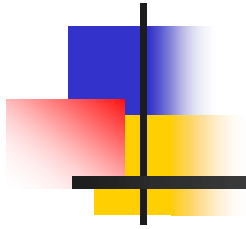


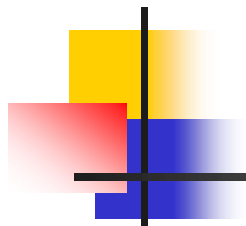
第三章 数据链路层





第三章 数据链路层

- 1、数据链路层的功能
- 2、数据链路和帧
- 3、链路层的三个基本问题
- 4、点对点信道链路层协议 PPP
- 5、广播信道的链路层协议
- 6、局域网
- 7、物理层扩展局域网—集线器
- 8、链路层扩展局域网—网桥、交换机
- 9、虚拟局域网
- 10、100M、1000M、10G以太网

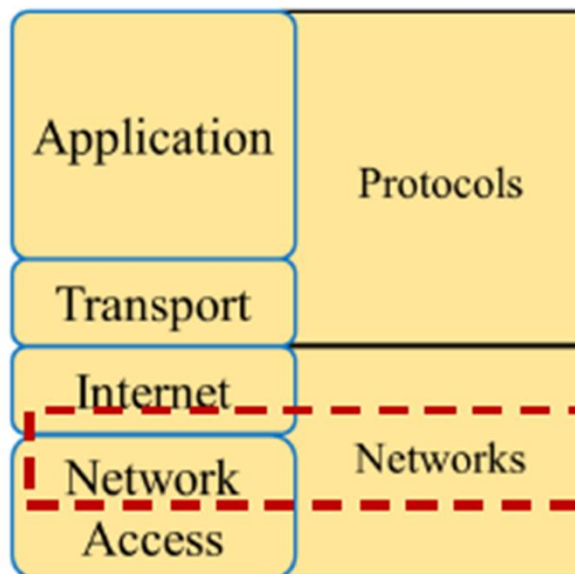


1、数据链路层的功能

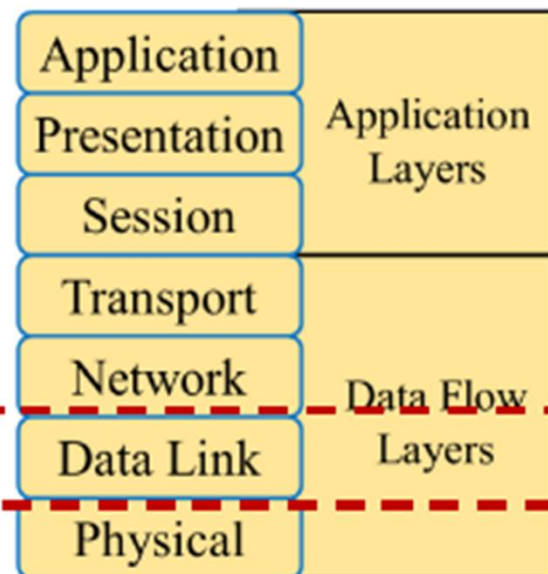
数据链路层有哪些功能？

- 为网络层提供服务，良好的服务接口
- 保证数据传输的有效、可靠：
 - 处理传输错误：差错检测和控制
 - 流量控制

TCP/IP Model

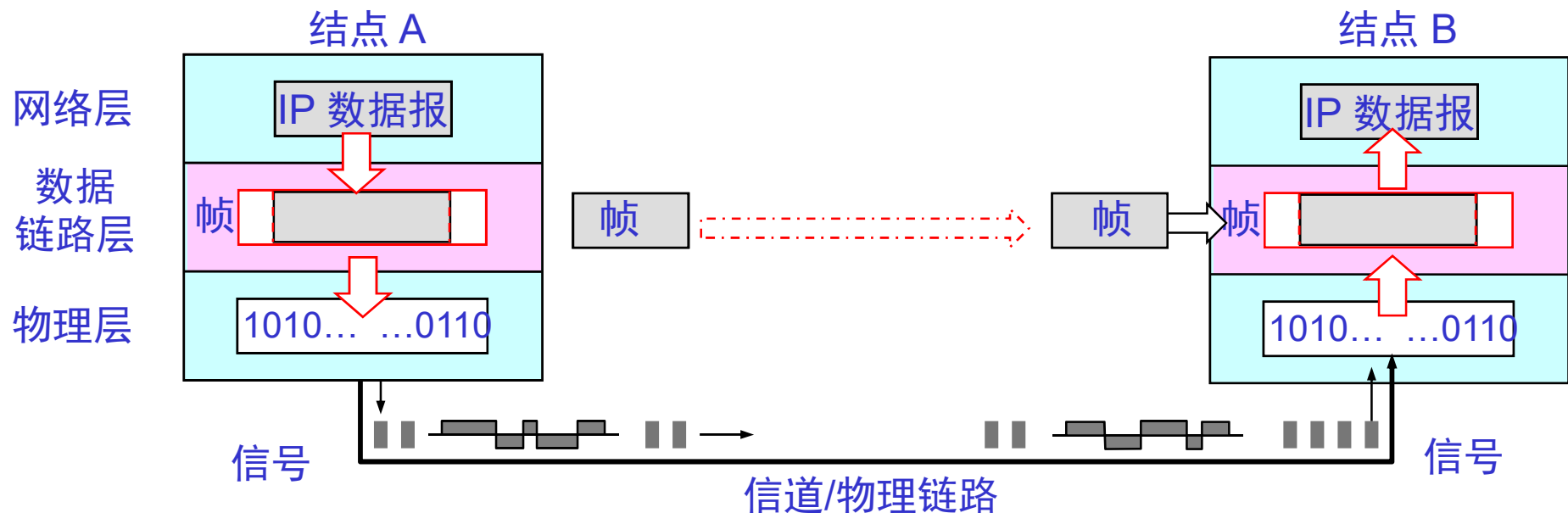


OSI Model



为网络层提供设备之间的数据帧传输

- 封装成帧：
 - 数据块与比特流之间的转换
 - 发送：数据帧 → 比特流；接收：比特流 → 数据帧
- 比特差错控制：差错检测、纠错等
- 流量控制？

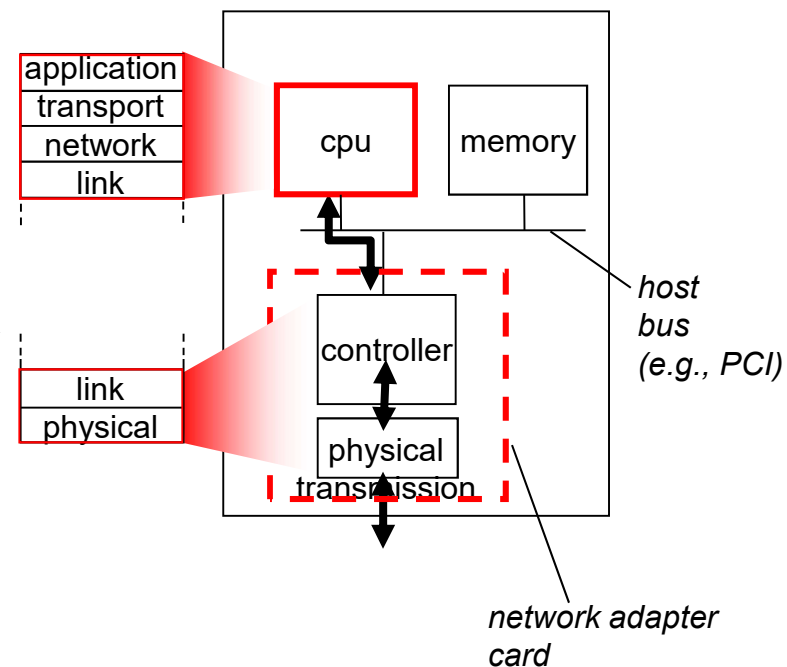


链路层在哪儿实现？

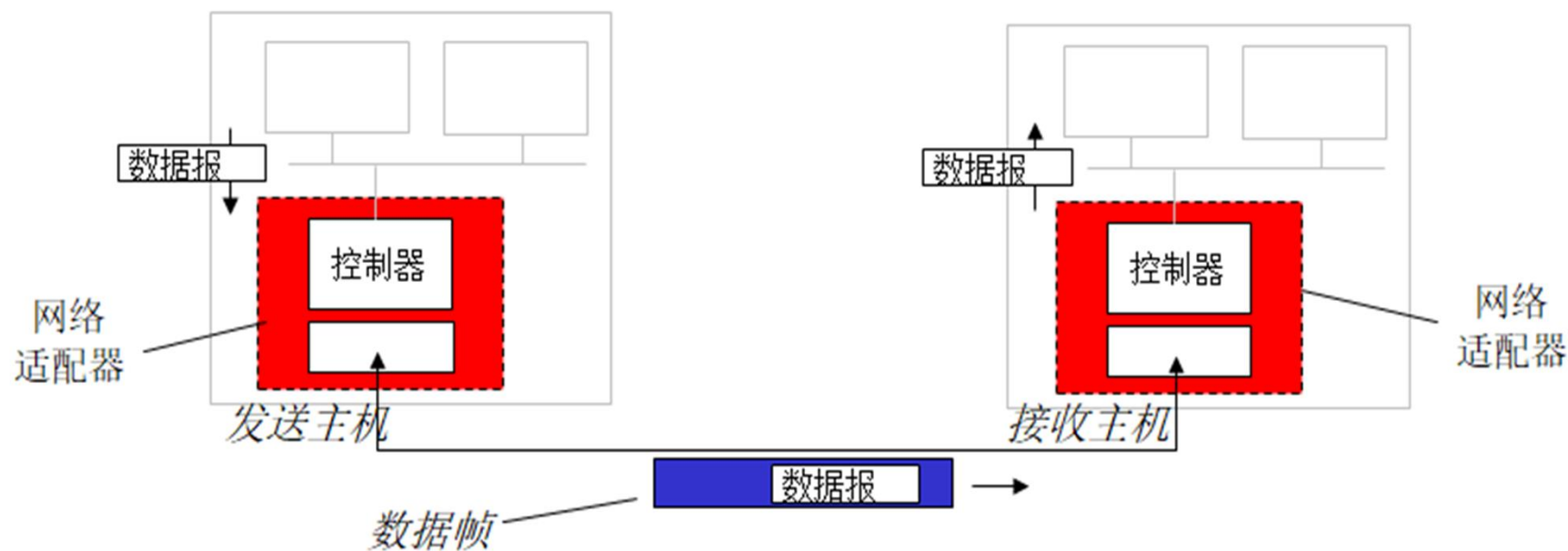


host schematic

- **路由器**：链路层在线卡中实现
- **主机**：链路层主体部分在网络适配器（网卡）中实现
 - 网络适配器连接物理媒体→包括实现物理层功能
- **链路层由硬件和软件实现：**
 - 网卡中的控制器芯片：组帧、链路接入、检错、可靠交付、流量控制；
 - 主机上的链路层软件：与网络层接口，激活控制器硬件、响应控制器中断；



链路层实现网络适配器之间的通信

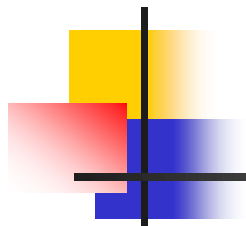


■ 发送方:

- 将数据报封装到帧中
- 生成校验比特
- 执行可靠传输和流量控制

■ 接收方:

- 提取帧，检测传输错误
- 执行可靠传输和流量控制
- 解封装数据报，交给上层协议



2、数据链路和帧

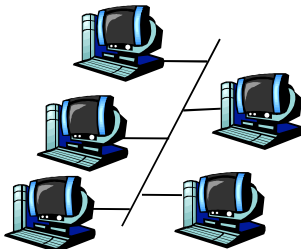


数据链路和帧

- **链路**是一段无源的点到点物理线路，中间没有任何其他的交换结点。
- **数据链路在**物理线路上，必须有通信协议来控制数据的传输，构成数据链路。
 - 使用适配器（即网卡）来实现协议的硬件和软件。
 - 适配器包括数据链路层和物理层功能。

数据链路层有两种信道（链路）

- **点对点信道：**信道使用一对一的点对点通信方式。
 - 仅连接了一个发送方和一个接收方的链路
 - 一条全双工链路可以看成是由两条单工链路组成
- **广播信道：**信道使用一对多的广播通信方式。
 - 连接了许多节点的单一共享信道，任何一个节点发送的数据可被链路上的其它节点接收到。



共享同轴电缆
(早期以太网)



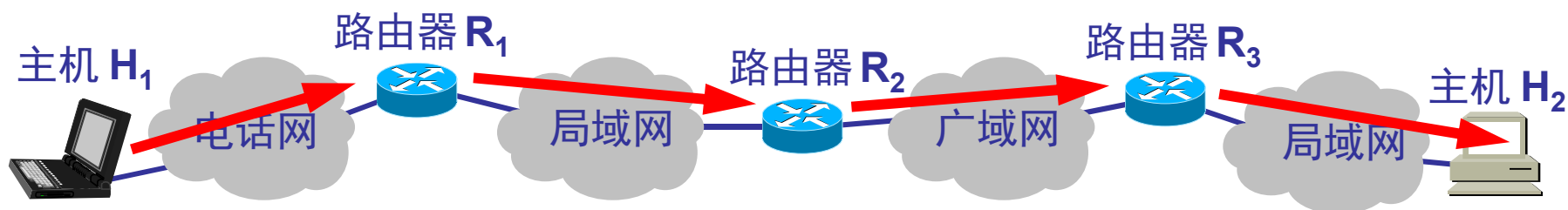
共享无线射频
(802.11 WiFi)



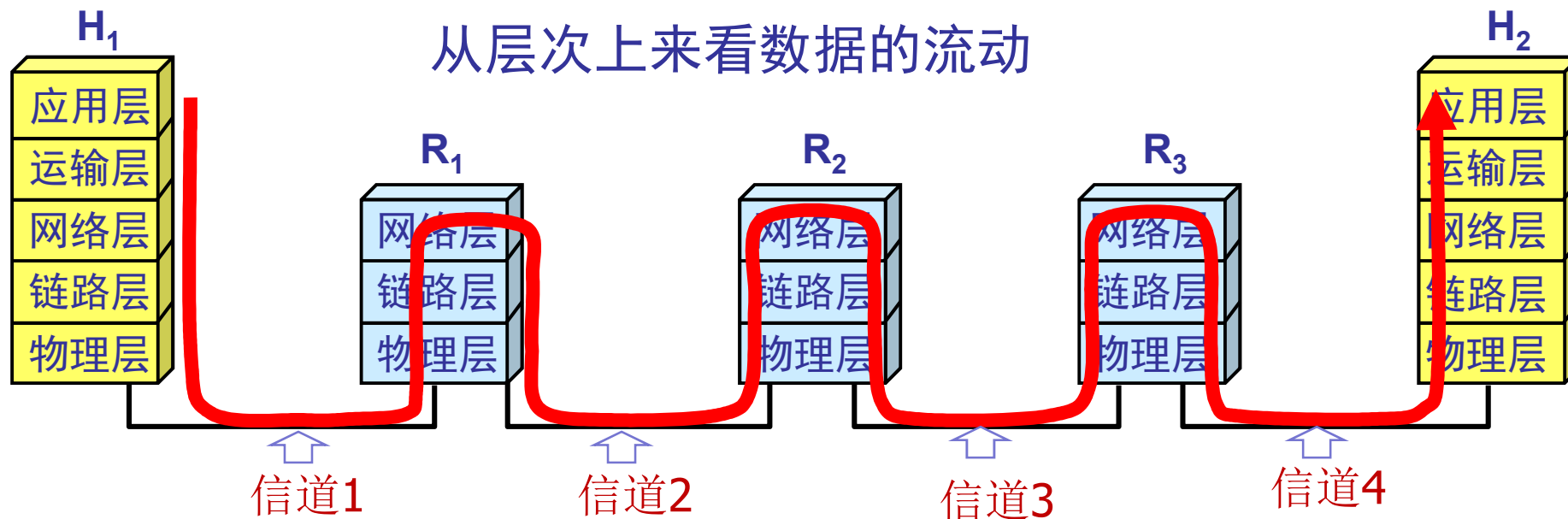
讨论组
(共享声波信道)

数据链路层实际数据传输过程

主机 H_1 向 H_2 发送数据

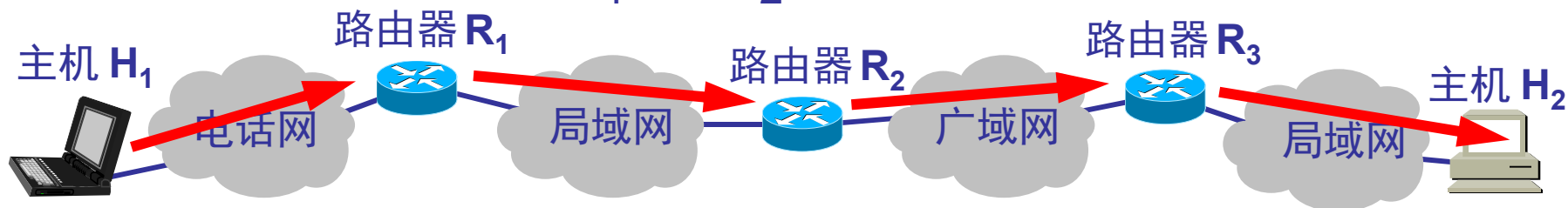


从层次上来看数据的流动

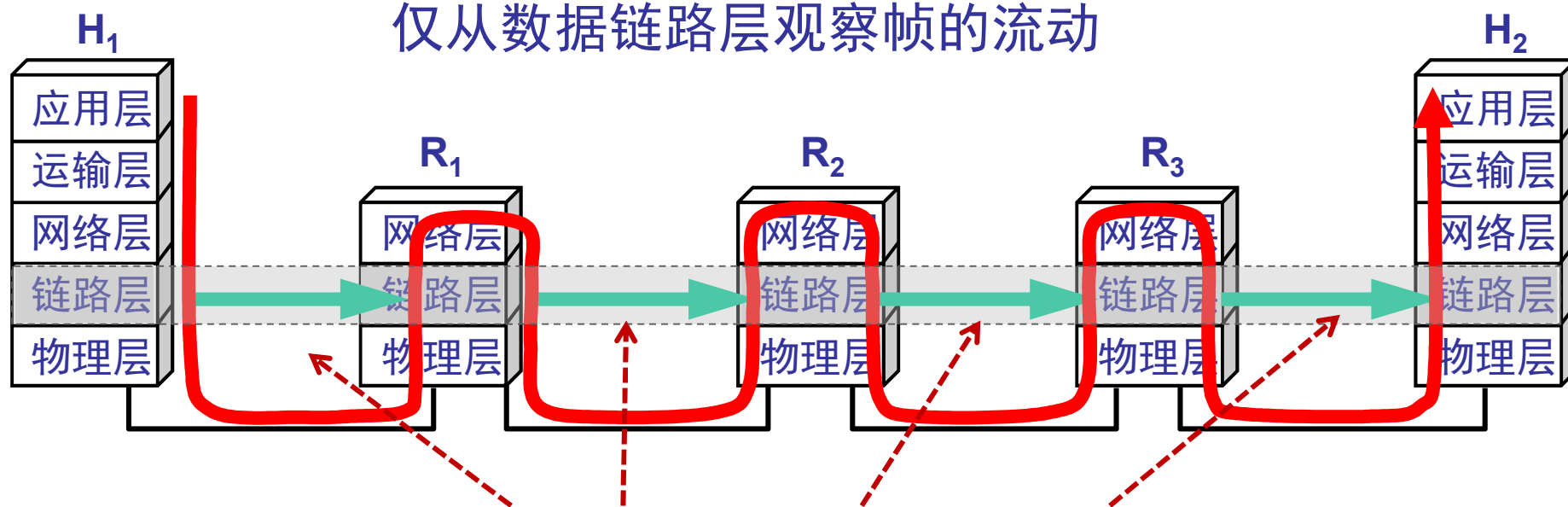


数据链路层的对等实体间的帧传输

主机 H_1 向 H_2 发送数据

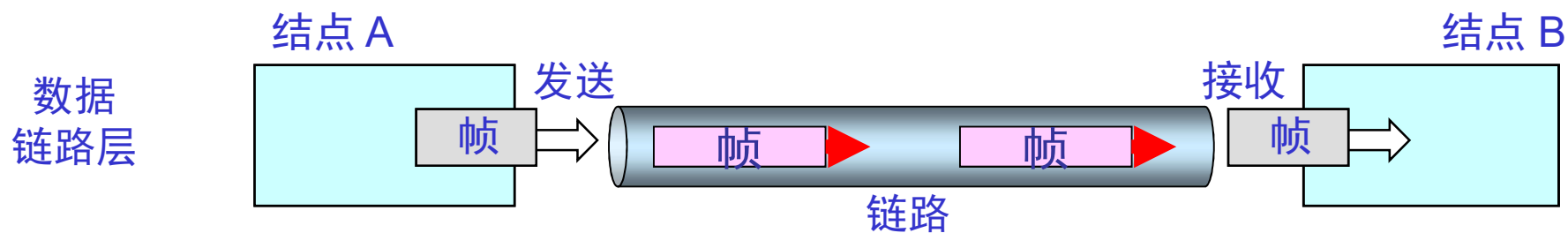
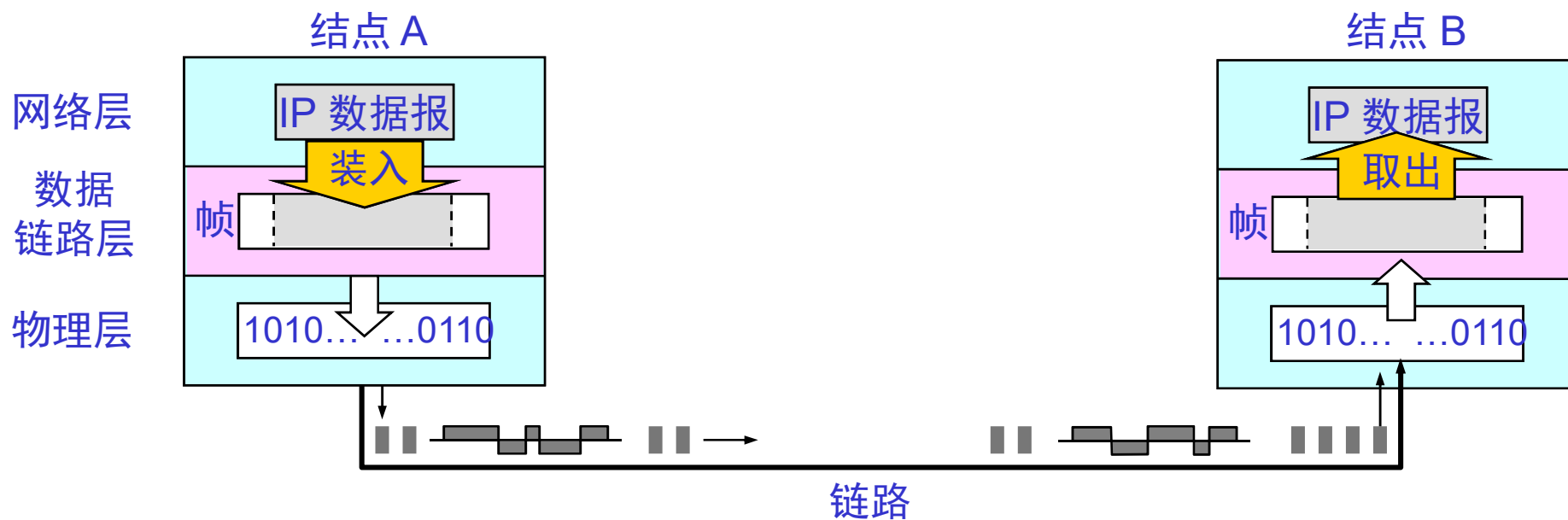


仅从数据链路层观察帧的流动

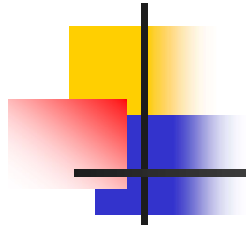


经过四段链路，链路层分段接力传输

数据链路层传送的是帧



数据链路层传送帧的数字管道



3、链路层的三个基本问题

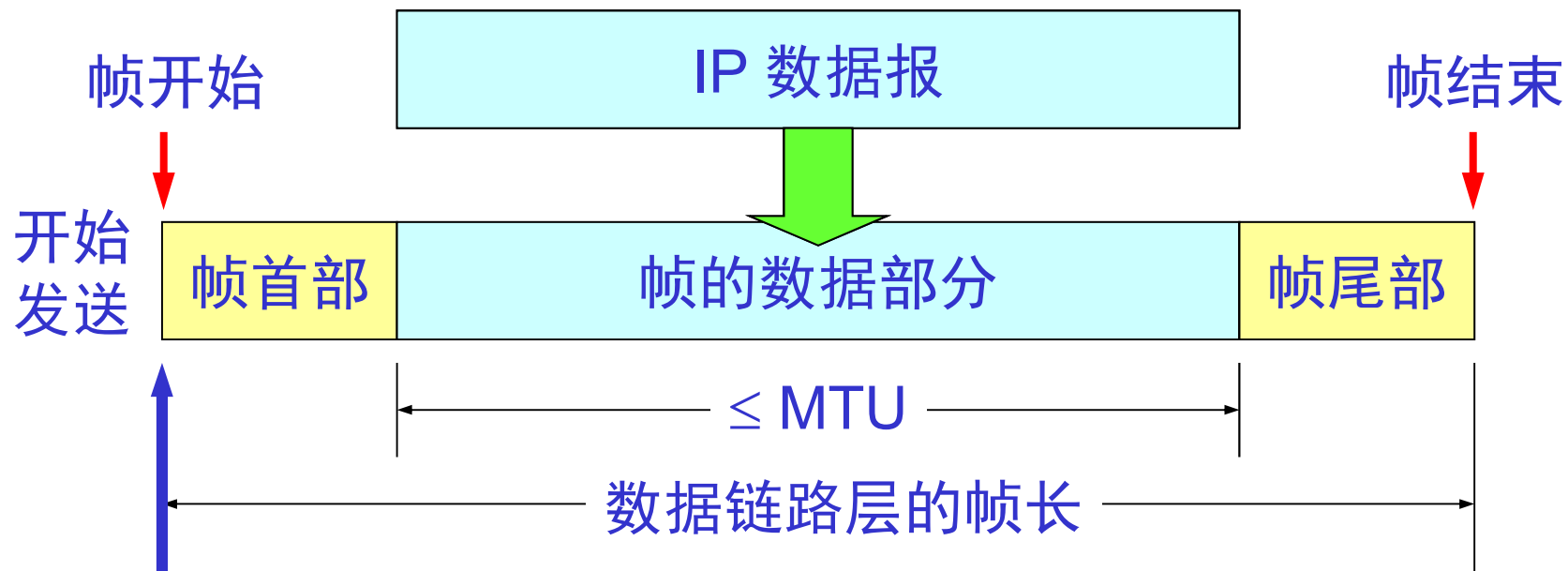


3、数据链路层的基本问题

- (1) 封装成帧
- (2) 透明传输
- (3) 差错控制

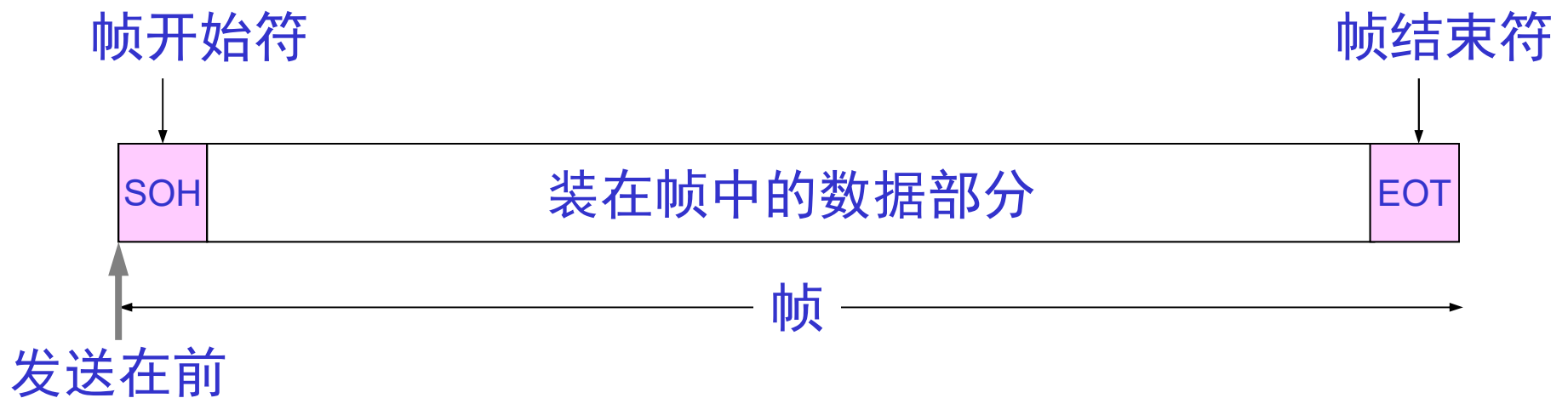
(1) 封装成帧

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

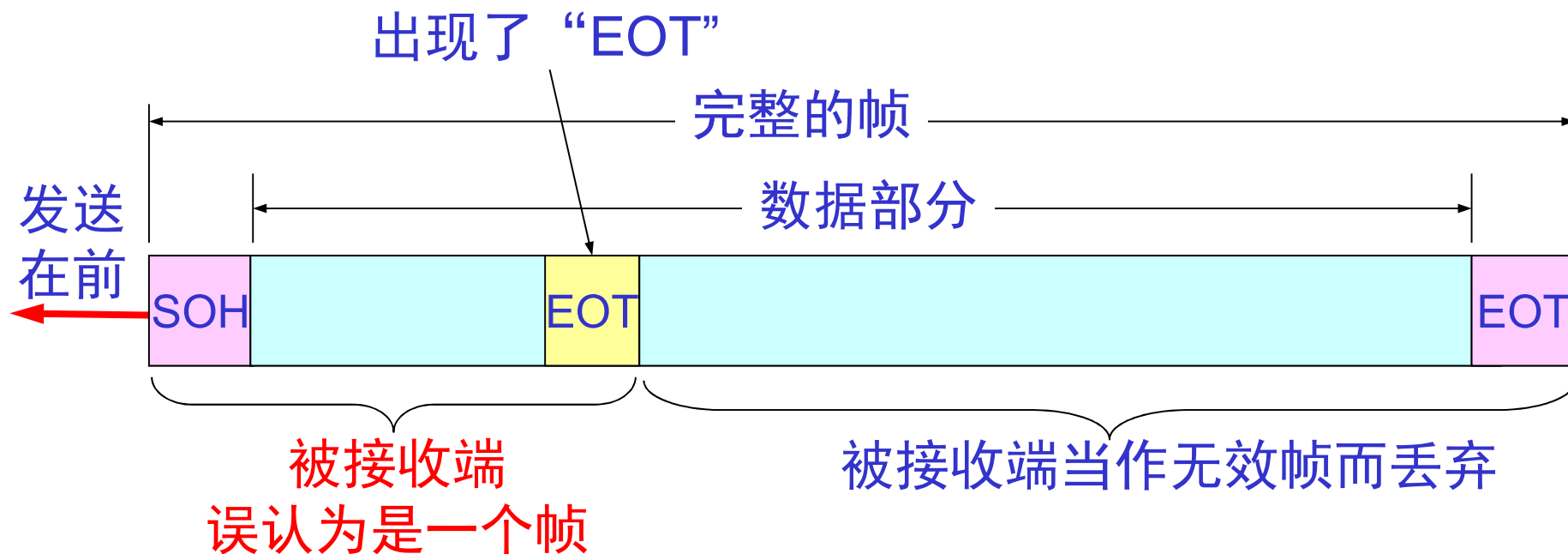




用控制字符进行帧定界的方法

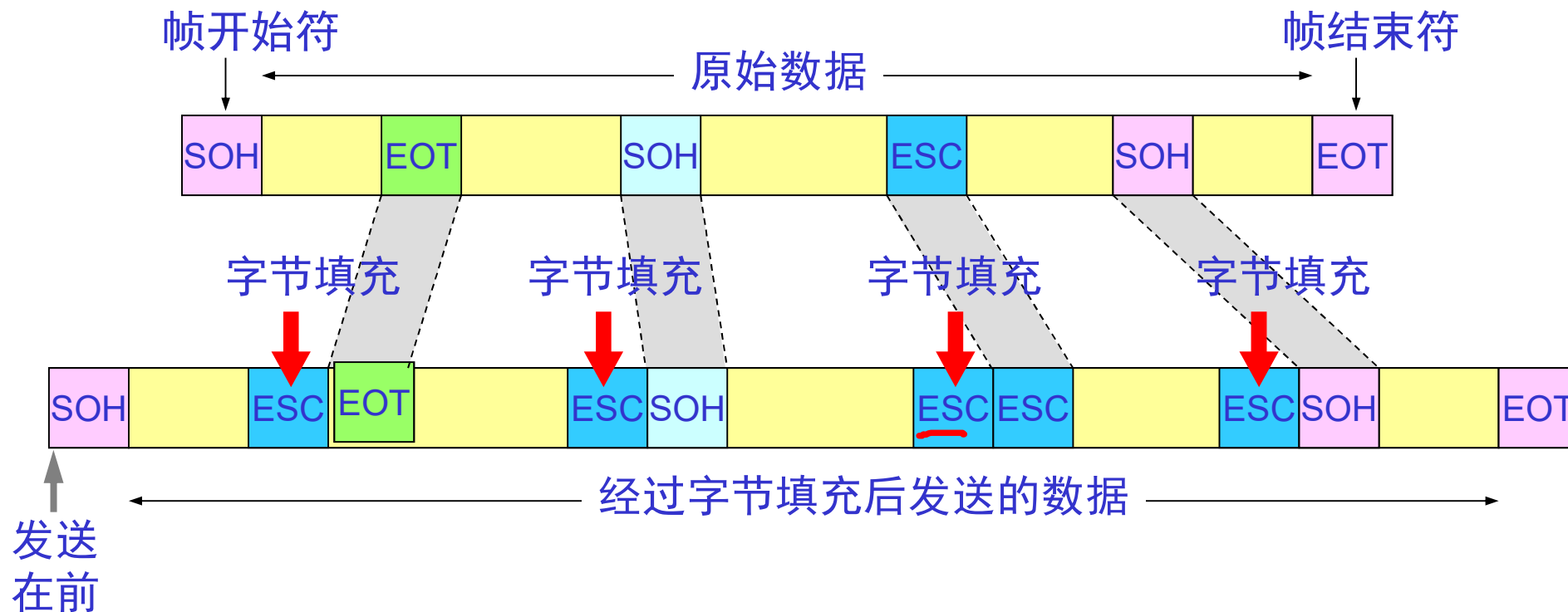


(2) 透明传输



字节填充法

- 发送端：在数据中出现控制字符“SOH”或“EOT”的前面插入一个转义字符“ESC”(其十六进制编码是 1B)。
- 接收端：将数据送往网络层之前删除插入的转义字符。





零比特填充法

- 面向二进制位的帧格式，以特殊位模式01111110（7E）作为帧标志，即一个帧的开始（也标志前一个帧的结束）
- 发送方：当帧内容中出现一个与帧标志相同的位串01111110，则在5个1后插入一个0 → 01111101
- 接收方：将自动删除第5个1后的0。
- 如果由于干扰，一个帧没有正确接收，则可扫描接收串，一旦扫描到01111110，开始新的一帧--再同步。



零比特填充例子

数据中出现了和标志字段 F 完全一样的 8 比特组合

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



会被误认为是标志字段 F

发送端：在 5 个连 1 之后
填入 0 比特再发送出去

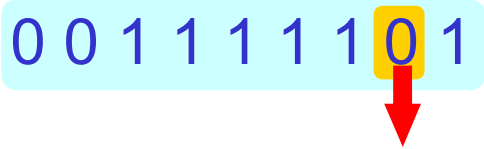
0 1 0 0 1 1 1 1 1 0 1 0 0 0 1 0 1 0



发送端填入 0 比特

接收端：把 5 个连 1 之
后的 0 比特删除

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



接收端删除填入的 0 比特

字符计数法





物理层编码违例法

- 在曼切斯特编码中，连续高电平或连续低电平可用作帧边界
 - 采用冗余编码技术，如曼切斯特编码，即两个脉冲宽来表示一个二进制位
 - 数据0：低-高电平对
 - 数据1：高-低电平对
 - 高-高电平对和低-低电平对没有使用，可用作帧边界



(3) 差错检测和纠错

■ 传输出错的类型

- 单个错：由随机的信道热噪声引起，一次只影响1位
- 突发错：由瞬间的脉冲噪声引起，一次影响许多位，使用突发长度表示突发错影响的最大数据位数

■ 差错控制编码的类型

- 检错码：只能检测出传输错误的编码，不能确定出错位置，通常与反馈重传机制结合进行差错恢复
- 纠错码：能够确定错误位置并自行纠正的编码



差错的处理

- 纠错码——前向纠错技术：发现错误，从错误中恢复出正确的来。
 - 因其需要太多的冗余位，纠错开销太大，在有线网络中极少使用，主要用于无线网络中(Why?)。
- 检错码：只能发现错误，不能从错误中恢复，但可采用重传)
 - 计算机网络中主要采用循环冗余码(CRC)。
- 两种不同的处理方法适用于不同的环境



如何检测与纠正错误？

- **码字** (codeword)：由 m 比特的数据加上 r 比特的冗余位（校验位）构成
- 有效编码集：由 2^m 个符合编码规则的码字组成
- **检错**：若收到的码字为无效码字，判定出现传输错误
- 海明距离 (Hamming Distance)：两个码字的对应位取值不同的位数
- **纠错**：将收到的无效码字纠正到距其最近的有效码字
- 检错码与纠错码的能力都是有限的！



编码集的检错与纠错能力

- 编码集的海明距离：编码集中任意两个有效码字的海明距离的最小值
- 检错能力：为检测出所有 d 比特错误，编码集的海明距离至少应为 $d+1$
- 纠错能力：为纠正所有 d 比特错误，编码集的海明距离至少应为 $2d+1$

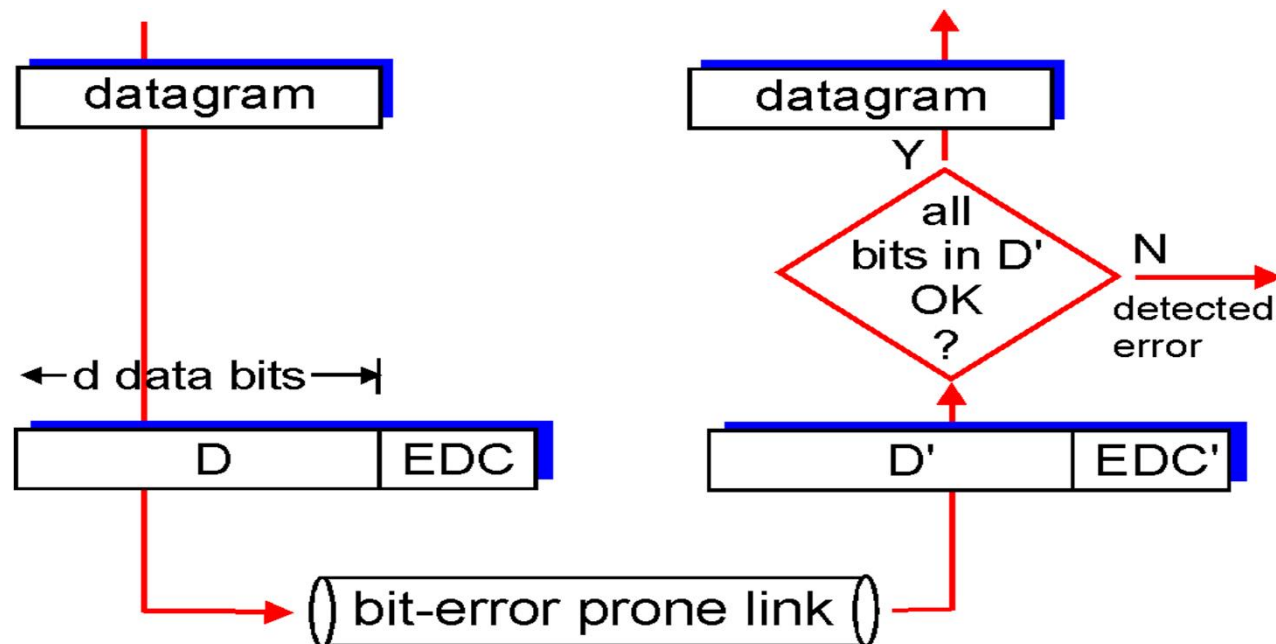


海明距离 (Hamming Distance)

- 码字：包含数据位和校验位的n位单元。
- 海明距离
 - 两个码字(codeword)的海明距离：两个码字之间不同位的数目。
如：10001001 和10110001 的海明距离为3。
 - 异或的结果中，1的个数

差错检测过程

- 发送端对要保护的数据D（包括帧头字段）生成校验位EDC，添加在帧头中
- 接收端对收到的数据D'计算校验位EDC'，根据EDC'判定是否有错





检错码

- 奇偶校验
- CRC循环冗余码
- 互联网校验和

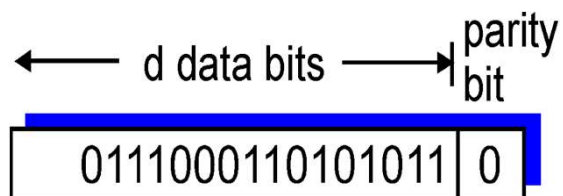
奇偶校验

单比特奇偶校验:

可检测单比特错误

检错率为50%

编码集海明距离为2

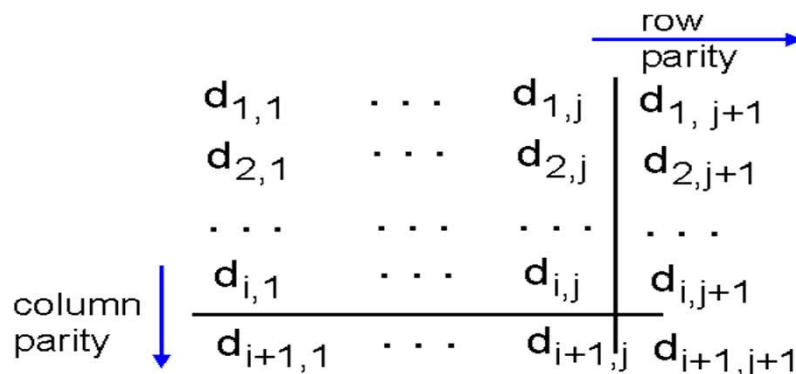


二维奇偶校验:

可检测2比特错和纠正单比特错

编码集海明距离为3

有利于检测突发错误



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

parity error

correctable
single bit error



奇偶校验

■ 查出偶数个错误的简单方法

- 例如: 1 error, 11100101; 5个1, 奇数个, 检出错
- 例如: 3 errors, 11011001; 5个1, 奇数个, 检出错
- 例如: 2 errors, 11101101; 6个1, 偶数个, 不能检出错误, 判定为正确。
- 出错误的概率为 $\frac{1}{2}$



循环冗余校验 (CRC)

- CRC是一种多项式编码，它将一个位串看成是某个一元多项式的系数，如1011看成是一元多项式 $X^3 + X + 1$ 的系数
- 信息多项式 $M(x)$ ：由m个信息比特为系数构成的多项式
- 冗余多项式 $R(x)$ ：由r个冗余比特为系数构成的多项式
- 码多项式 $T(x)$ ：在m个信息比特后加上r个冗余比特构成的码字所对应的多项式，表达式为 $T(x) = x^r \cdot M(x) + R(x)$
- 生成多项式 $G(x)$ ：双方确定用来计算 $R(x)$ 的一个多项式
- 编码方法： $R(x) = x^r \cdot M(x) \div G(x)$ 的余式（减法运算定义为异或操作）
- 检验方法：若 $T(x) \div G(x)$ 的余式为0，判定传输正确
- CRC码检错能力极强，可用硬件实现，是应用最广泛的检错码

CRC举例

模2运算：在二进制运算中，减法和加法都做异或运算，即相同得0，相异得1，比如：0+1=1，1+1=0，0-1=1，1-0=1

例1：取 $G(X) = X^3 + 1$ ，对信息比特101110计算CRC码。

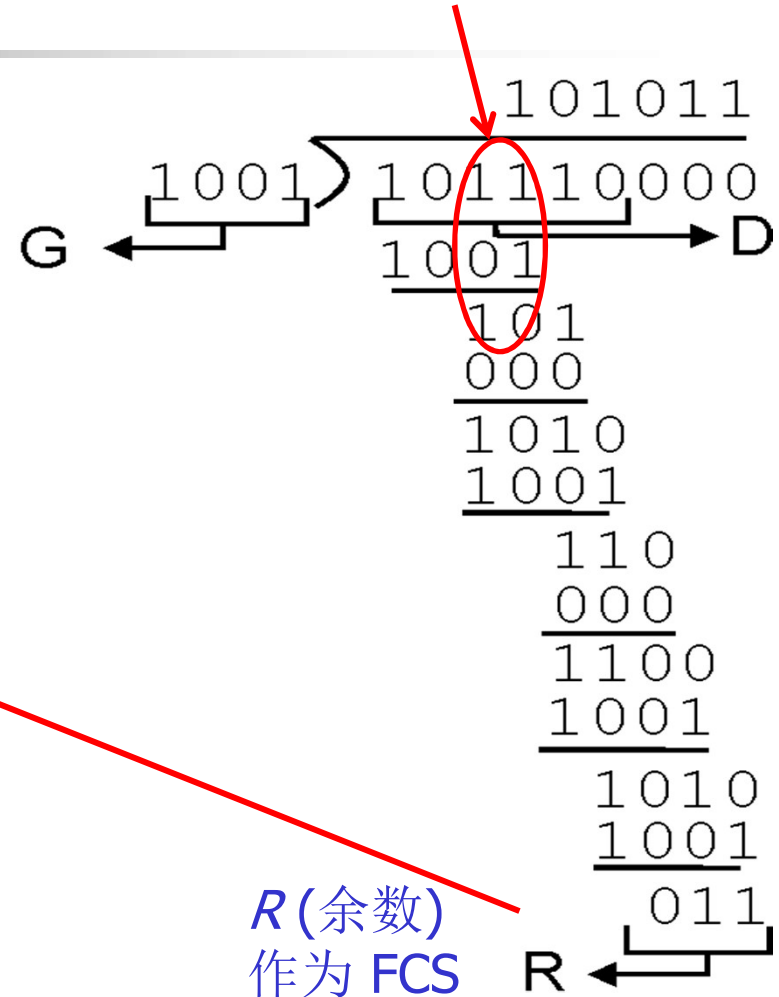
解答：

- 101110000 \div 1001的余式为R=011 (CRC code)
- 码字：101110011

例2：取 $G(X) = X^3 + 1$ ，接收端收到比特串1001001，问是否有错？

解答：

- 1001001 \div 1001的余式为001（不为0），有传输错误。



什么是模2运算？

模2加 以及 **模2减** 等同于异或运算，即相同得0，相异得1。

$$0 \oplus 0 = 0; 0 \oplus 1 = 1;$$

$$1 \oplus 0 = 1; 1 \oplus 1 = 0.$$

相同得0，不同得1

– **Modulo 2 multiplication:**

$$\begin{array}{r} 1010 \\ \times \quad 101 \\ \hline 1010 \\ 0000 \\ 1010 \\ \hline 100010 \end{array}$$

-- **Modulo 2 division:**

$$\begin{array}{r} 101 \overline{) 10000} \\ \underline{101} \\ 010 \\ \underline{000} \\ 100 \\ \underline{101} \\ 01 \end{array}$$



接收端对收到的每一帧进行 CRC 检验

- (1) 若得出的余数 $R = 0$ ，则判定这个帧没有差错，就**接受**(accept)。
- (2) 若余数 $R \neq 0$ ，则判定这个帧有差错，就**丢弃**。
- 但这种检测方法并不能确定究竟是哪一个或哪几个比特出现了差错。
- 只要经过严格的挑选，并使用位数足够多的除数 P ，那么出现检测不到的差错的概率就很小很小。

例：CRC码计算

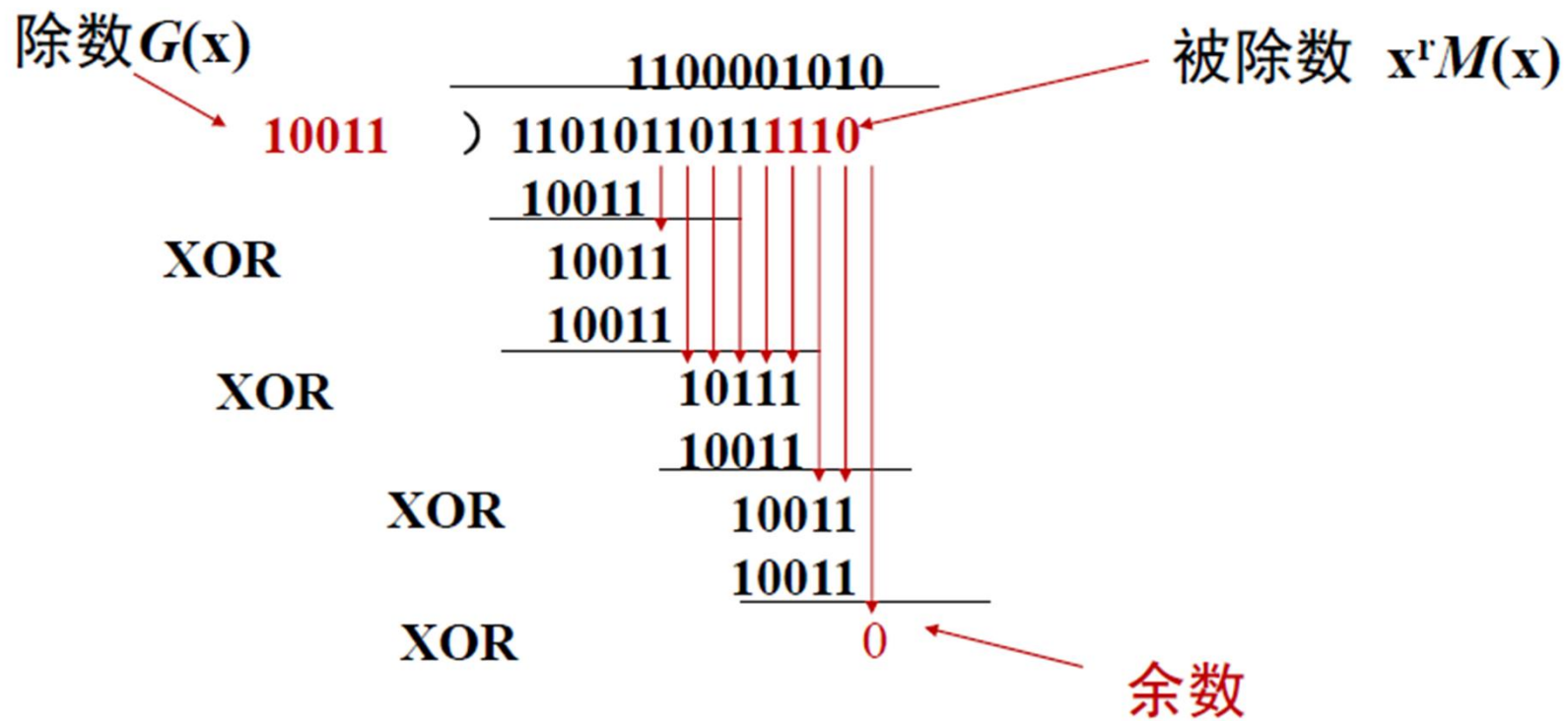
计算11010110110000/10011 得余数1110， $11010110110000-1110=$
 11010110111110 ，所以：

编码后得CRC码为：11010110111110

当这个码字达到接收方时：

- 如CRC码在接收端能被10011整除则说明接收正确。
- 如发送方发送的 $T(x)$ ，接收方收到的是 $T(x)+E(x)$ ，如果不能被整除，则被检测到已出错。

接收端的检查



$R(\text{余数})=0$

接收数据检测正确



生成多项式国际标准

□ CRC-12: $x^{12} + x^{11} + x^3 + x^2 + x^1 + 1$

用于字符长度为6位

□ CRC-16 : $x^{16} + x^{15} + x^2 + 1$

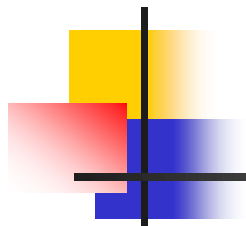
用于字符长度为8位

□ CRC-CCITT : $x^{16} + x^{12} + x^5 + 1$

用于字符长度为8

□ CRC32:

$$G(x) = x^{32} + x^{26} + x^{23} + x^{22} + x^{16} + x^{12} + x^{11} + x^{10} + x^8 + x^7 + x^5 + x^4 + x^3 + x + 1$$



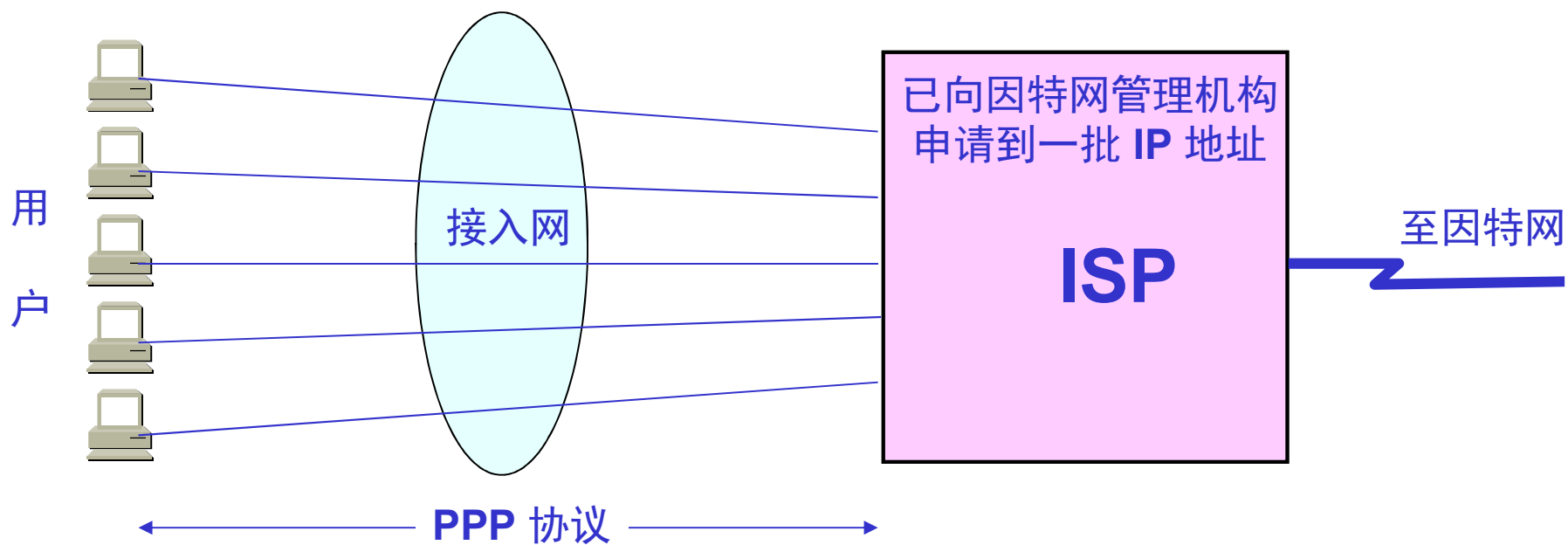
4、点对点信道链路层协议 PPP



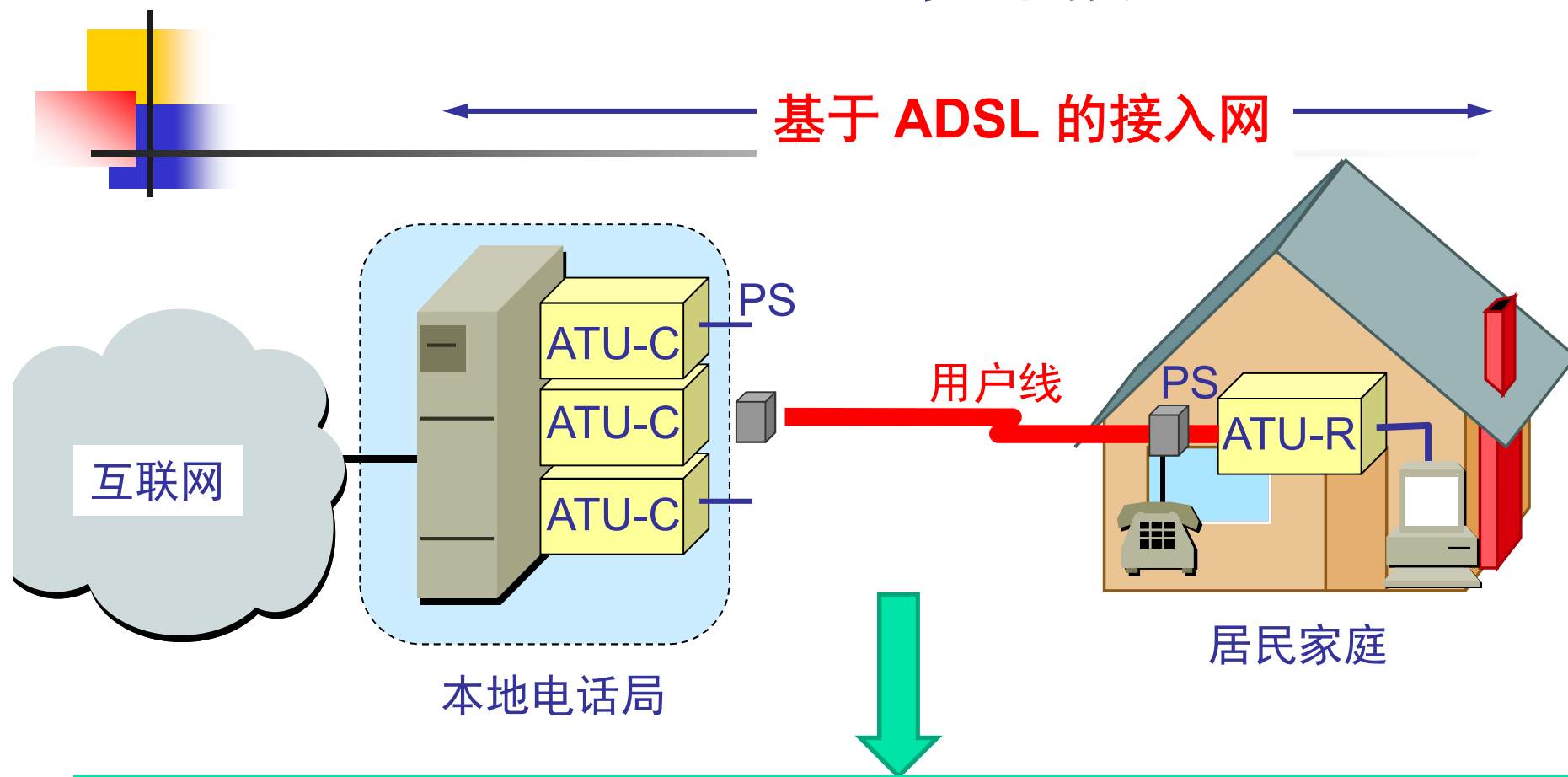
点对点协议 PPP

- 现在全世界使用得最多的数据链路层协议是**点对点协议** PPP (Point-to-Point Protocol)。
- 用户使用拨号电话线接入因特网时，一般都是使用 PPP 协议。

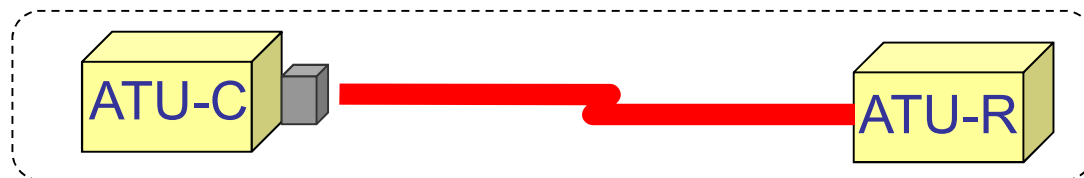
用户到 ISP 的链路使用 PPP 协议



PPP 链路——ADSL 宽带接入



PPP 链路





PPP 协议—基本设计要求

- 封装成帧
- 透明性
- 差错检测
- 最大传送单元

- 支持多种类型链路
- 检测连接状态

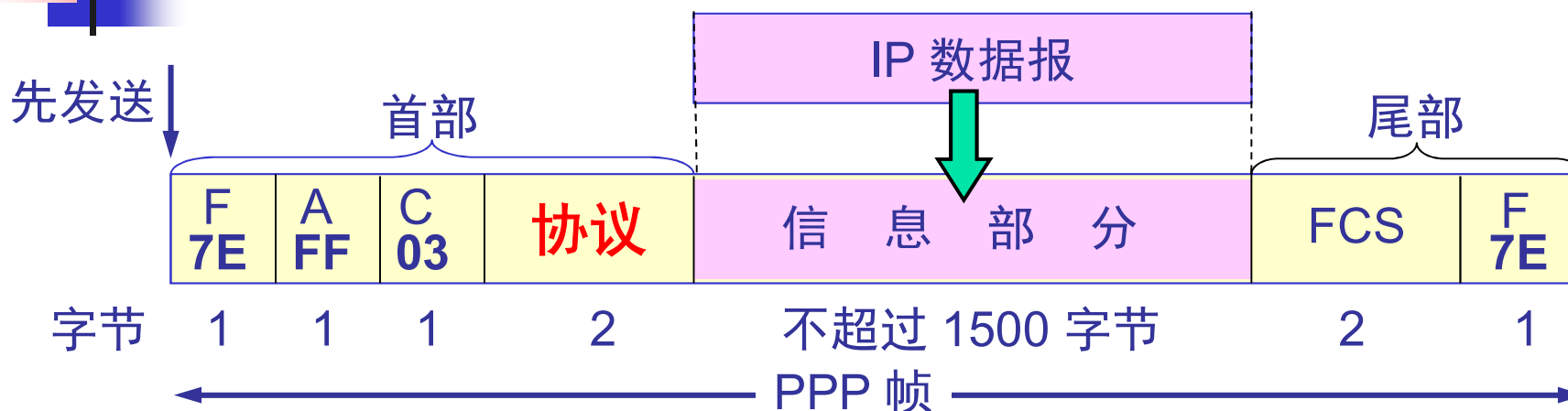
- 支持多种网络层协议
- 网络层地址协商



PPP 协议组成

- PPP 协议--因特网的正式标准[RFC 1661]。
- PPP 协议有三个组成部分
 - 一个将 IP 数据报封装到串行链路的方法。
 - 链路控制协议 LCP (Link Control Protocol)。
 - 网络控制协议 NCP (Network Control Protocol)。

PPP 协议的帧格式

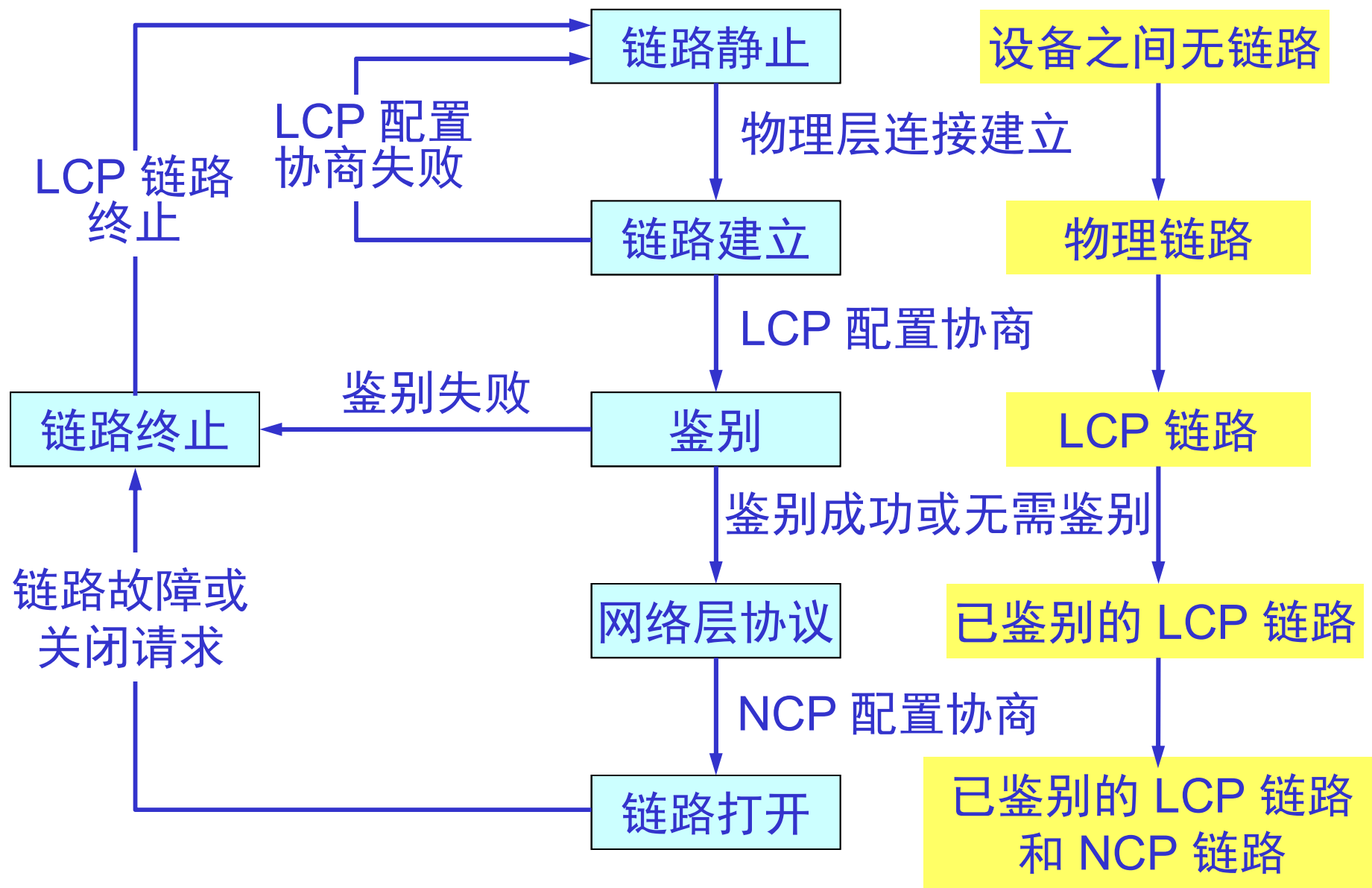


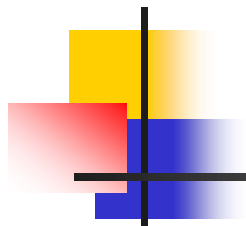
- PPP 有一个 2 个字节的协议字段。
 - 当协议字段为 0x0021 时，PPP 帧的信息字段就是 IP 数据报。
 - 若为 0xC021，则信息字段是 PPP 链路控制数据。
 - 若为 0x8021，则表示这是网络控制数据。



PPP 协议的工作状态

- 当用户拨号接入 ISP 时，路由器的调制解调器对拨号做出确认，并建立一条物理连接。
- PC 机向路由器发送一系列的 LCP 分组（封装成多个 PPP 帧）。
- 这些分组及其响应选择一些 PPP 参数，并进行网络层配置，NCP 给新接入的 PC 机分配一个临时的 IP 地址，使 PC 机成为因特网上的一个主机。
- 通信完毕时，NCP 释放网络层连接，收回原来分配出去的 IP 地址。接着，LCP 释放数据链路层连接。最后释放的是物理层的连接。





5、广播信道的链路层协议



使用广播信道的数据链路层

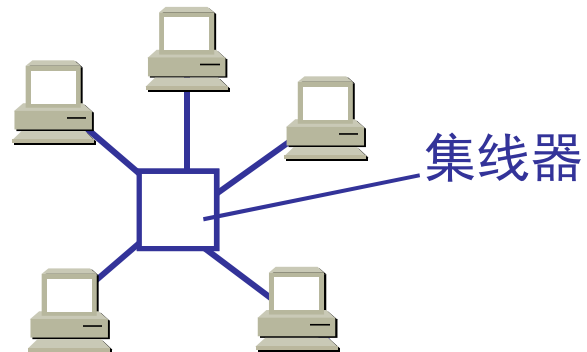
- 广播信道面临的问题

- 共享信道/多路访问信道/广播信道
- 可能两个（或更多）站点同时请求占用信道

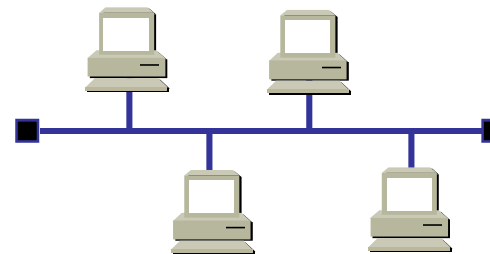
- 解决办法：介质的多路访问控制

- 在多路访问信道上确定下一个使用者

局域网多种共享信道拓扑

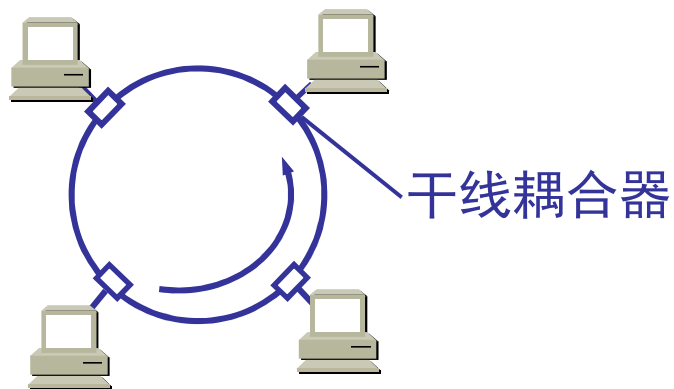


星形网

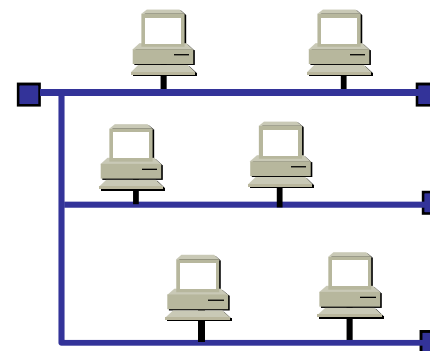


总线网

匹配电阻



环形网



树形网



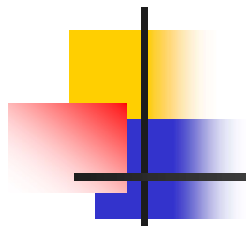
怎样分配信道（介质访问控制）？

■ 静态分配

- 只有一个站/用户使用信道
- 不用的就浪费了
 - 频分多路复用**FDM**，时分多路复用**TDM**

■ 动态分配

- 信道是开放的
- 没有预分配
- ？ ？

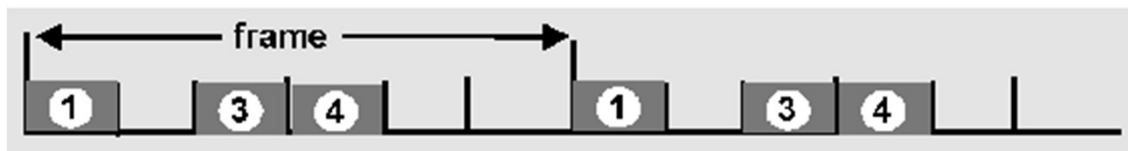


■ 信道静态划分协议

信道划分协议: TDMA

TDMA: 时分多址访问

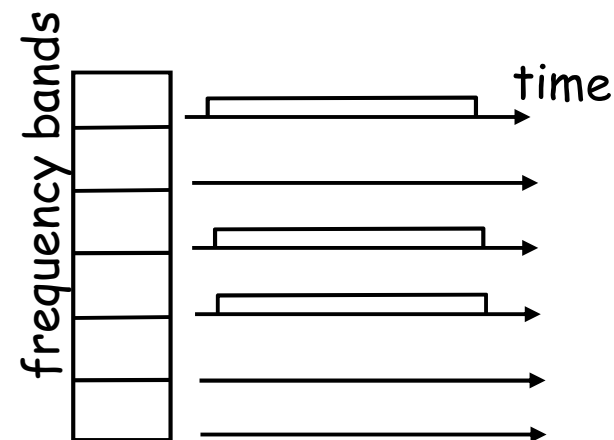
- TDM 将时间划分成时间帧, 然后每个帧又进一步划分成N个时隙, 然后将每个时隙分给N个节点中的一个
- 轮流访问信道
- 每个节点得到合适长度的时隙 (长度 = 分组传输时间), 没有用到的时隙闲置
- TDM 消除了碰撞, 而且非常公平 但有两个大的缺点
 - 每个节点的最大吞吐率为 R/N b/s
 - 每个节点受限于平均速度 R/N
 - 每个节点必须在自己所分配的时间槽内发送数据
- 举例: 6-station LAN, 1,3,4 有包需要发送, 时隙2,5,6 闲置

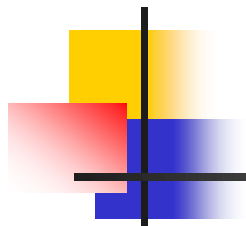


信道划分协议: FDMA

FDMA: 频分多路访问

- 信道带宽被划分成N个频带
- 每个节点分配合适的频率带宽
- 频带在没有用到的时间内闲置
- 例子: 6-station LAN, 1,3,4 有包需要发送, 频带2,5,6闲置
- 优缺点
 - 避免碰撞
 - 公平共享信道
 - 一个节点能够获得的带宽被限制在了 R/N
 - 节点不需要轮流传输数据





- 广播信道多址访问协议
 - Multiple access protocol



多址访问协议 (Multiple access protocol)

■ 随机访问协议 (Random Access)

- 特点：站点争用信道，可能出现站点之间的冲突
- 典型的随机访问协议
 - 纯ALOHA协议、时隙ALOHA协议
 - CSMA协议、CSMA/CD协议（以太网采用此协议）

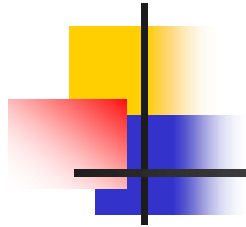
■ 受控访问协议 (Controlled Access)

- 特点：有数据的节点轮流发送，**不会出现冲突**
- 轮流使用信道：Token Ring



随机接入MAC协议

- 随机接入的基本思想：
 - 当节点有数据要发送时，以信道速率 R 发送，发送前不需要协调
 - 随机接入MAC协议规定如何检测冲突，以及如何从冲突中恢复
- 随机接入MAC协议的例子：
 - 发送前不监听信道：ALOHA家族
 - 发送前监听信道：CSMA家族



■ 随机访问协议--ALOHA协议



时隙 (Slotted) ALOHA

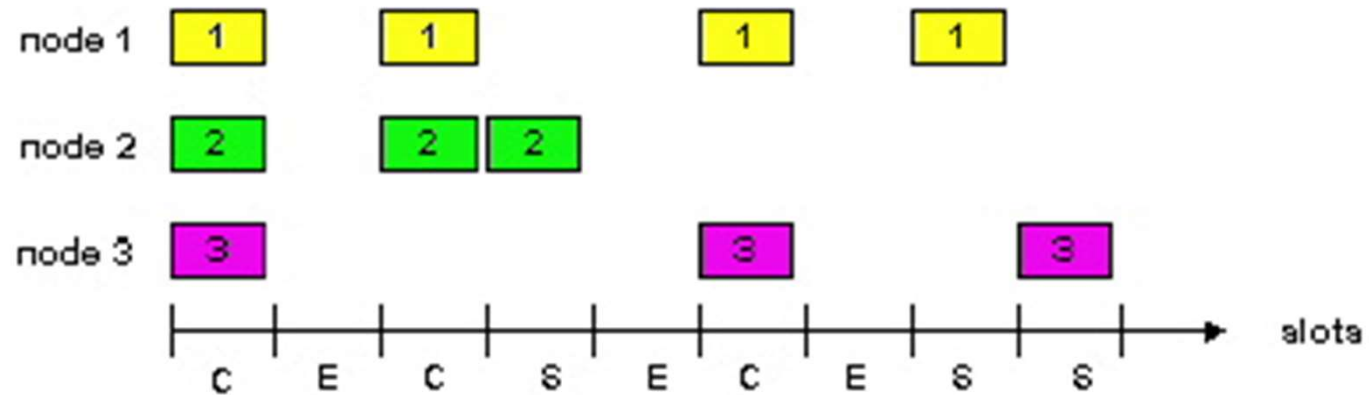
假设:

- 所有帧长度相同
- 时间被划分为等长的时隙，每个时隙传一帧
- 节点只能在时隙开始时发送
- 节点是时钟同步的（知道时隙何时开始）
- 所有节点可在时隙结束前检测到是否有冲突发生

操作:

- 节点从上层收到数据后，在下一个时隙发送
- 若时隙结束前未检测到冲突，节点可在下一个时隙发送新的帧
- 若检测到冲突，节点在随后的每一个时隙中以概率 P 重传，直至发送成功

时隙ALOHA



优点

- 单个活跃节点可以信道速率连续发送
- 分布式：节点自行决定什么时候发送
- 简单

缺点

- 发生冲突的时隙被浪费了
- 由于概率重传，有些时隙被闲置
- 需要时钟同步



时隙Aloha的效率

效率：当网络中存在大量活跃节点时，长期运行过程中成功时隙所占的比例

- 假设: 有 N 个活跃节点，每个节点在每个时隙开始时以概率 P 发送
- 给定节点在一个时隙中发送成功的概率 = $p(1-p)^{N-1}$
- 给定时隙中有节点发送成功的概率 = $Np(1-p)^{N-1}$

■ 最大效率:

- 找到令 $Np(1-p)^{N-1}$ 最大的概率 p^*
- 代入 $Np^*(1-p^*)^{N-1}$ ，并令 N 趋向于无穷，得到：
- 最大效率 = $1/e = 0.37$

最佳情况：信道用于有效传输的时间仅为**37%!**

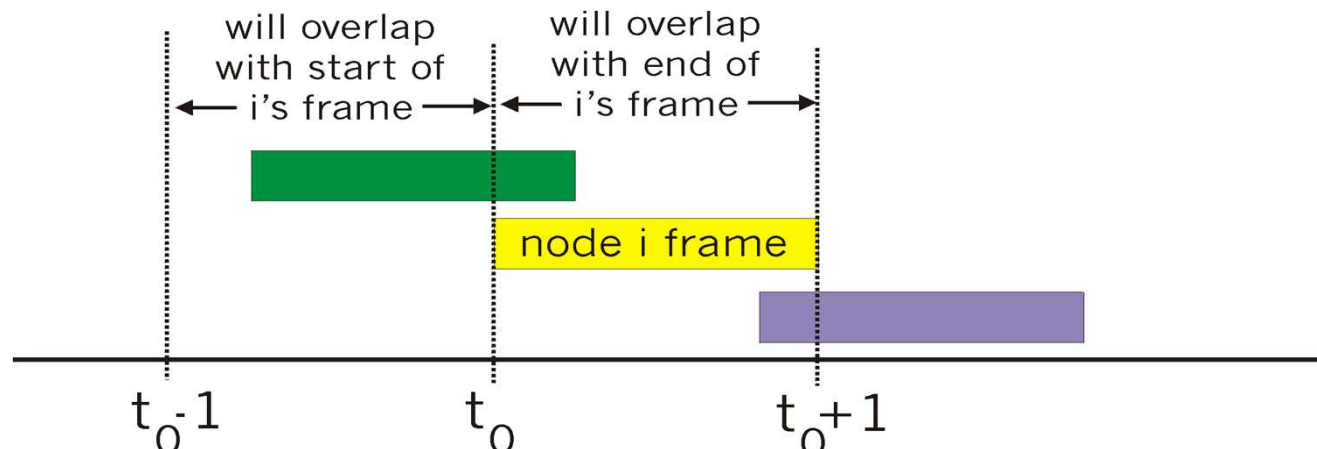
纯ALOHA

■ 基本思想:

- 取消同步时钟，任何节点有数据发送就可以立即发送
- 节点通过监听信道判断本次传输是否成功
- 若不成功，立即以概率 P 重传，以概率 $(1-P)$ 等待一个帧时后再决定。（帧时：发送一帧的时间，假设帧长度相同）

■ 发生冲突的情形:

- 在时刻 t_0 发送的帧与在 $[t_0-1, t_0+1]$ 时段内发送的其它帧冲突





纯Aloha的效率

$P(\text{给定节点发送成功}) = P(\text{节点发送}) \cdot$

$P(\text{无其它节点在}[t_0-1, t_0]\text{内发送}) \cdot$

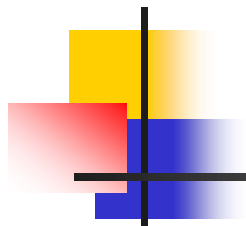
$P(\text{无其它节点在}[t_0, t_{0+1}]\text{内发送})$

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

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

求出令节点发送成功概率 $Np \cdot (1-p)^{2(N-1)}$ 最大的 p^* ，并令 $N \rightarrow \text{Infinity}$:

$$\text{最大效率} = 1/(2e) = 0.18$$



■ 随机访问协议--CSMA协议



载波侦听多址接入（CSMA）

- 发送前监听信道（carrier sensing）：
 - 信道空闲：发送整个帧
 - 信道忙：推迟发送

- 冲突仍可能发生：
 - 由于存在传输延迟，节点可能没有监听到其它节点正在发送
 - 即使忽略传输延迟，当两个（或多个）节点同时发现信道由忙变为空闲、并都决定立即发送时，仍会发生冲突



CSMA/CD (Collision Detection)

- 若在发送的过程中检测到冲突，怎么办？
 - 继续发送余下的部分（浪费带宽）
 - 停止发送余下的部分
- CSMA/CD的基本思想：
 - 在发送的过程中检测冲突（发生冲突时信号较强）
 - 检测到冲突后，立即停止发送剩余的部分
 - 立即启动冲突解决的过程



载波监听多点接入/碰撞检测 -CSMA/CD

- CSMA/CD 表示 Carrier Sense Multiple Access with Collision Detection。
- **多点接入**：表示许多计算机以多点接入的方式连接在一根总线上。
- **载波监听**：是指每一个站在发送数据之前先要检测一下总线上是否有其他计算机在发送数据，如果有，则暂时不要发送数据，以免发生碰撞。
- **碰撞检测**：在发送数据过程中进行碰撞冲突检测，而一旦检测到冲突立即停止发送数据。



CSMA/CD协议

- 以太网采用**CSMA/CD**协议:
 1. **NIC**从网络层接收数据报，构造以太帧
 2. 若**NIC**监听到信道空闲，立即发送帧；若信道忙，坚持监听直至发现信道空闲，然后发送帧
 3. 若**NIC**发送完整个帧而没有检测到冲突，认为发送成功！
 4. 若**NIC**在传输过程中检测到冲突，立即停止发送帧，并发送一个阻塞信号（加强冲突）



CSMA/CD协议（续）

5. NIC进入指数回退阶段，选择一个等待时间：

- 第一次冲突后：从 $\{0,1\}$ 中选择K，延迟 $K \cdot 512$ 比特时间
- 第二次冲突后：从 $\{0,1,2,3\}$ 中选择K，
- 第三次冲突后：从 $\{0,1,2,3,4,5,6,7\}$ 中选择K，
-
- 第10次冲突后，从 $\{0,1,2,3,4,...,1023\}$ 中选择K，

6. 返回Step 2

- 指数回退的目的是根据网络负载调整重传时间：
 - 负载越重（冲突次数越多），重传时间的选择范围越大，再次发生冲突的可能性越小

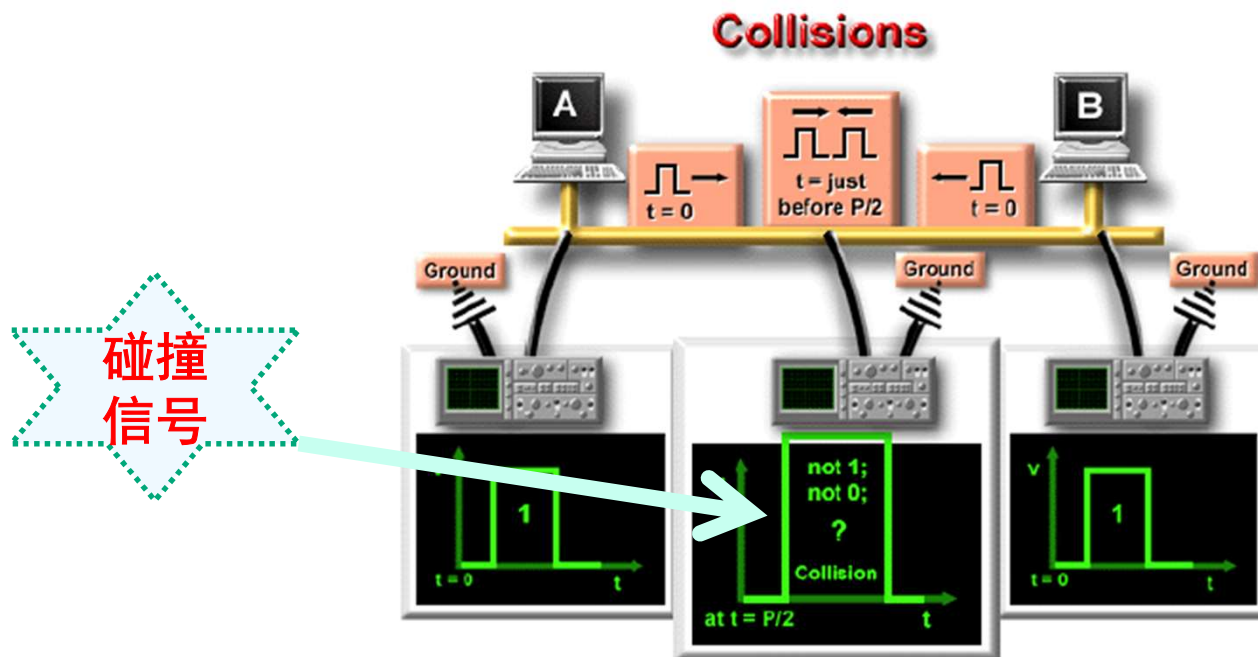


二进制指数类型退避算法

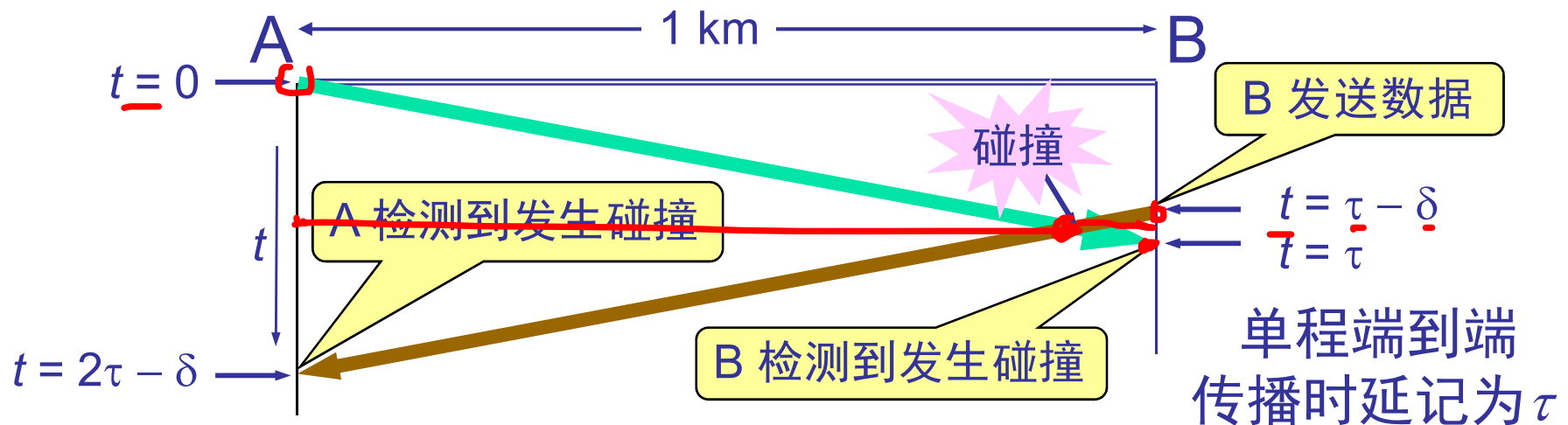
- 发生碰撞的站在停止发送数据后，要推迟（退避）一个随机时间才能再发送数据。
 - 确定基本退避时间，一般是取为争用期 2τ 。
 - 定义重传次数 k ， $k \leq 10$ ，即
$$k = \text{Min}[\text{重传次数}, 10]$$
 - 从整数集合 $[0, 1, \dots, (2^k - 1)]$ 中随机地取出一个数，记为 r 。重传所需的时延就是 r 倍的基本退避时间。
 - 当重传达 16 次仍不能成功时即丢弃该帧，并向高层报告。

碰撞检测

- “碰撞检测”就是计算机边发送数据边检测信道上的信号电压大小。
- 当几个站同时在总线上发送数据时，总线上的信号电压摆动值将会增大（互相叠加）。

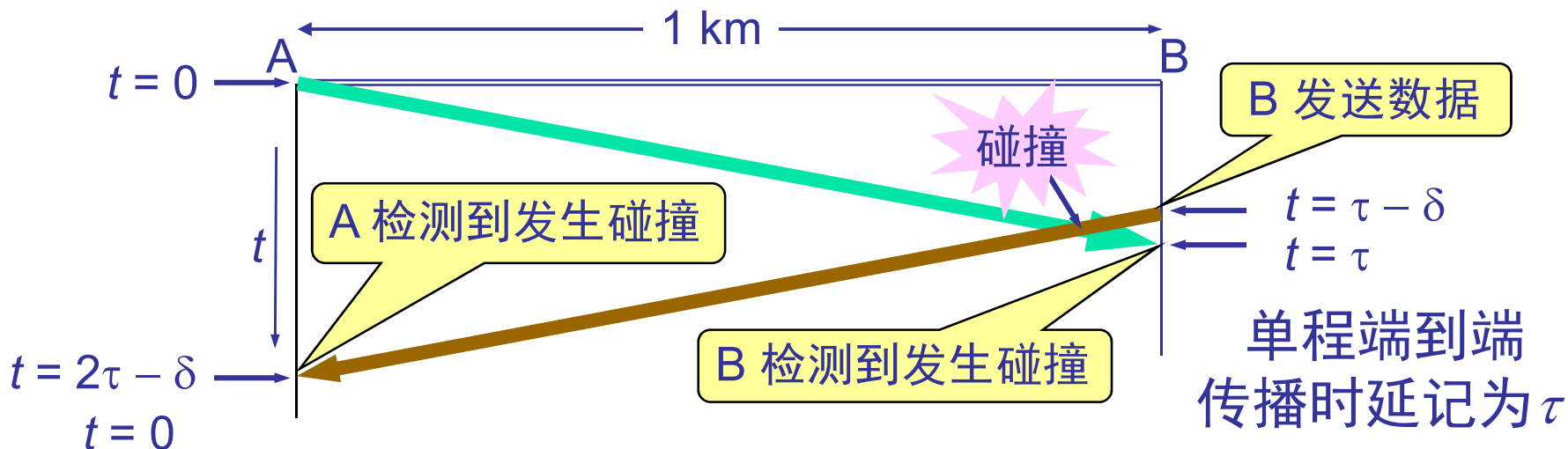


碰撞窗口：发送后检测到碰撞的时间？

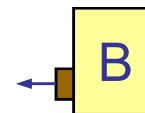
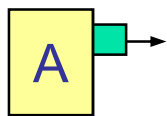


碰撞窗口--发生冲突时间的上限：

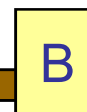
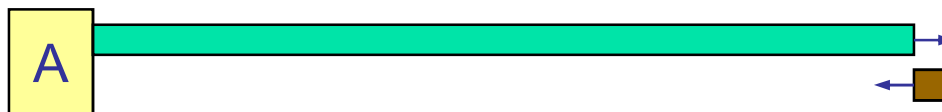
- 即发送站发出帧后能检测到碰撞的最长时间
- 数值上等于最远两站传播时间的两倍，即 2τ



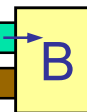
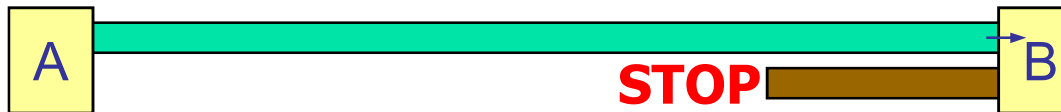
A 检测到
信道空闲
发送数据



$t = \tau - \delta$
B 检测到信道空闲
发送数据

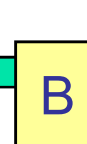
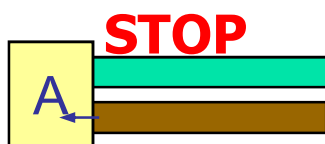


$t = \tau - \delta / 2$
发生碰撞



$t = \tau$
B 检测到发生碰撞
停止发送

$t = 2\tau - \delta$
A 检测到
发生碰撞



STOP

STOP



碰撞窗口值-以太网争用期 2τ

- 以太网的端到端往返时延 2τ 称为争用期（Contention Period），或碰撞窗口。
 - 最先发送数据帧的站，在发送数据帧后至多经过时间 2τ （两倍的端到端时延）就可知道发送的数据帧是否遭受了碰撞。
- 以太网的争用期取值？



争用期和最短有效帧长度

■ 10 M以太网取 $51.2 \mu\text{s}$ 为争用期的长度。

- 对于 10 Mb/s 以太网，在争用期内可发送 512 bit，即 64 字节。
- 比特时间： $0.1 \mu\text{s/bit}$

■ 最短有效帧长度：

- 以太网在发送数据时，若前 64 字节没有发生冲突，则后续的数据就不会发生冲突。
- 长度小于 64 字节的帧都是由于冲突而异常中止的无效帧。

■ 为什么最小帧长为64字节（不包括前导码）：

- 根据早期以太网的最大直径（2500米）和数据速率（10Mbps）计算得到



半双工信道

- 使用 CSMA/CD 协议的以太网不能进行全双工通信而只能进行双向交替通信（半双工通信）。
- 每个站在发送数据之后的一小段时间内，存在着遭遇碰撞的可能性。
- 发送的不确定性使整个以太网的平均通信量远小于以太网的最高数据率。



CSMA/CD的效率

- T_{prop} = 以太网中任意两个节点之间传播延迟的最大值
- t_{trans} = 最长帧的传输时间
 - $Efficiency = 1 / (1 + T_{prop} / t_{trans})$
- 在以下情况下，以太网的效率趋近于1：
 - t_{prop} 趋近于 0，或
 - t_{trans} 趋向于无穷
- 结论：应控制以太网的规模



CSMA (补充)

■ CSMA:

- ALOHA协议: 发送前不侦听信道
- CSMA: 发送前侦听信道

■ CSMA分类

- 非持续式 (发现信道忙, 等待一个随机时间后再侦听)
- 持续式 (持续侦听信道)
 - 1-持续CSMA (如果信道空闲, 立即发送数据)
 - P-持续CSMA (如果信道空闲, 以 p 的概率发送数据)



非持续式 CSMA

■ 特点：

- ①经侦听，如果介质空闲，开始发送。
 - ②如果介质忙，则等待一个随机分布的时间，然后重复步骤①。
- 等待一个随机时间可以减少再次碰撞冲突的可能性。
 - 但缺点是等待时间内介质上没有数据传送，这段时间是浪费的。



1-持续式 CSMA

■ 特点：

- ①经侦听，如介质空闲，则发送。
- ②如介质忙，持续侦听，一旦空闲立即发送。
- ③如果发生冲突，等待一个随机分布的时间再重复步骤①。

■ 持续式的延迟时间要少于非持续式。

- 主要问题是：如果两个以上的站等待发送，一旦介质空闲就一定会发生冲突。



p-持续式 CSMA

■ 特点:

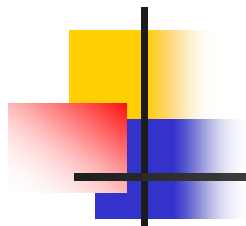
- ①经侦听，如介质空闲，那么以 p 的概率发送，以 $(1-p)$ 的概率延迟一个时间单元发送。
- ②如介质忙，持续侦听，一旦空闲重复①。
- ③如果发送已推迟一个时间单元，再重复步骤①。

■ 可见，1-持续式是p-持续式的特例。



CSMA/CD (1-持续)

- CSMA with Collision Detection
- “先听后发、边发边听”
- 特点：
 - ①经侦听，如介质空闲，则发送。
 - ②如介质忙，持续侦听，一旦空闲立即发送。
 - ③如果发生冲突，等待一个随机分布的时间再重复步骤①。

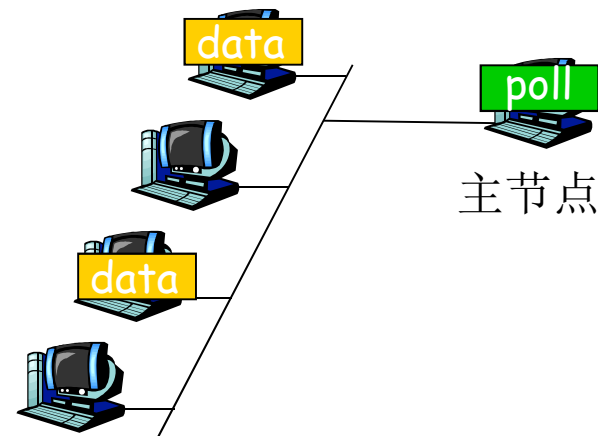


■ 受控访问协议（Controlled Access）

轮流MAC协议—轮询协议

轮询

- 主节点轮流“邀请”从节点发送，邀请到的从节点允许发送
- 缺点：
 - 引入轮询延迟
 - 单点失效（主节点）



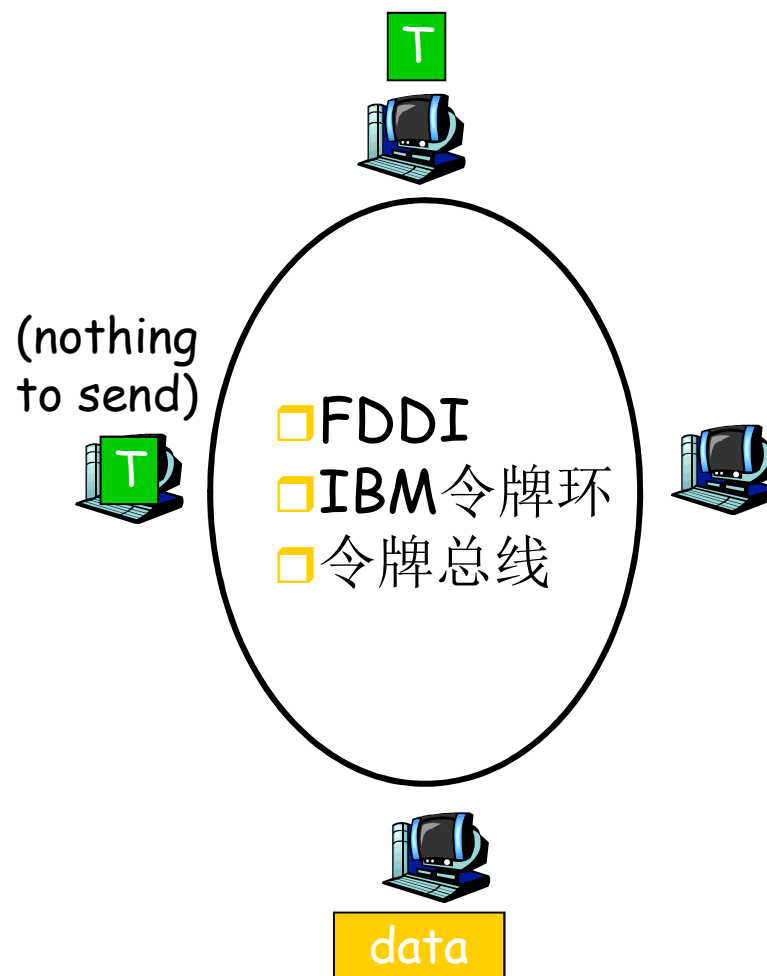
从节点

蓝牙

轮流MAC协议—令牌协议

令牌传递

- 网络中有一个令牌，按预定顺序在节点间传递
- 获得令牌的节点可以发送
- 发送完数据后释放令牌
- 缺点：
 - 令牌传递延迟
 - 单点失效（令牌）





多址访问协议 (Multiple access Protocol)

小结

- 按照时间、频率、编码划分信道：
 - 时分多址，频分多址，码分多址
- 随机接入：
 - 纯ALOHA, S-ALOHA (ALOHA网络)
 - CSMA/CD (早期以太网)
 - CSMA/CA (802.11) (无线网络)
- 轮流访问：
 - 中心节点轮询 (蓝牙)
 - 令牌传递 (FDDI, IBM令牌环, 令牌总线)



MAC协议比较

信道划分MAC协议:

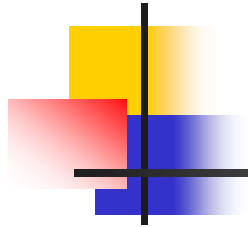
- 重负载下高效: 没有冲突, 节点公平使用信道
- 轻负载下低效: 即使只有一个活跃节点也只能使用 $1/N$ 的带宽

随机接入MAC协议:

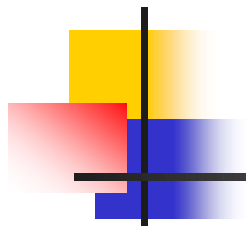
- 轻负载时高效: 单个活跃节点可以使用整个信道
- 重负载时低效: 频繁发生冲突, 信道使用效率低

轮流协议 (试图权衡以上两者):

- 按需使用信道 (避免轻负载下固定分配信道的低效)
- 消除竞争 (避免重负载下的发送冲突)



6、局域网



3.4 局域网



IEEE 802 局域网系列标准

- IEEE 802定义LAN的物理层、MAC子层的定义和描述。它的组成如下：
 - 802.1 基本介绍和接口原语定义
 - 802.2 逻辑链路控制（LLC）子层
 - 802.3 CSMA/CD协议的局域网
 - 802.4 令牌总线（Token Bus）的局域网
 - 802.5 采用令牌环（Token Ring）的局域网
 - 802.11 CSMA/CA协议的无线局域网



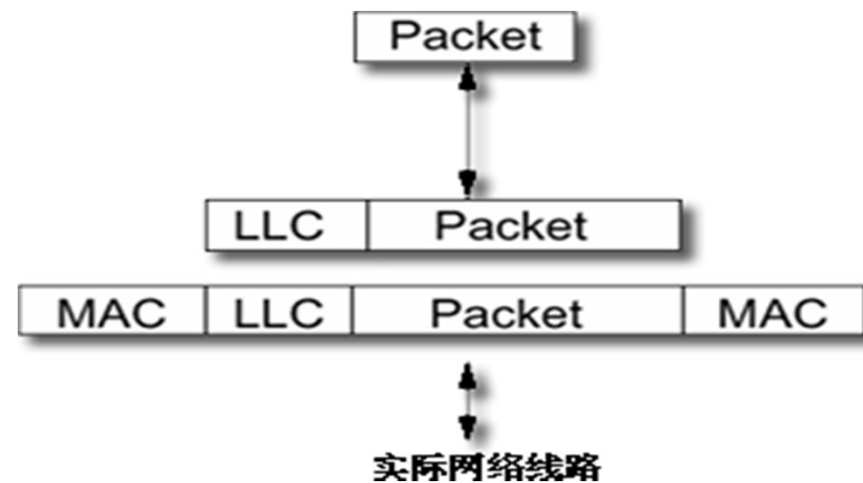
IEEE标准的LAN参考模型

- 数据链路层的两个子层
 - 逻辑链路控制 LLC (Logical Link Control)子层
 - 介质接入控制 MAC (Medium Access Control)子层。
- 与接入到传输媒体有关的内容都放在 MAC 子层
- 而 LLC 子层则与传输媒体无关，不管采用何种协议的局域网对 LLC 子层来说都是透明的

逻辑链路控制子层（现在不用了）

LLC（Logical Link Control）

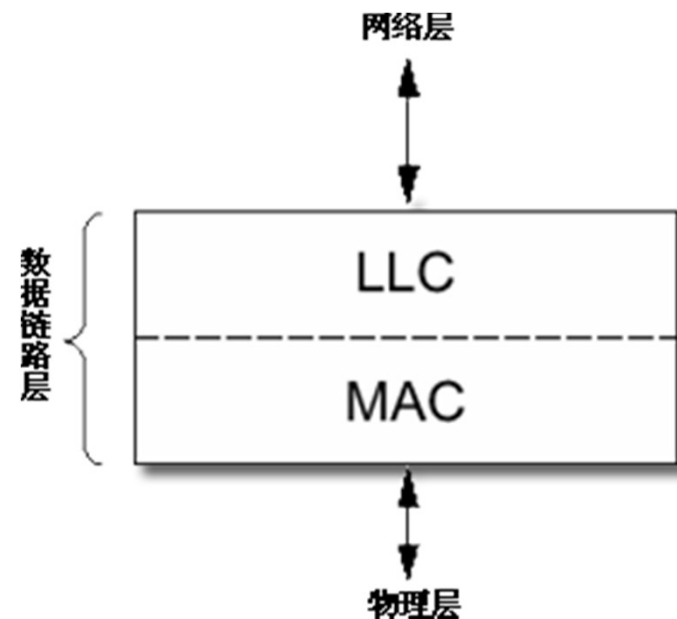
- LLC子层提供确认机制和流量控制；
- LLC隐藏了不同802MAC子层的差异，为网络层提供单一格式和接口；
- LLC提供三种服务选项：
 - ✓ 不可靠的数据报服务
 - ✓ 确认数据报服务
 - ✓ 面向连接的可靠服务
- LLC帧头基于HDLC协议



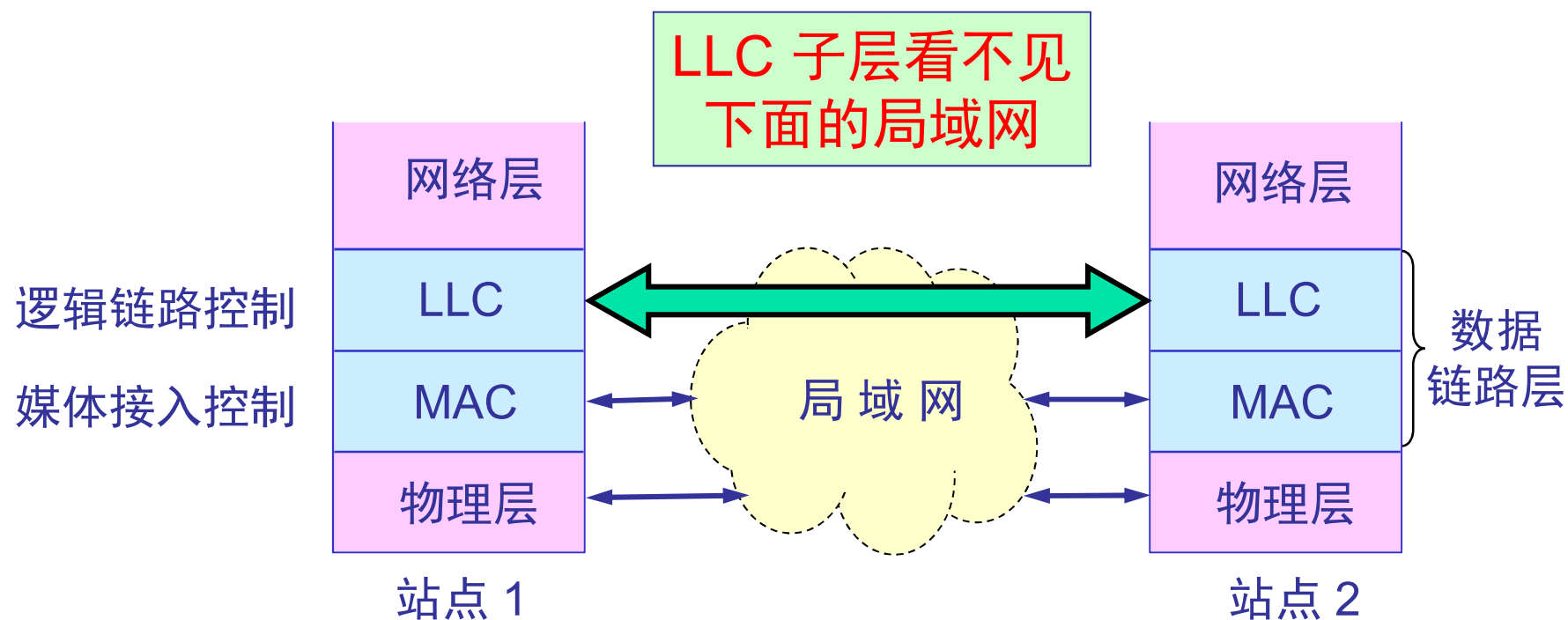
介质访问控制子层 MAC

(Medium Access Control)

- 数据封装 (transmit and receive)
 - 成帧Framing (帧定界、透明、同步)
 - 寻址Addressing
 - 检错Error detection
- 介质访问管理
 - 避免冲突
 - 处理冲突



局域网对 LLC 子层是透明的





现在一般不考虑 LLC 子层

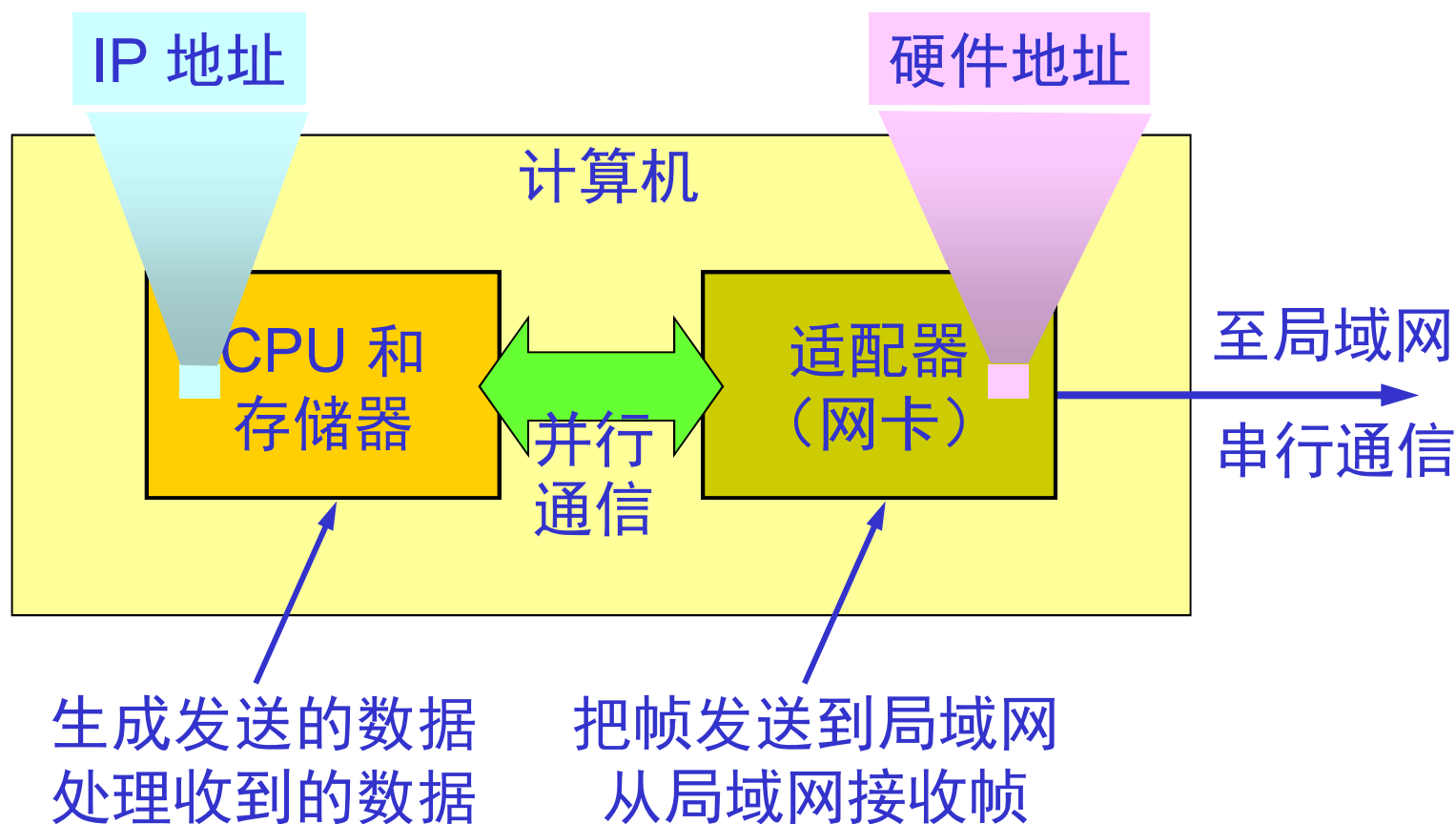
- 由于 TCP/IP 体系经常使用的局域网是 DIX Ethernet V2 而不是 802.3 标准中的几种局域网；
- 现在 802 委员会制定的逻辑链路控制子层 LLC（即 802.2 标准）的作用已经不大了。
- 很多厂商生产的适配器上就仅装有 MAC 协议而没有 LLC 协议。



适配器的作用

- 网络接口板又称为**通信适配器**(adapter)或**网络接口卡** NIC (Network Interface Card), 或“**网卡**”。
- 适配器的主要功能：
 - 进行串行/并行转换。
 - 对数据进行缓存。
 - 在计算机的操作系统安装设备驱动程序。
 - 实现以太网协议。

计算机通过适配器和局域网进行通信

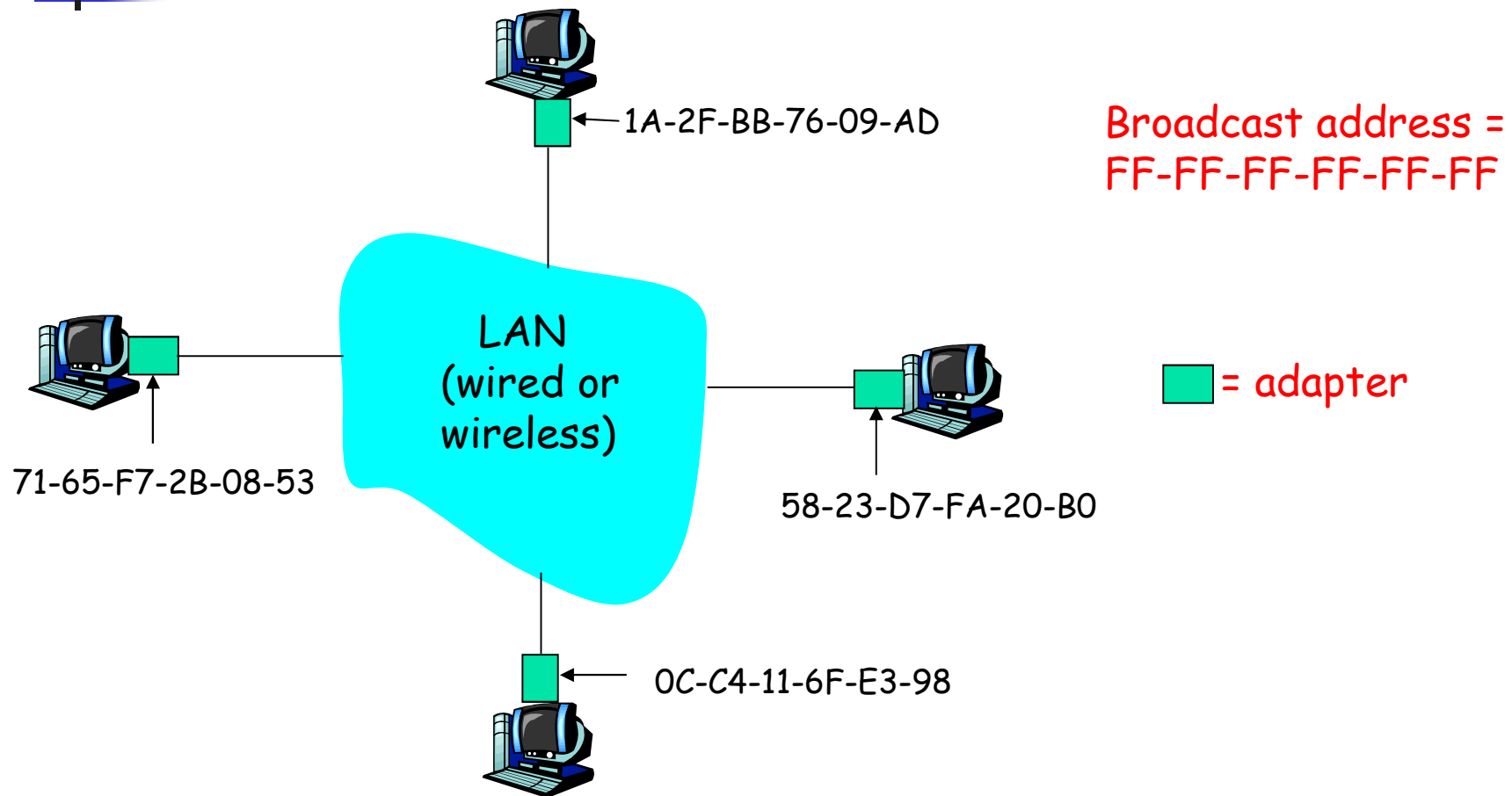




链路层编址—MAC地址

- 每一块网络适配器（网卡）固定分配一个地址，称为物理地址、硬件地址、链路层地址、MAC地址等
- MAC地址长6个字节，一般用由“:”或“-”分隔的6个十六进制数表示
- MAC地址由IEEE负责分配，每块适配器的地址是全球唯一的：
 - 网卡生产商向IEEE购买一块MAC地址空间（前3字节）
 - 生产商确保生产的每一块网卡有不同的MAC地址
 - MAC地址固化在网卡的ROM中
 - 现在用软件改变网卡的MAC地址也是可能的

每个适配器有一个MAC地址





网络适配器的帧接收处理

- 网络适配器仅将发送给本节点的帧交给主机
 - 目的地址为适配器MAC地址的单播帧
 - 所有广播帧
 - 指定接收的多播帧
- 若将适配器设置成混收模式，适配器将收到的所有帧交给主机



3类MAC地址

■ 单播地址：一对一

- 将数据帧传给同一网络上的对应地址主机
 - 34-02-86-EE-87-C9

■ 多播地址：

- 将数据帧传给同一网络上的多播组内的多个主机
- 最高字节的最低位置 1 的地址（以太网）
 - 01:00:00:00:00:00

■ 广播地址：同一网络上的所有主机

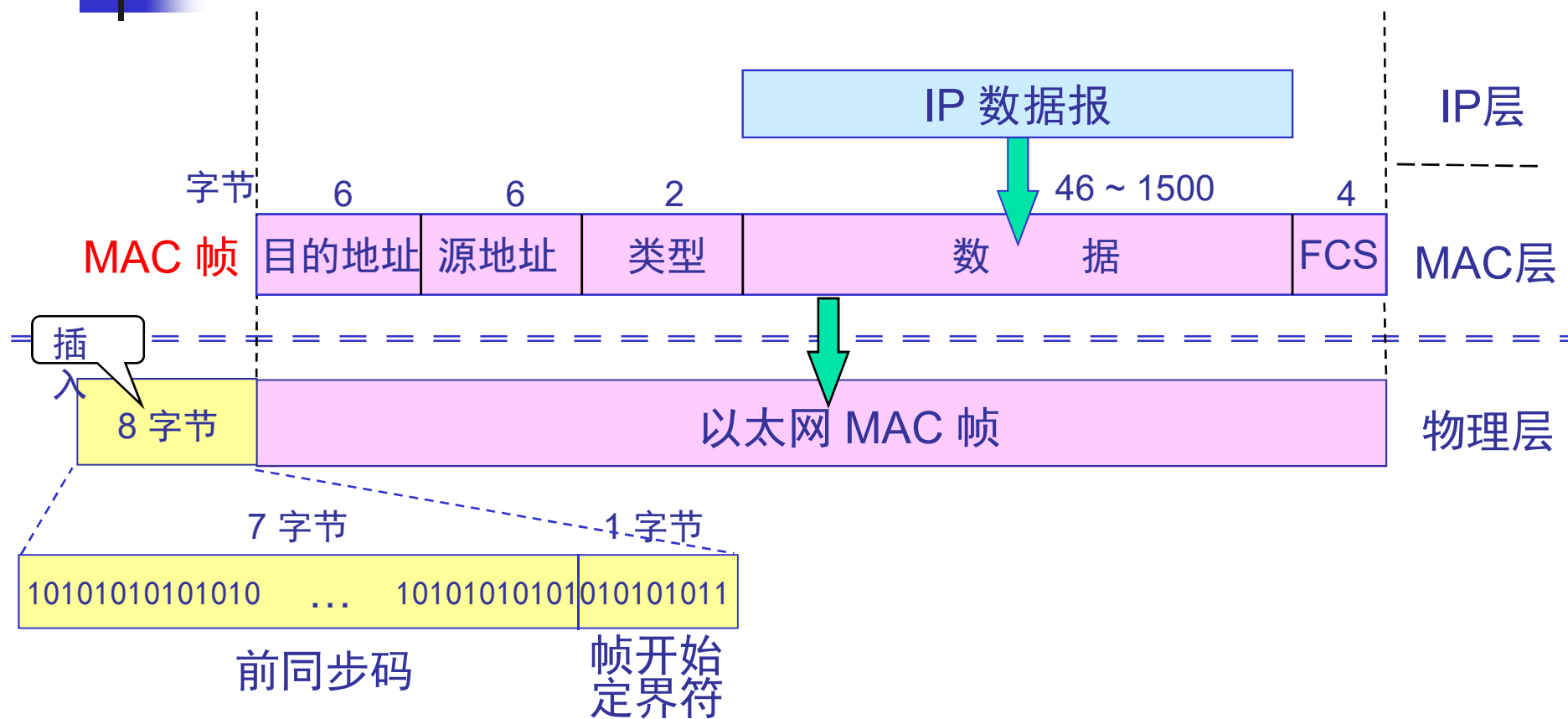
- 将数据帧传给同一网络上的所有主机
 - ff:ff:ff:ff:ff:ff



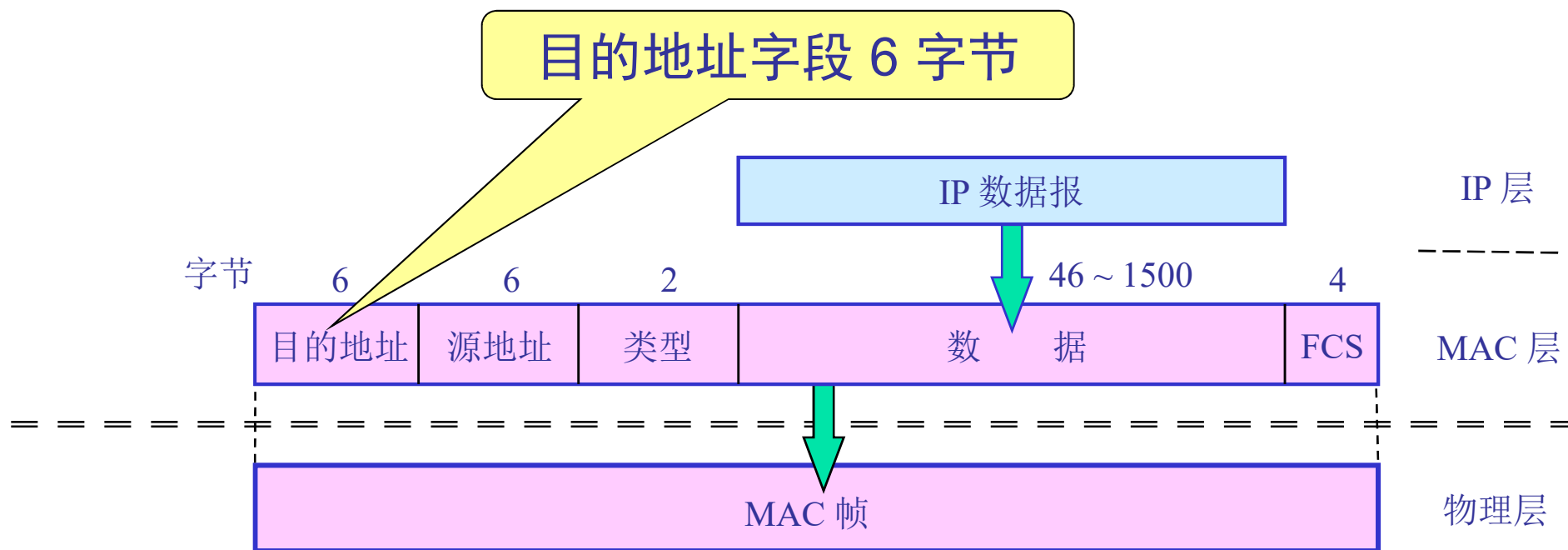
MAC 帧的格式

- 常用的以太网MAC帧格式有两种标准：
 - DIX Ethernet V2 标准
 - IEEE 的 802.3 标准
- 最常用的 MAC 帧是以太网 V2 的格式。

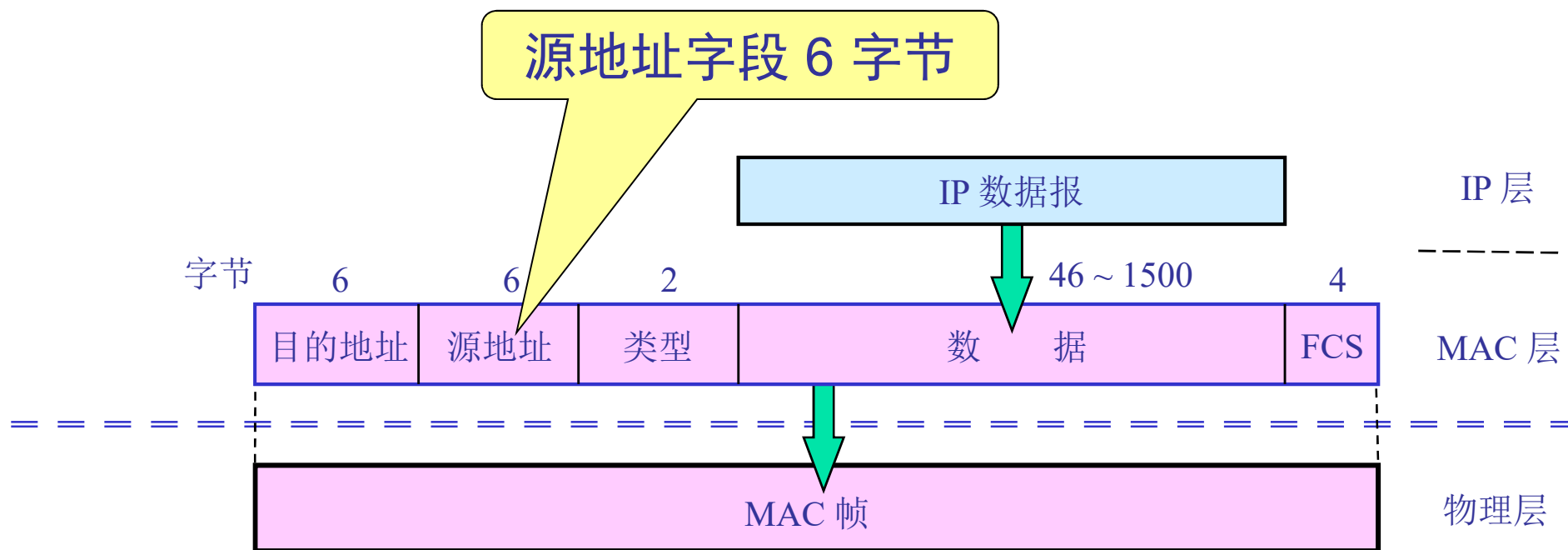
以太网的 MAC 帧格式



以太网 V2 的 MAC 帧格式

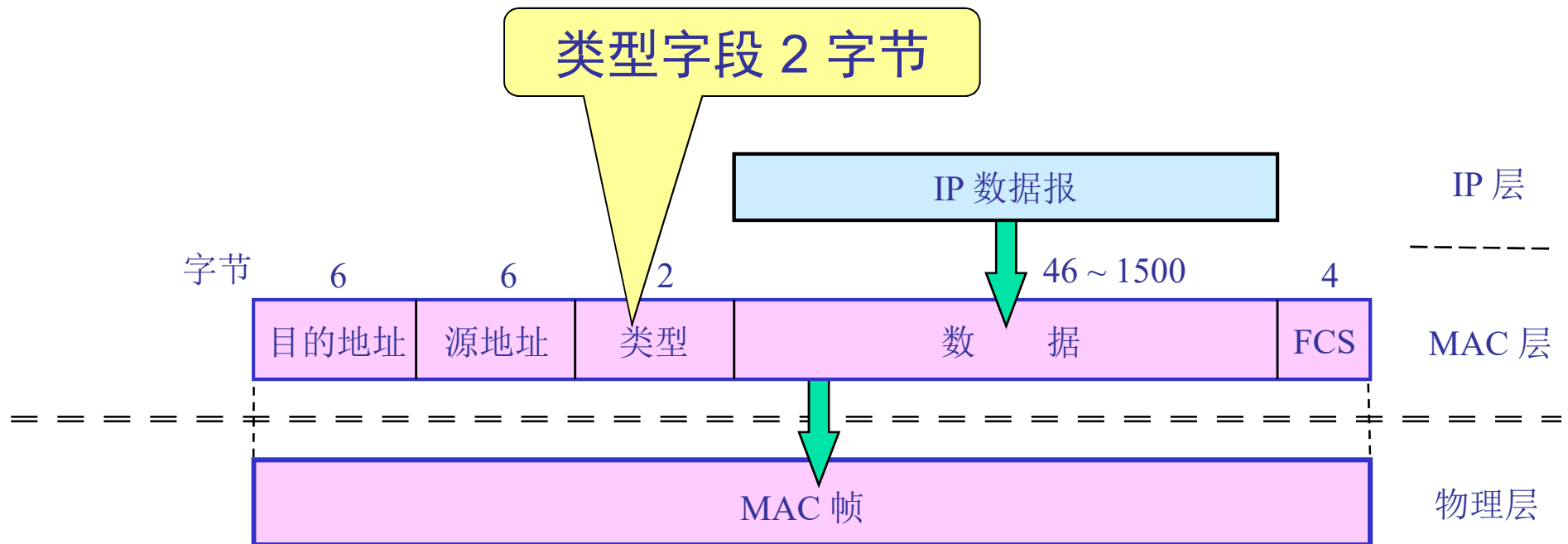


以太网 V2 的 MAC 帧格式



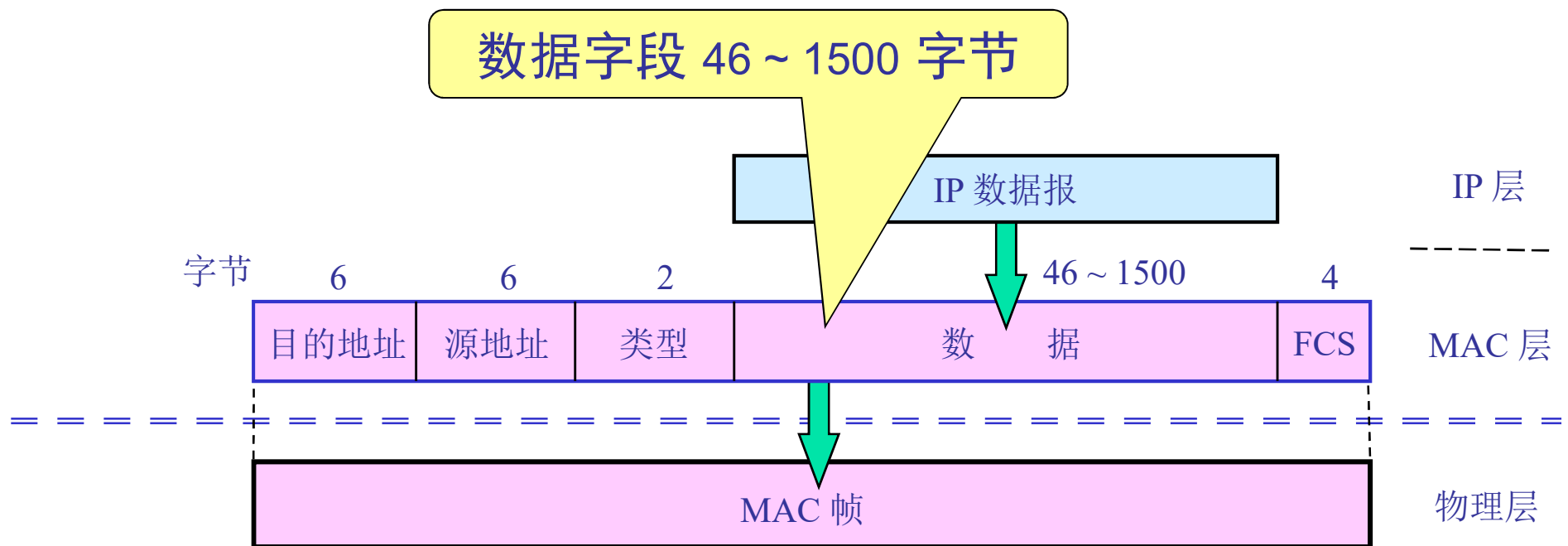
以太网 V2 的 MAC 帧格式

类型字段用来标志~~上~~一层使用的是什么协议，以便把收到的 MAC 帧的数据上交给上层的这个协议。



以太网 V2 的 MAC 帧格式

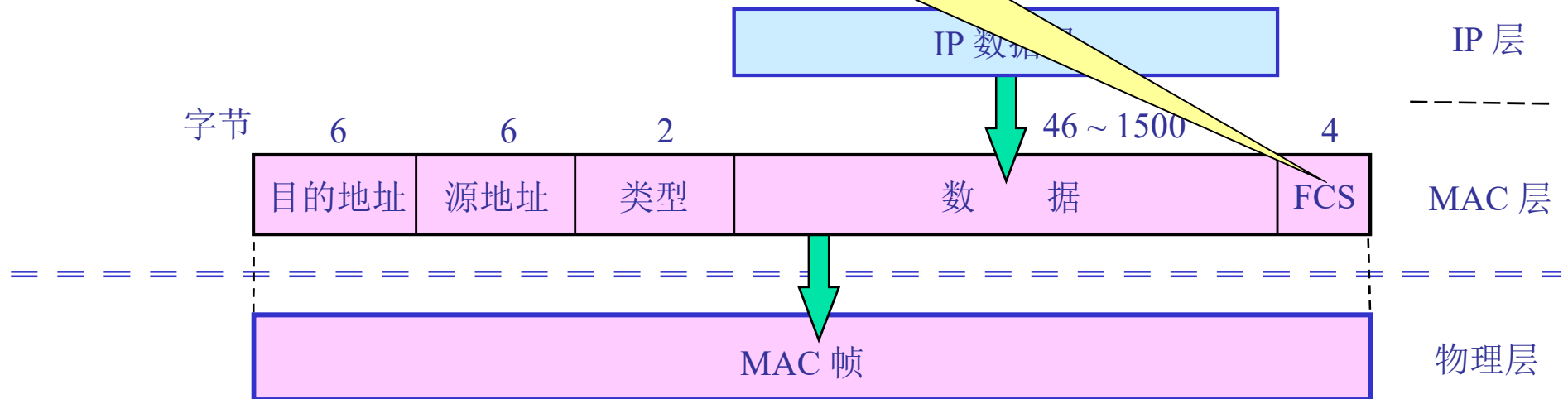
数据字段的正式名称是 **MAC 客户数据字段**
最小长度 64 字节 – 18 字节的首部和尾部 = 数据字段的最小长度



以太网 V2 的 MAC 帧格式

当传输媒体的误码率为 1×10^{-8} 时，
MAC 子层可使未检测到的差错小于 1×10^{-14} 。

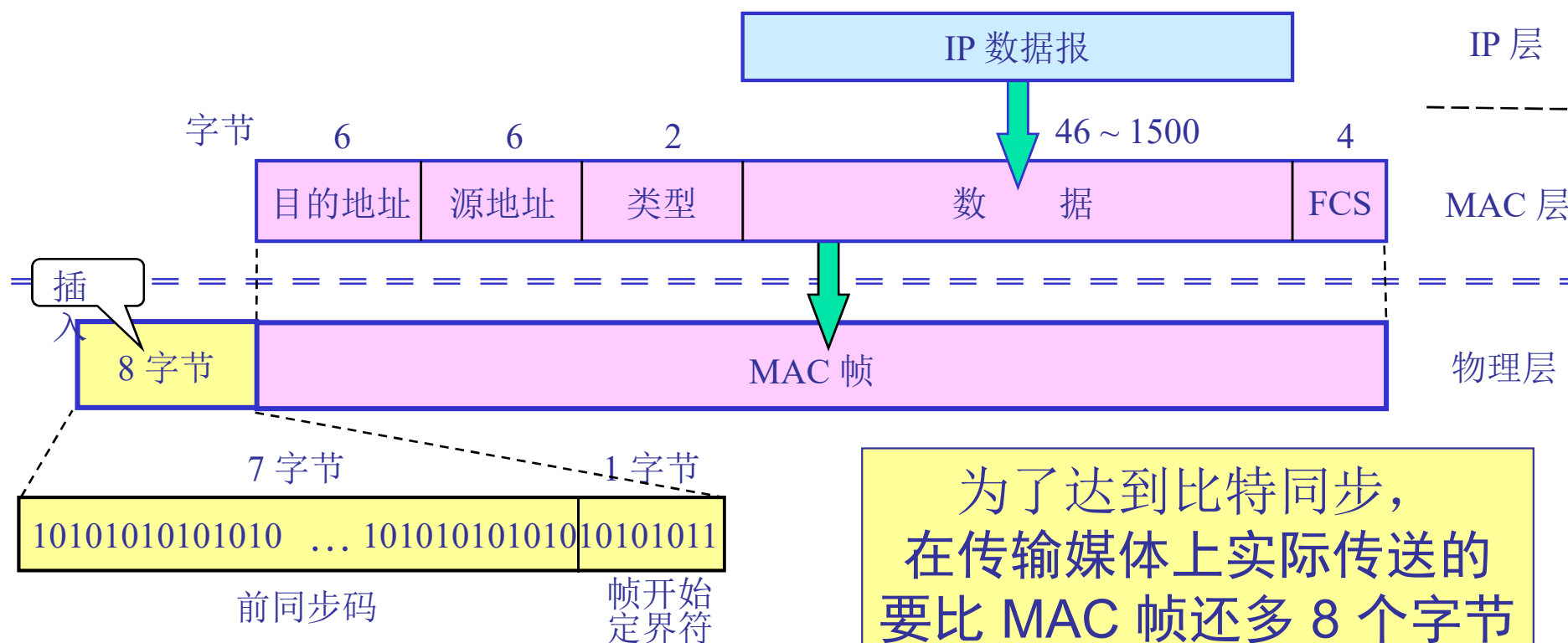
FCS 字段 4 字节



当数据字段的长度小于 46 字节时，
应在数据字段的后面加入整数字节的填充字段，
以保证以太网的 MAC 帧长不小于 64 字节。

以太网 V2 的 MAC 帧格式

在帧的前面插入的 8 字节中的第一个字段共 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。





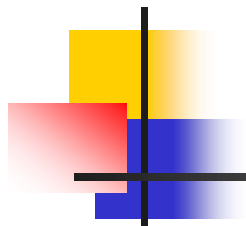
无效的 MAC 帧

- 数据字段的长度与长度字段的值不一致；
- 帧的长度不是整数个字节；
- 用收到的帧检验序列 FCS 查出有差错；
- 数据字段的长度不在 46 ~ 1500 字节之间。
- 有效的 MAC 帧长度为 64 ~ 1518 字节之间。
- 对于检查出的无效 MAC 帧就简单地丢弃。以太网不负责重传丢弃的帧。



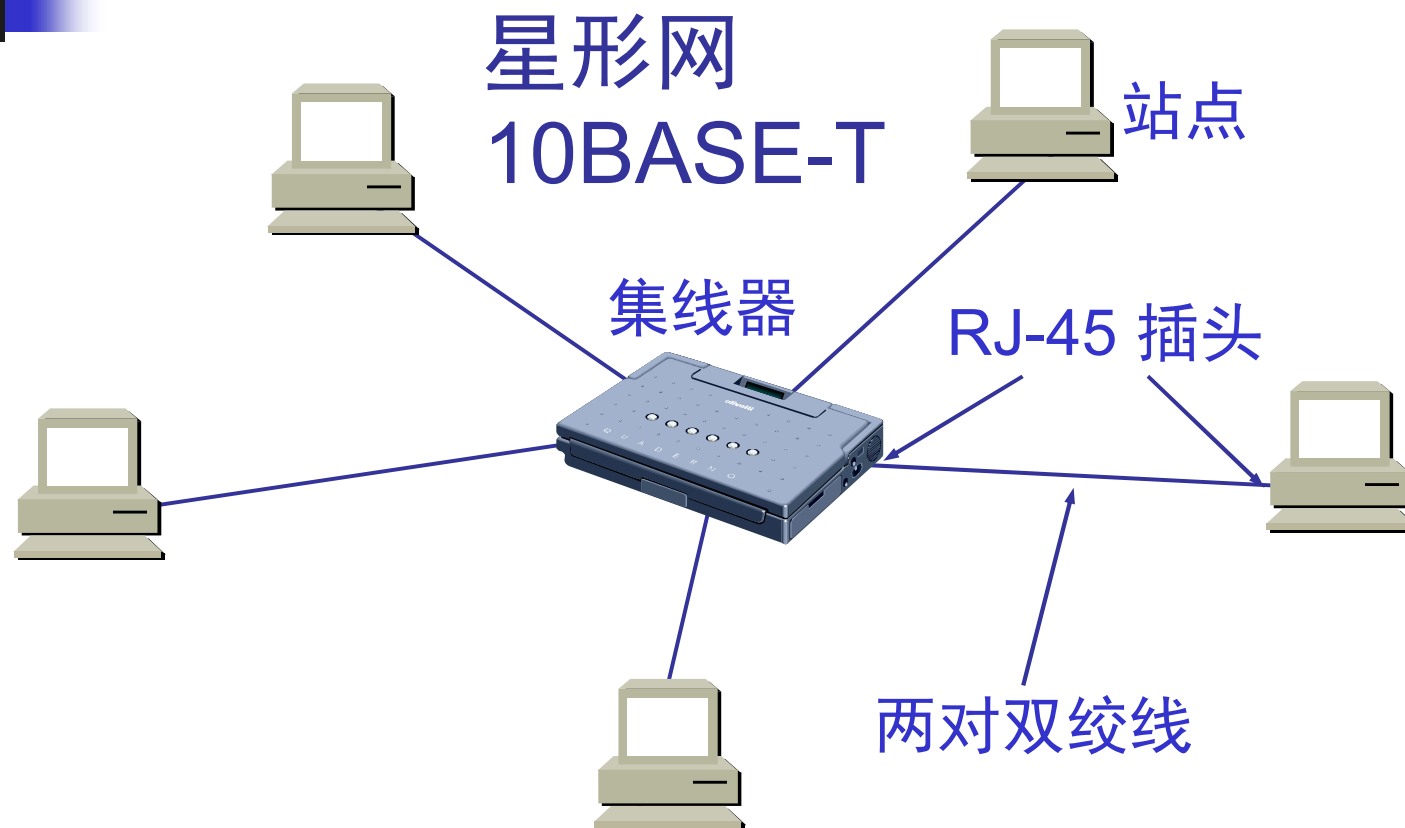
帧间最小间隔

- 帧间最小间隔为 $9.6\ \mu\text{s}$ ，相当于 96 bit 的发送时间。
- 一个站在检测到总线开始空闲后，还要等待 $9.6\ \mu\text{s}$ 才能再次发送数据。
- 这样做是为了使刚刚收到数据帧的站的接收缓存来得及清理，做好接收下一帧的准备。



7、物理层设备构造局域网—集线器

使用集线器构造以太网



以太网采用星形拓扑，在星形的中心则增加了一种可靠性非常高的设备，叫做**集线器**(hub)



10BASE-T以太网

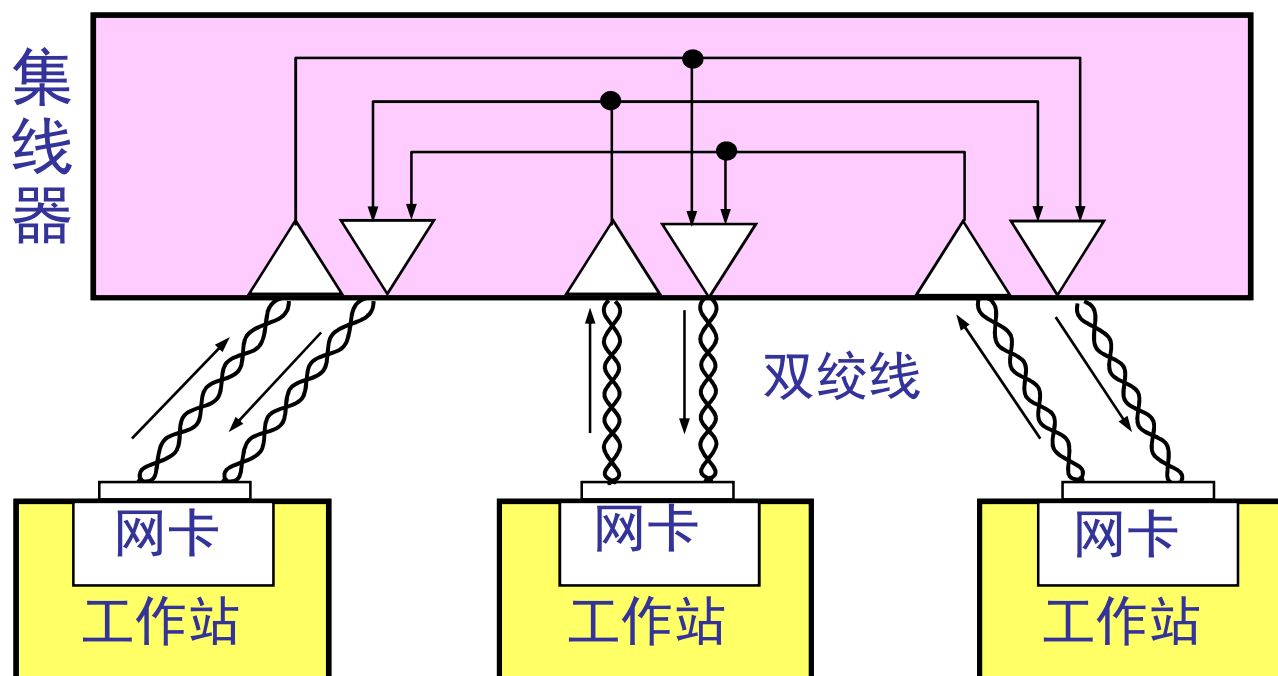
- 10BASE-T 的通信距离稍短，每个站到集线器的距离不超过 100 m。
- 10BASE-T 双绞线以太网的出现，是局域网发展史上的一个非常重要的里程碑，它为以太网在局域网中的统治地位奠定了牢固的基础。



10Mbps以太网（早期以太网）

- 10Base-5:
 - 基带同轴电缆（粗），每段电缆最大长度500米
- 10Base-2:
 - 基带同轴电缆（细），每段电缆最大长度约200米
- 10Base-T
 - 3类双绞线和集线器，双绞线最大长度100米
- 10Base-F
 - 多模光纤和集线器，光纤最大长度2000米

集线器构造的局域网实质是一种共享总线信道





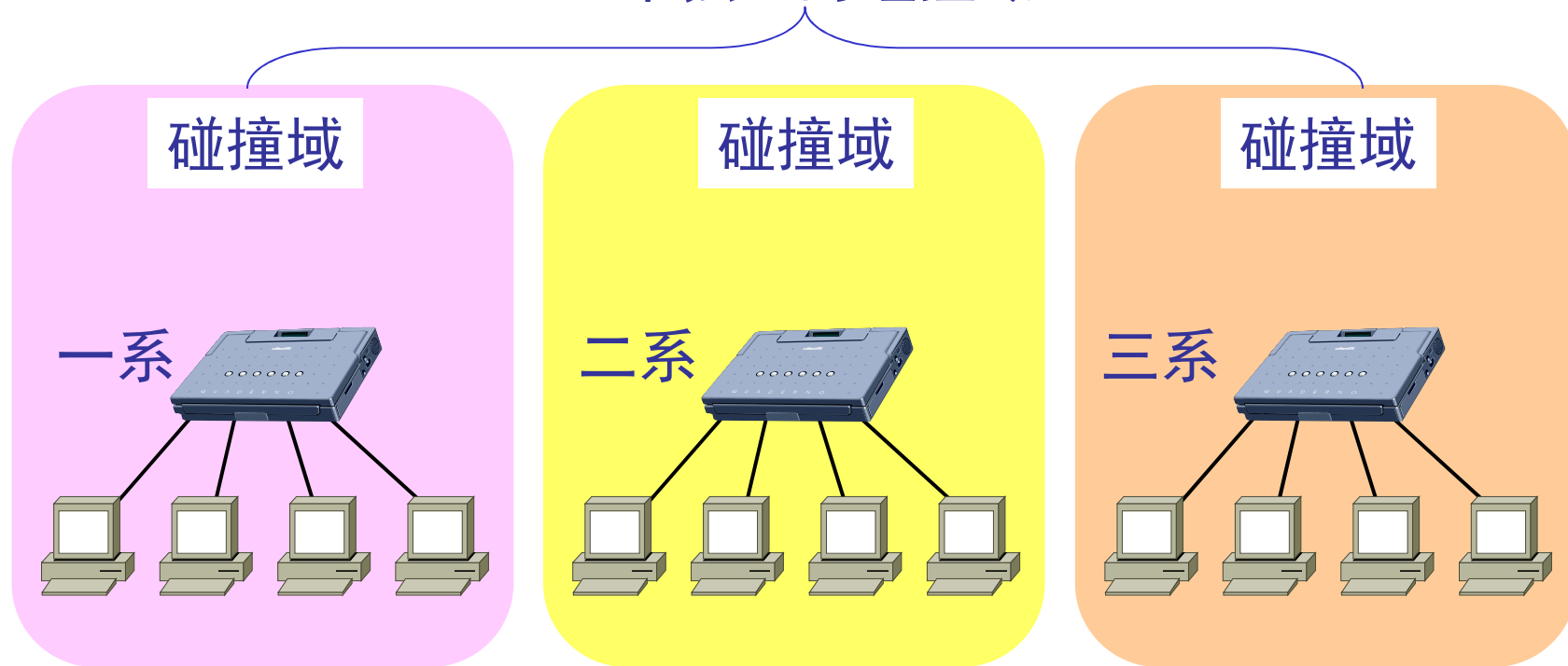
逻辑上是一个总线网

- 集线器是使用电子器件来模拟实际电缆线的工作。
- 使用集线器的以太网在逻辑上仍是一个总线网
- 集线器工作在物理层
- 使用CSMA/CD 协议实现共享信道介质访问。

用多个集线器级联成更大局域网

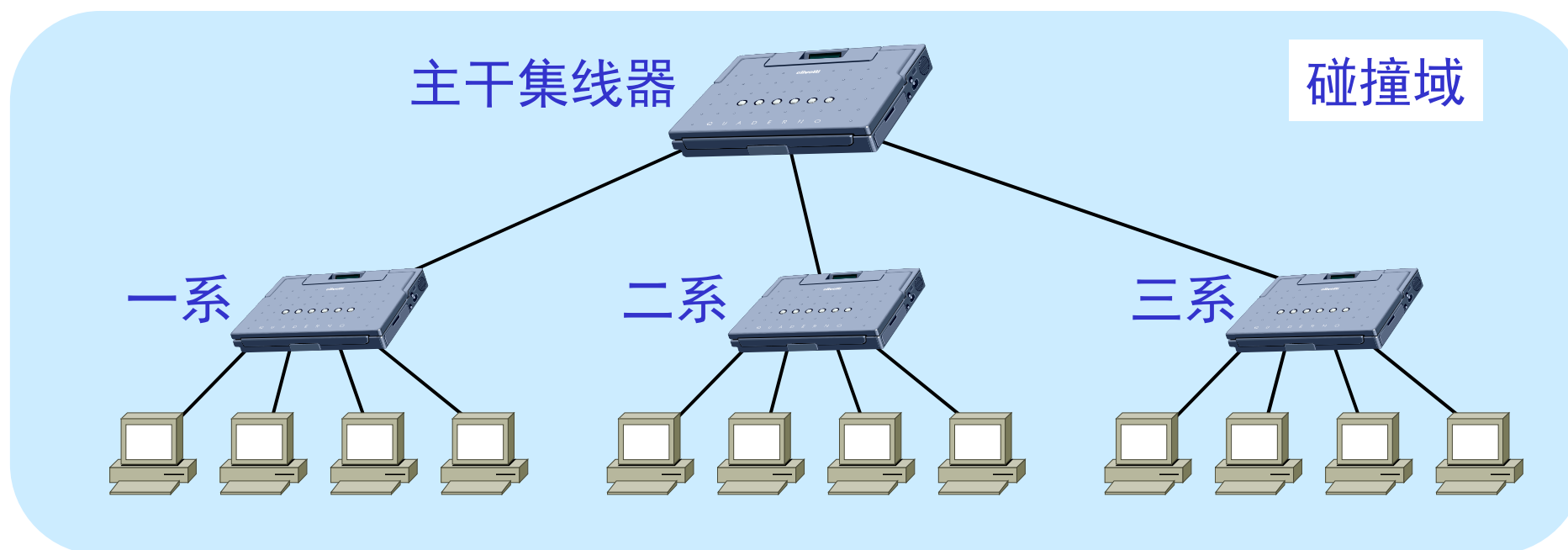
- 某大学有三个系，各自有一个局域网

三个独立的碰撞域



用集线器组成更大的局域网 都在一个碰撞域中

一个更大的碰撞域





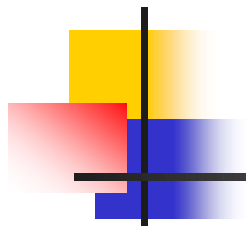
用集线器扩展局域网

■ 优点

- 使原来属于不同碰撞域的局域网上的计算机能够进行跨碰撞域的通信。
- 扩大了局域网覆盖的地理范围。

■ 缺点

- 碰撞域增大了，但总的吞吐量并未提高。
- 如果不同的碰撞域使用不同的数据率，那么就不能用集线器将它们互连起来。



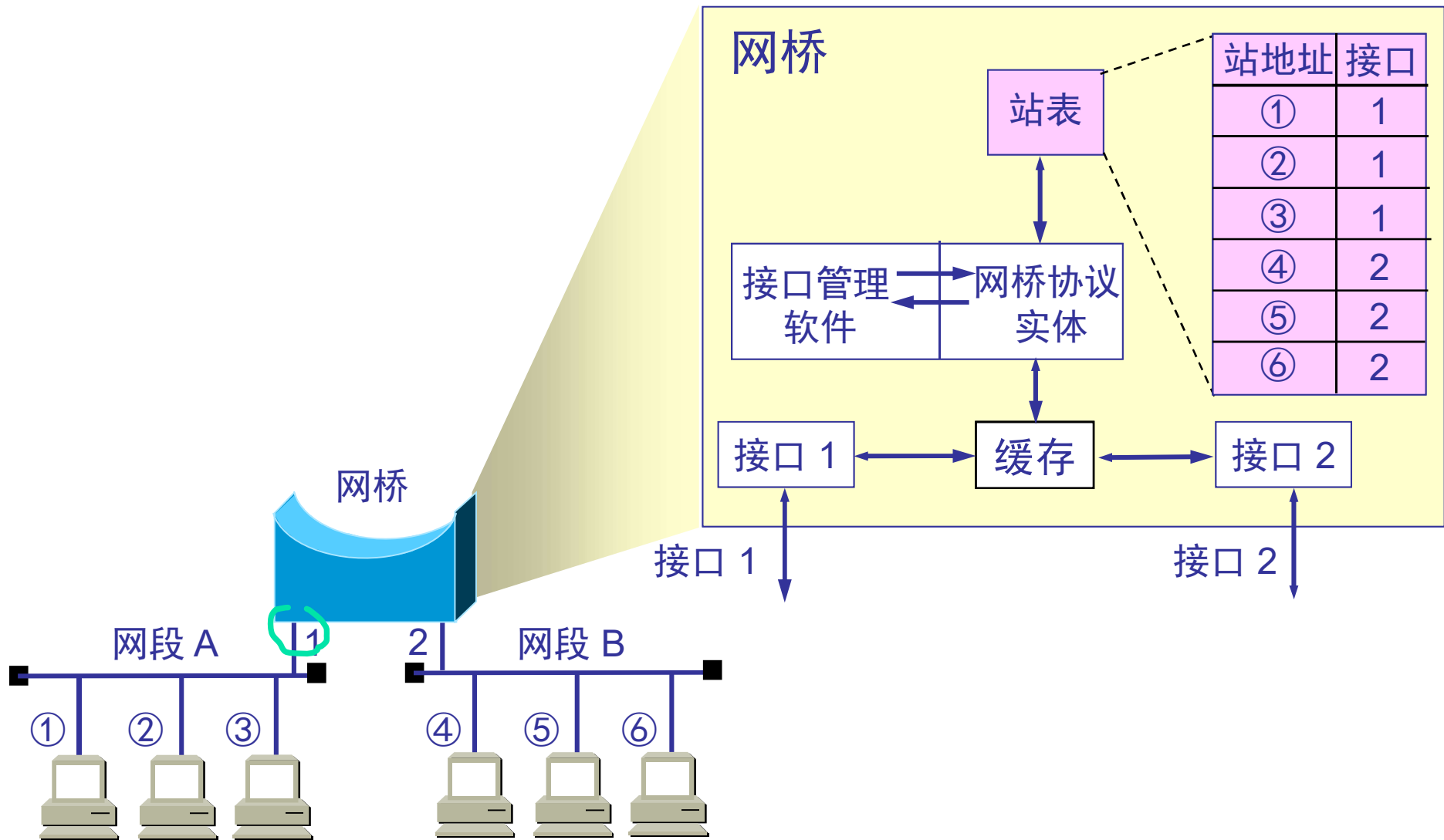
8、链路层设备扩展局域网—网桥、交换机



数据链路设备层扩展局域网

- 在数据链路层扩展局域网是使用网桥。
- 网桥工作在数据链路层，它根据 MAC 帧的地址对收到的帧进行转发。
- 网桥具有过滤帧的功能。当网桥收到一个帧时，并不是向所有的接口转发此帧，而是先检查此帧的目的 MAC 地址，然后再确定将该帧转发到哪一个接口

1. 网桥的内部结构

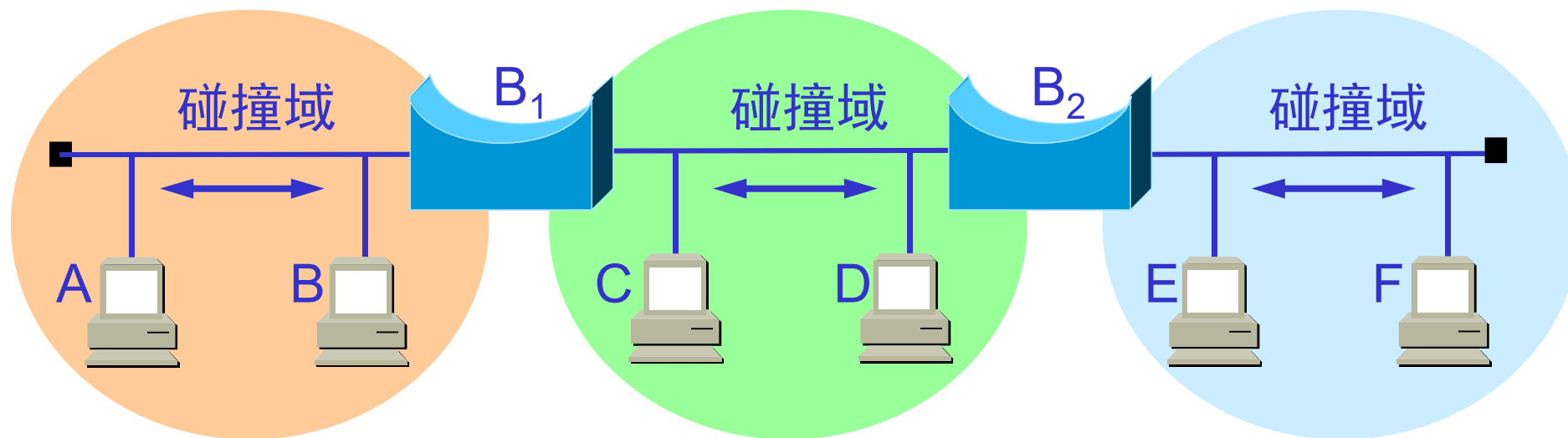




网桥的优点

- 过滤部分通信量。
 - 数据帧过滤
- 扩大了局域网物理范围。
- 不同局域网互联：
 - 可互连不同物理层、不同 **MAC** 子层和不同速率
 - 如10 Mb/s 和 100 Mb/s局域网
- 提高了可靠性。

网桥使各网段成为 隔离的碰撞域





网桥的缺点

■ 增加时延

- 存储转发增加了时延，具有不同 MAC 子层的网段桥接在一起时时延更大。

■ 适用小规模网络

- 网桥只适合于用户数不太多(不超过几百个)和通信量不太大的局域网，否则有时还会因传播过多的广播信息而产生网络拥塞--广播风暴。



网桥和集线器（或转发器）不同

■ 集线器属于物理层设备

- 不处理数据帧
- 不对传输媒体进行检测
- 属于同一碰撞域

■ 网桥属于链路层设备

- 接收并转发帧
- 转发帧需执行 CSMA/CD 算法
- 隔离多个碰撞域
- 可互连不同局域网：如10 Mb/s 和 100 Mb/s



透明网桥

- “透明”是指局域网上的站点并不知道所发送的帧将经过哪几个网桥，因为网桥对各站来说是看不见的。
- 透明网桥是一种**即插即用设备**，其标准是 IEEE 802.1D。



透明网桥

- 通过透明网桥（transparent bridges）将多个 LAN 连接起来，硬件和软件不需要做任何的变化
- 透明网桥接收所有跟它相联的 LAN 的帧
- 帧转发 or 丢弃：
 - 当一个帧到达网桥时，它必须作出丢弃（discard）还是转发（forward）的决策，如果是转发，它还要知道向哪个 LAN 转发
 - 通过在网桥内部的一张地址转发表中查找目的 MAC 地址而作出转发 or 丢弃的决策



网桥如何维护它的内部转发表？

- 初始时，转发表是空的
- 扩散算法（泛洪算法）
 - 当网桥不知道目的地址时（表中查不到），它会将这帧从除来的**LAN**外的所有**LAN**转发出去
- 逆向学习
 - 网桥从到达帧的源地址认识到源地址对应的那台机是在帧来的那个**LAN**上，所以，把它写入**MAC**地址表



网桥的转发表

- 网桥的转发表：
 - 包括地址、接口、帧进入该网桥的时间。
- 把每个帧到达网桥的时间登记下来，就可以在转发表中只保留网络拓扑的最新状态信息。反映当前网络的最新拓扑状态。



网桥的转发表自学习

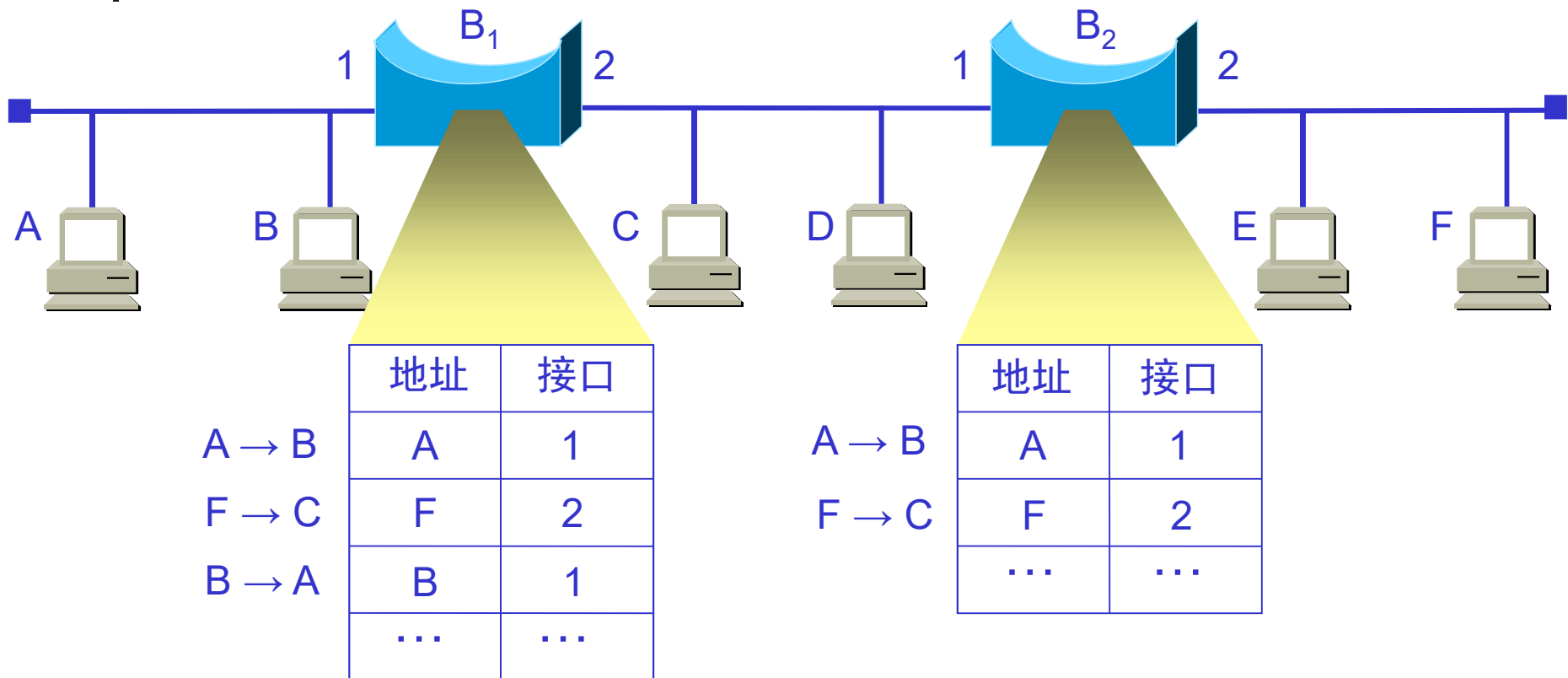
- 网桥收到一帧后先进行自学习。
 - 查找转发表中与收到帧的源地址有无相匹配的项目。
 - 如没有，就在转发表中增加一个项目（源地址、进入的接口和时间）。
 - 如有，则把原有的项目进行更新。
- 每当一帧到达，上述算法都将执行一遍



网桥的帧转发处理

- **转发帧。**查找转发表中与收到帧的目的地址有无相匹配的项目。
 - 如没有，则通过所有其他接口（但进入网桥的接口除外）按进行转发。
 - 如有，则按转发表中给出的接口进行转发。
 - 若转发表中给出的接口就是该帧进入网桥的接口，则应丢弃这个帧（因为这时不需要经过网桥进行转发）。
- 每当一帧到达，上述算法都将执行一遍

转发表的建立过程举例





网桥转发表更新

- 局域网拓扑是变化的：主机的接入和退出
- 网桥怎样适应这种变化？
 - 无论何时，凡往表中加入记录，也必须同时打下时戳
 - 到达帧的源地址在表中已有记录，将时戳更新为当前时间
 - 网桥周期性地扫描表，将那些超时的记录从表中删除



多接口网桥—以太网交换机

- 1990 年问世的**交换式集线器**(switching hub), 可明显地提高局域网的性能。
- 交换式集线器常称为**以太网交换机**(switch)或第二层交换机（表明此交换机工作在数据链路层）。
- 以太网交换机通常都有十几个接口。因此，以太网交换机实质上就是一个**多接口的网桥**，可见交换机工作在数据链路层。



以太网交换机的链路层特点

- 全双工：

- 以太网交换机的每个接口都直接与主机相连，并且一般都工作在全双工方式。

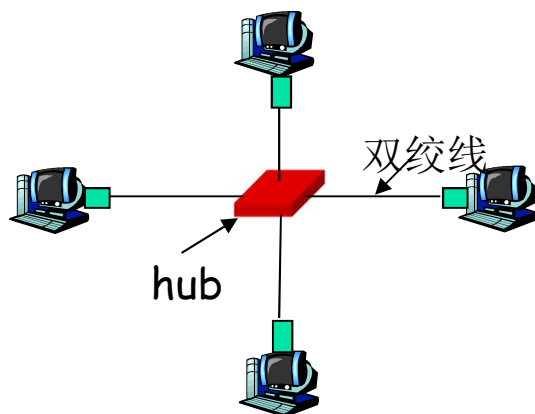
- 无碰撞：

- 交换机能同时连通许多对的接口，使每一对相互通信的主机都能像独占通信媒体那样，进行无碰撞地传输数据。

共享式和交换式以太网

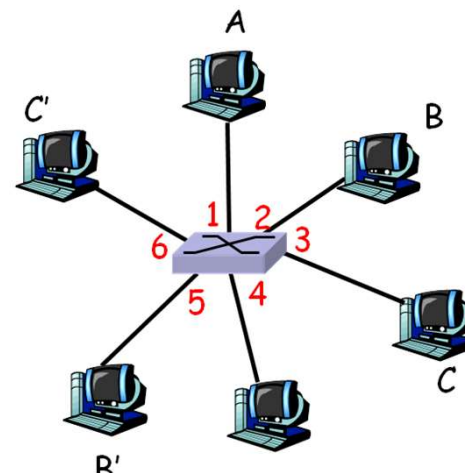
■ 共享式以太网：

- 集线器的所有端口位于同一个冲突域
- 任一时刻最多只允许一个主机发送
- 网络规模（节点数量）与网络性能的矛盾无法解决

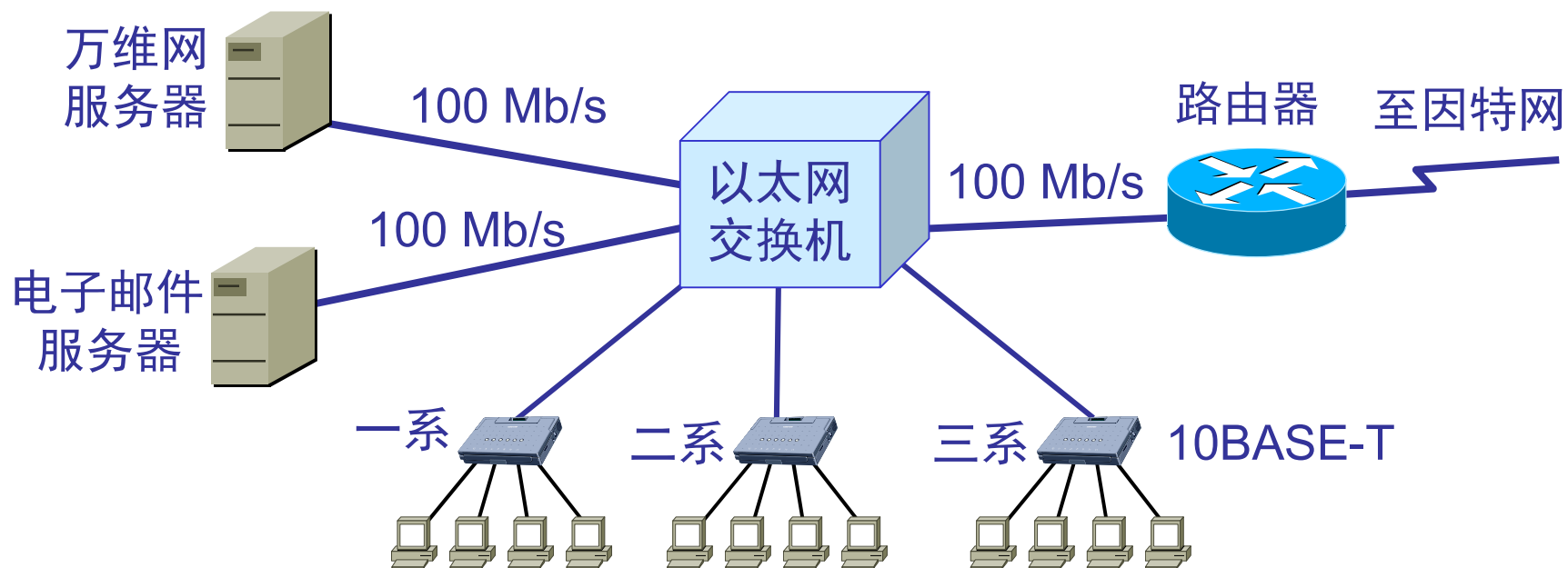


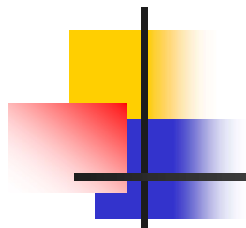
■ 交换式以太网：

- 交换机的每个端口为一个冲突域
- 多对端口可以同时通信
- 网络的集合带宽=各个端口的带宽之和
- 从根本上解决了网络规模与网络性能的矛盾



用以太网交换机扩展局域网





9、虚拟局域网



虚拟局域网（VLAN）

- 在大型机构网络中，管理员通常按部门将用户组织到不同的网络中
- 管理员配置网络遇到的困难：
 - 同一部门的人员在物理位置上可能很分散（他们的主机连接到在不同的交换机上），但是逻辑上，不同的网络可能需要隔离
 - 用大量的路由器来分割网段，成本很高
- 网络设备供应商提供用软件配置网络的方法，导致虚拟局域网概念的提出



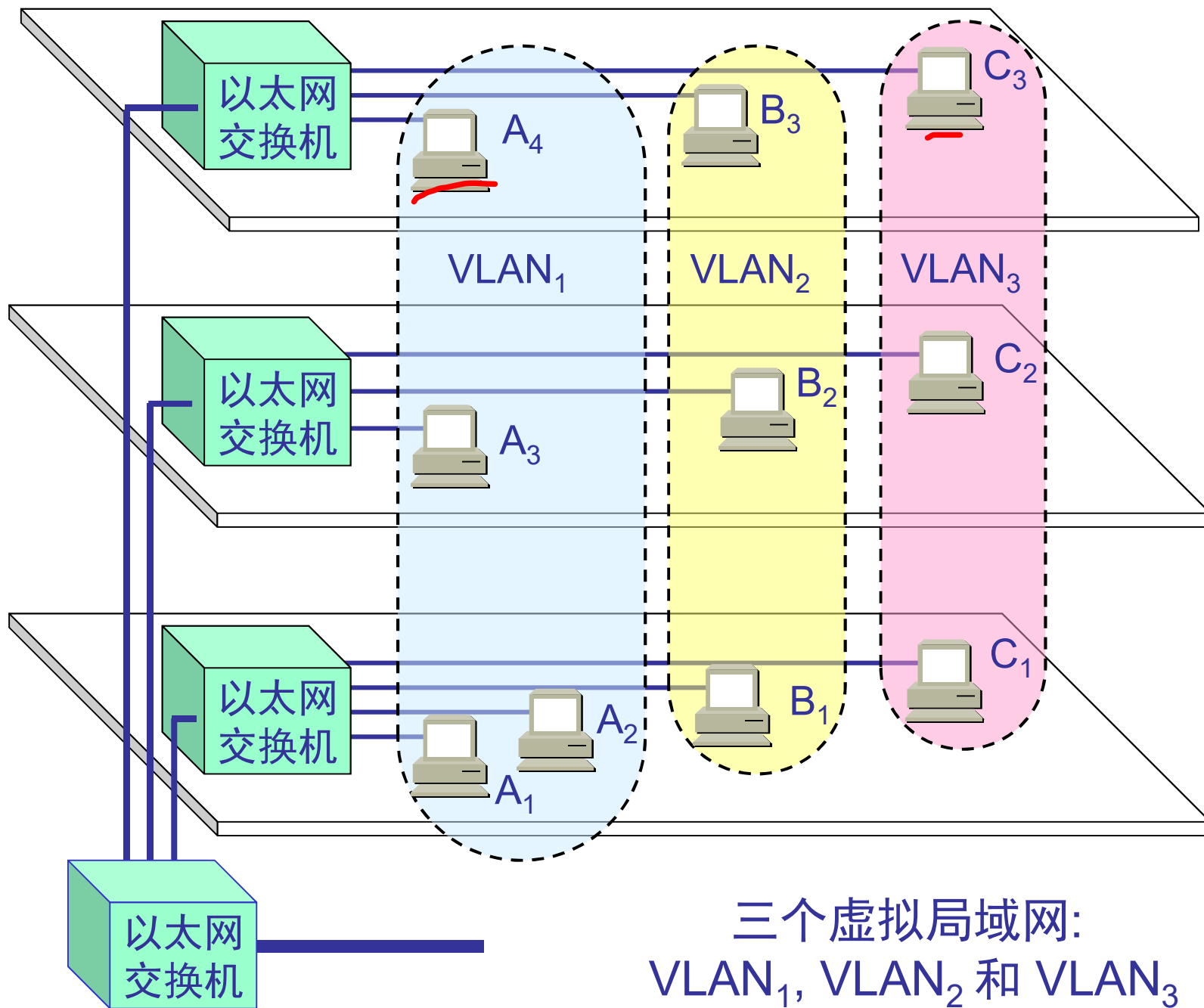
用以太网交换机实现**虚拟局域网**

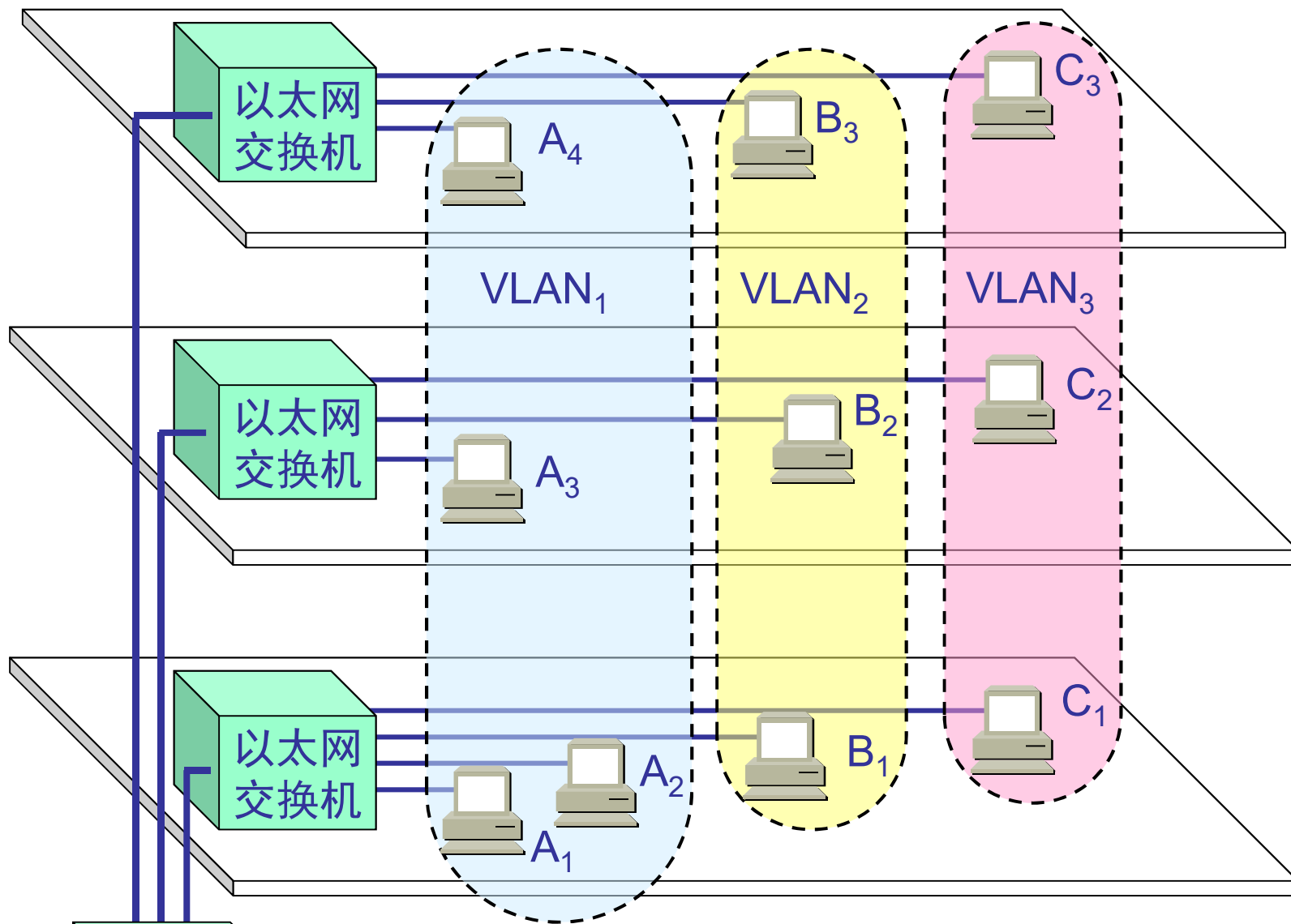
- **虚拟局域网** VLAN 是由一些局域网网段构成的与物理位置无关的逻辑组。
 - 这些网段具有某些共同的需求。
 - 每一个 VLAN 的帧都有一个明确的标识符，指明发送这个帧的工作站是属于哪一个 VLAN。
- 虚拟局域网其实只是局域网给用户提供服务的一种服务，而并不是一种新型局域网。



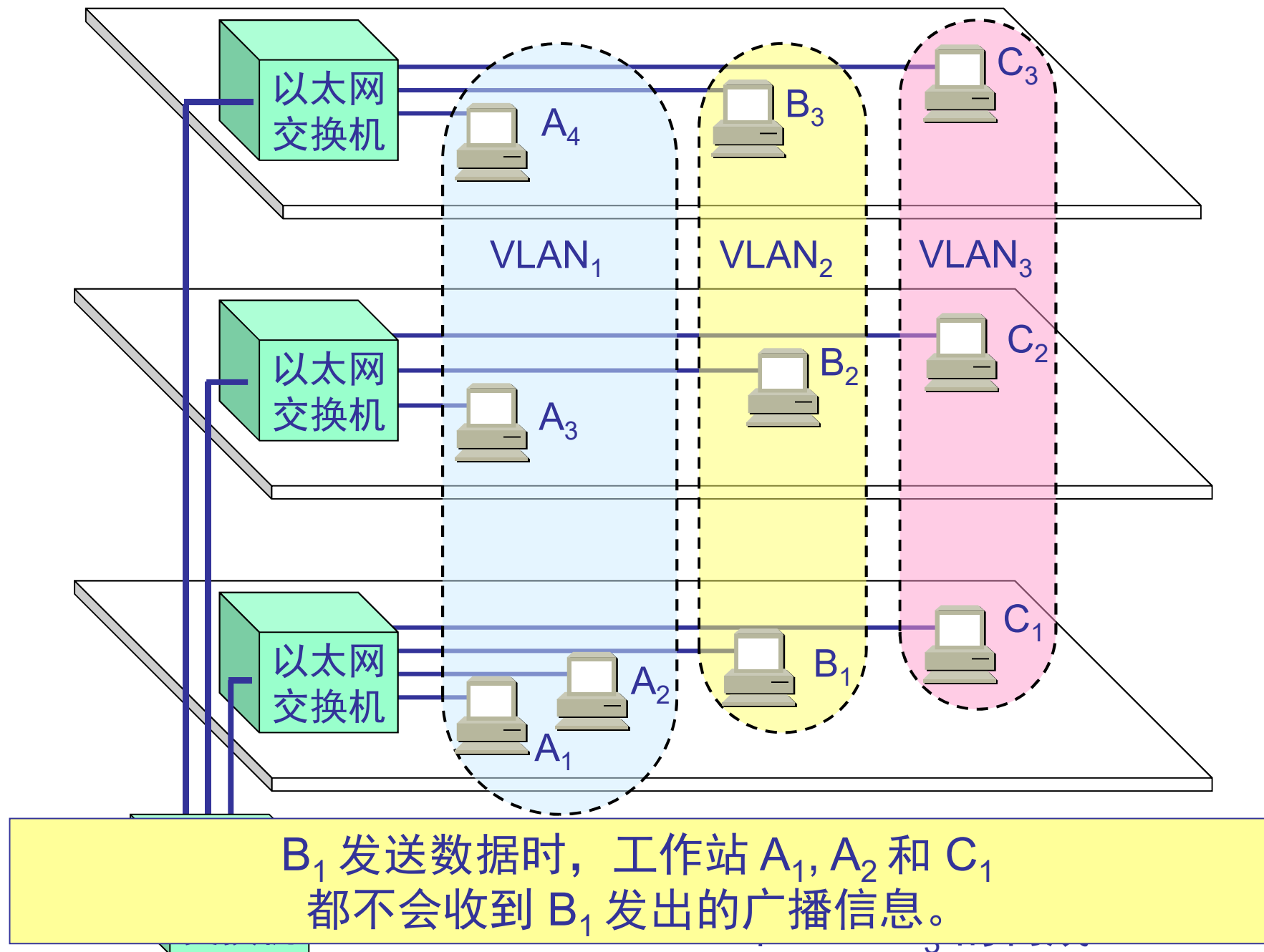
使用VLAN配置网络

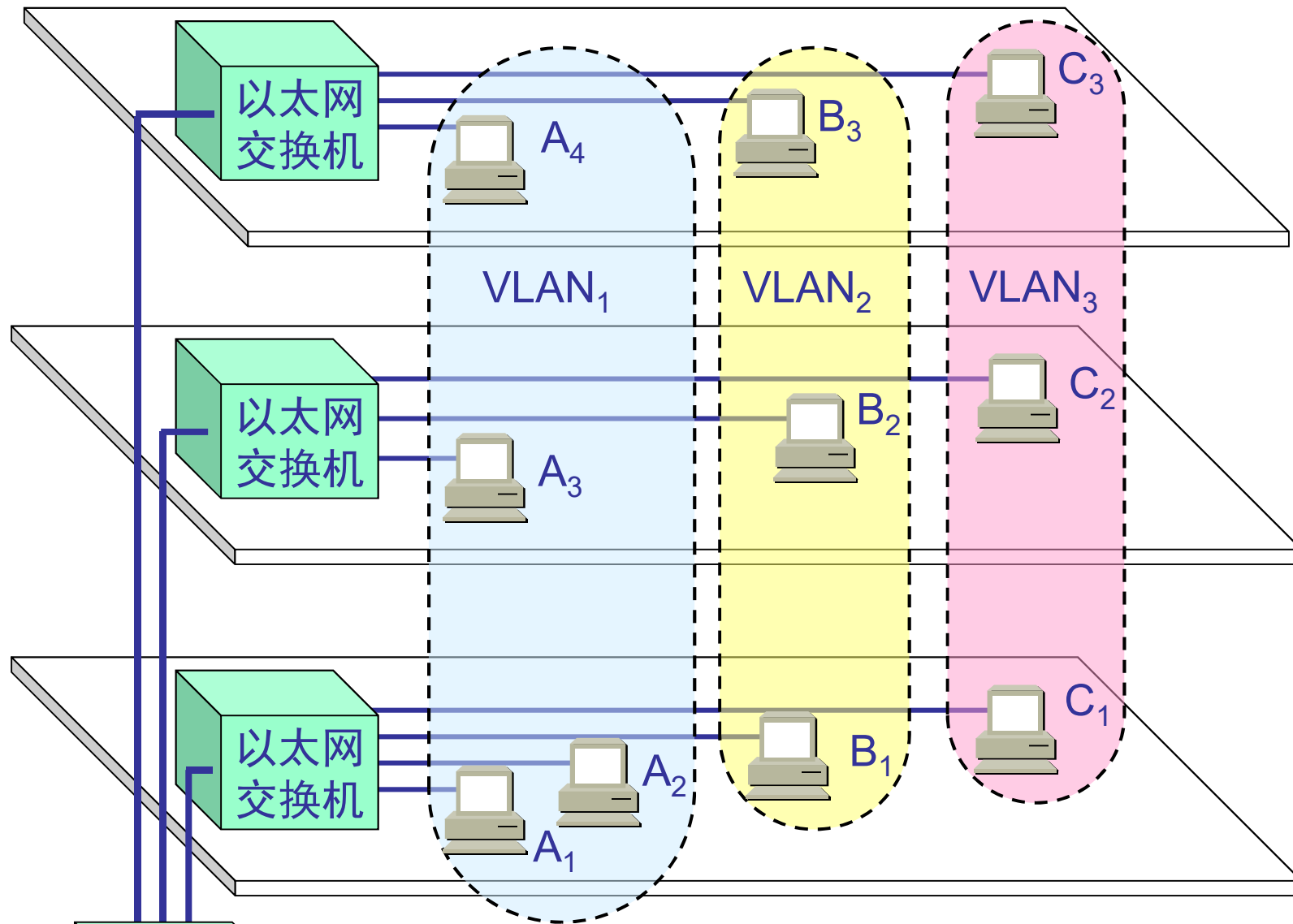
- VLAN的实现基础是支持VLAN功能的交换机
- 管理员配置VLAN:
 - 管理员决定将物理网络划分成几个VLAN、每个VLAN的名字、每个机器在哪个VLAN上
 - 在每个交换机上建立一个配置表，指出通过哪个端口可以到达哪些VLAN的成员（一个交换机端口可以到达多个VLAN的成员）





当 B₁ 向 VLAN₂ 工作组内成员发送数据时，
工作站 B₂ 和 B₃ 将会收到广播的信息。

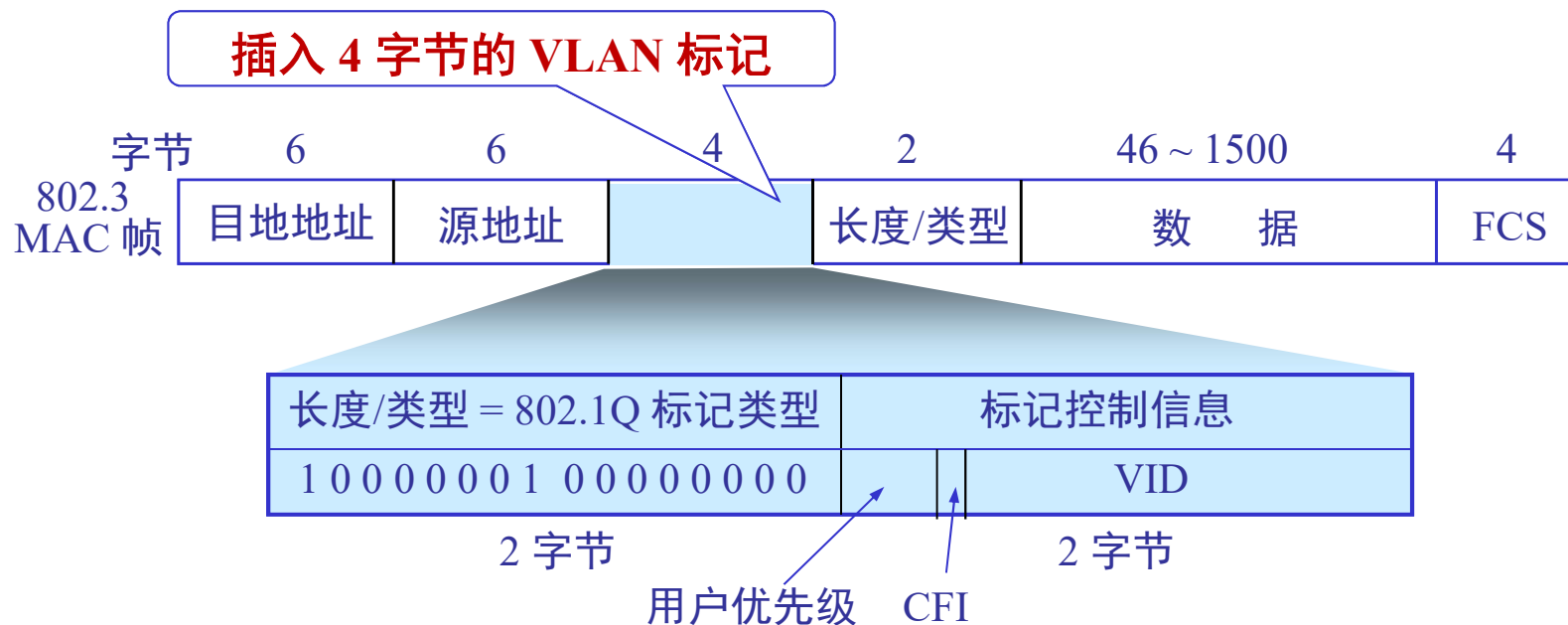




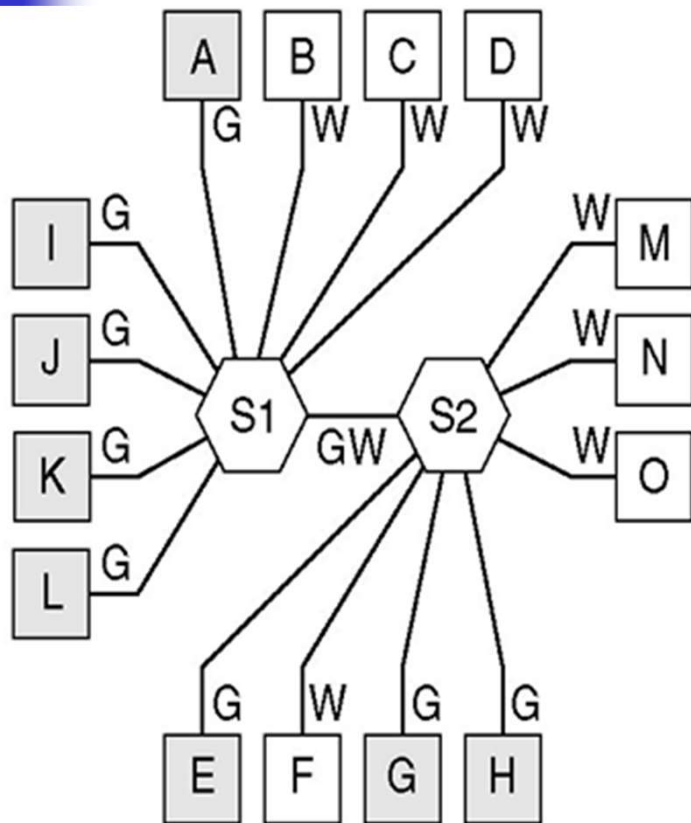
虚拟局域网限制了接收广播信息的工作站数，使得网络不会因传播过多的广播信息（即“广播风暴”）而引起性能恶化。

虚拟局域网的以太网帧格式

- 虚拟局域网协议允许在以太网的帧格式中插入一个 4 字节的标识符，称为 VLAN 标记(tag)，用来指明发送该帧的工作站属于哪一个虚拟局域网。

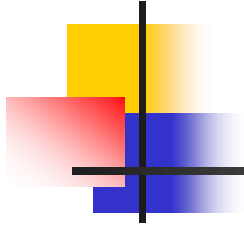


交换机如何在VLAN间转发帧？



标记为**G**的节点属于一个**VLAN**，标记为**W**的节点属于另一个**VLAN**

- 当一个帧到达时，
 - 交换机判断该帧属于哪个**VLAN**
 - 查找配置表得到该VLAN对应的端口
 - 在该VLAN对应的所有端口上转发帧



10、100M、1000M、10G以太网



100BASE-T 以太网

- 速率达到或超过 100 Mb/s 的以太网称为**高速以太网**。
- 在双绞线上传送 100 Mb/s 基带信号的星型拓扑以太网，仍使用 IEEE 802.3 的 CSMA/CD 协议。100BASE-T 以太网又称为**快速以太网**(Fast Ethernet)。



100Mbps以太网（快速以太网）

仅能使用光纤/双绞线，以及集线器/交换机

- 100Base-TX（可使用集线器或交换机）：
 - 5类双绞线（2对），不超过100米
- 100Base-T4（可使用集线器或交换机）：
 - 3类双绞线（4对），不超过100米
- 100Base-FX（只能使用交换机）：
 - 多模光纤（2条），不超过2000米



100BASE-T 以太网的特点

- 可在全双工方式下工作而无冲突发生。
- MAC 帧格式仍然是 802.3 标准规定的。
- 保持最短帧长不变，但将一个网段的最大电缆长度减小到 100 m。
- 帧间时间间隔从原来的 $9.6\ \mu\text{s}$ 改为现在的 $0.96\ \mu\text{s}$ 。



三种不同的物理层标准

- 100BASE-TX
 - 使用 2 对 UTP 5 类线或屏蔽双绞线 STP。
- 100BASE-FX
 - 使用 2 对光纤。
- 100BASE-T4
 - 使用 4 对 UTP 3 类线或 5 类线。



吉比特以太网

- 允许在 1 Gb/s 下全双工和半双工两种方式工作。
- 使用 802.3 协议规定的帧格式。
- 在半双工方式下使用 CSMA/CD 协议（全双工方式不需要使用 CSMA/CD 协议）。
- 与 10BASE-T 和 100BASE-T 技术向后兼容。



吉比特（千兆）以太网

使用交换机，并增加了对流量控制的支持

- 1000Base-SX:

- 多模光纤，不超过550米

- 1000Base-LX:

- 单模或多模光纤，不超过5000米

- 1000Base-CX（很少用）:

- 2对屏蔽双绞线，不超过25米

- 1000Base-T:

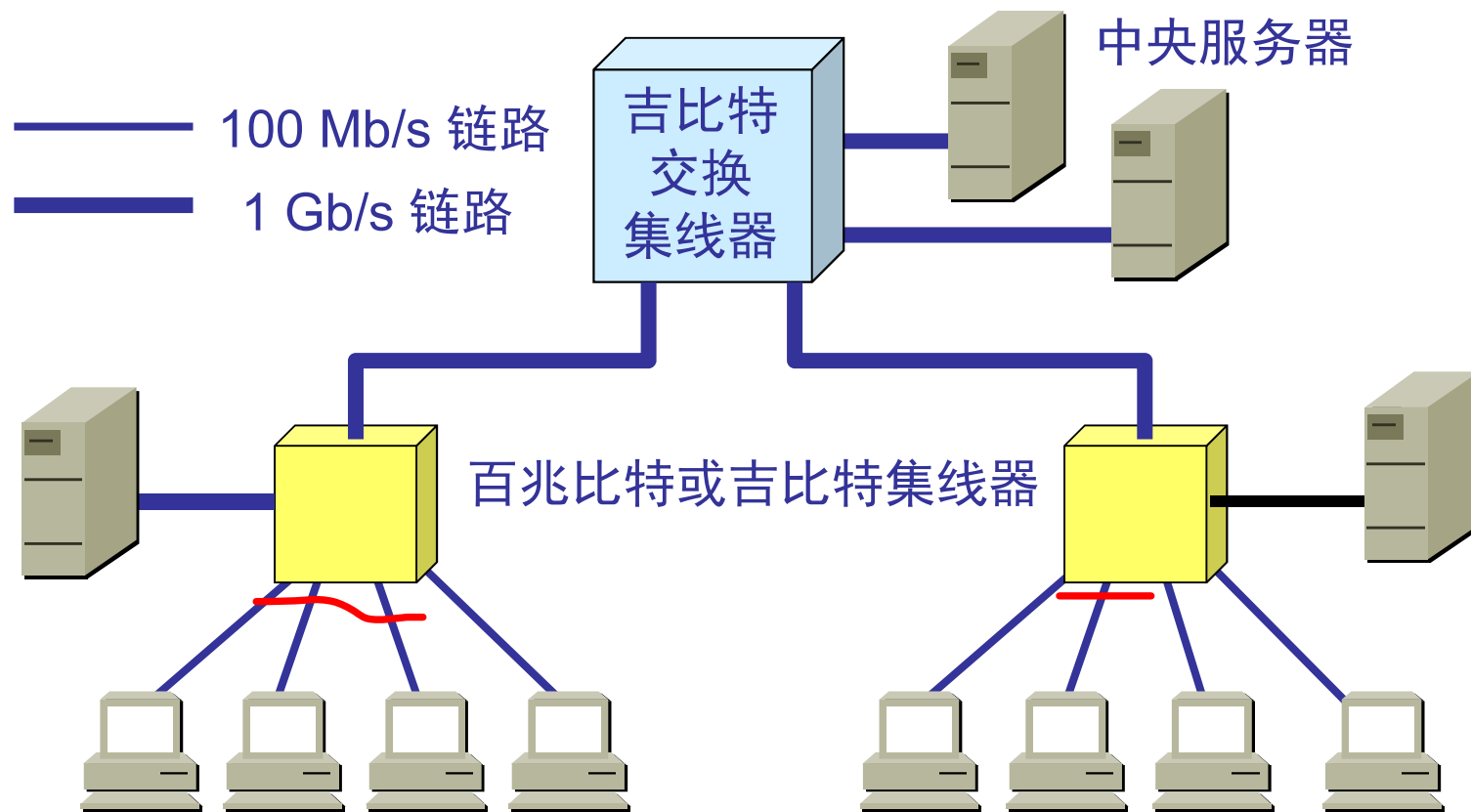
- 4对5类双绞线，不超过100米



全双工方式

- 当吉比特以太网工作在全双工方式时（即通信双方可同时进行发送和接收数据），不使用载波延伸和分组突发。

吉比特以太网的配置举例





10 吉比特以太网

- 10 吉比特以太网与 10 Mb/s, 100 Mb/s 和 1 Gb/s 以太网的帧格式完全相同。
- 10 吉比特以太网还保留了 802.3 标准规定的以太网最小和最大帧长, 便于升级。
- 10 吉比特以太网不再使用铜线而只使用光纤作为传输媒体。
- 10 吉比特以太网只工作在全双工方式, 因此没有争用问题, 也不使用 CSMA/CD 协议。



吉比特以太网的物理层

- 局域网物理层 LAN PHY。局域网物理层的数据率是 10.000 Gb/s。
- 可选的广域网物理层 WAN PHY。广域网物理层具有另一种数据率，这是为了和所谓的“Gb/s”的 SONET/SDH（即OC-192/STM-64）相连接。
 - 为了使 10 吉比特以太网的帧能够插入到 OC-192/STM-64 帧的有效载荷中，就要使用可选的广域网物理层，其数据率为 9.95328 Gb/s。



端到端的以太网传输

- 10 吉比特以太网的出现，以太网的工作范围已经从局域网（校园网、企业网）扩大到城域网和广域网，从而实现了端到端的以太网传输。
- 这种工作方式的好处是：
 - 成熟的技术
 - 互操作性很好
 - 在广域网中使用以太网时价格便宜。
 - 统一的帧格式简化了操作和管理。



使用高速以太网进行宽带接入

- 以太网已成功地把速率提高到 $1 \sim 10 \text{ Gb/s}$ ，所覆盖的地理范围也扩展到了城域网和广域网，因此现在人们正在尝试使用以太网进行宽带接入。
- 以太网接入的重要特点是它可提供双向的宽带通信，并且可根据用户对带宽的需求灵活地进行带宽升级。
- 采用以太网接入可实现端到端的以太网传输，中间不需要再进行帧格式的转换。这就提高了数据的传输效率和降低了传输的成本。

以太网接入举例：光纤到大楼 FTTB

高速汇接点 GigaPoP

