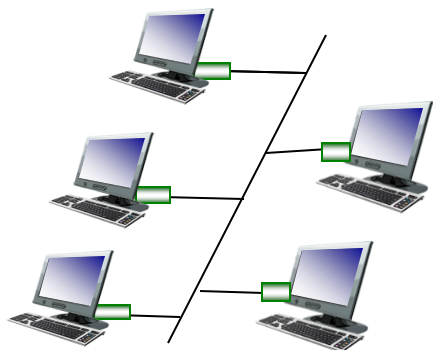
“统治” LAN的技术:

- 便宜, ¥ 50 for 100Mbps!
- 最早被广泛应用的 LAN 技术
- 较为简单, 比 token LANs 和ATM便宜
- 赶上了速率竞赛的步伐: 10, 100, 1000 Mbps

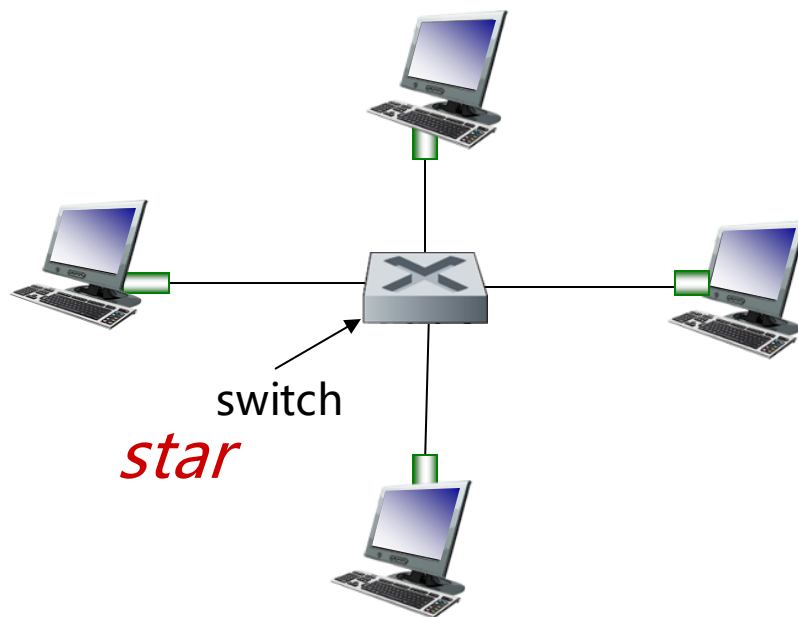


梅特卡夫的以太网草图

- **总线型**：90年代中期流行
  - 所有的结点在同一个冲突域中
- **星型**：目前的主流
  - 星型：目前的主流



总线：同轴电缆



发送适配器将IP分组封装在以太网帧中(或其他网络层协议分组)



preamble (前序) :

- 7 个 10101010 字节尾随一个 10101011 字节
- 用来同步收发双方的时钟速率

- ❖ **地址:** 6 个字节, 帧为某个LAN上的所有适配器接收, 但只要地址不匹配就被丢弃
  - 如果适配器接收到具有匹配目标地址或广播地址 (例如 ARP 数据包) 的帧, 它将帧中的数据传递到网络层协议
  - 否则, 适配器将丢弃帧
- ❖ **类型:** 说明其上层协议, 大部分为 IP, 但其他协议如Novell IPX和 AppleTalk 也支持
- ❖ **CRC:** 在接收端校验, 如果出错, 则将该帧丢弃



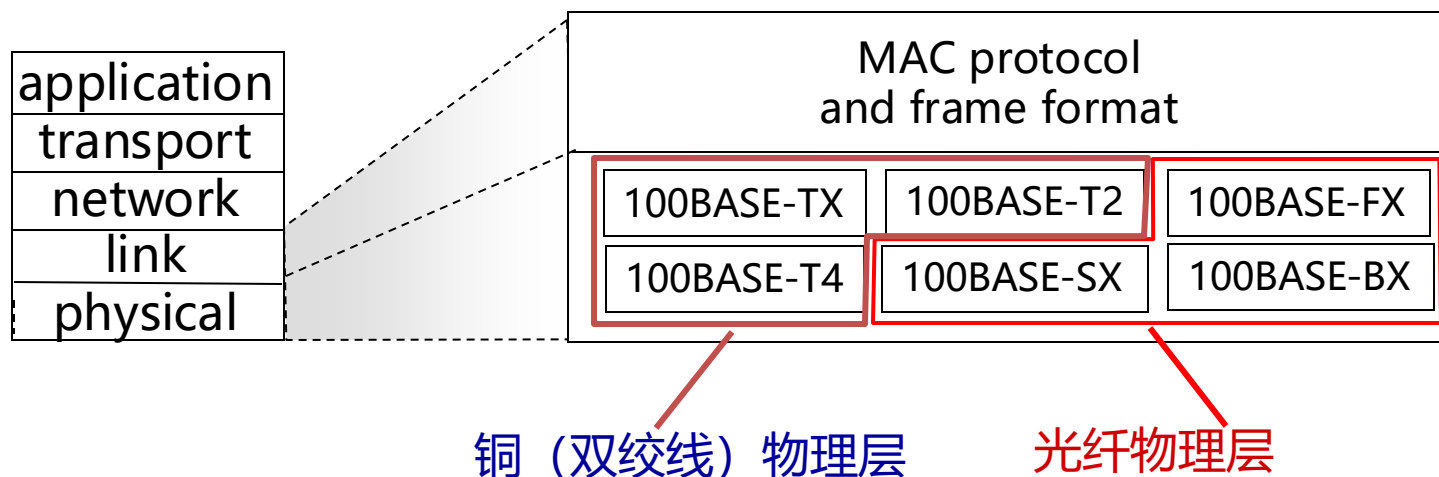


- **无连接**：发送端和接收端网卡都不需要握手
- **不可靠性**：接收端网卡不需要发送acks和nacks给发送网卡
  - 丢失帧中的数据只在高层的rdt机制中恢复
- 以太网的介质访问控制协议

# 802.3以太网标准：链路层和物理层

## ■ 许多不同以太网标准

- 公用的MAC协议（介质访问控制）和帧格式
- 不同的传输速度：2Mbps, 10Mbps, 100Mbps, 1Gbps, 10Gbps
- 不同的物理层介质：光纤，同轴电缆，双绞线



1

5.1引论和服务

2

5.2差错检测和纠正技术

3

5.3多路访问链路和协议

4

5.4链路层

1.寻址和ARP    2.以太网    3.交换机    4.虚拟局域网

5

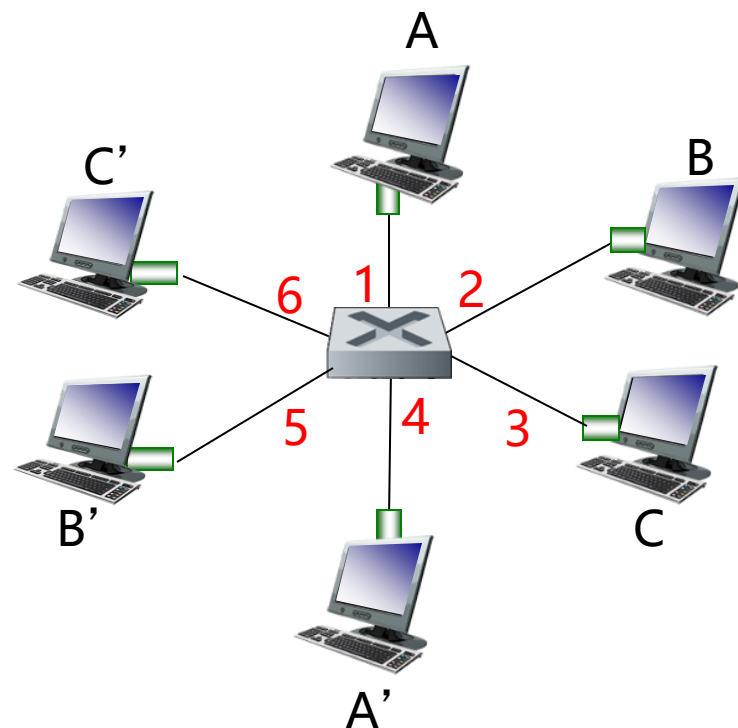
5.5链路虚拟化: MPLS

6

5.6 Web 请求生命周期中的一天

- 链路层设备
  - 存储、转发以太网帧
  - 检查进入设备的帧的MAC地址，选择性将帧转发至1个或多个输出链路，采用CSMA/CD策略
- 透明性-主机和路由器感觉不到交换机的存在
- 即插即用，自学习-交换机不需要配置

- 主机直连至交换机
- 交换机有缓存能力
- 采用以太网协议的各相连链路（可能是异构的），彼此隔离，不会产生冲突，全双工
  - 每条链路是一个冲突域
- **交换**- A-to-A' and B-to-B同时传输，无冲突



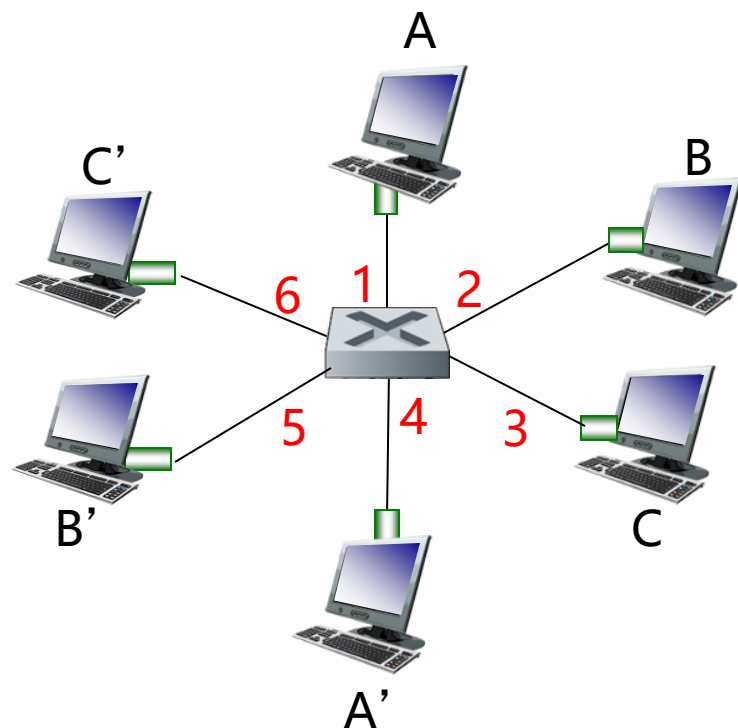
带六个接口的交换机  
(1,2,3,4,5,6)

**Q:** 交换机如何知道A' 通过接口4到达, B' 通过接口5到达?

**A:** 每台交换机有一个交换表, 每条表现记录:

主机的MAC地址, 通向MAC地址的  
交换机接口, 表现放置在表中时间

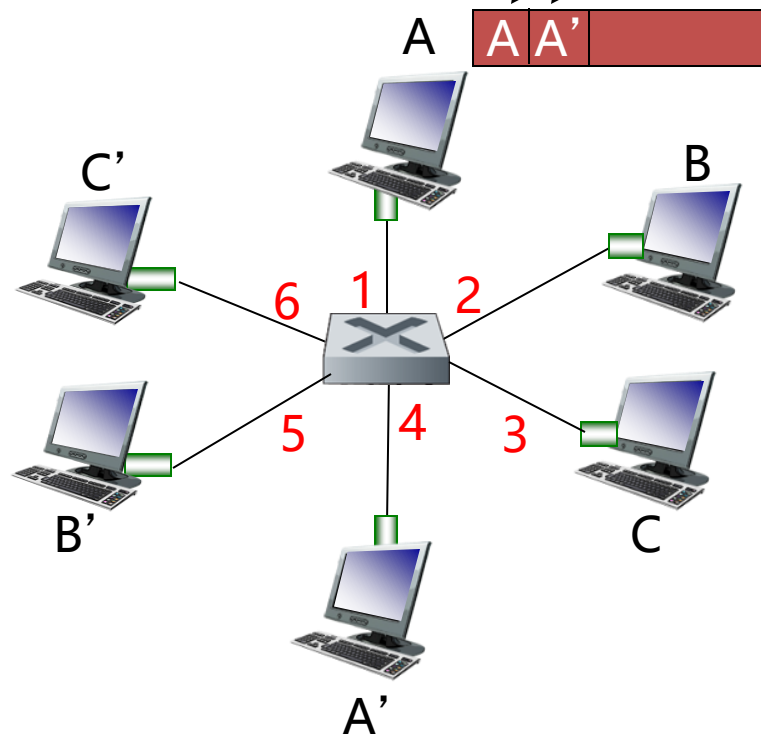
类似于**路由表**



带六个接口的交换机  
(1,2,3,4,5,6)

**Q:** 表项如何创建, 并在交换表中维持?

Source: A  
Dest: A'



- 交换机了解可以通过哪些接口访问哪些主机
  - 收到帧时，切换“学习”发送方的位置：传入 LAN 段
  - 在交换机表中记录发件人/位置对

MAC addr	interface	TTL
<i>A</i>	<i>1</i>	<i>60</i>

开关表  
(最初为空)

在交换机接收帧时：

1.记录传入链路，发送主机的MAC地址

2.使用 MAC 目标地址的索引交换机表

3. if entry found for destination

then {

if destination on segment from which frame arrived

then drop frame

else forward frame on interface indicated by entry

}

else flood /\* forward on all interfaces except arriving

interface \*/

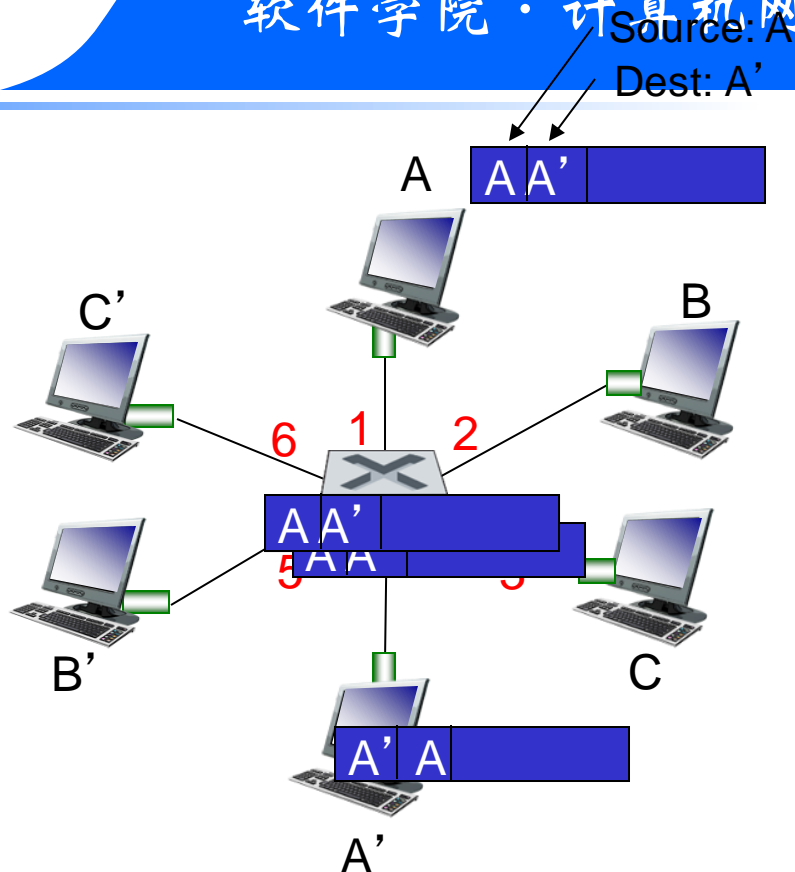


- 帧目的地，A'，定位未知：

- ❖ 目的地 已知位置：

有选择地发送

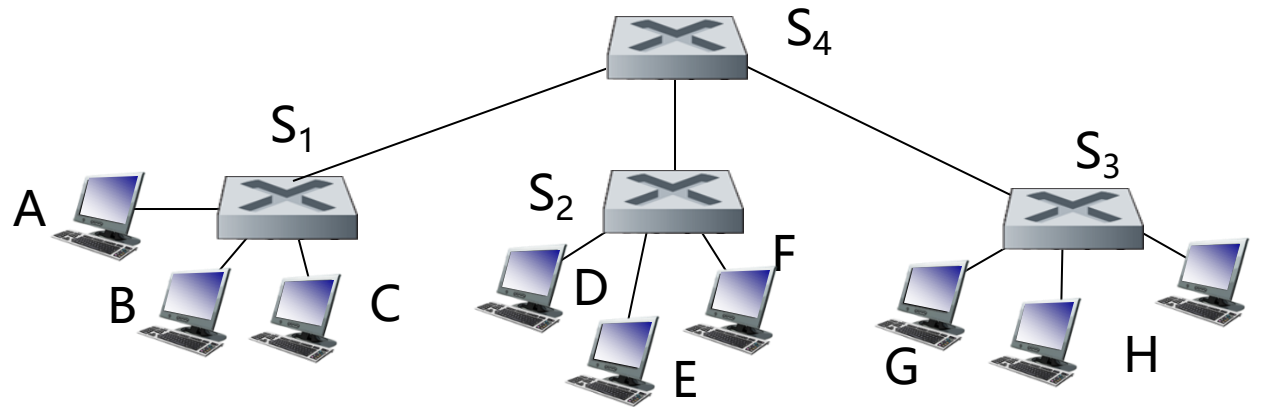
仅在一个链接上



MAC addr	interface	TTL
A	1	60
A'	4	60

switch table  
(initially empty)

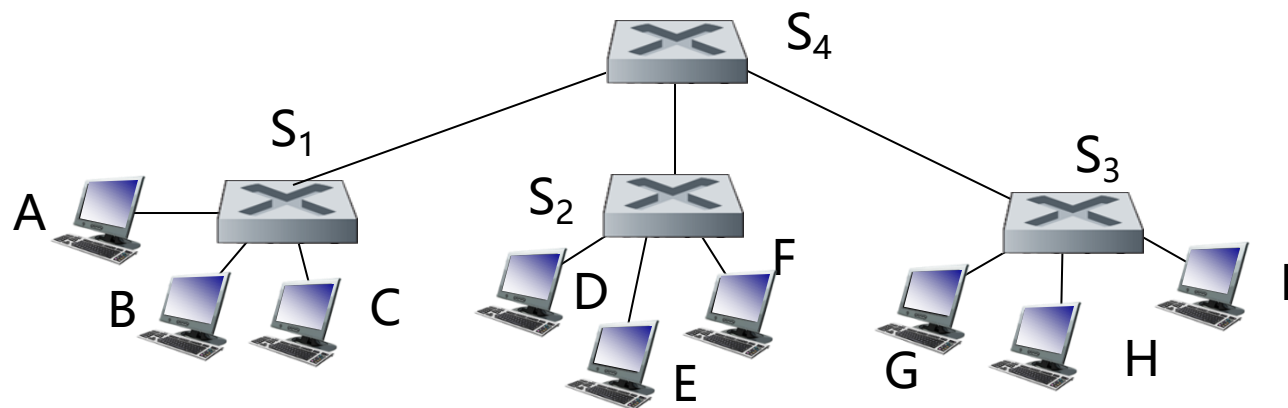
## ❖ 交换机之间也能彼此互联



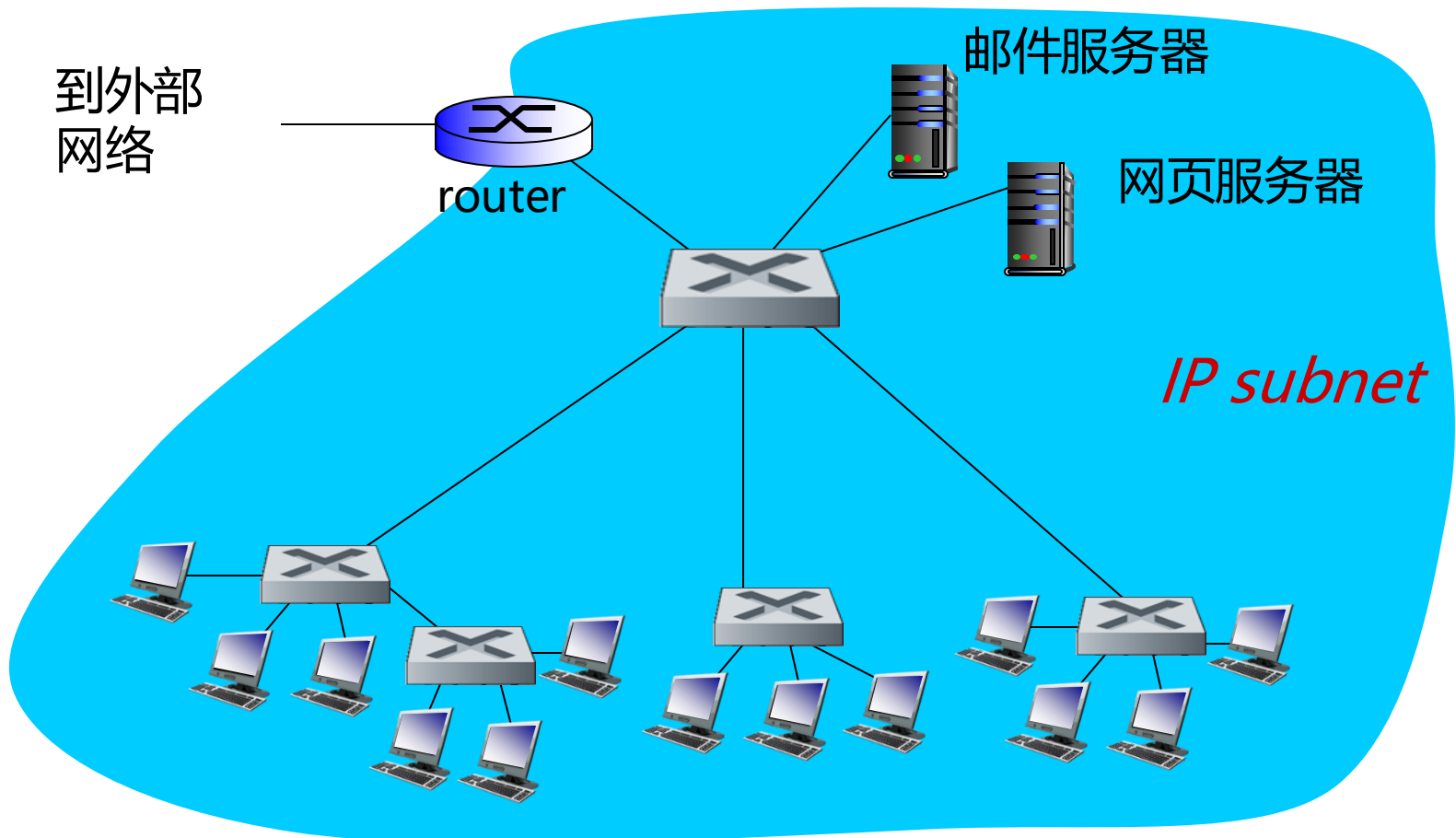
Q: 从A向G发送，S1如何知道将帧通过S4和S3转发至G？

A: 自学习! (工作原理与单开关情况完全相同)

假设 C 将帧发送给 I, I 响应 C



- ❖ Q: 显示 S1、S2、S3、S4 中的交换表和数据包转发



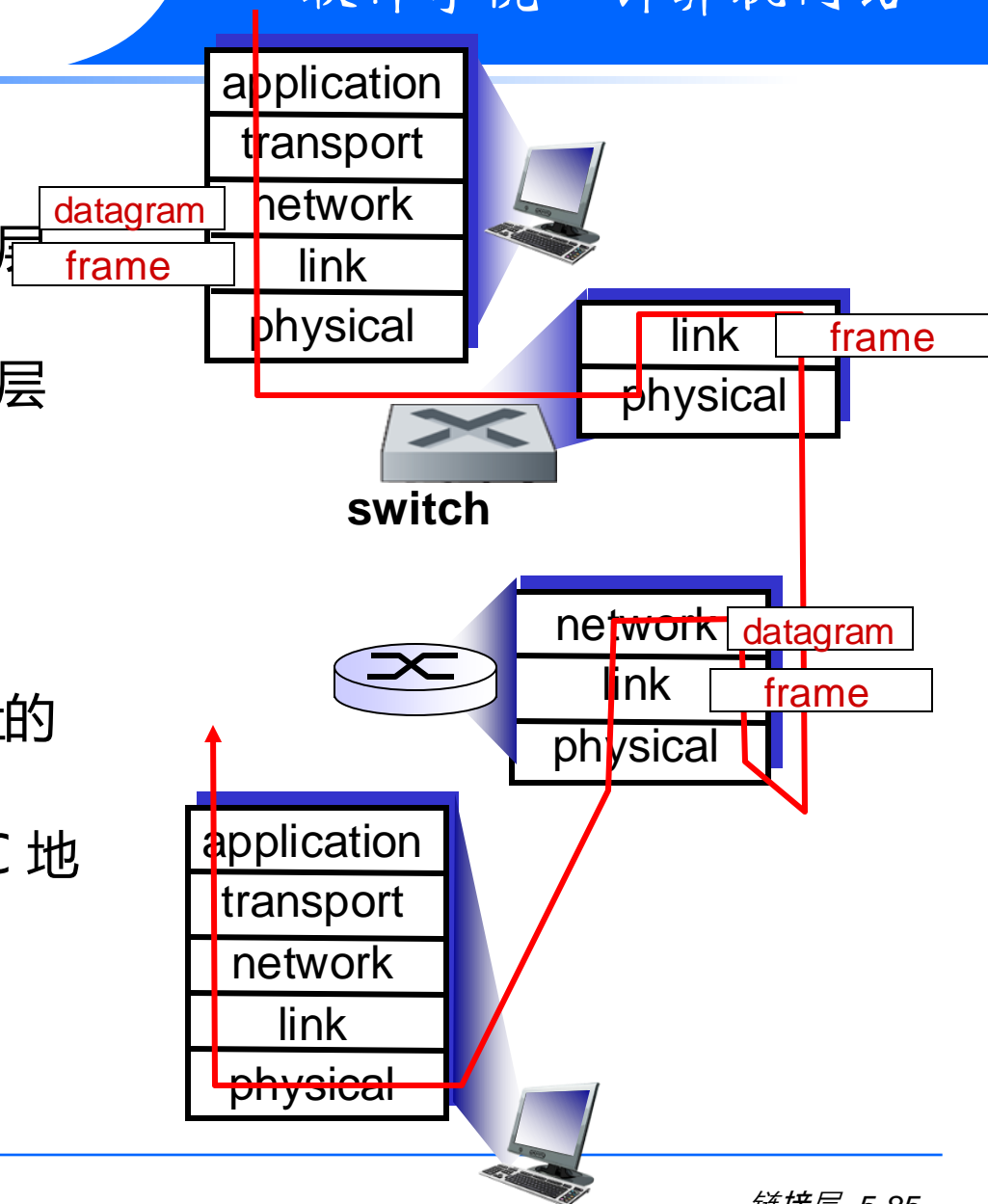
# 交换机与路由器

## 两者都是存储转发:

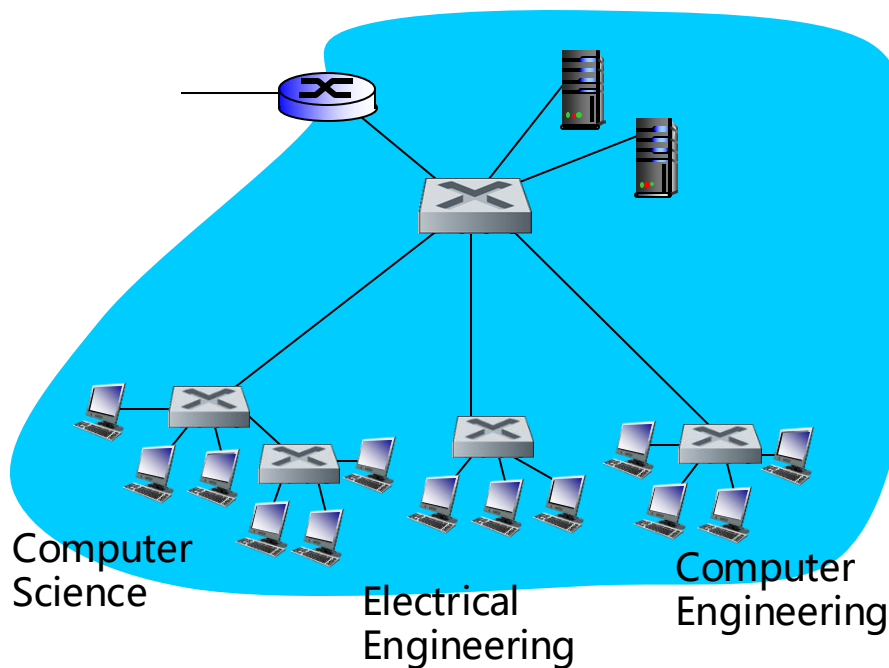
- 路由器: 网络层设备 (检查网络层头部)
- 交换机: 链路层设备 (检查链路层头部)

## 两者都有转发表:

- 路由器: 使用路由算法、IP 地址的计算表
- 交换机: 使用泛洪、学习、MAC 地址学习转发表



思考:

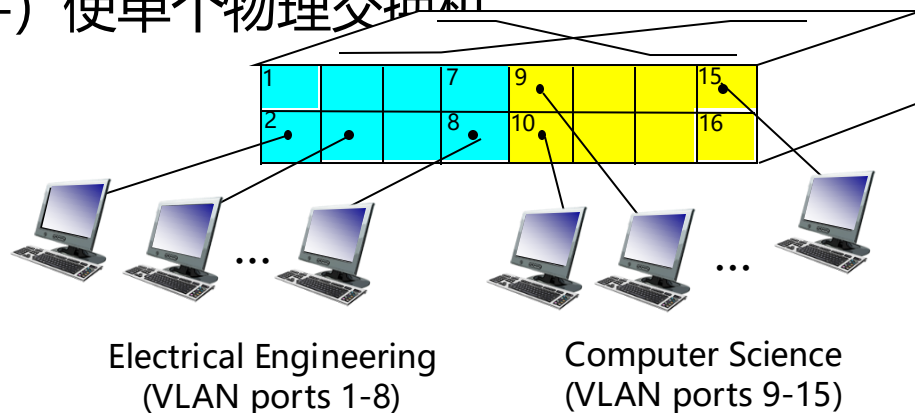


- CS 用户将办公室迁移到 EE，但想要连接到 CS 交换机？
- 单个广播域：
  - 所有第 2 层广播流量（ARP、DHCP、目标 MAC 地址的未知位置）必须跨越整个 LAN
  - 安全/隐私、效率问题

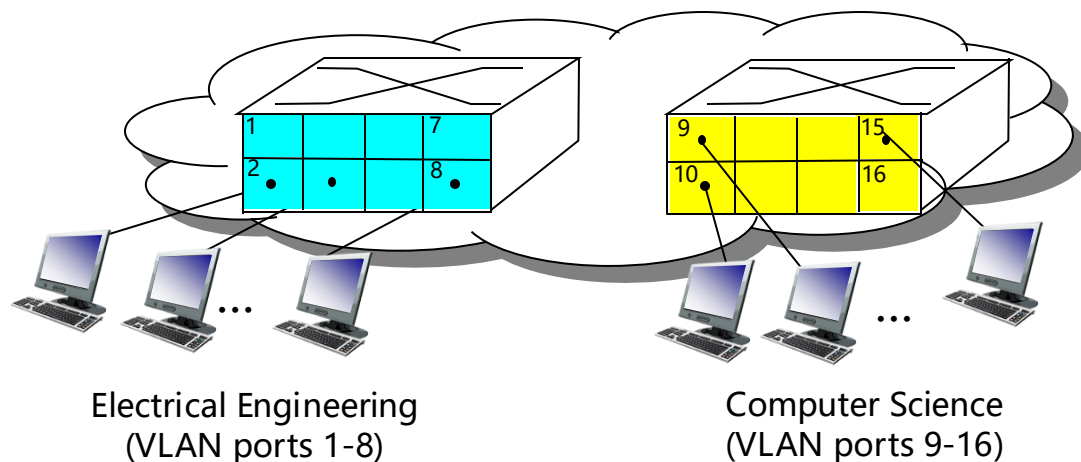
**基于端口的VLAN：**将交换机端口分组（通过交换机管理软件）使单个物理交换机

## 虚拟本地局域网

可以将支持 VLAN 功能的交换机配置为在单个物理 LAN 基础架构上定义多个虚拟 LAN。



...作为**多个**虚拟交换机运行



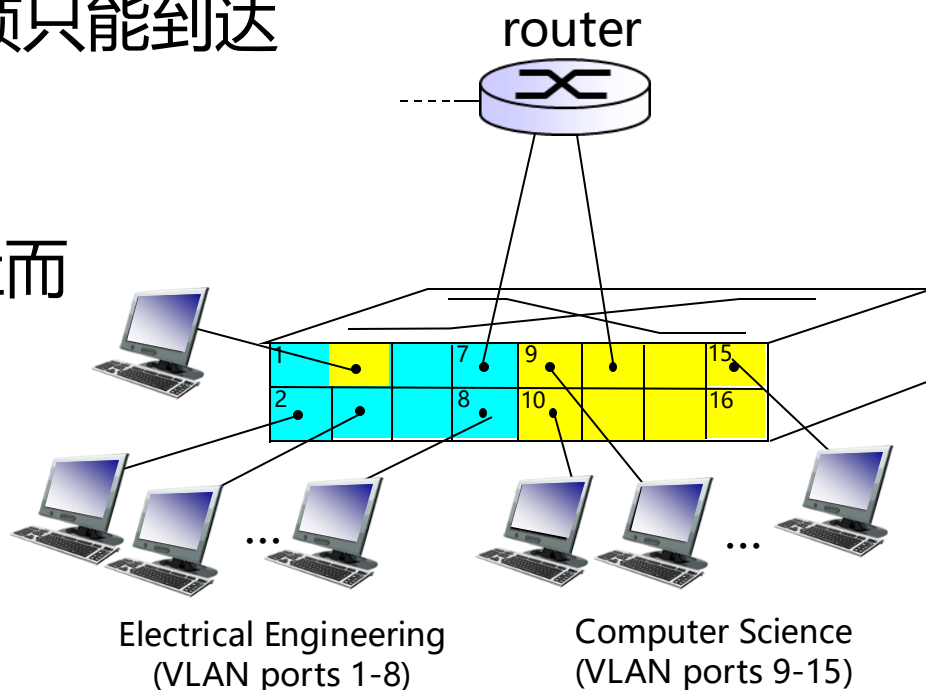
- ❖ **流量隔离**：进出端口 1-8 的帧只能到达端口 1-8

- ❖ 还可以根据端点的 MAC 地址而不是交换机端口定义 VLAN

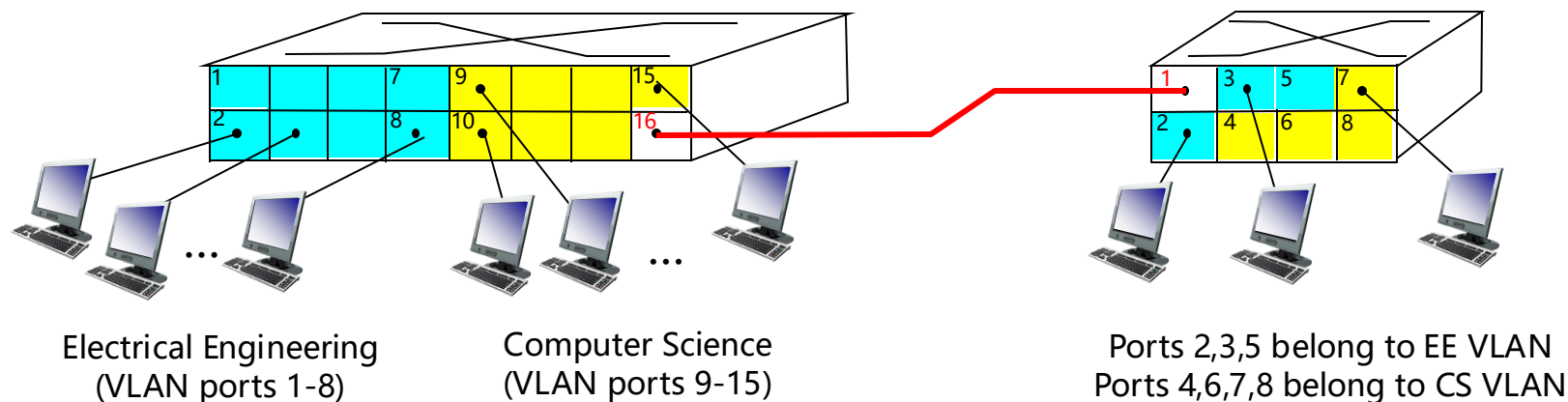
- ❖ **动态成员资格**：端口可以在 VLAN 之间动态分配

- ❖ **在 VLAN 之间转发**：通过路由完成（就像使用单独的交换机一样）

- ❖ 实际上，供应商销售组合交换机和路由器

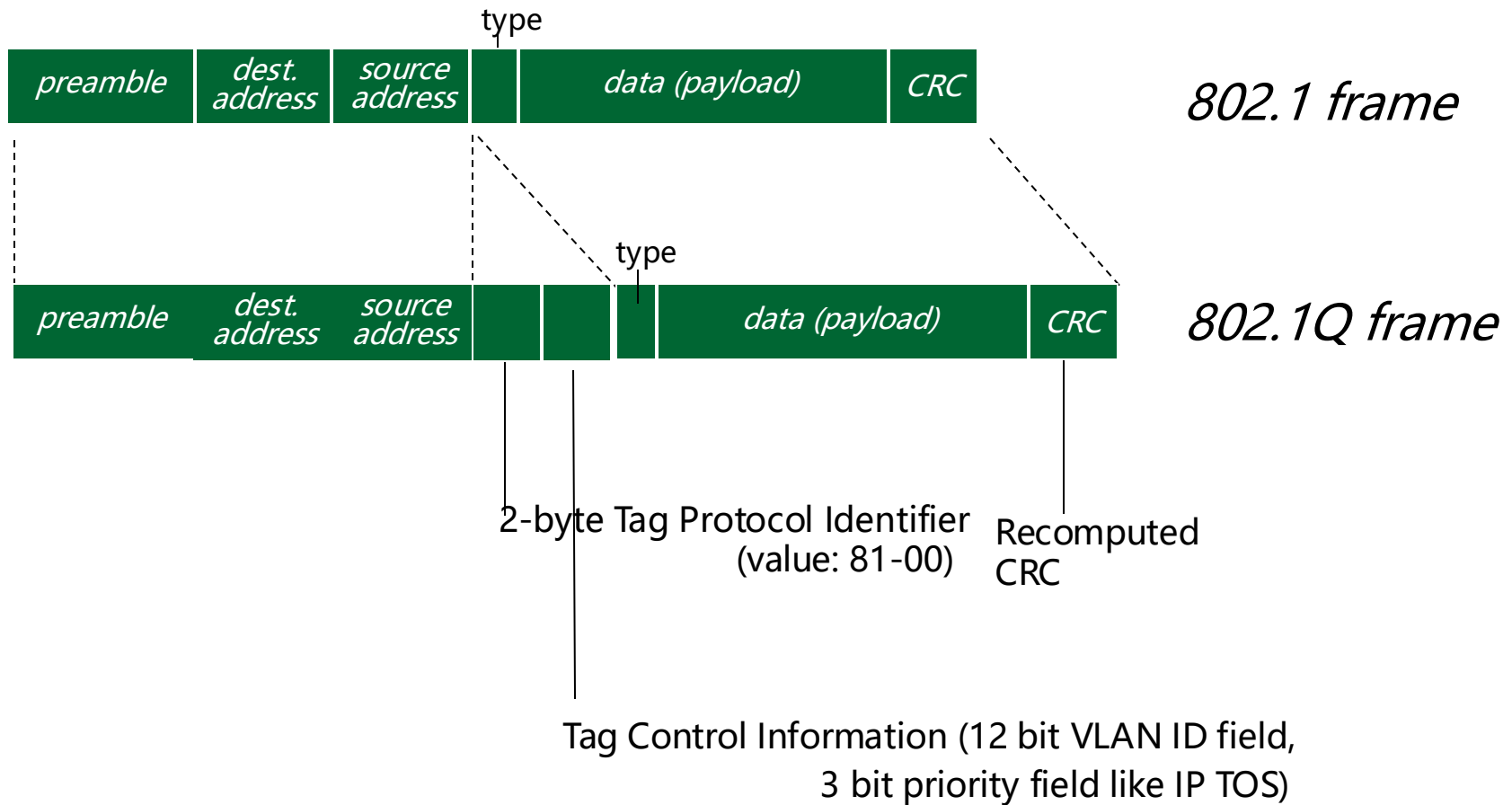






- **中继端口：**在通过多个物理交换机定义的 VLAN 之间传输帧
  - 在交换机之间的 VLAN 内转发的帧不能是原版 802.1 帧（必须携带 VLAN ID 信息）
  - 802.1q 协议为中继端口之间转发的帧添加/删除了其他标头字段

# 802.1Q VLAN 帧格式



1

5.1引论和服务

2

5.2差错检测和纠正技术

3

5.3多路访问链路和协议

4

5.4链路层

1.寻址和ARP    2.以太网    3.交换机    4.虚拟局域网

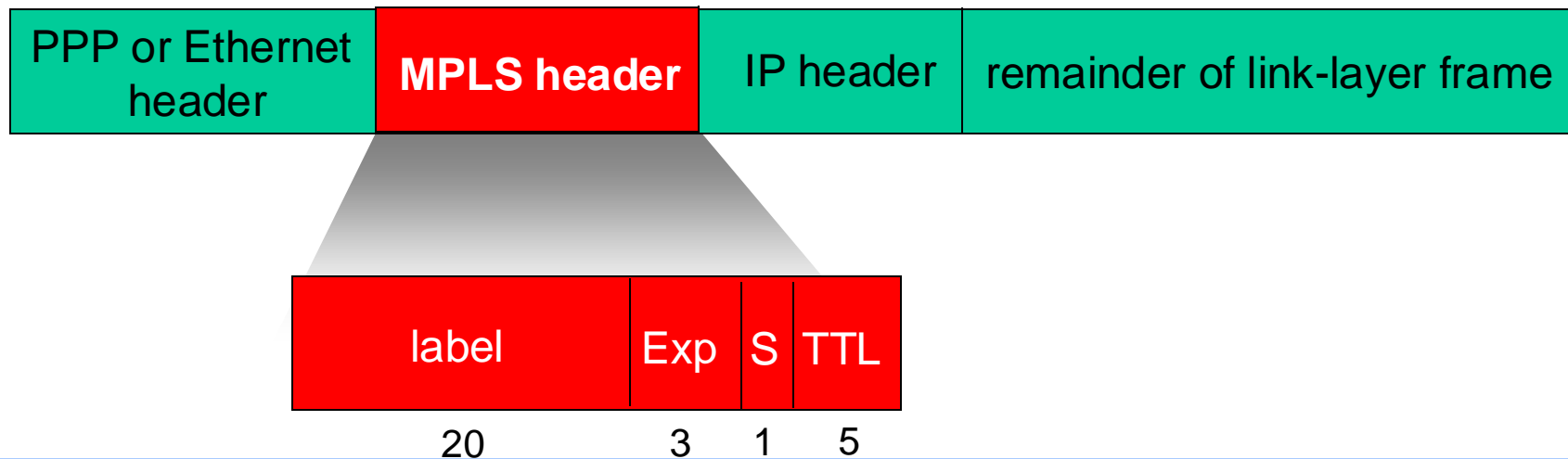
5

5.5链路虚拟化: MPLS

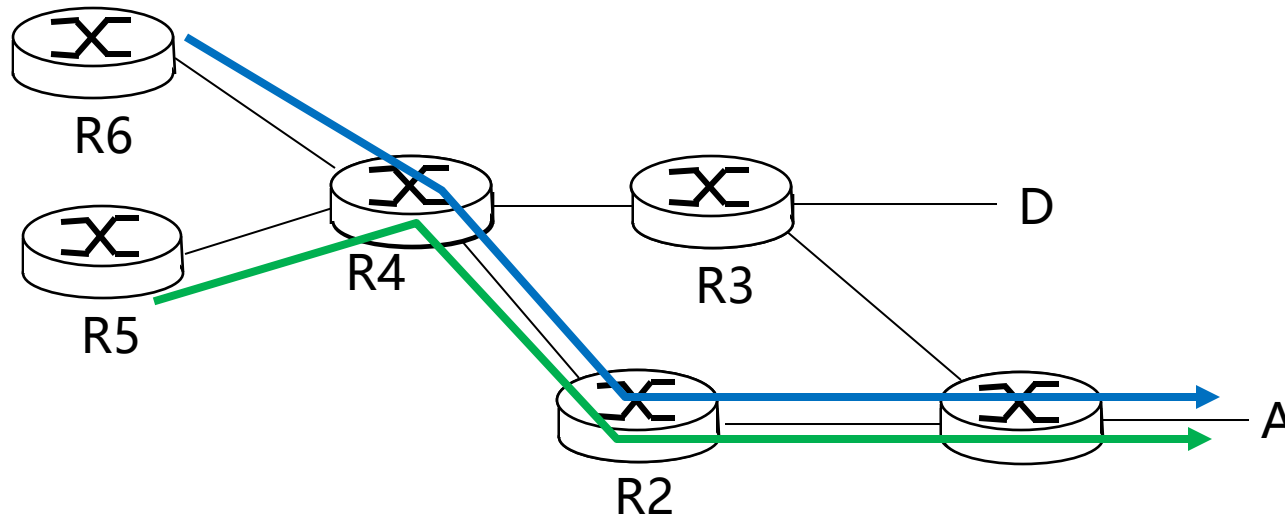
6

5.6 Web 请求生命周期中的一天

- 最初的目标：使用固定长度标签（不是目的IP地址）进行高速的IP转发
  - 相比最短前缀匹配，使用定长标识更快
  - 借用了虚电路的一些思路
  - IP数据中IP地址依然存在

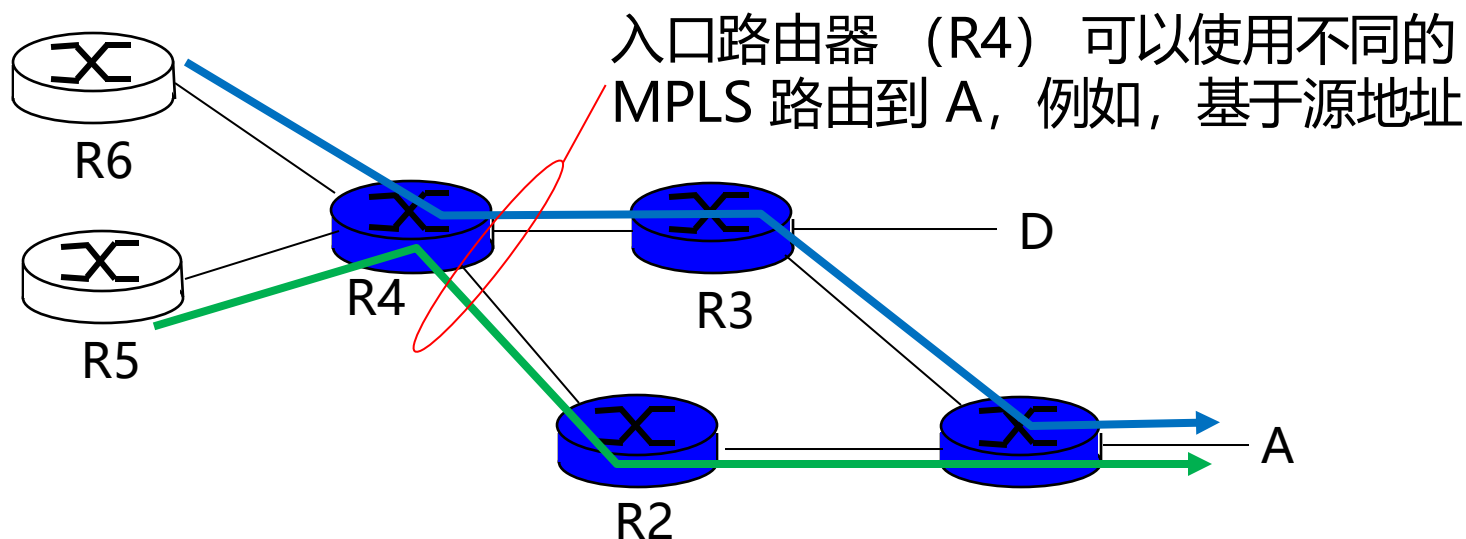


- 也称为标签交换路由器
- 转发分组的依据是标签值（而不是IP地址）-MPLS转发表和路由表是有区别的
- **灵活性**：MPLS转发的决策不同于IP路由
  - 提供了沿着多条路由转发分组的能力
  - 经过一条预计算的无故障路径重路由流量来对链路故障做出反应（对VoIP有用）



❖ **IP 路由:**仅由目标地址确定的到达目的地的路径





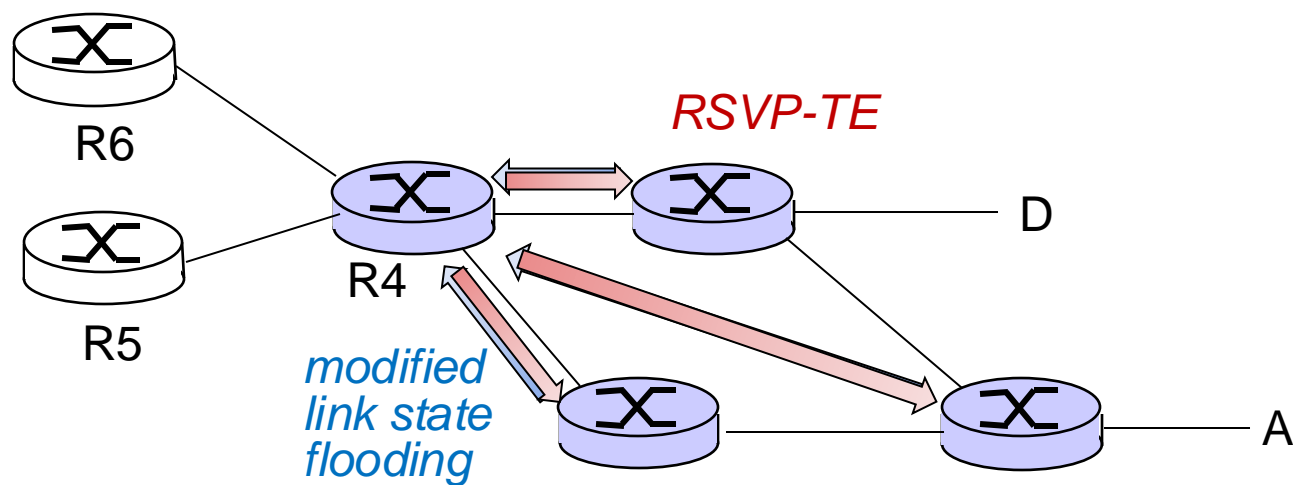
❖ **IP 路由**: 仅由目标地址确定到目标的路径

❖ **MPLS routing**: 到目标的路径可以基于源地址和目标地址



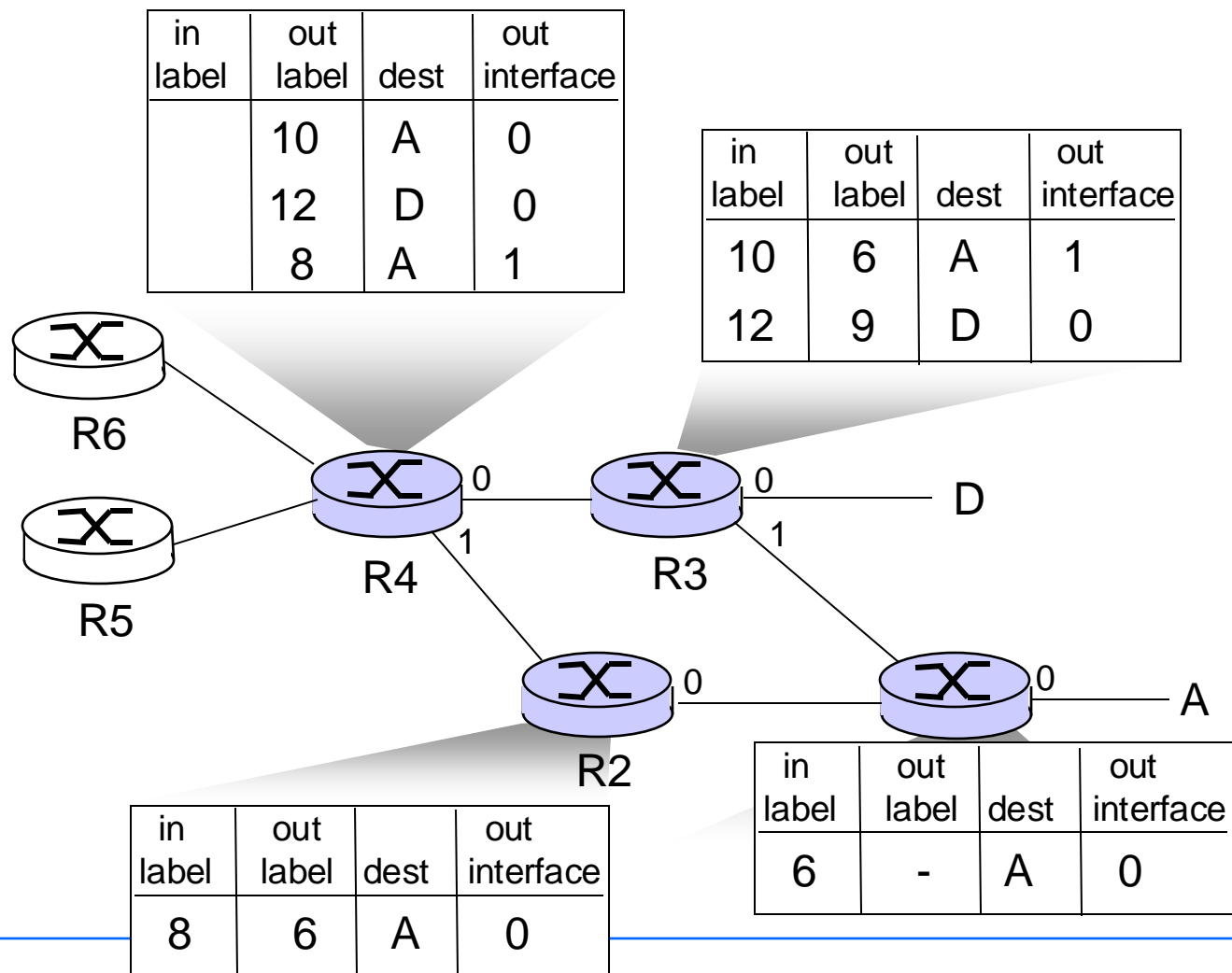
- **快速重新路由**: 在链路发生故障时预先计算备份路由

- 现有的链路状态路由选择算法如OSPF，ISIS已经扩展为向MPLS使能路由器“洪泛”
- MPLS路由器利用RSVP-TE（RSVP扩展）信令建立与其下游路由器的MPLS转发





# MPLS 转发表



1

5.1引论和服务

2

5.2差错检测和纠正技术

3

5.3多路访问链路和协议

4

5.4链路层

1.寻址和ARP    2.以太网    3.交换机    4.虚拟局域网

5

5.5链路虚拟化: MPLS

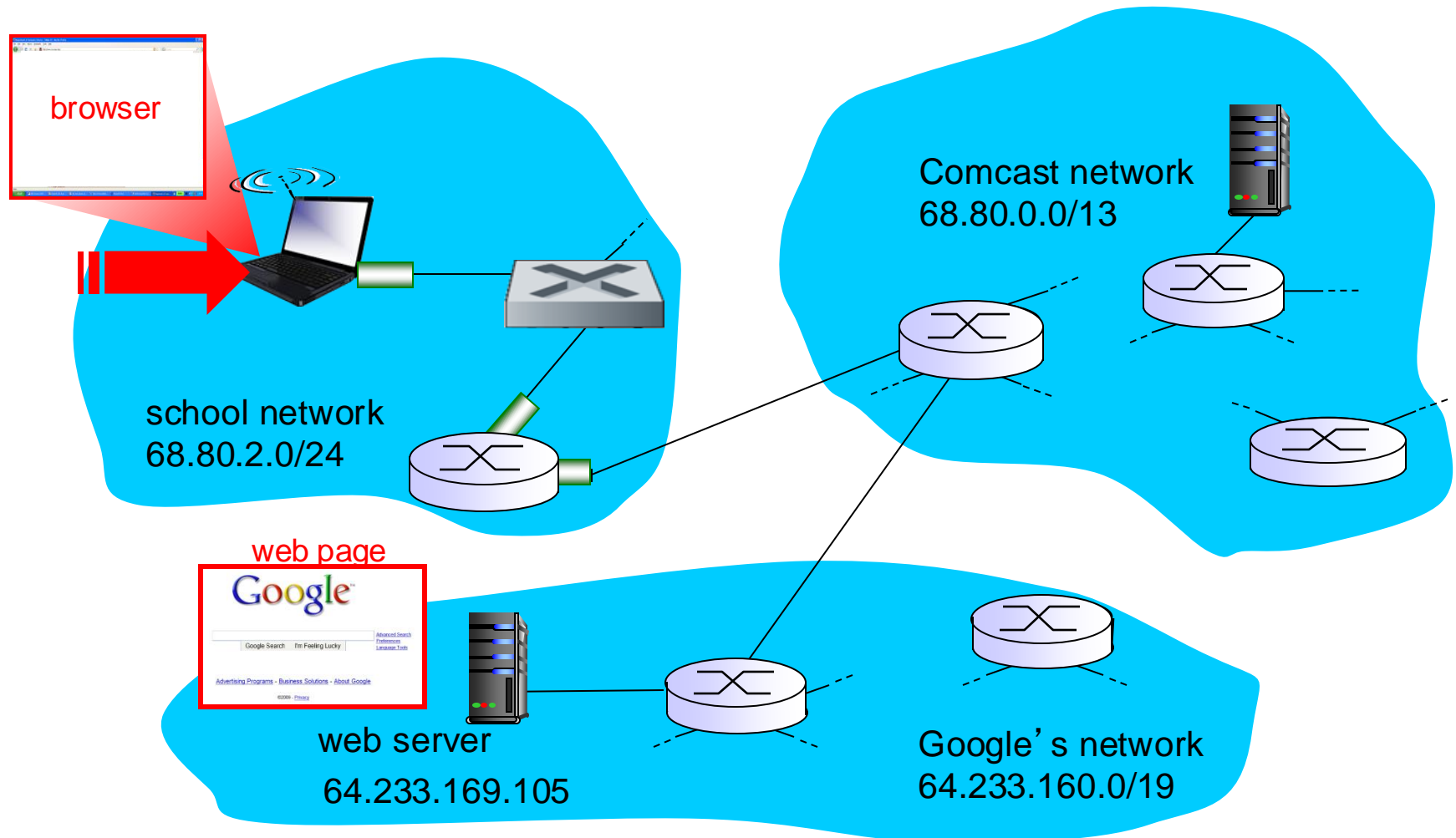
6

5.6 Web 请求生命周期中的一天

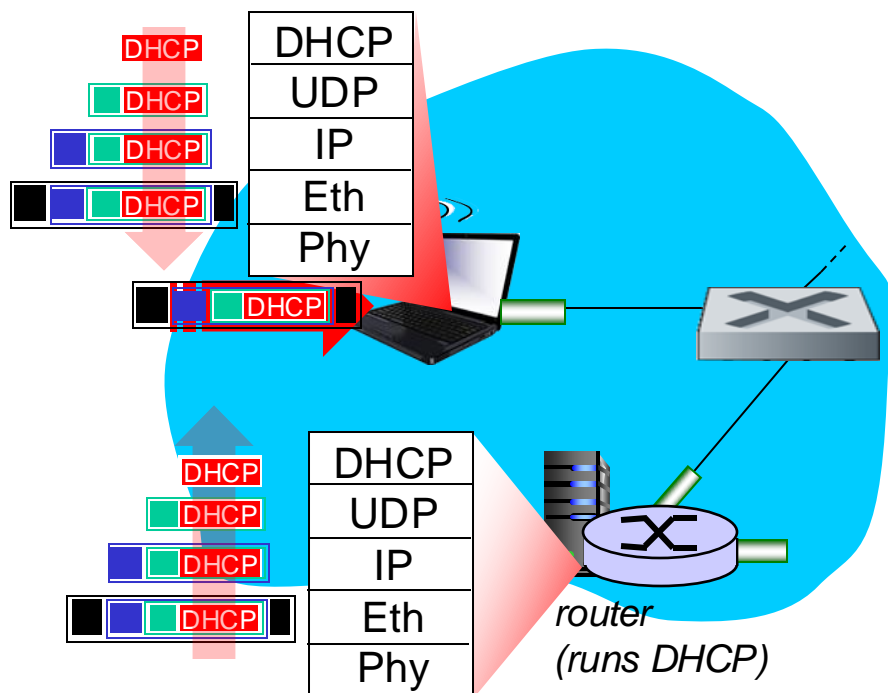
- 协议栈下行完成!
  - 应用， 传输， 网络， 链接
- 把它放在一起：综合!
  - **目标：** 复习、理解、区别不同的协议在一个简单的应用场合的作用
  - **场景：** 使用浏览器访问 [www.google.com](http://www.google.com)

# 综合应用：Web协议请求的历程

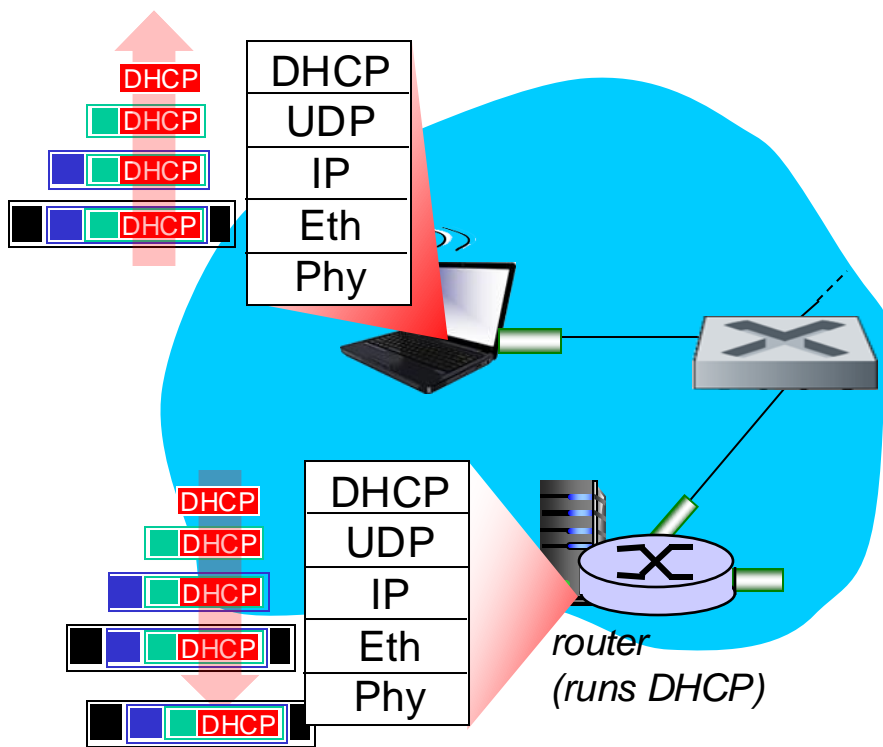
软件学院 · 计算机网络



# 生命中的一天...连接到互联网



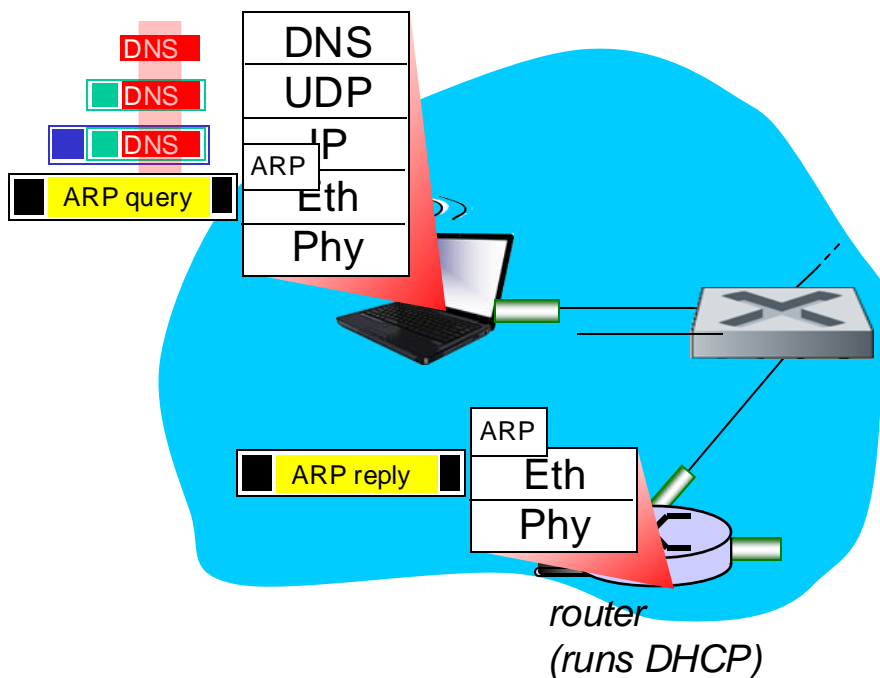
- ❖ 1.使用**DHCP**获取本机IP地址, 网关和DNS服务器地址
- ❖ 2.DHCP请求封装在UDP中, **UDP**封装在IP中, **IP**封装在**802.3**以太网帧中
- ❖ 3.以太网帧以广播方式发送给**DHCP**服务器
- ❖ 4.到达服务器后, 逐层抽取出DHCP请求报文



- ❖ 5.DHCP 服务器返回**DHCP ACK**，携带有客户机分配地址和网关地址及DNS服务器地址
- ❖ 6.服务器完成最终封装，将帧通过局域网发送至客户机，在客户机完成分解
- ❖ 7.客户机接收DHCP ACK后，reply服务器，服务器ack，完成四次握手

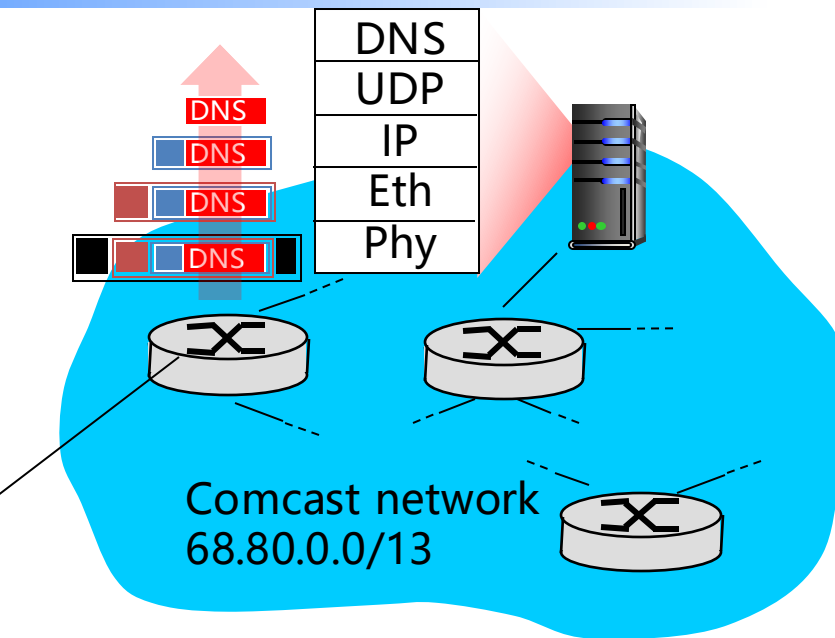
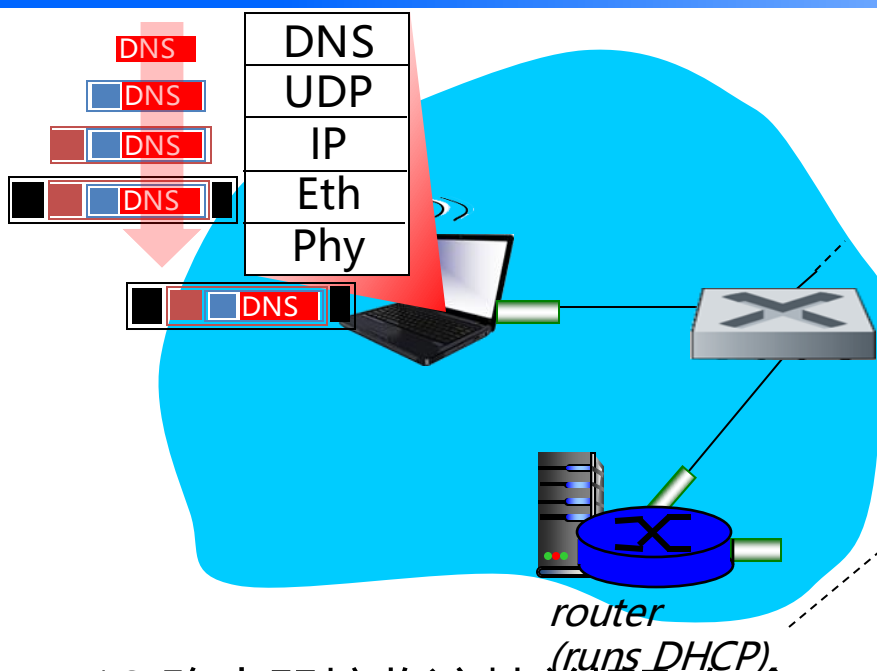
客户机最终获得IP，DNS服务器名字和IP，及网关

# 生命中的一天...ARP(在 DNS 之前, 在 HTTP 之前)



- ❖ 8.在HTTP请求之前, 需要获取 www.google.com的IP地址, 域名-》IP
- ❖ 9.DNS查询请求报文创建, 封装于 UDP, IP, 以太网帧, 之前知道网关 IP地址, 单不知道网关MAC地址, 如果需要向外寻址, 需要ARP
- ❖ 10.ARP查询报文广播, 路由器收到后, 返回ARP reply报文包含有网关的 MAC地址
- ❖ 11.客户机知道网关MAC地址后, 可以发送包含DNS请求的帧

# 生命中的一天...使用域名解析



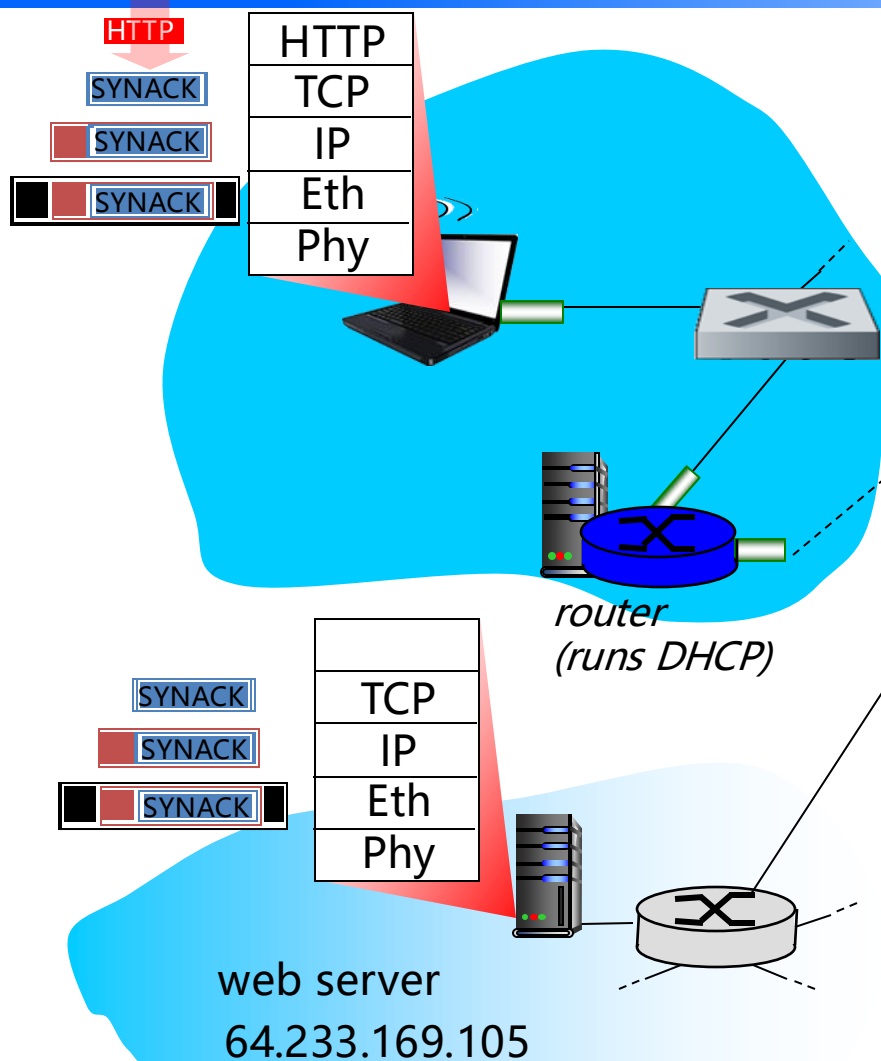
- ❖ 12.路由器接收该帧并提取包含DNS查询的IP数据报
- ❖ 13.查询目的地址，查询路由表确定出接口，向DNS服务器转发数据报，路由表依据是（RIP, OSPF等域内路由协议，BGP域间路由协议）

- ❖ 14.DNS服务器抽取DNS查询报文，经查找后，找到对应IP地址
- ❖ 15.服务器生成回答报文，包含有查询域名的IP，返回给客户机



# 生命中的一天...携带 HTTP 的 TCP 连接

软件学院 · 计算机网络



❖ 16. 客户机发送HTTP请求前，客户端首先建立到web服务器的TCP连接

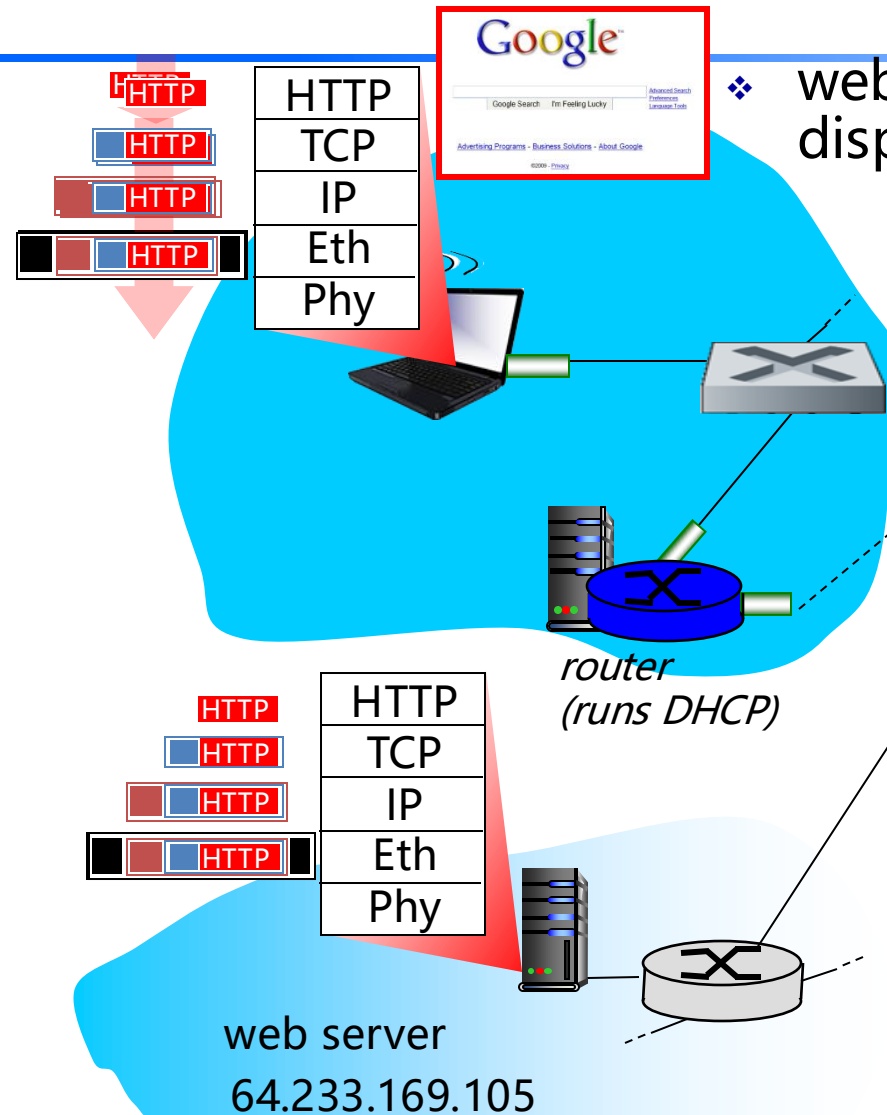
❖ 17. TCP SYN 请求经过路由到达web服务器

❖ 18. Web服务器产生回应，TCP SYN + ack

❖ 19. 客户机接收，返回ack，tcp三次握手完成，连接建立

# 生命中的一天...HTTP请求

软件学院 · 计算机网络



❖ web page *finally (!!!)* displayed

- ❖ 20. HTTP请求通过tcp socket 发送
- ❖ 21. IP 数据报包含有http请求, 最终路由至google服务器
- ❖ 22. 服务器收到request请求后, 产生response-》 http reply
- ❖ 23. Reply封装至IP数据报中, 路由返回至客户机

- 数据链路层服务背后的原则:
  - 错误检测, 纠正
  - 共享广播频道: 多次访问
  - 链路层地址在
- 各种链路层技术的实例化和实现
  - 以太网
  - 交换式局域网、虚拟局域网
  - 作为链路层的虚拟网络: MPLS
  - 综合: Web 请求生命周期中的一天

- 协议栈完成 (PHY 除外)
- 对网络原理和实践的深刻理解
- .....可以停在这里....但有**很多**有趣的话题!
  - 无线电
  - 多媒体
  - 安全
  - 网络管理

- R3, R5, R6, R9, R11, R12
- P5, P14, P15