第 3 章　数据链路层

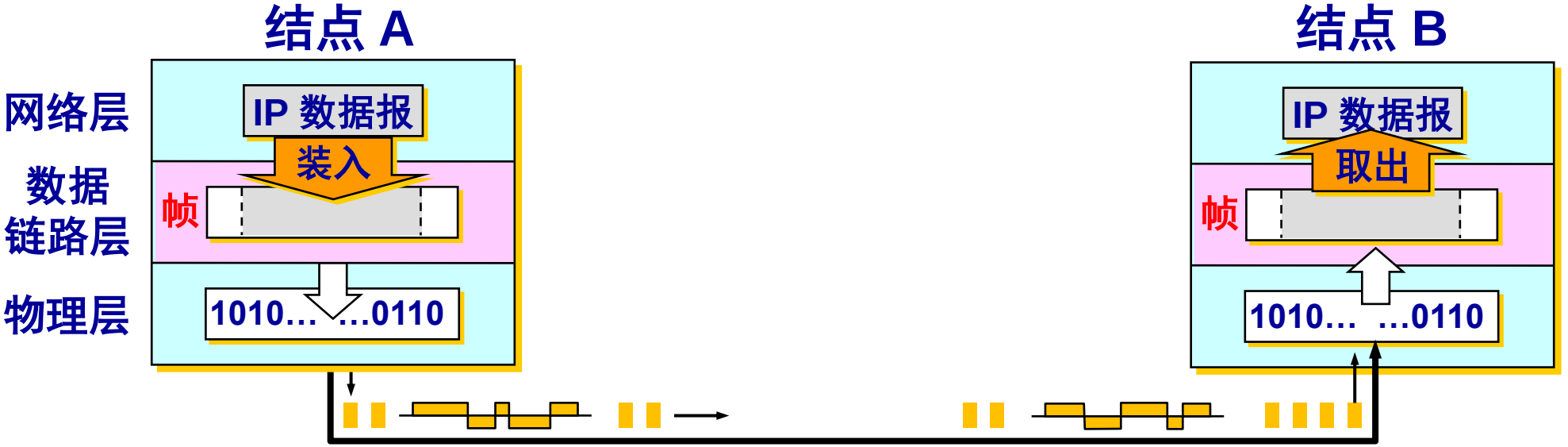
数据链路层的基本概念

- **链路** (link) 是一条无源的结点到**相邻结点**的**物理线路**（有线或无线），中间没有任何其他的交换结点。
  - 一条链路只是一条通路的一个组成部分。
- **数据链路** (data link) 除了物理线路外，还必须有通信协议来控制这些数据的传输。若把实现这些协议的硬件和软件加到链路上，就构成了数据链路。
  - 现在最常用的方法是使用**网络适配器（即网卡）**来实现这些协议的硬件和软件。
  - 一般的适配器都包括了数据链路层和物理层这两层的功能。

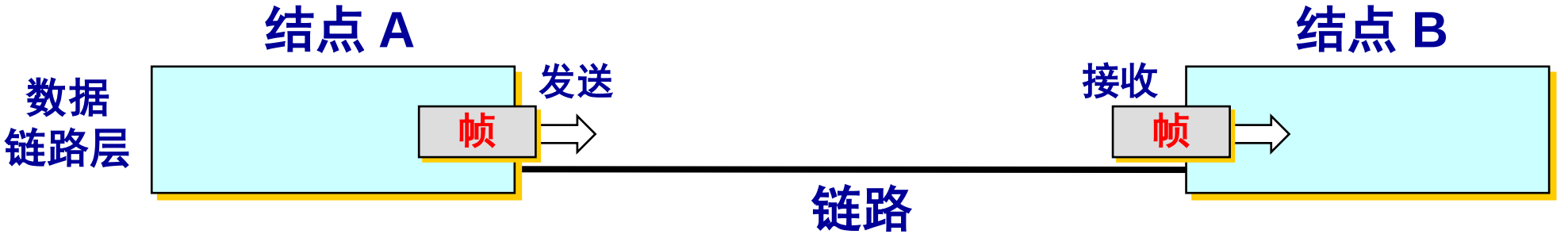
数据链路层的基本概念

- 也有人采用另外的术语。这就是把链路分为物理链路和逻辑链路。
- **物理链路**就是上面所说的**链路**。
- **逻辑链路**就是上面的**数据链路**，是物理链路加上必要的通信协议。
- 早期的数据通信协议曾叫做**通信规程**(procedure)。因此在数据链路层，规程和协议是同义语。

数据链路层的协议数据单元—帧



(a) 三层的简化模型

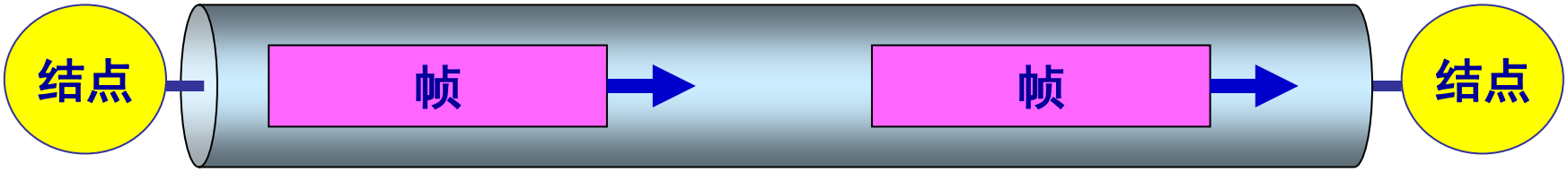


(b) 只考虑数据链路层

使用点对点信道的数据链路层

数据链路层像个数字管道

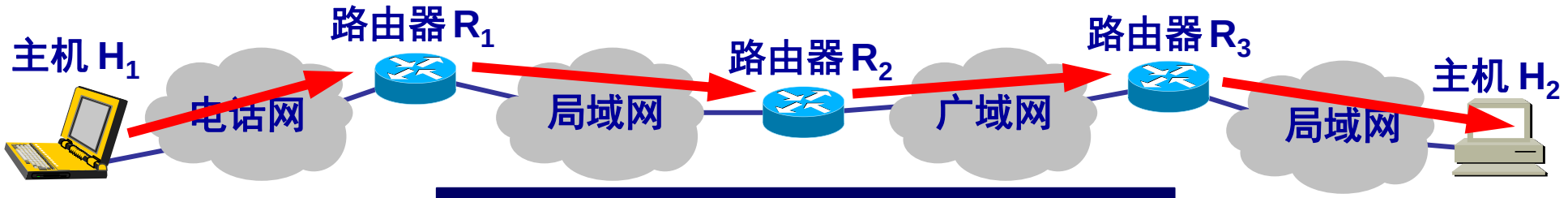
- 常常在两个对等的数据链路层之间画出一个**数字管道**，而在这条数字管道上**传输的数据单位**是**帧**。



- **数据链路层不必考虑物理层如何实现比特传输的细节。**甚至还可以更简单地设想好像是沿着两个数据链路层之间的水平方向把帧直接发送到对方。

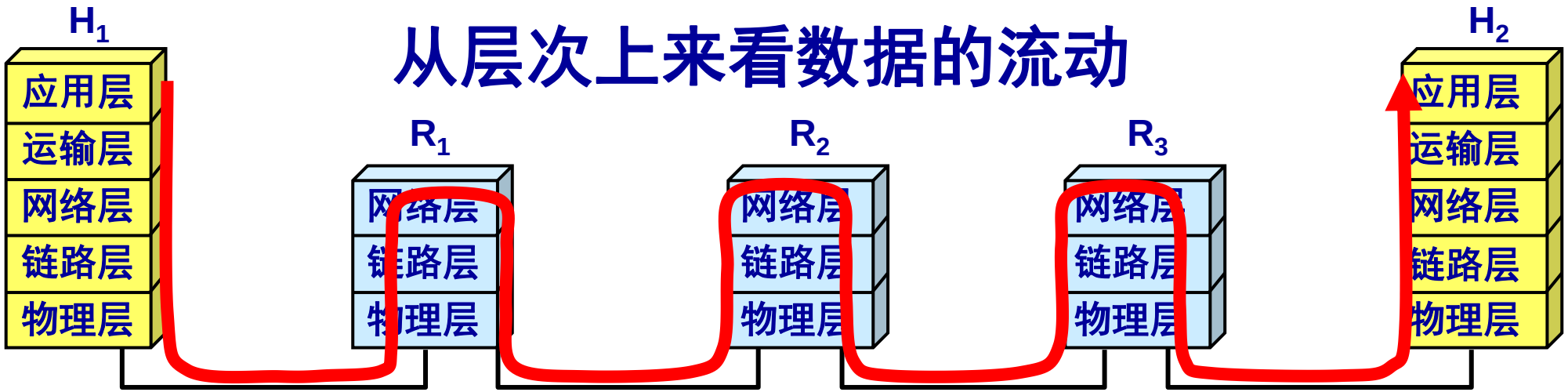
数据链路层的简化模型

主机 H_1 向 H_2 发送数据



H_1 到 H_2 所经过的网络可以是多种的

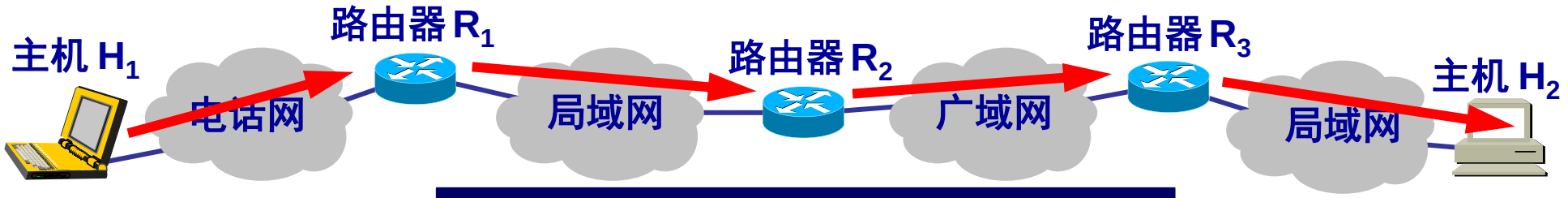
从层次上来看数据的流动



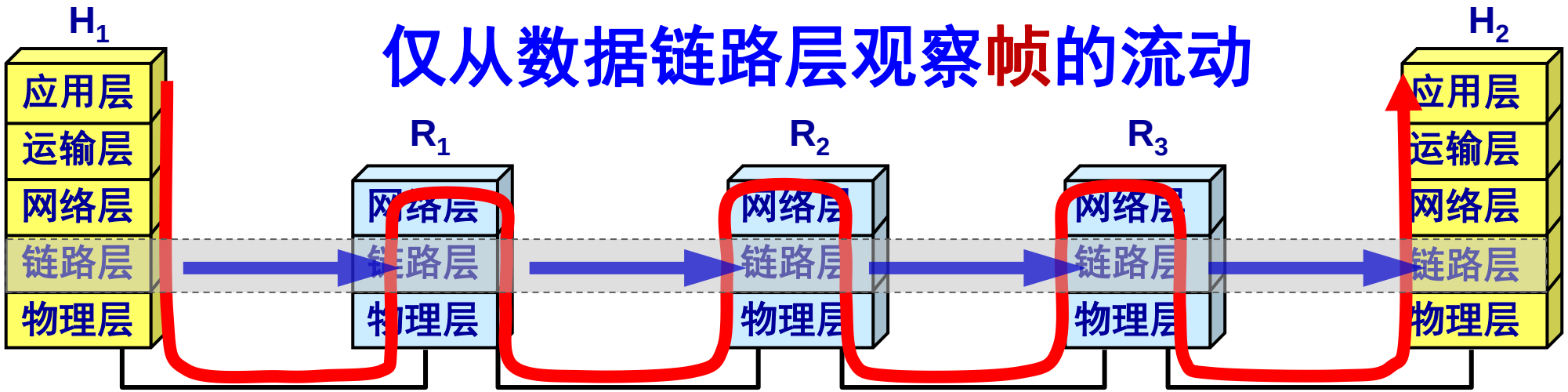
数据链路层的地位

数据链路层的简化模型

主机 H_1 向 H_2 发送数据



H_1 到 H_2 所经过的网络可以是多种的



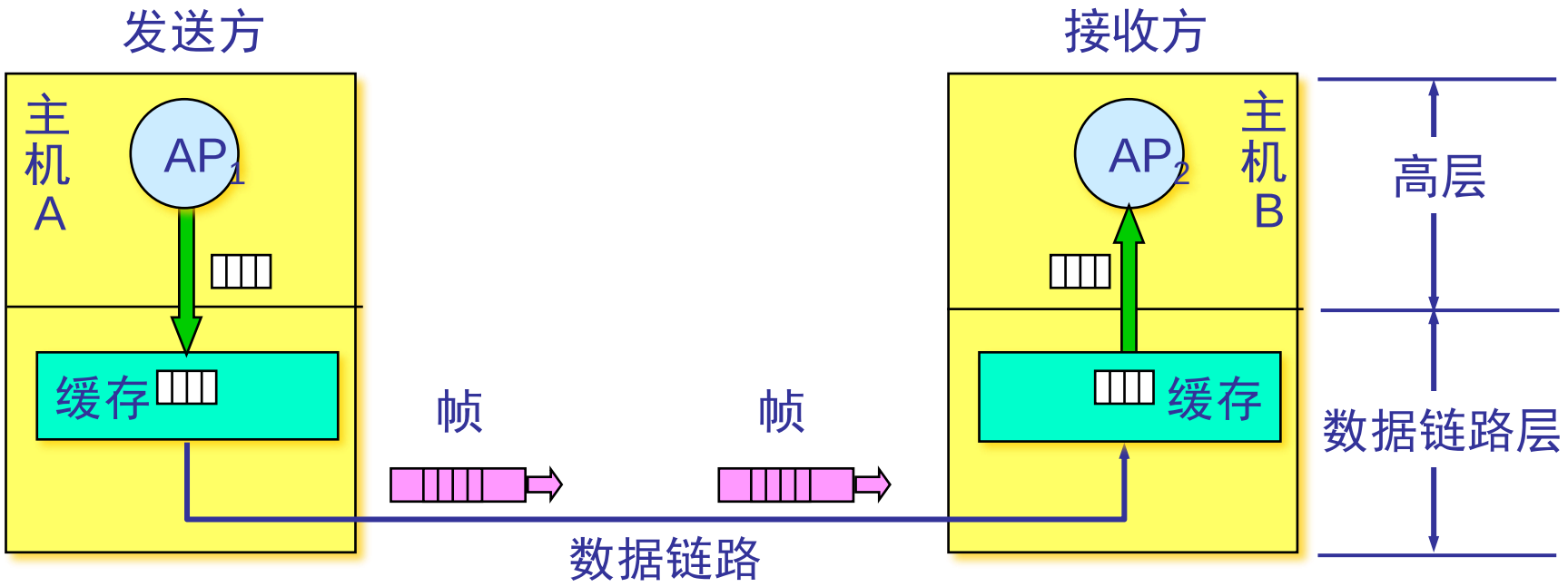
仅从数据链路层观察帧的流动

不同的链路层可能采用不同的数据链路层协议

只考虑数据在数据链路层的流动

数据链路层的简化模型

- 为了分析链路层协议，采用简化的链路层模型
  - 数据链路层以上的各层用一个主机代替；
  - 物理层和通信线路等效成一条简单数据链路；



数据链路层使用的信道

数据链路层使用的信道主要有以下两种类型：

- 点对点信道。这种信道使用一对一的点对点通信方式。（PPP协议）
- 广播信道。这种信道使用一对多的广播通信方式，因此过程比较复杂。广播信道上连接的主机很多，因此必须使用专用的共享信道协议来协调这些主机的数据发送。（CSMA/CD协议）

第 3 章　数据链路层

- 3.1　使用点对点信道的数据链路层
- 3.2　点对点协议 PPP
- 3.3　使用广播信道的数据链路层
- 3.4　以太网

3.1 数据链路层的三个基本问题

- 数据链路层协议有许多种，但有三个基本问题则是共同的。这三个基本问题是：

1. 封装成帧

数据的传送以帧为单位
帧定界

2. 透明传输

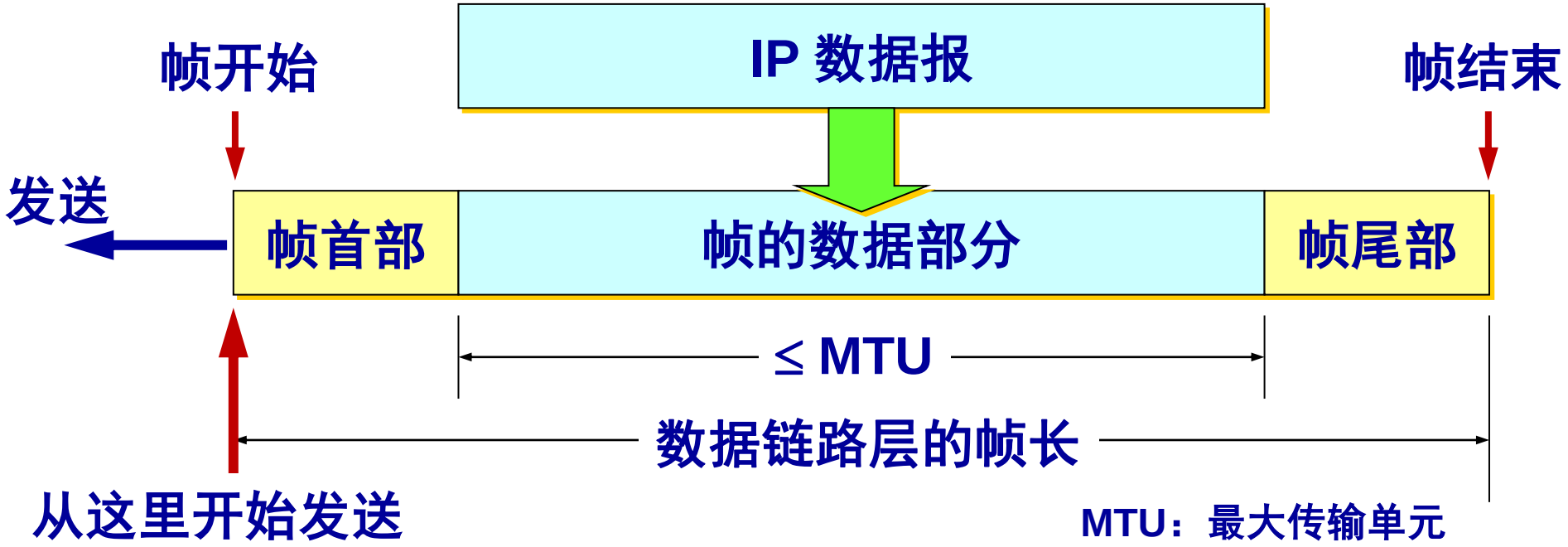
若所传的数据的比特片段与某一个控制信息相同，要有可靠机制，保证收方能正确识别

3. 差错控制

纠错：通过编码技术，接收方自动将差错改正过来
检错：检测出帧有错误，要么忽略或重传

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**，然后就构成了一个帧。使接收方能**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**。



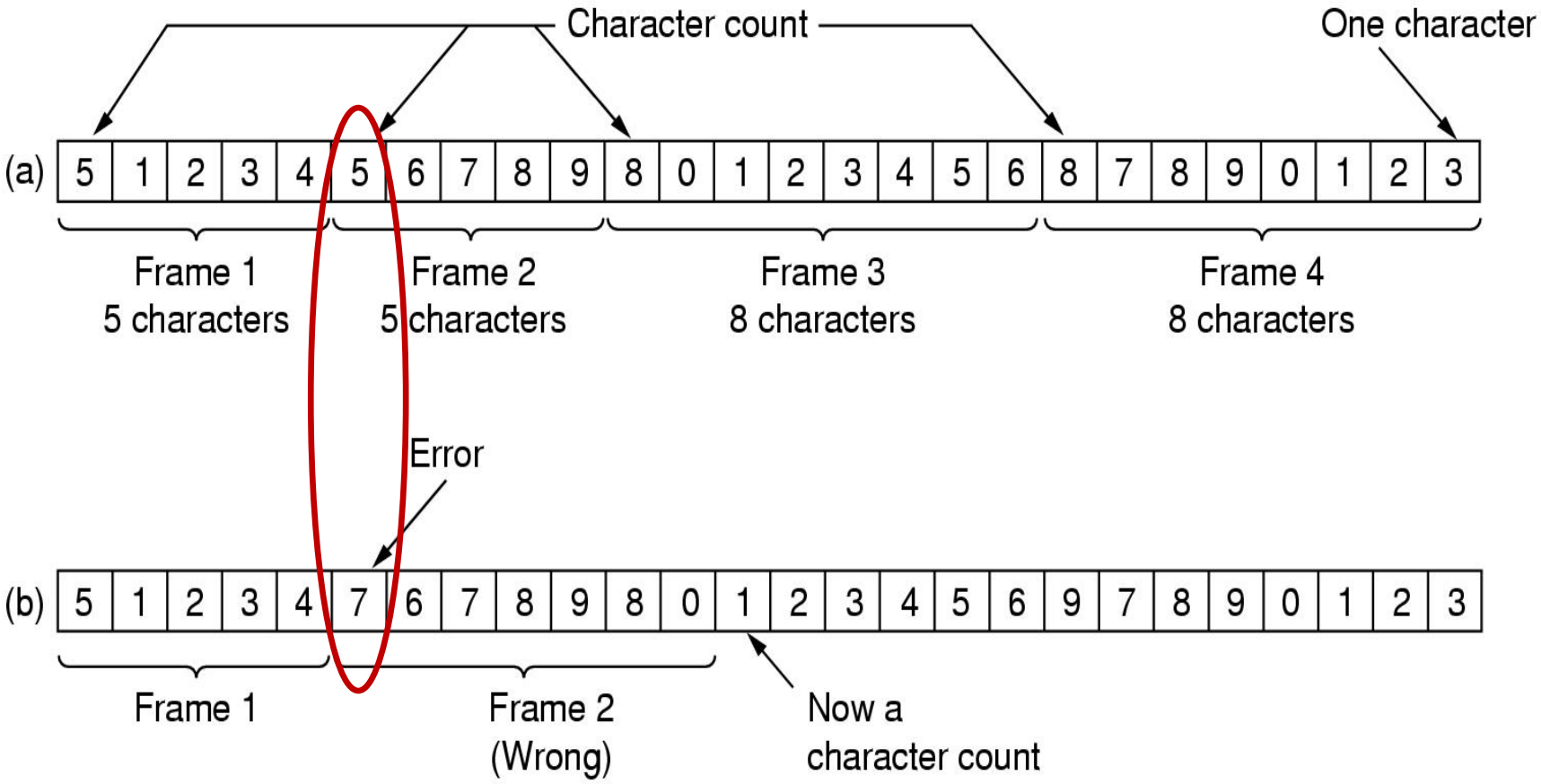
帧定界（帧同步）的方法

- 1、字节计数法
- 2、使用字符填充的首尾定界法
- 3、使用比特填充的首尾定界法
- 4、违法编码法*

1）字节计数法

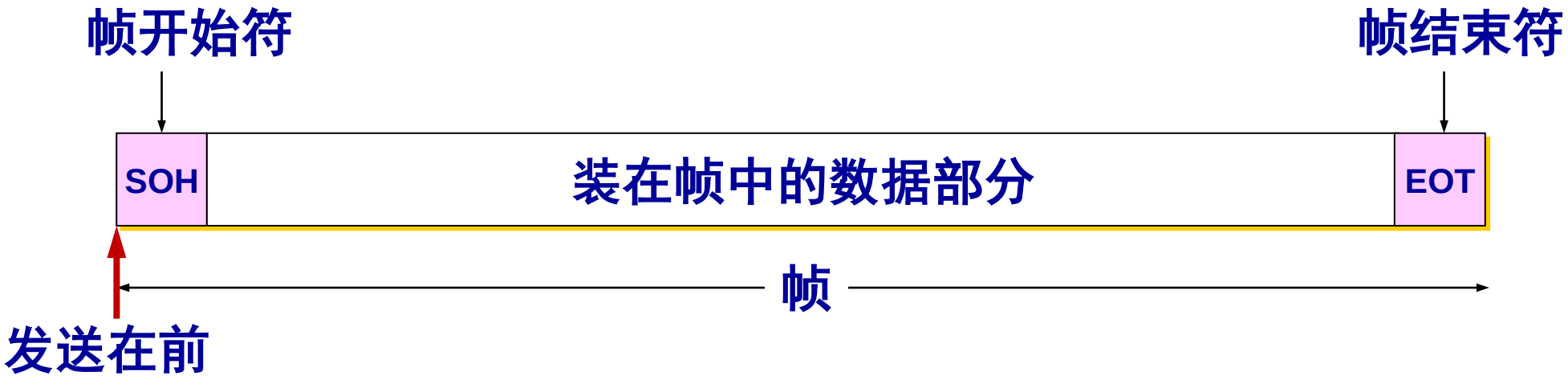
- 思想
  - 在帧头设置一个长度域，放置该帧的字节数，当收方收到帧后，通过帧的长度，确定帧的开始。
- 问题
  - 当帧的长度域出错，帧同步完全丢失；
  - 该方法很少单独使用。

字节计数法举例


Character count
One character
(a)
5 1 2 3 4 5 6 7 8 9 8 0 1 2 3 4 5 6 8 7 8 9 0 1 2 3
Frame 1
5 characters
Frame 2
5 characters
Frame 3
8 characters
Frame 4
8 characters
Error
(b)
5 1 2 3 4 7 6 7 8 9 8 0 1 2 3 4 5 6 9 7 8 9 0 1 2 3
Frame 1
Frame 2
(Wrong)
Now a
character count

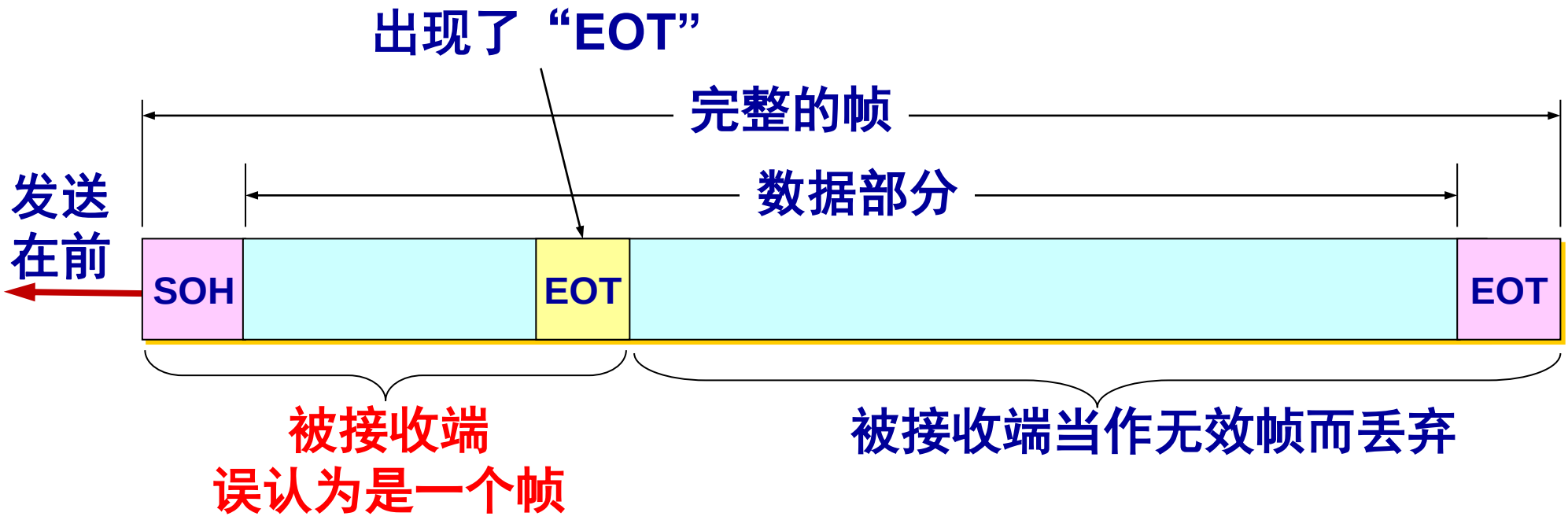
2）字符填充法

- 当数据是由可打印的ASCII码组成的文本文件时，帧定界可以使用特殊的ASCII码（不可打印的控制字符）作为帧定界符。
  - 控制字符 SOH (Start Of Header) 放在一帧的最前面，表示帧的首部开始。另一个控制字符 EOT (End Of Transmission) 表示帧的结束。



2.透明传输

- 如果数据中的某个字节的二进制代码恰好和 SOH 或 EOT 一样，数据链路层就会错误地“找到帧的边界”。

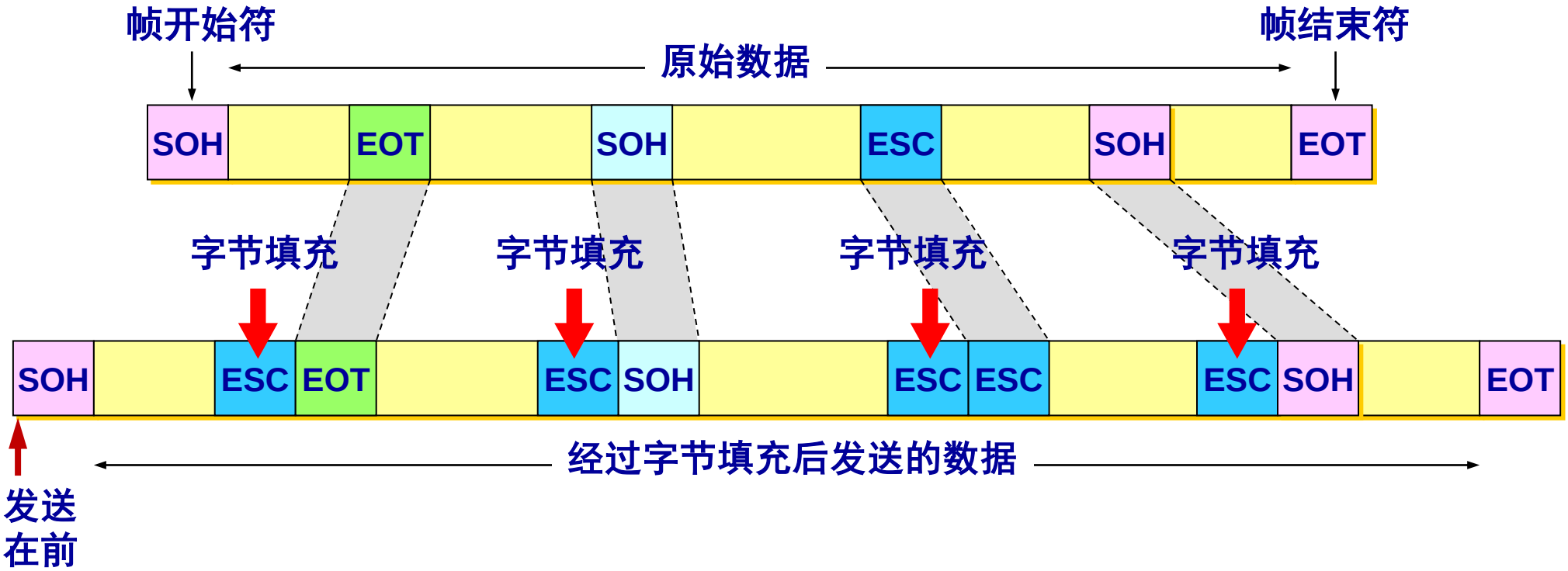


数据部分恰好出现与 EOT 一样的代码

解决方法

- 发送端的数据链路层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)。
- 接收端的数据链路层在将数据送往网络层之前删除插入的转义字符。
- 如果转义字符也出现在数据当中，那么应在转义字符前面插入一个转义字符 ESC。当接收端收到连续的两个转义字符时，就删除其中前面的一个。

用字节填充法解决透明传输的问题



用字节填充法解决透明传输的问题

3）比特填充法

- 思想
  - 使用一个特殊的比特模式01111110作为帧的起始和结束标志。
  - 发送方边发送边检查数据，每连续发送5个“1”后在后面自动插入一个“0”。这样数据中只会连续出现5个“1”，而不会出现定界符。
  - 接收方在收到5个连续的“1”后将后面的“0”删掉而恢复出原始数据。
- 好处
  - 数据传输的基本单位是比特而不是字符，可用来传输任意长度的二进制比特串，通用性强。

零比特填充

信息字段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 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

零比特的填充与删除

4）违法编码法

- 前提
  - 物理介质上使用的信号编码有冗余码字时，使用这些冗余的码字来作为帧的定界。
- 举例
  - 如曼彻斯特编码或差分曼彻斯特编码中，有效电平是“低－高”或“高－低”，而“低－低”和“高－高”电平没有定义，这种违法编码可以作为帧的边界。

3. 差错检测

- 在传输过程中可能会产生**比特差错**：1 可能会变成 0 而 0 也可能变成 1。
- 在一段时间内，传输错误的比特占所传输比特总数的比率称为**误码率** BER (Bit Error Rate)。
- **误码率与信噪比有很大的关系。**
- 为了保证数据传输的可靠性，在计算机网络传输数据时，必须采用各种**差错检测**措施。

差错控制

- 差错编码技术：如何发现差错？
  - 检错码（奇偶校验码、CRC）
  - 纠错码（海明码）

数据块中插入冗余信息的过程，使得数据块中的各个比特建立某种形式的关联，接收端通过验证这种关联关系来判断是否有传输错误

能检测出错误，但不能纠正错误

能知道错误，且知道错误的位置

- 差错控制技术：发现差错如何处理？
  - 前向纠错

由接收方来检查并纠正错误

  - 自动重发请求

不能纠正，接收方反馈。若有错误则重发，否则给肯定应答

差错控制技术（一）

- 前向纠错（FEC，Forward Error Correct）
  - 即发送方发送能使接收方检错并纠错的冗余位，纠错任务由接收方完成；常采用海明码。
  - 主要应用于没有反向信道或反向传输时间很长的场合
- 缺点：为纠错附加的冗余码较多，传输效率低；
- 优点：实时性好。

差错控制技术（二）

- 自动重发请求(ARQ – Automatic Repeat reQuest)
  - 即发送方发送能使接收方检错的冗余位，若无差错，则接收方回送一个肯定应答(ACK)；若有差错，则接收方回送一个否定应答(NAK)，要求发送方重发。
- 缺点：信息传递连贯性差
- 优点：接收端设备简单，只要请求重发，无需纠正错误。

检错码

- 检错码的构造
  - 检错码(码字、传输帧) ＝ 信息位＋冗余校验位
  - 码字长n = K(信息位位数)＋r(校验位位数)
  - 编码效率R ＝ 有效数据位K / 码字长n
- 信息字段 和 校验字段 之间的对应关系
  - 校验字段越长，编码的检错能力越强，编码/解码越复杂；附加的冗余信息在整个编码中所占的比例越大，传输的有效成分越低，传输的效率下降。
  - 检错码一旦形成，整个检错码将作为一个整体被发往线路，通常的发送顺序是信息字段在前，校验字段在后。

1）奇/偶校验码

- 奇校验：使码字中“1”的总个数为奇数。
- 偶校验：使码字中“1”的总个数为偶数。
- 奇/偶校验码：最常用的一种检验码，包括：
  - 水平奇/偶校验码
  - 垂直奇/偶校验码
  - 水平垂直奇/偶校验码

水平奇/偶校验

- 其信息字段以字符为单位，校验字段仅含一个比特称为校验比特或校验位。

- 例如：
  - 使用七比特的ASCII码来构造成八比特的检错码时，
  - 若采用奇/偶校验，校验位的取值应使整个码字包括校验位，1的比特个数为奇数或偶数。

水平奇/偶校验

- 例： 信息字段　　奇校验码　　偶校验码

  0110001　01100010　01100011

- 编码效率： Q/(Q+1) (信息字段占Q个比特)
- 应用： 通常在异步传输方式中采用偶校验， 同步传输方式中采取奇校验。

- 偶校验位的产生直接对发送数据依次做异或运算就可以得到，
- 而产生奇校验位还要在偶校验电路的输出上取非，
- 相对而言，产生奇校验位的代价高，速度也相对慢（慢一个逻辑门的时延）

垂直奇/偶校验

- 被传输的信息进行分组，并排列为若干行和若干列。组中每行的相同列进行奇/偶校验，最终产生由校验位形成的校验字符（校验行），并附加在信息分组之后传输
- 举例：4个字符（4行）组成一信息组

信息组	0	1	1	1	0	0	1
	0	0	1	0	1	0	1
	0	1	0	1	0	1	1
	1	0	1	0	1	0	1
奇校验	0	1	0	1	1	0	1
偶校验	1	0	1	0	0	1	0

发往线路顺序（垂直奇校验）
0111001|0010101|0101011|10
10101|0101101

- 编码效率： PQ/P(Q+1) （假设信息分组占Q行P列）

水平垂直奇/偶校验

- 发往线路顺序(偶校验字符)：
  01110010|00101011|01010110|10101010|10100101
- 第1字符 | 第2字符 |第3字符 | 第4字符 |偶校验字符
- 编码效率：
  - PQ/(P+1)(Q+1) （假设被传信息分组占Q行P列）

2）循环冗余码（CRC）
(Cyclic Redundancy Check)

- 在发送端，先把数据划分为组
- 假定每组 ***k*** 个比特
- 假设待传送的一组数据 ***M*** = 101001（现在 ***k*** = 6）。我们在 ***M*** 的后面再添加供差错检测用的 ***n*** 位冗余码一起发送。

冗余码的计算方法

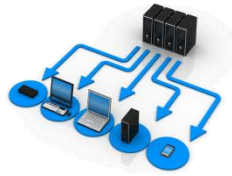
- 用二进制的模2运算进行 2^n 乘 M 的运算，这相当于在 M 后面添加 n 个 0。
- 得到的 $(k + n)$ bit 的数除以事先选定好的长度为 $(n + 1)$ bit 的除数 P，得出商是 Q 而余数是 R，余数 R 比除数 P 少 1 位，即 R 是 n 位。
- 将余数 R 作为冗余码（帧检验序列FCS (Frame Check Sequence)）拼接在数据 M 后面发送出去。

模2运算

- 用模2运算进行加法时不进位。
- 减法和加法一样，按加法规则进行运算。

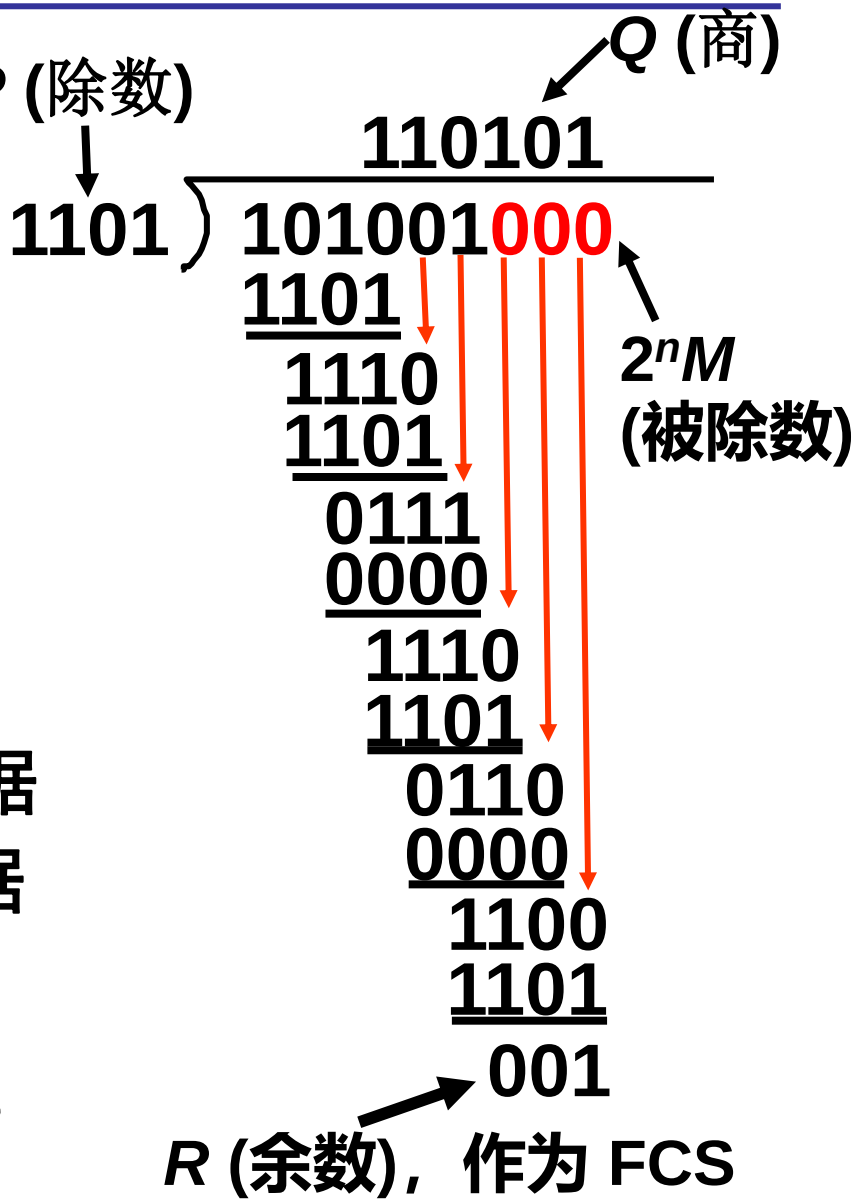
举例: 1111＋1010＝？答案为0101

冗余码的计算举例



- 现在 $k = 6$, $M = 101001$。
- 设 $n = 3$, 除数 $P = 1101$,
- 被除数是 $2^n * M = 101001000$。
- 模 2 运算的结果是：
  - 商 $Q = 110101$,
  - 余数 $R = 001$。
- 把余数 R 作为冗余码添加在数据 M 的后面发送出去。发送的数据是：$2^nM + R$ （即 MR）

即：101001001，共 $(k + n)$ 位。



帧检验序列 FCS

- 在数据后面添加上的冗余码称为帧检验序列 FCS (Frame Check Sequence)。
- 循环冗余检验 CRC 和帧检验序列 FCS 并不等同。
  - CRC 是一种常用的检错方法，而 FCS 是添加在数据后面的冗余码。
  - FCS 可以用 CRC 这种方法得出，但 CRC 并非用来获得 FCS 的唯一方法。

接收端对收到的每一帧进行 CRC 检验

- 把收到的每一个帧都除以相同的除数*P*（模2运算），然后检查得到的余数*R*
- (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。
- 漏检：CRC不能保证检测出所有的传输错误，但是只要选择位数足够的P，可以使得差错的概率足够小。

CRC也称多项式编码

- 任意一个由二进制位串组成的代码都可以和一个系数仅为‘0’和‘1’取值的多项式一一对应。

- 多项式表示：即将k比特的数据用k项多项式表示，它的各项为X^{k-1}，…，X^0，它的系数为数据中对应位的0或1。
- 例如：
  - 代码1010111对应的多项式为$x^6+x^4+x^2+x+1$
  - 多项式为$x^5+x^3+x^2+x+1$对应的代码101111

生成多项式P

- 除数 *P* 可表示成生成多项式 *P(X)*
  - 例如：*P*＝110101，即$P(X)=X^5+X^4+X^2+1(X^0)$（P为5阶多项式）；
  - 生成多项式的最高位和最低位都必须为1；
  - 若 *P* 为 *r* 阶（ *r*+1 bit），将产生 *r* 位冗余位；

发送端帧检验序列FCS的生成和接收端CRC检验都是用硬件完成的，处理速度很快，不会延误数据的传输

- ✓ *M(X)* -------- 信息多项式
- ✓ *R(X)* -------- 冗余多项式
- ✓ *T(X)* -------- 传输帧多项式

P的确定

- P为生成多项式，已有的国际标准。
  - CRC-12 = $X^{12}+X^{11}+X^3+X^2+X^1+1$
  - CRC-16 = $X^{16}+X^{15}+X^2+1$
  - CRC-CCITT = $X^{16}+X^{12}+X^5+1$
    - HDLC和X.25采用
  - CRC32 =

    $X^{32}+X^{26}+X^{23}+X^{22}+X^{16}+X^{12}+X^{11}+X^{10}+X^8+X^7+X^5+X^4+X^2+X+1$
    - CSMA/CD LAN采用

应当注意

- 仅用循环冗余检验 CRC 差错检测技术只能做到**无差错接受** (accept)。
- **“无差错接受”是指：**“凡是接受的帧（即不包括丢弃的帧），我们都能以**非常接近于1的概率**认为这些帧在传输过程中没有产生差错”。
- 也就是说：“凡是接收端数据链路层接受的帧都无差错”（**有差错的帧就丢弃而不接受**）。

- 区分**“无比特差错”**与**“无传输差错”**(在运输层实现)
- 要做到**“可靠传输”（即发送什么就收到什么）**就必须再加上**确认和重传机制。**

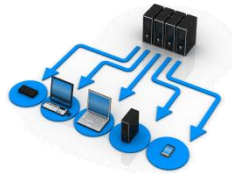
应当注意

- 传输差错 [#1]-[#2]-[#3]
  - 帧丢失 [#1]-[#3]
  - 帧重复 [#1]-[#2]-[#2]-[#3]
  - 帧失序 [#1]-[#3]-[#2]
- 可靠传输
  - 帧编号
  - 确认：收到正确帧要发送确认
  - 重传机制：一定期限内，发送端未收到确认，则重传
- 在数据链路层使用 CRC 检验，能够实现无比特差错的传输，但这还不是可靠传输。
- 本章介绍的数据链路层协议都不是可靠传输的协议
- 对于通信质量较差的无线传输链路，数据链路层协议使用确认和重传机制可以提高通信效率

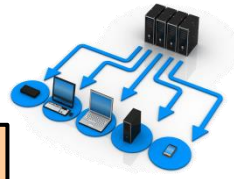
3）海明码（Hamming code）

- 海明码是R.Hamming在1959年提出的，基本思想是：
  - 在 *k* 比特信息 上附加 *r* 比特冗余信息（校验比特），构成 *n=k+r* 比特的码字，其中 每个校验比特 和 某几个特定的信息比特 构成 偶校验 关系。
  - 接收端对这 *r* 个偶校验关系进行校验，即将每个校验比特和与它关联的信息比特进行相加（异或），相加的结果为校正因子。
    - 如果没有错，则 *r* 个校正因子都为0；
    - 若校正因子不全为0，根据校正因子的取值，确定错误发生的位置。
- 主要介绍单比特纠错海明码

码距



- 码距（海明距离Hamming Distance）
  - 一个编码系统中任意两个合法编码（码字）之间不同的二进位（bit）数叫这两个码字的码距。
  - 而整个编码系统中任意两个码字的的最小距离就是该编码系统的码距。
- 两个结论
  - 如果要检测出d个比特的错，则编码集的海明距离至少为d+1。
  - 如果要纠正 d个比特的错，则编码集的海明距离至少应为2d+1。



000(晴)
001(雪)
010(霜)
011(多云)
100(雾)
101(阴)
110(雨)
111(雹)

如果任一比特出错，接收端无法发现

000(晴)
001(不可用)
010(不可用)
011(多云)
100(不可用)
101(阴)
110(雨)
111(不可用)

接收端可以发现1个比特的传输差错

000(晴)
001(不可用)
010(不可用)
011(不可用)
100(不可用)
101(不可用)
110(不可用)
111(雨)

可以发现2个比特的差错，纠正1个比特的差错

发送方冗余位计算

- A、根据信息位长度(如每帧K位)，计算出所需冗余位位数r：
  - 求海明码时的一项基本考虑是确定所需最少校验位数 *r*，若需纠正1bit错，需满足： $2^r \geq K+r+1$
    - 考虑长度为 K 位的信息，若附加了 r 个校验位，则所发送信息的总长度为 K+r
    - 在接收端中要进行 r 个奇偶检查，每个检查结果或是真或是伪。这个奇偶检查的结果确定最多 2^r 种不同状态
    - 这些状态中必有一个是判定信息正确的条件
    - 剩下的（2r - 1）种状态，可以用来判定误码的位置
    - 则导出关系： $2^r-1 \geq K+r$
  - 例如：如果K＝4，则r=3，则n=K+r=7

某公司笔试题

- 实验室里有1000个一模一样的瓶子，但是其中的一瓶有毒。可以用实验室的小白鼠来测试哪一瓶是毒药。如果小白鼠喝掉毒药的话，会在一个星期的时候死去，其他瓶子里的药水没有任何副作用。请问最少用多少只小白鼠可以查出哪瓶是毒药？

  a. 1　　b. 10　　c. 999

- 请给出正确答案，并解释原因。

海明码计算

- **B、确定 校验比特 和 信息比特 的位置**
  - 理论上校验比特可在任何位置，但习惯都是将校验比特放在1、2、4、8、16…位置上。
  - 通常是将2^k位置上，放校验位，其余位置放信息位。
  - 例：7 bit 的二进制数

要传输的比特流：(高位)0110101(低位)

（低位）　　（高位）

位置	1	2	3	4	5	6	7	8	9	10	11
内容	x1	x2	1	x3	0	1	0	x4	1	1	0
	R_0	R_1	I_1	R_2	I_2	I_3	I_4	R_3	I_5	I_6	I_7

海明码的计算

恰好是校验位所在位置

- 将每个信息比特的位置写成2的次幂之和的形式有：
  - I_4：$7=2^2+2^1+2^0$（说明I_4参与R_2、R_1和R_0的生成）
  - I_3：$6=2^2+2^1$（说明I_3参与R_2、R_1的生成）
  - I_2：$5=2^2+2^0$（说明I_2参与R_2、R_0的生成）
  - I_1：$3=2^1+2^0$（说明I_1参与R_1、R_0的生成）

目的：计算每一个信息位与哪些校验位有关联

位置	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011
内容	x1	x2	1	x3	0	1	0	x4	1	1	0
	R_0	R_1	I_1	R_2	I_2	I_3	I_4	R_3	I_5	I_6	I_7

海明码计算

- 从另一个方面说：
  - R_1参与校验I_1、I_3、I_4、I_6、I_7，
  - 即R_1和I_1、I_3、I_4、I_6、I_7构成偶校验关系
  - 这样可以写成如下比特计算公式（XOR运算）：
    - $R_1 = I_1 \oplus I_3 \oplus I_4 \oplus I_6 \oplus I_7$

异或也叫半加运算，其运算法则相当于不带进位的二进制加法

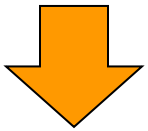
位置	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011
内容	x1	x2	1	x3	0	1	0	x4	1	1	0
	R_0	R_1	I_1	R_2	I_2	I_3	I_4	R_3	I_5	I_6	I_7

x2 ⊕ 1 ⊕ 1 ⊕ 0 ⊕ 1 ⊕ 0 = 0 => x2 = 1

海明码计算

- 同理可得x1 = 0, x3 = 1, x4 = 0

位置	1	2	3	4	5	6	7	8	9	10	11
内容	x1	x2	1	x3	0	1	0	x4	1	1	0



位置	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011
内容	0	1	1	1	0	1	0	0	1	1	0

（低位）　　（高位）

- 因此，0110101 的海明码为 01100101110

接收方验证

- 接收端利用相应的偶关系进行验证：
  - $S_1 = R_1 \oplus I_1 \oplus I_3 \oplus I_4 \oplus I_6 \oplus I_7$
  - 同理可得S_0、 S_2、 S_3
- 这里S_1为校正因子，若校正因子为0，则无错；
- 校正因子不为0，有错，错误位置为$S = S_3S_2S_1S_0$处。
- 校正方式：将S位置的比特取反
  - 例如：若S＝0101＝5，则将位置5的比特取反
- 最后去掉校验比特即可得到正确的信息。

接收方验证

	R_0	R_1	I_1	R_2	I_2	I_3	I_4	R_3	I_5	I_6	I_7
位置	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011
内容	0	1	1	1	1	1	0	0	1	1	0

经计算 S =0101

位置	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011
内容	0	1	1	1	0	1	0	0	1	1	0

位置			0011		0101	0110	0111		1001	1010	1011
内容			1		0	1	0		1	1	0

第 3 章　数据链路层

- 3.1　使用点对点信道的数据链路层
- 3.2　点对点协议 PPP
- 3.3　使用广播信道的数据链路层
- 3.4　以太网

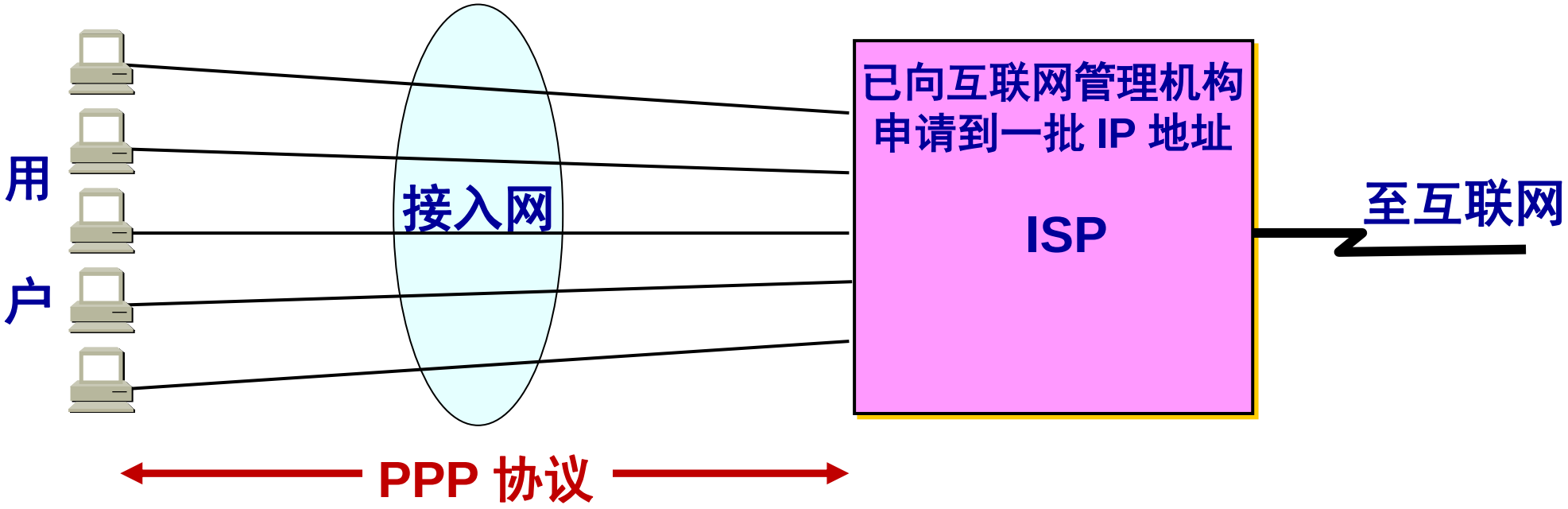
3.2 点对点协议 PPP

- 3.2.1 PPP 协议的特点
- 3.2.2 PPP 协议的帧格式
- 3.2.3 PPP 协议的工作状态

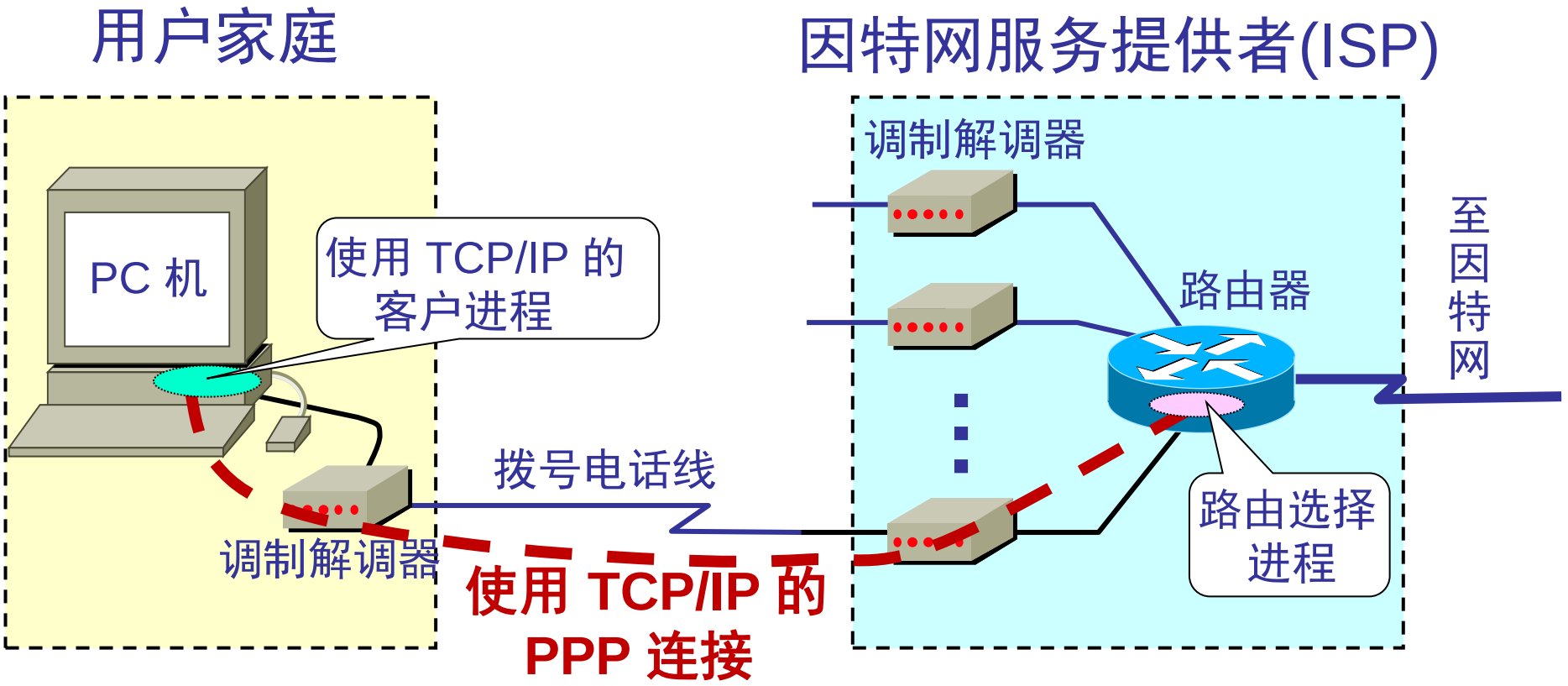
3.2.1 PPP 协议的特点

- 对于点对点的链路，目前使用得最广泛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互联网时， 用户计算机 和 ISP 进行通信时所使用的数据链路层协议就是 PPP 协议。
- 1992年制订了 PPP 协议。经过 1993 年和 1994 年的修订，现在的 PPP 协议已成为因特网的正式标准[RFC1661]。

用户到 ISP 的链路使用 PPP 协议



用户拨号上网示意图


用户家庭
因特网服务提供者(ISP)
PC 机
使用 TCP/IP 的
客户进程
调制解调器
拨号电话线
调制解调器
路由器
至
因
特
网
路由选择
进程
使用 TCP/IP 的
PPP 连接

3.2.1 PPP 协议的特点

- **封装成帧**
  - 必须规定特殊的字符作为帧定界符。
- **透明性**
  - 保证数据传输的透明性。
- **支持多种网络层协议**
  - 能够在同一条物理链路上同时支持多种网络层协议。
- **差错检测**
  - 能够对接收端收到的帧进行检测，并立即丢弃有差错的帧。
- **允许身份验证**
- **允许网络层地址协商**
  - 提供一种机制使通信的两个网络层实体能够通过协商知道或能够配置彼此的网络层地址

PPP 协议的组成

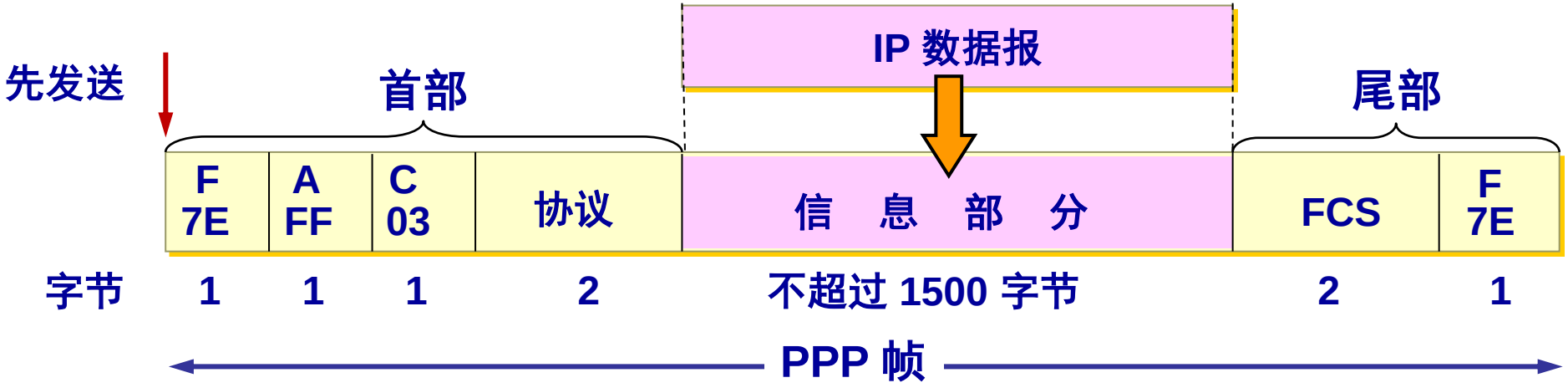
- PPP 协议有三个组成部分：
  - (1) 一个将 IP 数据报封装到串行链路的方法。
  - (2) 链路控制协议 LCP (Link Control Protocol)：建立、配置和测试数据链路的协议。
  - (3) 网络控制协议 NCP (Network Control Protocol)：如为IP协议分配临时IP地址，支持多个网络层协议。

3.2.2 PPP 协议的帧格式

- PPP 帧的首部和尾部分别为 4 个字段和 2 个字段。
- 标志字段 F = 0x7E （符号“0x”表示后面的字符是用十六进制表示。十六进制的 7E 的二进制表示是 01111110）。
- 地址字段 A 只置为 0xFF。地址字段实际上并不起作用
- 控制字段 C 通常置为 0x03。
- 帧检验序列FCS采用CRC。

- PPP帧 是面向字节的，所有的 PPP 帧的长度都是整数字节

PPP 协议的帧格式



PPP 有一个 2 个字节的协议字段。其值

- 若为 0x0021，则信息字段就是 IP 数据报。
- 若为 0x8021，则信息字段是PPP网络控制协议NCP的数据。
- 若为 0xC021，则信息字段是PPP链路控制协议LCP的数据。

透明传输问题

- 异步传输时：字符填充
  - 将信息字段中出现的每一个 0x7E 字节转变成为 2 字节序列 (0x7D, 0x5E)。
  - 若信息字段中出现一个 0x7D 的字节, 则将其转变成为 2 字节序列 (0x7D, 0x5D)。
  - 若信息字段中出现 ASCII 码的控制字符（即数值小于 0x20 的字符），则在该字符前面要加入一个 0x7D 字节，同时将该字符的编码加以改变
- 同步传输时：比特填充
  - 在5个连续 1 的后面插入0

PPP字符填充举例

- 若PPP帧的数据部分为：
  - 7D 5E FE 27 7D 5D 7D 5D 65 7D 5E
- 0x7E -> 0x7D 0x5E
- 0x7D -> 0x7D 0x5D

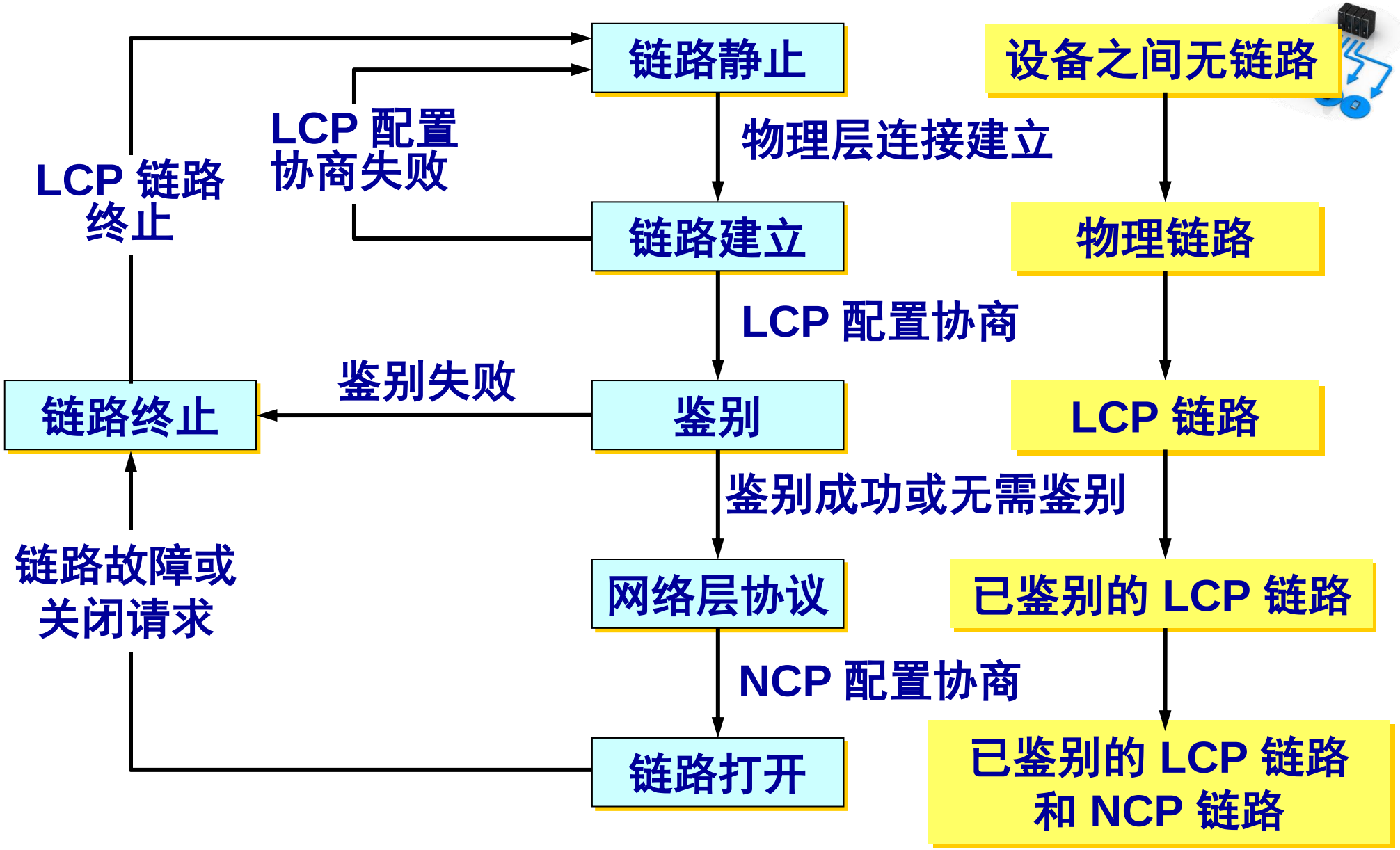
- 则原始数据为：7E FE 27 7D 7D 65 7E

不提供使用序号和确认的可靠传输

- PPP 协议之所以不使用序号和确认机制是出于以下的考虑：
  - ① 在数据链路层出现差错的概率不大时，使用比较简单的 PPP 协议较为合理。
  - ② 在因特网环境下，PPP 的信息字段放入的数据是 IP 数据报。数据链路层的可靠传输并不能够保证网络层的传输也是可靠的。
  - ③ 帧检验序列 FCS 字段可保证无差错接受。

3.2.3 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
- 接着进行网络层配置，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。
- 可见，PPP 协议已不是纯粹的数据链路层的协议，它还包含了物理层和网络层的内容。


链路静止
设备之间无链路
LCP 配置
协商失败
物理层连接建立
LCP 链路
终止
链路建立
物理链路
LCP 配置协商
鉴别失败
链路终止
鉴别
LCP 链路
鉴别成功或无需鉴别
链路故障或
关闭请求
网络层协议
已鉴别的 LCP 链路
NCP 配置协商
链路打开
已鉴别的 LCP 链路
和 NCP 链路

PPP 协议的状态图

PPP协议的特点

- PPP帧中增加了校验字段，PPP在链路层具有差错检测功能；
- PPP的LCP协议提供通信双方进行参数协商的手段；
  - 协商参数有：数据帧的最大帧长、身份认证、NCP协议、数据压缩方式等。
- PPP帧中增加了协议字段，使得PPP可以支持多种网络层协议，有IP、IPX、OSI、CLNP等。
- 支持IP的NCP可以在建立连接时动态分配IP地址，解决了家庭用户拨号上网的问题。

第 3 章　数据链路层

- 3.1　使用点对点信道的数据链路层
- 3.2　点对点协议 PPP
- 3.3　使用广播信道的数据链路层
- 3.4　以太网

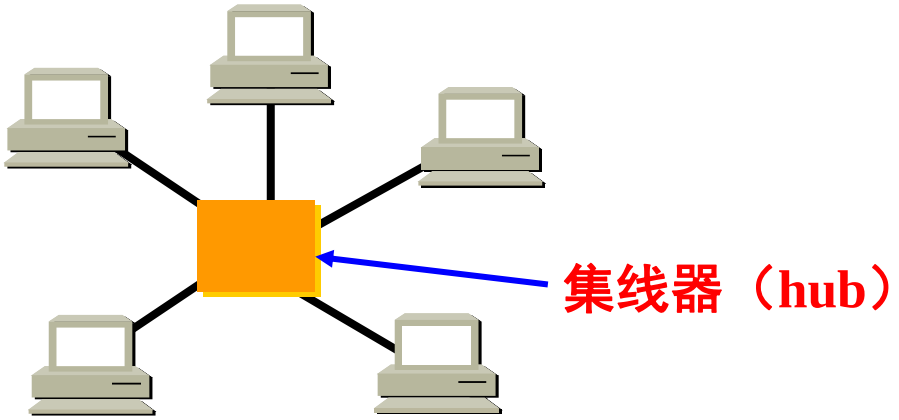
3.3 使用广播信道的数据链路层

- 3.3.1 局域网的数据链路层
- 3.3.2 CSMA/CD 协议
- 3.3.3 信道利用率

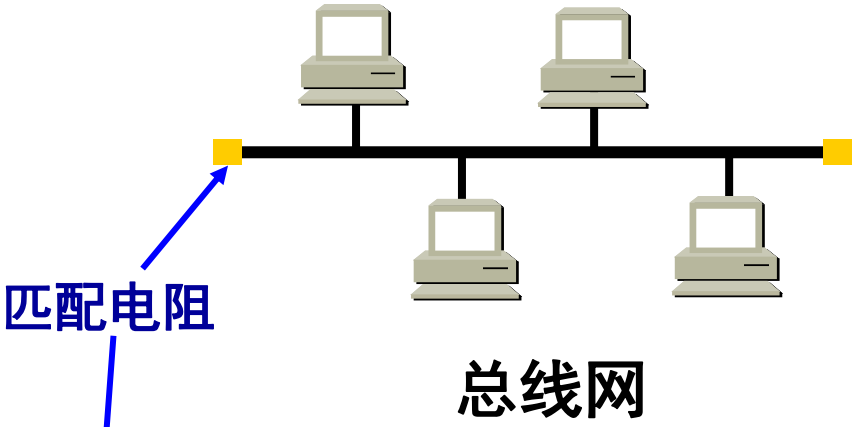
3.3.1 局域网的数据链路层

- 局域网的概念
  - 局域网(Local Area Network，简称LAN)是计算机网络的一种。局域网是在一个较小的范围(一个办公室、一幢楼、一家工厂等)，利用通信线路将众多计算机(一般为微机)及外围设备连接起来，达到数据通信和资源共享的目的。
- 局域网最主要的特点是：
  - 网络为一个单位所拥有；地理范围和站点数目均有限。
  - 具有较高的数据率、较低的时延和较小的误码率。
- 局域网具有如下主要优点：
  - 具有广播功能，从一个站点可很方便地访问全网。局域网上的主机可共享连接在局域网上的各种硬件和软件资源。
  - 便于系统的扩展和逐渐地演变，各设备的位置可灵活调整和改变。
  - 提高了系统的可靠性、可用性和生存性。

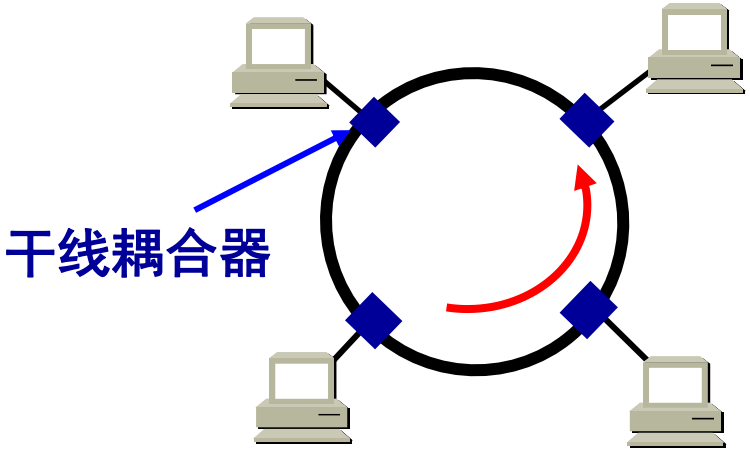
局域网拓扑结构



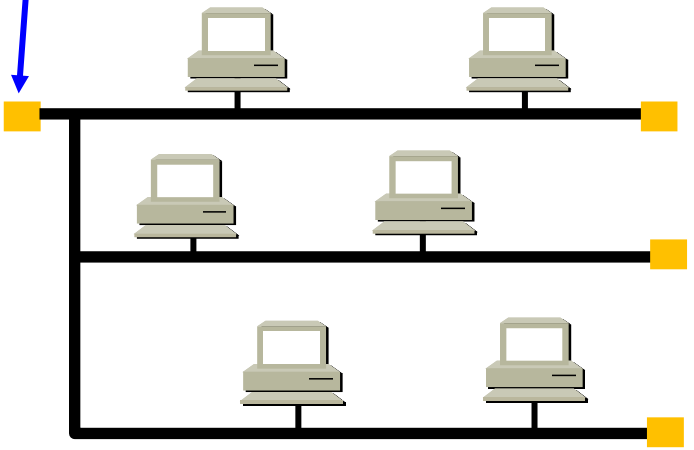
星形网



总线网



环形网



树形网

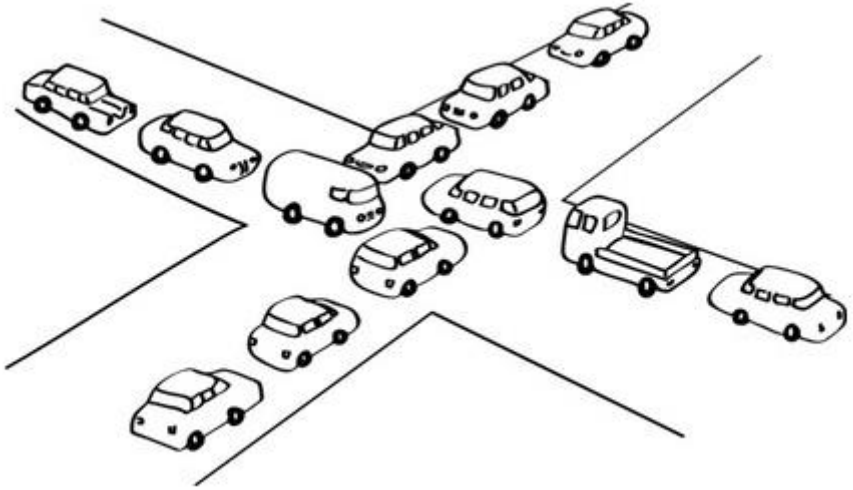
局域网信道分配策略

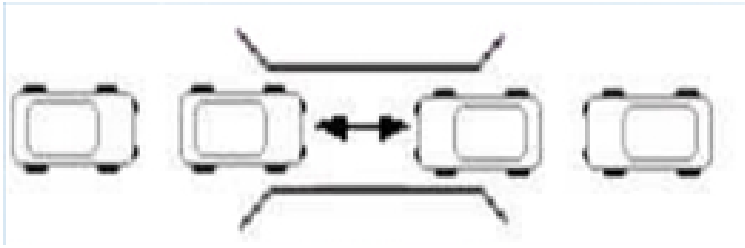
- 广播网中所有站点共享同一个信道，任一站点发送的信息能被所有其他站点接收到。

- 问题
  - 若有两个或两个以上的站点同时发送数据，则信号在信道中发生碰撞，数据发送失败，为冲突。
  - 广播网中，如何将单一的信道分配各各个不同的用户，是个重要的问题。

- 用户使用的信道称为媒体（介质），决定由谁来使用信道的协议为“媒体（介质）访问控制协议”。

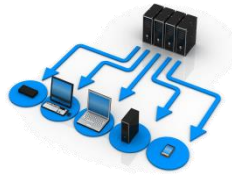
冲突的产生





单一车道上同时有两个方向的车行驶时，如果要让两个方向的车通过，就必须是不同时刻经过。

媒体（介质）共享技术



- **静态划分信道**
  - 频分复用
  - 时分复用
  - 波分复用
  - 码分复用

用户只要分配到了信道就不会和其他用户发生冲突。

- **静态分配的特点**
  - 站点数目少且固定，且每个站点有大量数据发送，控制协议简单且传输的效率高。
  - 对于大部分计算机网络，站点数目多且不固定，数据传输有突发性，信道的利用率低。
  - 代价较高，不适合于局域网使用

媒体共享技术—动态分配

- **动态媒体接入控制**
  - 信道不是在用户通信时固定分配给用户。
    - 例如：异步时分多路复用STDM，各站点仅当有数据发送时，才占用信道发送数据。
- **动态接入控制类型**
  - 随机接入
    - 用户发送前不需要取得发送权，有数据就发送，发生冲突（碰撞）后采取措施解决冲突
  - 控制接入
    - 用户首先获得发送权，再发送数据，不会产生冲突
    - 令牌环局域网的多点线路探询（polling）

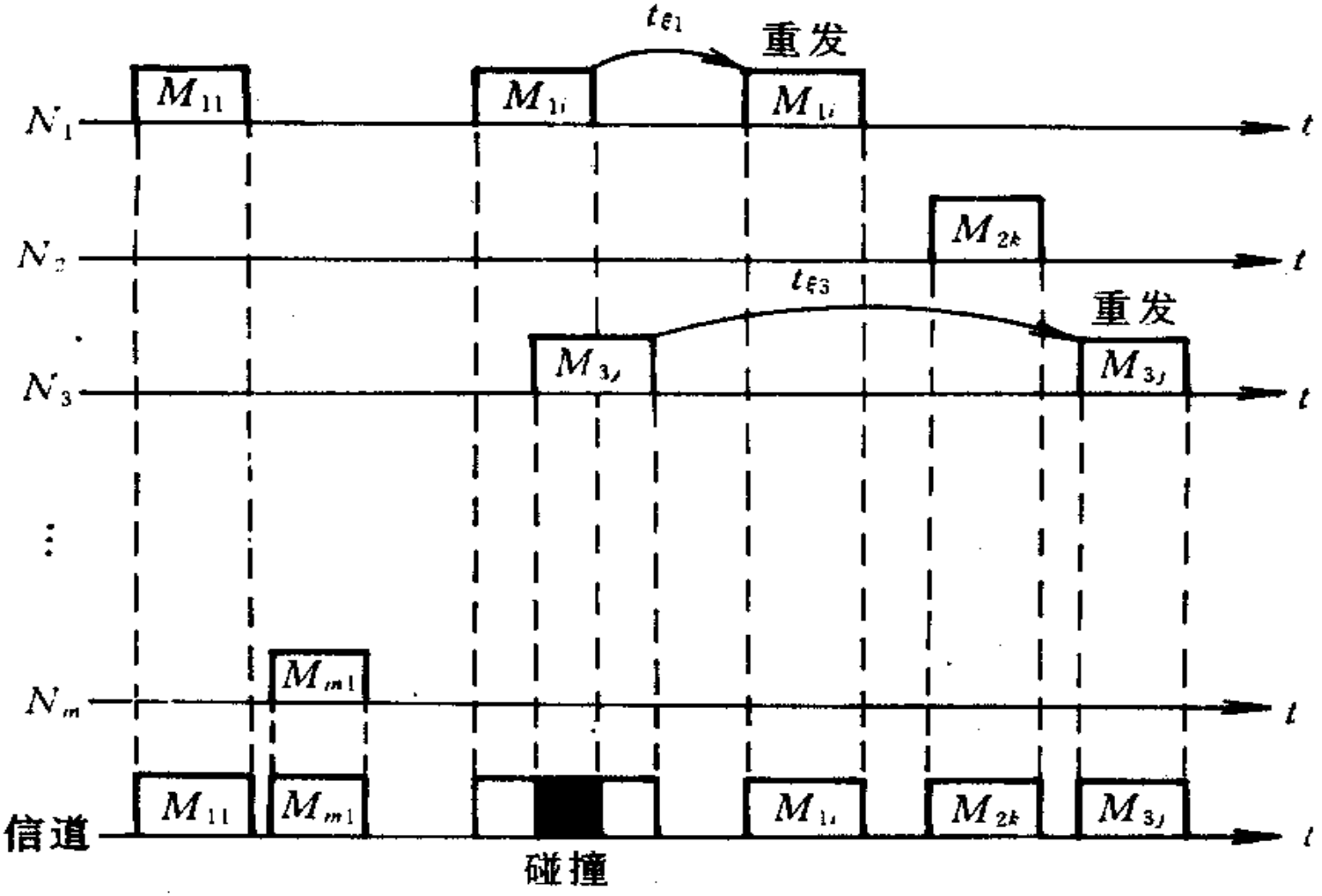
代表性媒体访问控制方法

- 争用协议
  - ALOHA协议
  - CSMA/CD协议
  - 随机访问：意味着对任何站都无法预计其发送的时刻；
  - 竞争发送：是指所有发送的站自由竞争信道的使用权
- 无冲突协议（略）
  - 比特映像介质访问控制协议（先预约然后传输）
  - 小时间片轮换优先优先权介质访问控制协议
  - 二进制地址相加协议
- 有限争用协议（略）
  - 思想：网络轻负载时使用竞争策略，重负载时使用无冲突策略
  - 自适应步进树协议

ALOHA系统

- 争用协议最早起源于夏威夷大学的ALOHA系统，该网络通过无线信道将各个分校的终端连接到本部的主机上。
- ALOHA是陆地无线网，其基本思想是如何实现多个用户竞争使用单一信道的系统。
- ALOHA系统的思想
  - 任何用户有数据发送就可以发送（会带来冲突）；
  - 每个用户通过监听信道来获知数据传输是否成功；
  - 发现数据传输失败后，各自等待一段随机时间，再重新发送。

ALOHA系统工作原理

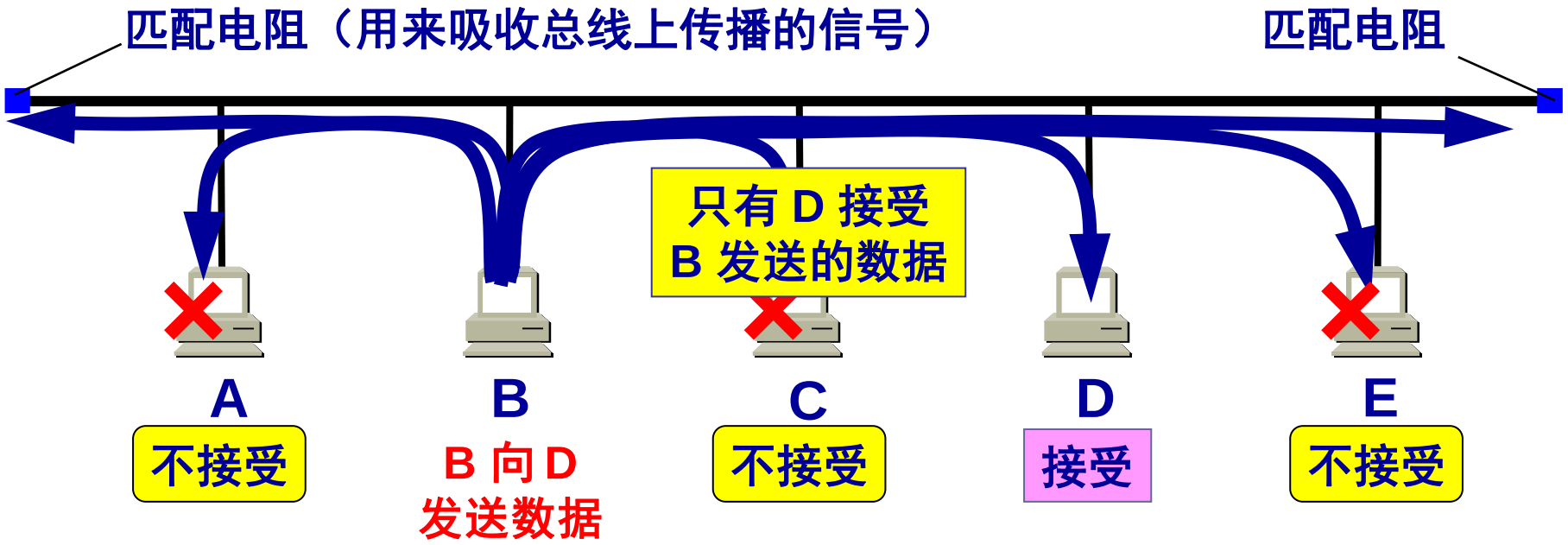


ALOHA系统信道分析

- 竞争系统中，一方面不断有新的数据帧发送，另一方面冲突帧需要重发，系统的吞吐量是一个重要的指标。
- 吞吐量：单位时间内系统能够成功发送的新的数据帧的平均数量。
- 结论：
  - ALOHA系统最大的信道利用率为18.4%；
  - 对ALOHA系统改进的时分ALOHA系统的最大信道利用率为36.8%.
  - ALOHA系统的信道利用率是非常低的。原因主要是各个站自由发送数据，碰撞概率增大。

CSMA/CD 协议

- 最初的以太网是将许多计算机都连接到一根总线上。当初认为这样的连接方法既简单又可靠，因为总线上没有有源器件。



以太网采用广播方式发送

- 总线上的每一个工作的计算机都能检测到 B 发送的数据信号。
- 由于只有计算机 D 的地址与数据帧首部写入的地址一致，因此只有 D 才接收这个数据帧。
- 其他所有的计算机（A, C 和 E）都检测到不是发送给它们的数据帧，因此就丢弃这个数据帧而不能够收下来。
- 在具有广播特性的总线上实现了一对一的通信。

以太网采取了两种重要的措施

为了通信的简便，以太网采取了两种重要的措施：

(1) 采用较为灵活的无连接的工作方式

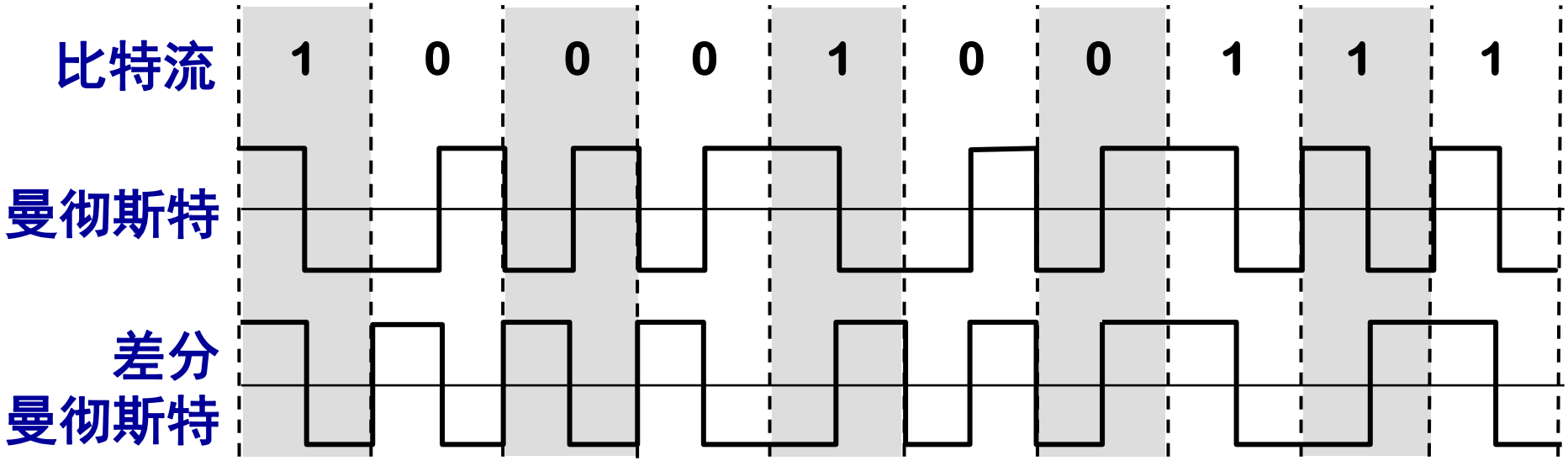
- 不必先建立连接就可以直接发送数据。
- 对发送的数据帧不进行编号，也不要求对方发回确认。
- 这样做的理由是局域网信道的质量很好，因信道质量产生差错的概率是很小的。

以太网提供的服务

- 以太网提供的服务是不可靠的交付，即尽最大努力的交付。
- 当目的站收到有差错的数据帧时就丢弃此帧，其他什么也不做。差错的纠正由高层来决定。
- 如果高层发现丢失了一些数据而进行重传，但以太网并不知道这是一个重传的帧，而是当作一个新的数据帧来发送。

以太网采取了两种重要的措施

(2) 以太网发送的数据都使用曼彻斯特 (Manchester) 编码



曼彻斯特编码缺点是：它所占的频带宽度比原始的基带信号增加了一倍。

CSMA/CD协议

- CSMA/CD 含义：载波监听 多点接入 / 碰撞检测
  (Carrier Sense Multiple Access with Collision Detection)

- “多点接入”指总线型网络，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。
- “载波监听”就是用电子技术检测总线上有没有其他计算机发送的数据信号。

CSMA/CD协议

- “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）。
- 当一个站检测到的信号电压摆动值超过一定的门限值时，就认为总线上至少有两个站同时在发送数据，表明产生了碰撞。
- 所谓“碰撞”就是发生了冲突。因此“碰撞检测”也称为“冲突检测”。

先听后发，边发边听

检测到碰撞后

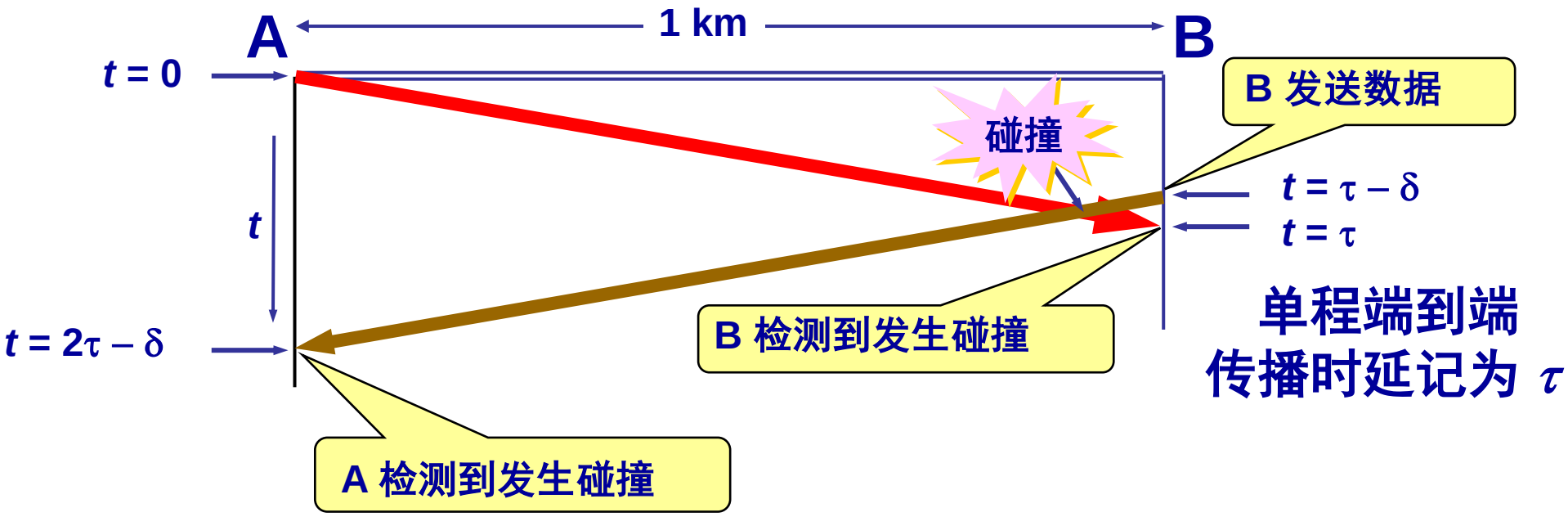
- 在发生碰撞时，总线上传输的信号产生了严重的失真，无法从中恢复出有用的信息来。
- 每一个正在发送数据的站，一旦发现总线上出现了碰撞，就要立即停止发送，免得继续浪费网络资源，然后等待一段随机时间后再次发送。

- 为什么会发生“碰撞”？
  - 每个站点都是在监听到信道“空闲”时才发送数据的，为什么还会发生碰撞？根本原因是因为电磁波在媒体上的传播速度总是有限的。

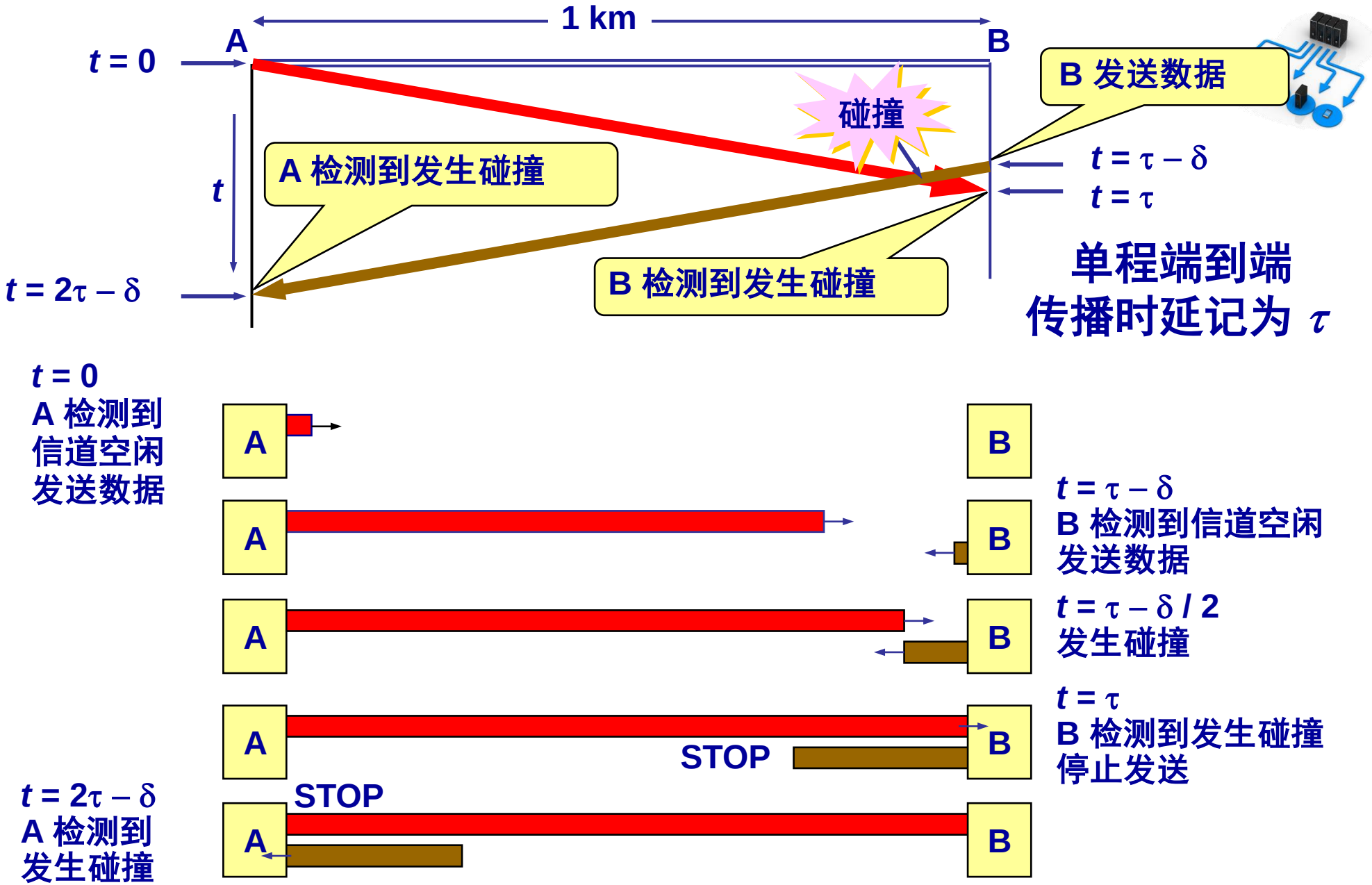
为什么要进行碰撞检测？

- 由于电磁波在总线上的传播速率是有限的，当某个站监听到总线是空闲时，也可能总线并非真正是空闲的。
  - A 向 B 发出的信息，要经过一定的时间后才能传送到 B。
  - B 若在 A 发送的信息到达 B 之前发送自己的帧 (因为这时 B 的载波监听检测不到 A 所发送的信息)，则必然要在某个时间和 A 发送的帧发生碰撞。
- 碰撞的结果是两个帧都变得无用。
- 所以需要在发送期间进行碰撞检测，以检测冲突。

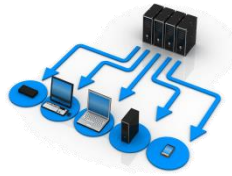
信号传播时延对载波监听的影响



A需要单程传播时延的 2 倍的时间，
才能检测到与 B 的发送产生了冲突


1 km
A
B
t = 0
碰撞
B 发送数据
A 检测到发生碰撞
t
t = τ – δ
t = τ
t = 2τ – δ
B 检测到发生碰撞
单程端到端
传播时延记为 τ
t = 0
A 检测到
信道空闲
发送数据
A
B
t = τ – δ
B 检测到信道空闲
发送数据
t = τ – δ / 2
发生碰撞
t = τ
B 检测到发生碰撞
停止发送
STOP
STOP
t = 2τ – δ
A 检测到
发生碰撞

CSMA/CD 重要特性



- 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。
- 这种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。

- 两个重要问题
  - 什么时候可以检测到碰撞？（争用期）
  - 检测到冲突后，等待多长时间再重试？（退避算法）

争用期

- 最先发送数据帧的站，在发送数据帧后至多经过时间 2τ（两倍的端到端往返时延）就可知道发送的数据帧是否遭受了碰撞。
- 以太网的端到端往返时延 2τ 称为争用期，或碰撞窗口。
- 经过争用期这段时间还没有检测到碰撞，才能肯定这次发送不会发生碰撞。

二进制指数类型退避算法 (truncated binary exponential type)

- 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。
  - 基本退避时间取为争用期 2τ。
  - 从整数集合 [0, 1, 2，3，4，… , $(2^k - 1)$] 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。
  - 参数 k 按下面的公式计算：

$$k = \text{Min}[\text{重传次数}, 10]$$

  - 当 $k \le 10$ 时，参数 k 等于重传次数。
  - 当重传达 16 次仍不能成功时即丢弃该帧，并向高层报告。

争用期的长度

- 10 Mbit/s 以太网 取 51.2 μs 为争用期的长度。
- 对于 10 Mbit/s 以太网，在争用期内可发送 512 bit，即 64 字节（最短有效帧长）。

> 这意味着：
> 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。

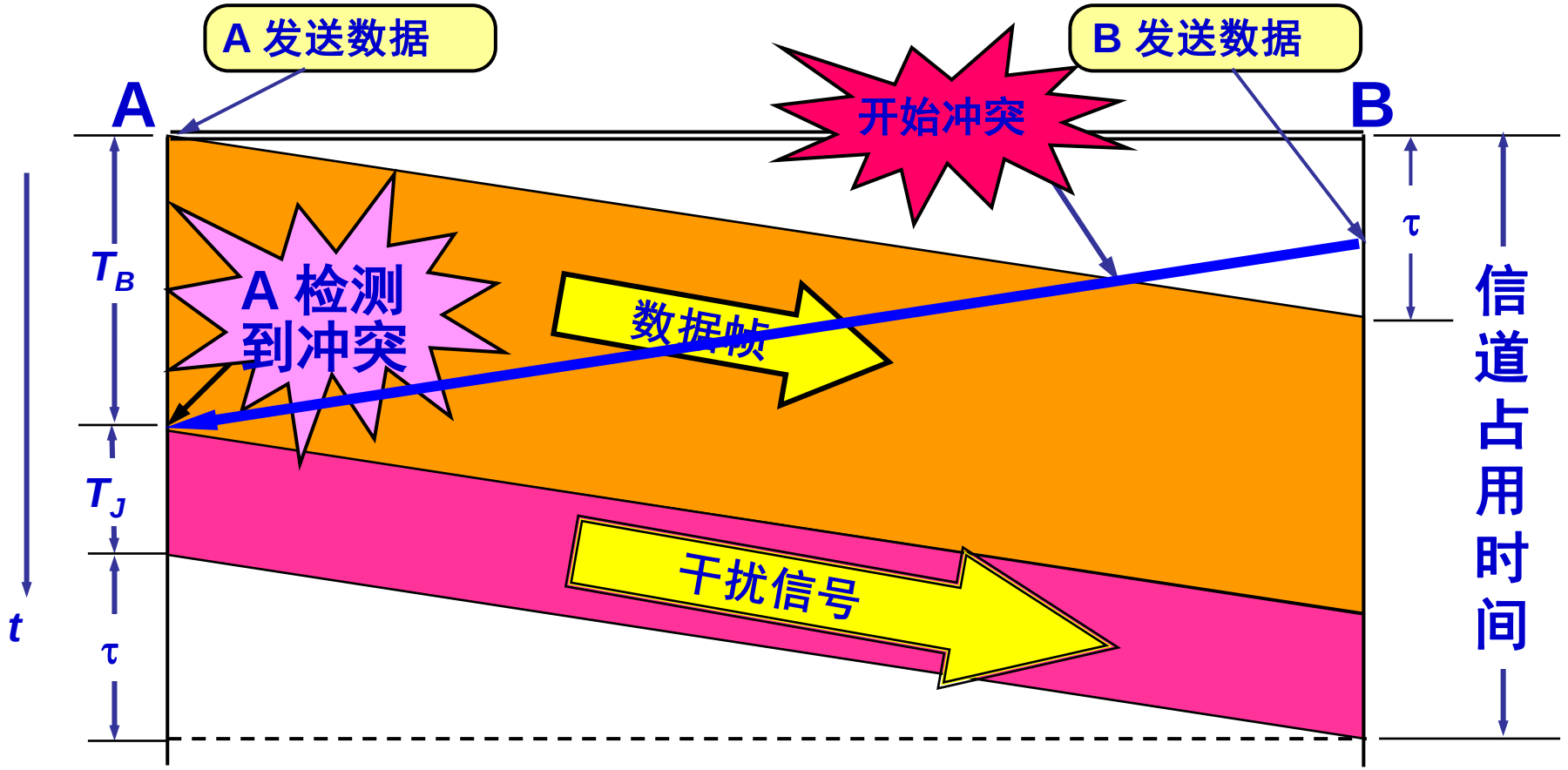
- 如果发生冲突，就一定是在发送的前 64 字节之内。
- 由于一检测到冲突就立即中止发送，这时已经发送出去的数据一定小于 64 字节。
- 以太网规定了最短有效帧长为 64 字节，凡长度小于 64 字节的帧都是由于冲突而异常中止的无效帧。

强化碰撞

当发送数据的站一旦发现发生了碰撞时：

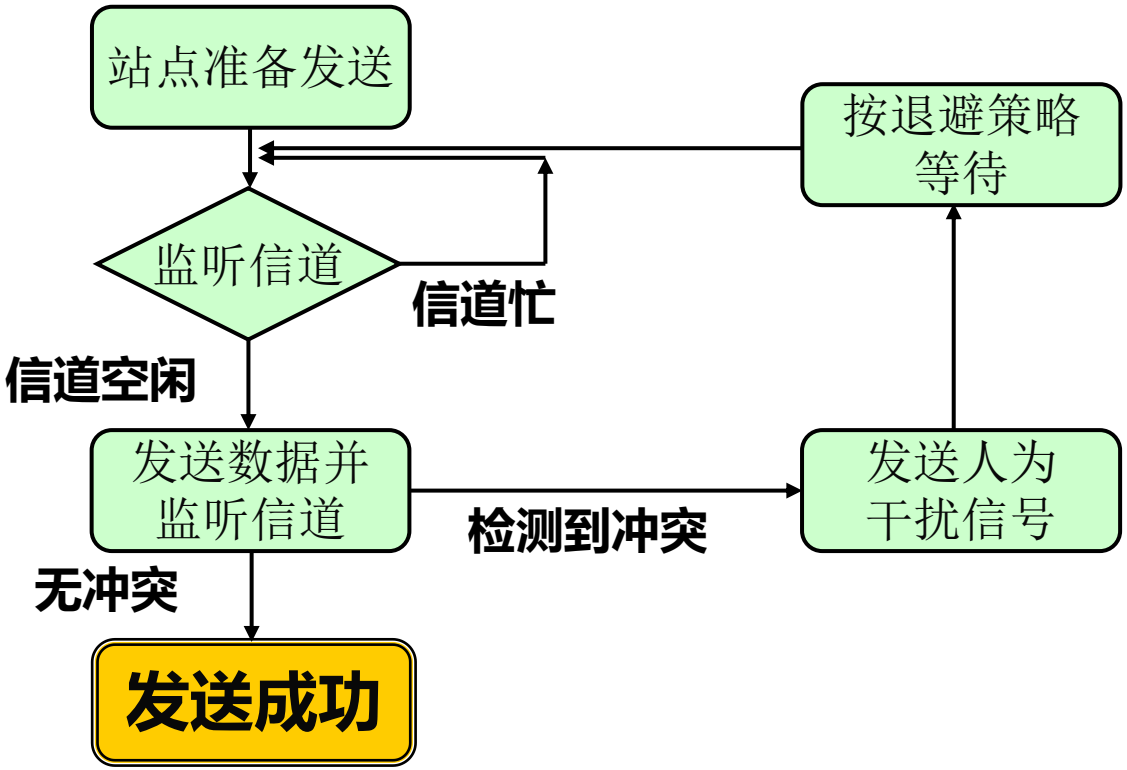
- (1) 立即停止发送数据；
- (2) 再继续发送若干比特的人为干扰信号 (jamming signal)，以便让所有用户都知道现在已经发生了碰撞。

人为干扰信号

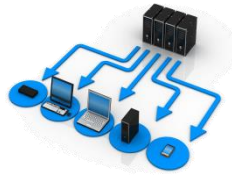


B 也能够检测到冲突，并立即停止发送数据帧，接着就发送干扰信号。这里为了简单起见，只画出 A 发送干扰信号的情况。

CSMA/CD工作流程图



CSMA/CD协议的要点

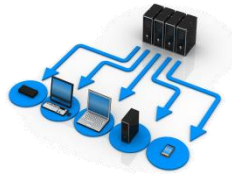


- (1) 准备发送。但在发送之前，必须先检测信道。
- (2) 检测信道。若检测到信道忙，则应不停地检测，一直等待信道转为空闲。
  - 若检测到信道空闲，并在 96 比特时间内信道保持空闲（保证了帧间最小间隔 9.6 μs），就发送这个帧。
- (3) 检查碰撞。在发送过程中仍不停地检测信道，即网络适配器要边发送边监听。这里只有两种可能性：
  - ①发送成功：在争用期内一直未检测到碰撞。这个帧肯定能够发送成功。发送完毕后，其他什么也不做。然后回到 (1)。
  - ②发送失败：在争用期内检测到碰撞。这时立即停止发送数据，并按规定发送人为干扰信号。适配器接着就执行指数退避算法，等待 *r* 倍 512 比特时间后，返回到步骤 (2)，继续检测信道。但若重传达 16 次仍不能成功，则停止重传而向上报错。

帧间最小间隔

- 帧间最小间隔为 9.6 μs，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 μs 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。

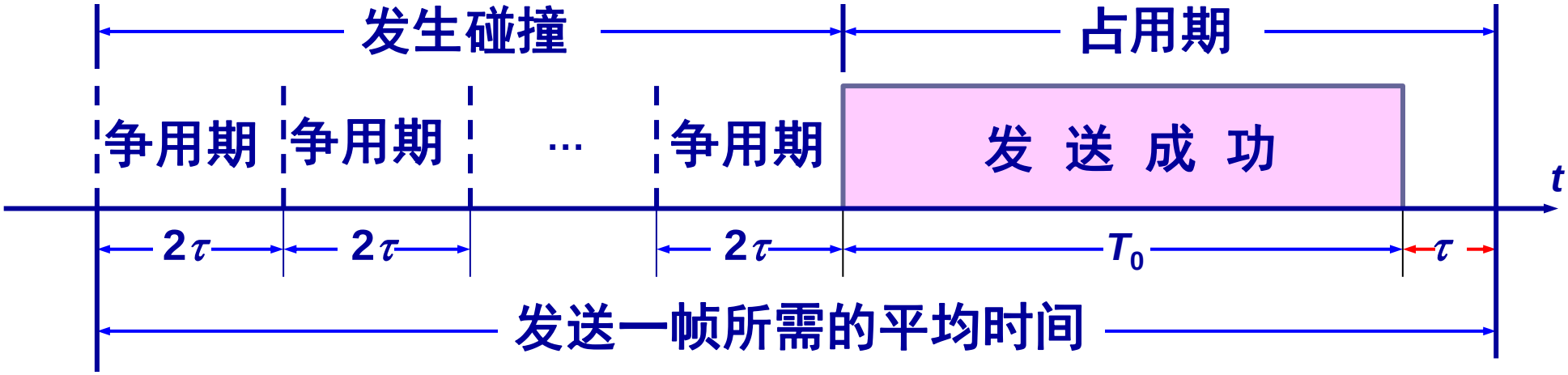
3.3.3 信道利用率



- 多个站在以太网上同时工作就可能会发生碰撞。
- 当发生碰撞时，信道资源实际上是被浪费了。因此，当扣除碰撞所造成的信道损失后，**以太网总的信道利用率并不能达到 100%。**
- 假设 τ 是以太网单程端到端传播时延。则争用期长度为 2τ，即端到端传播时延的两倍。检测到碰撞后不发送干扰信号。
- 设帧长为 L (bit)，数据发送速率为 C (bit/s)，则帧的发送时间为 $T_0 = L/C$ (s)。

以太网信道被占用的情况

- 一个站在发送帧时出现了碰撞。经过一个争用期 2τ 后，可能又出现了碰撞。这样经过若干个争用期后，一个站发送成功了。假定发送帧需要的时间是 T_0。



以太网信道被占用的情况

- 注意到，成功发送一个帧需要占用信道的时间是 $T_0 + \tau$，比这个帧的发送时间要多一个单程端到端时延τ。
- 这是因为当一个站发送完最后一个比特时，这个比特还要在以太网上传播。
- 在最极端的情况下，发送站在传输媒体的一端，而比特在媒体上传输到另一端所需的时间是 τ。

参数 α 与利用率

- 要提高以太网的信道利用率，就必须减小 τ 与 T_0 之比。
- 在以太网中定义了**参数 α**，它是以太网单程端到端时延 τ 与帧的发送时间 T_0 之比：

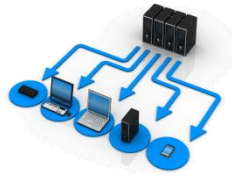
$$a = \tau / T_0$$

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

对以太网参数 α 的要求

- 为提高利用率，以太网的参数**α** 的值应当尽可能小些。
- **对以太网参数 α 的要求是：**
  - 当数据率一定时，**以太网的连线的长度受到限制**，否则 τ 的数值会太大。
  - **以太网的帧长不能太短**，否则 T_0 的值会太小，使 α 值太大。

信道利用率的最大值 S_{max}



- 在**理想化**的情况下，以太网上的各站发送数据都不会产生碰撞（这显然已经不是 CSMA/CD，而是需要使用一种特殊的调度方法），即总线一旦空闲就有某一个站立即发送数据。
- 发送一帧占用线路的时间是 $T_0 + \tau$，而帧本身的发送时间是 T_0。于是我们可计算出**理想情况下的极限信道利用率 S_{max} 为：**

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

- 只有当参数 a 远小于 1 才能得到尽可能高的极限信道利用率。
- 据统计，当以太网的利用率达到 30% 时就已经处于重载的情况。很多的网络容量被网上的碰撞消耗掉了。

第 3 章　数据链路层

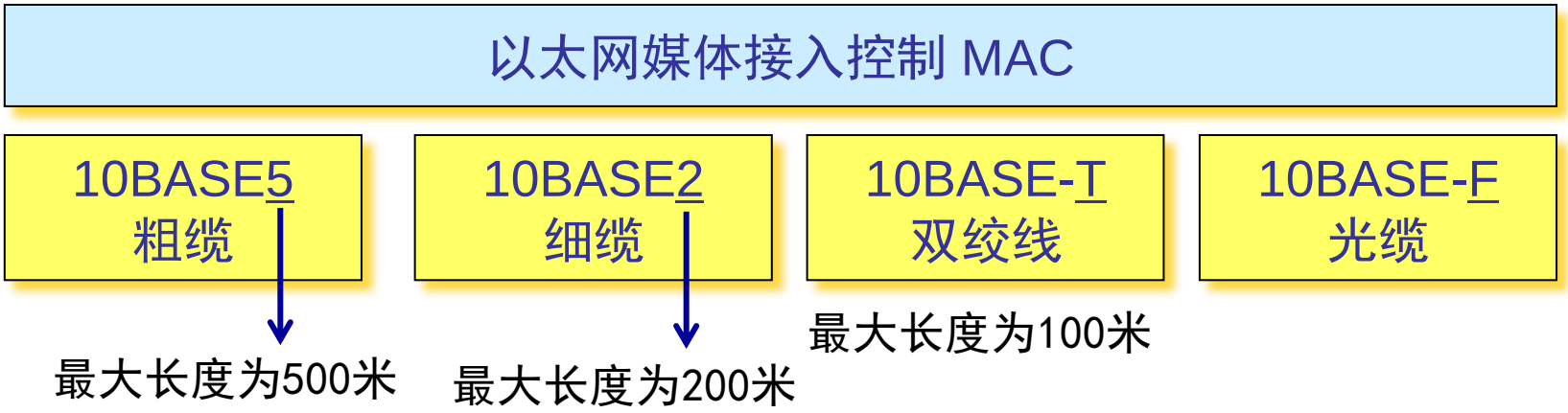
- 3.1　使用点对点信道的数据链路层
- 3.2　点对点协议 PPP
- 3.3　使用广播信道的数据链路层
- 3.4　以太网

3.4 以太网

- 3.4.1 传统以太网
- 3.4.2 以太网的层次结构
- 3.4.3 以太网的MAC层
  - MAC层的硬件地址
  - MAC帧格式
- 3.4.4 扩展以太网
- 3.4.5 虚拟以太网
- 3.4.6 高速以太网

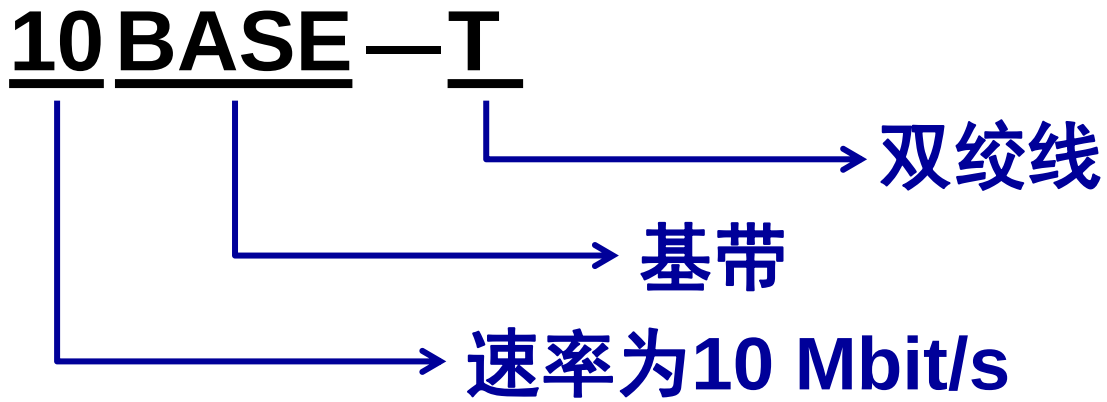
3.4.1 传统以太网

- **以太网(Ethernet)**指的是由美国施乐(Xerox)公司创建并由Xerox、Intel和DEC公司联合开发的**基带局域网规范**，是当今现有局域网采用的最通用的通信协议标准。
- **传统以太网（10Mbits/s速率）**最初是使用**粗同轴电缆**，后来演进到使用比较便宜的**细同轴电缆**，最后发展为使用更便宜和更灵活的**双绞线**。
- 以太网有四种不同的物理层



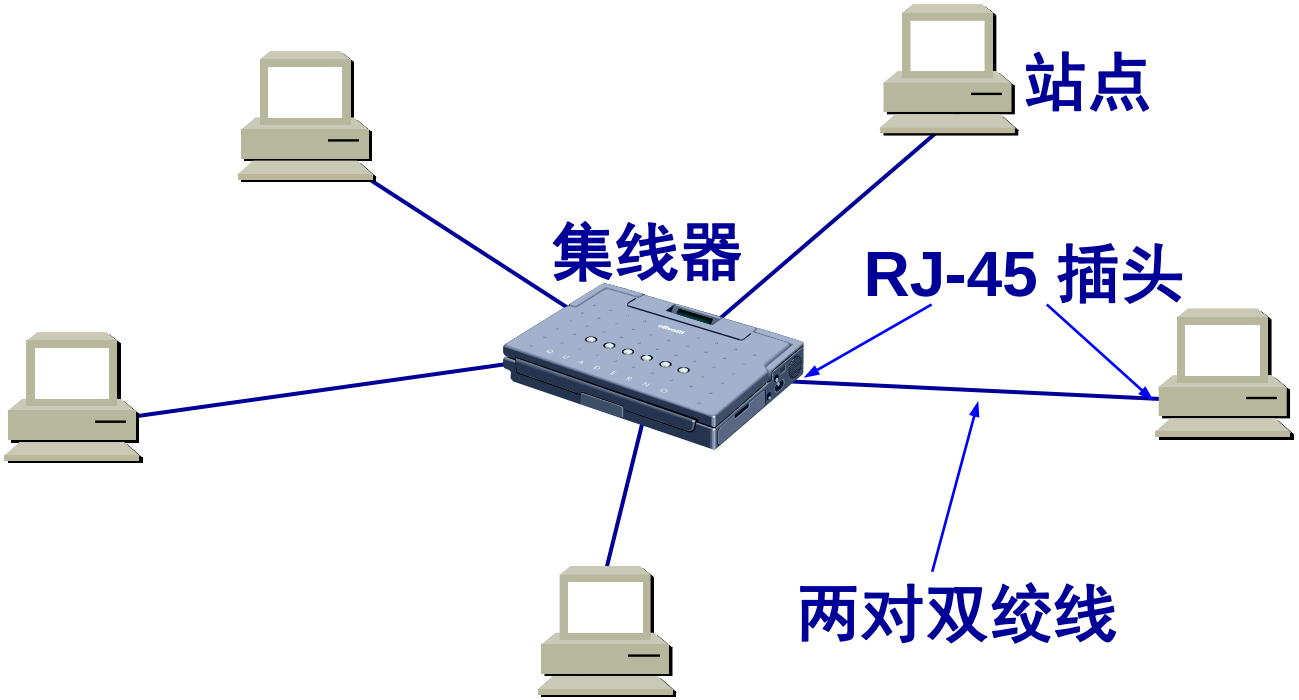
星形以太网 10BASE-T

- 1990 年，IEEE 制定出星形以太网 10BASE-T 的标准 802.3i。

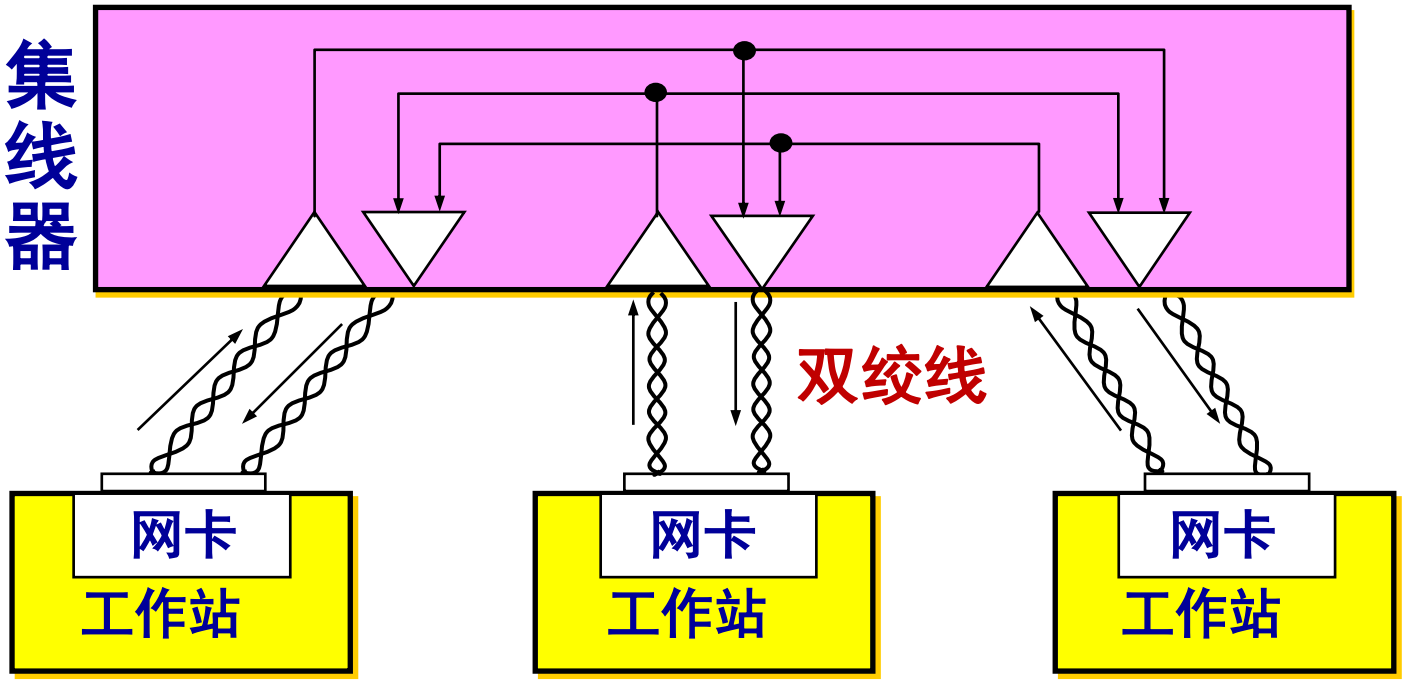


使用集线器的双绞线以太网

- 采用双绞线的以太网采用**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)。



具有三个接口的集线器



星形以太网 10BASE-T

- 使用无屏蔽双绞线，采用星形拓扑。
- 每个站需要用两对双绞线，分别用于发送和接收。
- 双绞线的两端使用 RJ-45 插头。
- 集线器使用了大规模集成电路芯片，因此集线器的可靠性提高。
- 10BASE-T 的通信距离稍短，每个站到集线器的距离不超过 100 m。
  - 以太网上主机之间距离不能太远，否则主机发送的信号经过传输衰减，使CSMA/CD协议无法正常工作。

集线器的特点

- (1) 集线器是使用电子器件来模拟实际电缆线的工作，因此整个系统仍然像一个传统的以太网那样运行。
- (2) 使用集线器的以太网在逻辑上仍是一个总线网，各工作站使用的还是 CSMA/CD 协议，并共享逻辑上的总线。
- (3) 集线器很像一个多接口的转发器，工作在物理层，不进行碰撞检测。
- (4) 集线器采用了专门的芯片，进行自适应串音回波抵消，减少了近端串音。每个比特在转发之前还要进行再生整形并重新定时。

10BASE-T 以太网在局域网中的统治地位

- 这种 10 Mbit/s 速率的无屏蔽双绞线星形网的出现，既降低了成本，又提高了可靠性。 具有很高的性价比。
- 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。
- 从此以太网的拓扑就从总线形变为更加方便的星形网络，而以太网也就在局域网中占据了统治地位。

以太网（Ethernet）的两个标准

- DIX Ethernet V2 是DEC、Intel、Xerox公司联合提出的世界上第一个局域网产品（以太网）的规约 ——10M/s的以太网规约。
- IEEE 802.3 是第一个 IEEE 802委员会制定的局域网标准

- DIX Ethernet V2 标准与 IEEE 的 802.3 标准只有很小的差别，因此可以将 802.3 局域网简称为“以太网”。
- 严格说来，“以太网”应当是指符合 DIX Ethernet V2 标准的局域网 。

IEEE 802标准

- 80年代局域网迅速发展，各种标准层出不穷，为了使得不同厂家生产的局域网能够通信，IEEE于1980年2月成立一个局域网标准委员会，形成一系列的标准为IEEE 802 标准。
- IEEE802标准已被ANSI接收为美国国家标准，并于84年3月被ISO采纳为局域网的国际标准。
- 由于厂商的竞争，IEEE没有制定一个统一的局域网的标准，而是制定不同的局域网标准。

IEEE802标准间的关系

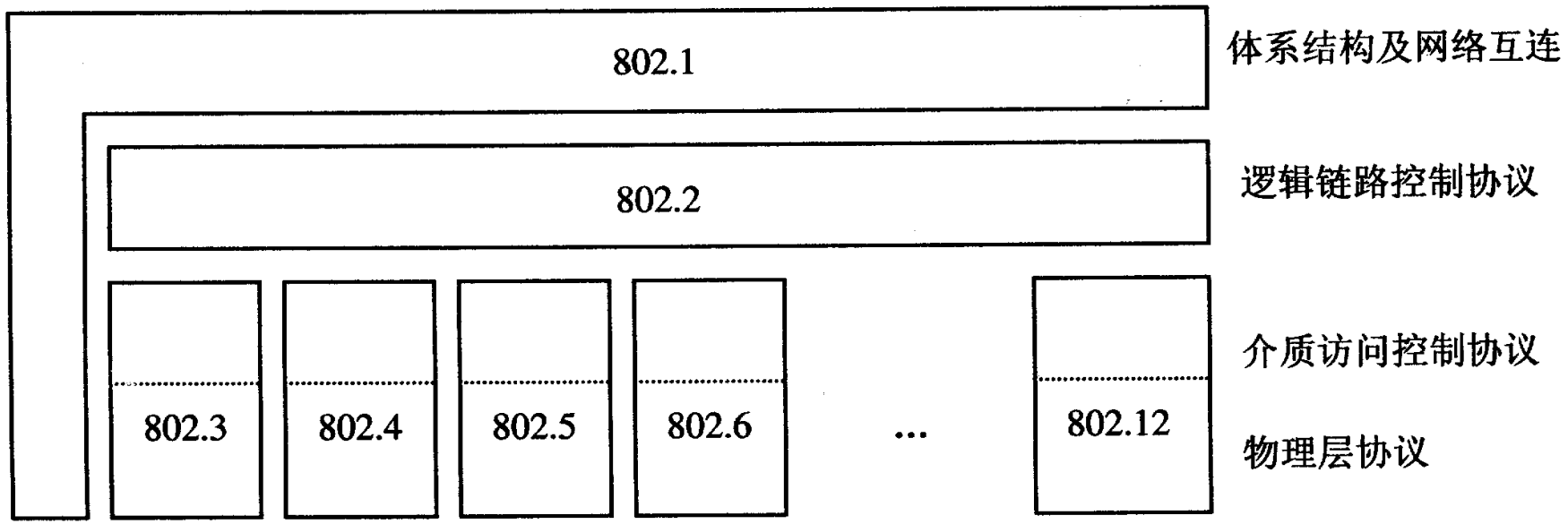


图 3.1　IEEE 802 局域网标准系列间的关系

802.1局域网概述，体系结构，网络管理和性能测量等；
802.2逻辑链路控制协议；
802.3总线网介质访问控制协议CSMA/CD及物理层技术规范；
802.11无线局域网的介质访问控制协议及其物理层技术规范。

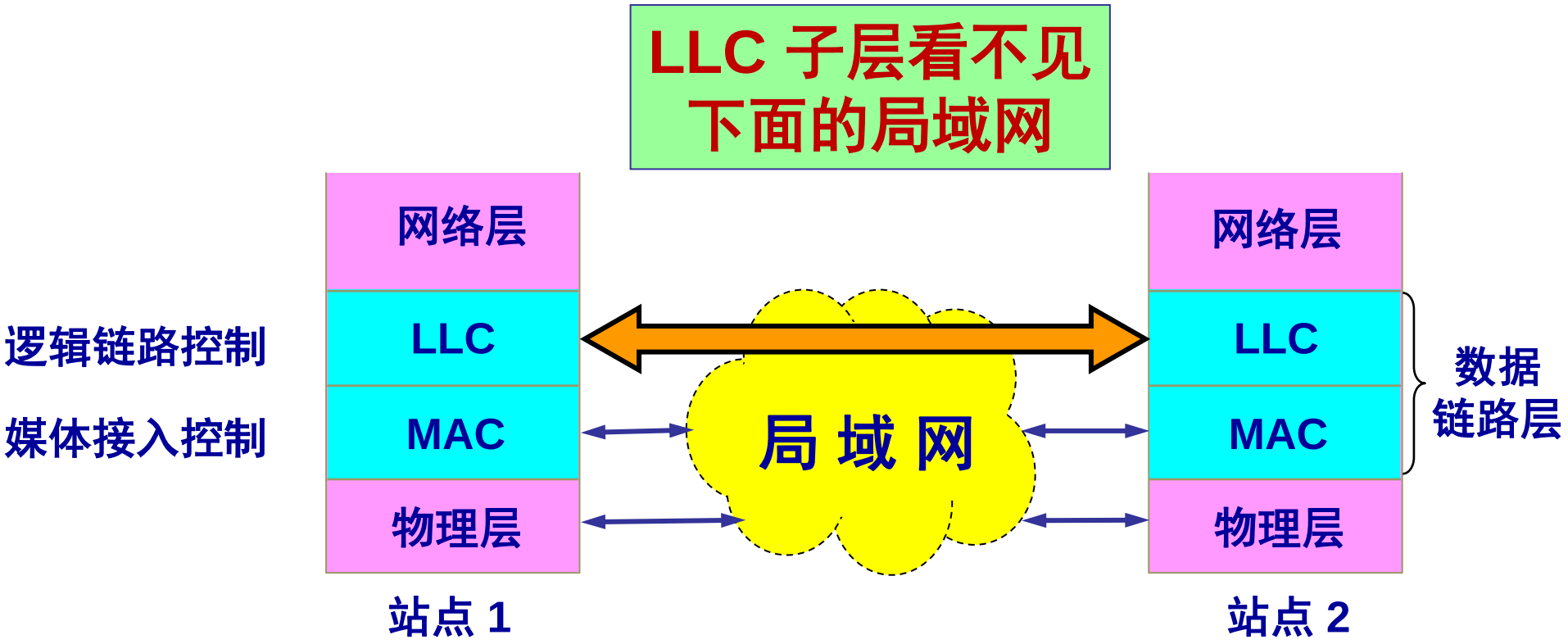
3.4.2 以太网的层次结构

- IEEE 802委员会是局域网标准的主要制定者，它提出的局域网参考模型主要定义了物理层和数据链路层的规范，即相当于OSI模型的下两层。
  - 物理层：与OSI模型中类似，物理层负责与传输介质的连接，并在传输介质上传输比特流，因此，它描述和规定了与传输介质接口的特性。
  - 数据链路层：OSI模型中数据链路层的功能在IEEE 802模型中分成了两个子层(MAC和LLC)。

数据链路层的两个子层

- 为了使数据链路层能更好地适应多种局域网标准，IEEE 802 委员会就将局域网的数据链路层拆成两个子层：
  - 逻辑链路控制 LLC (Logical Link Control)子层：屏蔽对各种不同物理网络的访问方法的差异，向上提供数据传输服务的统一的逻辑接口
  - 媒体接入控制 MAC (Medium Access Control)子层：控制对传输介质的访问，并在物理层的基础上实现无差错通信。该子层随不同的物理网络差异较大
- 好处：
  - 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。

局域网对 LLC 子层是透明的



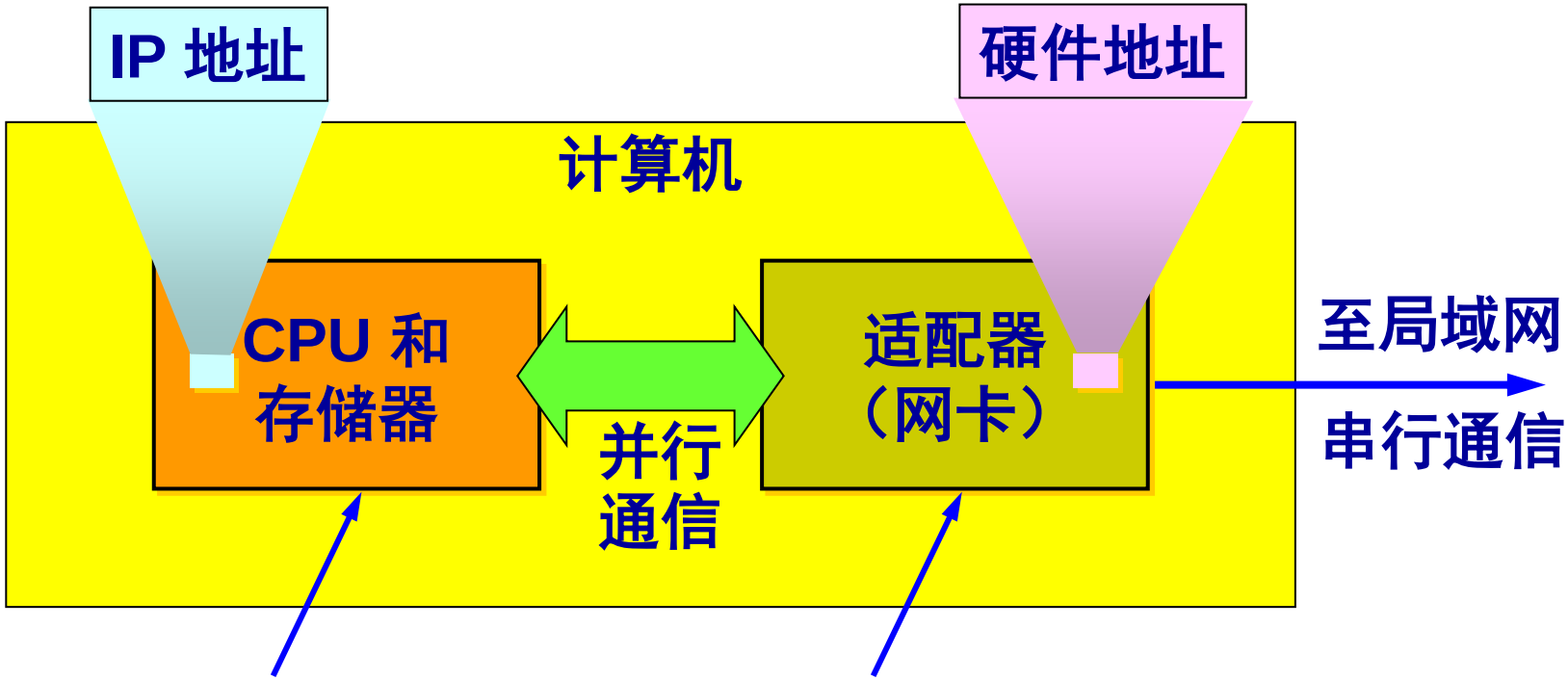
TCP/IP一般不考虑 LLC 子层

- 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，因此现在 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。

3.4.3 以太网的MAC层

- 计算机要连接到局域网需要依靠网络接口板。
- 网络接口板又称为适配器 (adapter) 或网络接口卡 NIC (Network Interface Card)，或“网卡”
- 适配器的重要功能：
  - 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



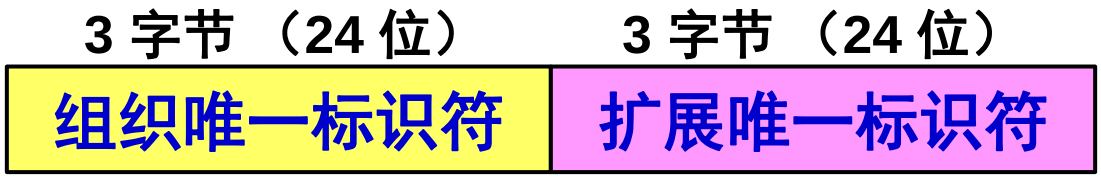
1. MAC 层的硬件地址

- 在局域网中，硬件地址又称为物理地址，或 MAC地址。
- IEEE 802 标准所说的“地址”严格地讲应当是每一个站的“名字”或标识符，采用 6 字节 (48位) ，是固化在网卡ROM中的。
- 但鉴于大家都早已习惯了将这种 48 位的“名字”称为“地址”，所以本书也采用这种习惯用法，尽管这种说法并不太严格。

请注意，如果连接在局域网上的主机或路由器安装有多个适配器，那么这样的主机或路由器就有多个“地址”。更准确些说，这种 48 位“地址”应当是某个接口的标识符。

48 位的 MAC 地址

- IEEE 的注册管理机构 RA 负责向生产局域网网卡的厂家分配地址字段 6 个字节中的前三个字节 (即高位 24 位)，称为组织唯一标识符OUI。
- 地址字段 6 个字节中的后三个字节 (即低位 24 位) 由厂家自行指派，称为扩展唯一标识符EUI，必须保证生产出的适配器没有重复地址。



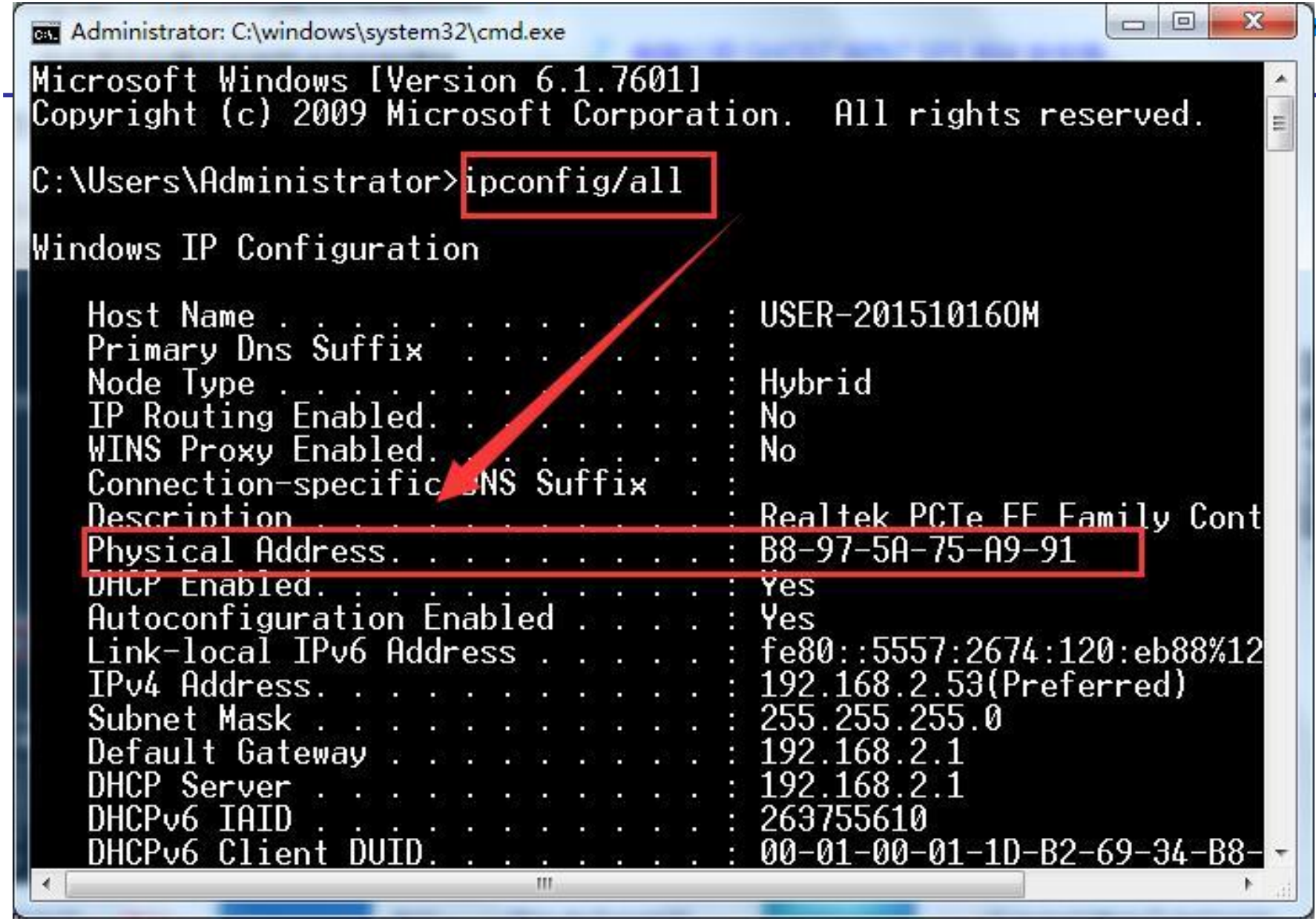
48 位的 MAC 地址

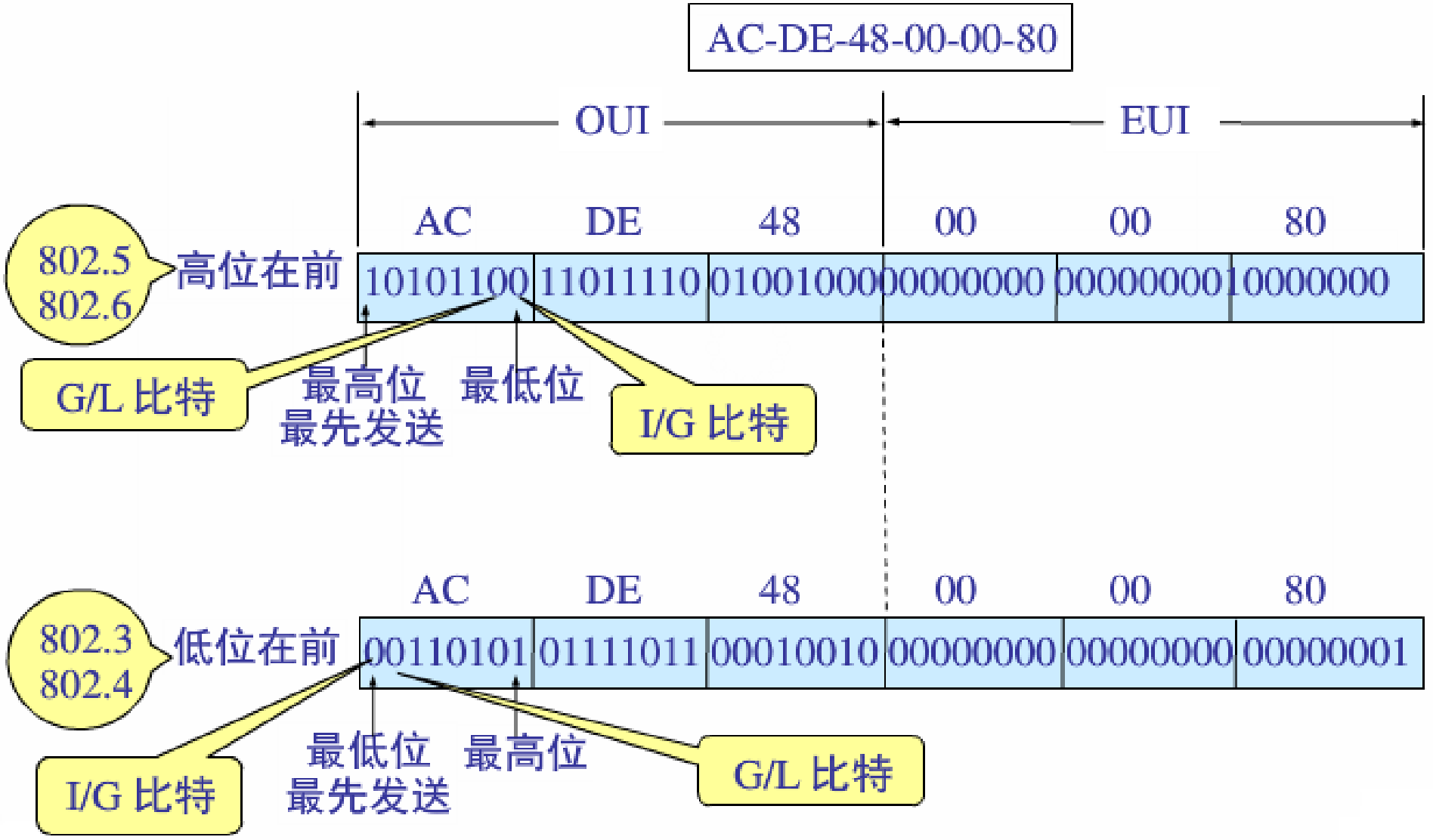
单站地址，组地址，广播地址

- IEEE 规定地址字段的第一字节的最低位为 I/G 位。I/G 表示 Individual / Group。
- 当 I/G位 = 0 时，地址字段表示一个单站地址。
- 当 I/G位 = 1 时，表示组地址，用来进行多播（以前曾译为组播）。此时，IEEE 只分配地址字段前三个字节中的 23 位。
- 当 I/G 位分别为 0 和 1 时，一个地址块可分别生成 2^{23} 个单个站地址和 2^{23} 个组地址。
- 所有 48 位都为 1 时，为广播地址。只能作为目的地址使用。

全球管理与本地管理

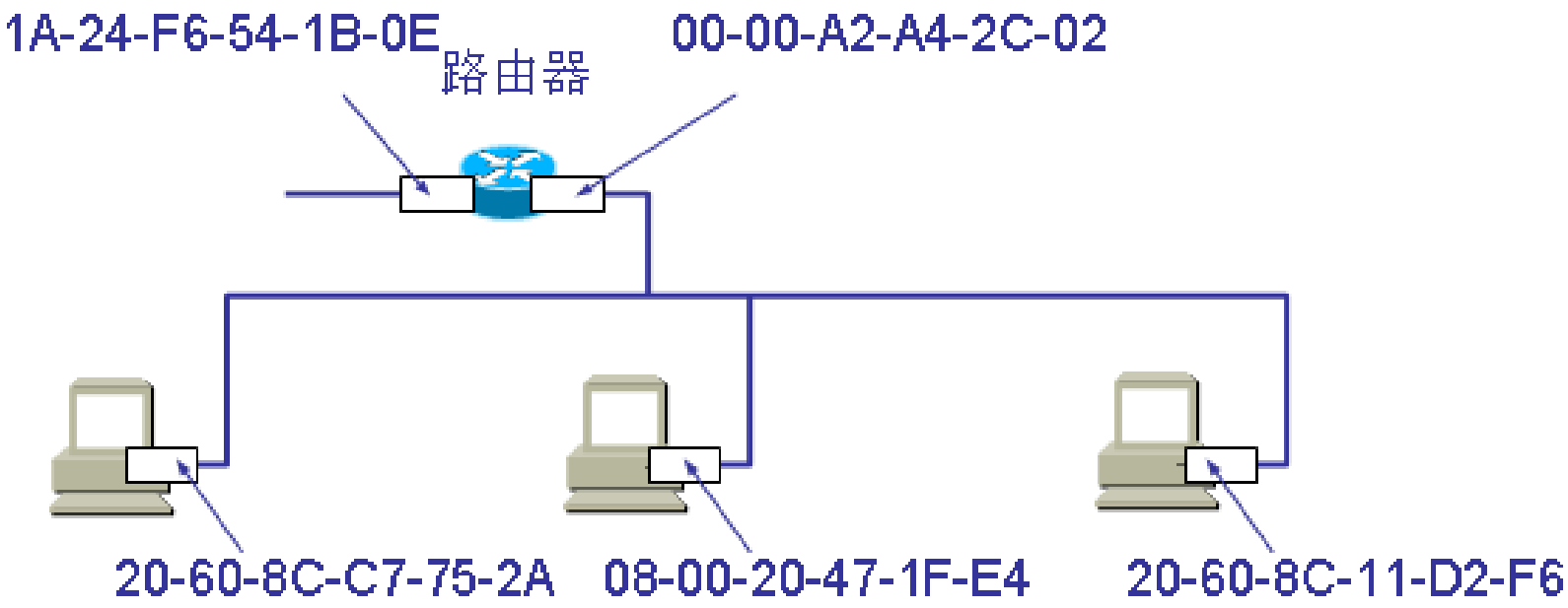
- IEEE 把地址字段第一字节的最低第 2 位规定为 G/L 位，表示 Global / Local。
- 当 G/L位 = 0 时，是全球管理（保证在全球没有相同的地址），厂商向IEEE购买的 OUI 都属于全球管理。
- 当 G/L位 = 1 时， 是本地管理，这时用户可任意分配网络上的地址。


Administrator: C:\windows\system32\cmd.exe
Microsoft Windows [Version 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. All rights reserved.
C:\Users\Administrator>ipconfig/all
Windows IP Configuration
Host Name . . . . . : USER-201510160M
Primary Dns Suffix . . . . . :
Node Type . . . . . : Hybrid
IP Routing Enabled. . . . . : No
WINS Proxy Enabled. . . . . : No
Connection-specific DNS Suffix . :
Description . . . . . : Realtek PCIe FE Family Cont
Physical Address. . . . . : B8-97-5A-75-A9-91
DHCP Enabled. . . . . : Yes
Autoconfiguration Enabled . . . . : Yes
Link-local IPv6 Address . . . . . : fe80::5557:2674:120:eb88%12
IPv4 Address. . . . . : 192.168.2.53(Preferred)
Subnet Mask . . . . . : 255.255.255.0
Default Gateway . . . . . : 192.168.2.1
DHCP Server . . . . . : 192.168.2.1
DHCPv6 IAID . . . . . : 263755610
DHCPv6 Client DUID. . . . . : 00-01-00-01-1D-B2-69-34-B8-


AC-DE-48-00-00-80
OUI
EUI
AC
DE
48
00
00
80
802.5
802.6
高位在前
10101100
11011110
01001000
00000000
00000000
10000000
G/L 比特
最高位
最先发送
最低位
I/G 比特
AC
DE
48
00
00
80
802.3
802.4
低位在前
00110101
01111011
00010010
00000000
00000000
00000001
I/G 比特
最低位
最先发送
最高位
G/L 比特

网卡上的硬件地址

路由器由于同时连接到两个网络上，
因此它有两块网卡和两个硬件地址。



适配器检查 MAC 地址

- 适配器从网络上每收到一个 MAC 帧就首先用硬件检查 MAC 帧中的 MAC 地址。
  - 如果是发往本站的帧则收下，然后再进行其他的处理。
  - 否则就将此帧丢弃，不再进行其他的处理。
- “发往本站的帧”包括以下三种帧：
  - 单播 (unicast) 帧（一对一）
  - 广播 (broadcast) 帧（一对全体）
  - 多播 (multicast) 帧（一对多）

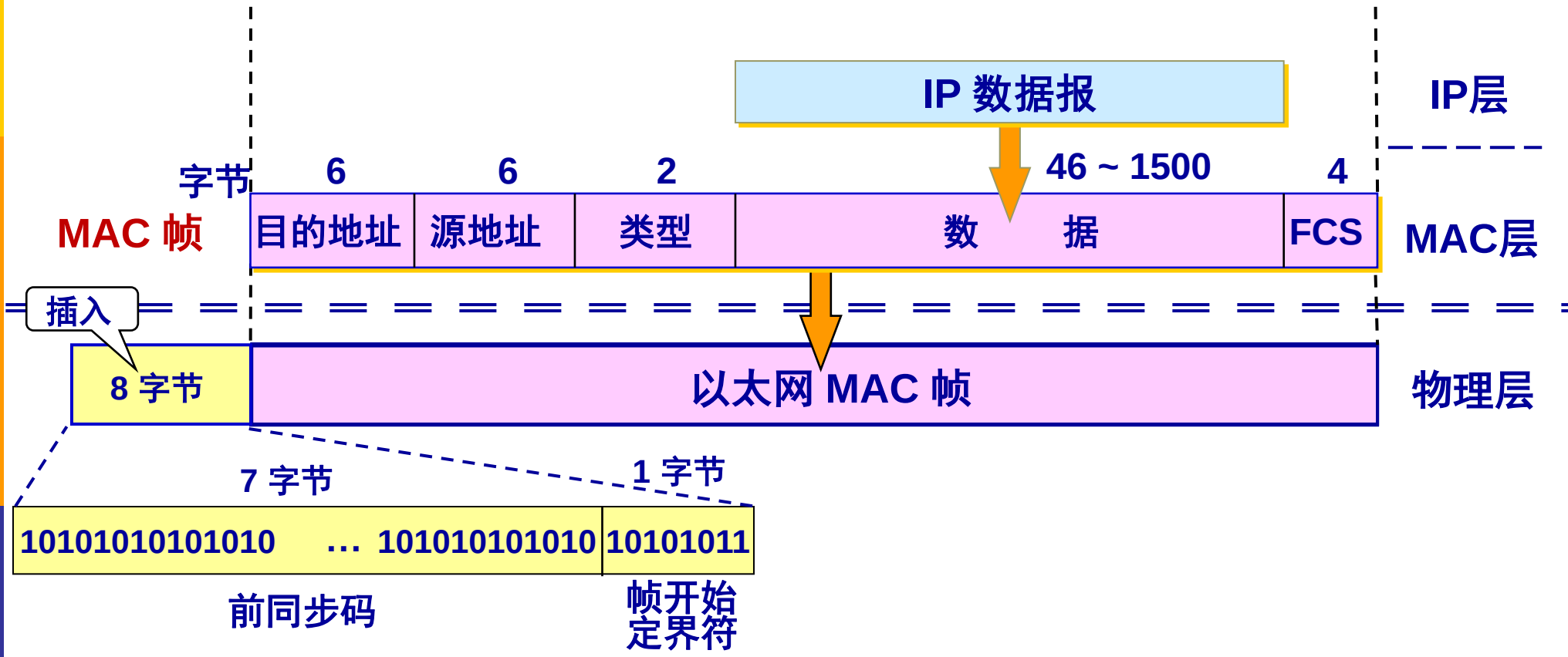
适配器检查 MAC 地址

- 所有的适配器都至少能够识别前两种帧，即能够识别单播地址和广播地址。
- 有的适配器可用编程方法识别多播地址。
- 只有目的地址才能使用广播地址和多播地址。
- 以混杂方式 (promiscuous mode) 工作的以太网适配器只要“听到”有帧在以太网上传输就都接收下来。

2. 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。

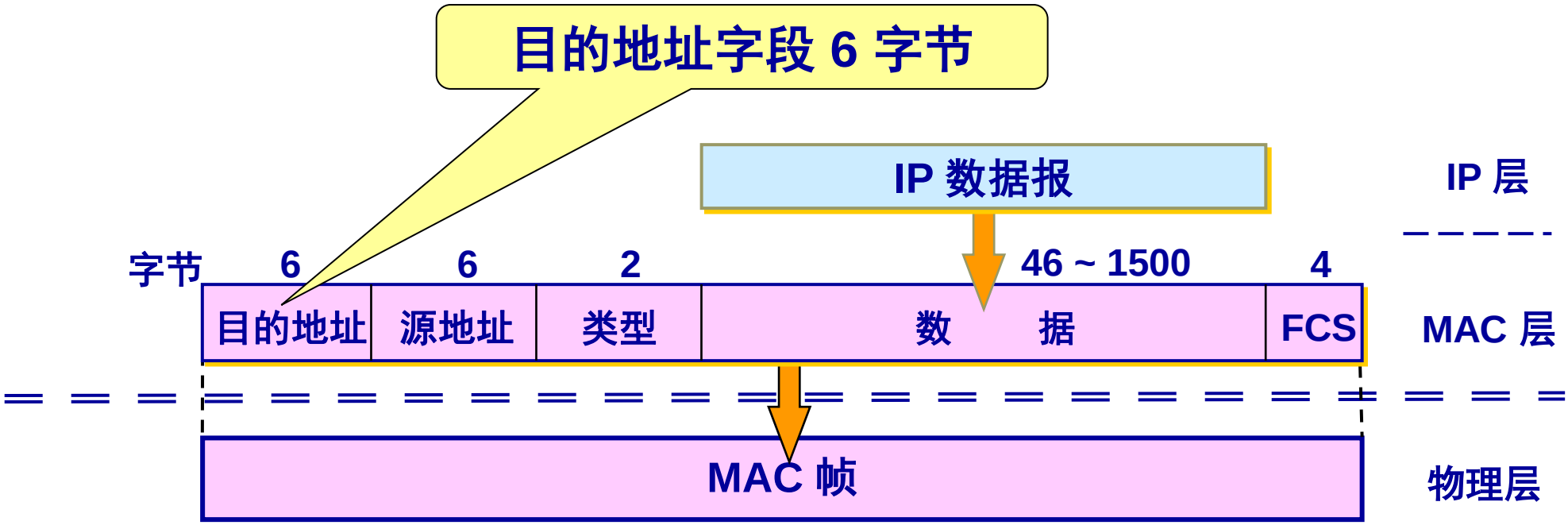
以太网V2的 MAC 帧格式



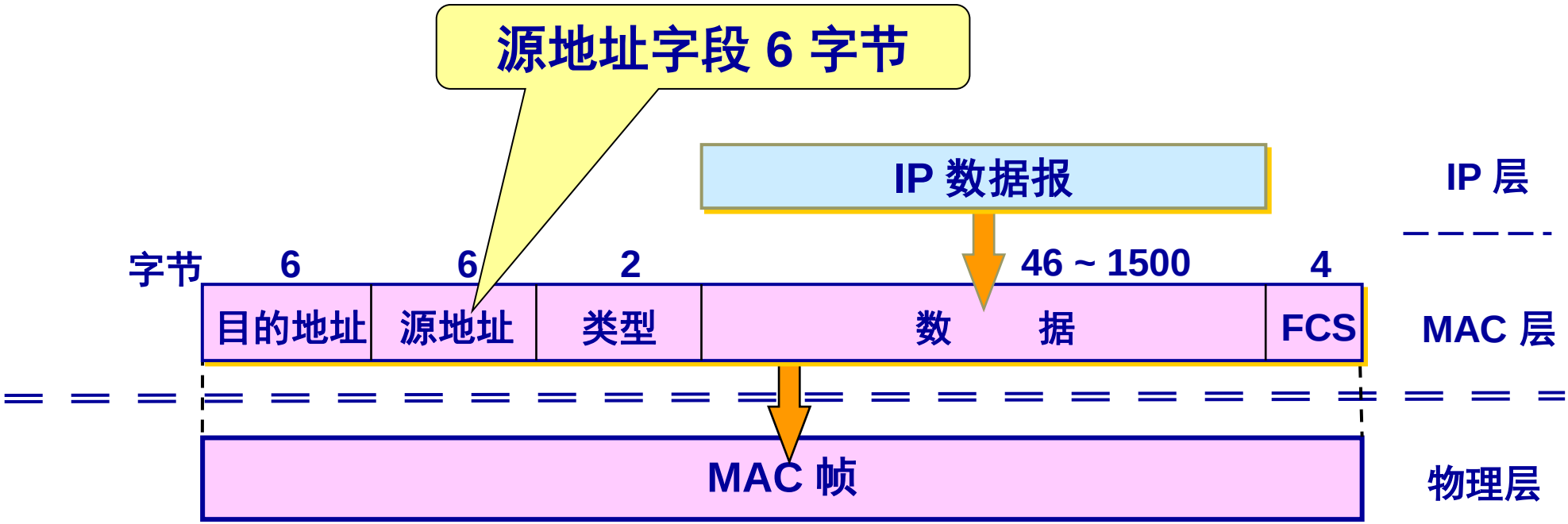
MAC帧格式

- **类型字段**
  - 用来标识上一层使用的是什么协议，以便把收到的MAC帧交给上一层的协议。
  - Xerox公司负责类型字段的代码分配：
    - 类型字段为0x0800表示上层使用IP数据报；
    - 类型字段为0x8137表示上层使用Novell IPX数据报；
- **同步码**
  - 实际传送的比MAC帧多8个字节；
  - 因为当一个站开始接收MAC帧时，没有与到达的比特流同步，因此MAC帧的开始若干比特无法接收，这样使得整个帧无效。这样需要插入同步码。

以太网 V2 的 MAC 帧格式

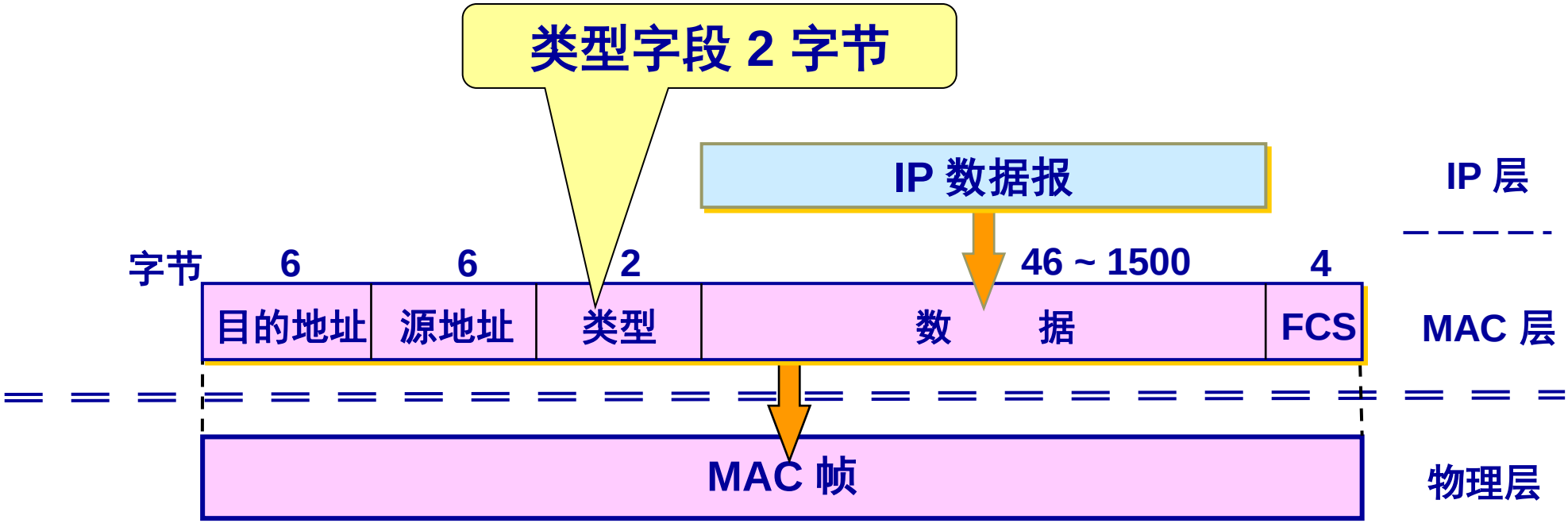


以太网 V2 的 MAC 帧格式



以太网 V2 的 MAC 帧格式

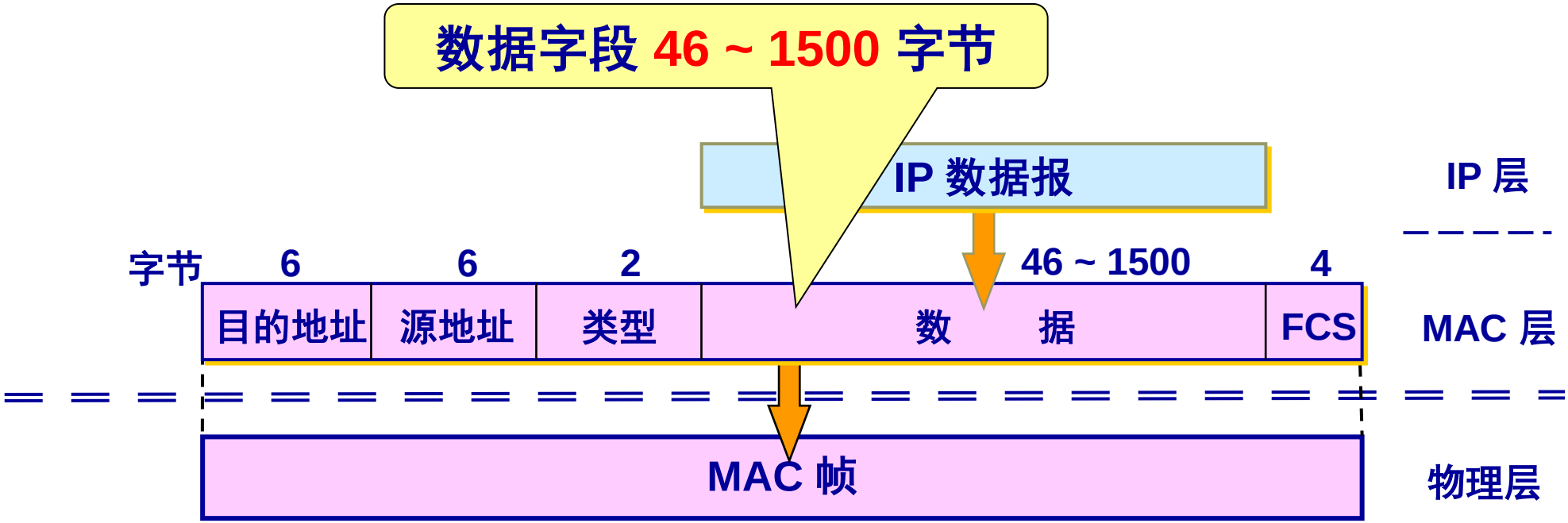
类型字段用来标志上一层使用的是什么协议，
以便把收到的 MAC 帧的数据上交给上一层的这个协议。



以太网 V2 的 MAC 帧格式

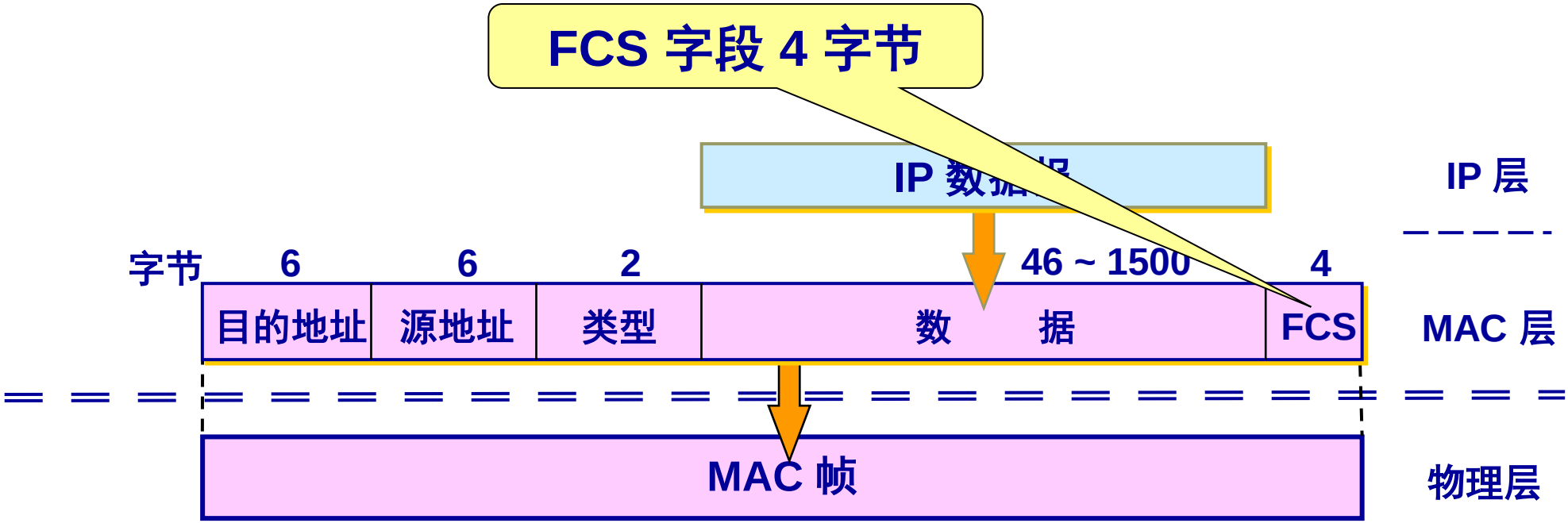
数据字段的正式名称是 MAC 客户数据字段。

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（46字节）



以太网 V2 的 MAC 帧格式

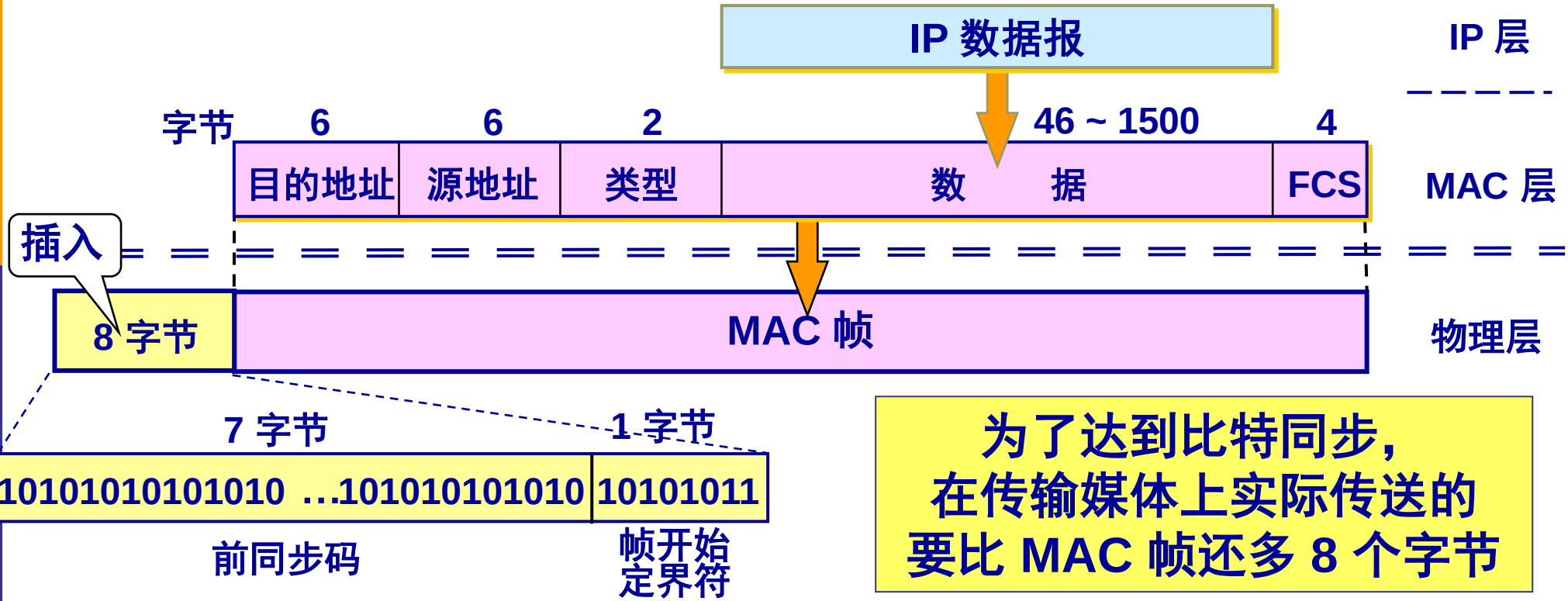
当传输媒体的误码率为 1×10^{-8} 时，
MAC 子层可使未检测到的差错小于 1×10^{-14}。



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（硬件生成）的 8 字节中，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 1 个字节是帧开始定界符，表示后面的信息就是 MAC 帧。



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

无效的 MAC 帧

- 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。

IEEE 802.3 MAC 帧格式

与以太网V2 MAC 帧格式相似，区别在于：

- (1) IEEE 802.3 规定的 MAC 帧的第三个字段是“长度 / 类型”。
  - 当这个字段值大于 0x0600 时（相当于十进制的 1536），就表示“类型”。这样的帧和以太网 V2 MAC 帧完全一样。
  - 当这个字段值小于 0x0600 时才表示“长度”。
- (2) 当“长度/类型”字段值小于 0x0600 时，数据字段必须装入上面的逻辑链路控制 LLC 子层的 LLC 帧。

现在市场上流行的都是以太网V2 的 MAC 帧，但大家也常常把它称为 IEEE 802.3 标准的 MAC 帧。

3.4.4 扩展以太网

- 在物理层扩展以太网
- 在数据链路层扩展以太网

局域网互联与互联设备

- 为什么要进行局域网互联？
  - (1) 局域网覆盖的距离有限
    - 单个局域网覆盖的距离往往不能满足应用的需要。
  - (2) 局域网能支持的连网计算机数目有限
    - 单个局域网所能连接的计算机数目往往不能满足应用的需要
  - (3) 局域网上能传输的通信量有限
    - 单个局域网所容许的通信量往往不能满足应用的需要。

局域网互联设备

- 互联的本质
  - 由于网络是分层次实现的，而局域网又各有不同的标准，因此网络互联的本质就是在不同的协议层次上实现协议的彼此转换。
- 互联设备
  - 转发器（Repeater，重发器，中继器）
  - 集线器（Hub）
  - 网桥（Bridge）
  - 交换机（Switch）

在物理层上扩展

在数据链路层上扩展

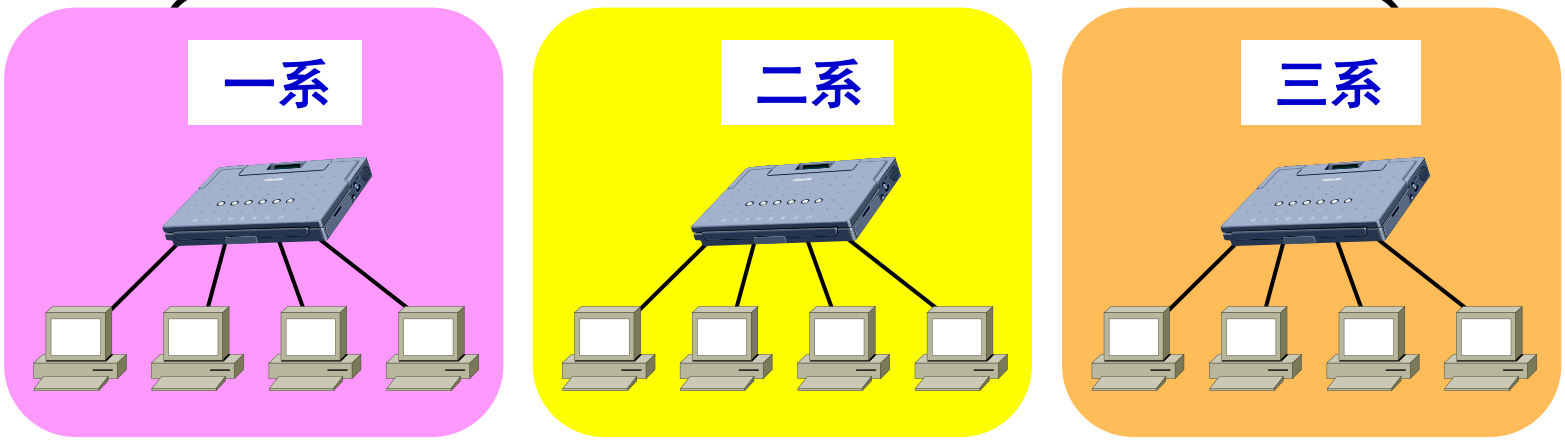
转发器

- 转发器又称重发器、中继器，它工作于在物理层上，用来扩展以太网的地理覆盖范围。
- 转发器的功能
  - 常用于连接两个同轴电缆以太网，将信号放大整形后，以延伸网络的传输距离。
  - 不具有信号通路的选择功能
  - 随着双绞线成为以太网的主流，已很少使用转发器

集线器

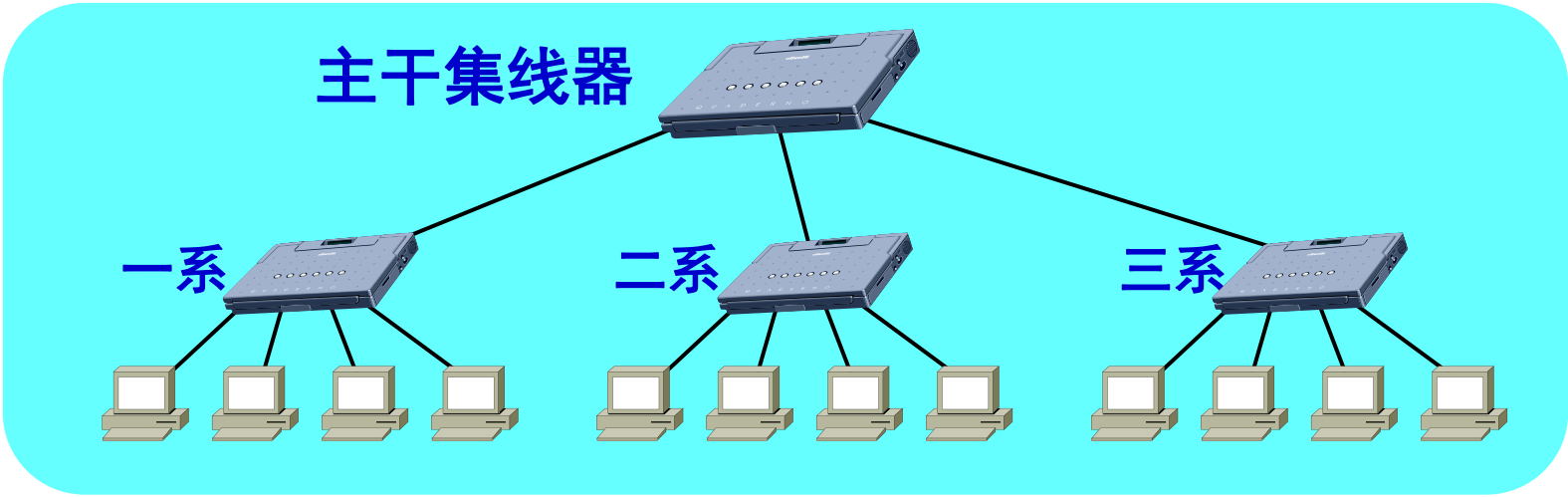
- 使用多个集线器可连成更大的、多级星形结构的以太网。
  - 例如，一个学院的三个系各有一个 10BASE-T 以太网，可通过一个主干集线器把各系的以太网连接起来，成为一个更大的以太网。
  - 好处：使这个学院不同系的以太网上的计算机能够进行跨系通信；扩大了以太网覆盖的地理范围。
  - 主机与集线器之间的最大距离是100m，两台主机之间的最大距离是200m，但集线器之间的距离可以是100m（使用双绞线）或者更远（使用光纤）。

三个独立的碰撞域



三个独立的以太网

一个更大的碰撞域



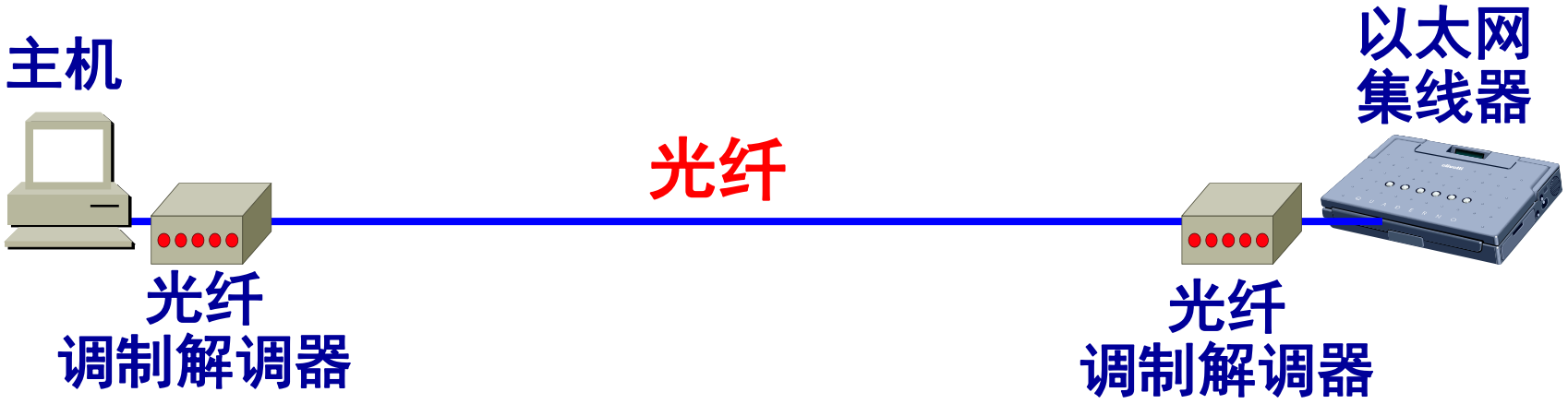
一个扩展的以太网

集线器

- (1) 集线器是使用电子器件来模拟实际电缆线的工作，因此整个系统仍然像一个传统的以太网那样运行。
- (2) 使用集线器的以太网在逻辑上仍是一个总线网，各工作站使用的还是 CSMA/CD 协议，并共享逻辑上的总线。
- (3) 集线器很像一个多接口的转发器，工作在物理层。
- (4) 集线器采用了专门的芯片，进行自适应串音回波抵消，减少了近端串音。

集线器

- **使用光纤扩展**
  - 主机使用光纤（通常是一对光纤）和一对光纤调制解调器连接到集线器。
  - 很容易使主机和几公里以外的集线器相连接。



主机使用光纤和一对光纤调制解调器连接到集线器

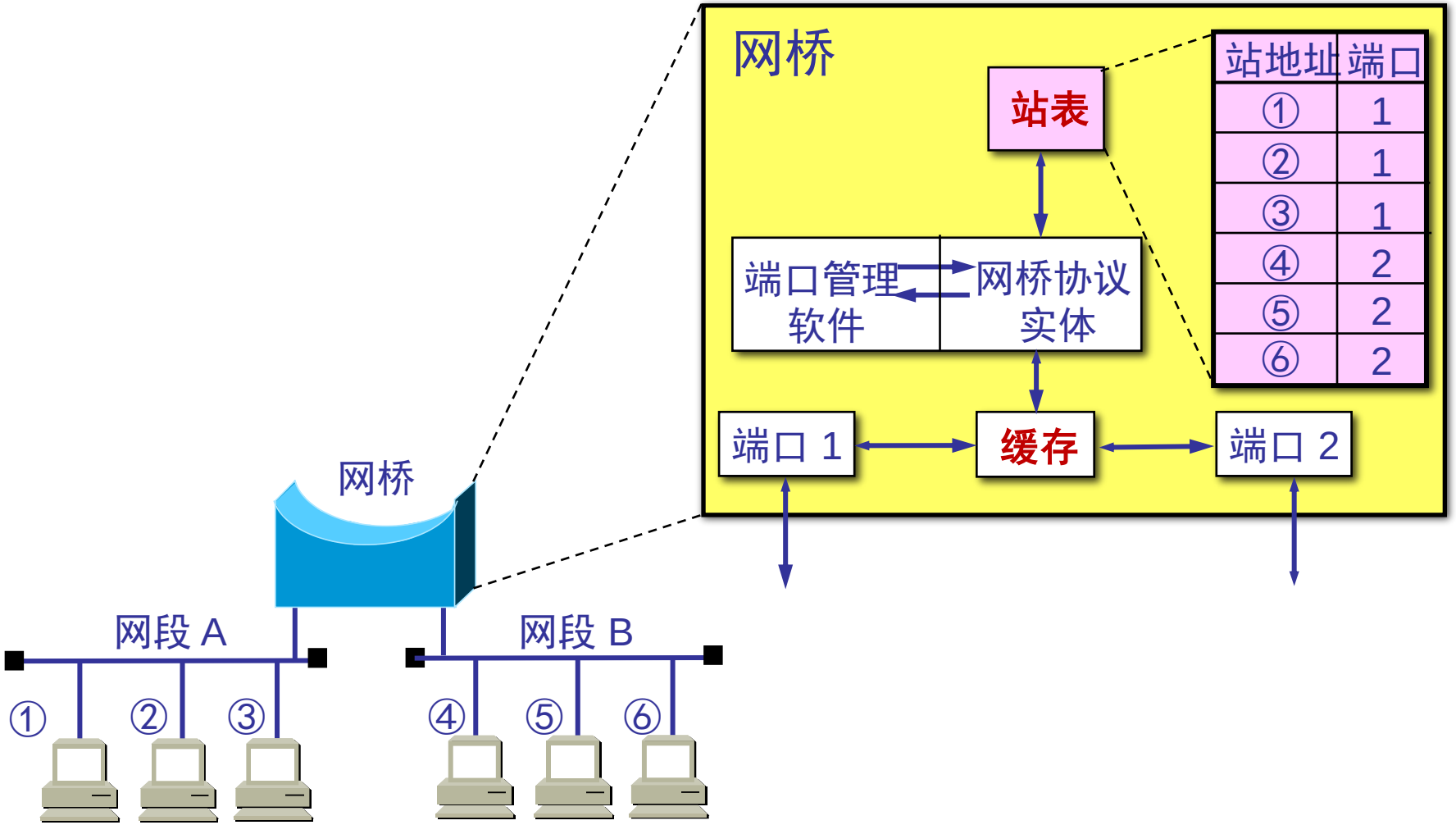
用集线器扩展以太网

- 优点
  - 使原来属于不同碰撞域的以太网上的计算机能够进行跨碰撞域的通信。
  - 扩大了以太网覆盖的地理范围。
- 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
  - 集线器是个多借口的转发器，不能把帧进行缓存

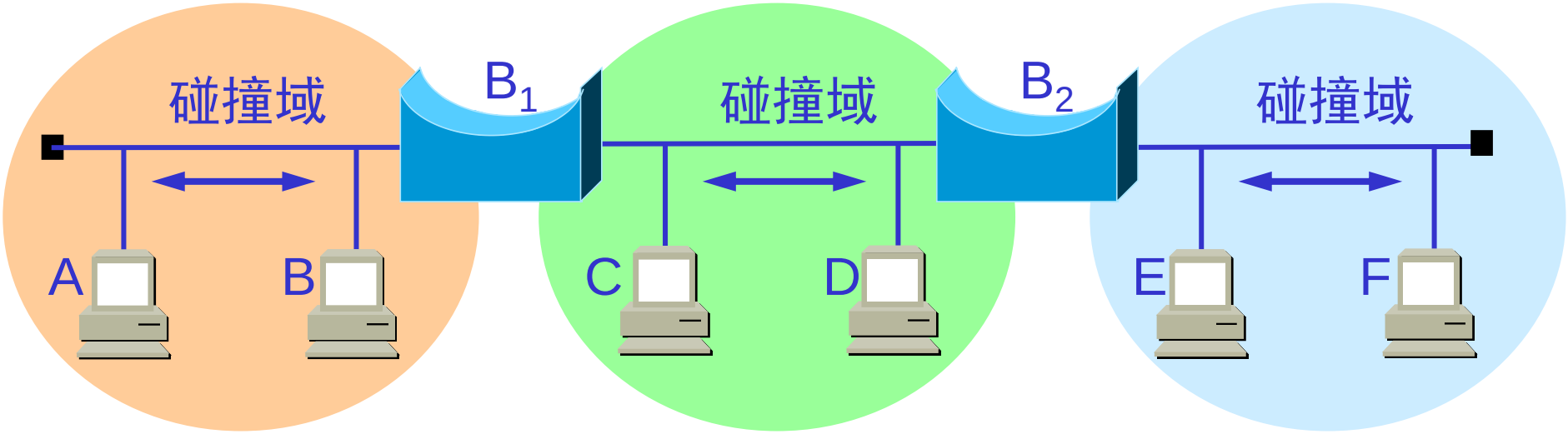
网桥

- 网桥工作在数据链路层。
- 根据 MAC 帧的目的地址对收到的帧进行转发和过滤
- 网桥的工作原理
  - 网桥从端口接收网段上传送的各种帧；
  - 每当收到一个帧时，先暂存在缓存中。
  - 若此帧未出错，且欲发送的目的站的MAC地址属于另外一个网段，则通过查找“转发表”，将收到的帧送往对应的端口转发。
  - 若此帧出错，则丢弃该帧。
  - 同一个网段内的帧，不会被网桥转发，不会增加网络负担

网桥的内部结构


网桥
站表
站地址 端口
① 1
② 1
③ 1
④ 2
⑤ 2
⑥ 2
端口管理软件
网桥协议实体
端口 1
缓存
端口 2
网桥
网段 A
网段 B
①
②
③
④
⑤
⑥

网桥使各网段成为隔离开的碰撞域

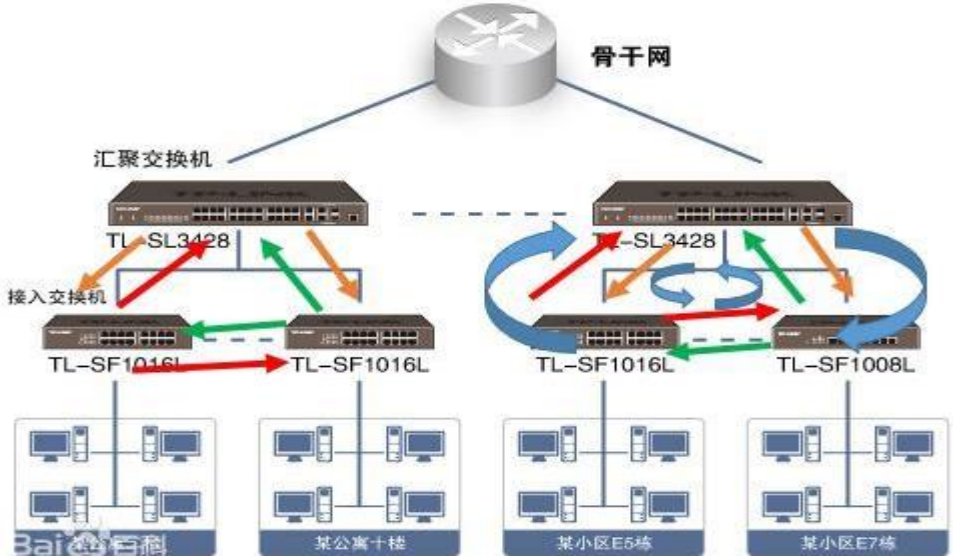


使用网桥的优缺点

- 优点
  - 过滤通信量、扩大了物理范围、提高了可靠性。
  - 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 以太网）的局域网。
- 缺点
  - 存储转发增加了时延。
  - 在 MAC 子层并没有流量控制功能。
  - 具有不同 MAC 子层的网段桥接在一起时时延更大。
  - 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。这就是所谓的广播风暴。

广播风暴

- 广播风暴（broadcast storm）简单的讲是指当广播数据充斥网络无法处理，并占用大量网络带宽，导致正常业务不能运行，甚至彻底瘫痪，这就发生了“广播风暴”。
- 一个数据帧或包被传输到本地网段 （由广播域定义）上的每个节点就是广播；由于网络拓扑的设计和连接问题，或其他原因导致广播在网段内大量复制，传播数据帧，导致网络性能下降，甚至网络瘫痪，这就是广播风暴。



广播风暴

- 广播风暴的产生有多种原因，如蠕虫病毒、交换机端口故障、网卡故障、链路冗余没有启用生成树协议、网线线序错误或受到干扰等。从目前来看，蠕虫病毒和ARP攻击是造成网络广播风暴最主要的原因。

- 在一些使用集线器的网络中仍然非常常见。解决网络广播风暴最快捷的方法是给集线器断电然后上电启动即可，但这只是治标不治本的方法，要彻底解决，最好使用交换机设备，并划分VLAN、通过端口控制网络广播风暴

网桥和集线器的不同

- 集线器在转发帧时，不对传输媒体进行检测。
- 网桥在转发帧之前必须执行 CSMA/CD 算法。
  - 若在发送过程中出现碰撞，就必须停止发送和进行退避。
  - 在这一点上网桥的接口很像一个网卡。但网桥却没有网卡。
- 由于网桥没有网卡，因此网桥并不改变它转发的帧的源地址。

多端口网桥——以太网交换机

- 1990 年问世的交换式集线器 (switching hub) 可明显地提高以太网的性能。
- 交换式集线器常称为以太网交换机 (switch) 或第二层交换机 (L2 switch)，强调这种交换机工作在数据链路层。
- 以太网交换机实质上就是一个多接口的网桥。
  - 通常都有十几个或更多的接口。

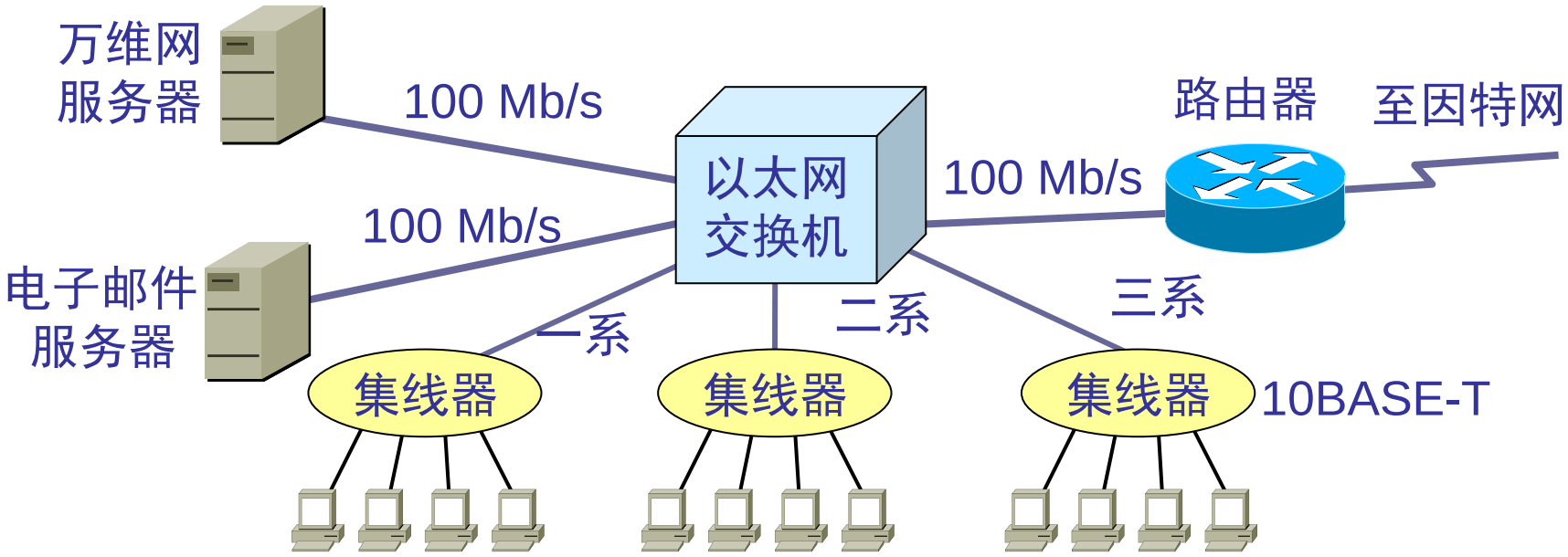
以太网交换机的特点

- 每个接口都直接与一个单台主机或另一个以太网交换机相连，并且一般都工作在全双工方式。
- 以太网交换机具有并行性。
  - 能同时连通多对接口，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。
- 相互通信的主机都是独占传输媒体，无碰撞地传输数据。
- 以太网交换机的接口有存储器，能在输出端口繁忙时把到来的帧进行缓存。
- 以太网交换机是一种即插即用设备，其内部的帧交换表（又称为地址表）是通过自学习算法自动地逐渐建立起来
- 以太网交换机使用了专用的交换结构芯片，用硬件转发，其转发速率要比使用软件转发的网桥快很多。

以太网交换机的优点

- **用户独享带宽，增加了总容量。**
  - 对于普通 10 Mbit/s 的共享式以太网，若共有 *N* 个用户，则每个用户占有的平均带宽只有总带宽 (10 Mbit/s)的 *N* 分之一。
  - 使用以太网交换机（交换式以太网）时，虽然在每个接口到主机的带宽还是 10 Mbit/s，但由于一个用户在通信时是独占而不是和其他网络用户共享传输媒体的带宽，因此对于拥有 *N* 个接口的交换机的总容量为 *N*×10Mbit/s

用以太网交换机扩展局域网



以太网交换机 v.s. 集线器

- 有10个站连接在以太网上。试计算以下三种情况下每一个站所能得到的带宽。
  - （1） 10个站都连接到一个10Mb/s以太网集线器。
  - （2） 10个站都连接到一个100Mb/s以太网集线器。
  - （3） 10个站都连接到一个10Mb/s以太网交换机。

- 答：（1）10个站共享10Mb/s。 （2）10个站共享100Mb/s。 （3）每个站独占10Mb/s。

以太网交换机的交换方式

- 存储转发方式
  - 把整个数据帧先缓存后再进行处理。
- 直通 (cut-through) 方式
  - 接收数据帧的同时就立即按数据帧的目的 MAC 地址决定该帧的转发接口，因而提高了帧的转发速度。
  - 缺点是它不检查差错就直接将帧转发出去，因此有可能也将一些无效帧转发给其他的站。

在某些情况下，仍需要采用基于软件的存储转发方式进行交换，例如，当需要进行线路速率匹配、协议转换或差错检测时。

以太网交换机的自学习功能

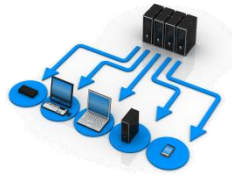
- 以太网交换机运行自学习算法自动维护交换表。
- 开始时，以太网交换机里面的交换表是空的。

交换表一开始是空的

自学习算法

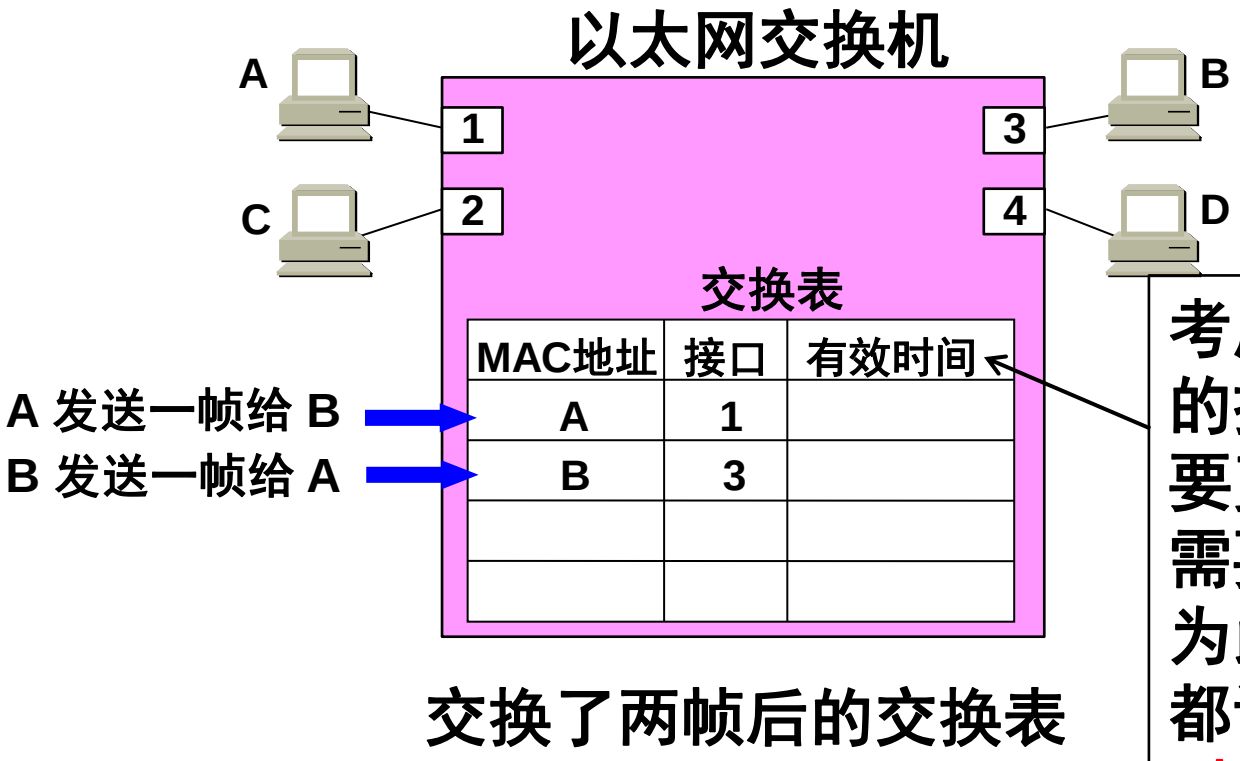
- A 先向 B 发送一帧，从接口 1 进入到交换机。
- 交换机收到帧后，先查找交换表，没有查到应从哪个接口转发这个帧。
- 交换机把这个帧的源地址 A 和接口 1 写入交换表中，并向除接口1以外的所有的接口广播这个帧。
- C 和 D 将丢弃这个帧，因为目的地址不对。只 B 才收下这个目的地址正确的帧。这也称为过滤。
- 从新写入交换表的项目 (A, 1) 可以看出，以后不管从哪一个接口收到帧，只要其目的地址是A，就应当把收到的帧从接口1转发出去。

自学习算法



- B 通过接口 3 向 A 发送一帧。
- 交换机查找交换表，发现交换表中的 MAC 地址有 A。表明要发送给A的帧（即目的地址为 A 的帧）应从接口1转发。于是就把这个帧传送到接口 1 转发给 A。显然，现在已经没有必要再广播收到的帧。
- 交换表这时新增加的项目 (B, 3)，表明今后如有发送给 B 的帧，就应当从接口 3 转发出去。
- 经过一段时间后，只要主机 C 和 D 也向其他主机发送帧，以太网交换机中的交换表就会把转发到 C 或 D 应当经过的接口号（2 或 4）写入到交换表中。

自学习算法



交换了两帧后的交换表

考虑到可能有时要在交换机的接口更换主机，或者主机要更换其网络适配器，这就需要更新交换表中的项目。为此，在交换表中每个项目都设有一定的有效时间（计时器）。过期的项目就自动被删除。

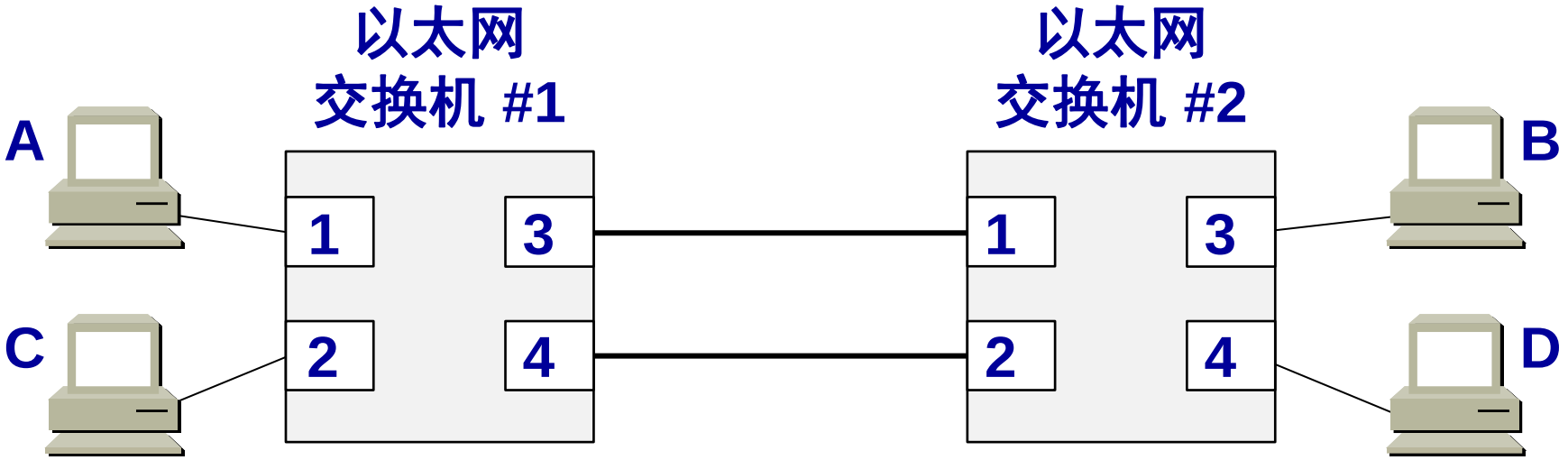
以太网交换机的这种自学习方法使得以太网交换机能够即插即用，不必人工进行配置，因此非常方便。

交换机自学习和转发帧的步骤归纳

- 交换机收到一帧后先查找交换表中与收到帧的**源地址**有无相匹配的项目。
  - 如没有，就在交换表中增加一个项目（源地址、进入的接口和有效时间）。
  - 如有，则把原有的项目进行更新（进入的接口或有效时间）。
- **转发帧。**查找交换表中与收到帧的**目的地址有无相匹配**的项目。
  - 如没有，则向所有其他接口（进入的接口除外）转发(Flooding)。
  - 如有，则按交换表中给出的接口进行转发。
  - 若交换表中给出的接口就是该帧进入交换机的接口，则应丢弃这个帧（因为这时不需要经过交换机进行转发）。

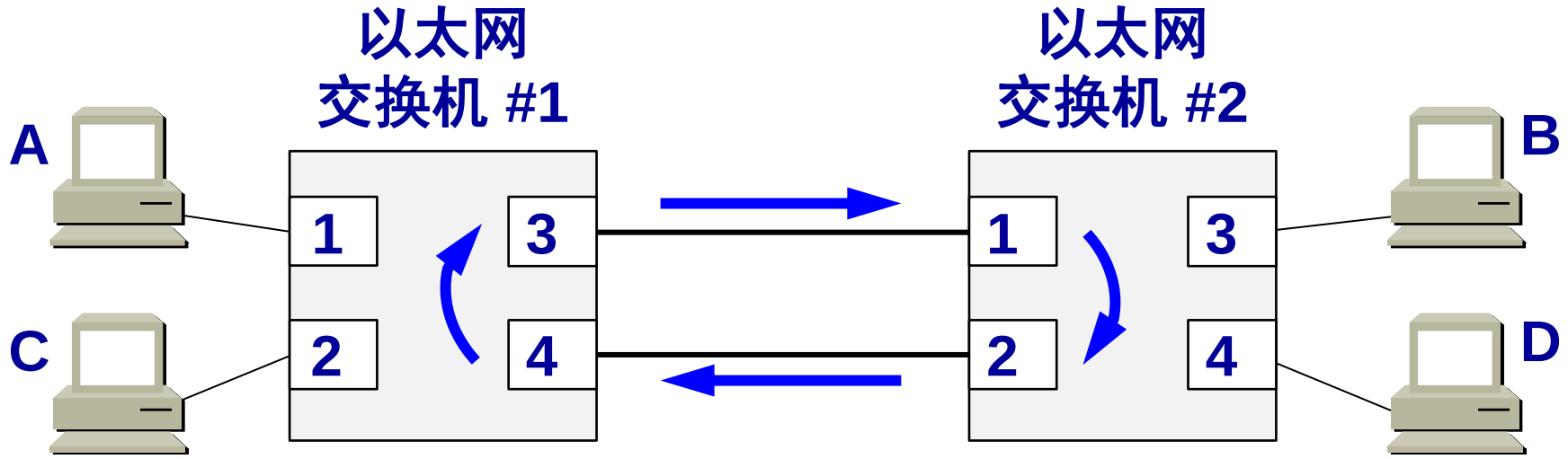
交换机使用了生成树协议

- 为了增加网络的可靠性，会增加冗余链路。自学习的过程就可能导致以太网帧在网络的某个环路中无限制地兜圈子。
- 如图，假定开始时，交换机 #1 和 #2 的交换表都是空的，主机 A 通过接口交换机 #1 向主机 B 发送一帧。



交换机使用了生成树协议

- 按交换机自学习和转发方法，该帧的某个走向如下：离开交换机 #1 的接口 3 → 交换机 #2 的接口 1 → 接口 2 → 交换机 #1 的接口 4 → 接口 3 → 交换机 #2 的接口 1 →……。**这样就无限制地循环兜圈子下去，白白消耗了网络资源。**

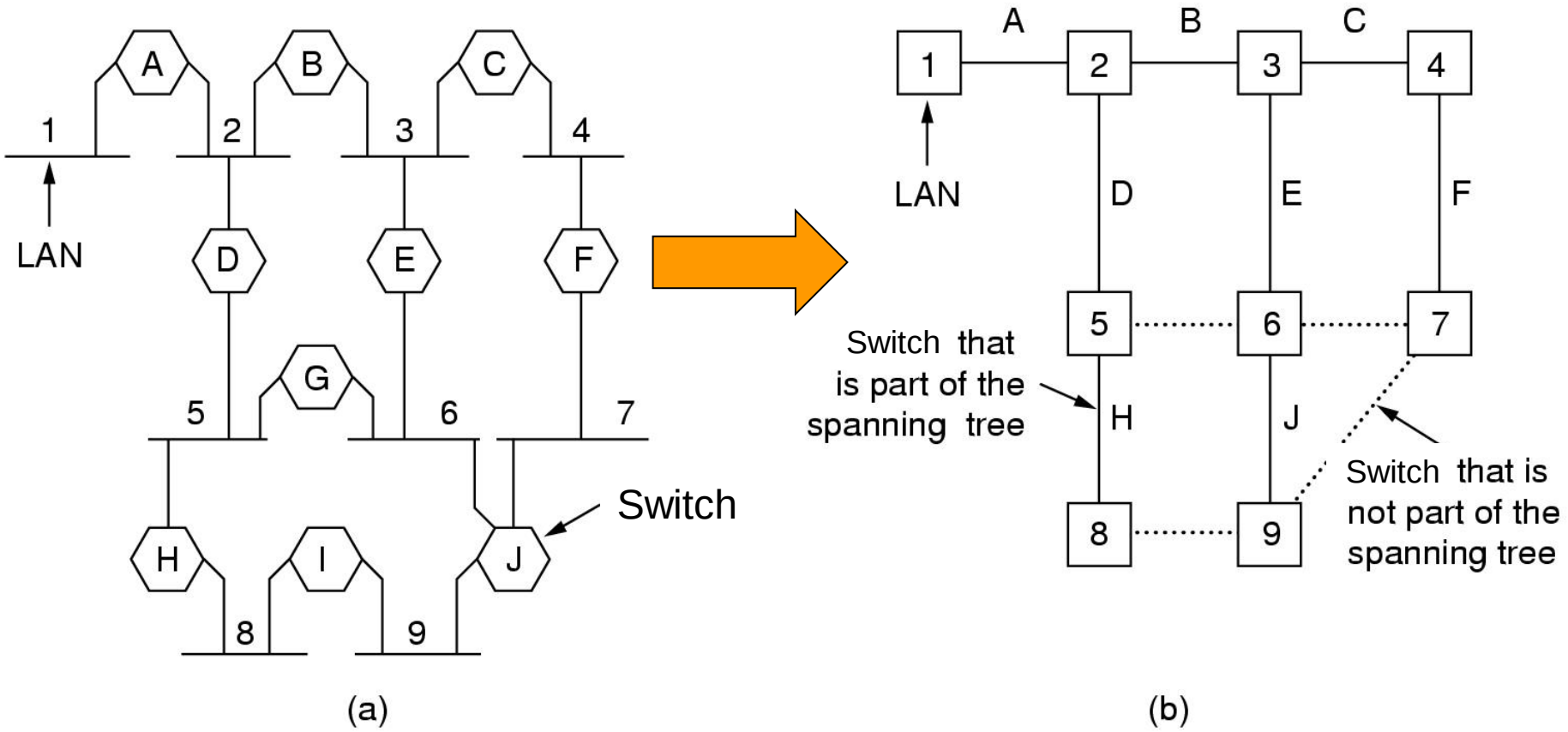


在两个交换机之间兜圈子的帧

交换机使用了生成树协议

- IEEE 802.1D 标准制定了一个生成树协议 STP (Spanning Tree Protocol)。
- 其要点是：不改变网络的实际拓扑，但在逻辑上则切断某些链路，使得从一台主机到所有其他主机的路径是无环路的树状结构，从而消除了兜圈子现象。

生成树算法



基本原理：选择一个交换机作为生成树的根，然后以最短路径为依据，找到树上的每一个结点，使整个连通的网络中不存在回路

从总线以太网到星形以太网

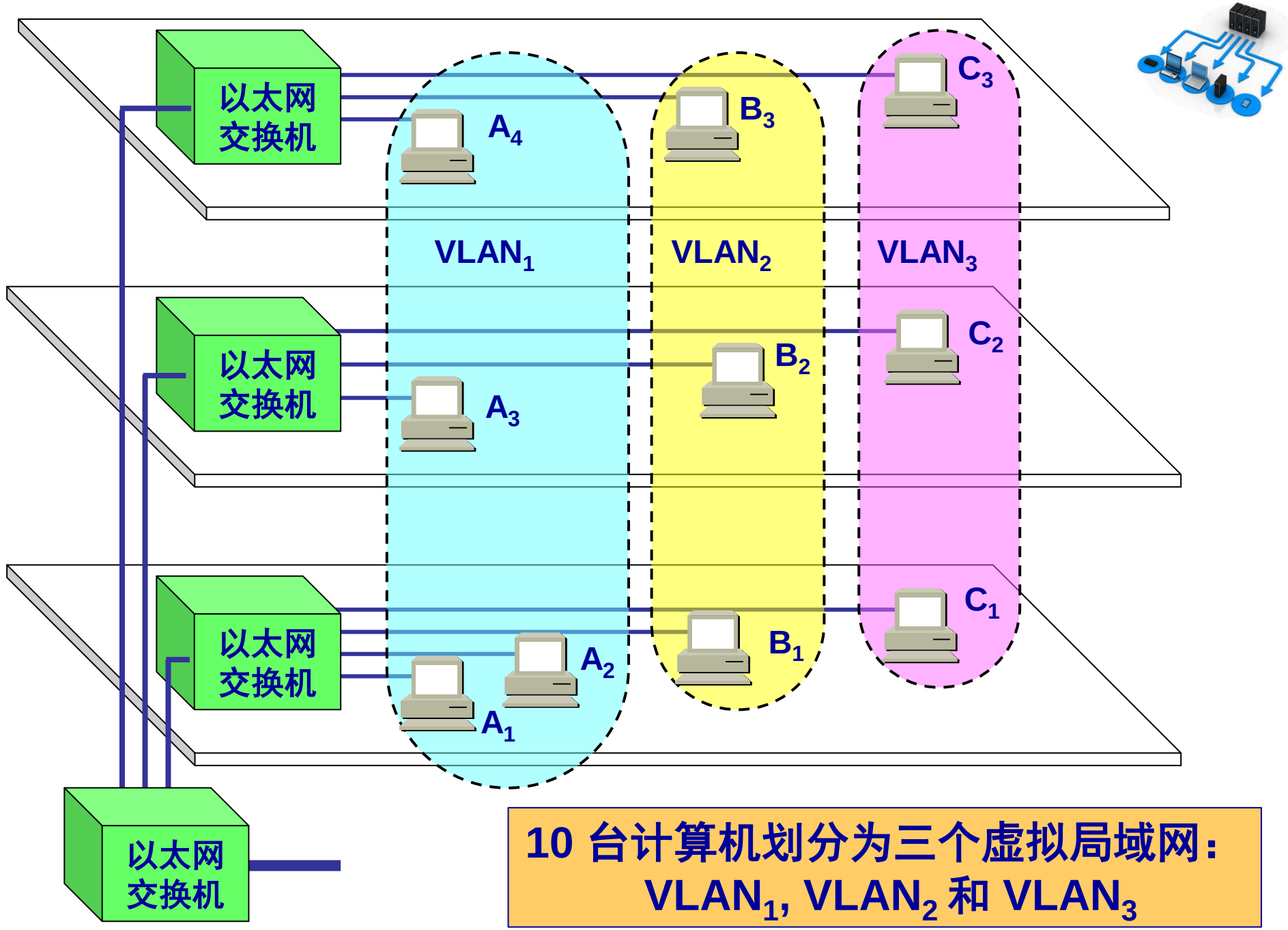
- 早期，以太网采用无源的总线结构。
- 现在，采用以太网交换机的星形结构成为以太网的首选拓扑。

- 总线以太网使用 CSMA/CD 协议，以半双工方式工作。
- 以太网交换机不使用共享总线，没有碰撞问题，因此不使用 CSMA/CD 协议，而是以全双工方式工作。但仍然采用以太网的帧结构。

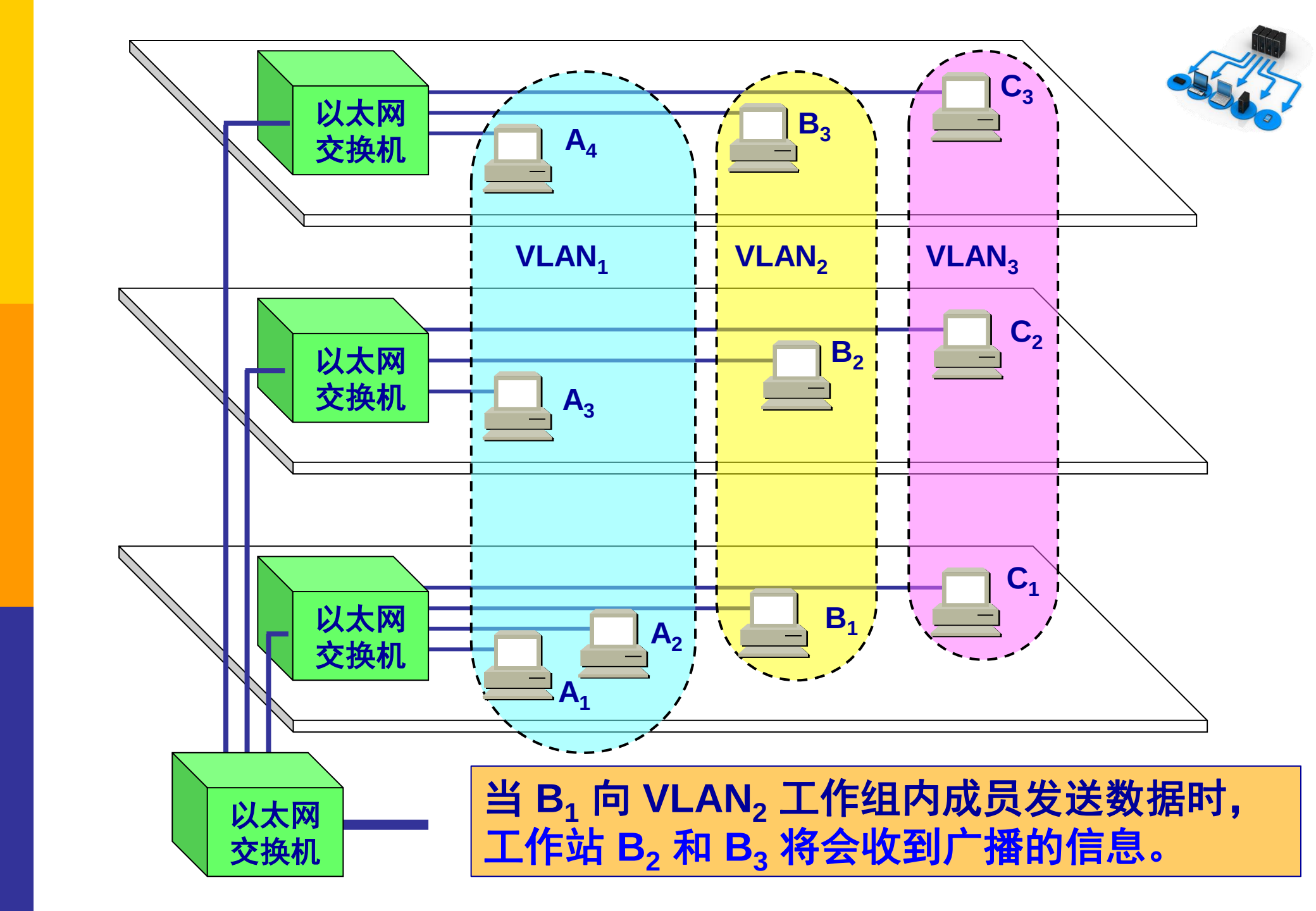
3.4.5 虚拟局域网

- 利用以太网交换机可以很方便地实现虚拟局域网 VLAN (Virtual LAN)。
- 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，指明发送这个帧的计算机是属于哪一个 VLAN。

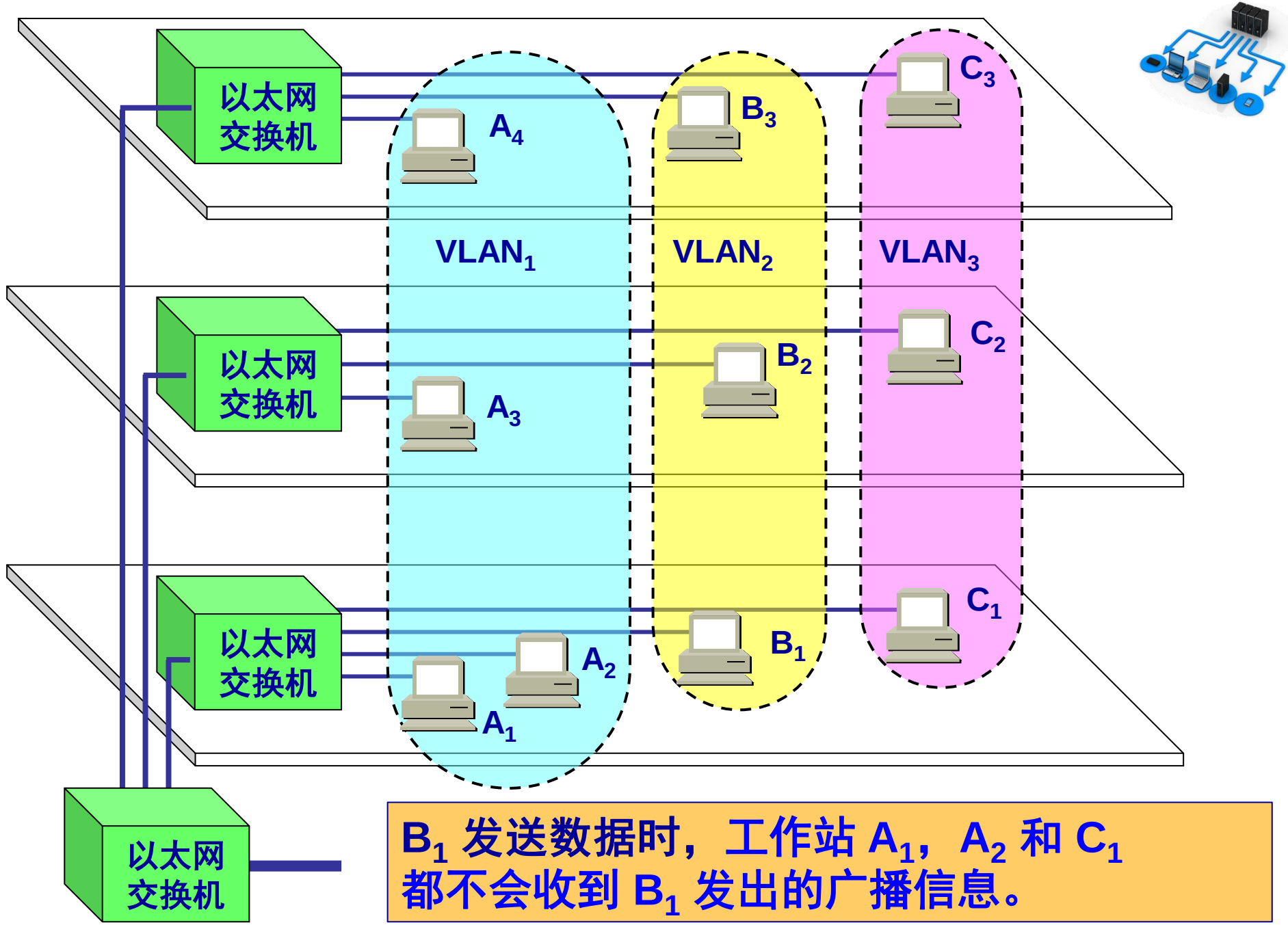
- 虚拟局域网其实只是局域网给用户提供的一种服务，而并不是一种新型局域网。



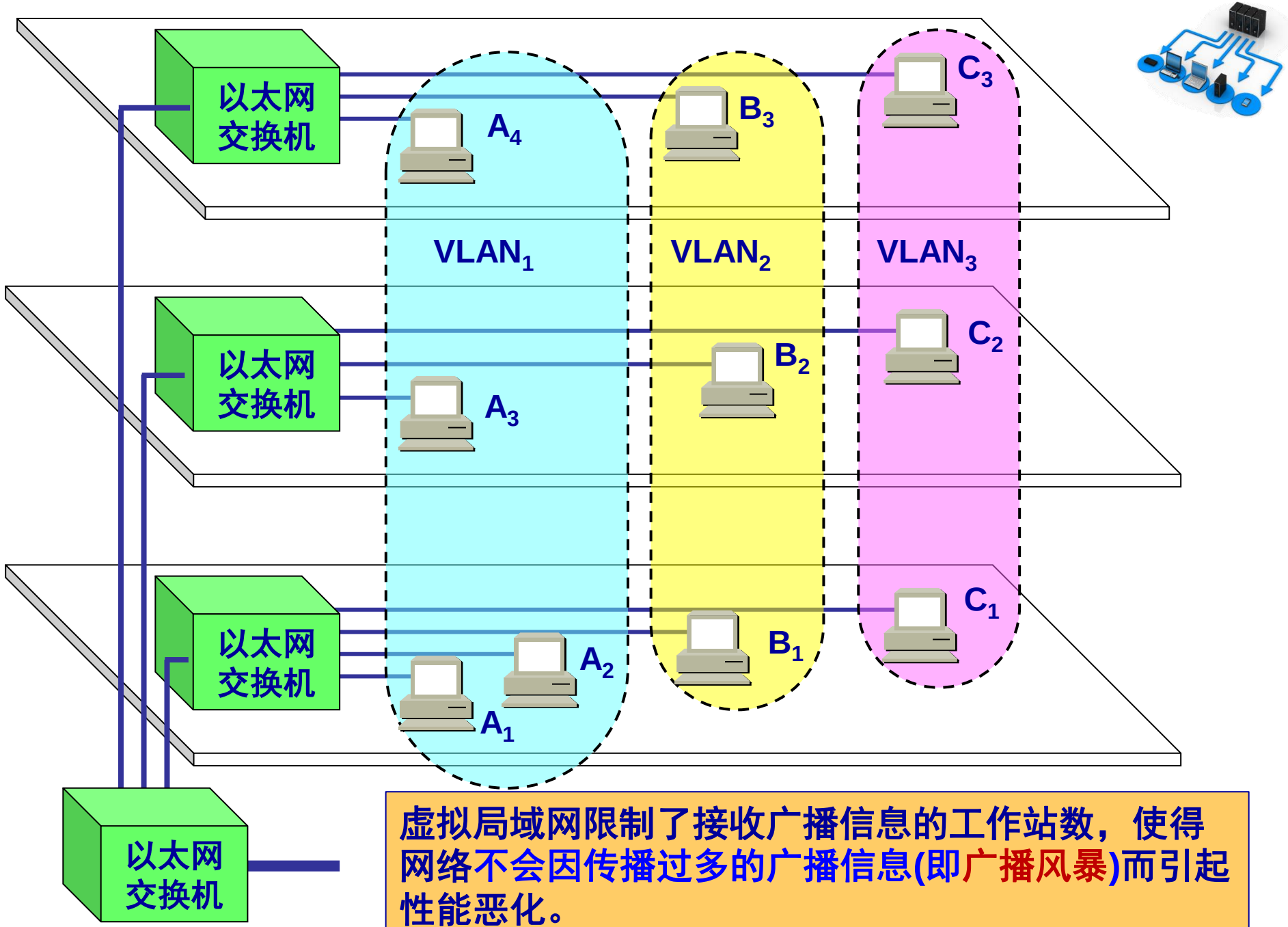
10 台计算机划分为三个虚拟局域网：
$VLAN_1$, $VLAN_2$ 和 $VLAN_3$



当 B_1 向 $VLAN_2$ 工作组内成员发送数据时，工作站 B_2 和 B_3 将会收到广播的信息。



B_1 发送数据时，工作站 A_1，A_2 和 C_1 都不会收到 B_1 发出的广播信息。

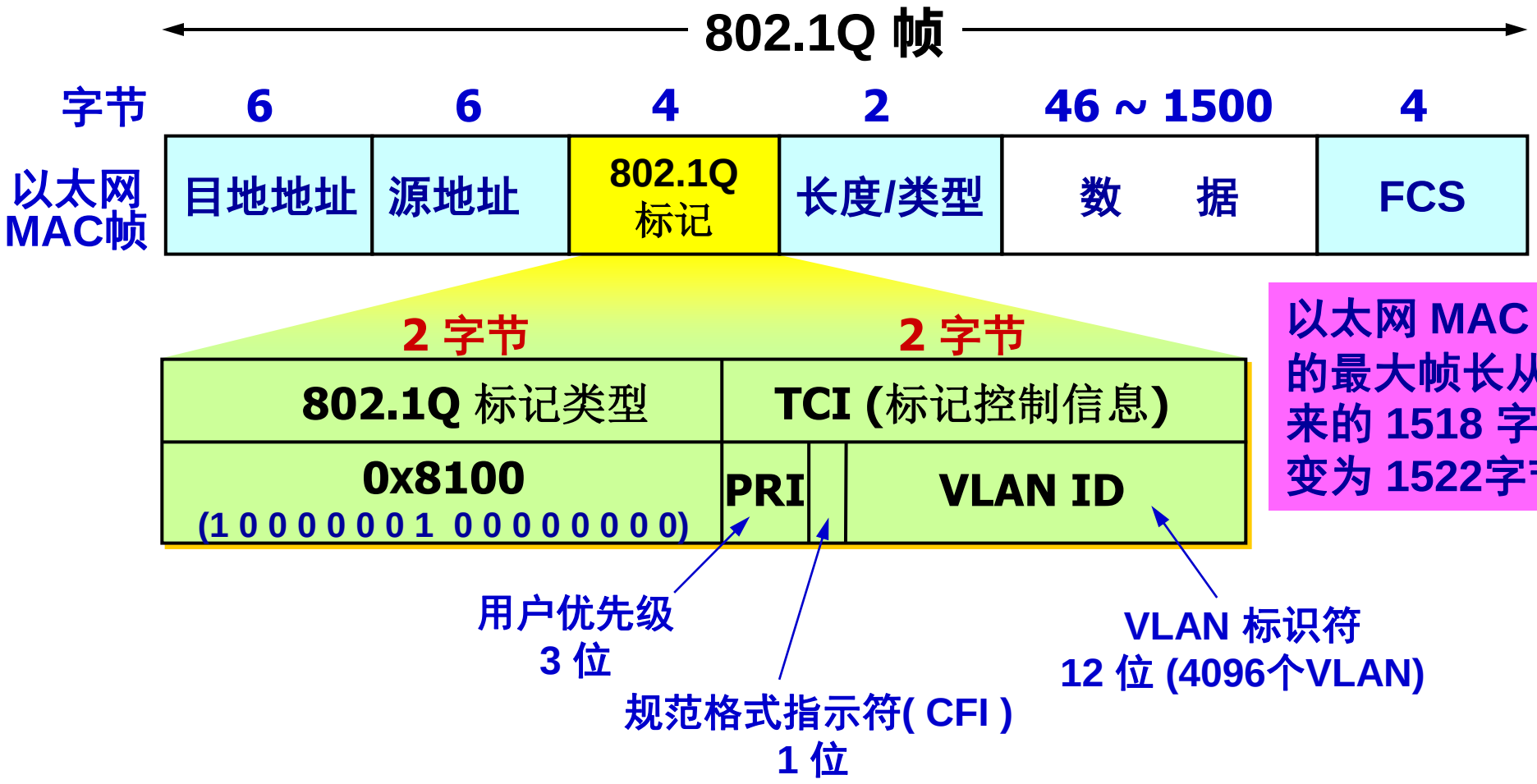


虚拟局域网限制了接收广播信息的工作站数，使得网络不会因传播过多的广播信息(即广播风暴)而引起性能恶化。

虚拟局域网使用的以太网帧格式

- IEEE 批准了 802.3ac 标准，该标准定义了以太网的帧格式的扩展，以支持虚拟局域网。
- 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计算机属于哪一个虚拟局域网。
- 插入 VLAN 标记得出的帧称为 802.1Q 帧 或 带标记的以太网帧。

虚拟局域网使用的以太网帧格式



以太网 MAC 帧的最大帧长从原来的 1518 字节变为 1522字节。

插入 VLAN 标记后变成了 802.1Q 帧

虚拟局域网

- **冲突域 / 碰撞域 与 广播域**
  - 连接在同一个网桥或交换机端口的计算机构成一个冲突域，即处于同一个端口的计算机在某一个时刻只能有一台计算机发送数据，其他处于监听状态，如果出现两台或两台以上的计算机同时发送数据，便会冲突。
  - 网桥或交换机的本质是通过将网络分割成多个冲突域来增强网络服务。
  - 因为网桥会向所有端口转发未知目的端口的数据帧，所以网桥/交换网络会产生广播风暴。

虚拟局域网的优点

- 安全性好
  - 没有路由的情况下，不同虚拟局域网间不能相互通信
- 网络分段
  - 可将物理网络按逻辑分段，而不是按物理分段。
  - 可将不同地点、不同部门的计算机划分在一个虚拟局域网上
- 提供较好的灵活性
  - 方便地将一个站点加入或从一个VLAN中删除。

3.4.6 高速以太网

- 100BASE-T 以太网
- 吉比特以太网
- 10吉比特以太网 (10GE) 和更快的以太网
- 使用以太网进行宽带接入

1. 100BASE-T 以太网

- 速率达到或超过 100 Mbit/s 的以太网称为高速以太网。
- 100BASE-T 在双绞线上传送 100 Mbit/s 基带信号的星形拓扑以太网，仍使用 IEEE 802.3 的 CSMA/CD 协议，工作方式为半双工。
- 100BASE-T 以太网又称为快速以太网 (Fast Ethernet)。
- 用户使用100Mbit的适配器或100Mbit/s的集线器或交换机，可以升级到100Mbit/s
- 1995 年IEEE已把 100BASE-T 的快速以太网定为正式标准，其代号为 IEEE 802.3u。

100BASE-T 以太网的特点

- 可在全双工方式下工作而无冲突发生。在全双工方式下工作时，不使用 CSMA/CD 协议。
- MAC 帧格式仍然是 802.3 标准规定的。
- 保持最短帧长不变，仍为64字节
- 一个网段的最大电缆长度为 100 m
- 争用期使5.12 μs
- 帧间时间间隔从原来的 9.6 μs 改为现在的 0.96 μs。

100 Mbit/s 以太网的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。
  - 网段最大程度：100米。
- 100BASE-T4
  - 使用 4 对 UTP 3 类线 或 5 类线。
  - 网段最大程度：100米。
- 100BASE-FX
  - 使用 2 根光纤。
  - 网段最大程度：2000米。

2. 吉比特以太网

- 允许在 1 Gbit/s 下以全双工和半双工两种方式工作。
- 使用 IEEE 802.3 协议规定的帧格式。
- 在半双工方式下使用 CSMA/CD 协议，全双工方式不使用 CSMA/CD 协议。
- 与 10BASE-T 和 100BASE-T 技术向后兼容。

吉比特以太网可用作现有网络的主干网，也可在高带宽（高速率）的应用场合中。

吉比特以太网的物理层

- **使用两种成熟的技术：**一种来自现有的以太网，另一种则是美国国家标准协会 ANSI 制定的**光纤通道 FC (Fiber Channel)**。

吉比特以太网物理层标准

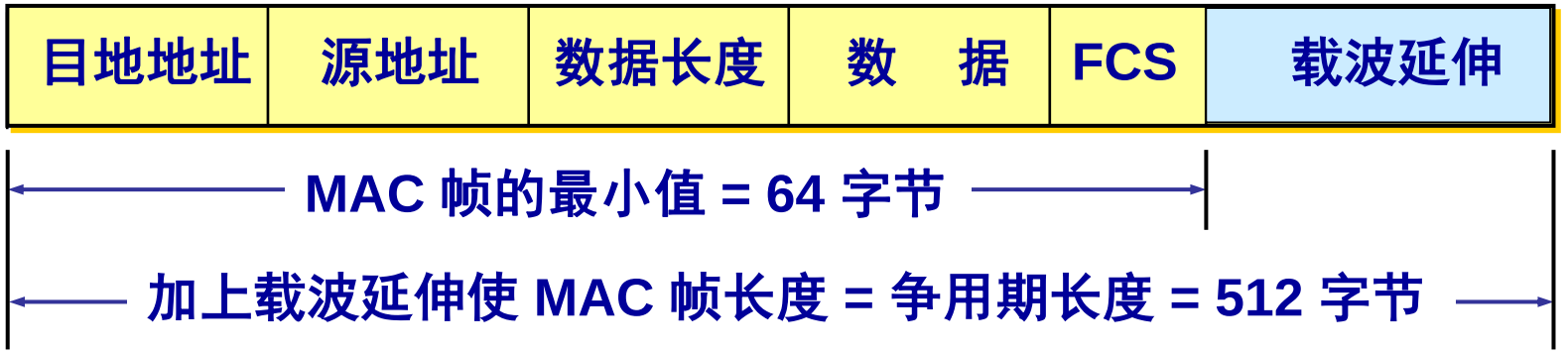
名称	媒体	网段最大长度	特点
1000BASE-SX	光缆	550 m	多模光纤（50 和 62.5 μm）
1000BASE-LX	光缆	5000 m	单模光纤（10 μm）多模光纤（50 和 62.5 μm）
1000BASE-CX	铜缆	25 m	使用 2 对屏蔽双绞线电缆 STP
1000BASE-T	铜缆	100 m	使用 4 对 UTP 5 类线

半双工方式工作的吉比特以太网

- 吉比特以太网工作在半双工方式时，就必须进行碰撞检测。
- 由于数据率提高了，因此只有减小最大电缆长度或增大帧的最小长度。
- 为保持 64 字节最小帧长度，以及 100 米的网段的最大长度，吉比特以太网增加了两个功能：
  - 载波延伸 (carrier extension)
  - 分组突发 (packet bursting)
- 当吉比特以太网工作在全双工方式时，不使用载波延伸和分组突发。

载波延伸

- 使最短帧长仍为 64 字节（这样可以保持兼容性），同时将争用时间增大为 512 字节。
- 凡发送的 MAC 帧长不足 512 字节时，就用一些特殊字符填充在帧的后面，使MAC 帧的发送长度增大到 512 字节。接收端在收到以太网的 MAC 帧后，要将所填充的特殊字符删除后才向高层交付。



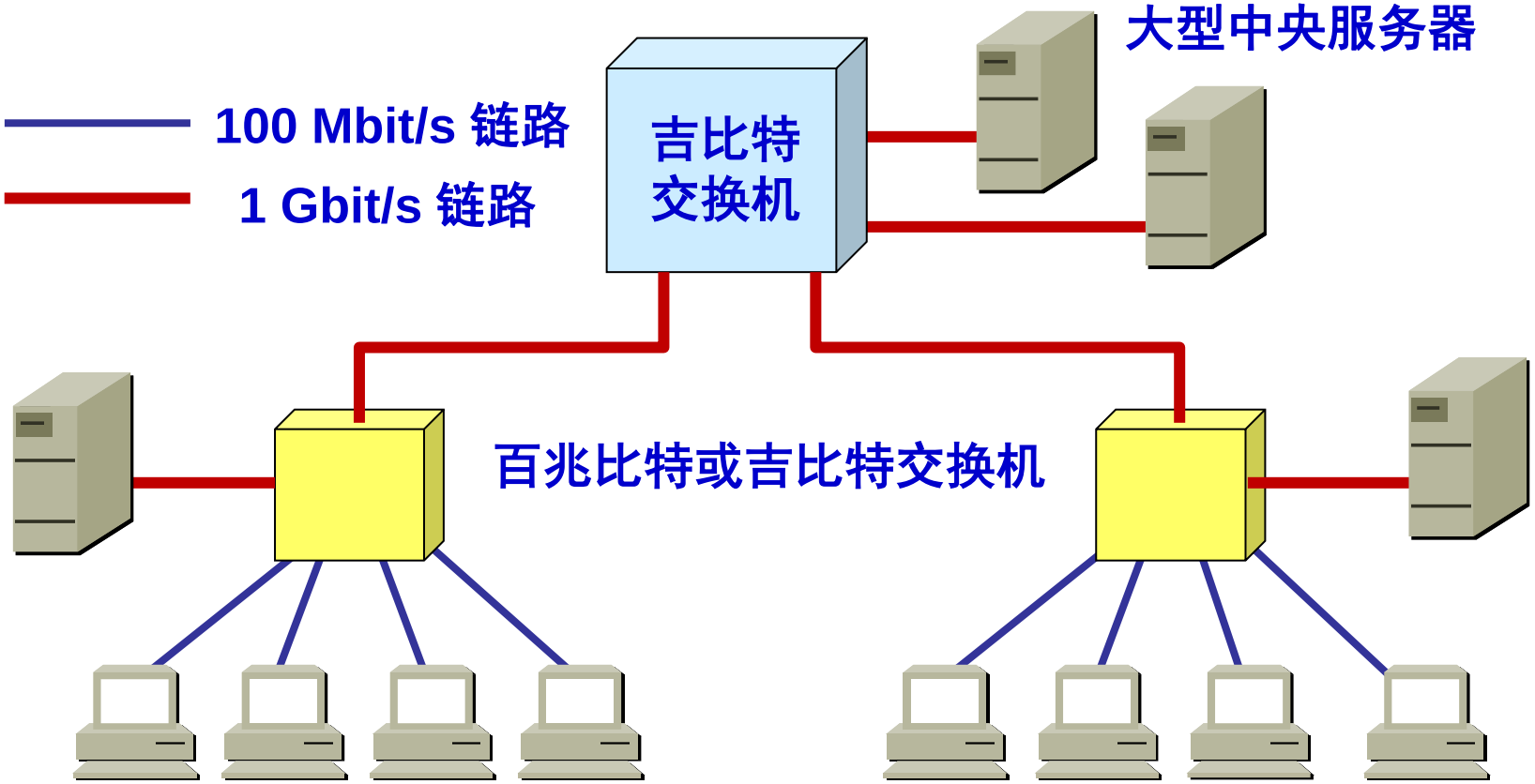
载波延伸

分组突发

- 当很多短帧要发送时，第一个短帧要采用载波延伸方法进行填充，随后的一些短帧则可一个接一个地发送，只需留有必要的帧间最小间隔即可。这样就形成可一串分组的突发，直到达到1500 字节或稍多一些为止。

分组突发

吉比特以太网的配置举例



吉比特以太网交换机可以与多个图形工作站相连，也可用作百兆以太网的主干网，与百兆比特或吉比特交换机相连

3. 10吉比特以太网和更快的以太网

- 10 吉比特以太网（10GE）并非把吉比特以太网的速率简单地提高到 10 倍，其主要特点有：
  - 与 10 Mbit/s、100 Mbit/s 和 1 Gbit/s 以太网的帧格式完全相同。
  - 保留了 802.3 标准规定的以太网最小和最大帧长，便于升级。
  - 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。

10 吉比特以太网的物理层

10GE 的物理层标准

名称	媒体	网段最大长度	特点
10GBASE-SR	光缆	300 m	多模光纤（0.85 μm）
10GBASE-LR	光缆	10 km	单模光纤（1.3 μm）
10GBASE-ER	光缆	40 km	单模光纤（1.5 μm）
10GBASE-CX4	铜缆	15 m	使用 4 对双芯同轴电缆 (twinax)
10GBASE-T	铜缆	100 m	使用 4 对 6A 类 UTP 双绞线

更快的以太网

- 以太网的技术发展得很快。
- 在 10GE 之后又制订了 40GE/100GE（即 40 吉比特以太网和 100 吉比特以太网）的标准 IEEE 802.3ba-2010 和 802.3bm-2015。
- 40GE/100GE 只工作在全双工的传输方式（因而不使用 CSMA/CD 协议），并仍保持了以太网的帧格式以及 802.3 标准规定的以太网最小和最大帧长。
- 100GE 在使用单模光纤传输时，仍然可以达到 40 km 的传输距离，但这是需要波分复用（使用 4 个波长复用一根光纤，每一个波长的有效传输速率是 25 Gbit/s）。

40GE/100GE 的物理层

40GE/10GE 的物理层标准

物理层	40GE	100GE
在背板上传输至少超过 1 m	40GBASE-KR4	
在铜缆上传输至少超过 7 m	40GBASE-CR4	100GBASE-CR10
在多模光纤上传输至少 100 m	40GBASE-SR4	100GBASE-SR10， *100GBASE-SR4
在单模光纤上传输至少 10 km	40GBASE-LR4	100GBASE-LR4
在单模光纤上传输至少 40 km	*40GBASE-ER	100GBASE-ER4

端到端的以太网传输

- 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有：
  - **技术成熟；**
  - **互操作性很好，**不同厂商生产的以太网都能可靠地进行互操作；
  - **在广域网中使用以太网时价格便宜；**
  - **采用统一的以太网帧格式，简化了操作和管理，**不需要在进行帧的格式转换。

以太网从 10 Mbit/s 到100 Gbit/s 的演进

- 以太网从 10 Mbit/s 到 100 Gbit/s 的演进证明了以太网是：
  - 可扩展的（从 10 Mbit/s 到 100 Gbit/s）；
  - 灵活的（多种传输媒体、全/半双工、共享/交换）；
  - 易于安装；
  - 稳健性好。

4. 使用以太网进行宽带接入

- IEEE 在 2001 年初成立了 802.3 EFM 工作组，专门研究高速以太网的宽带接入技术问题。
- 以太网宽带接入具有以下特点：
  - 可以提供双向的宽带通信。
  - 可以根据用户对带宽的需求灵活地进行带宽升级。
  - 可以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且降低了传输的成本。
  - 但是不支持用户身份鉴别。

PPPoE

- PPPoE (PPP over Ethernet) 的意思是“在以太网上运行 PPP”，它把 PPP 协议与以太网协议结合起来 —— 将 PPP 帧再封装到以太网中来传输。
- 现在的光纤宽带接入 FTTx 都要使用 PPPoE 的方式进行接入，不需要使用调制解调器，只有一个RJ-45插口。
  - 在 PPPoE 弹出的窗口中键入在网络运营商购买的用户名和密码，就可以进行宽带上网了。
- 利用 ADSL 进行宽带上网时，从用户个人电脑到家中的 ADSL 调制解调器之间，也是使用 RJ-45 和 5 类线（即以太网使用的网线）进行连接的，并且也是使用 PPPoE 弹出的窗口进行拨号连接的。ADSL调制解调器将以太网帧转换成PPP帧。

总结

- 数据链路层的基本概念
- 点对点的数据链路层：PPP协议
  - 封装成帧
  - 透明传输
  - 差错检测
- 广播信道的数据链路层：以太网
  - CSMA/CD 协议
  - 以太网 MAC层的硬件地址和帧格式
  - 扩展以太网、虚拟以太网、高速以太网