



复习大纲

- 本PPT只是重要内容的提示，具体请结合书本和作业
- 有问题发邮件王娟老师邮箱：
jjmao2009@163.com



1.2.2 因特网发展的三个阶段

- 第一阶段是从单个网络 **ARPANET** 向互联网发展的过程。
- 第二阶段：Internet 网



多层次 ISP 结构的因特网

- 第三阶段的特点是逐渐形成了多层次 ISP 结构的因特网。
 - 出现了因特网服务提供者 **ISP (Internet Service Provider)**。
- 传统的三网：电信网，有线电视网，计算机网络



1.3 因特网的组成

从因特网的工作方式上看，可以划分为以下的两大块（也称两类子网）：

- (1) **边缘部分** 由所有连接在因特网上的主机组成。这部分是用户直接使用的，用来进行通信（传送数据、音频或视频）和资源共享。
- (2) **核心部分** 由大量网络和连接这些网络的路由器组成。这部分是为边缘部分提供服务的（提供连通性和交换）。



两种通信方式

在网络边缘的端系统中运行的程序之间的通信方式通常可划分为两大类：

- 客户-服务器方式（C/S 方式）

即Client/Server方式

- 对等方式（P2P 方式）

即 Peer-to-Peer方式，不区分哪一个是服务请求方还是服务提供方

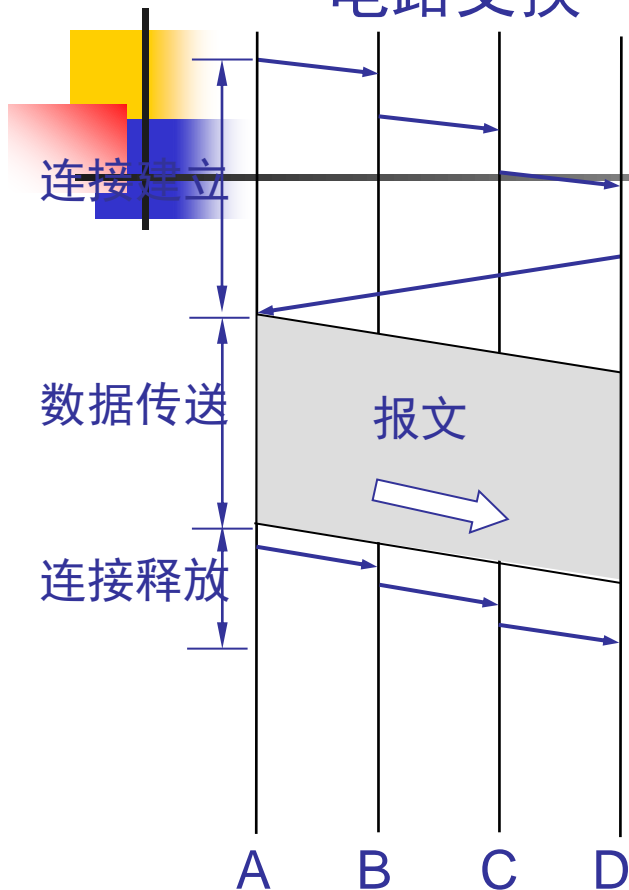


路由器的重要任务

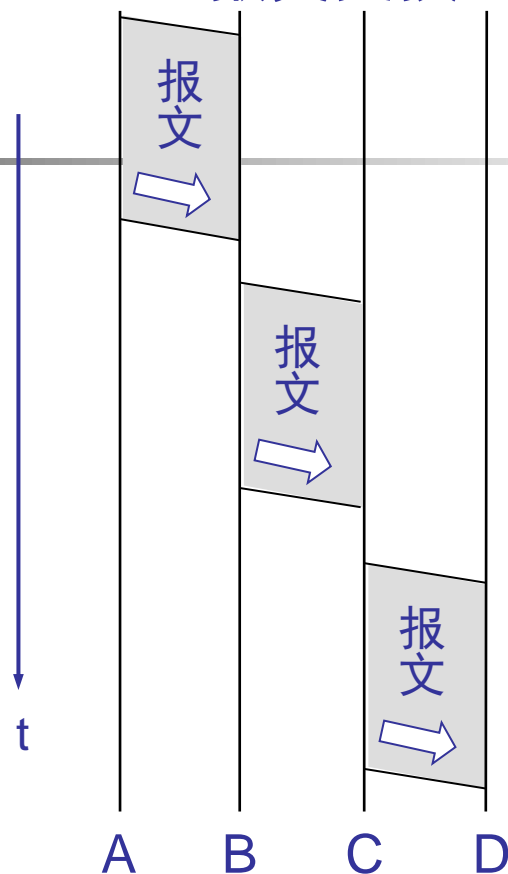
- 在网络核心部分起特殊作用的是**路由器**(router)。
- 路由器是实现**分组交换**(packet switching)的关键构件，其任务是转发收到的分组，这是网络核心部分最重要的功能。

三种交换的比较

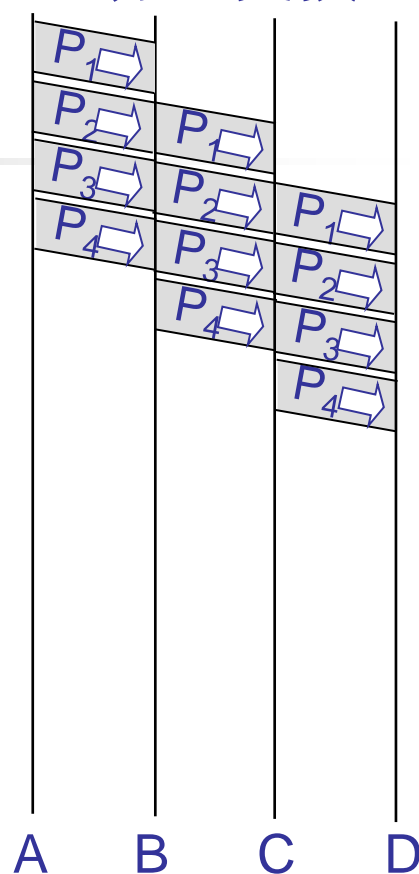
电路交换



报文交换

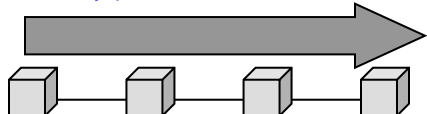


分组交换

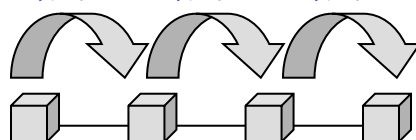


数据传送
的特点

比特流直达终点

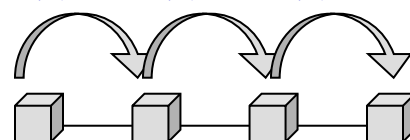


报文 报文 报文



存储
转发 存储
转发

分组 分组 分组



存储
转发 存储
转发



1.5.2 几种不同类别的网络

1. 从网络的作用范围进行分类

- 广域网 **WAN** (Wide Area Network)
- 城域网 **MAN** (Metropolitan Area Network)
- 局域网 **LAN** (Local Area Network)
- 个人区域网 **PAN** (Personal Area Network)



2. 不同使用者的网络

- 从网络的使用者进行分类
 - 公用网 (public network)
 - 专用网 (private network)
 - 虚拟专用网络 (Virtual Private Network: VPN)



1.6 计算机网络的性能

1.6.1 计算机网络的性能指标

1. 速率

- **比特** (bit) 是计算机中数据量的单位，也是信息论中使用的信息量的单位。
- Bit 来源于 binary digit，意思是一个“**二进制数字**”，因此一个比特就是二进制数字中的一个 1 或 0。
- **速率**即**数据率**(data rate)或**比特率**(bit rate)是计算机网络中最重要的一个性能指标。速率的单位是 b/s，或kb/s, Mb/s, Gb/s 等
- 速率往往是指**额定速率**或**标称速率**。



2. 带宽

- “**带宽**” (bandwidth)本来（通讯领域）是指信号具有的频带宽度，单位是赫（或千赫、兆赫、吉赫等）。
- 现在“带宽”是数字信道所能传送的“**最高数据率**”的同义语，单位是“比特每秒”，或 b/s (bit/s)。



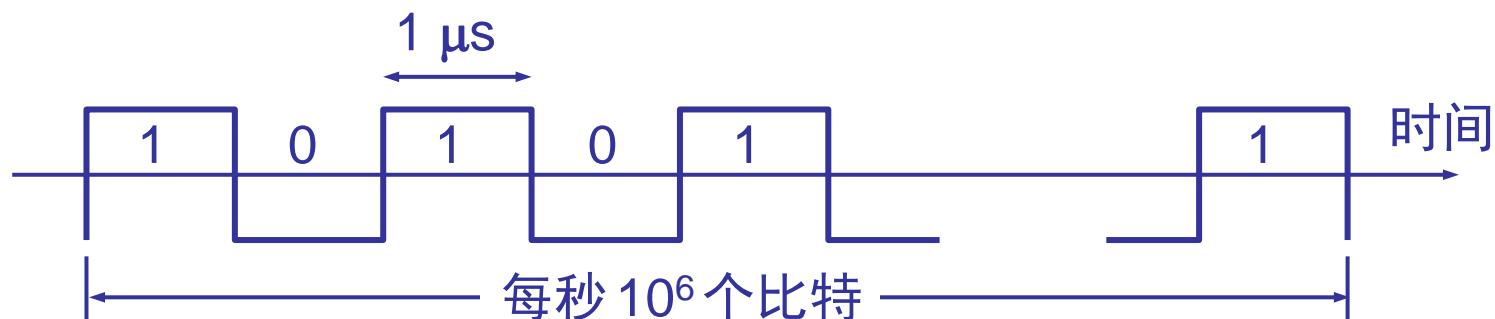
常用的带宽单位

- 更常用的带宽单位是
 - 千比每秒，即 kb/s (10^3 b/s)
 - 兆比每秒，即 Mb/s (10^6 b/s)
 - 吉比每秒，即 Gb/s (10^9 b/s)
 - 太比每秒，即 Tb/s (10^{12} b/s)
- 请注意：在计算机界， $K = 2^{10} = 1024$
 $M = 2^{20}$, $G = 2^{30}$, $T = 2^{40}$ 。

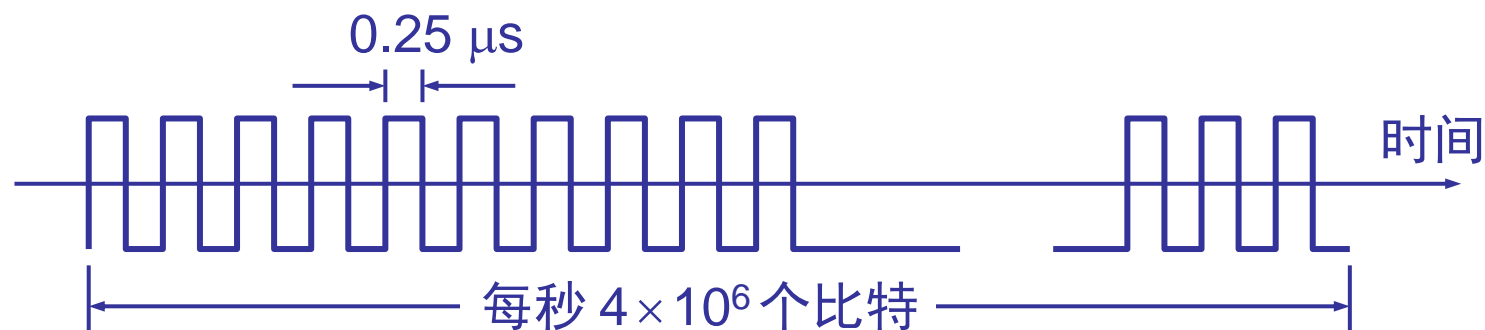
数字信号流随时间的变化

- 在时间轴上信号的宽度随带宽的增大而变窄。

带宽为
1 Mb/s



带宽为
4 Mb/s





4. 时延(delay 或 latency)

- **发送时延** 发送数据时，数据帧从结点进入到传输媒体所需要的时间。
- 也就是从发送数据帧的第一个比特算起，到该帧的最后一个比特发送完毕所需的时间。(结合书后题,会计算)

$$\text{发送时延} = \frac{\text{数据帧长度 (b)}}{\text{发送速率 (b/s)}}$$



时延(delay 或 latency)

- **传播时延** 电磁波在信道中需要传播一定的距离而花费的时间（介质的物理性质决定P）。（结合书后题,会计算）
- 信号发送速率和信号在信道上的**传播速率**是完全不同的概念。

$$\text{传播时延} = \frac{\text{信道长度（米）}}{\text{信号在信道上的传播速率（米/秒）}}$$



时延(delay 或 latency)

- **处理时延** 交换结点为存储转发而进行一些必要的处理所花费的时间。
- **排队时延** 结点缓存队列中分组**排队**所经历的时延。
- 排队时延的长短往往取决于网络中**当时**的**通信量**。



时延(delay 或 latency)

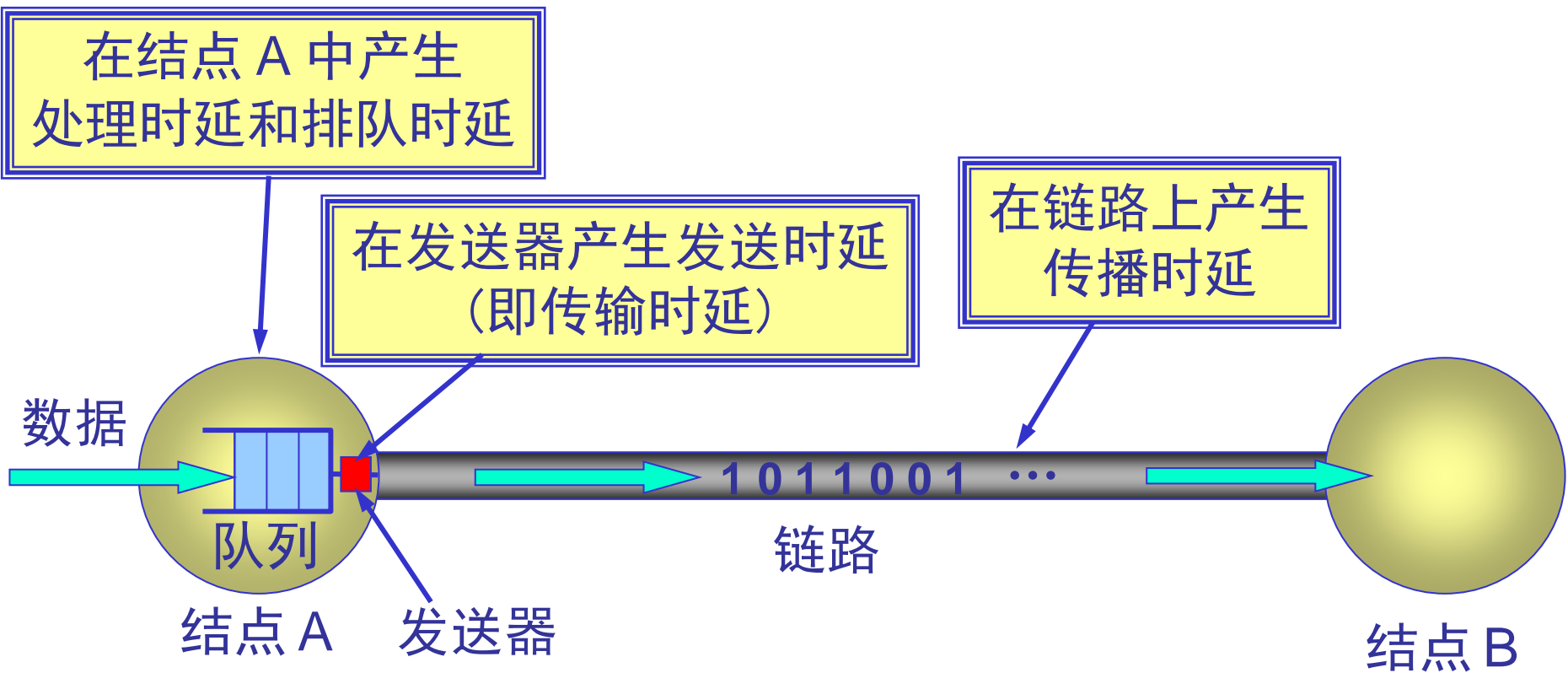
- 数据经历的总时延就是发送时延、传播时延、处理时延和排队时延之和：

总时延 = 发送时延 + 传播时延 + 处理时延 + 排队时延

会做书后,发送时延,传播时延的题:1-17,1-18

四种时延所产生的地方

从结点 A 向结点 B 发送数据





6.往返时间RTT

- 从发送方发送开始，到发送方收到接收方的确认消息的时间。
- ping

7. 利用率

- **信道利用率**指出某信道有百分之几的时间是被利用的（有数据通过）。完全空闲的信道的利用率是零。
- **网络利用率**则是全网络的信道利用率的加权平均值。
- 信道利用率并非越高越好。
- 自己总结下所有的对网络的衡量指标



1.7.2 协议与划分层次

- 网络协议(network protocol), 简称为协议, 是为进行网络中的数据交换而建立的规则、标准或约定。
- 协议是控制两个对等实体间通信规则的集合
- 要实现本层协议, 必须要下一层提供的服务



1.7.2 协议与划分层次

- 协议组成要素:
 - **语法**:数据与控制信息的结构或格式
 - **语义**:需要发出何种控制信息, 完成何种动作以及做出何种响应
 - **同步**:事件实现顺序的详细说明

五层协议的体系结构



- 应用层(application layer)-主机
- 运输层(transport layer)
- 网络层(network layer) –路由器
- 数据链路层(data link layer)
- 物理层(physical layer)（不属于TCPIP）
- 下层为上层提供服务
- 每层的主要功能和协议,自己总结下



第二章 物理层



2.1 物理层的基本概念

- 物理层考虑的是怎样才能在连接各种计算机的传输媒体上**传输数据比特流**



(2) 信噪比

- **信噪比**就是信号的平均功率和噪声的平均功率之比。常记为 S/N ，并用分贝 (dB) 作为度量单位。即：

$$\text{信噪比(dB)} = 10 \log_{10}(S/N) \quad (\text{dB})$$

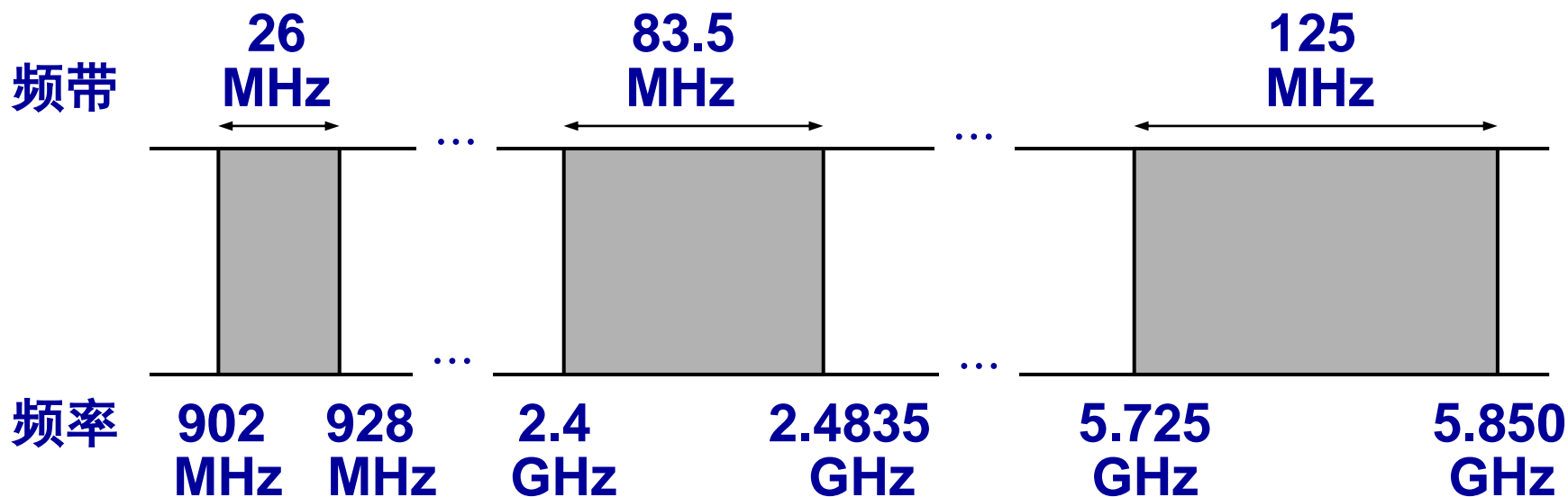
- 信道的极限信息传输速率 C 可表达为：

$$C = W \log_2(1+S/N) \quad (\text{bit/s})$$

其中：
 W 为信道的带宽（以 Hz 为单位）；
 S 为信道内所传信号的平均功率；
 N 为信道内部的高斯噪声功率。

无线局域网使用的 ISM 频段

要使用某一段无线电频谱进行通信，通常必须得到本国政府有关无线电频谱管理机构的许可证。但是，也有一些无线电频段是可以自由使用的。例如：ISM。各国的 ISM 标准有可能略有差别。



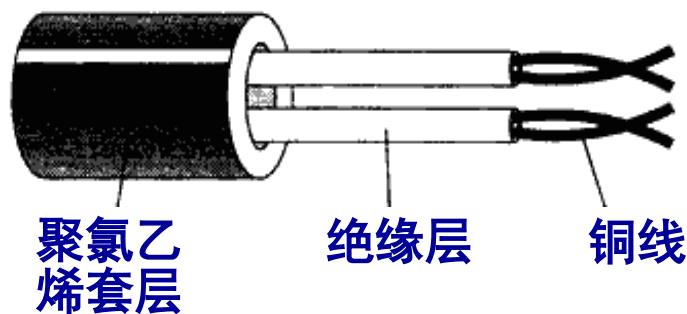
无线局域网使用的 ISM 频段



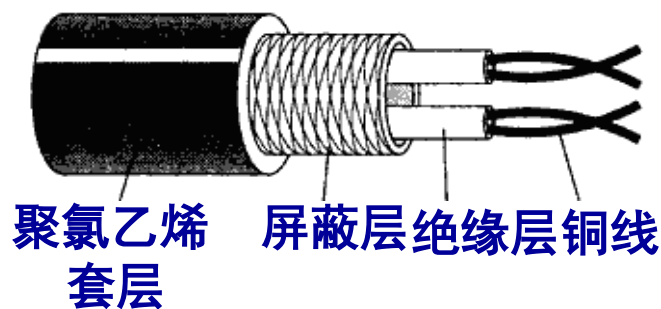
2.3 物理层下面的传输媒体

- 传输媒体也称为传输介质或传输媒介，它就是数据传输系统中在发送器和接收器之间的物理通路。
- 传输媒体可分为两大类，即导引型传输媒体和非导引型传输媒体。
- 在导引型传输媒体中，电磁波被导引沿着固体媒体（铜线或光纤）传播。有线
- 非导引型传输媒体就是指自由空间。在非导引型传输媒体中，电磁波的传输常称为无线传输。

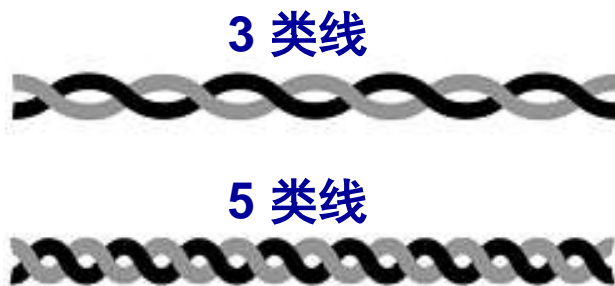
2.3.1 导引型传输媒体-会认图



(a) 无屏蔽双绞线



(b) 屏蔽双绞线



(c) 不同的绞合度的双绞线

双绞线的示意图



2.3.1 导引型传输媒体

■ 光缆

- 光纤体积小、重量轻、抗雷电和电磁干扰性能好
- 光纤通信系统的传输带宽远远大于目前其他各种传输媒体的带宽。
- 单模距离比多模光纤远,但是速度比多模光纤慢.



几个术语

- 数据(data)——运送消息的实体。
- 信号(signal)——数据的电气的或电磁的表现。
- “模拟的” (analogous)——代表消息的参数的取值是连续的。
- “数字的” (digital)——代表消息的参数的取值是离散的。
- 码元(code)——在使用时间域（或简称为时域）的波形表示数字信号时，代表不同离散数值的基本波形。
- 信噪比：信号的平均功率和噪声的平均功率之比



2.2.2 信道的通信方式

- **单向通信**（单工通信）——只能有一个方向的通信而没有反方向的交互。
- **双向交替通信**（半双工通信）——通信的双方都可以发送信息，但不能双方同时发送(当然也就不能同时接收)。
- **双向同时通信**（全双工通信）——通信的双方可以同时发送和接收信息。



2.2.2 有关信道的几个基本概念

调制分为两大类：

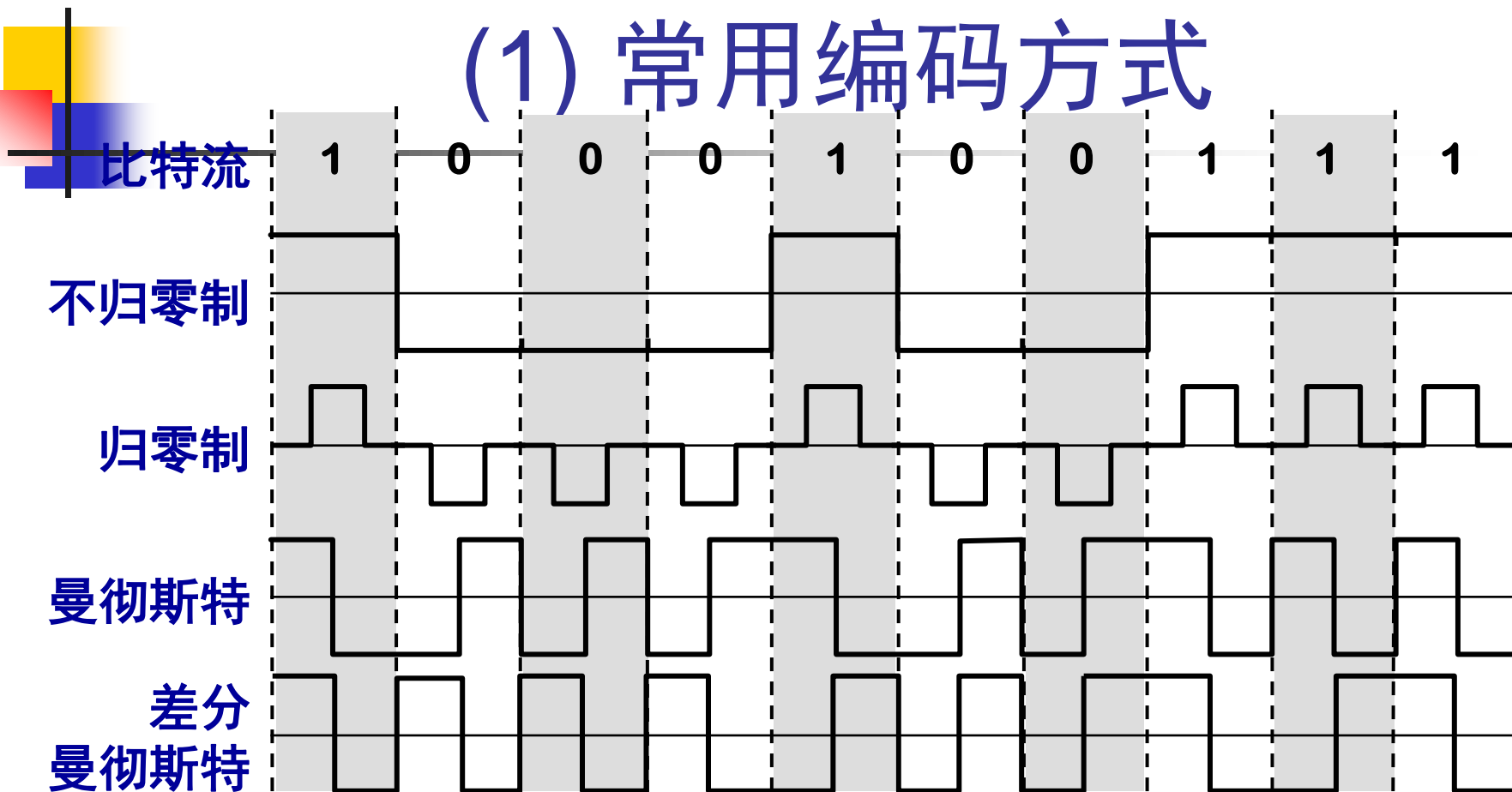
- **基带调制**：仅对基带信号的波形进行变换，使它能够与信道特性相适应。变换后的信号仍然是基带信号。把这种过程称为**编码 (coding)**。
- **带通调制**：使用**载波 (carrier)**进行调制，把基带信号的频率范围搬移到**较高的频段**，并**转换为模拟信号**，这样就能够更好地在模拟信道中传输（即仅在一段频率范围内能够通过信道）。
- **带通信号**：经过载波调制后的信号。



几种最基本的调制方法

- 最基本的二元制调制方法有以下几种：
 - 调幅(AM)：载波的振幅随基带数字信号而变化。
 - 调频(FM)：载波的频率随基带数字信号而变化。
 - 调相(PM)：载波的初始相位随基带数字信号而变化。

(1) 常用编码方式



数字信号常用的编码方式



(1) 常用编码方式

- 从信号波形中可以看出，曼彻斯特 (Manchester) 编码和差分曼彻斯特编码产生的信号频率比不归零制高。
- 从自同步能力来看，不归零制不能从信号波形本身中提取信号时钟频率（这叫作没有自同步能力），而曼彻斯特编码和差分曼彻斯特编码具有自同步能力。



频分复用 FDM

(Frequency Division Multiplexing)

- 用户在分配到一定的频带后，在通信过程中自始至终都占用这个频带。
- **频分复用**的所有用户在同样的时间占用不同的带宽资源（请注意，这里的“带宽”是频率带宽而不是数据的发送速率）。
- **波分复用是一种特殊的频分复用**



时分复用TDM (Time Division Multiplexing)

- 时分复用则是将时间划分为一段段等长的时分复用帧（TDM 帧）。每一个时分复用的用户在每一个 TDM 帧中占用固定序号的时隙。



2.4.3 码分复用 CDM

(Code Division Multiplexing)

- 常用的名词是码分多址 CDMA (Code Division Multiple Access)。
- 每一个比特时间划分为 m 个短的间隔，称为码片(chip)。



码片序列(chip sequence)

- 每个站被指派一个唯一的 m bit 码片序列。
 - 如发送比特 1, 则发送自己的 m bit 码片序列。
 - 如发送比特 0, 则发送该码片序列的二进制反码。
- 例如, S 站的 8 bit 码片序列是 00011011。
 - 发送比特 1 时, 就发送序列 00011011,
 - 发送比特 0 时, 就发送序列 11100100。
- S 站的码片序列: $(-1 -1 -1 +1 +1 -1 +1 +1)$
- 不要背, 会做书后的计算题!!! 就是 ABCD 站那个发了 1, 0 的那道题。



2.6.1 ADSL 技术

- 非对称数字用户线 ADSL (Asymmetric Digital Subscriber Line)
- ADSL 技术就把 0~4 kHz 低端频谱留给传统电话使用，而把原来没有被利用的高端频谱留给用户上网使用。
- 上行和下行带宽做成不对称的。
 - 上行指从用户到 ISP，而下行指从 ISP 到用户



第三章 数据链路层

3.1 使用点对点信道的数据链路层

3.1.1 数据链路和帧

- **链路(link)**是一条无源的点到点的物理线路段，**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+协议
- MTU,链路层规定的最大传输单元,即数据部分最大值



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。
- **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。



3.1.2 三个基本问题

- 封装成帧：就是在一段数据的前后分别添加首部和尾部，然后就构成了一个帧。确定帧的界限
- 透明传输：解决在数据中出现控制字符的问题
- 差错控制：检测数据帧是否出错，一般采用CRC校验



循环冗余检验的原理

- 在数据链路层传送的帧中，广泛使用了**循环冗余检验 CRC** 的检错技术。
- 课后题3-07！！！！ 会从多项式写余数.只能检测一个bit错误！！ 接收方余数不是0就是有错！ 检测错误不等于可靠！



3.2 点对点协议 PPP

3.2.1 PPP 协议的特点

- 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。



3. 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)。
- (3) 网络控制协议 NCP (Network Control Protocol)

PPP协议的流程:

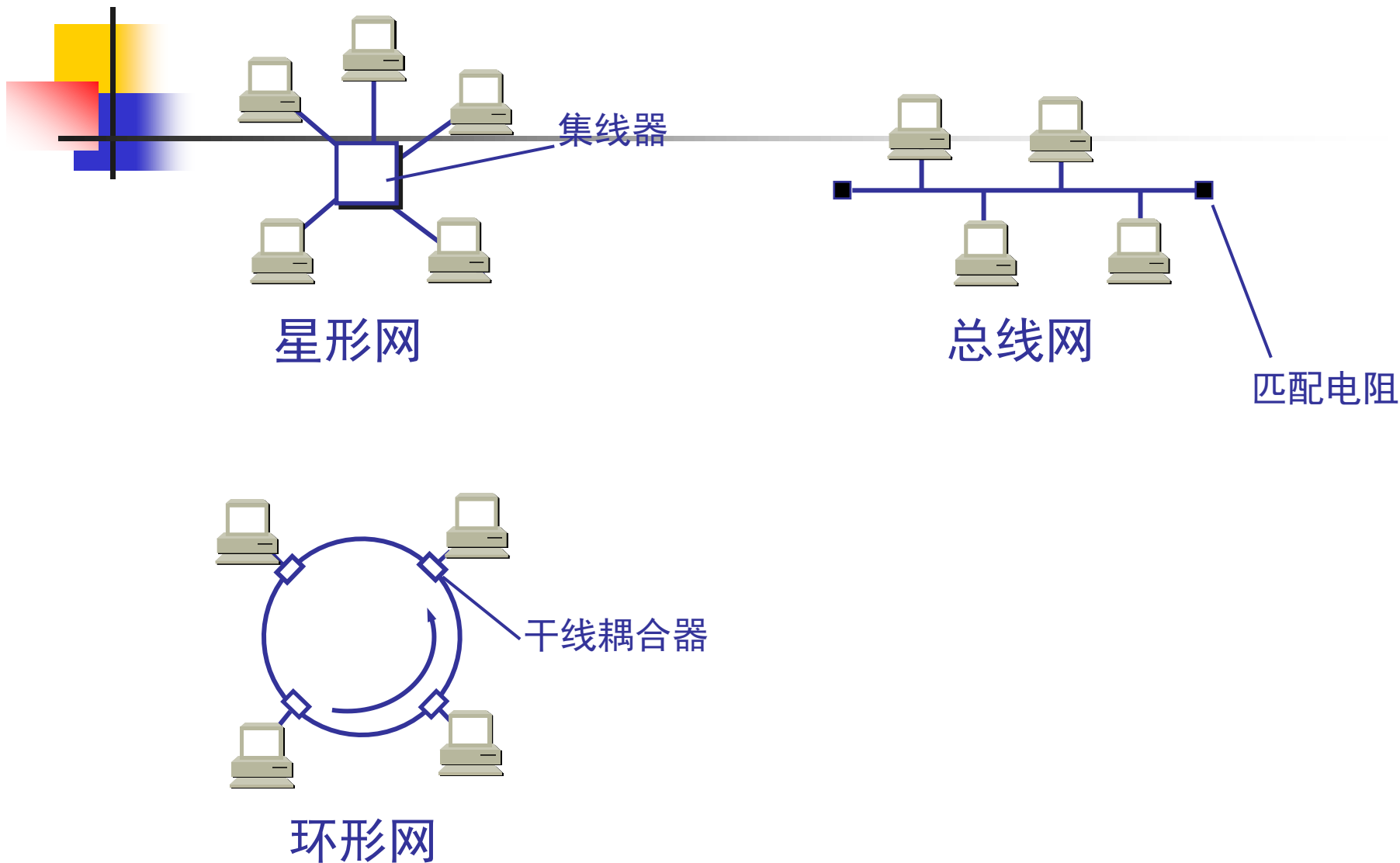
- 链路协商（建立链路），用户名密码鉴别，网络层协商配置



1. 以太网的两个标准

- **DIX Ethernet V2** 是世界上第一个局域网产品（以太网）的规约。
- **IEEE 802.3** 是第一个 IEEE 的以太网标准。
- 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。

局域网的拓扑





2. 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。



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。
- CS: 载波监听
- MA: 多点接入
- CD: 碰撞检测(冲突检测)



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。



最短有效帧长

- 发送最短帧要经过争用期这段时间，此时间段内没有检测到碰撞，才能肯定这次发送不会发生碰撞。小于最短帧的都是无效帧。
- 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。



48 位的 MAC 地址

■ MAC:媒体介入控制

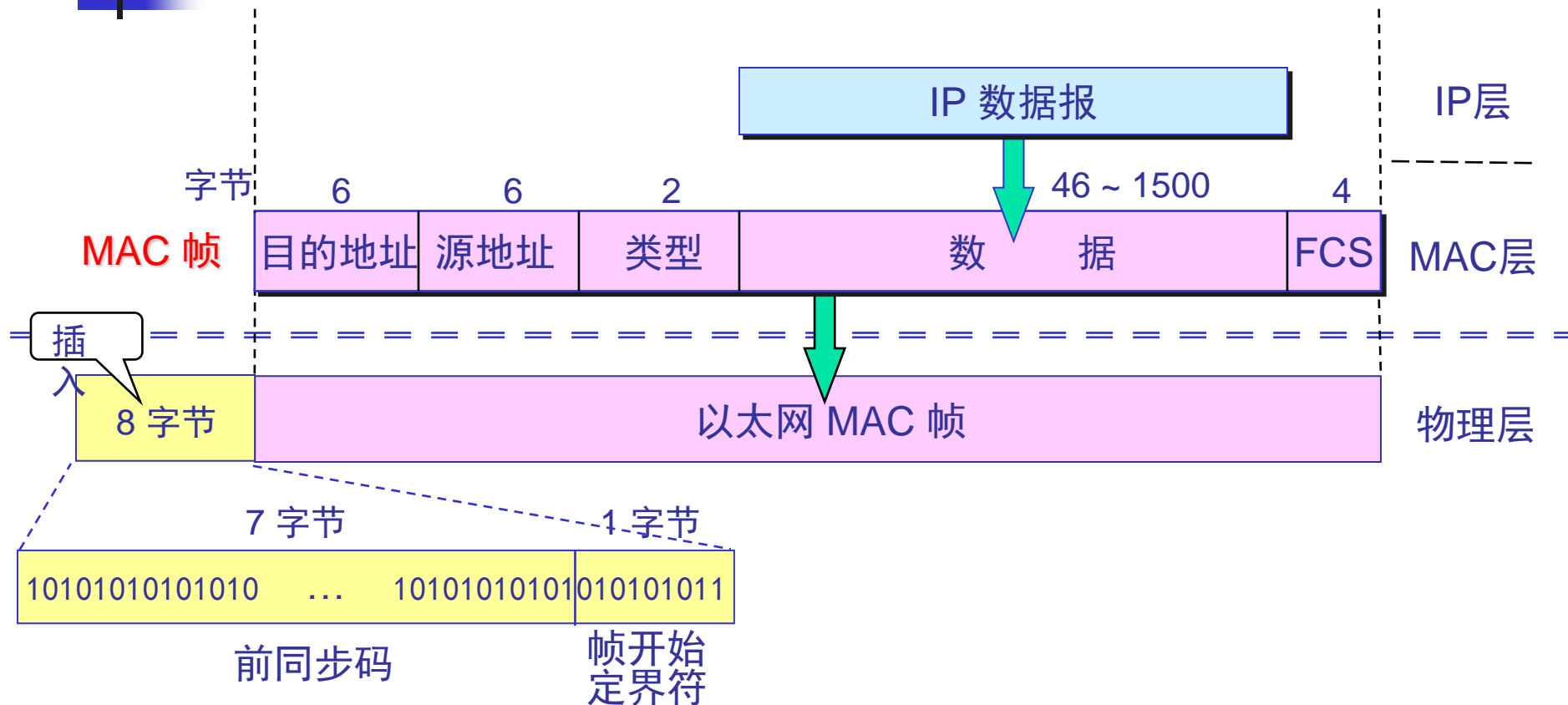
- 在局域网中，硬件地址又称为物理地址，或 MAC 地址。
- IEEE 的注册管理机构 RA 负责向厂家分配地址字段的前三个字节(即高位 24 位)。
- 地址字段中的后三个字节(即低位 24 位)由厂家自行指派，称为扩展标识符，必须保证生产出的适配器没有重复地址。



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）

以太网的 MAC 帧格式



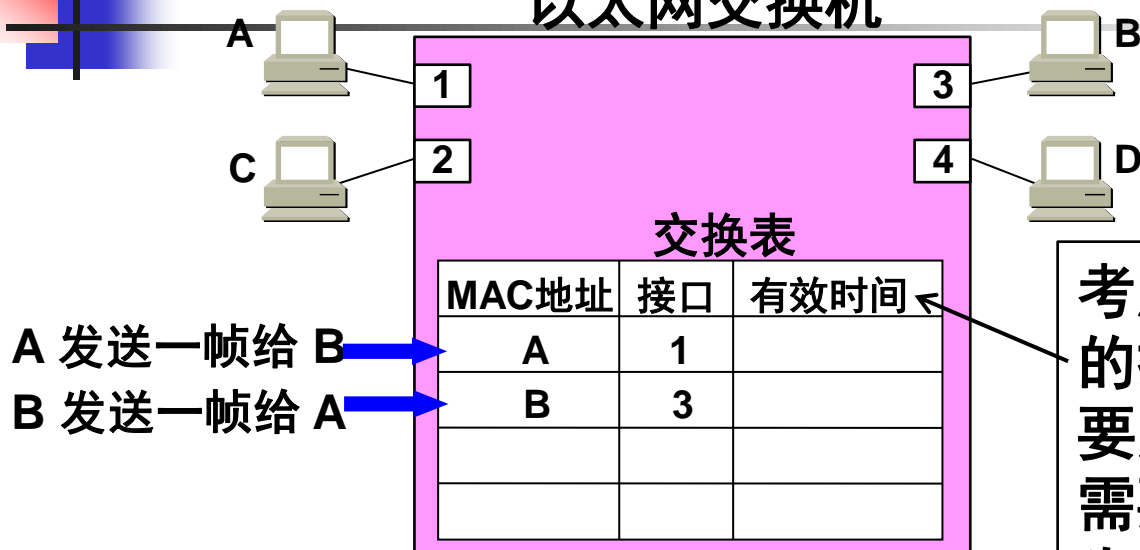


无效的 MAC 帧

- 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。

按照以下自学习算法 处理收到的帧和建立交换表

以太网交换机



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，这就需要更改交换表中的项目。为此，在交换表中每个项目都设有一定的**有效时间**。**过期的项目就自动被删除**。

不要背会做题3-33！！

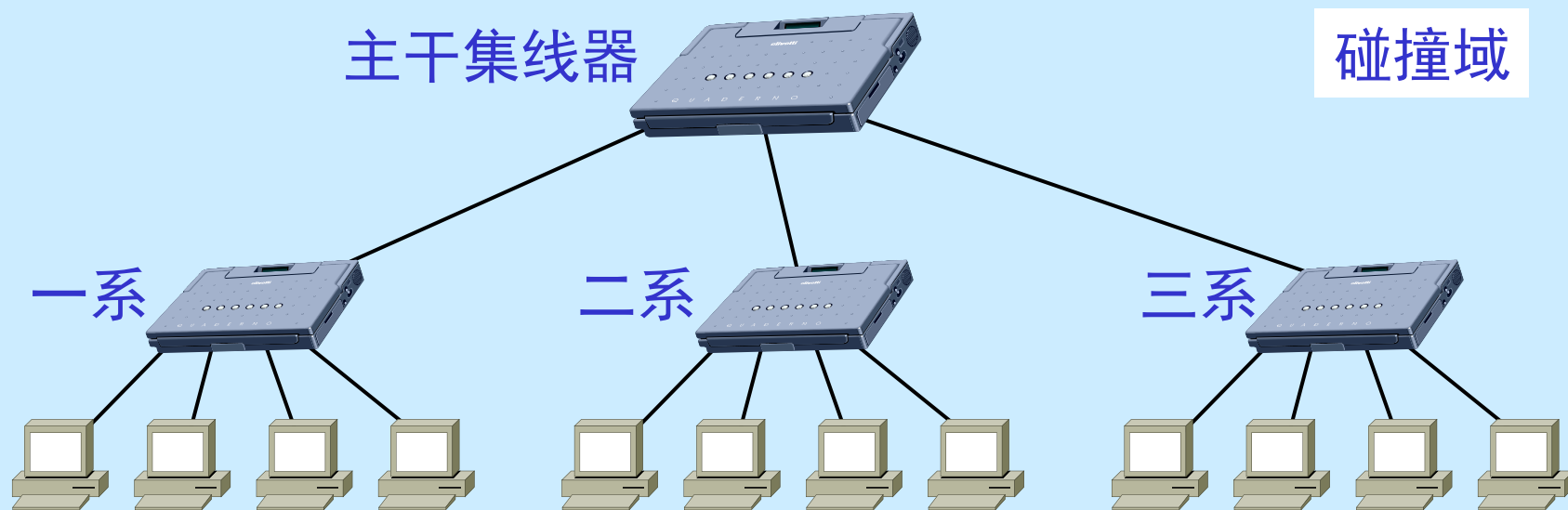


交换机自学习和转发帧的步骤归纳

- 交换机收到一帧后先进行**自学习**。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。
 - 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）。

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





独占传输媒体的带宽

- 对于普通 10 Mb/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 Mb/s)的 N 分之一。
- 使用以太网交换机时，虽然在每个接口到主机的带宽还是 10 Mb/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 对接口的交换机的总容量为 $N \times 10$ Mb/s。这正是交换机的最大优点。



交换器和集线器的不同

- (1) 数据转发给哪个端口,交换机基于MAC地址作出决定,集线器根本不做决定,而是将数据转发给所有端口.数据在交换机内部可以采用独立路径,在集线器中所有的数据都可以在所有的路径上流动.
- (2) 集线器所有端口共享一个带宽,交换即每个端口有自己独立的带宽,互不影响.
- (3) 集线器所有端口均是同一个冲突域,而交换机每个端口下是一个独立的冲突域.

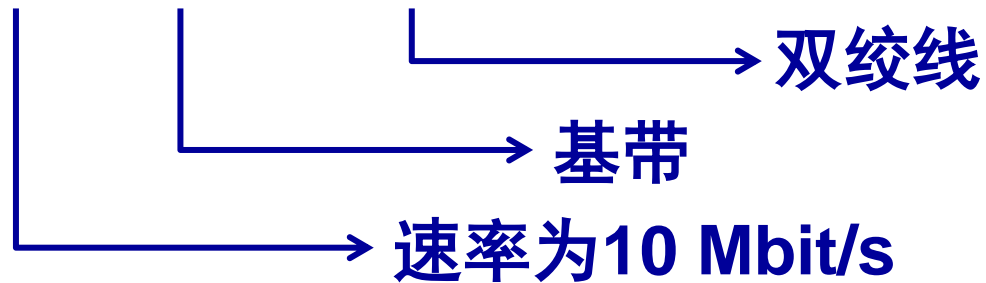


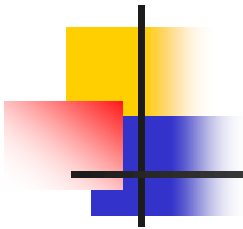
星形以太网 10BASE-T

1990 年，IEEE 制定出星形以太网

10BASE-T 的标准 802.3i。

10BASE-T





第四章 网络层

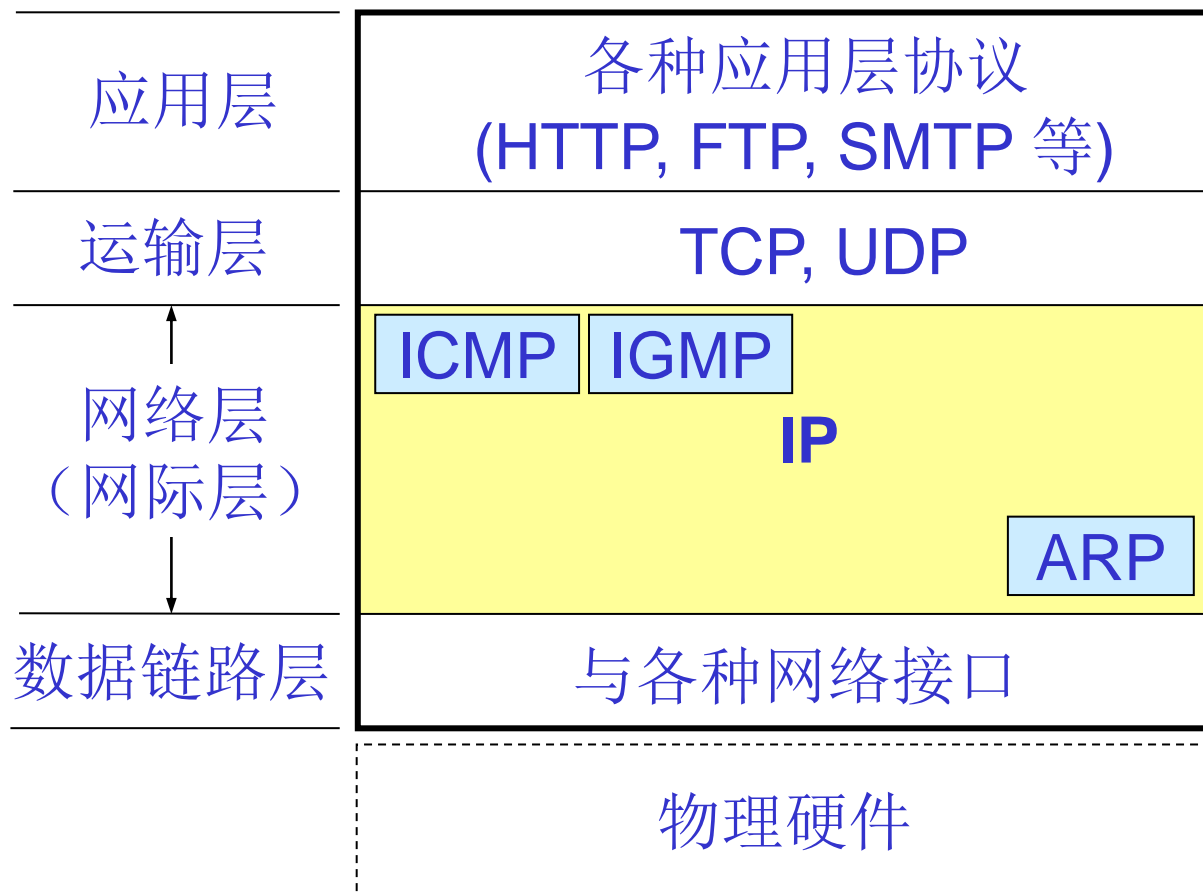
网络层向上只提供简单灵活的、无连接的、尽最大努力交付的数据报服务。



4.2 网际协议IP

- 网际协议 IP 是 TCP/IP 体系中两个最主要的协议之一。
与 IP 协议配套使用的还有三个协议：
- 地址解析协议 ARP
(Address Resolution Protocol)
- 网际控制报文协议 ICMP
(Internet Control Message Protocol)
- 网际组管理协议 IGMP
(Internet Group Management Protocol)

网际层的 IP 协议及配套协议



网络互相连接起来 要使用一些中间设备

- 中间设备又称为中间系统或中继(relay)系统。
 - 物理层中继系统：转发器又叫中继器(repeater), 。
 - 数据链路层中继系统：网桥或桥接器(bridge)。
 - 网络层中继系统：路由器(router)。
 - 网络层以上的中继系统：网关(gateway)-通过高层协议转换链接两个不兼容的系统（无线，有线；以太网，光纤）。



分类 IP 地址

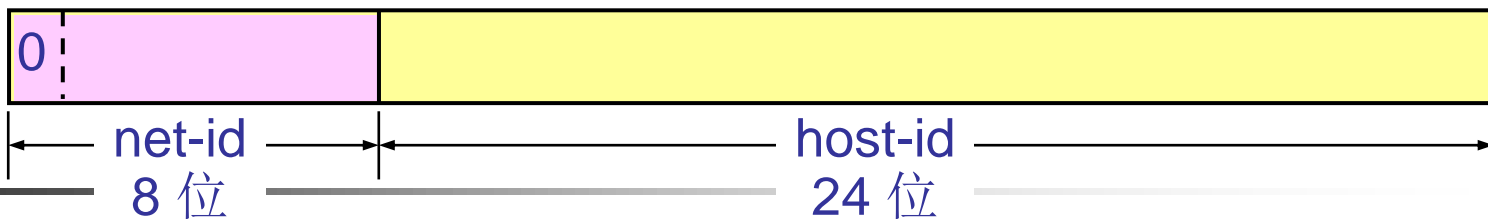
- 每一类地址都由两个固定长度的字段组成，其中一个字段是**网络号 net-id**，它标志主机（或路由器）所连接到的网络，而另一个字段则是**主机号 host-id**，它标志该主机（或路由器）。
- 两级的 IP 地址可以记为：

IP 地址 ::= { <网络号>, <主机号> } (4-1)

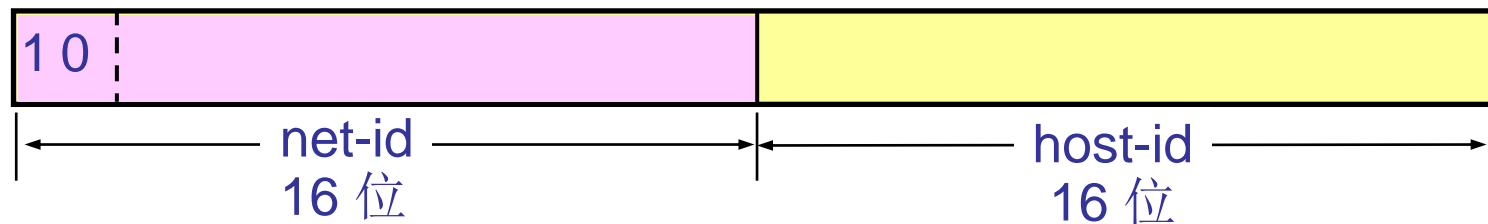
::= 代表 “**定义为**”

会判断地址类型！！！！

A 类地址



B 类地址



C 类地址



D 类地址



E 类地址



一般不使用的特殊的 IP 地址

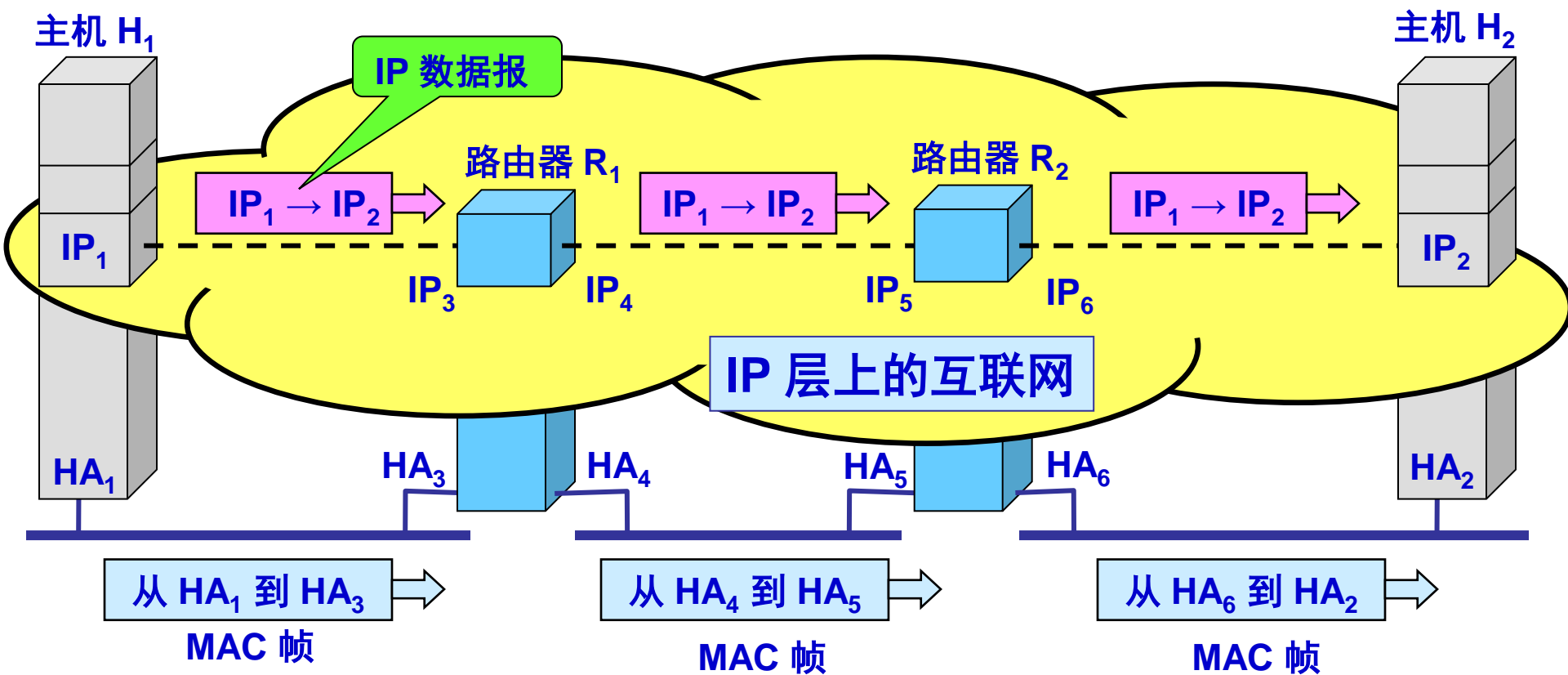
网络号	主机号	源地址使用	目的地址使用	代表的意思
0	0	可以	不可	在本网络上的本主机（见 6.6 节 DHCP 协议）
0	host-id	可以	不可	在本网络上的某台主机 host-id
全 1	全 1	不可	可以	只在本网络上进行广播（各路由器均不转发）
net-id	全 1	不可	可以	对 net-id 上的所有主机进行广播
127	非全 0 或全 1 的任何数	可以	可以	用作本地软件环回测试之用，仅在网络层交互，不经过链路和物理层



注意

- 一般全0地址和全1地址都不可用作主机地址。因此C类地址最大可用主机数为：
- 254个（255-1个全1广播）

IP 层抽象的互联网屏蔽了下层很复杂的细节。
在抽象的网络层上讨论问题，就能够使用
统一的、抽象的 IP 地址
研究主机和主机或主机和路由器之间的通信。

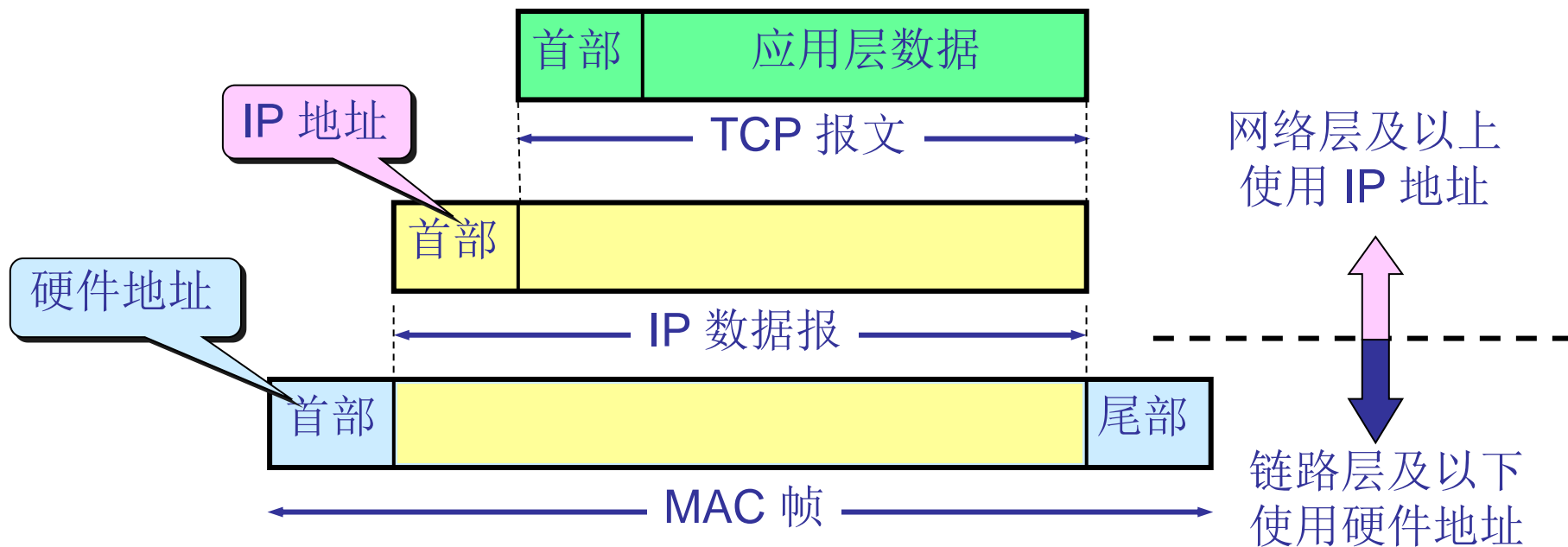




主机 H_1 与 H_2 通信中使用的 IP地址 与 硬件地址 HA

	在网络层 写入 IP 数据报首部的地址		在数据链路层 写入 MAC 帧首部的地址	
	源地址	目的地址	源地址	目的地址
从 H_1 到 R_1	IP_1	IP_2	HA_1	HA_3
从 R_1 到 R_2	IP_1	IP_2	HA_4	HA_5
从 R_2 到 H_2	IP_1	IP_2	HA_6	HA_2

4.2.3 IP 地址与硬件地址





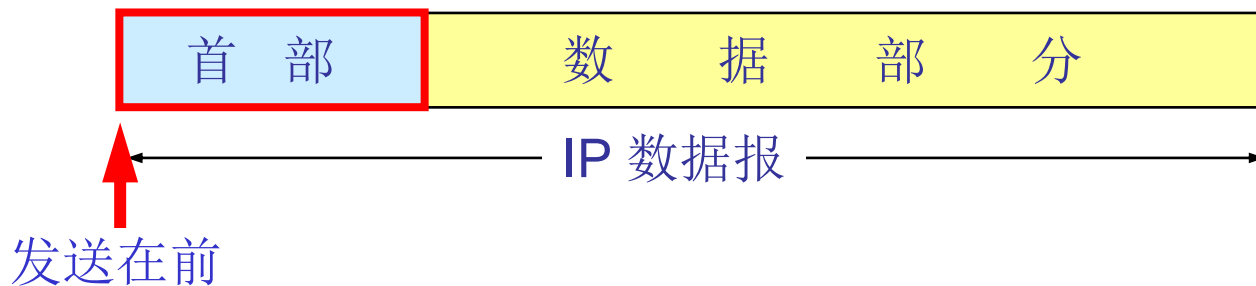
地址解析协议 ARP

- ARP: IP-->MAC, IP地址与MAC地址的映射问题
- ARP由于ARP协议没有相关的安全验证之类的设计,导致了ARP欺骗的产生

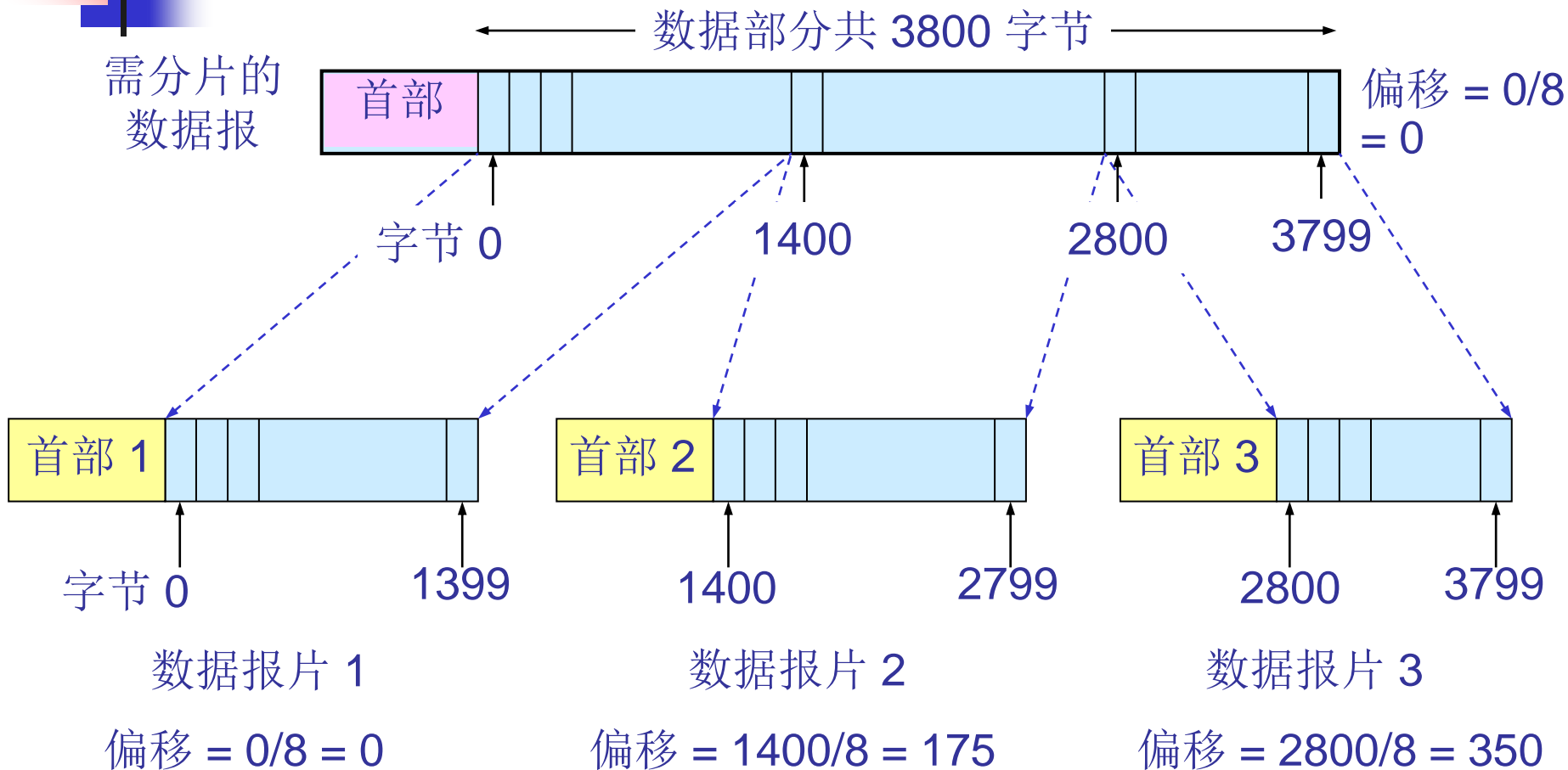


4.2.5 IP 数据报的格式

- 一个 IP 数据报由首部和数据两部分组成。
 - 首部的前一部分是固定长度，共 20 字节，是所有 IP 数据报必须具有的。
 - 在首部的固定部分的后面是一些可选字段，其长度是可变的。



【例4-1】 IP 数据报分片





特定主机路由

- 这种路由是为特定的目的主机指明一个路由。
- 采用特定主机路由可使网络管理人员能更方便地控制网络和测试网络，同时也可在需要考虑某种安全问题时采用这种特定主机路由。



默认路由(default route)

- 路由器还可采用**默认路由**以减少路由表所占用的空间和搜索路由表所用的时间。
- 这种转发方式在一个网络只有很少的对外连接时是很有用的。
- 默认路由在主机发送 IP 数据报时往往更能显示出它的好处。
- 如果一个主机连接在一个小网络上，而这个网络只用一个路由器和因特网连接，那么在这种情况下使用默认路由是非常合适的。



划分子网的基本思路

- 划分子网纯属一个单位内部的事情。单位对外仍然表现为没有划分子网的网络。
- 从主机号借用若干个位作为子网号 subnet-id, 而主机号 host-id 也就相应减少了若干个位。

IP地址 ::= {<网络号>, <子网号>, <主机号>} (4-2)



2. 子网掩码

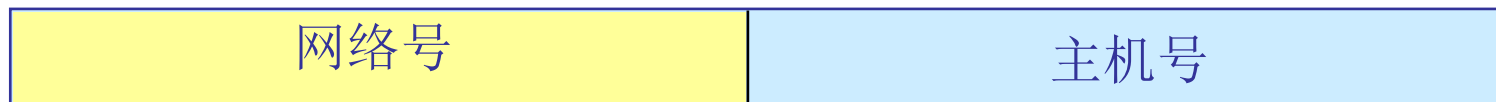
- 从一个 IP 数据报的首部并**无法判断**源主机或目的主机所连接的网络是否进行了子网划分。
- 使用**子网掩码** (subnet mask) 可以找出 IP 地址中的子网部分。

IP 地址的各字段和子网掩码

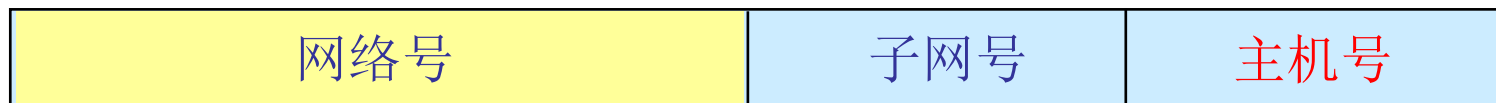


(IP 地址) AND (子网掩码) = 网络地址

两级 IP 地址

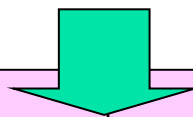
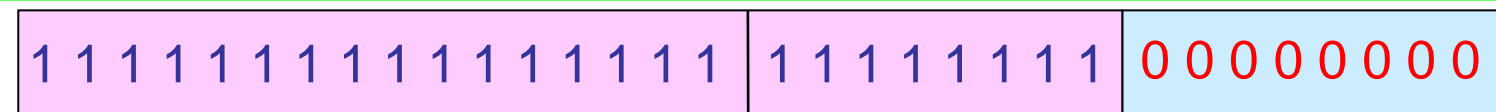


三级 IP 地址

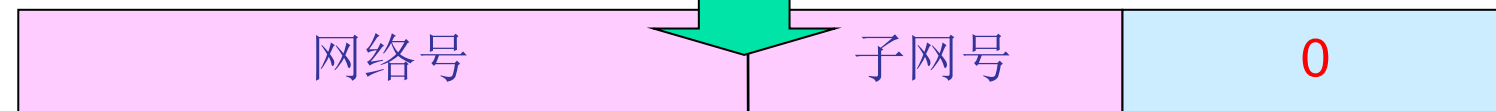


逐位进行 AND 运算

三级 IP 地址
的子网掩码



子网的
网络地址



【例4-2】 已知 IP 地址是 141.14.72.24，子网掩码是 255.255.192.0。试求网络地址。

(a) 点分十进制表示的 IP 地址

141 . 14 . 72 . 24

(b) IP 地址的第 3 字节是二进制

141 . 14 . 01001000 . 24

(c) 子网掩码是 255.255.192.0

11111111 11111111 11000000 00000000

(d) IP 地址与子网掩码逐位相与

141 . 14 . 01000000 . 0

(e) 网络地址（点分十进制表示）

141 . 14 . 64 . 0

【例4-3】 在上例中，若子网掩码改为 255.255.224.0。试求网络地址，讨论所得结果。

(a) 点分十进制表示的 IP 地址

141 . 14 . 72 . 24

(b) IP 地址的第 3 字节是二进制

141 . 14 . 01001000 . 24

(c) 子网掩码是 255.255.224.0

11111111 11111111 11100000 00000000

(d) IP 地址与子网掩码逐位相与

141 . 14 . 01000000 . 0

(e) 网络地址（点分十进制表示）

141 . 14 . 64 . 0

不同的子网掩码得出**相同**的网络地址。
但不同的掩码的效果是不同的。

在划分子网的情况下路由器转发分组的算法

不用背会做书后题

- (1) 从收到的分组的首部提取目的 IP 地址 D 。
- (2) 先用各网络的子网掩码和 D 逐位相“与”，看是否和相应的网络地址匹配。若匹配，则将分组直接交付。否则就是间接交付，执行(3)。
- (3) 若路由表中有目的地址为 D 的特定主机路由，则将分组传送给指明的下一跳路由器；否则，执行(4)。
- (4) 对路由表中的每一行的子网掩码和 D 逐位相“与”，若其结果与该行的目的网络地址匹配，则将分组传送给该行指明的下一跳路由器；否则，执行(5)。
- (5) 若路由表中有一个默认路由，则将分组传送给路由表中所指明的默认路由器；否则，执行(6)。
- (6) 报告转发分组出错。



IP 编址问题的演进

- 二级地址分类-->三级地址(划分子网)->无分类域间路由选择 (Classless Inter-Domain Routing)。



无分类的两级编址

- 无分类的两级编址的记法是：

IP地址 ::= {<网络前缀>, <主机号>} (4-3)

- CIDR 还使用“斜线记法” (slash notation), 它又称为**CIDR记法**, 即在 IP 地址面加上一个斜线 “/”, 然后写上网络前缀所占的位数 (这个数值对应于三级编址中子网掩码中 1 的个数)。
- CIDR 把网络前缀都相同的连续的 IP 地址组成“**CIDR 地址块**”。



CIDR 地址块

- 128.14.32.0/20 表示的地址块共有 2^{12} 个地址（因为斜线后面的 20 是网络前缀的位数，所以这个地址的主机号是 12 位）。
- 这个地址块的起始地址是 128.14.32.0。
- 在不需要指出地址块的起始地址时，也可将这样的地址块简称为“/20 地址块”。
- 128.14.32.0/20 地址块的最小地址：128.14.32.0
- 128.14.32.0/20 地址块的最大地址：128.14.47.255
- 全 0 和全 1 的主机号地址一般不使用。

128.14.32.0/20 表示的地址 (2^{12} 个地址)

最小地址



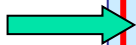
```
10000000 00001110 00100000 00000000
10000000 00001110 00100000 00000001
10000000 00001110 00100000 00000010
10000000 00001110 00100000 00000011
10000000 00001110 00100000 00000100
10000000 00001110 00100000 00000101
```

所有地址
的 20 位
前缀都是
一样的

```
10000000 00001110 00101111 11111011
10000000 00001110 00101111 11111100
10000000 00001110 00101111 11111101
10000000 00001110 00101111 11111110
10000000 00001110 00101111 11111111
```

...

最大地址

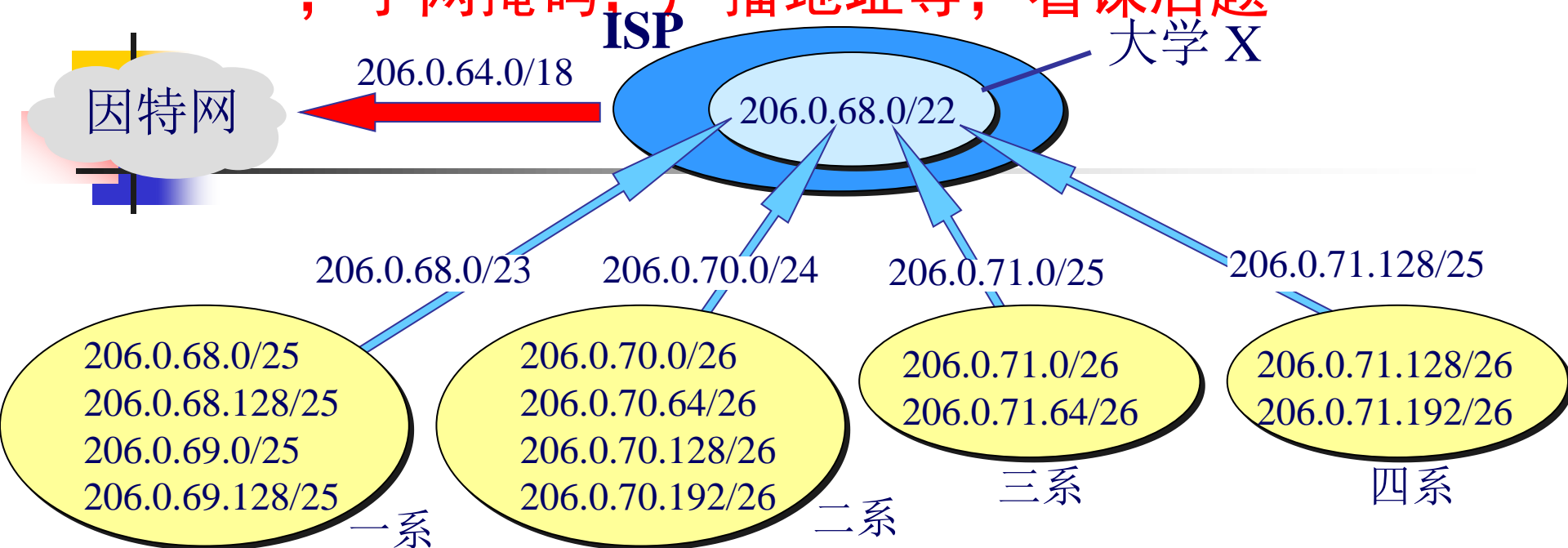




CIDR 记法的其他形式

- 10.0.0.0/10 可简写为 10/10，也就是将点分十进制中低位连续的 0 省略。
- 10.0.0.0/10 相当于指出 IP 地址 10.0.0.0 的掩码是 255.192.0.0，即
11111111 11000000 00000000 00000000
- 网络前缀的后面加一个星号 * 的表示方法
如 00001010 00*，在星号 * 之前是网络前缀，而星号 * 表示 IP 地址中的主机号，可以是任意值。

CIDR 地址块划分：画二叉树，地址块开始结束，子网掩码，广播地址等，看课后题



单位	地址块	二进制表示	地址数
ISP	206.0.64.0/18	11001110.00000000.01*	16384
大学	206.0.68.0/22	11001110.00000000.010001*	1024
一系	206.0.68.0/23	11001110.00000000.0100010*	512
二系	206.0.70.0/24	11001110.00000000.01000110.*	256
三系	206.0.71.0/25	11001110.00000000.01000111.0*	128
四系	206.0.71.128/25	11001110.00000000.01000111.1*	128



2. 最长前缀匹配

- 使用 CIDR 时，路由表中的每个项目由“网络前缀”和“下一跳地址”组成。在查找路由表时可能会得到不止一个匹配结果。
- 应当从匹配结果中选择具有最长网络前缀的路由：**最长前缀匹配**(longest-prefix matching)。
- 网络前缀越长，其地址块就越小，因而路由就越具体(more specific)。
- 最长前缀匹配又称为**最长匹配**或**最佳匹配**。

最长前缀匹配举例

收到的分组的地址 $D = 206.0.71.128$

路由表中的项目: 206.0.68.0/22 (ISP)
206.0.71.128/25 (四系)

查找路由表中的第 1 个项目

第 1 个项目 206.0.68.0/22 的掩码 M 有 22 个连续的 1。

$M = 11111111\ 11111111\ 11111100\ 00000000$

因此只需把 D 的第 3 个字节转换成二进制。

$M = 11111111\ 11111111\ 11111100\ 00000000$

AND $D =$ 206. 0. 01000111. 0

206. 0. 01000100. 0

与 206.0.68.0/22 匹配

最长前缀匹配举例

收到的分组的地址 $D = 206.0.71.128$

路由表中的项目: $206.0.68.0/22$ (ISP)
 $206.0.71.128/25$ (四系)

再查找路由表中的第 2 个项目

第 2 个项目 $206.0.71.128/25$ 的掩码 M 有 25 个连续的 1。

$M = 11111111\ 11111111\ 11111111\ 10000000$

因此只需把 D 的第 4 个字节转换成二进制。

$M = 11111111\ 11111111\ 11111111\ 10000000$

AND $D = 206.\quad 0.\quad 71.\ 10000000$

$206.\quad 0.\quad 71.\ 10000000$

与 $206.0.71.128/25$ 匹配

最长前缀匹配

$D \text{ AND } (11111111 \ 11111111 \ 11111100 \ 00000000)$
= 206.0.68.0/22 匹配

$D \text{ AND } (11111111 \ 11111111 \ 11111111 \ 10000000)$
= 206.0.71.128/25 匹配

- 选择两个匹配的地址中更具体的一个，即选择**最长前缀的地址**。



4.4 网际控制报文协议 ICMP

- 为了提高 IP 数据报交付成功的机会，在网际层使用了网际控制报文协议 ICMP (Internet Control Message Protocol)。
- ICMP 允许主机或路由器**报告差错情况和提供有关异常情况的报告**。
- ICMP 不是高层协议，而是 **IP 层的协议**。
- ICMP 报文作为 IP 层数据报的数据，加上数据报的首部，组成 IP 数据报发送出去。



4.4.1 ICMP 报文的种类

- ICMP 报文的种类有两种，即 ICMP 差错报告报文和 ICMP 询问报文。



ICMP 差错报告报文共有 5 种

- 终点不可达（出现丢包就发这个）
- 源点抑制(Source quench)
- 时间超过
- 参数问题
- 改变路由（重定向）(Redirect)



不应发送 ICMP 差错报告报文的几种情况

- 对 ICMP 差错报告报文不再发送 ICMP 差错报告报文。
- 对第一个分片的数据报片的所有后续数据报片都不发送 ICMP 差错报告报文。
- 对具有多播地址的数据报都不发送 ICMP 差错报告报文。
- 对具有特殊地址（如127.0.0.0 或 0.0.0.0）的数据报不发送 ICMP 差错报告报文。

4.4.2 ICMP的应用举例

PING (Packet InterNet Groper)

- PING 用来测试两个主机之间的连通性。
- PING 使用了 ICMP 回送请求与回送回答报文。
- PING 是应用层直接使用网络层 ICMP 的例子，它没有通过运输层的 TCP 或UDP。



Traceroute 的原理

追踪IP数据包所经过的路由器的命令。

原理是利用ICMP的差错报文，类型为时间超过
设置IP首部TTL值为n，当到达第n路由器时，路
由器会因为超时而丢弃这个IP报文，并发回一个
ICMP差错报告给源端，源端就可以从这个报告
的IP首部取到源IP地址即路由器的地址



从路由算法的自适应性考虑

- 静态路由选择策略——人工配置。
- 动态路由选择策略——路由算法自动生成。



自治系统 AS (Autonomous System)

- 自治系统 AS 的定义：在单一的技术管理下的一组路由器，使用一致的路由选择策略。

自治系统和 内部网关协议、外部网关协议



自治系统之间的路由选择也叫做
域间路由选择(interdomain routing),
在自治系统内部的路由选择叫做
域内路由选择(intradomain routing)



“距离”的定义

- 从一路由器到**直接连接**的网络的距离定义为 1。
- 从一个路由器到非直接连接的网络的距离定义为所经过的路由器数加 1。
- RIP 协议中的“距离”也称为“**跳数**” (hop count)，因为每经过一个路由器，跳数就加 1。
- 这里的“距离”实际上指的是“**最短距离**”，



“距离” 的定义

- RIP 认为一个好的路由就是它通过的路由器的数目少，即“距离短”。
- RIP 允许一条路径最多只能包含 15 个路由器。
- “距离”的最大值为16 时即相当于不可达。可见 RIP 只适用于小型互联网。
- RIP 不能在两个网络之间同时使用多条路由。RIP 选择一个具有最少路由器的路由（即最短路由），哪怕还存在另一条高速(低时延)但路由器较多的路由。



RIP 协议的三个要点

- 仅和相邻路由器交换信息。
- 交换的信息是当前本路由器所知道的全部信息，即自己的路由表。
- 按固定的时间间隔交换路由信息，例如，每隔 30 秒。
- 特点：好消息传得快,坏消息传得慢

2. 距离向量算法—不要背会做题

收到相邻路由器（其地址为 X）的一个 RIP 报文：

(1) 先修改此 RIP 报文中的所有项目：把“下一跳”字段中的地址都改为 X，并把所有的“距离”字段的值加 1。

(2) 对修改后的 RIP 报文中的每一个项目，重复以下步骤：

若项目中的目的网络不在路由表中，则把该项目加到路由表中。

否则

若下一跳字段给出的路由器地址是同样的，则把收到的项目替换原路由表中的项目。

否则

若收到项目中的距离小于路由表中的距离，则进行更新，
否则，什么也不做。

(3) 若 3 分钟还没有收到相邻路由器的更新路由表，则把此相邻路由器记为不可达路由器，即将距离置为16（距离为16表示不可达）。

(4) 返回。

4.5.3 内部网关协议 OSPF

(Open Shortest Path First)

- 注意自己总结下RIP和OSPF的不同
- 例如：
- RIP是一种距离向量协议，而OSPF是一种链路状态协议
- RIP只跟邻居路由器交换信息，OSPF是泛洪
- RIP最大距离16，只适合小规模网络。OSPF适合大规模网络
- ○ ○ ○ ○



三个要点

- 向本自治系统中所有路由器发送信息，这里使用的方法是洪泛法。
- 发送的信息就是与本路由器相邻的所有路由器的链路状态，但这只是路由器所知道的部分信息。
 - “链路状态”就是说明本路由器都和哪些路由器相邻，以及该链路的“度量”(metric)。
- 只有当链路状态发生变化时，路由器才用洪泛法向所有路由器发送此信息。



链路状态数据库

(link-state database)

- 由于各路由器之间频繁地交换链路状态信息，因此所有的路由器最终都能建立一个链路状态数据库。
- 这个数据库实际上就是**全网的拓扑结构图**，它在全网范围内是一致的（这称为链路状态数据库的同步）。
- OSPF 的链路状态数据库能较快地进行更新，使各个路由器能及时更新其路由表。OSPF 的更新过程收敛得快是其重要优点。



4.5.4 外部网关协议 BGP

- BGP 是不同自治系统的路由器之间交换路由信息的协议。
- BGP交换:自治系统间的可达性

典型的路由器的

注意路由表和转发表的区别和联系，影子副本的作用

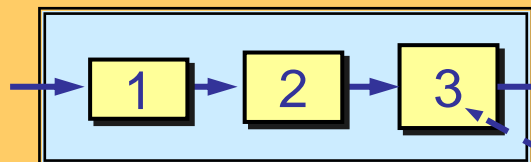
3——网络层
2——数据链路层
1——物理层

路由选择处理机

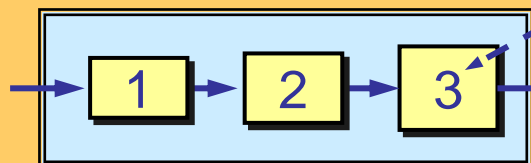
路由选择协议

路由表

输入端口



输入端口

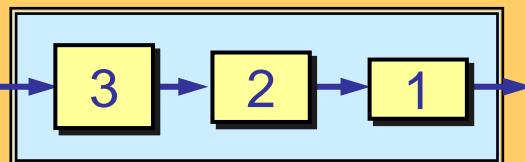


分组处理

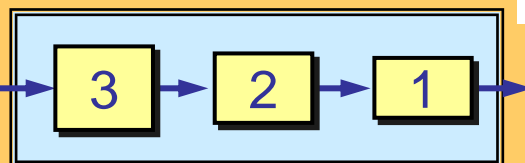
转发表

交换结构

输出端口



输出端口



选择

分组
转发



2. 交换结构

- 交换结构是路由器的关键构件。
- 正是这个交换结构把分组从一个输入端口转移到某个合适的输出端口。
- 实现交换有多种方法。常用交换方法有三种：
 - 通过存储器
 - 通过总线
 - 通过纵横交换结构



4.6.3 网际组管理协议 IGMP 和多播路由选择协议

1. IP多播需要两种协议

- 为了使路由器知道多播组成员的信息，需要利用网际组管理协议 IGMP (Internet Group Management Protocol)。
- 连接在局域网上的多播路由器还必须和因特网上的其他多播路由器协同工作，以便把多播数据报用最小代价传送给所有的组成员。这就需要使用多播路由选择协议。

4.7 虚拟专用网 VPN 和网络地址转换 NAT

4.7.1 虚拟专用网 VPN

- **本地地址**——仅在机构内部使用的 IP 地址，可以由本机构自行分配，而不需要向因特网的管理机构申请。
- **全球地址**——全球唯一的IP地址，必须向因特网的管理机构申请。

RFC 1918 指明的专用地址 (private address)

- 10.0.0.0 到 10.255.255.255
- 172.16.0.0 到 172.31.255.255
- 192.168.0.0 到 192.168.255.255
- 这些地址只能用于一个机构的内部通信，而不能用于和因特网上的主机通信。
- 专用地址只能用作本地地址而不能用作全球地址。在因特网中的所有路由器对目的地址是专用地址的数据报一律不进行转发。



虚拟专用网 VPN

- 利用公用的互联网作为本机构各专用网之间的通信载体，这样的专用网又称为**虚拟专用网VPN** (Virtual Private Network)。
- “**专用网**”为本机构的主机用于机构内部通信的网络，比外网安全。



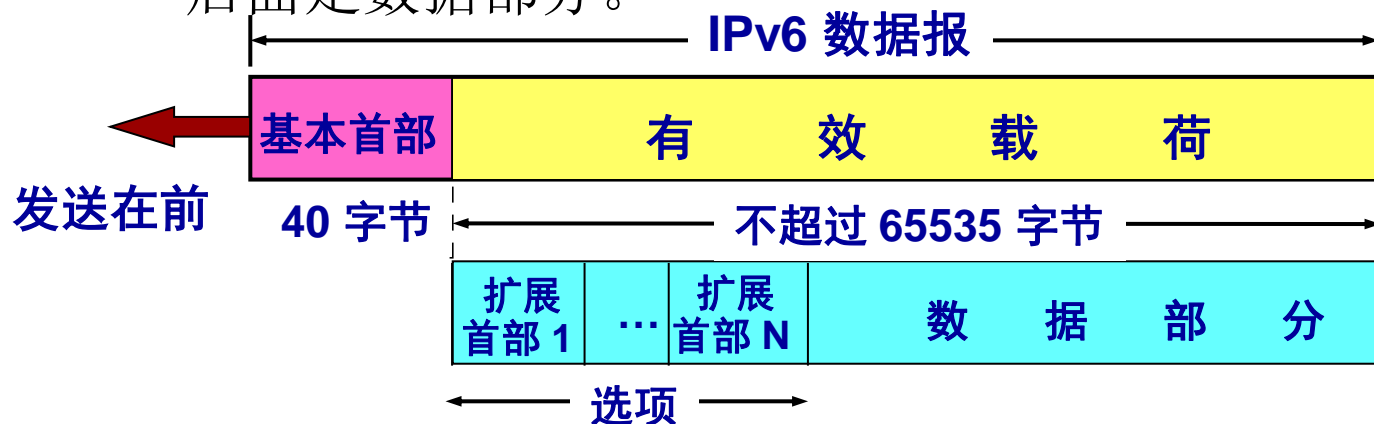
4.6.1 IPv6 的基本首部

- IPv6 仍支持无连接的传送。
- 所引进的主要变化如下：
 - 更大的地址空间。IPv6 将地址从 IPv4 的 32 位增大到了 128 位。
 - 扩展的地址层次结构。
 - 灵活的首部格式。IPv6 定义了许多可选的扩展首部。
 - 改进的选项。IPv6 允许数据报包含有选项的控制信息，其选项放在有效载荷中。

IPv6 数据报的一般形式

- IPv6 数据报由两大部分组成：

- 基本首部 (base header)
- 有效载荷 (payload)。有效载荷也称为净负荷。有效载荷允许有零个或多个扩展首部 (extension header)，再后面是数据部分。



具有多个可选扩展首部的 IPv6 数据报的一般形式

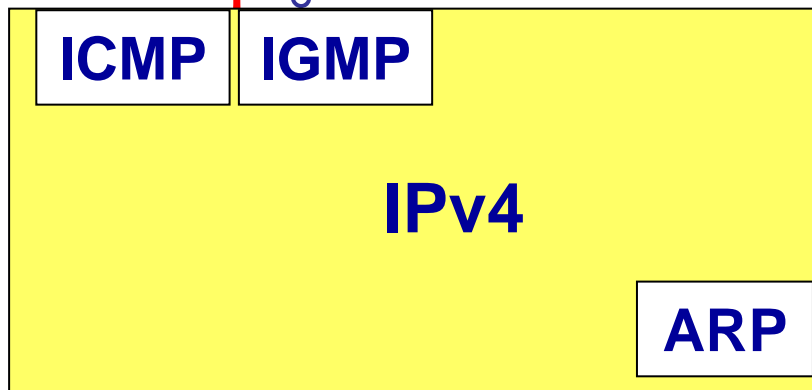


4.6.3 从 IPv4 向 IPv6 过渡

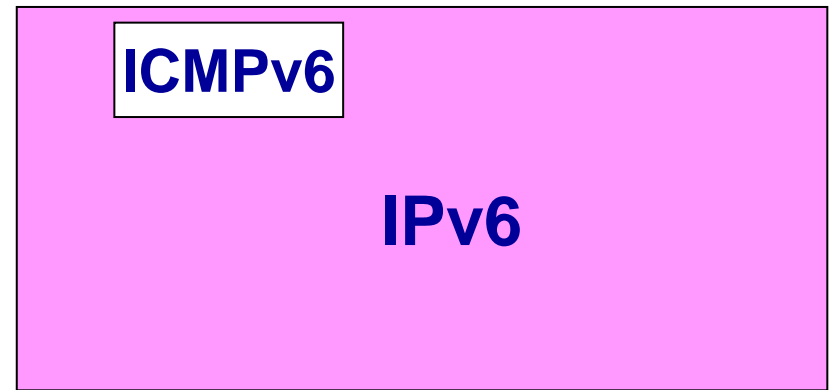
- 两种向 IPv6 过渡的策略：
 - 使用双协议栈
 - 使用隧道技术

4.6.4 ICMPv6

- 地址解析协议 ARP 和网际组管理协议 IGMP 协议的功能都已被合并到 ICMPv6 中。



版本 4 中的网络层



版本 6 中的网络层

新旧版本中的网络层的比较



4.8.1 虚拟专用网 VPN

- 利用公用的互联网作为本机构各专用网之间的通信载体，这样的专用网又称为**虚拟专用网**VPN (Virtual Private Network)。
- “**专用网**”是因为这种网络是为本机构的主机用于**机构内部的通信**，而不是用于和网络外非本机构的主机通信。



本地地址与全球地址

- **本地地址**——仅在机构内部使用的 IP 地址，可以由本机构自行分配，而不需要向互联网的管理机构申请。
- **全球地址**——全球唯一的 IP 地址，必须向互联网的管理机构申请。
- **问题：**在内部使用的本地地址就有可能和互联网中某个 IP 地址重合，这样就会出现地址的二义性问题。



本地地址与全球地址

- **问题：**在内部使用的本地地址就有可能和互联网中某个 IP 地址重合，这样就会出现地址的**二义性**问题。
- **解决：**RFC 1918 指明了一些**专用地址** (private address)。专用地址只能用作本地地址而不能用作全球地址。在互联网中的所有路由器，对目的地址是专用地址的数据报一律不进行转发。





第五章 运输层



应用进程之间的通信

- 两个主机进行通信实际上就是两个主机中的**应用进程互相通信**。
- 应用进程之间的通信又称为**端到端的通信**。
- 运输层的一个很重要的功能就是**复用和分用**。
- 应用层不同进程的报文通过不同的**端口**向下交到运输层，再往下就共用网络层提供的服务。



5.1.2 运输层的两个主要协议

TCP/IP 的运输层有两个不同的协议：

- (1) 用户数据报协议 UDP
(User Datagram Protocol)
- (2) 传输控制协议 TCP
(Transmission Control Protocol)



三类端口

- **熟知端口**，数值一般为 0~1023。
- **登记端口号**，数值为1024~49151，为没有熟知端口号的应用程序使用的。使用这个范围的端口号必须在 **IANA 登记**，以防止重复。
- **客户端口号或短暂端口号**，数值为 49152~65535，留给客户进程选择暂时使用。当服务器进程收到客户进程的报文时，就知道了客户进程所使用的动态端口号。通信结束后，这个端口号可供其他客户进程以后使用。

UDP 的主要特点

- (1) UDP 是无连接的，发送数据之前不需要建立连接，因此减少了开销和发送数据之前的时延。
- (2) UDP 使用尽最大努力交付(无连接,没有确认)，即不保证可靠交付，因此主机不需要维持复杂的连接状态表。
- (3) UDP 是面向报文的。UDP 对应用层交下来的报文，既不合并，也不拆分，而是保留这些报文的边界。UDP 一次交付一个完整的报文。
- (4) UDP 没有拥塞控制，因此网络出现的拥塞不会使源主机的发送速率降低。这对某些实时应用是很重要的。很适合多媒体通信的要求。



UDP 的主要特点

- (5) UDP 支持一对一、一对多、多对一和多对多的交互通信。
- (6) UDP 的首部开销小，只有 8 个字节，比 TCP 的 20 个字节的首部要短。

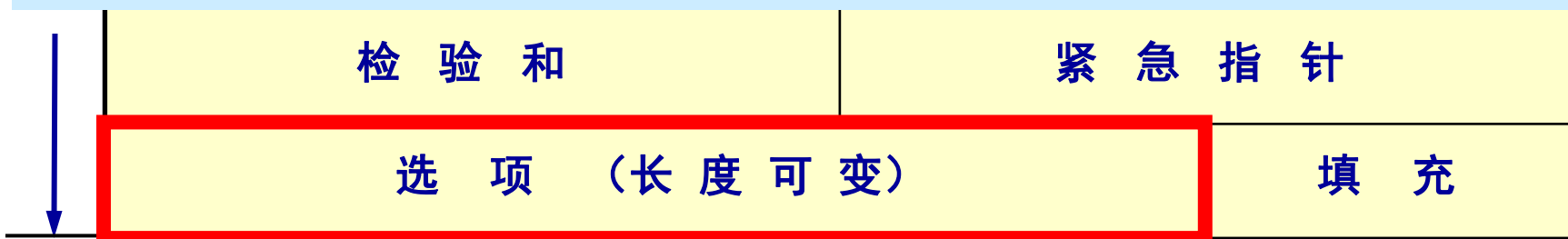
位 0 8 16 24 31



MSS (Maximum Segment Size)

是 TCP 报文段中的**数据字段**的最大长度。

数据字段加上 TCP 首部才等于整个的 TCP 报文段。
所以，MSS 是“TCP 报文段长度减去 TCP 首部长度的”。



选项字段 —— 长度可变。TCP 最初只规定了一种选项，即**最大报文段长度 MSS**。MSS 告诉对方 TCP：“我的缓存所能接收的报文段的数据字段的最大长度是 MSS 个字节。”



其他选项

- 窗口扩大选项 —— 占 3 字节，其中有一个字节表示移位值 S 。新的窗口值等于 TCP 首部中的窗口位数增大到 $(16 + S)$ ，相当于把窗口值向左移动 S 位后获得实际的窗口大小。
- 时间戳选项 —— 占 10 字节，其中最主要的字段时间戳值字段（4 字节）和时间戳回送回答字段（4 字节）。
- 选择确认选项 —— 在后面的 5.6.3 节介绍。

5.3 传输控制协议 TCP 概述

5.3.1 TCP 最主要的特点

- TCP 是面向连接的运输层协议。
- 每一条 TCP 连接只能有两个端点(endpoint)，每一条 TCP 连接只能是点对点的（一对一）。
- TCP 提供可靠交付的服务。
- TCP 提供全双工通信。
- 面向字节流。
- 总结：TCP提供面向连接的，可靠的，字节流服务



套接字 (socket)

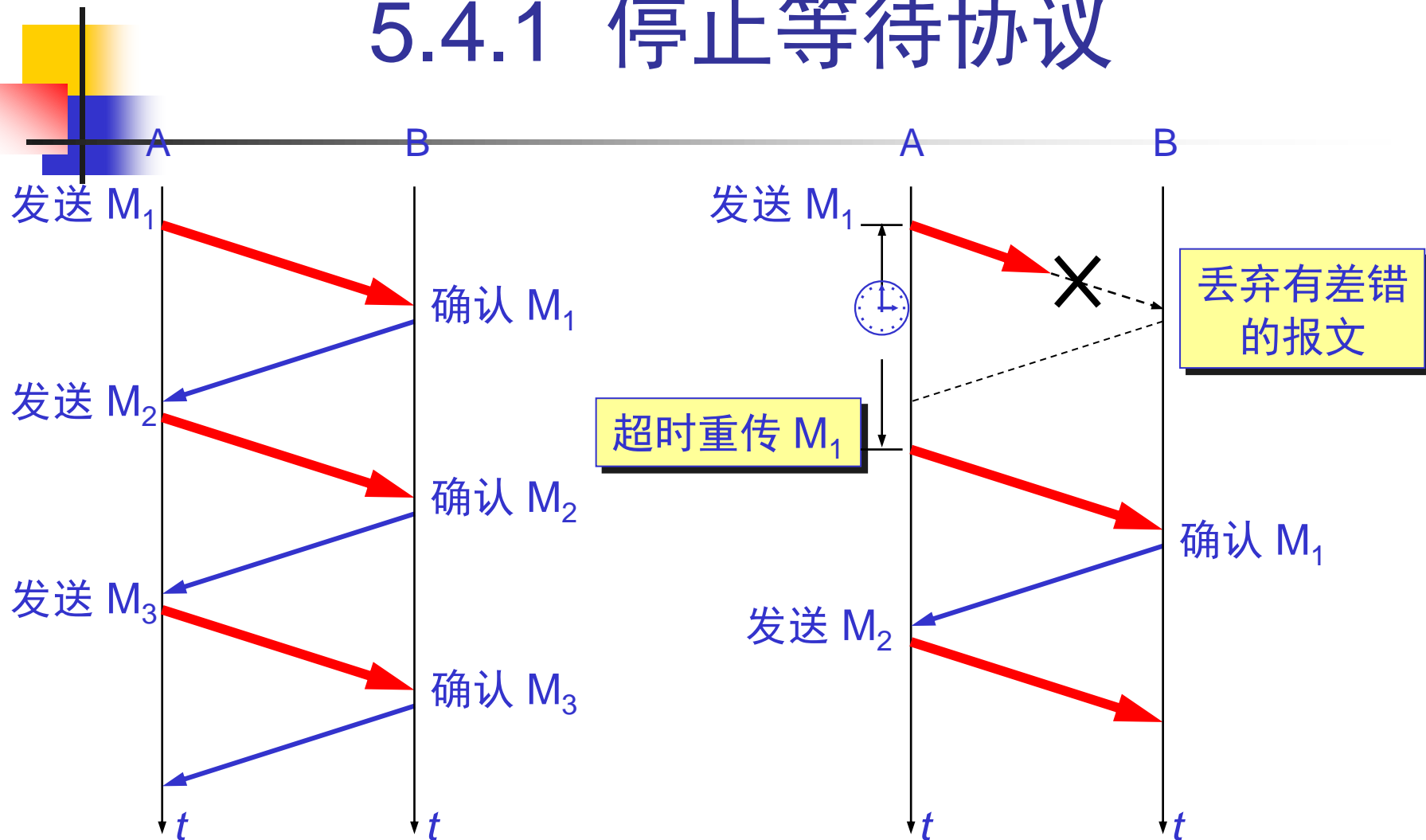
套接字 socket = (IP地址: 端口号) (5-1)

- 每一条 **TCP** 连接唯一地被通信两端的两个端点（即两个套接字）所确定。即：

TCP 连接 ::= {socket1, socket2}
= {(IP1: port1), (IP2: port2)} (5-2)

5.4 可靠传输的工作原理

5.4.1 停止等待协议



(a) 无差错情况

(b) 超时重传



可靠通信的实现

- 使用上述的确认和重传机制，我们就可以在不可靠的传输网络上实现可靠的通信。
- 这种可靠传输协议常称为自动重传请求 ARQ (Automatic Repeat reQuest)。
- ARQ 表明重传的请求是自动进行的。接收方不需要请求发送方重传某个出错的分组。



累积确认

- 接收方一般采用**累积确认**的方式。即不必对收到的分组逐个发送确认，而是对按序到达的最后一个分组发送确认，这样就表示：**到这个分组为止的所有分组都已正确收到了**。
- 累积确认有的优点是：容易实现，即使确认丢失也不必重传。缺点是：不能向发送方反映出接收方已经正确收到的所有分组的信息。



TCP 可靠通信的具体实现

- TCP 连接的每一端都必须设有两个窗口——一个发送窗口和一个接收窗口。
- TCP 的可靠传输机制用字节的序号进行控制。TCP 所有的确认都是基于序号而不是基于报文段。
- TCP 两端的四个窗口经常处于动态变化之中。
- TCP连接的往返时间 RTT 也不是固定不变的。需要使用特定的算法估算较为合理的重传时间。

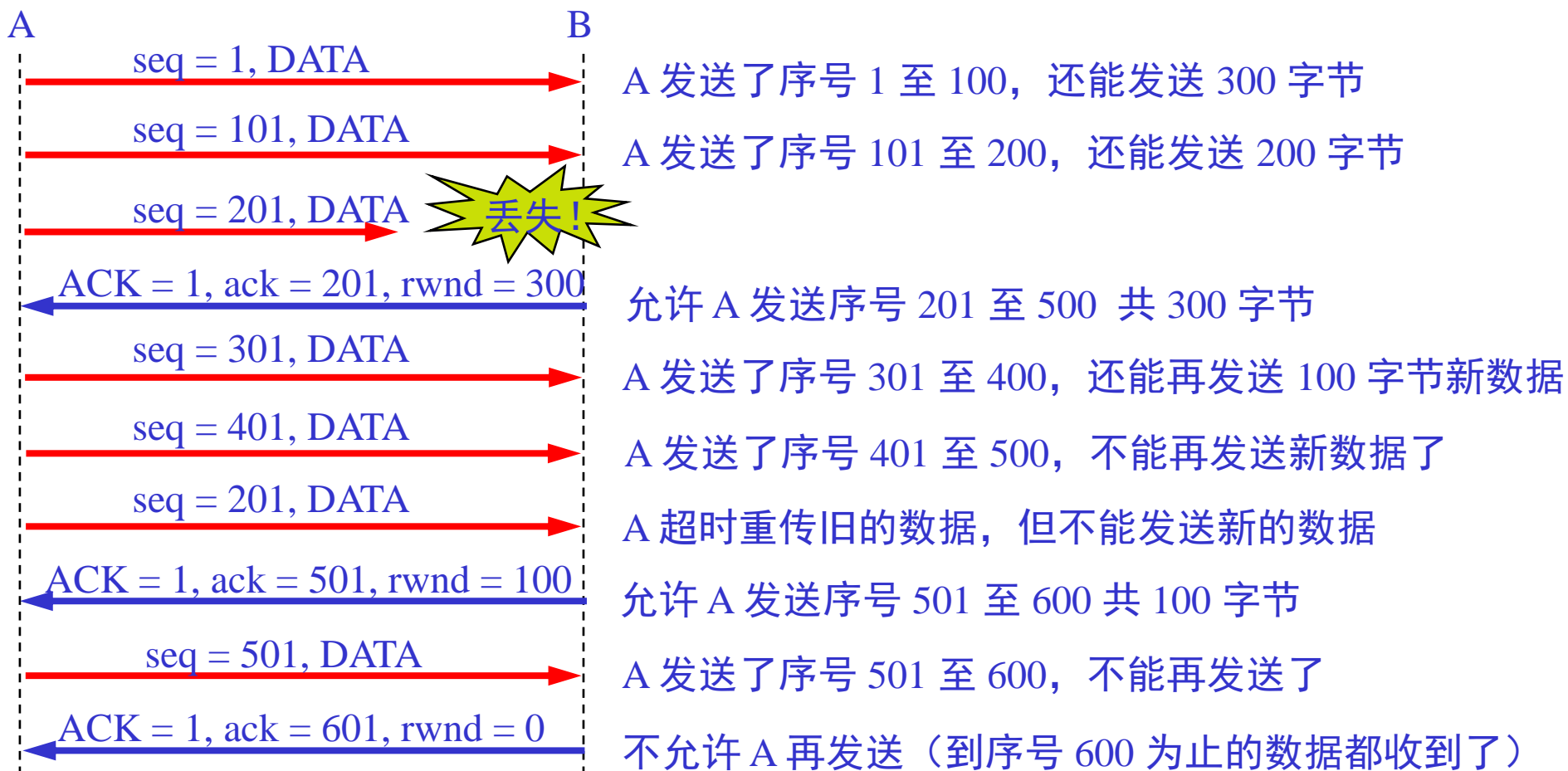
5.7 TCP 的流量控制

5.7.1 利用滑动窗口实现流量控制

- 一般说来，我们总是希望数据传输得更快一些。但如果发送方把数据发送得过快，接收方就可能来不及接收，这就会造成数据的丢失—接收缓存溢出了。
- **流量控制**(flow control)就是让发送方的发送速率不要太快，既要让接收方来得及接收，也不要使网络发生拥塞。
- 利用滑动窗口机制可以很方便地在 TCP 连接上实现流量控制。

流量控制举例

A 向 B 发送数据。在连接建立时，
B 告诉 A：“我的接收窗口 $\text{rwnd} = 400$ （字节）”。



零窗口与持续计时器

(作用,死锁的概念与解决方法)

- TCP 为每一个连接设有一个持续计时器。
- 只要 TCP 连接的一方收到对方的零窗口通知，就启动持续计时器。
- 若持续计时器设置的时间到期，就发送一个零窗口探测报文段（仅携带 1 字节的数据），而对方就在确认这个探测报文段时给出了现在的窗口值。
- 若窗口仍然是零，则收到这个报文段的一方就重新设置持续计时器。
- 若窗口不是零，则死锁的僵局就可以打破了。



拥塞控制与流量控制的关系

- 拥塞控制所要做的都有一个前提，就是网络能够承受现有的网络负荷。
- 拥塞控制是一个全局性的过程，涉及到所有的主机、所有的路由器，以及与降低网络传输性能有关的所有因素。
- 流量控制往往指在给定的发送端和接收端之间的点对点通信量的控制。
- 流量控制所要做的就是抑制发送端发送数据的速率，以便使接收端来得及接收。

5.8.2 几种拥塞控制方法

1. 慢开始和拥塞避免

- 发送方维持一个叫做**拥塞窗口 cwnd** (congestion window)的状态变量。拥塞窗口的大小取决于网络的拥塞程度，并且动态地在变化。发送方让自己的**发送窗口等于拥塞窗口**。如再考虑到接收方的接收能力，则发送窗口还可能**小于**拥塞窗口。
- 发送方控制拥塞窗口的原则是：只要网络没有出现拥塞，拥塞窗口就再增大一些，以便把更多的分组发送出去。但只要网络出现拥塞，拥塞窗口就减小一些，以减少注入到网络中的分组数。



慢开始算法的原理

- 在主机刚刚开始发送报文段时可先设置拥塞窗口 $cwnd = 1$ ，即设置为一个最大报文段 MSS 的数值。
- 在每收到一个对新的报文段的确认后，将拥塞窗口加 1，即增加一个 MSS 的数值。
- 用这样的方法逐步增大发送端的拥塞窗口 $cwnd$ ，可以使分组注入到网络的速率更加合理。



设置慢开始门限状态变量 ssthresh

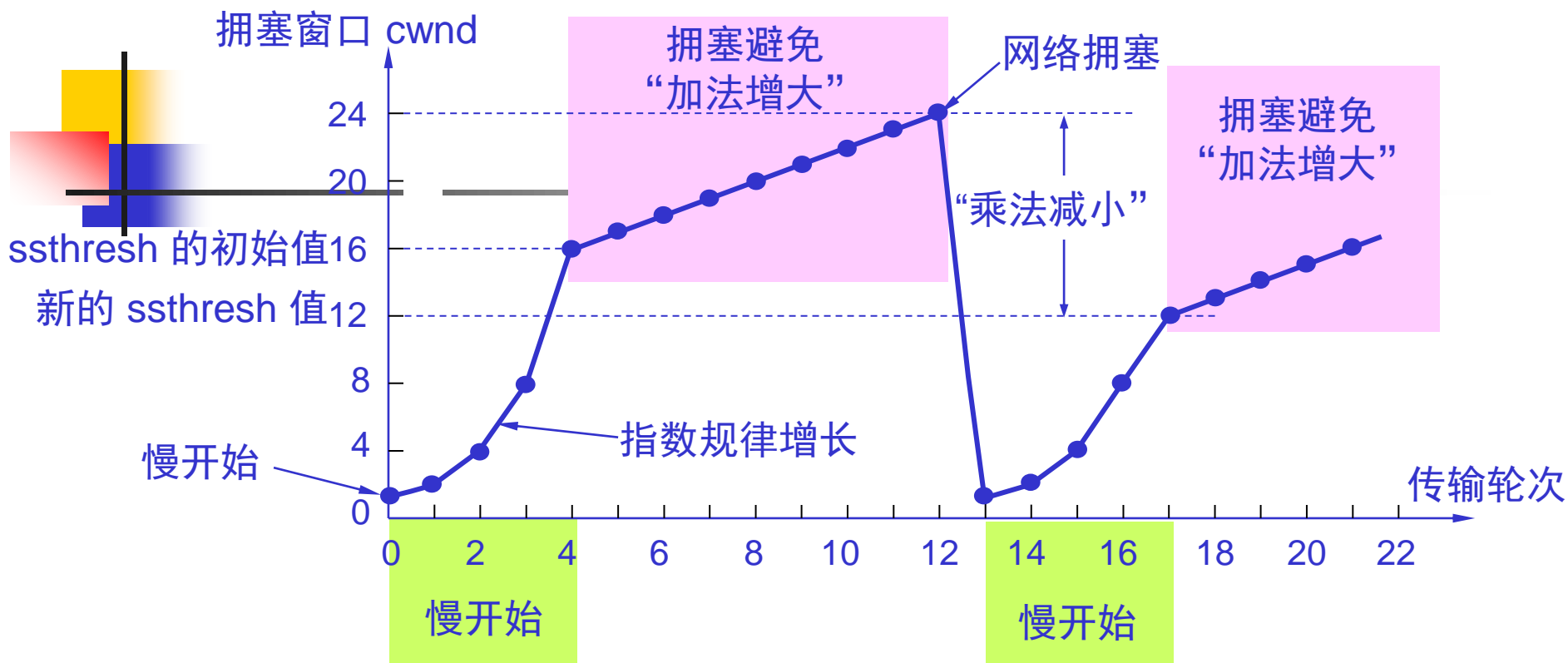
- 慢开始门限 ssthresh 的用法如下：
- 当 $cwnd < ssthresh$ 时，使用慢开始算法。
- 当 $cwnd > ssthresh$ 时，停止使用慢开始算法而改用拥塞避免算法。
- 当 $cwnd = ssthresh$ 时，既可使用慢开始算法，也可使用拥塞避免算法。
- 拥塞避免算法的思路是让拥塞窗口 $cwnd$ 缓慢地增大，即每经过一个往返时间 **RTT 就把发送方的拥塞窗口 $cwnd$ 加 1，而不是加倍**，使拥塞窗口 $cwnd$ 按线性规律缓慢增长。



当网络出现拥塞时

- 无论在慢开始阶段还是在拥塞避免阶段，只要发送方判断网络**出现拥塞**（其根据就是没有按时收到确认），就要把慢开始**门限** ssthresh 设置为出现拥塞时的**发送方窗口值的一半**（但不能小于2）。
- 然后把拥塞窗口 cwnd **重新设置为 1**，执行慢开始算法。
- 这样做的目的就是要迅速减少主机发送到网络中的分组数，使得发生拥塞的路由器有足够时间把队列中积压的分组处理完毕。

慢开始和拥塞避免算法的实现举例



当 TCP 连接进行初始化时，将拥塞窗口置为 1。图中的窗口单位不使用字节而使用**报文段**。

慢开始门限的初始值设置为 16 个报文段，即 $ssthresh = 16$ 。

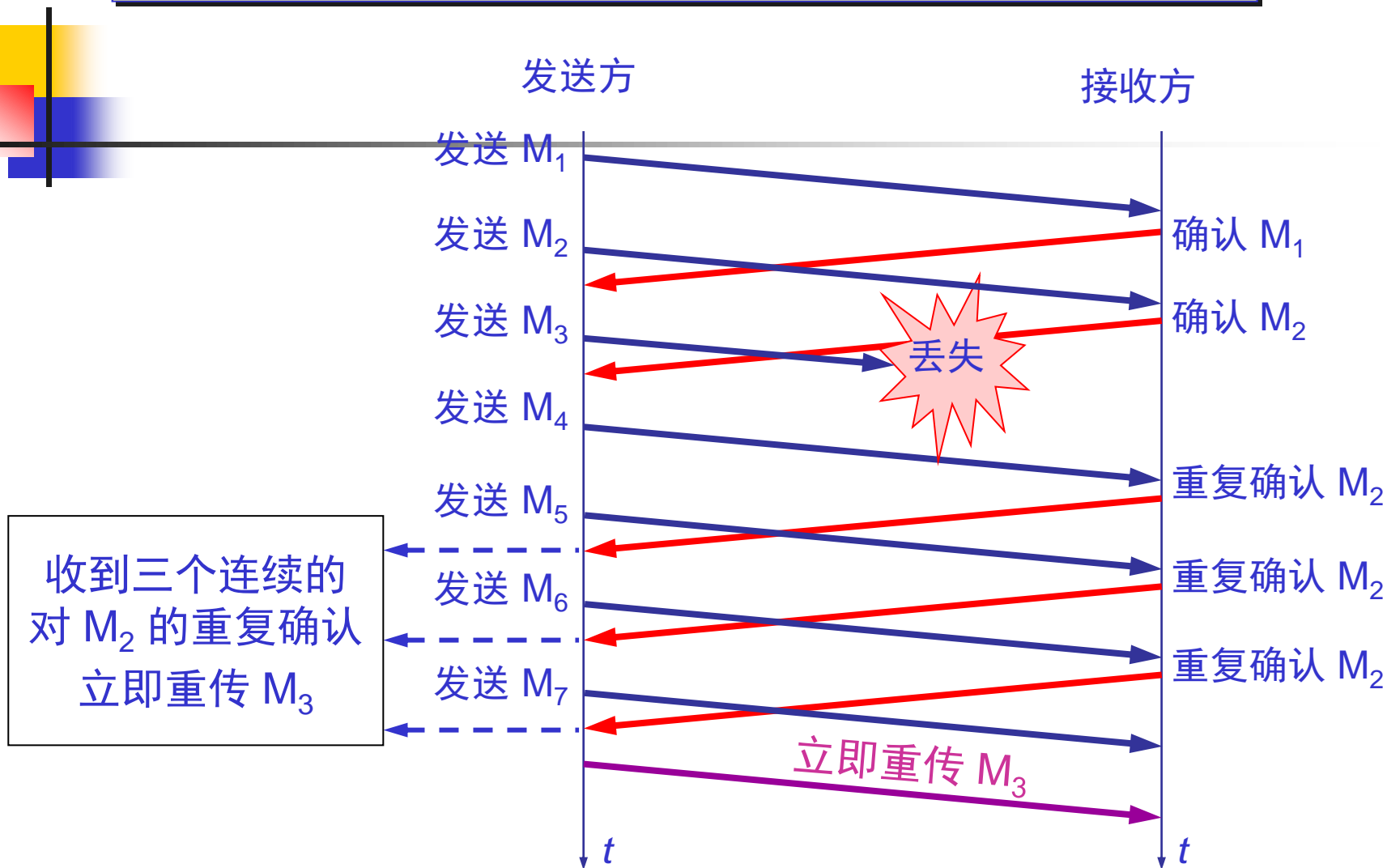
不要背,会做书后5-39题



2. 快重传和快恢复

- 快重传算法首先要求接收方每收到一个失序的报文段后就立即发出重复确认。这样做可以让发送方及早知道有报文段没有到达接收方。
- 发送方只要一连收到三个重复确认就应当立即重传对方尚未收到的报文段。
- 不难看出，快重传并非取消重传计时器，而是在某些情况下可更早地重传丢失的报文段。

快重传举例

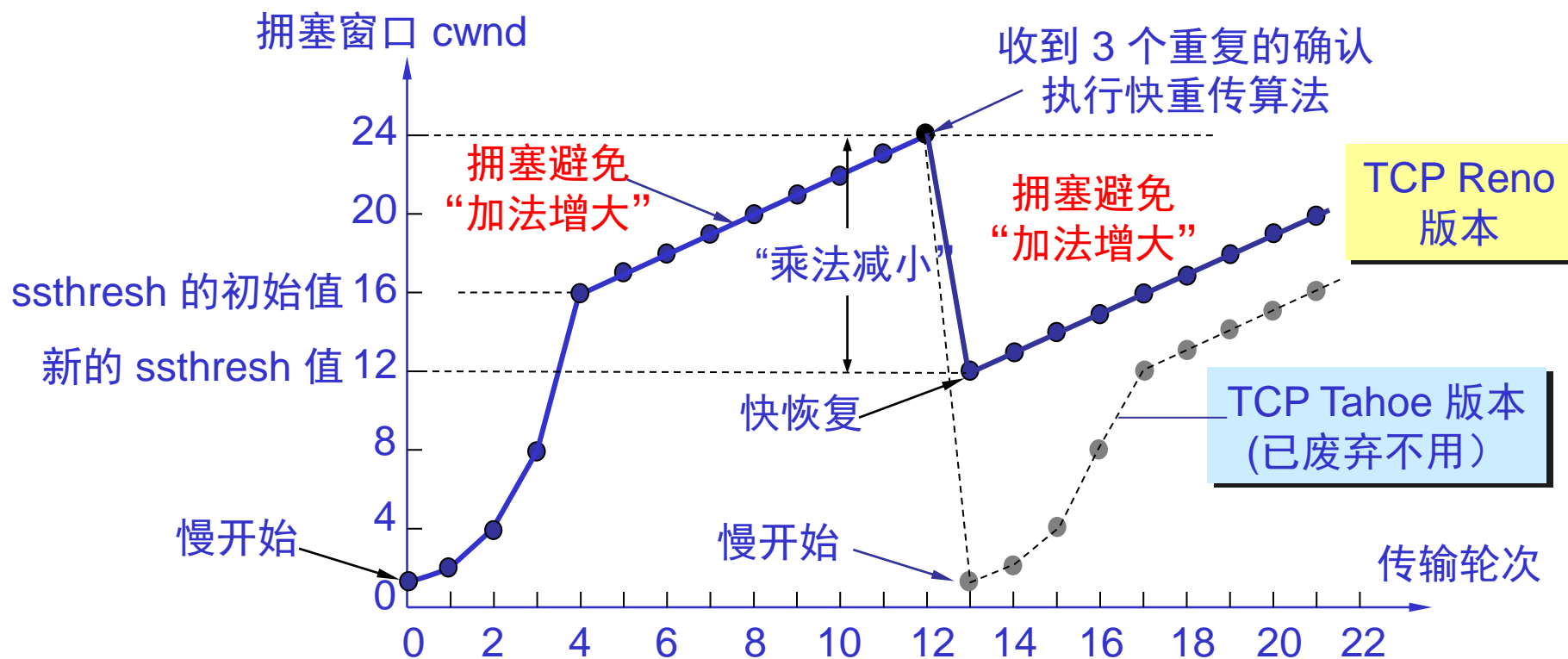




快恢复算法

- (1) 当发送端收到连续三个重复的确认时，就执行“乘法减小”算法，把慢开始门限 $ssthresh$ 减半。但接下去不执行慢开始算法。
- (2) 由于发送方现在认为网络很可能没有发生拥塞，因此现在不执行慢开始算法，即拥塞窗口 $cwnd$ 现在不设置为 1，而是设置为慢开始门限 $ssthresh$ 减半后的数值，然后开始执行拥塞避免算法（“加法增大”），使拥塞窗口缓慢地线性增大。

从连续收到三个重复的确认 转入拥塞避免





发送窗口的上限值

- 发送方的发送窗口的上限值应当取为接收方窗口 $rwnd$ 和拥塞窗口 $cwnd$ 这两个变量中较小的一个，即应按以下公式确定：

$$\text{发送窗口的上限值} = \text{Min} [rwnd, cwnd] \quad (5-8)$$

- 当 $rwnd < cwnd$ 时，是接收方的接收能力限制发送窗口的最大值。
- 当 $cwnd < rwnd$ 时，则是网络的拥塞限制发送窗口的最大值。



5.8.3 主动队列管理 AQM

- 所谓“主动”就是不要等到路由器的队列长度已经达到最大值时才不得不丢弃后面到达的分组。这样就太被动了。应当在队列长度达到某个值得警惕的数值时（即当网络拥塞有了某些拥塞征兆时），就**主动丢弃**到达的分组。

5-9 TCP 的运输连接管理

1. 运输连接的三个阶段

- 运输连接就有三个阶段，即：**连接建立**、**数据传送**和**连接释放**。运输连接的管理就是使运输连接的建立和释放都能正常地进行。
- 连接建立过程中要解决以下三个问题：
 - 要使每一方能够确知对方的存在。
 - 要允许双方协商一些参数（如最大报文段长度，最大窗口大小，服务质量等）。
 - 能够对运输实体资源（如缓存大小，连接表中的项目等）进行分配。

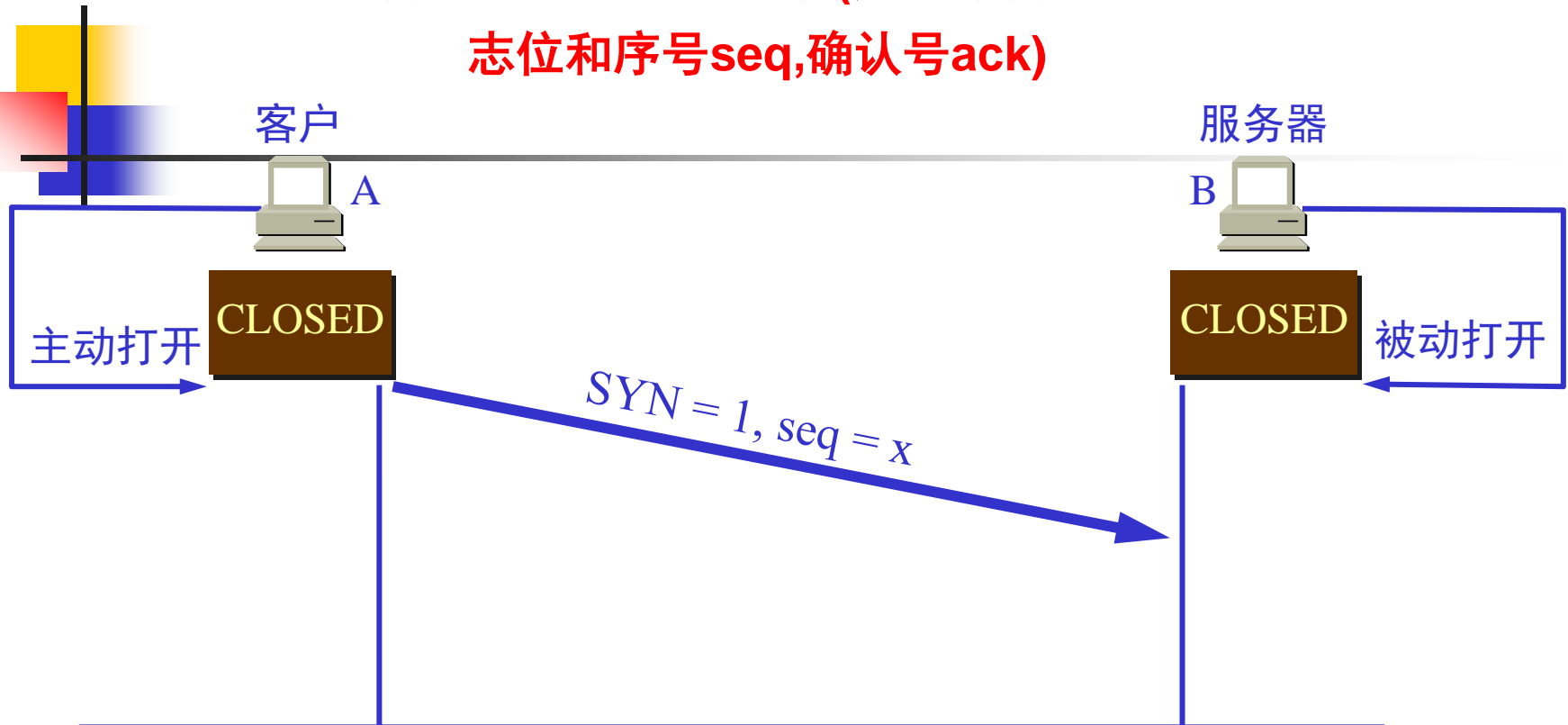


客户-服务器方式

- TCP 连接的建立都是采用客户服务器方式。
- 主动发起连接建立的应用进程叫做**客户**(client)。
- 被动等待连接建立的应用进程叫做**服务器**(server)。

5.9.1 TCP 的连接建立

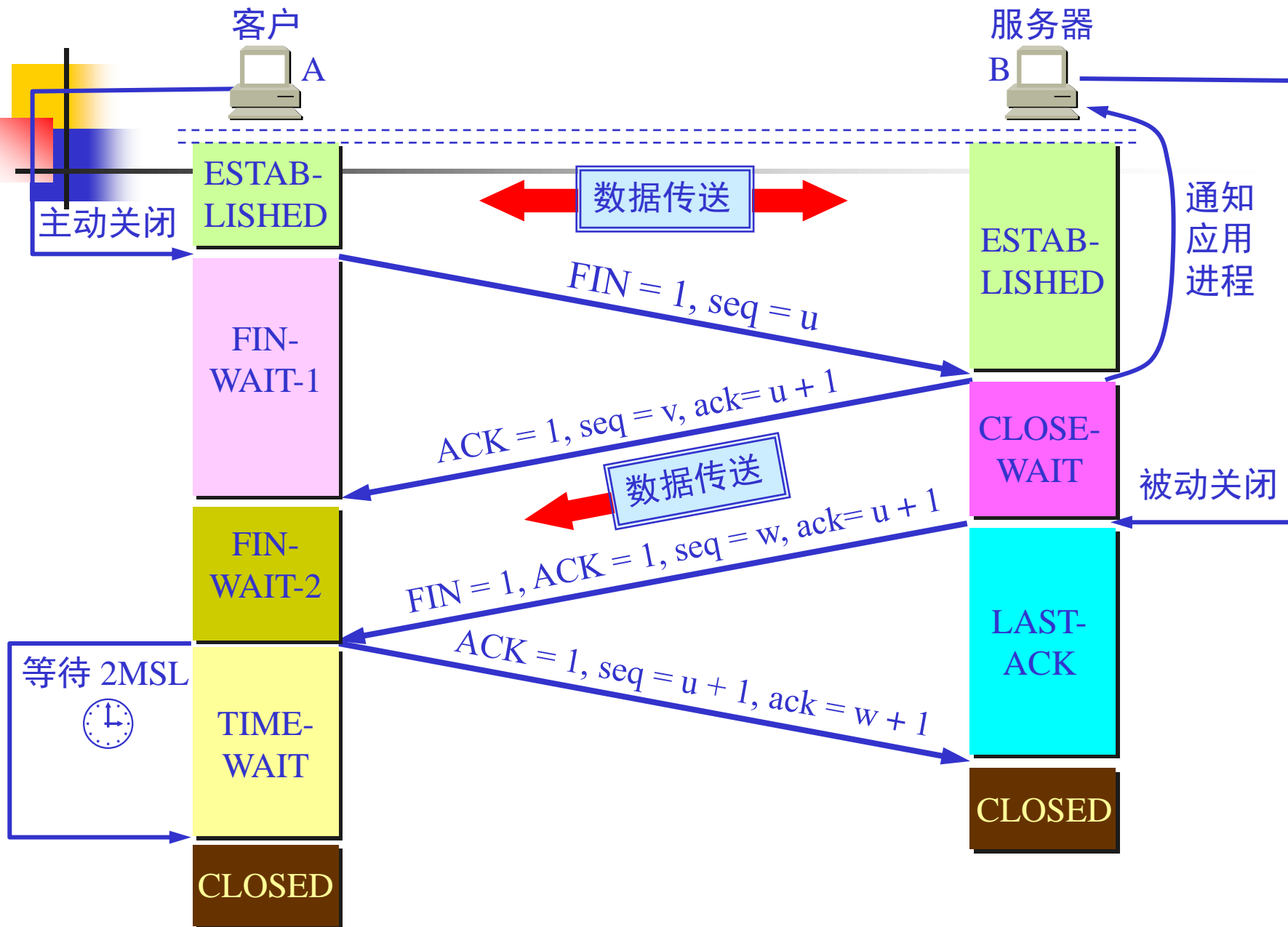
用三次握手建立 TCP 连接(注意每次握手的标志位和序号seq,确认号ack)

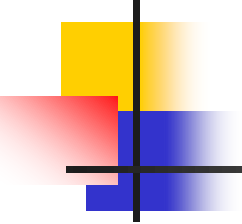


A 的 TCP 向 B 发出连接请求报文段，其首部中的同步位 $SYN = 1$ ，并选择序号 $seq = x$ ，表明传送数据时的第一个数据字节的序号是 x 。

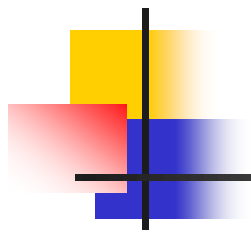
总结：SYN=1，ACK=0 连接的请求，即连接的第一个报文

TCP 连接必须经过时间 2MSL 后才真正释放掉。





以上拥塞控制，三次握手建立连接，四次握手断开连接注意每次的标志位，序号和确认号不要背，会做题



第六章 应用层



第 6 章 应用层

- 6.1 域名系统 DNS
- 6.2 文件传送协议FTP
- 6.3 远程终端协议 TELNET
- 6.4 万维网 WWW (HTTP)
- 6.5 电子邮件(SMTP POP3)
- 6.6 动态主机配置协议 DHCP
- 6.7 简单网络管理协议 SNMP



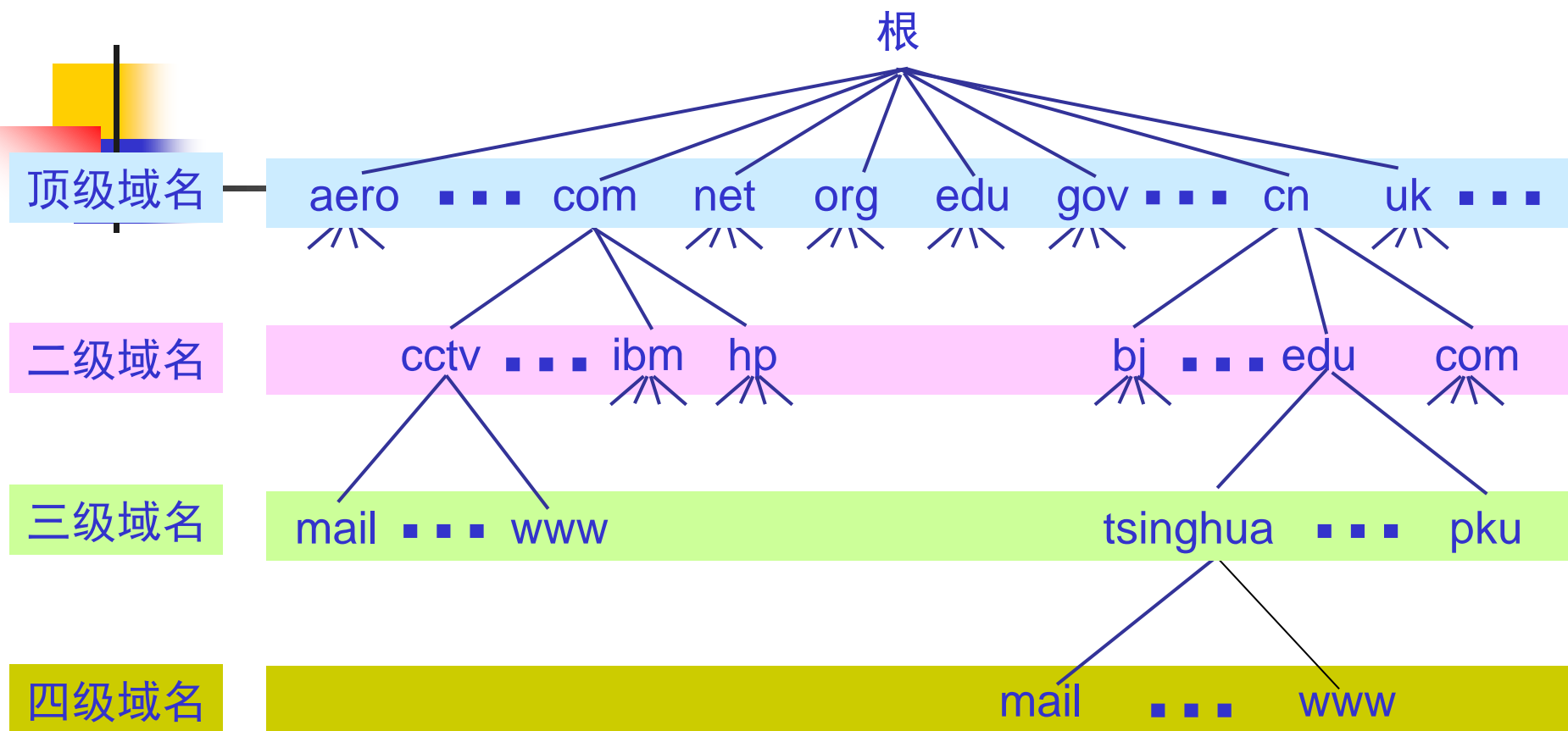
6.1.2 互联网的域名结构

- 互联网采用了层次树状结构的命名方法。
- 任何一个连接在互联网上的主机或路由器，都有一个**唯一**的层次结构的**名字**，即**域名**。
- 域名的结构由标号序列组成，各标号之间用**点**隔开：

... . 三级域名 . 二级域名 . 顶级域名

- 各标号分别代表不同级别的域名。

因特网的域名空间





域名服务器有以下四种类型

- 根域名服务器是最高层次的域名服务器，也是最重要的域名服务器。所有的根域名服务器都知道所有的顶级域名服务器的域名和 IP 地址。
- 顶级域名服务器（即 TLD 服务器）负责管理在该顶级域名服务器注册的所有二级域名。
- 权限域名服务器负责一个区的域名服务器。
- 本地域名服务器：缓存已经访问过的域名对应的 IP

6.2 文件传送协议

6.2.1 FTP概述

- 文件传送协议 FTP (File Transfer Protocol) 是因特网上使用得最广泛的文件传送协议。
- 传输层使用TCP协议，是一种可靠传输协议,因为需要保证文件的传输的可靠,完整,无差错.



万维网必须解决的问题

(1) 怎样标志分布在整个因特网上的万维网文档？

- 使用统一资源定位符 URL (Uniform Resource Locator) 来标志万维网上的各种文档。
- 使每一个文档在整个因特网的范围内具有唯一的标识符 URL。



URL 的一般形式

- 由以冒号隔开的两大部分组成，并且在 URL 中的字符对大写或小写没有要求。
- URL 的一般形式是：

<协议>://<主机>:<端口>/<路径>

ftp —— 文件传送协议 FTP

http —— 超文本传送协议 HTTP

News —— USENET 新闻



万维网必须解决的问题

(2) 用何协议实现万维网上各种超链的链接？

- 在万维网客户程序与万维网服务器程序之间进行交互所使用的协议，是超文本传送协议 HTTP (HyperText Transfer Protocol)。
- HTTP 是一个应用层协议，它使用 TCP 连接进行可靠的传送。



3. HTTP 的报文结构

HTTP 有两类报文：

- **请求报文**——从客户向服务器发送请求报文。
- **响应报文**——从服务器到客户的回答。
- 由于 HTTP 是面向正文的 (text-oriented), 因此在报文中的每一个字段都是一些 **ASCII 码串**, 因而每个字段的长度都是不确定的。

HTTP 的报文结构（请求报文）

开始行

方 法 URL 版 本 CRLF 请求行

首部字段名 : 值 CRLF

⋮

首部字段名 : 值 CRLF

CRLF

} 首部行

实体主体
(通常不用)

报文由三个部分组成，即**开始行**、**首部行**和**实体主体**。
在请求报文中，开始行就是请求行。



HTTP 请求报文的一些方法

方法（操作）

意义

OPTION

请求一些选项的信息

GET

请求读取由 URL 所标志的信息

HEAD

请求读取由 URL 所标志的信息的

首部

POST

给服务器添加信息（例如，注释）

PUT

在指明的 URL 下存储一个文档

DELETE

删除指明的 URL 所标志的资源

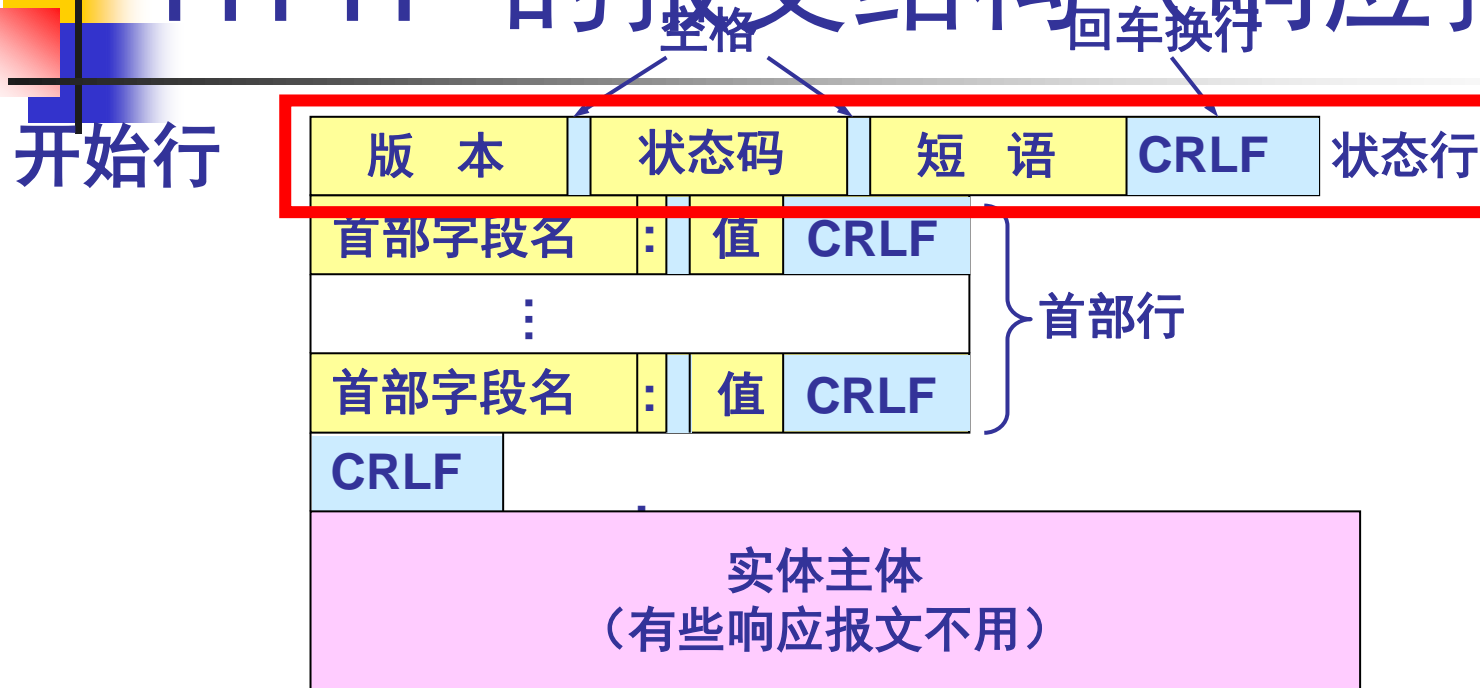
TRACE

用来进行环回测试的请求报文

CONNECT

用于代理服务器

HTTP 的报文结构（响应报文）



响应报文的开始行是**状态行**。

状态行包括三项内容，即 **HTTP** 的版本，**状态码**，以及解释状态码的**简单短语**。



状态码都是三位数字

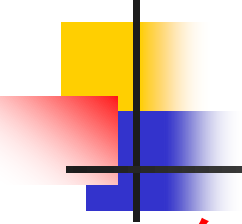
- 1xx 表示通知信息的，如请求收到了或正在进行处理。
- 2xx 表示成功，如接受或知道了。
- 3xx 表示重定向，表示要完成请求还必须采取进一步的行动。
- 4xx 表示客户的差错，如请求中有错误的语法或不能完成。
- 5xx 表示服务器的差错，如服务器失效无法完成请求。



其他应用协议

- 发送邮件的协议：SMTP(下层为TCP)
- 读取邮件的协议：POP3 和 IMAP
- 动态主机配置协议 DHCP(IPv4地址配置)
- 简单网络管理协议SNMP 使用无连接的UDP

自己总结下应用层协议，那些传输时候用TCP，哪些用UDP

- 
-
- 每层的数据单位：
 - 物理层： bit
 - 数据链路层： 帧
 - 网络层： 分组（报文）
 - 传输层： TCP报文段， UDP数据报

- 
-
- 每层的ID
 - 物理层：无
 - 数据链路层：MAC地址
 - 网络层：IP地址
 - 传输层：PORT端口



最后：

- 所有的作业都是重点！
- 所有课上做过的题都是重点！ 特别是大题！
- 1-17, 2-16, 3-07, 3-20, 3-27, 3-33, 4-10, 4-20, 4-22, 4-37, 4-41, 4-55, 5-39



END, 祝好运!