网络技术基础

高智刚

M.P. & WeChat：13572460159

E-mail: gaozhigang@nwpu.edu.cn

第五章：网络层

本章主要内容

- 网络层基本概述
- 网际协议（IP）
- 地址解析协议（ARP）
- 因特网控制报文协议（ICMP）
- 路由选择协议（RP）
- 因特网管理协议（IGMP）
- 最新网际协议（IPv6）

5.1 网络层基本概述

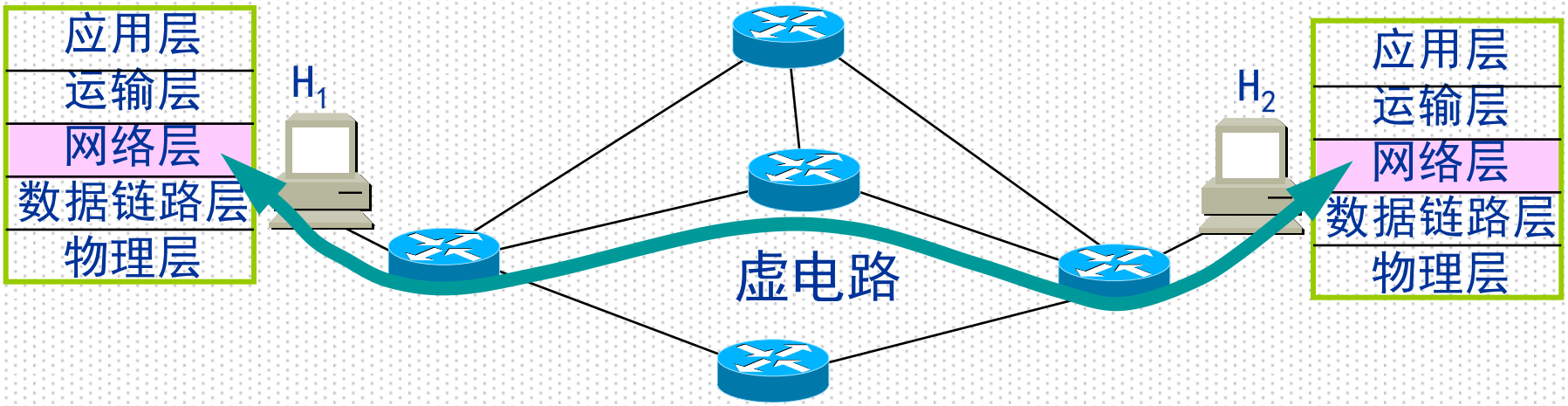
- 网络层提供的两种服务
  - 在计算机网络领域，网络层应该向传输层提供怎样的服务（“面向连接”还是“无连接”）曾引起了长期的争论
  - 争论焦点的实质就是：在计算机通信中，可靠交付应当由谁来负责？是网络还是端系统？

5.1 网络层基本概述

- **电信网的经验：让网络负责可靠交付**
  - 面向连接的通信方式
  - 建立虚电路(Virtual Circuit)，以保证双方通信所需的一切网络资源
  - 如果再使用可靠传输的网络协议，就可使所发送的分组无差错按序到达终点

5.1 网络层基本概述

- 虚电路服务



H_1 发送给 H_2 的所有分组都沿着同一条虚电路传送

- 虚电路是逻辑连接
  - 虚电路表示这只是一条逻辑上的连接，分组都沿着这条逻辑连接按照存储转发方式传送，而并不是真正建立了一条物理连接
  - 请注意，电路交换的电话通信是先建立了一条真正的连接。因此分组交换的虚连接和电路交换的连接只是类似，但并不完全一样

- 因特网采用的设计思路
  - 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
    - 网络在发送分组时不需要先建立连接。每一个分组（即 IP 数据报）独立发送，与其前后分组无关（不编号）
    - 不提供完善的可靠性措施，仅对数据报报头进行检验，数据不检验，出错简单丢弃
    - 尽力发送每个分组，只有资源用尽或底层网络故障时才丢弃分组
    - 网络层不提供服务质量的承诺。即所传送的分组可能出错、丢失、重复和失序（不按序到达终点），也不保证分组传送的时限

5.1 网络层基本概述

- 尽最大努力交付的好处
  - 由于传输网络不提供端到端的可靠传输服务，使网络中的路由器可以比较简单，且价格低廉
  - 如果主机（即端系统）中的进程之间的通信需要是可靠的，那么就由网络的主机中的传输层负责（包括差错处理、流量控制等）
  - 采用这种设计的好处：网络造价大大降低，运行方式灵活，能够适应多种应用
  - 因特网能够发展到今日的规模，充分证明了当初采用这种设计思路的正确性

5.1 网络层基本概述

- 网络核心功能
  - 实现网络互联
  - 网络层承担了实现网络互联的基础和核心任务，它提供的跨网络分组传输服务是Internet所有应用赖以存在的基础

5.1 网络层基本概述

- 已有的网络互联方式
  - 物理层：中继器、集线器
  - 数据链路层：网桥、交换机
  - 网络层：路由器（通过转换不同的分组格式）
- 数据链路层与网络层的区别
  - 数据链路层：利用帧的MAC地址
  - 网络层：利用数据报中分组内的网络层地址（如IP地址）决定输出线路

5.2 数据交换技术

- **什么是交换？**
  - 由中间节点进行转接的通信称为交换
    - 例如，电话交换机在用户呼叫时为用户选择一条可用线路进行接续；用户挂机后则断开线路，该线路又可分配给其它用户
    - 最初的交换：人工转接交换
- **为什么要采用交换技术？**
  - 节省线路投资，提高线路利用率
- **实现交换的方法主要有：电路交换、报文交换和分组交换**

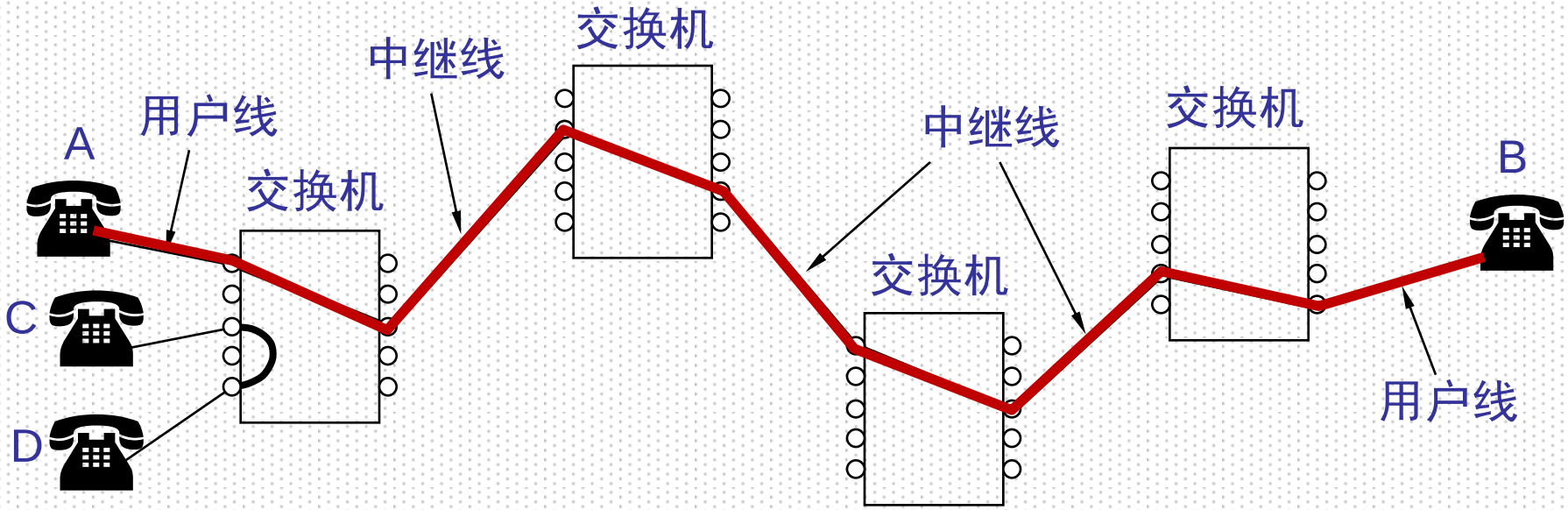
5.2.1 电路交换

- 在通信双方之间建立一条临时专用线路的过程
  - 可以是真正的物理线路，也可是一个复用信道
- 特点：数据传输前需要建立一条端到端的通路

  ——称为“面向连接的”（典型例子：电话）
- 过程：建立连接→通信→释放连接
- 优缺点：
  - 建立连接的时间长
  - 一旦建立连接就独占线路，线路利用率低
  - 无纠错机制
  - 建立连接后，传输延迟小

5.2.1 电路交换

- 电话网络中的电路交换



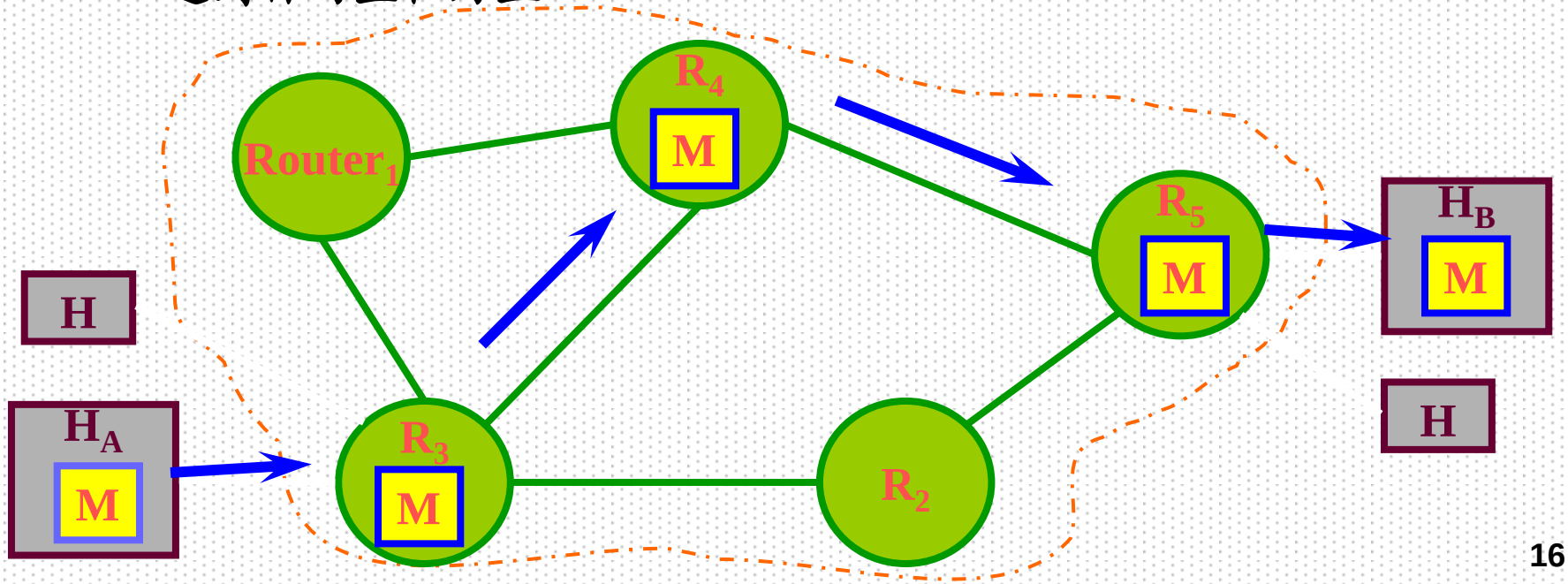
- 电路交换也能在多路复用信道上实现
  - 在物理线路的某个信道上建立连接

5.2.2 报文交换

- 以报文为单位进行“存储-转发”的交换技术
  - 在交换过程中，交换设备将接收到的报文先存储，待信道空闲时再转发出去，一级一级中转，直到目的地。这种数据传输技术称为存储-转发
- 传输之前不需要建立端到端的连接，仅在相邻结点传输报文时建立 — 称为“无连接的”
- 整个报文（Message）作为一个整体一起发送
- 优缺点：
  - 没有建立和拆除连接所需的等待时间
  - 线路利用率高
  - 传输可靠性较高
  - 报文大小不一，造成存储管理复杂
  - 大报文造成存储转发的延时过长，且对存储容量要求较高
  - 出错后整个报文全部重发

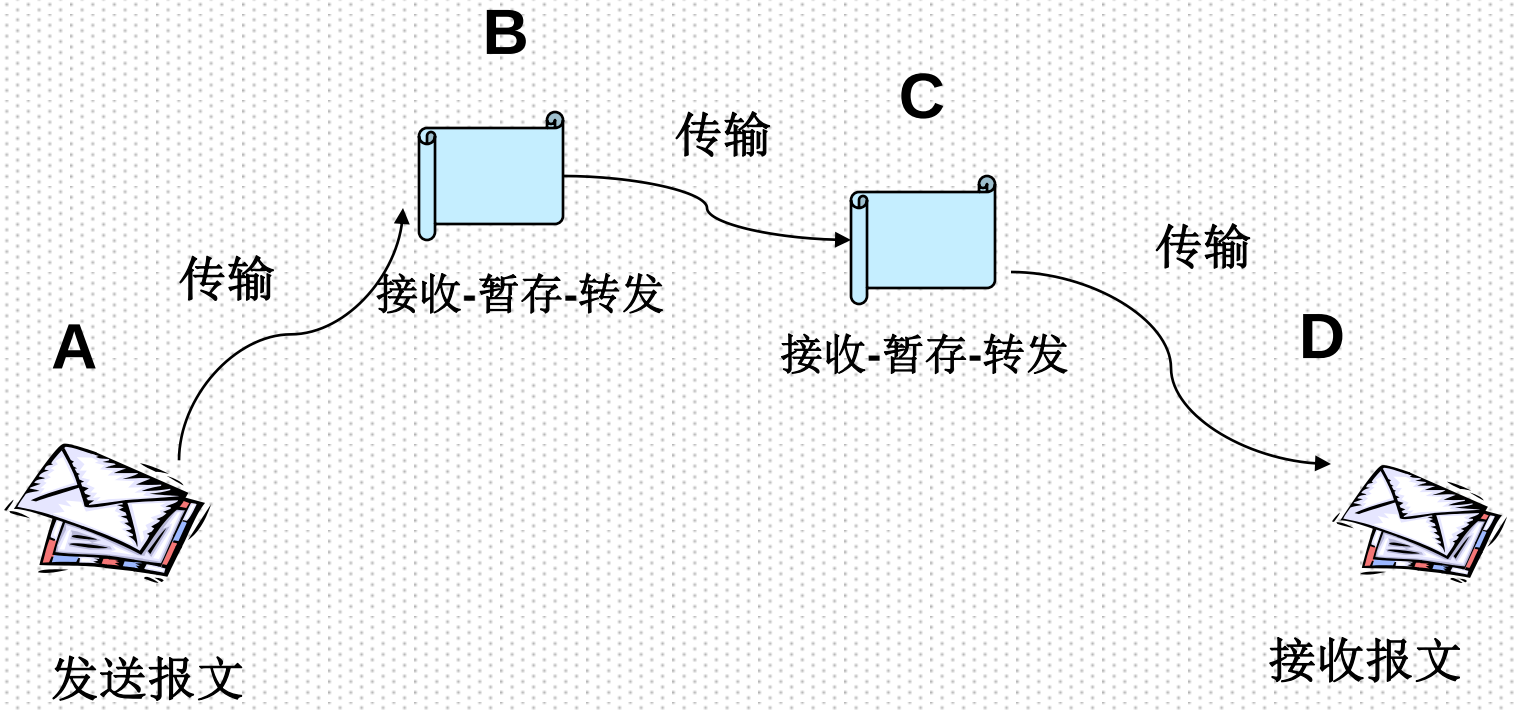
5.2.2 报文交换

- **存储转发** (Store and Forward)
  - 发送报文的主机在发送之前，要将报文的目的地址附加在报文前面。然后将报文发送到网络中的结点中
  - 每个网络中的结点将完整地接收报文，暂存报文，然后将报文发送到下一个更接近目的主机的结点中。如此操作，直至将报文发送到目的主机为止



5.2.2 报文交换

- 报文“存储-转发”过程

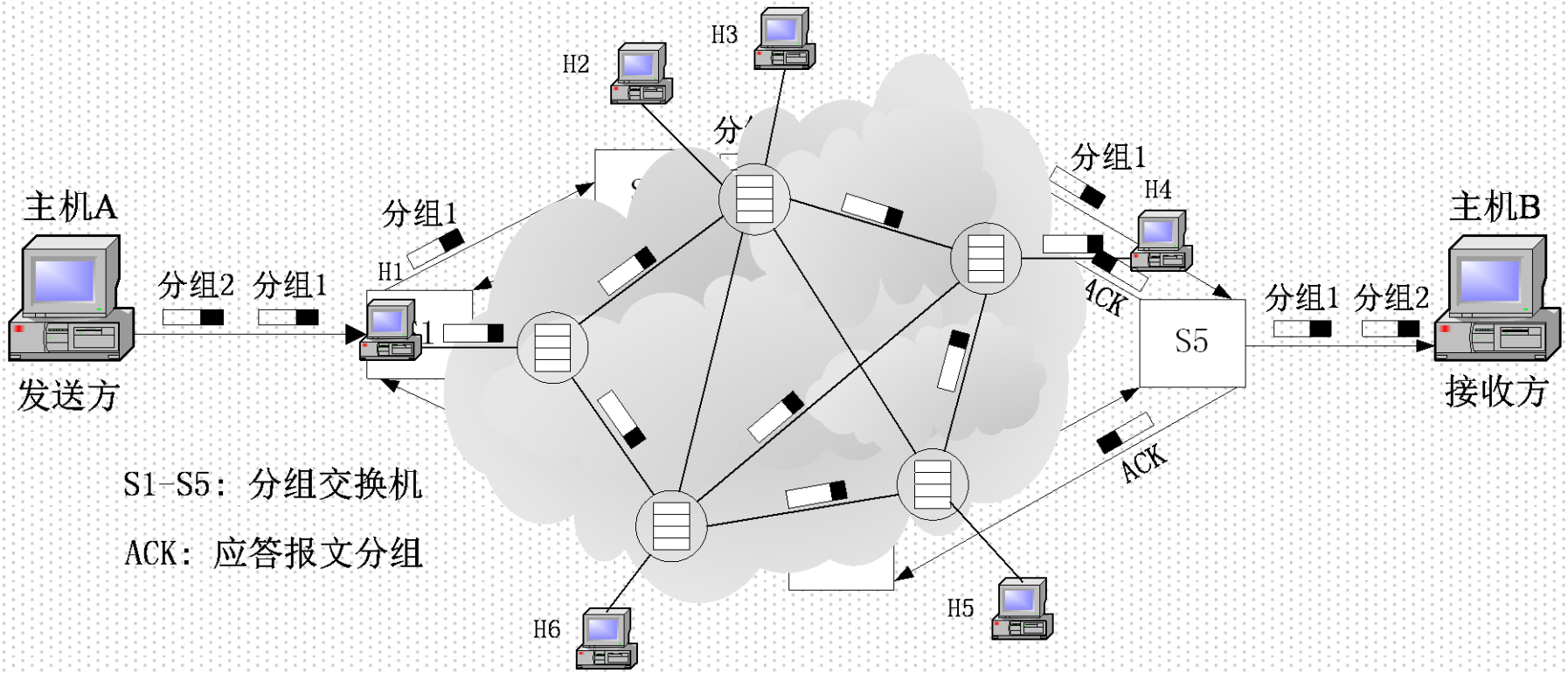


5.2.3 分组交换

- **分组交换**
  - 将报文分割成若干个大小相等的分组（Packet）进行存储转发
  - 数据传输前不需要建立一条端到端的物理通路
  - 有强大的纠错机制、流量控制、拥塞控制和路由选择功能
- **优缺点**
  - 对转发结点的存储要求较低，可以用内存来缓冲分组——速度快
  - 转发延时小——适用于交互式通信
  - 某个分组出错可以仅重发出错的分组——效率高
  - 各分组可通过不同路径传输，容错性好
  - 需要分割报文和重组报文，增加了端站点的负担
- **分组交换有两种交换方式**
  - 数据报方式和虚电路方式

5.2.3 分组交换

- **数据报方式**（Datagram）
  - 各分组**独立地确定路由**（传输路径）
  - **不能保证分组按序到达**，所以目的站点需要按分组编号重新排序和组装



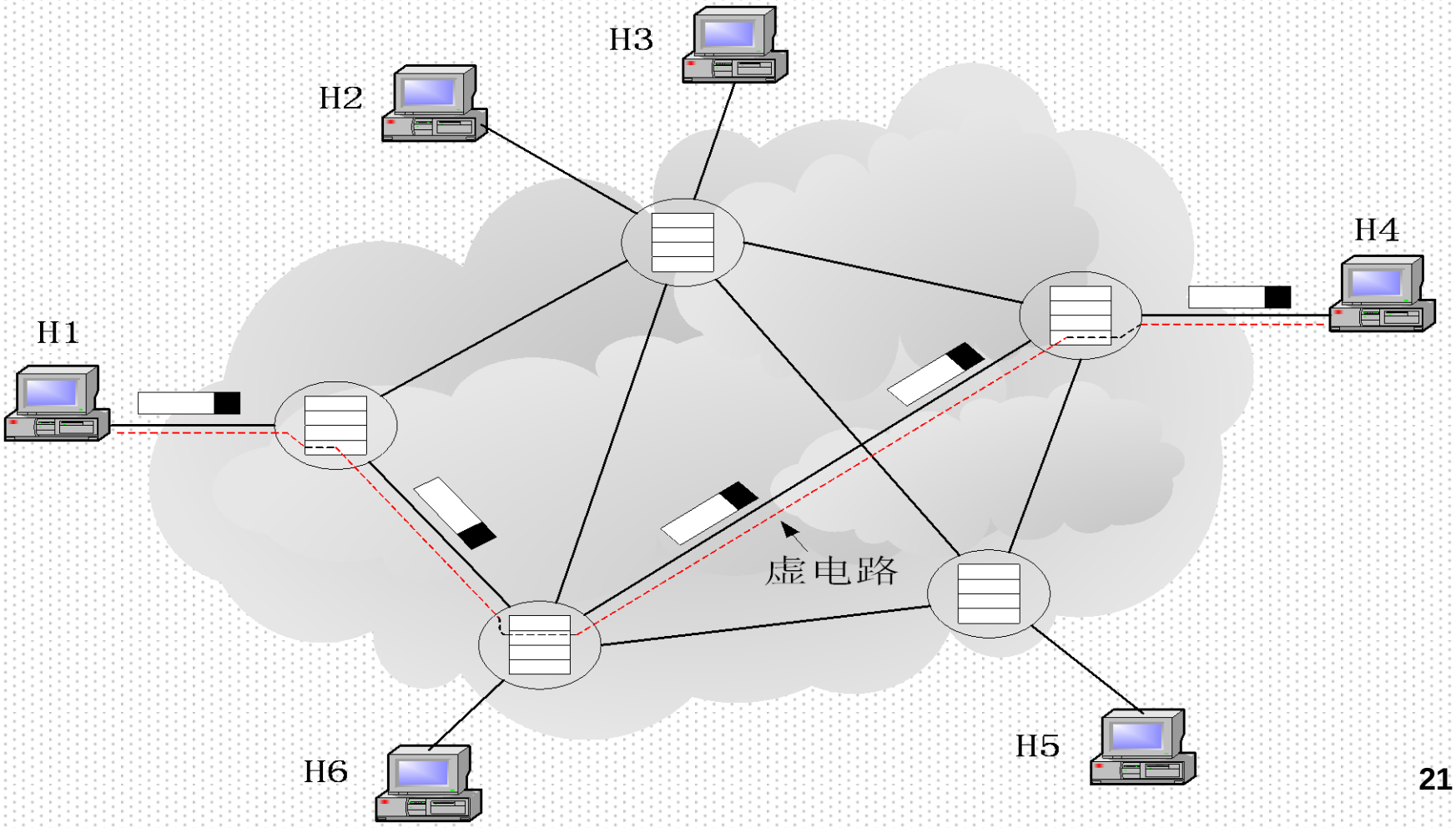
数据报方式不能保证分组按序到达
分组可能通过多个路径穿越网络

5.2.3 分组交换

- **虚电路方式** (Virtual Circuit)
  - 通信前预先建立一条逻辑连接——虚电路
    - 虚电路是由其路径上的所有交换机中的路由表定义的
  - 也需要三个过程：建立－数据传输－拆除
    - 建立虚电路时，交换机将预留传输时所需的所有资源
  - 虚电路的路由在建立时确定，传输数据时则不再需要
    - 数据传输时只需指定虚电路号，分组即可按虚电路的路由穿越网络——“数字管道”
  - 提供的是“面向连接”的服务
    - 但却没有像电路交换那样始终占用一条端到端的物理通道，只是断续地依次占用传输路径上各个链路段
    - 可以看成是采用了电路交换思想的分组交换
    - 能够保证分组按序到达
  - 永久虚电路PVC和交换虚电路SVC

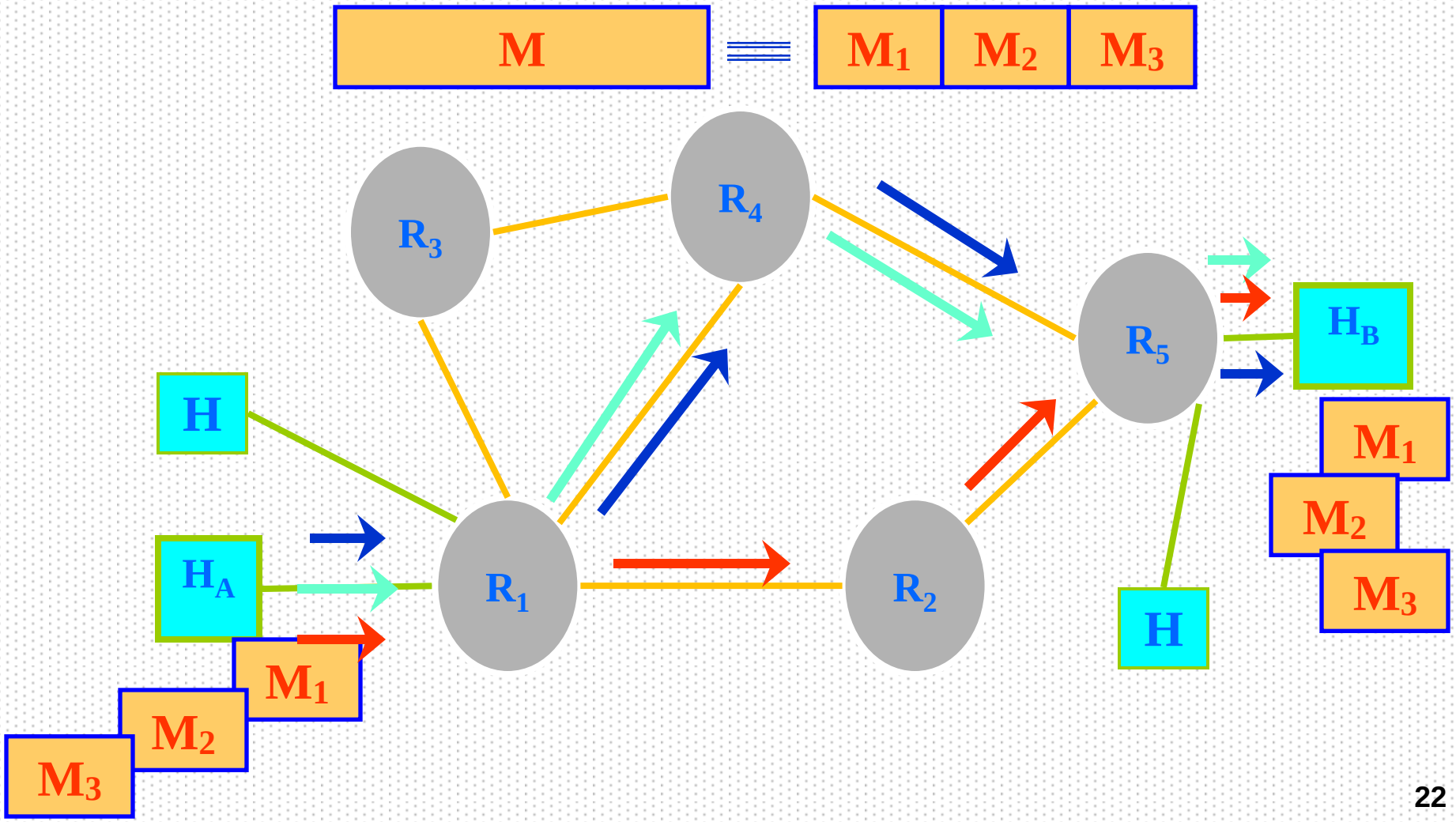
5.2.3 分组交换

- 分组通过预先建立好的虚电路穿越网络

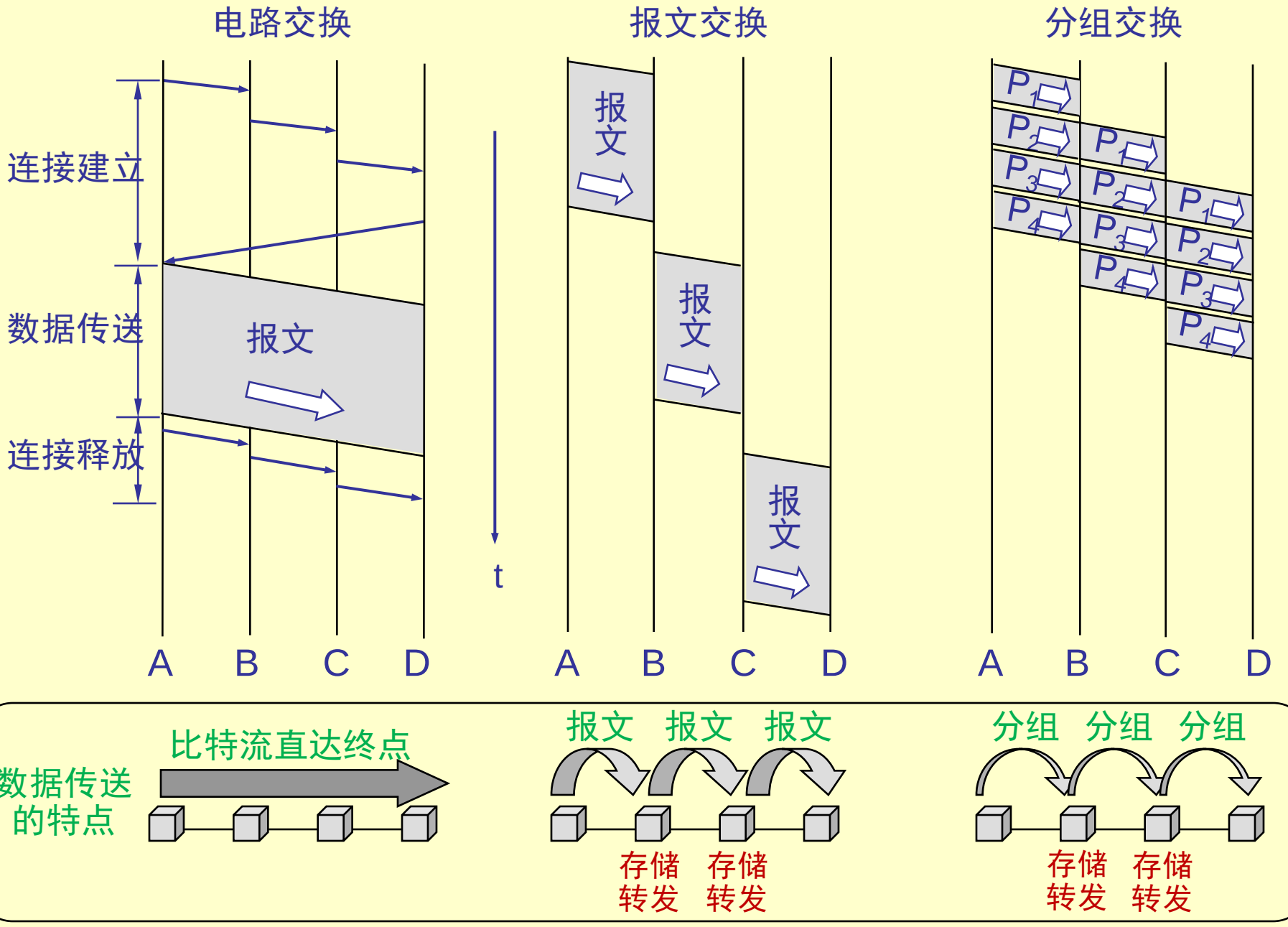


5.2.3 分组交换

- 【例】请判断是虚电路还是数据报？



三种数据交换方法


电路交换
报文交换
分组交换
连接建立
数据传送
连接释放
报文
报文
报文
报文
t
A
B
C
D
A
B
C
D
A
B
C
D
P1
P2
P3
P4
P1
P2
P3
P4
P1
P2
P3
P4
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

5.2.4 三种数据交换技术总结

(1) 电路交换

- 在数据传送之前需建立一条物理通路，在线路被释放之前，该通路将一直被用户完全占有

(2) 报文交换

- 报文从发送方传送到接收方采用存储转发方式

(3) 分组交换

- 与报文交换类似，但报文被分成组传送，并规定了分组的最大长度，到达目的地后需重新将分组组装成报文

5.3 路由器-Router

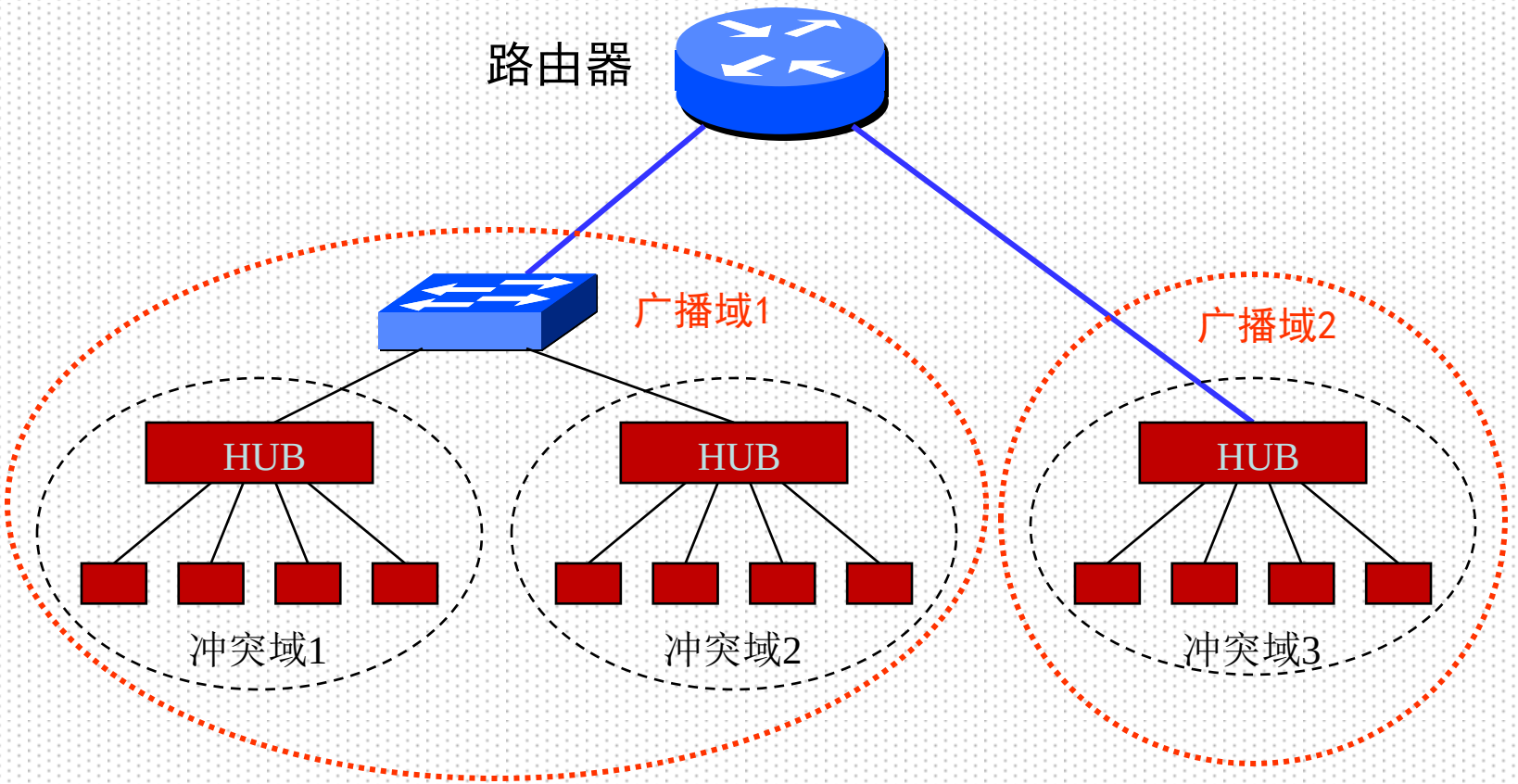
- 工作在OSI第三层（网络层）
- 功能
  - 在网络之间转发网络分组
  - 为网络分组寻找最佳传输路径
  - 实现子网隔离，限制广播风暴(目的地址无法识别时，路由器将其丢弃，而不是广播——*比较网络交换机*)
  - 提供逻辑地址，以识别互联网上的主机
  - 提供广域网服务

5.3 路由器-Router

- 应用
  - 把LAN连入广域网或作为广域网的核心连接设备
- 路由器与交换机的主要区别
  - 用路由器连接起来的若干个网络，它们仍是各自独立的。要想从一个网络访问用路由器连接起来的另一个网络中的站点，必须指定该站点的逻辑地址（IP地址），通过广播是无法与之进行通信

5.3 路由器-Router

- 用路由器进行网络互联



路由器可以隔离冲突域和广播域

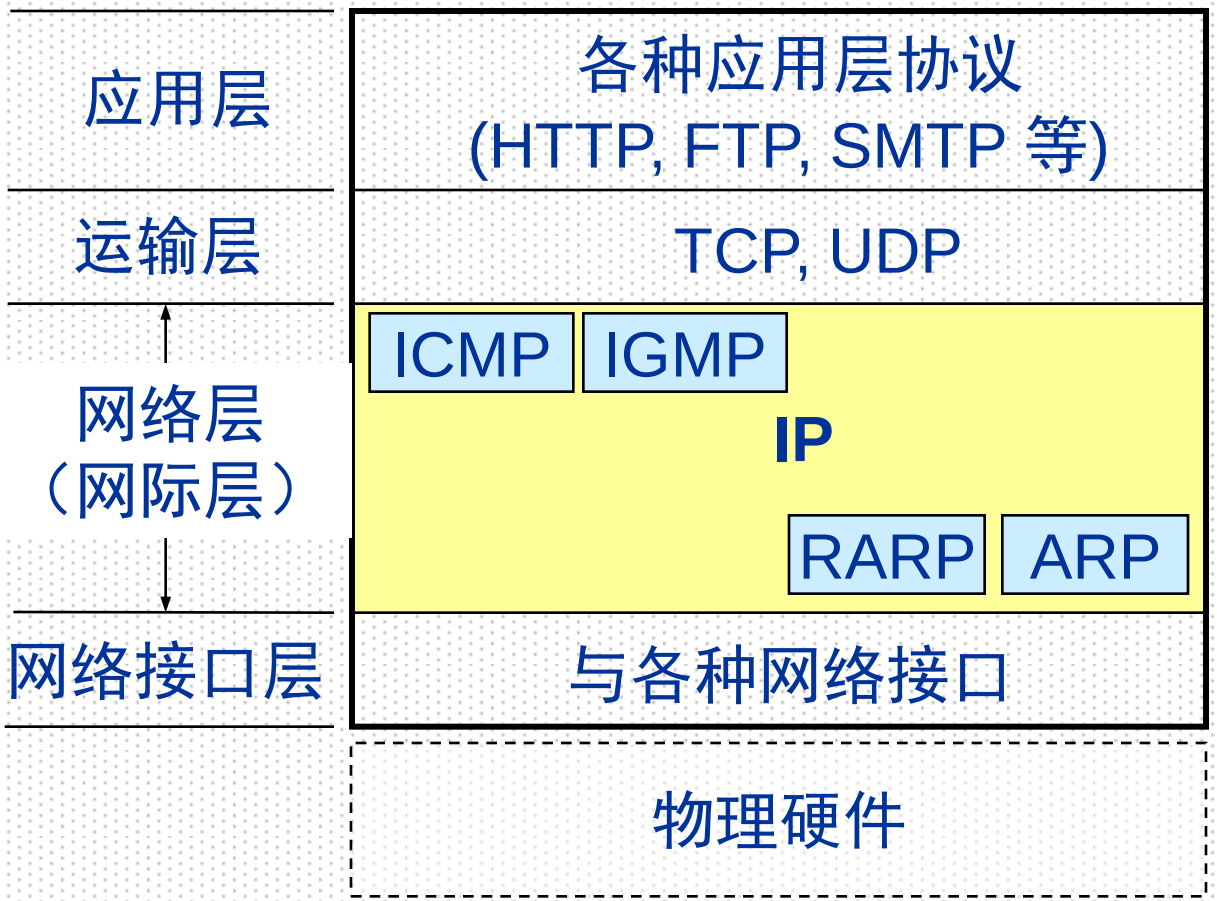
5.4 网络层特点及相关协议

- 网络层提供的服务和特点
  - 主要负责将分组数据报从源主机传送到目的主机，并提供无连接、不可靠但尽力而为的分组传输服务
    - 无连接：传输之前，通信双方并不建立连接，使用分组交换方式，每个分组都被独立地进行转发，可能不按顺序到达
    - 不可靠：只是对数据报报头进行检验，而数据不检验，报头有差错就直接丢弃
    - 尽力而为：尽力发送每个分组，并不轻易放弃，只有当资源用尽或底层网络出现故障时才会放弃分组

- 网络层的协议
  - 网络层最重要的协议就是网际协议即IP协议，IPv4已转向IPv6
    （1）IP协议定义了网络层的数据报格式
    （2）IP软件实现数据转发功能，选择数据报发送的路由并转发
    （3）包括了一组体现了不可靠、尽力分组传送特点的规则

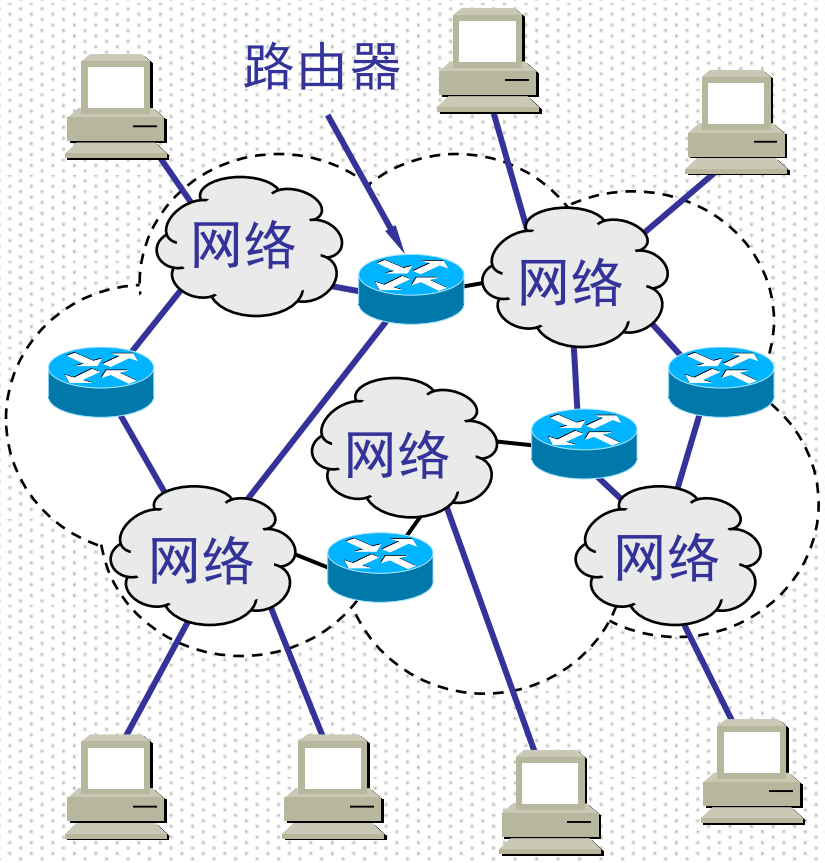
- 网络层的协议
  - 另外，与IP配套使用的网络层协议还有
    - 地址解析协议 **ARP**
      (Address Resolution Protocol)
    - 逆地址解析协议 **RARP**
      (Reverse Address Resolution Protocol)
    - 网际控制报文协议 **ICMP**
      (Internet Control Message Protocol)
    - 网际组管理协议 **IGMP**
      (Internet Group Management Protocol)

- 网络层的IP协议及配套协议



5.4.1 网络层特点

- 互连网络与虚拟互连网络



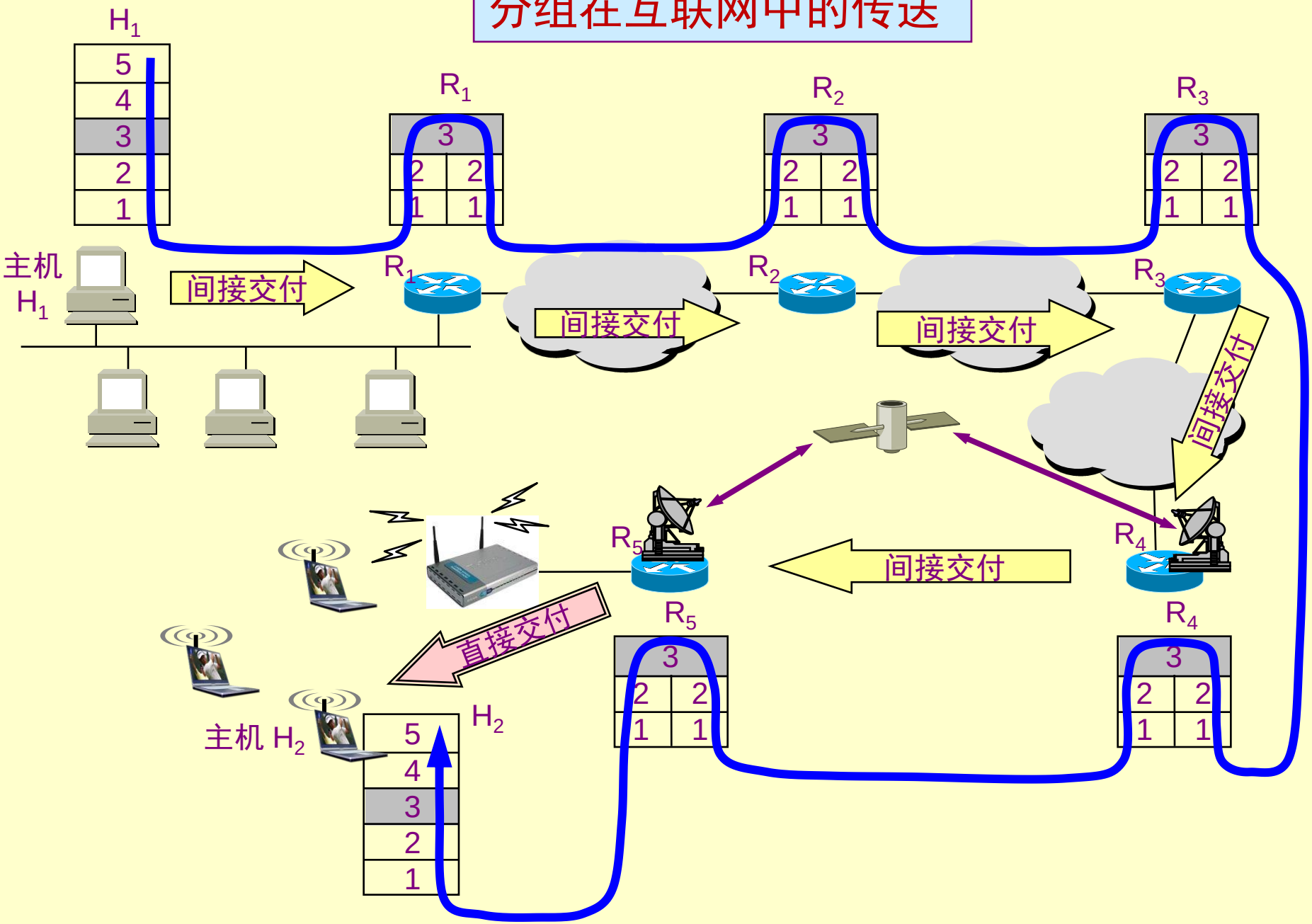
(a) 互连网络

(b) 虚拟互连网络

5.4.1 网络层特点

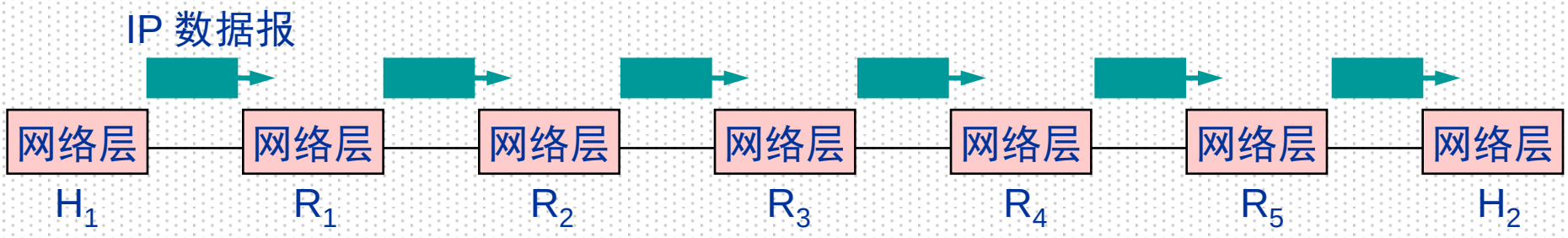
- 虚拟互连网络的意义
  - 所谓虚拟互连网络也就是逻辑互连网络，它的意思就是互连起来的各种物理网络的异构性本来是客观存在的，但是我们利用IP协议就可以使这些性能各异的网络从用户看起来好像是一个统一的网络
  - 使用IP协议的虚拟互连网络可简称为IP网
  - 使用虚拟互连网络的好处是：当互联网上的主机进行通信时，就好像在一个网络上通信一样，而看不见互连的各具体的网络异构细节

分组在互联网中的传送


H1
5
4
3
2
1
R1
3
2 2
1 1
R2
3
2 2
1 1
R3
3
2 2
1 1
主机
H1
间接交付
R1
R2
R3
间接交付
间接交付
间接交付
R5
R4
间接交付
R5
R4
3
2 2
1 1
3
2 2
1 1
直接交付
主机 H2
H2
5
4
3
2
1

5.4.1 网络层特点

- 从网络层看IP数据报的传送
  - 如果我们只从网络层考虑问题，那么 IP 数据报就可以想象是在网络层中传送



5.4.2 网际协议(IP)

- IP地址及其表示方法
  - 我们把整个因特网看成为一个单一的、抽象的网络。IP地址就是给每个连接在因特网上的主机（或路由器）分配一个在全世界范围内唯一的32位标识符

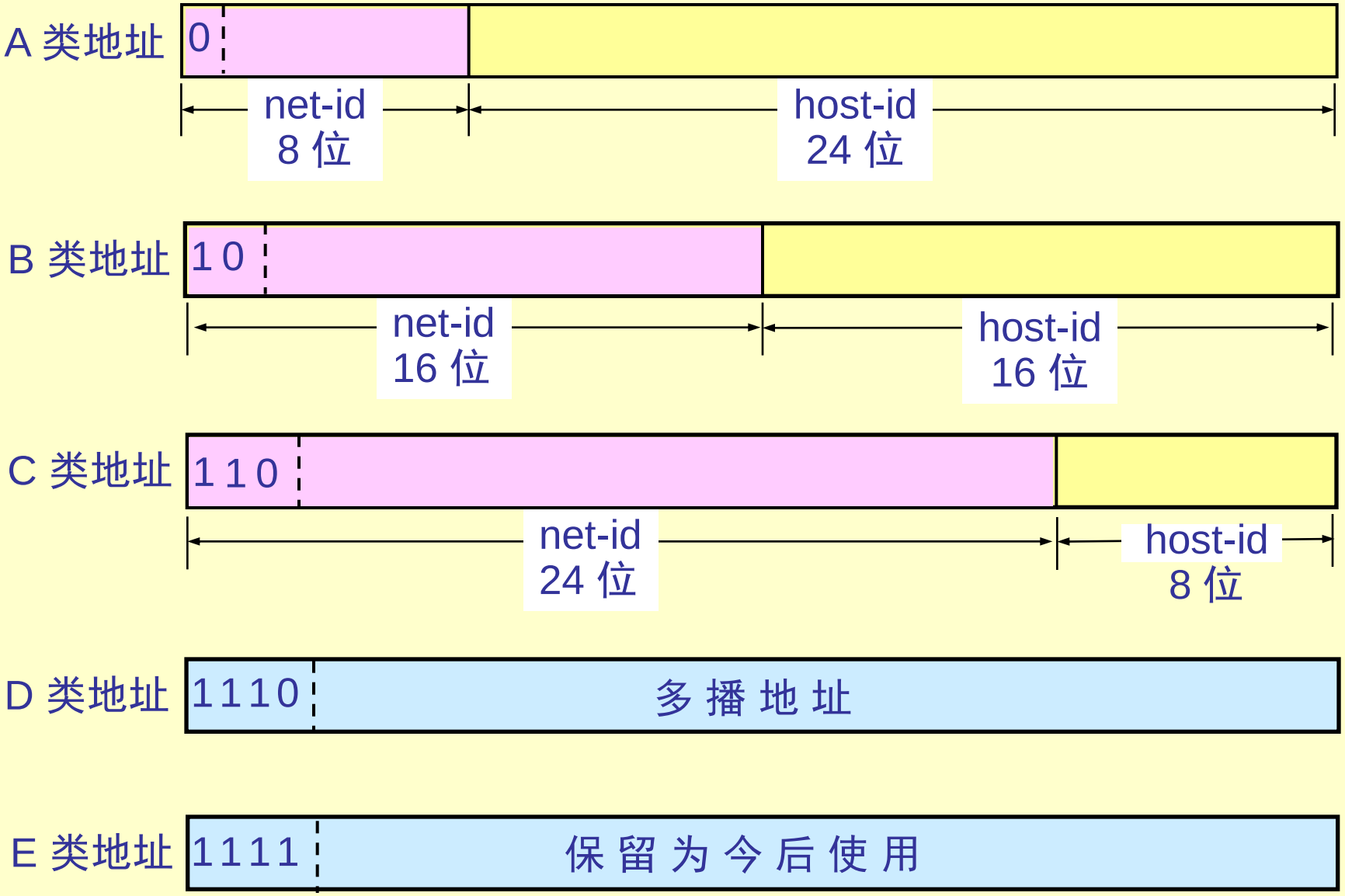
5.4.2 网际协议(IP)

- 分类IP地址
  - 共5类IP地址：A类、B类、C类、D类、E类
  - 每一类地址都由三个字段组成，其中类别字段表示属于哪个类别，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）
- 采用二级的IP地址寻址方式

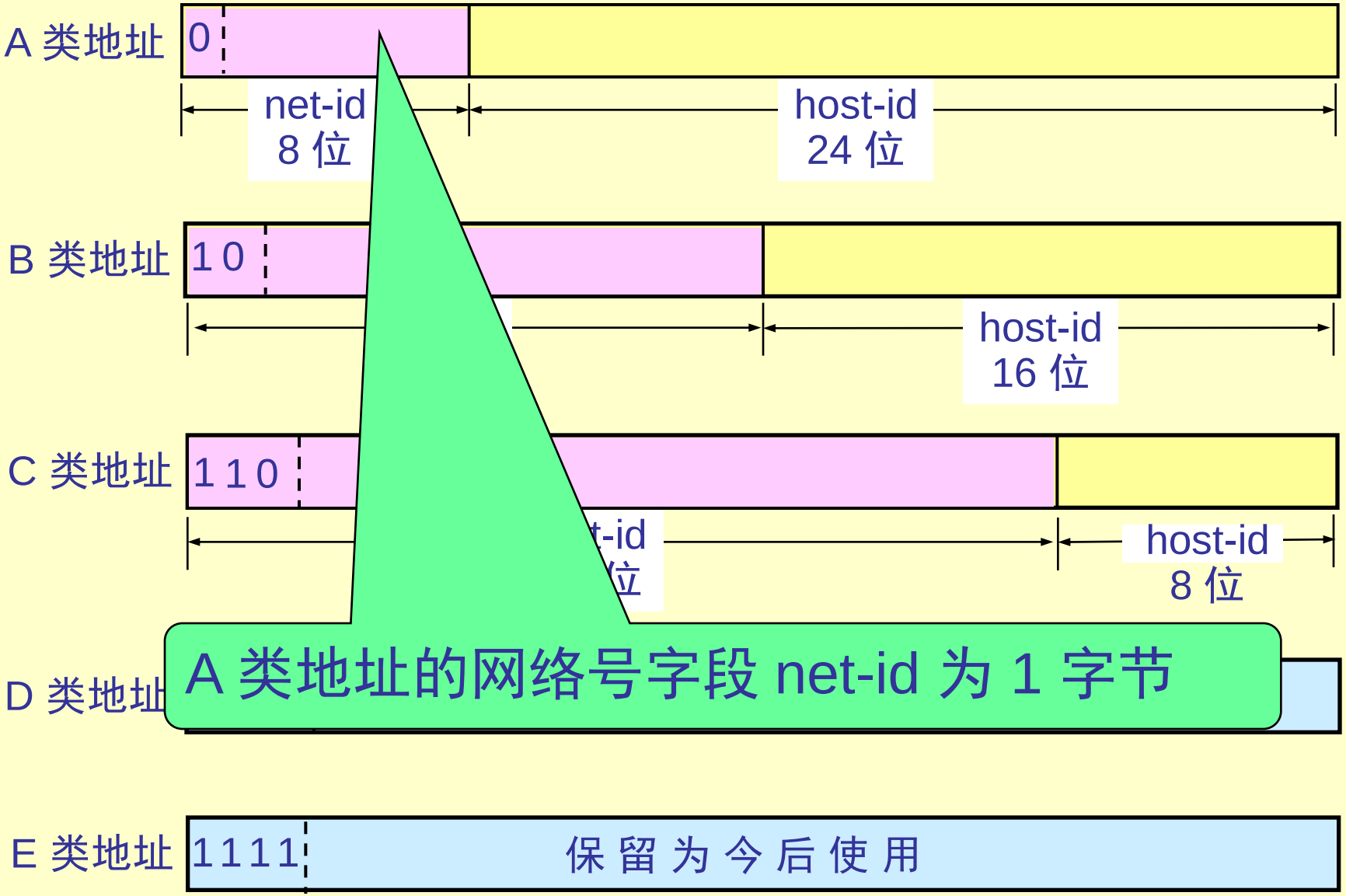
(1) 使用网络号字段将数据报先转发到目的网络

(2) 采用主机号将数据报交付到目的网络上的目的主机

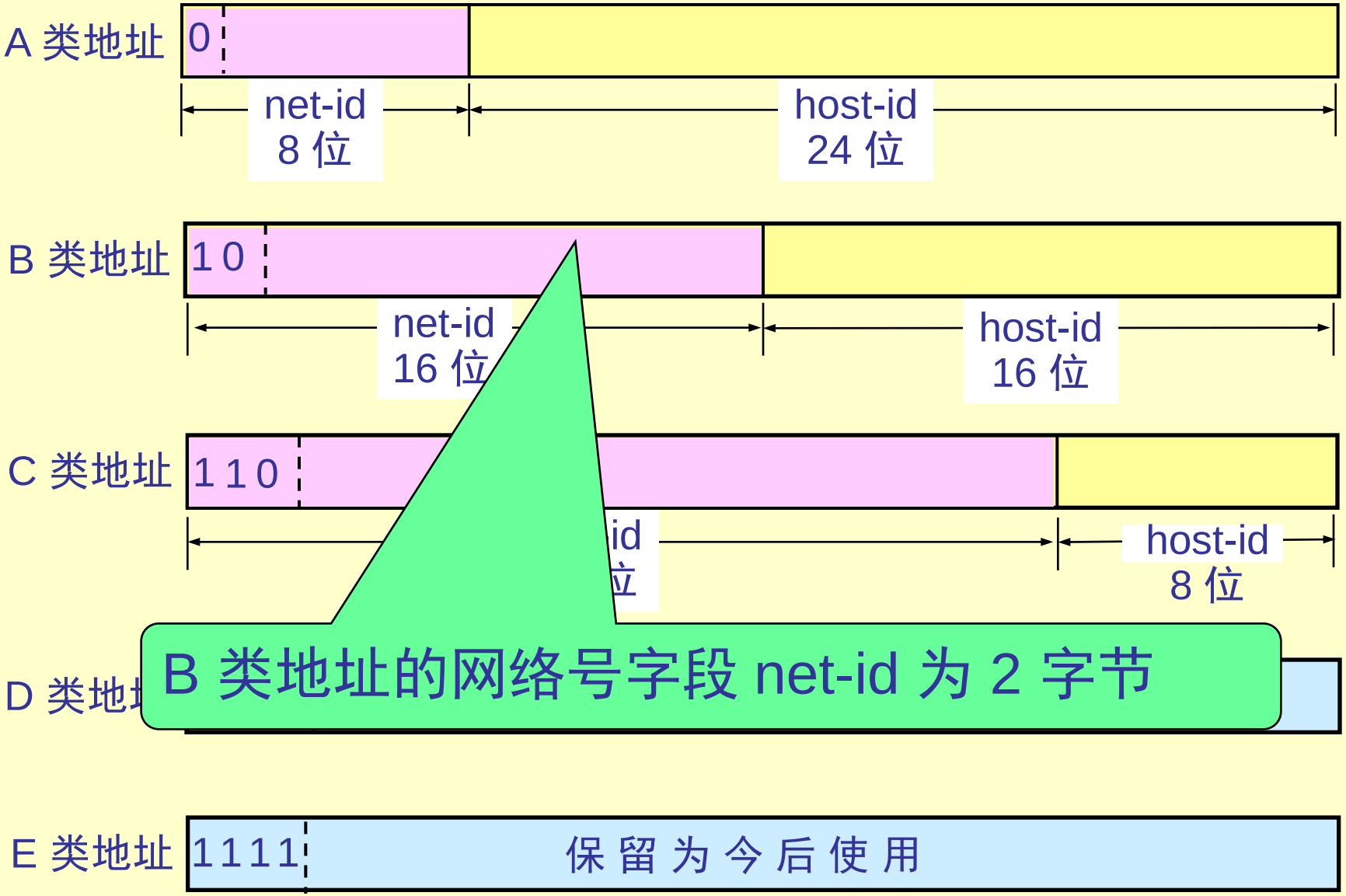
IP 地址中的网络号字段和主机号字段



IP 地址中的网络号字段和主机号字段

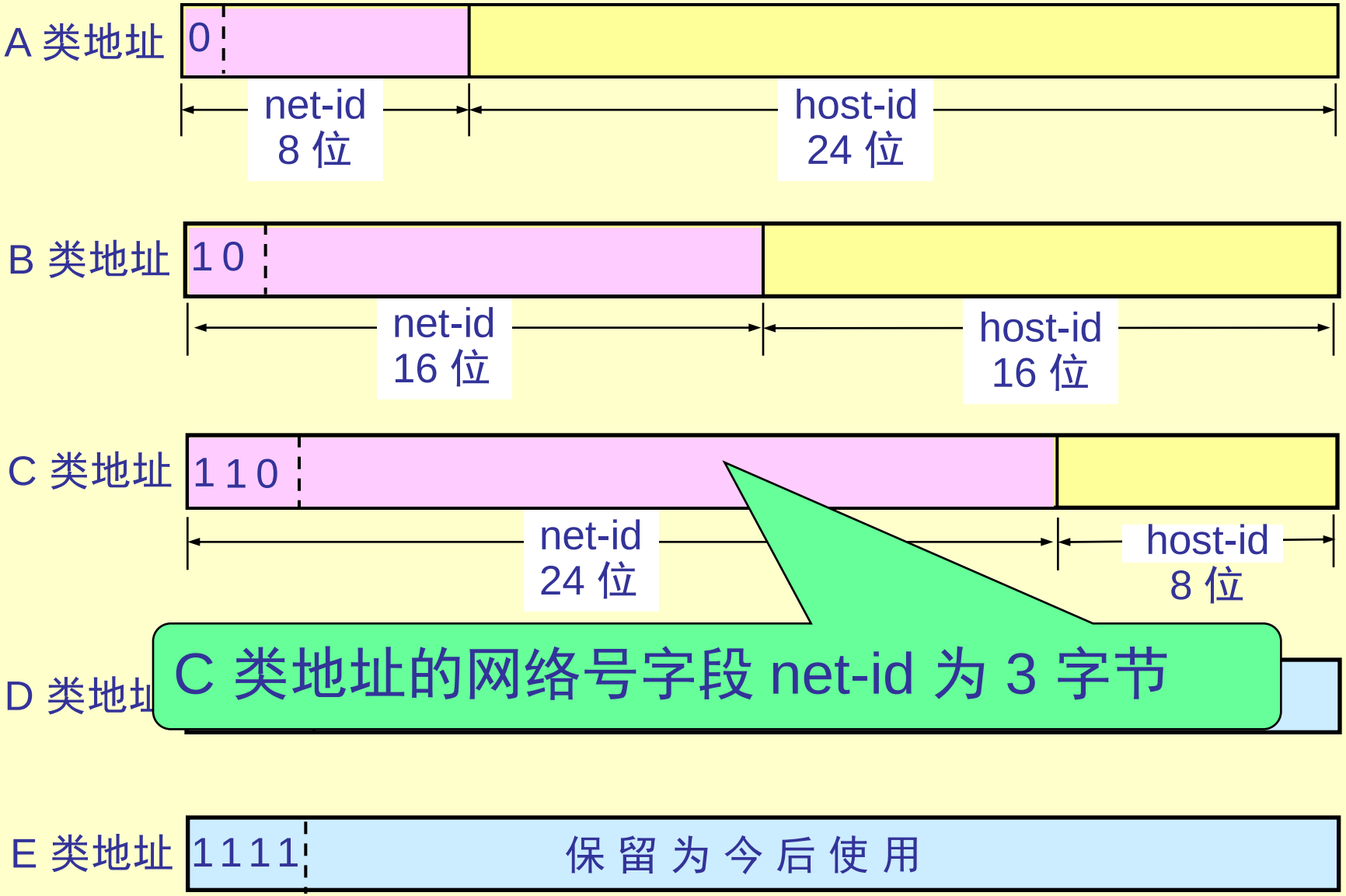


IP 地址中的网络号字段和主机号字段

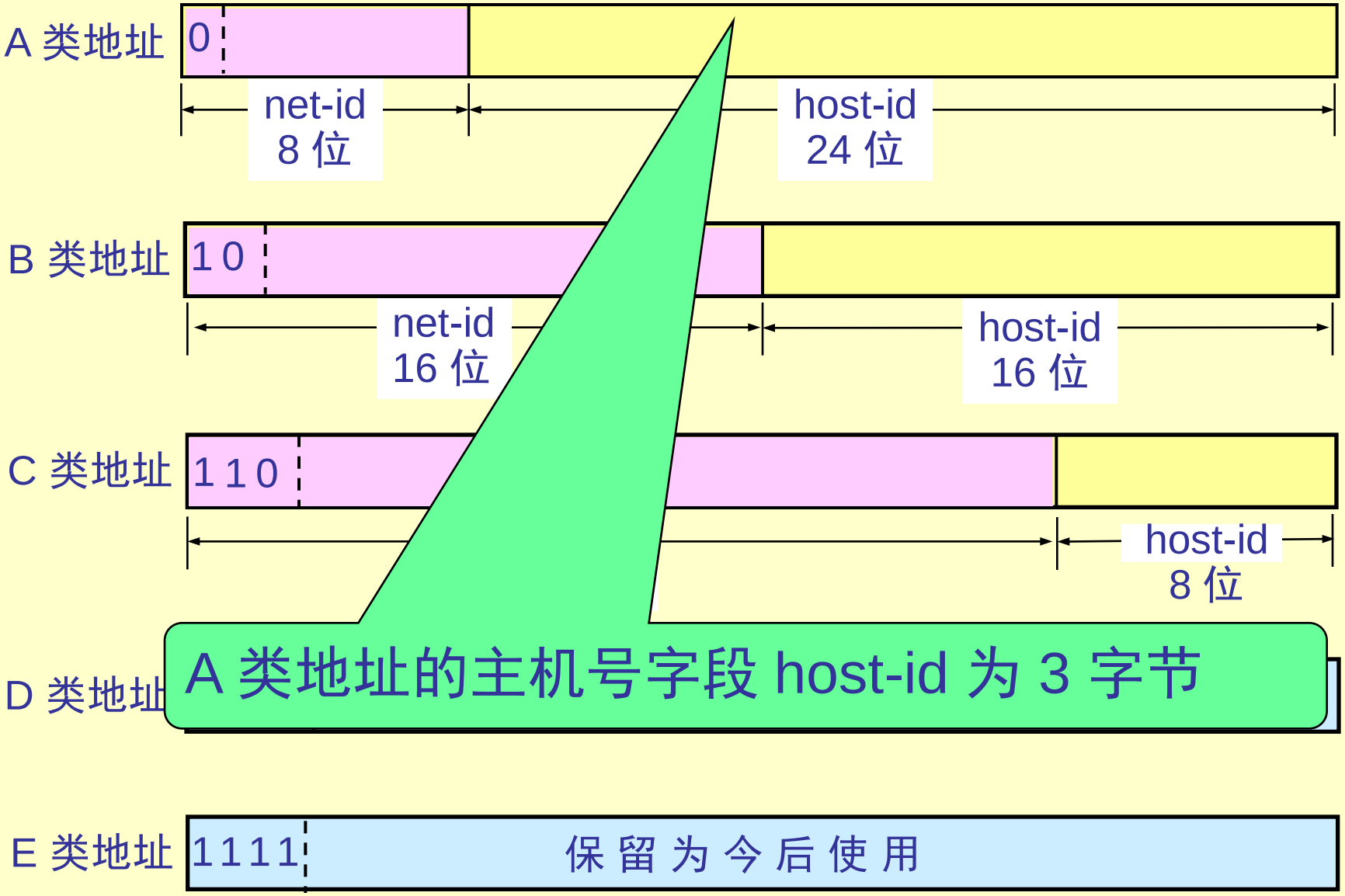


B 类地址的网络号字段 net-id 为 2 字节

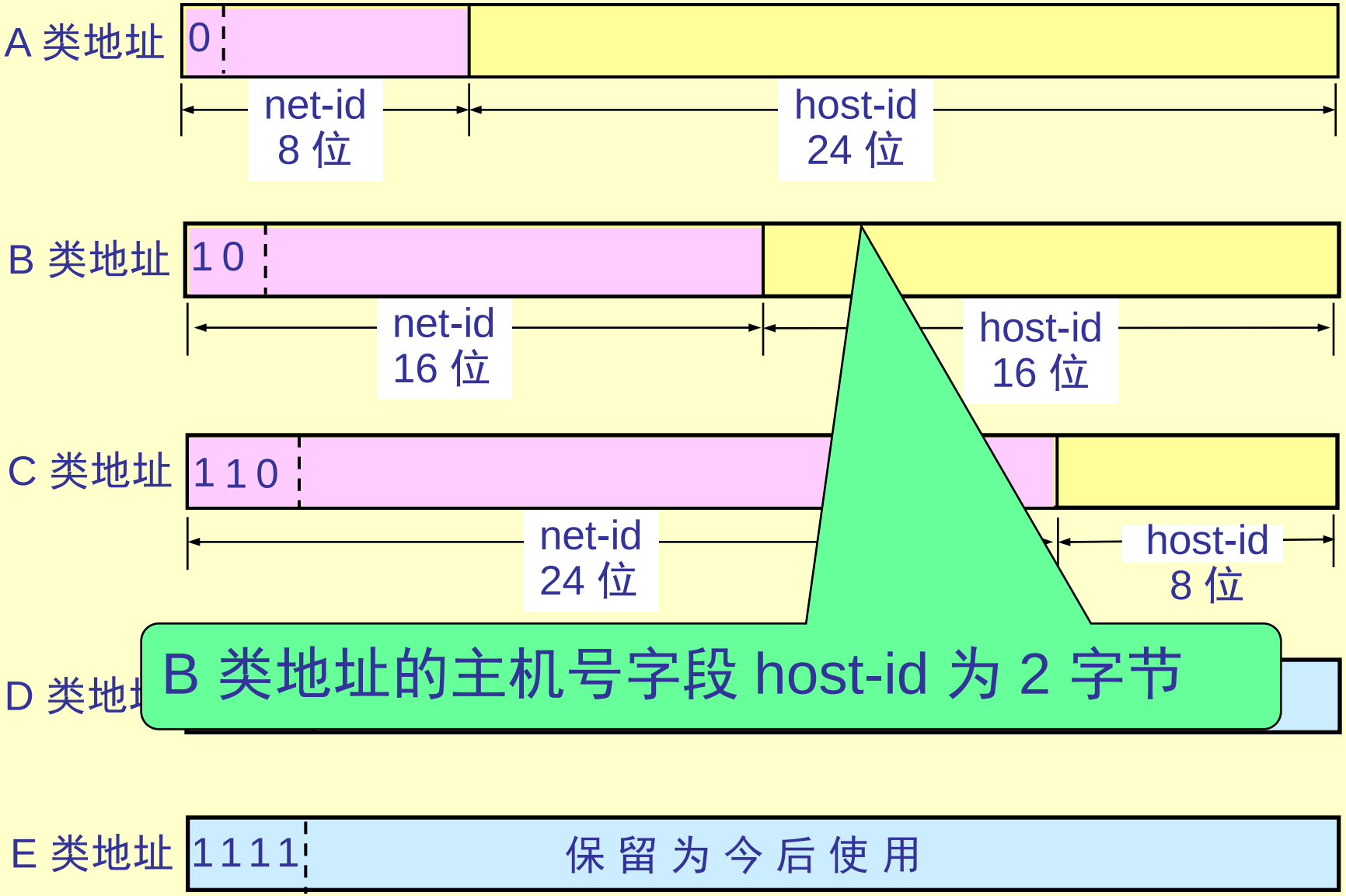
IP 地址中的网络号字段和主机号字段



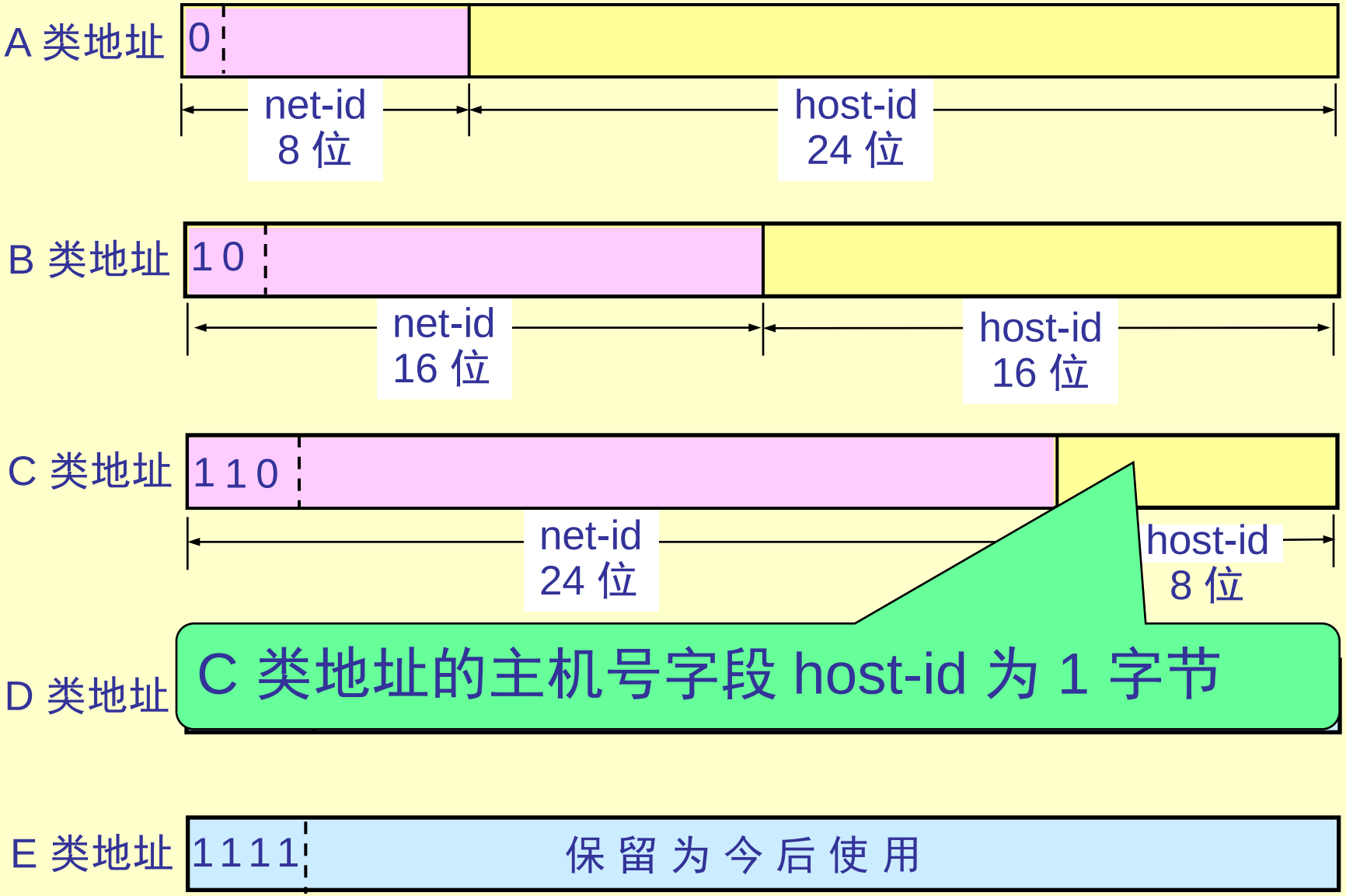
IP 地址中的网络号字段和主机号字段



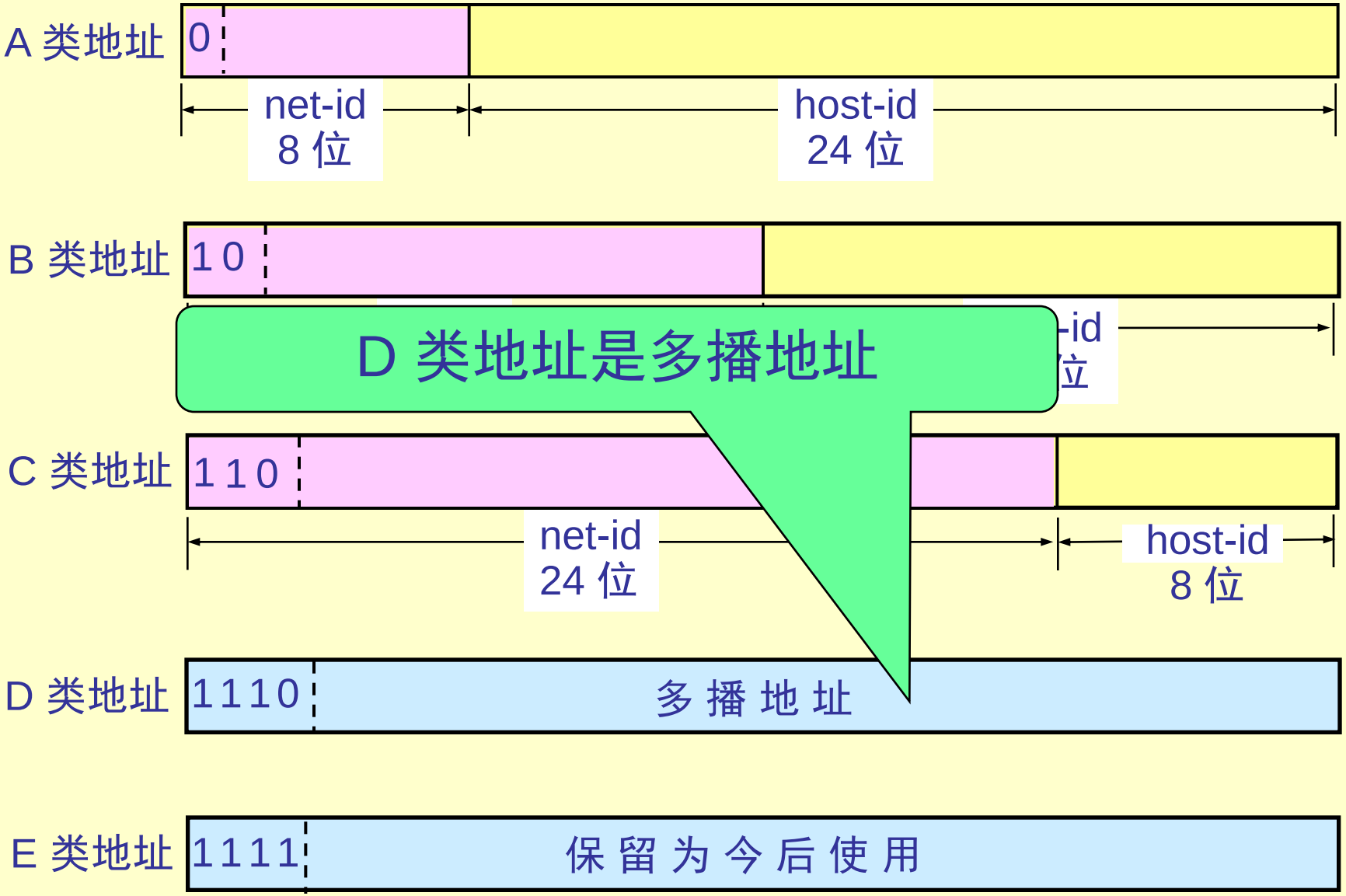
IP 地址中的网络号字段和主机号字段



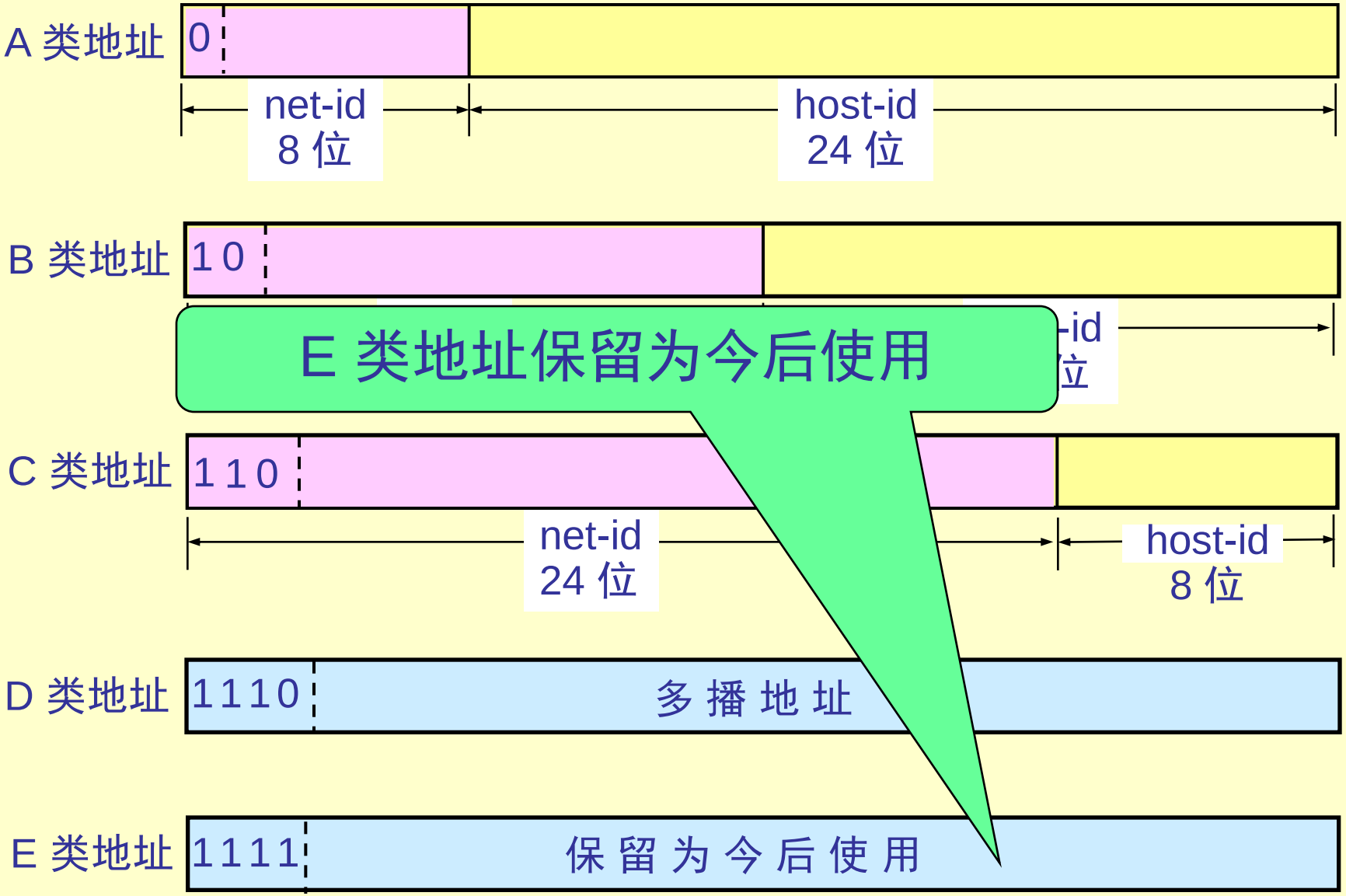
IP 地址中的网络号字段和主机号字段



IP 地址中的网络号字段和主机号字段



IP 地址中的网络号字段和主机号字段



5.4.2 网际协议(IP)

- 点分十进制记法

机器中存放的 IP 地址是 32 位 二进制代码 → 10000000000010110000001100011111

每隔 8 位插入一个空格能够提高可读性 → 10000000 00001011 00000011 00011111

将每 8 位的二进制数转换为十进制数 → 128 11 3 31

采用点分十进制记法则进一步提高可读性 → 128.11.3.31

5.4.2 网际协议(IP)

- 常用的三种类别的IP地址

IP地址的使用范围

网络类别	最大网络数	第一个可用的网络号	最后一个可用的网络号	每个网络中最大的主机数
A	126 ($2^7 - 2$)	1	126	16,777,214
B	16,383($2^{14} - 1$)	128.1	191.255	65,534
C	2,097,151 ($2^{21} - 1$)	192.0.1	223.255.255	254

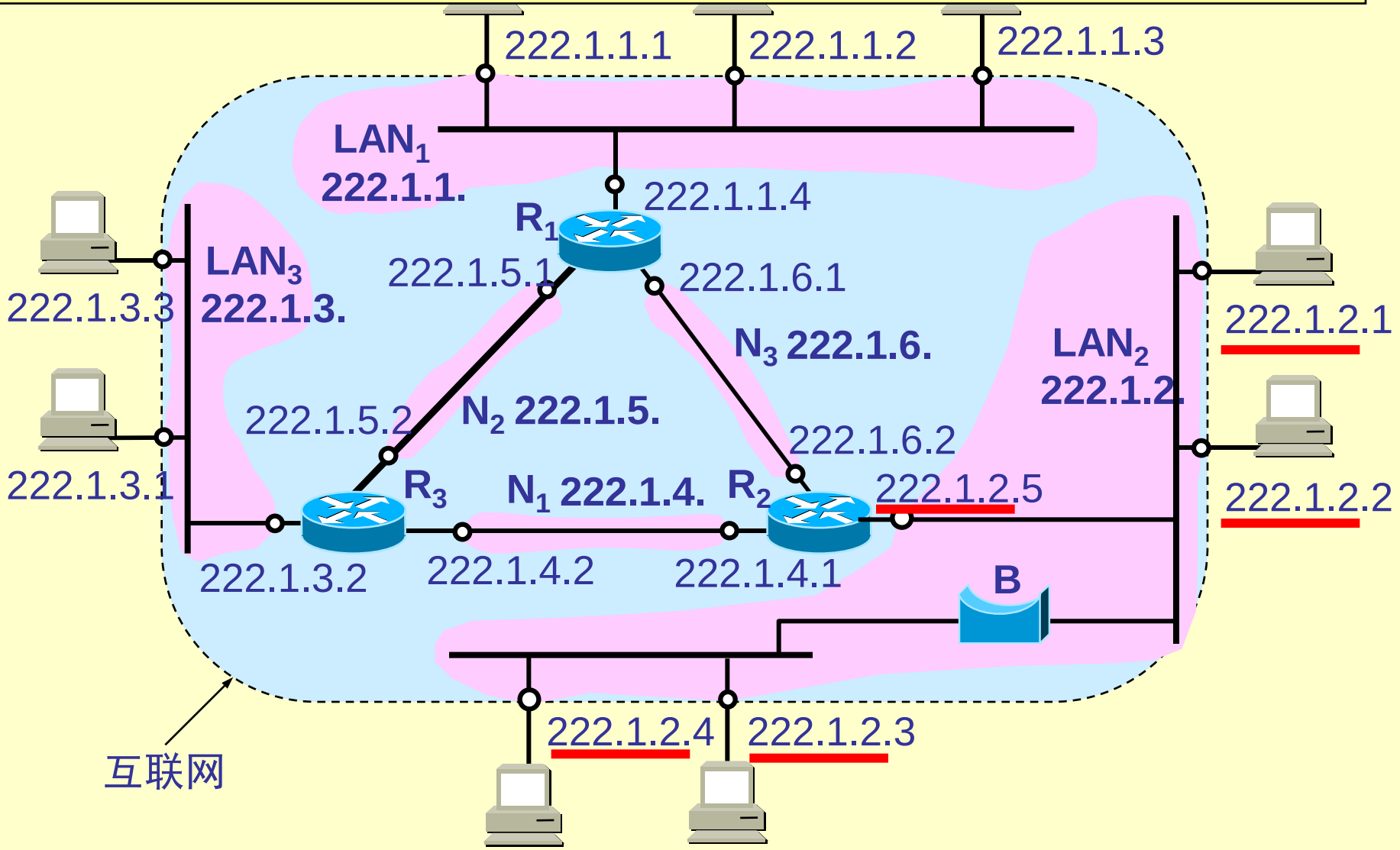
IP地址的一些重要特点

- IP地址是一种分等级的地址结构
- 分两个等级的好处
  - IP地址管理机构在分配IP地址时只分配网络号，而剩下的主机号则由得到该网络号的单位自行分配。这样就方便了IP地址的管理
  - 路由器仅根据目的主机所连接的网络号来转发分组（而不考虑目的主机号），这样就可以使路由表中的项目数大幅度减少，从而减小了路由表所占的存储空间

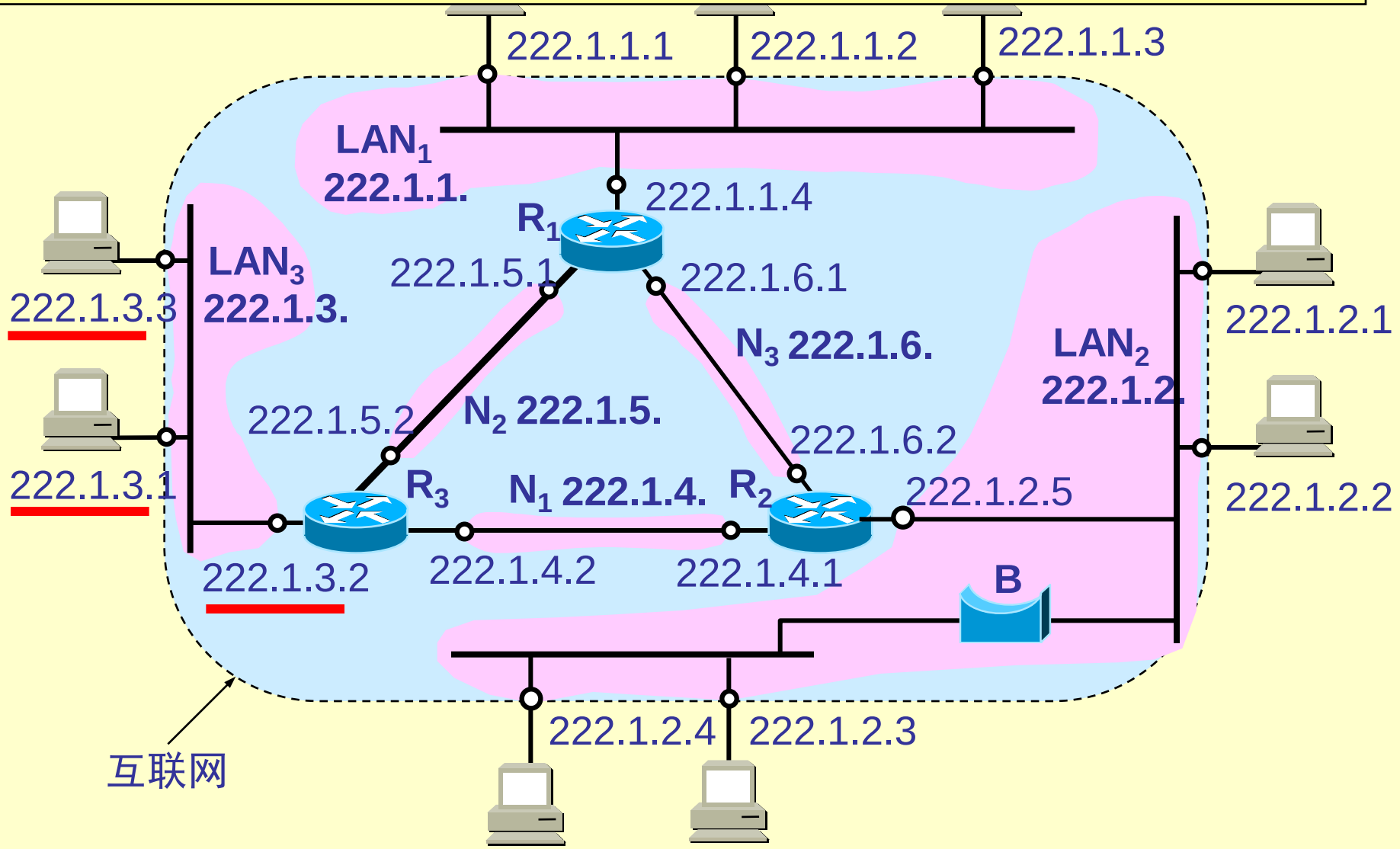
IP地址的一些重要特点

- 实际上IP地址标志一个主机（或路由器）和一条链路的接口
  - 当一个主机同时连接到两个网络上时，该主机就必须同时具有两个相应的IP地址，其网络号net-id必须是不同的。这种主机称为多归属主机（multihomed host）
  - 由于一个路由器至少应当连接到两个网络（这样它才能将IP数据报从一个网络转发到另一个网络），因此一个路由器至少应当有两个不同的IP地址

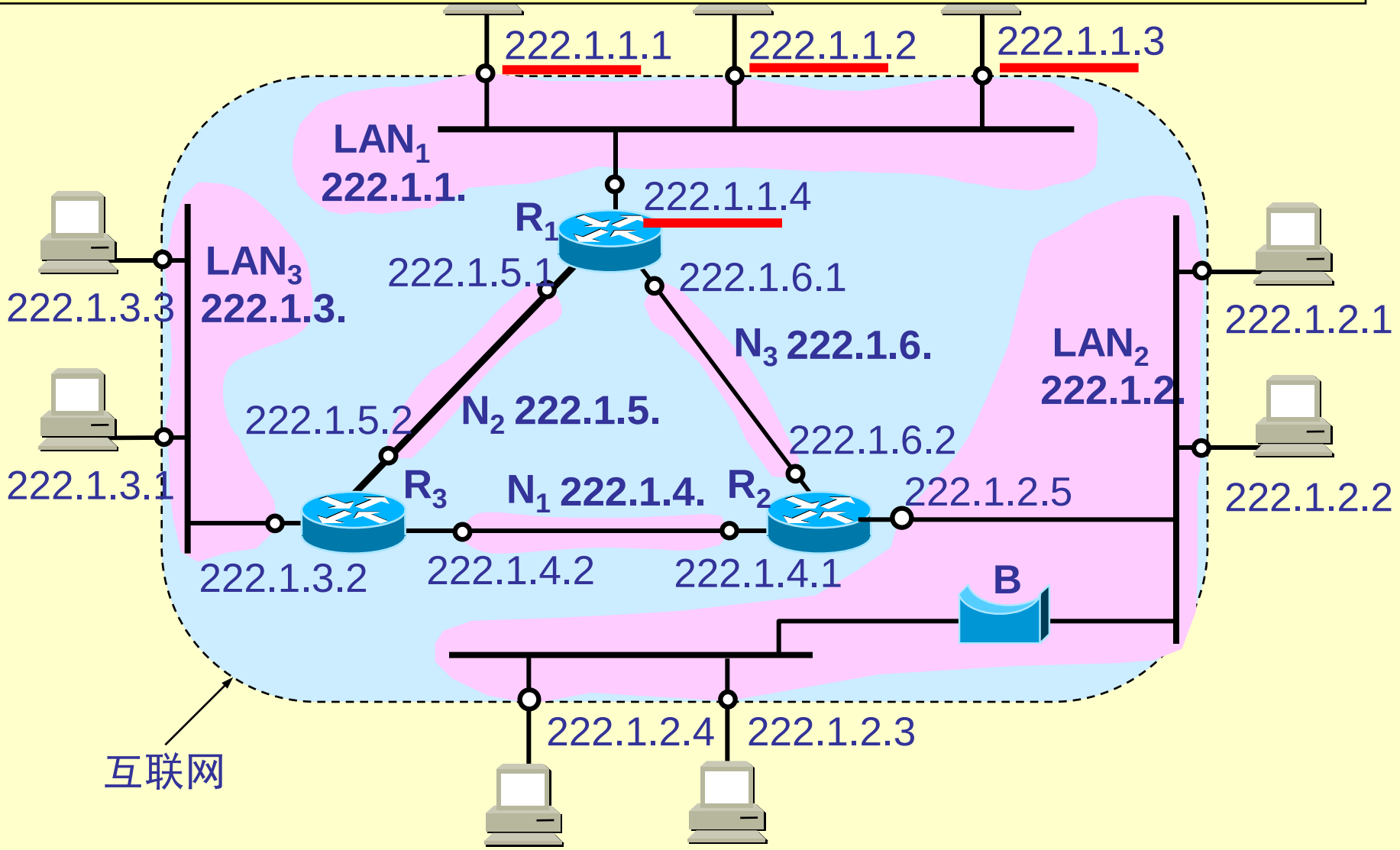
在同一个局域网上的主机或路由器的IP 地址中网络号必须是一样的。图中网络号就是 IP 地址中的 net-id。



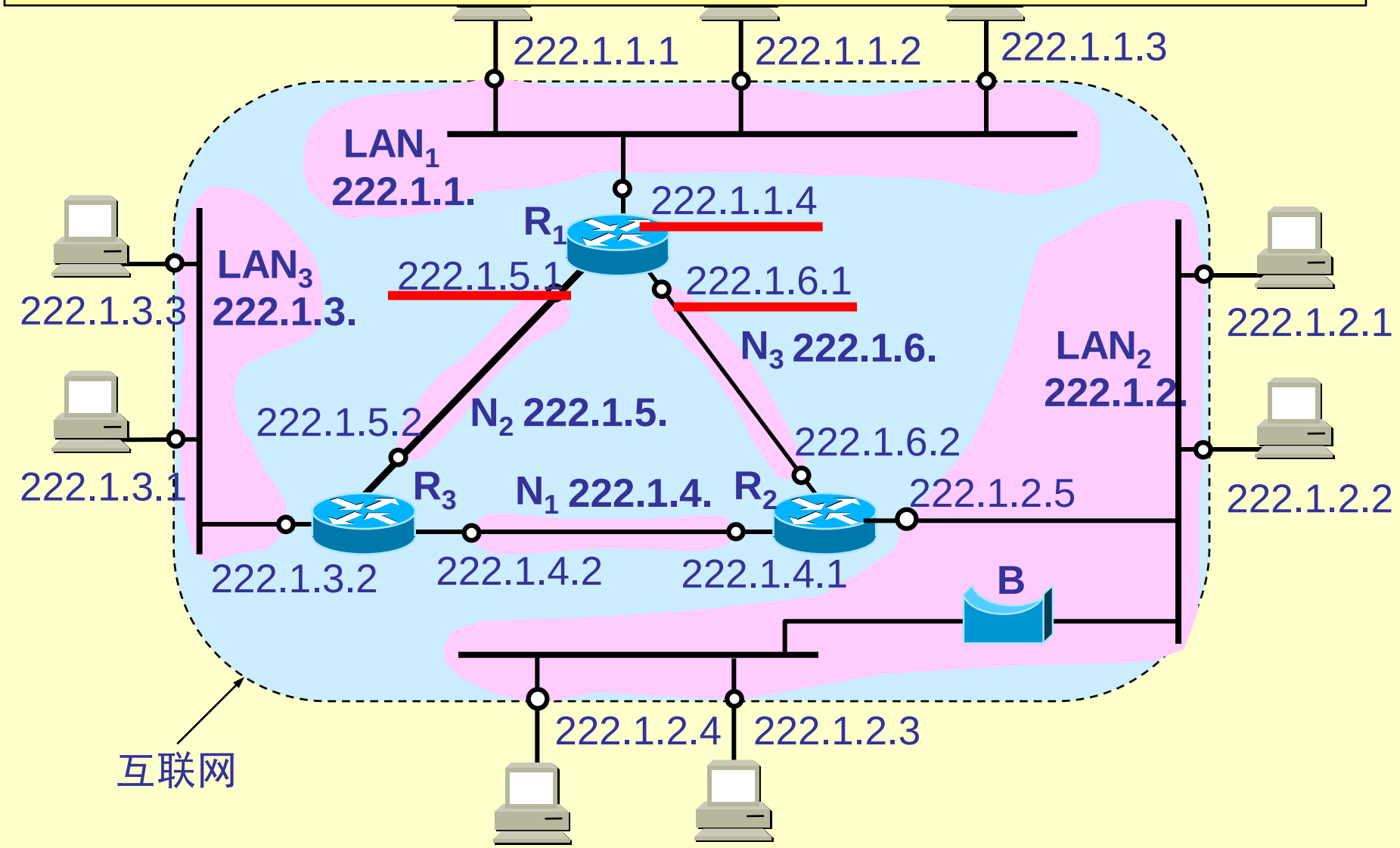
在同一个局域网上的主机或路由器的IP 地址中网络号必须是一样的。图中网络号就是 IP 地址中的 net-id。



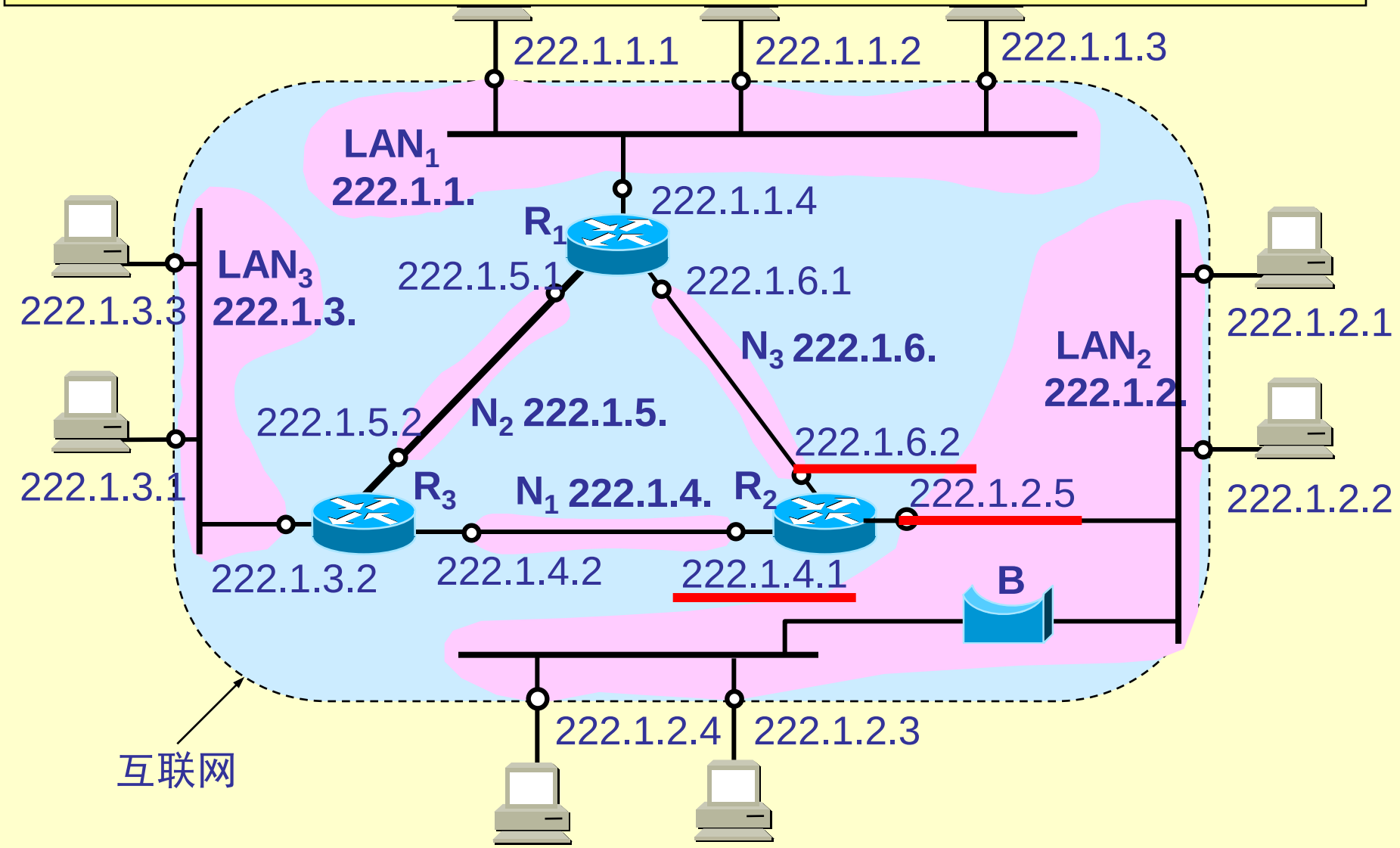
在同一个局域网上的主机或路由器的IP 地址中网络号必须是一样的。图中网络号就是 IP 地址中的 net-id。



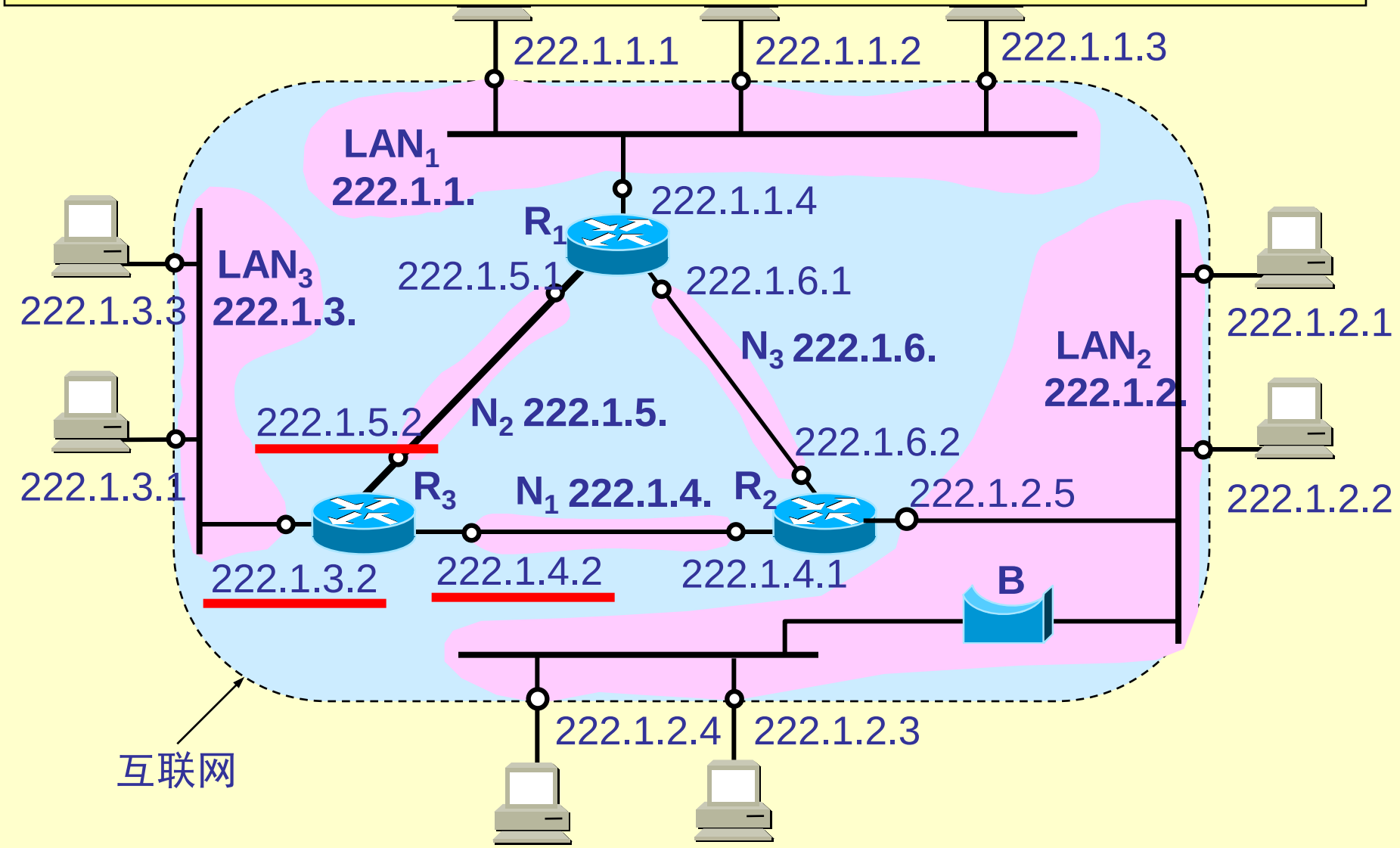
路由器总是具有两个或两个以上的 IP 地址。
路由器的每一个接口都有一个不同网络号的 IP 地址。



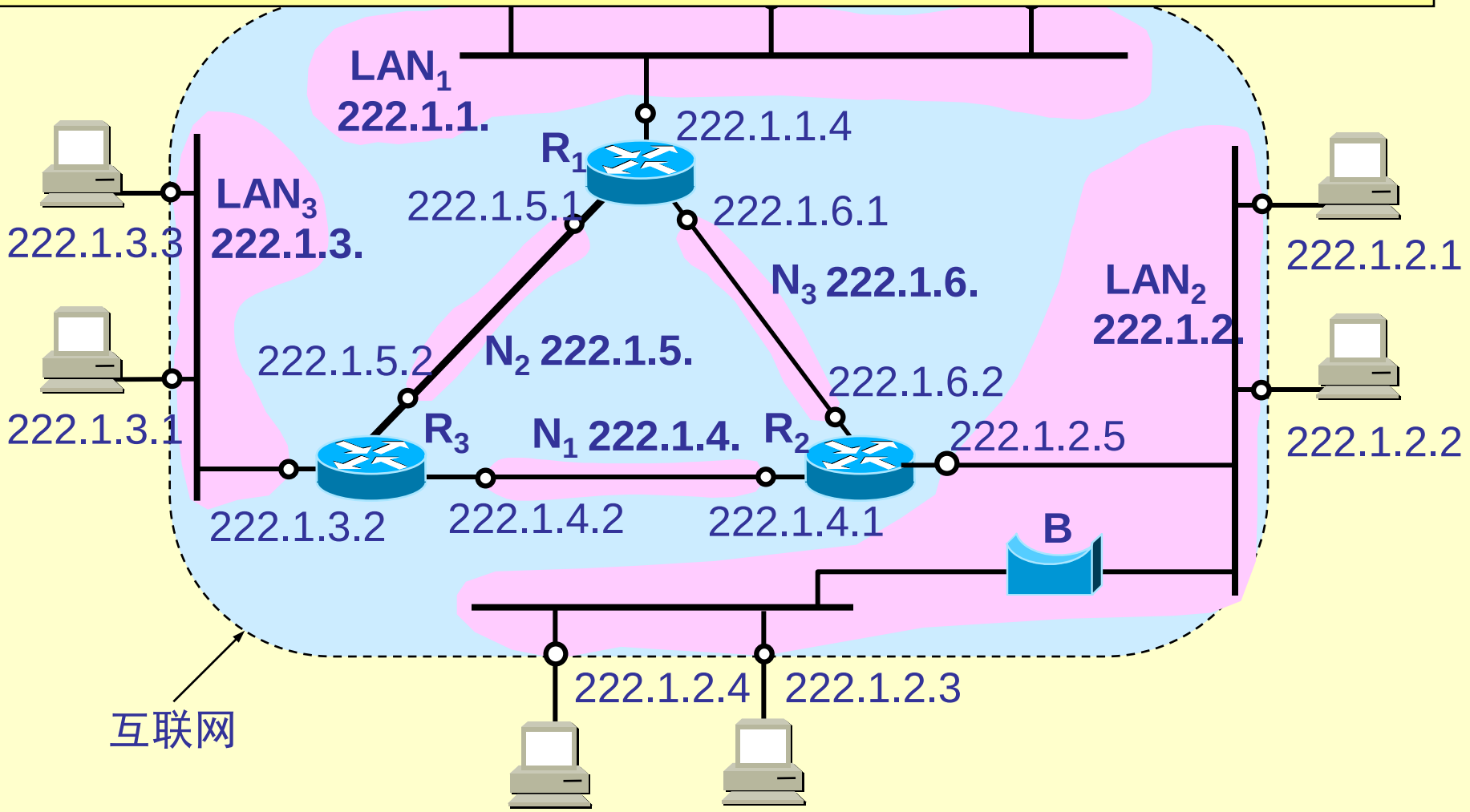
路由器总是具有两个或两个以上的 IP 地址。
路由器的每一个接口都有一个不同网络号的 IP 地址。



路由器总是具有两个或两个以上的 IP 地址。
路由器的每一个接口都有一个不同网络号的 IP 地址。



两个路由器直接相连的接口处，可指明也可不指明 IP 地址。如指明 IP 地址，则这一段连线就构成了一种只包含一段线路的特殊“网络”。现常不指明 IP 地址。



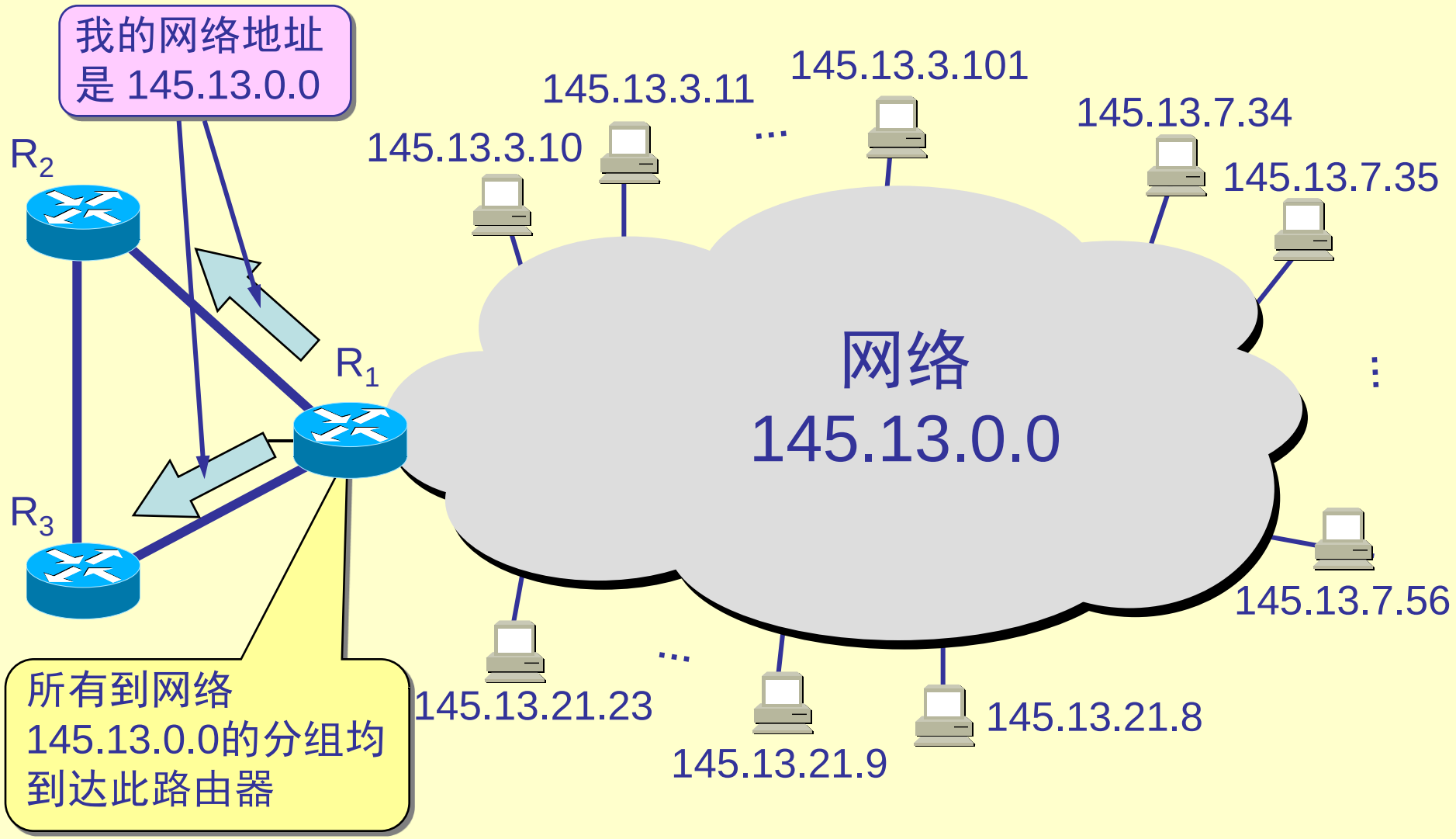
5.4.2 网际协议(IP)

- 划分子网
  - 一个单位在一个网络号下有大量计算机并不便于管理，因此可以根据单位的所属部门及其地理分布位置等划分子网
  - IP地址允许将单位自己控制的主机号字段中的前若干比特划分出来作为子网号
  - 子网号使用多少比特，单位根据自己需要决定

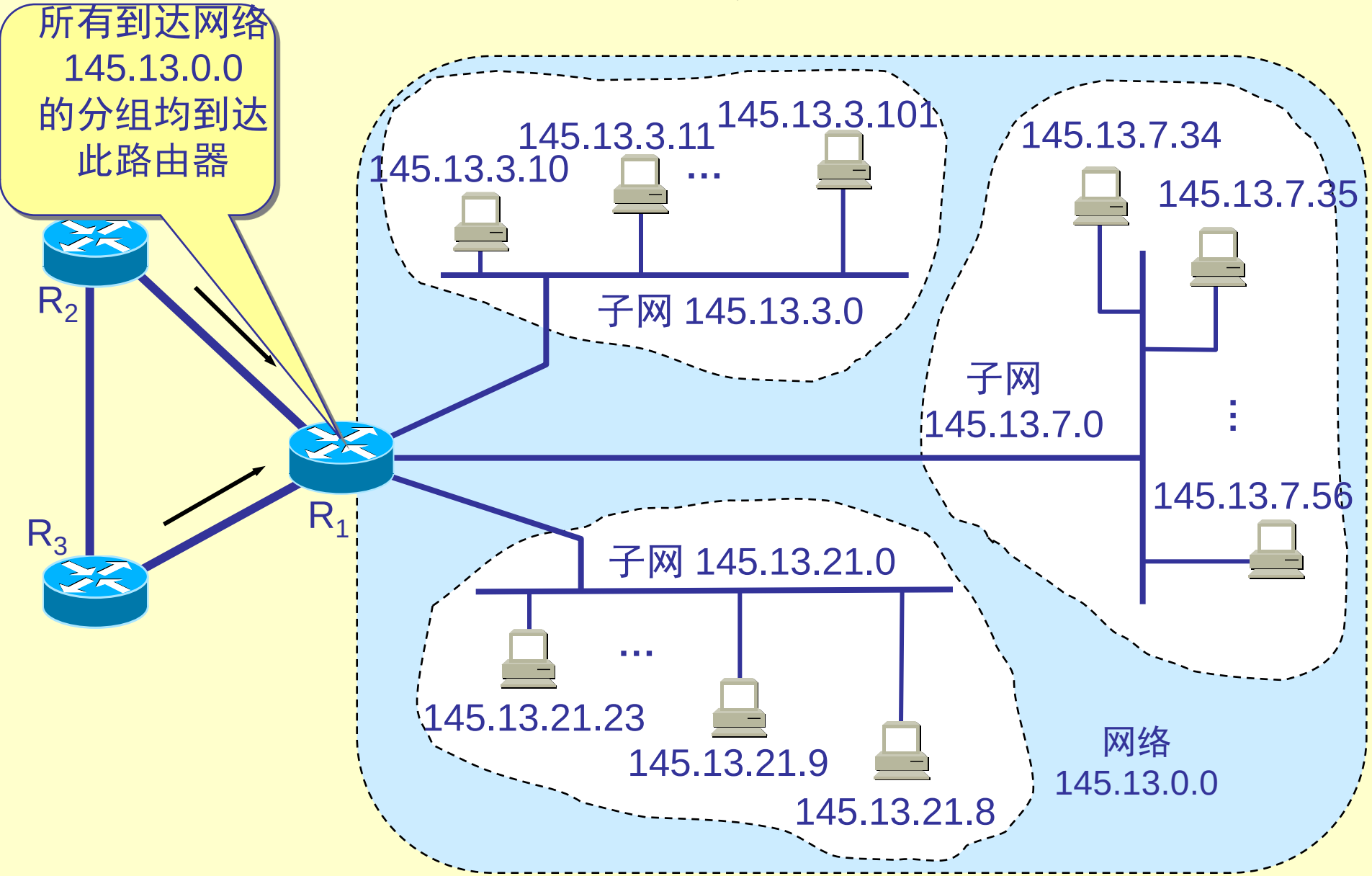
5.4.2 网际协议(IP)

- 划分子网
  - 凡是从其他网络发送给本单位某个主机的IP数据报，仍然是根据IP数据报的目的网络号net-id，先找到连接在本单位网络上的路由器
  - 然后此路由器在收到IP数据报后，再按目的网络号net-id和子网号subnet-id找到目的子网
  - 最后就将IP数据报直接交付给目的主机

一个未划分子网的 B 类网络145.13.0.0



划分为三个子网后对外仍是一个网络，对内根据子网号交付至对应子网的目的主机

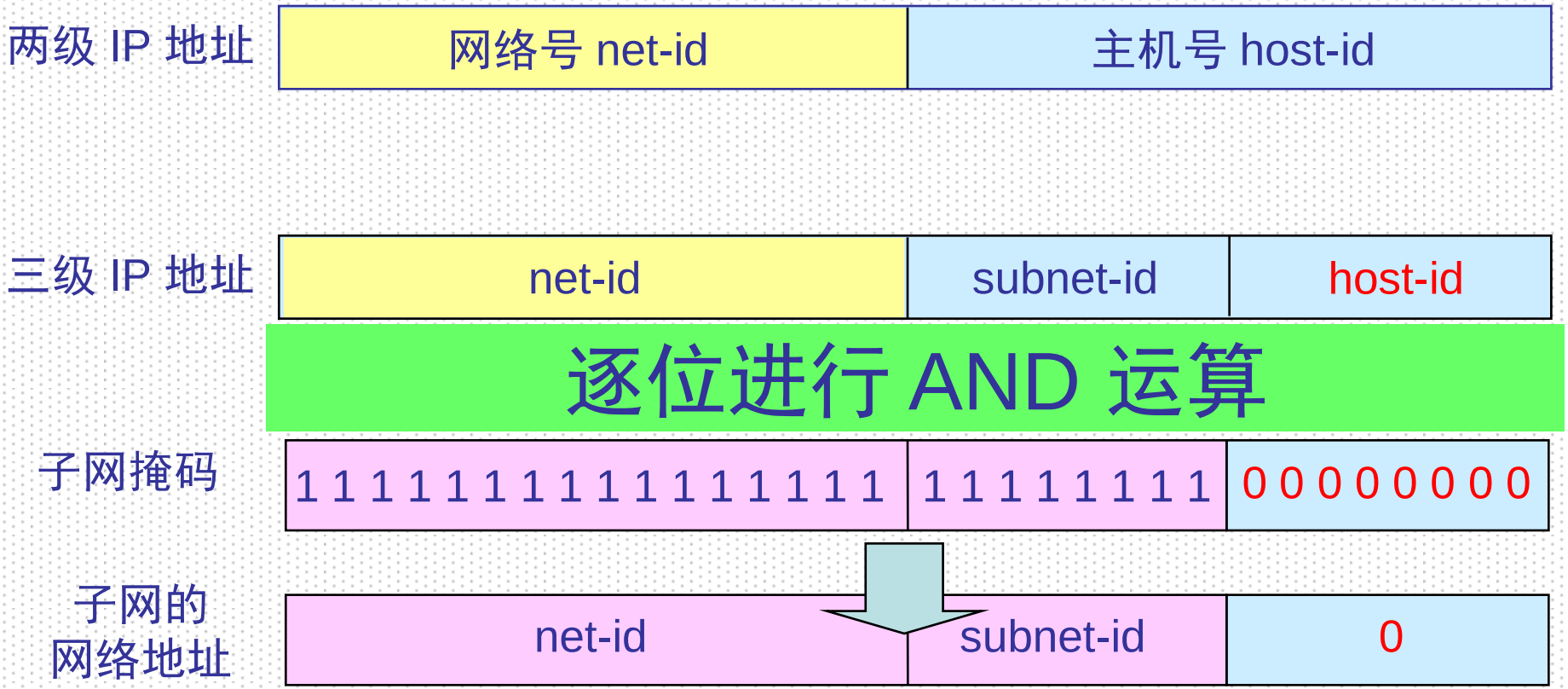


5.4.2 网际协议(IP)

- 子网掩码的定义
  - 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

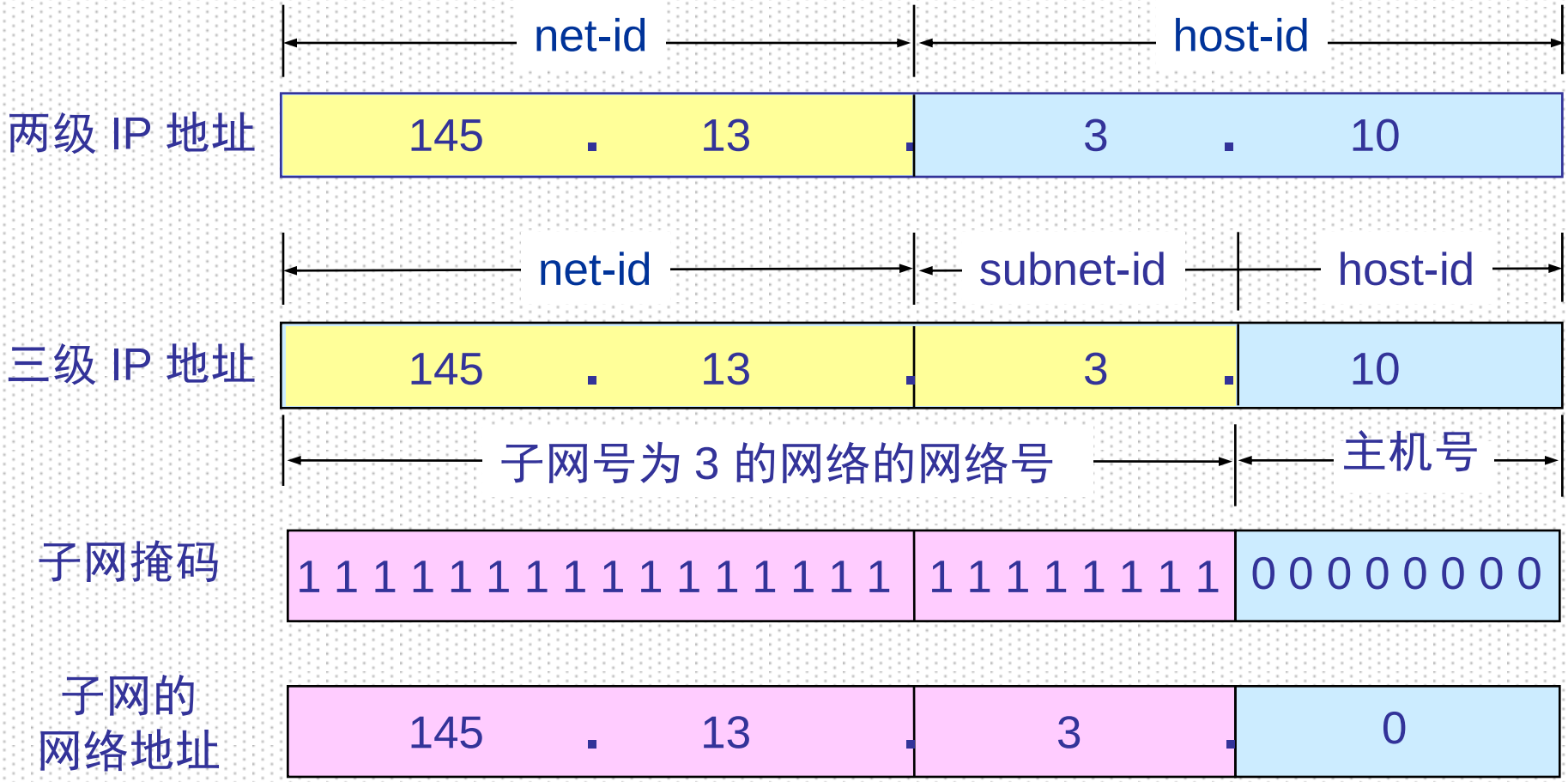
5.4.2 网际协议(IP)

- (IP地址) AND (子网掩码) = 网络地址

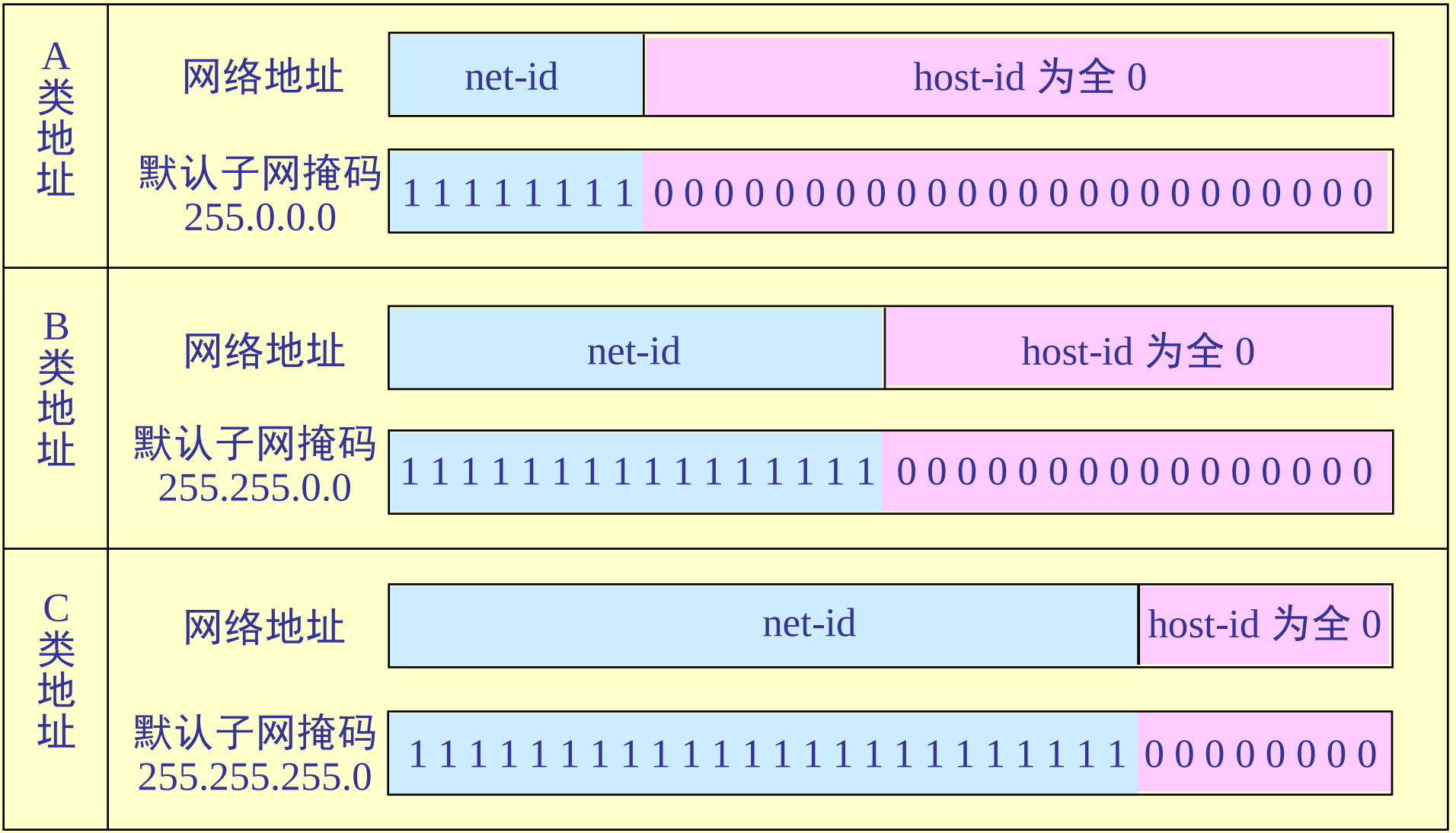


5.4.2 网际协议(IP)

- IP地址的各字段和子网掩码



默认子网掩码


A 类地址
网络地址
net-id
host-id 为全 0
默认子网掩码
255.0.0.0
1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
B 类地址
网络地址
net-id
host-id 为全 0
默认子网掩码
255.255.0.0
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
C 类地址
网络地址
net-id
host-id 为全 0
默认子网掩码
255.255.255.0
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0

5.4.2 网际协议(IP)

- **子网掩码是一个重要属性**
  - 子网掩码是一个网络或一个子网的重要属性
  - 路由器在和相邻路由器交换路由信息时，必须把自己所在网络（或子网）的子网掩码告诉相邻路由器
  - 路由器的路由表中的每一个项目，除了要给出目的网络地址外，还必须同时给出该网络的子网掩码
  - 若一个路由器连接在两个子网上就拥有两个网络地址和两个子网掩码

子网掩码-实例

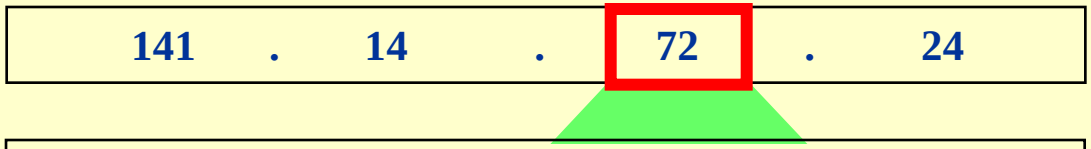
- 无子网的表示法
  - 如果一个IP网络无子网，则屏蔽码中的网络号字段各位全为1，主机号字段各位全为0
  - IP地址：202.114.80.5
  - 子网掩码：255.255.255.0
  - IP地址202.114.80.5标识了202.114.80号网络中的5号主机，并且202.114.80号网络中没有设置子网

子网掩码-实例

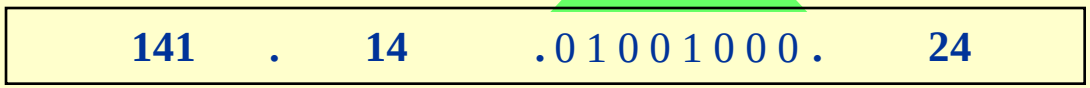
- 有子网的表示法
  - 如果一个IP网络有子网，则子网号用主机号字段的前几位来表示，所占的位数与子网的数量相对应，并且屏蔽码和IP地址必须成对出现，屏蔽码中的网络号字段各位全为1，主机号字段中的子网号各位也全为1，而主机号各位全为0

【例】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。

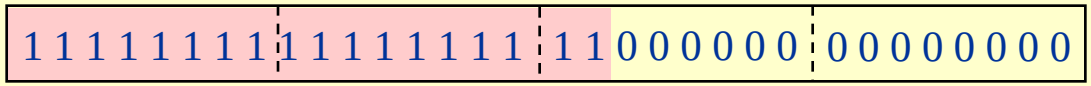
(a) 点分十进制表示的 IP 地址



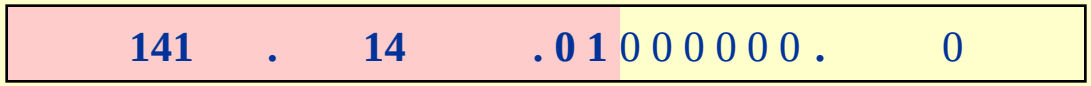
(b) IP 地址的第 3 字节是二进制



(c) 子网掩码是 255.255.192.0

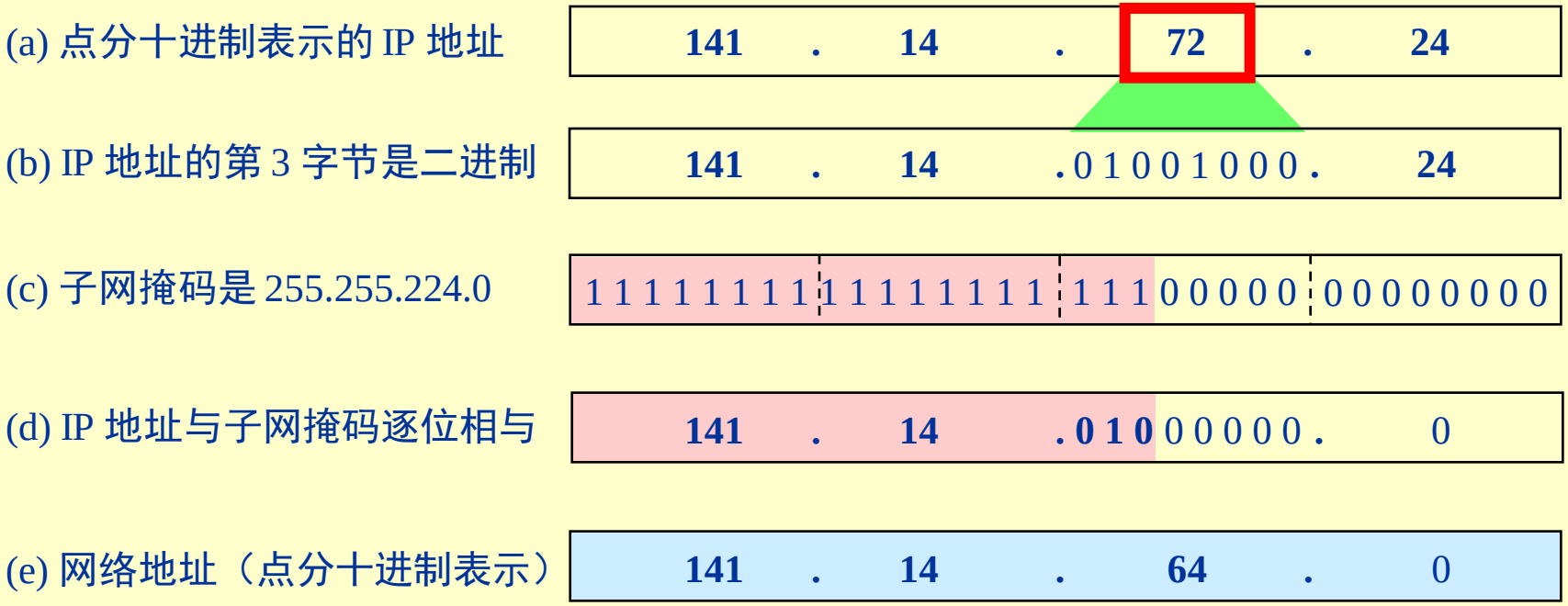


(d) IP 地址与子网掩码逐位相与



(e) 网络地址（点分十进制表示）

【例】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。



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。

IP地址转发过程

- 使用子网掩码的分组转发过程
  - 在不划分子网的两级IP地址下，从IP地址得出网络地址是个很简单的事
  - 但在划分子网的情况下，从IP地址却不能唯一地得出网络地址来，这是因为网络地址取决于那个网络所采用的子网掩码，但数据报的首部并没有提供子网掩码的信息
  - 因此分组转发的算法也必须做相应的改动

IP地址转发过程

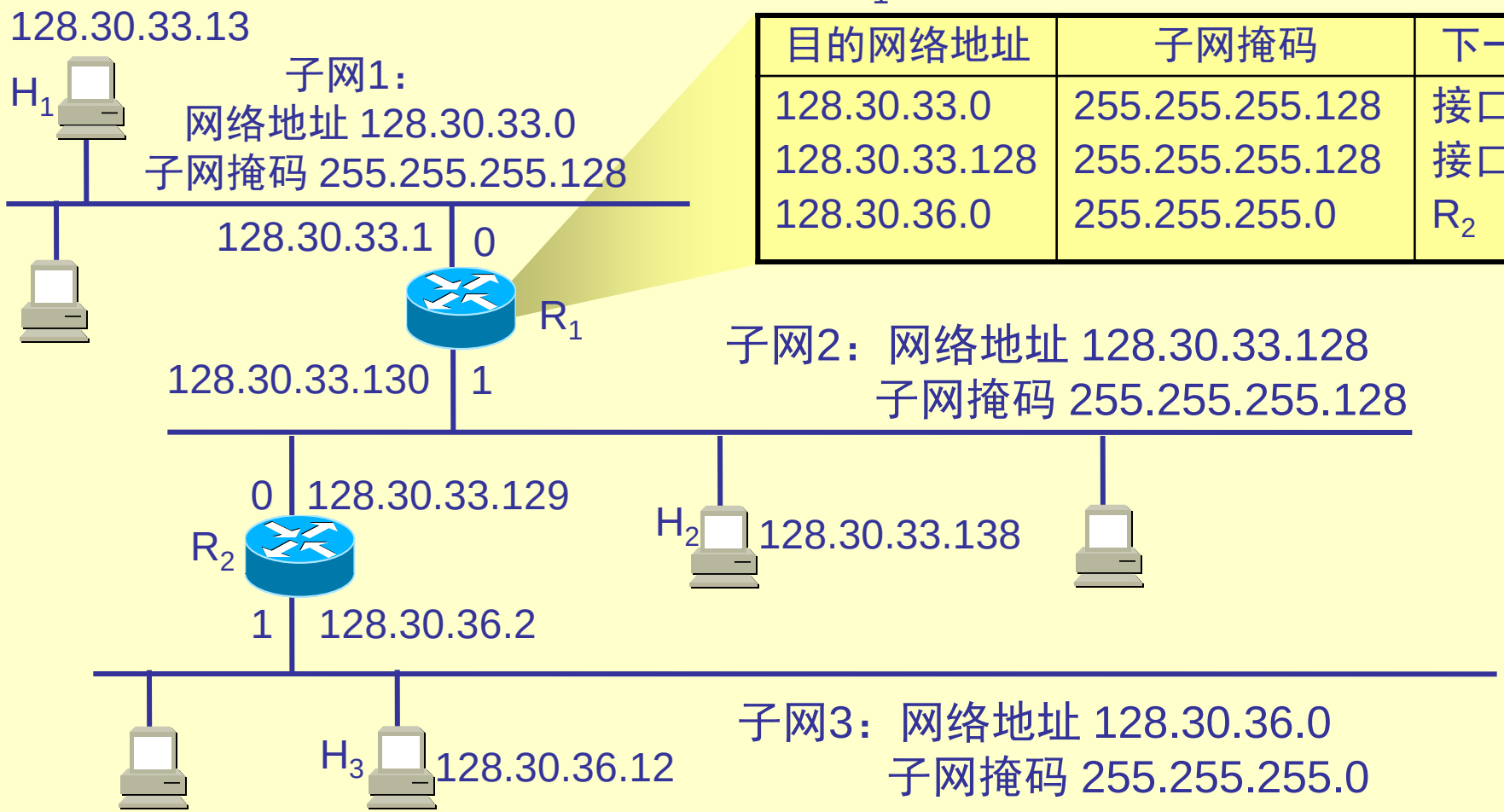
- 在划分子网的情况下路由器转发分组的算法
  - 此时路由表中必须包含三项内容：目的网络地址、子网掩码和下一跳地址

(1) 从收到的分组的首部提取目的 IP 地址 D 。
(2) 对路由器直接相连的网络逐个检查，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。

【例】已知互联网和路由器 R_1 中的路由表。主机 H_1 向 H_2 发送分组。试讨论 R_1 收到 H_1 向 H_2 发送的分组后查找路由表的过程。

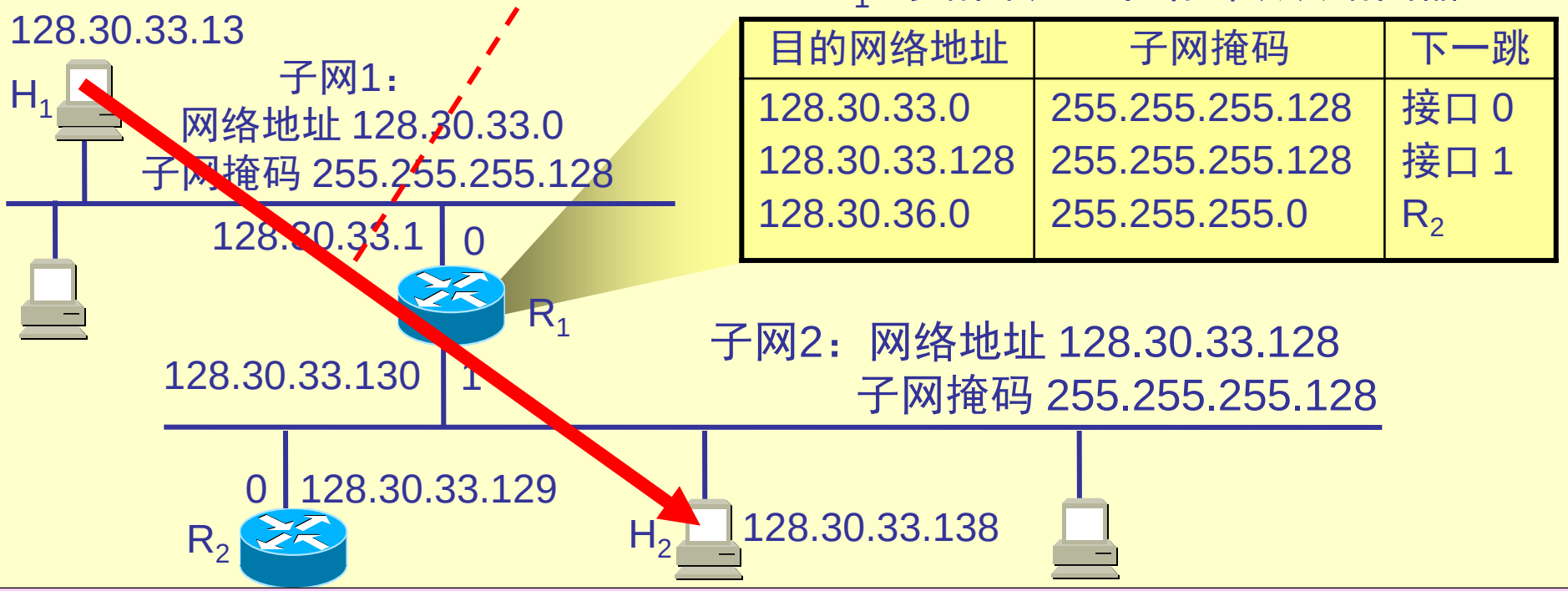
R_1 的路由表（未给出默认路由器）

目的网络地址	子网掩码	下一跳
128.30.33.0	255.255.255.128	接口 0
128.30.33.128	255.255.255.128	接口 1
128.30.36.0	255.255.255.0	R_2



主机 H_1 要发送分组给 H_2

要发送的分组的目的 IP 地址：128.30.33.138



R_1 的路由表（未给出默认路由器）

目的网络地址	子网掩码	下一跳
128.30.33.0	255.255.255.128	接口 0
128.30.33.128	255.255.255.128	接口 1
128.30.36.0	255.255.255.0	R_2

因此 H_1 首先检查主机 128.30.33.138 是否连接在本网络上

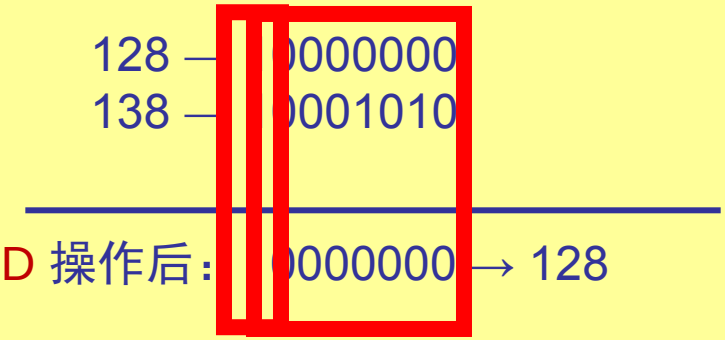
如果是，则直接交付；

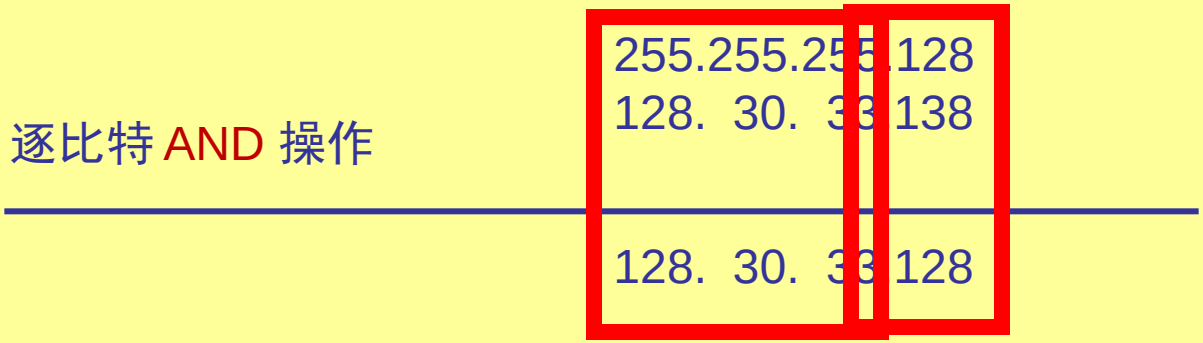
否则，就送交路由器 R_1，并逐项查找路由表。

主机 H_1 首先将
本子网的子网掩码 255.255.255.128
与分组的 IP 地址 128.30.33.138 逐比特相“与”(AND 操作)

255.255.255.128 AND 128.30.33.138 的计算

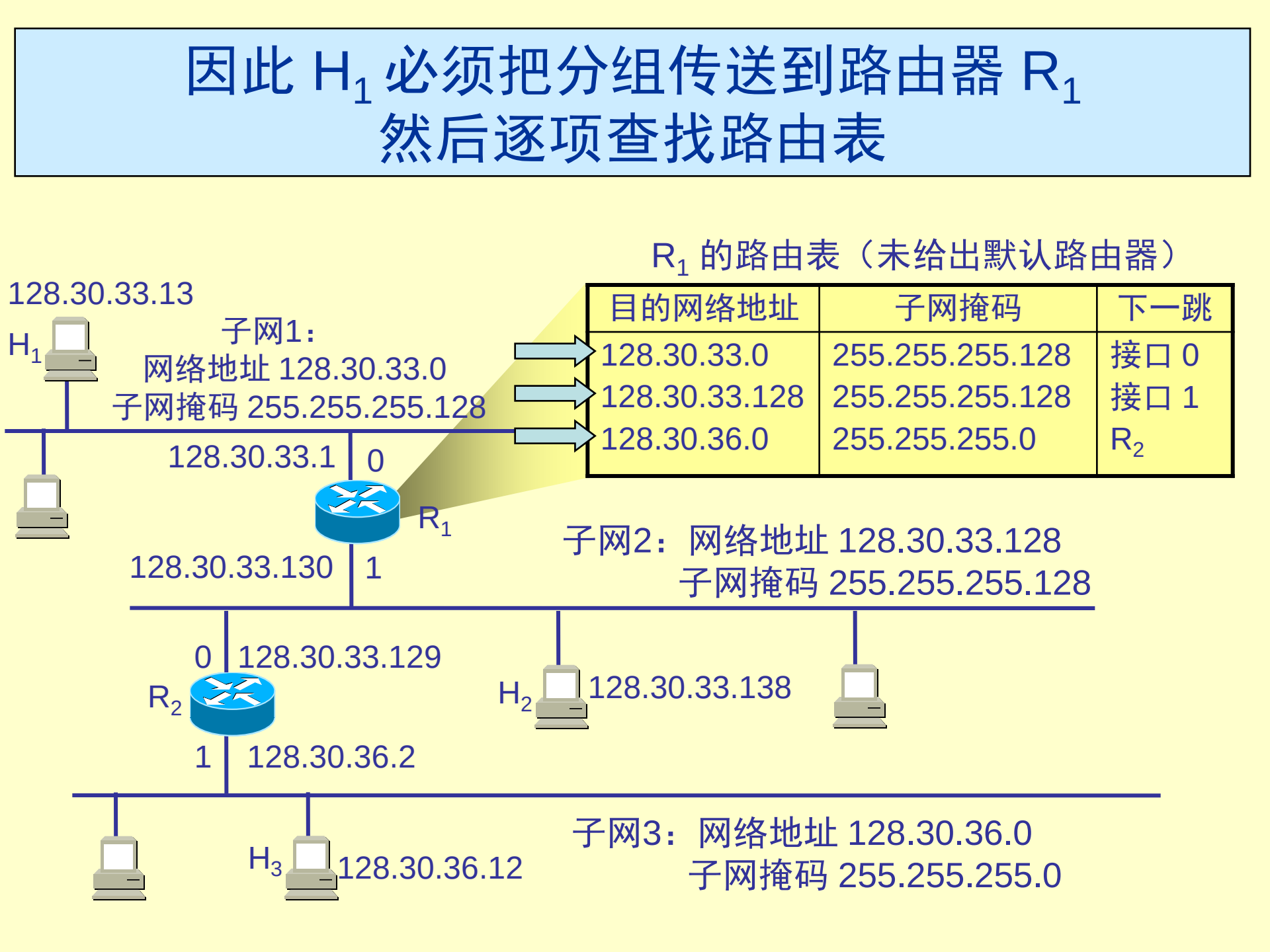
255 就是二进制的全 1，因此 255 AND xyz = xyz，
这里只需计算最后的 128 AND 138 即可。





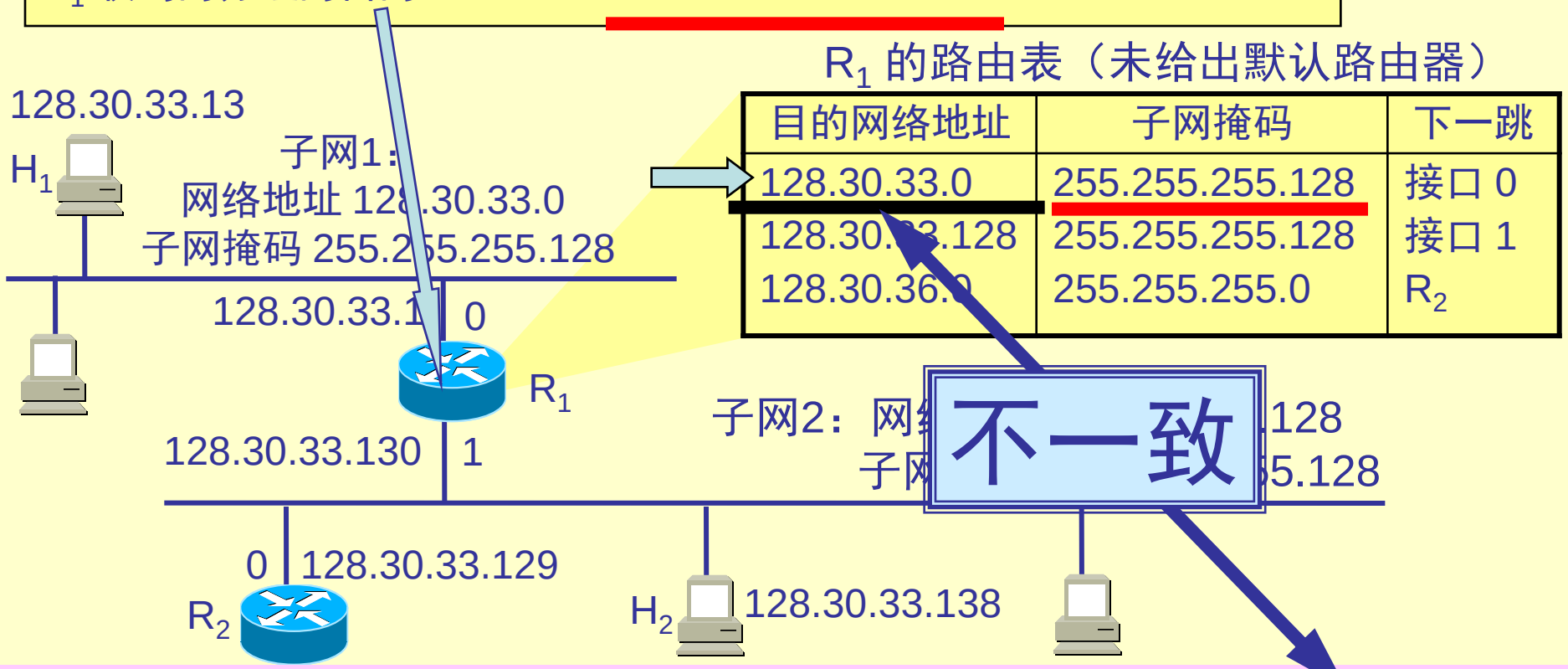
≠H_1 的网络地址
128.30.33.0

因此 H_1 必须把分组传送到路由器 R_1 然后逐项查找路由表



路由器 R_1 收到分组后就用路由表中第 1 个项目的子网掩码和 128.30.33.138 逐比特 AND 操作

R_1 收到的分组的目的 IP 地址：128.30.33.138



R_1 的路由表（未给出默认路由器）

目的网络地址	子网掩码	下一跳
128.30.33.0	255.255.255.128	接口 0
128.30.33.128	255.255.255.128	接口 1
128.30.36.0	255.255.255.0	R_2

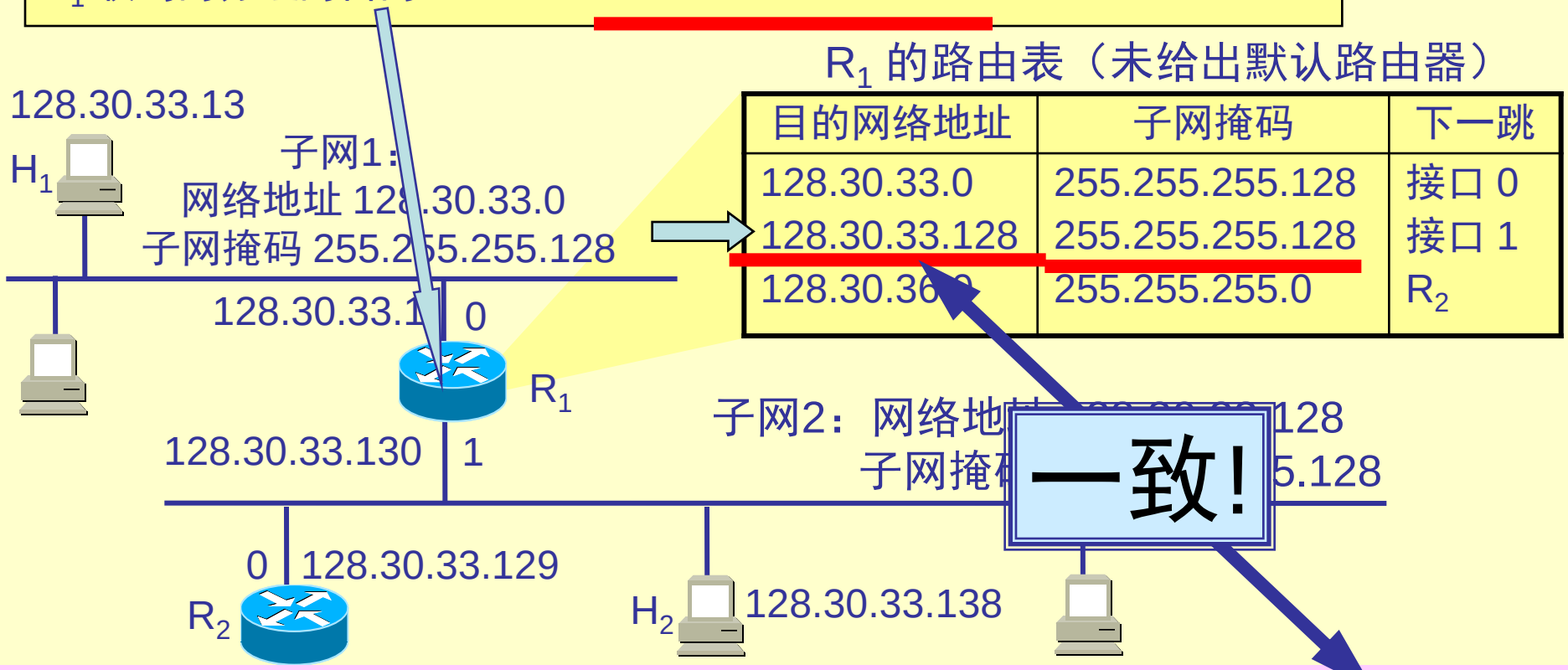
255.255.255.128 AND 128.30.33.138 = 128.30.33.128

不匹配!

（因为128.30.33.128 与路由表中的 128.30.33.0 不一致）

路由器 R_1 再用路由表中第 2 个项目的子网掩码和 128.30.33.138 逐比特 AND 操作

R_1 收到的分组的目的 IP 地址：128.30.33.138



R_1 的路由表（未给出默认路由器）

目的网络地址	子网掩码	下一跳
128.30.33.0	255.255.255.128	接口 0
128.30.33.128	255.255.255.128	接口 1
128.30.36.0	255.255.255.0	R_2

255.255.255.128 AND 128.30.33.138 = 128.30.33.128

匹配!

这表明子网 2 就是收到的分组所要寻找的目的网络

- IP地址分配
  - IP地址的分配必须由国际组织统一分配，以保持IP地址惟一性，避免IP地址冲突
  - 分配A类（最高一级）IP地址的国际组织是国际网络信息中心NIC（Network Information Center）。它负责分配A类IP地址，并授权分配B类IP地址的组织—自治区系统。它有权重新刷新IP地址

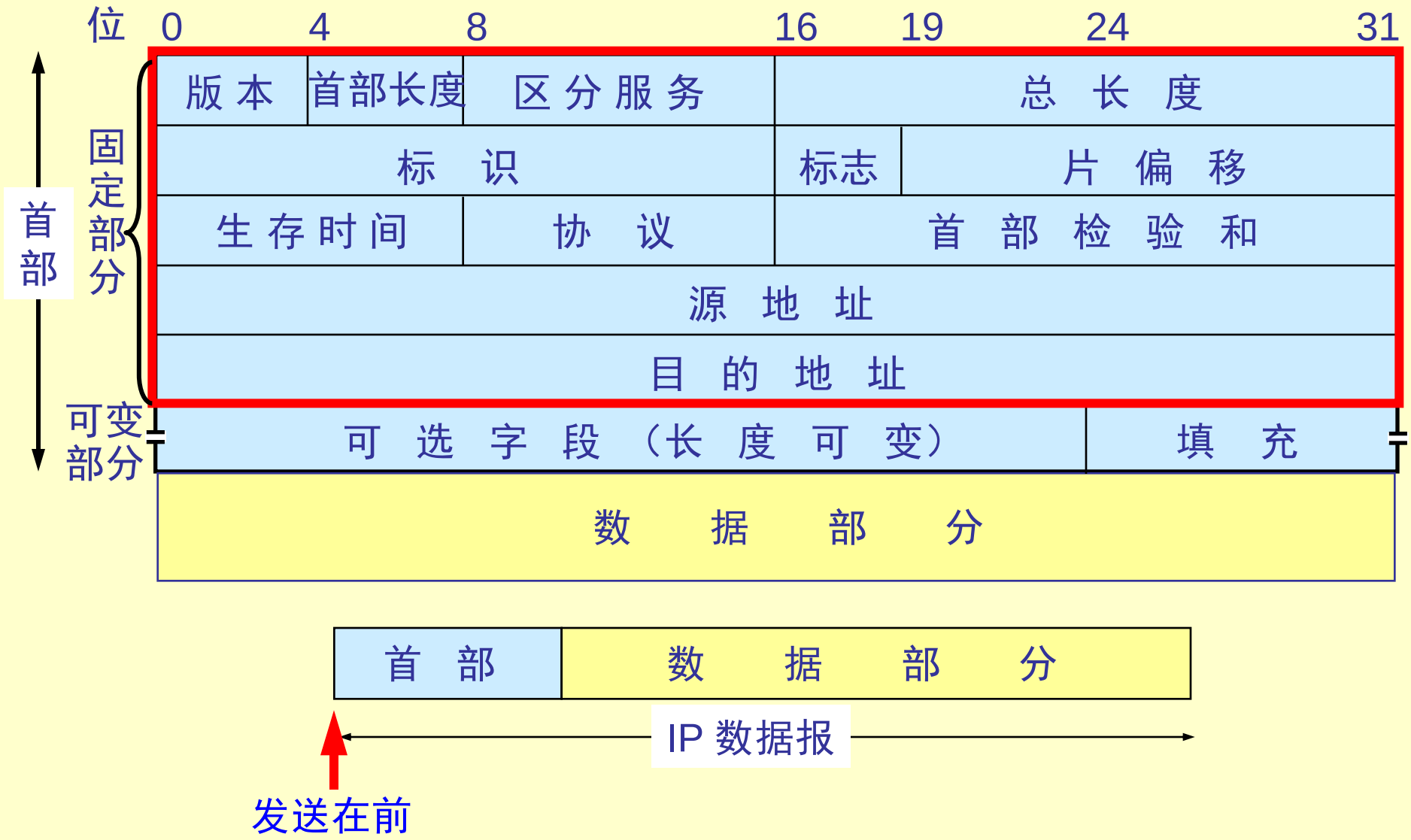
5.4.2 网际协议(IP)

- IP地址分配
  - 分配B类IP地址的组织是InterNIC、APNIC和ENIC。ENIC负责欧洲地址分配工作；InterNIC负责北美地区；而APNIC负责亚太地区，设在日本东京大学。我国属于APNIC，由它来分配B类地址。例如，APNIC给中国CERNET分配了10个B类地址
  - 分配C类IP地址的组织是国家或地区网络的NIC。例如CERNET的NIC设在清华大学，CERNET各地区的网管中心需向CERNET NIC申请分配C类地址

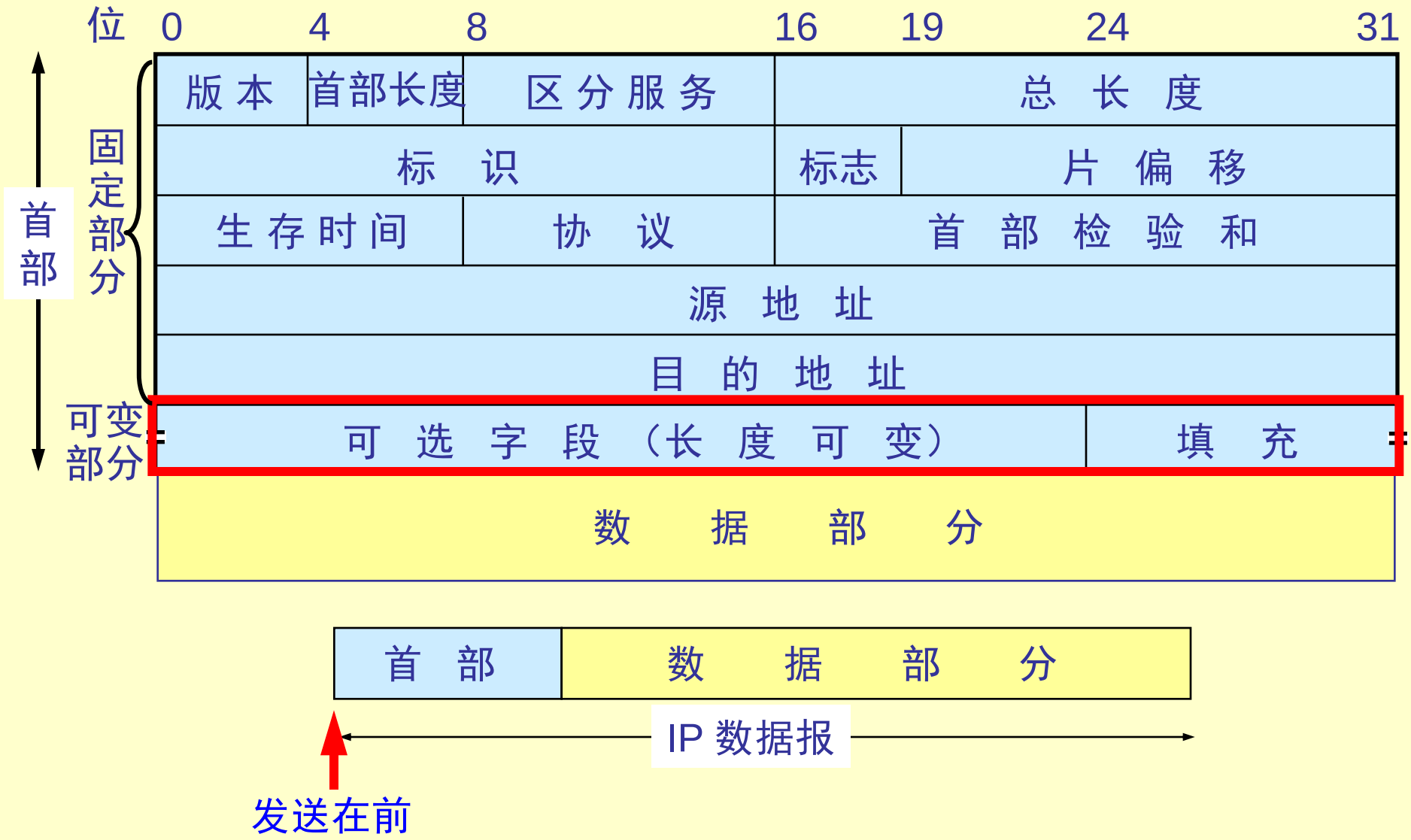
IP数据报

- 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


位
0
4
8
16
19
24
31
首部
固定部分
可变部分
版本
首部长度
区分服务
总长度
标识
标志
片偏移
生存时间
协议
首部检验和
源地址
目的地址
可选字段（长度可变）
填充
数据部分
首部
数据部分
IP 数据报
发送在前


位
0
4
8
16
19
24
31
首部
固定部分
版本
首部长度
区分服务
总长度
标识
标志
片偏移
生存时间
协议
首部检验和
源地址
目的地址
可变部分
可选字段（长度可变）
填充
数据部分

首部
数据部分
IP 数据报
发送在前


位
0
4
8
16
19
24
31
首部
固定部分
可变部分
版本
首部长度
区分服务
总长度
标识
标志
片偏移
生存时间
协议
首部检验和
源地址
目的地址
可选字段（长度可变）
填充
数据部分
首部
数据部分
IP 数据报
发送在前

1. IP 数据报首部的固定部分中的各字段

版本——占 4 位，指 IP 协议的版本
目前的 IP 协议版本号为 4 (即 IPv4)

首部长度——占 4 位，可表示的最大数值是 15 个单位(一个单位为 4 字节)

因此 IP 的首部长度的最大值是 60 字节。

区分服务——占 8 位，用来获得更好的服务

在旧标准中叫做服务类型，但实际上一直未被使用过。

1998 年这个字段改名为区分服务。

只有在使用区分服务（DiffServ）时，这个字段才起作用。

在一般的情况下都不使用这个字段

总长度——占 16 位，指首部和数据之和的长度，单位为字节，因此数据报的最大长度为 65535 字节。传输时总长度必须不超过数据帧最大传送单元 MTU。（以太网MTU为1500字节，FDDI为4500字节）

标识(identification)　占 16 位，
它是一个计数器，用来产生数据报的标识。

标志(flag)　占 3 位，目前只有前两位有意义。

标志字段的最低位是 **MF** (More Fragment)。

MF = 1 表示后面“还有分片”。MF = 0 表示最后一个分片。

标志字段中间的一位是 **DF** (Don't Fragment) 。

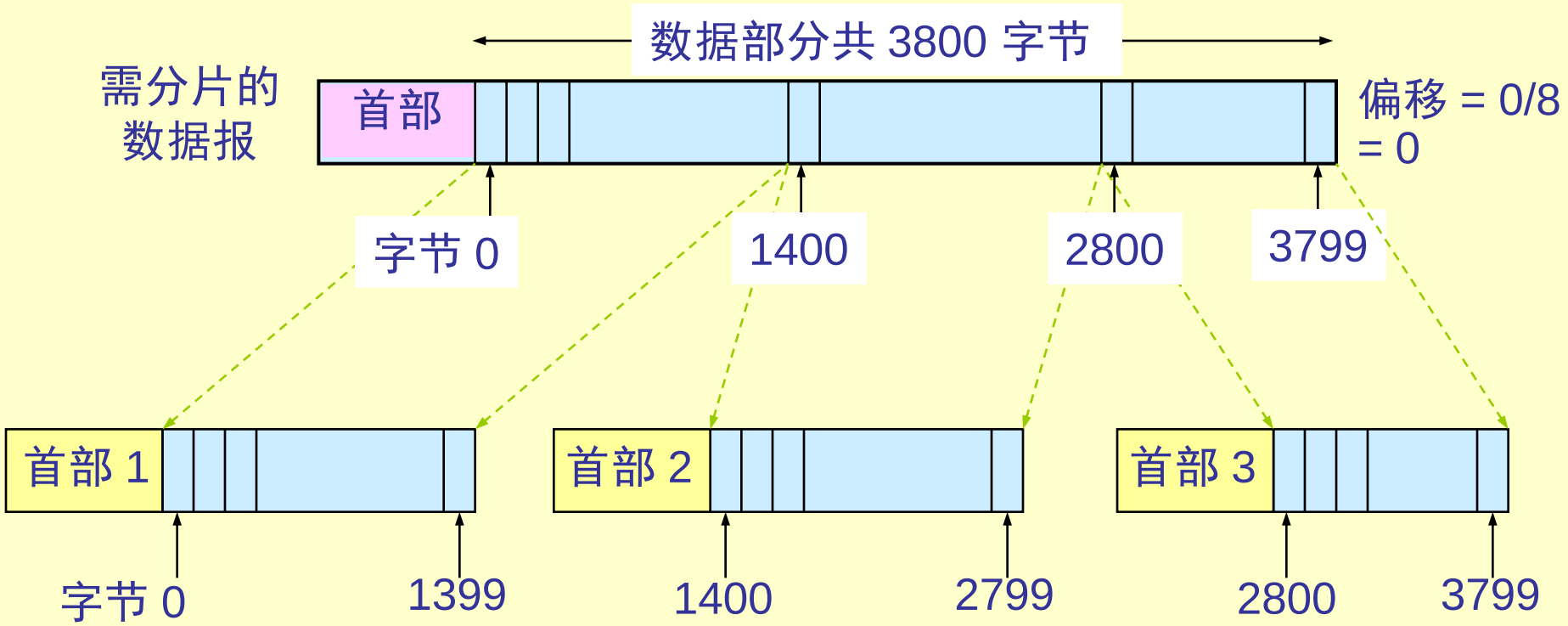
只有当 DF = 0 时才允许分片。

片偏移(13 位)指出：较长的分组在分片后某片在原分组中的相对位置。

片偏移以 8 个字节为偏移单位。

片应在MTU范围内尽可能大，偏移量应为8的整数倍。

【例】 IP 数据报分片



数据报片 1

标志=001
偏移 = 0/8 = 0

数据报片 2

标志=001
偏移 = 1400/8 = 175

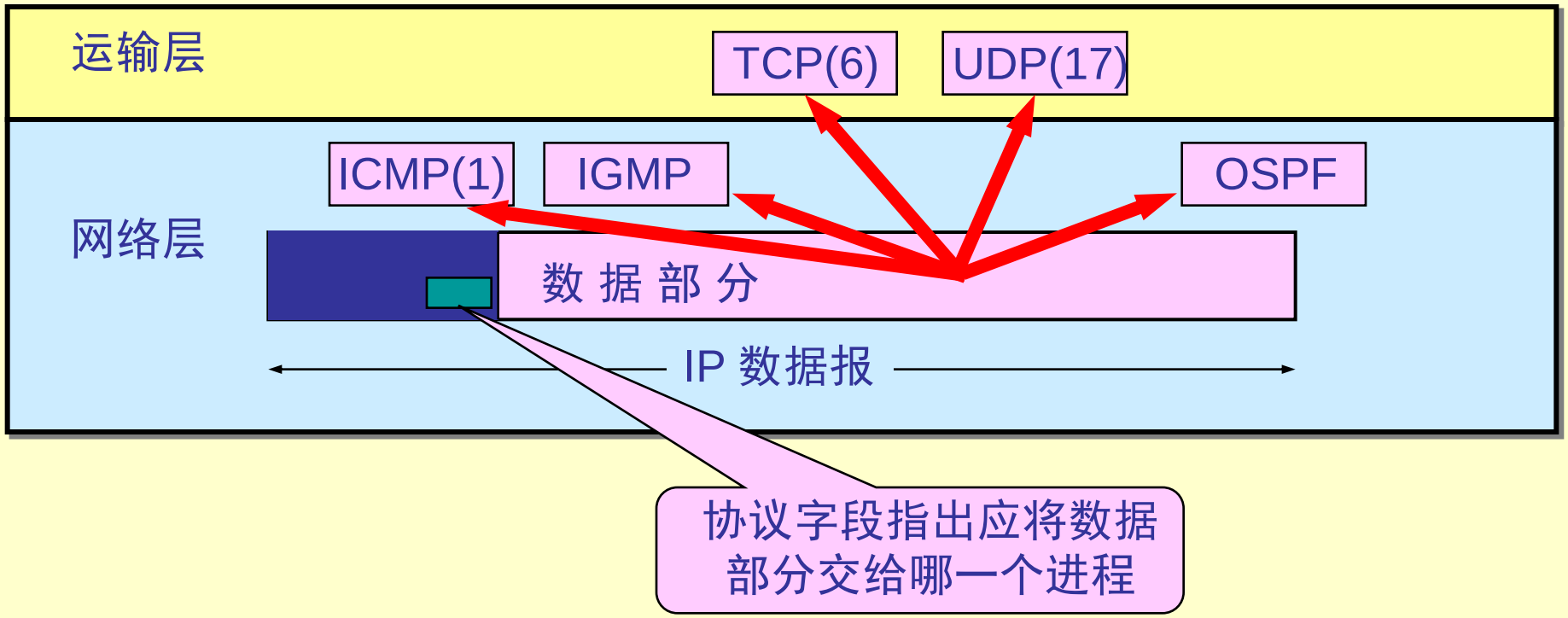
数据报片 3

标志=000
偏移 = 2800/8 = 350

生存时间(8 位)记为 TTL (Time To Live)

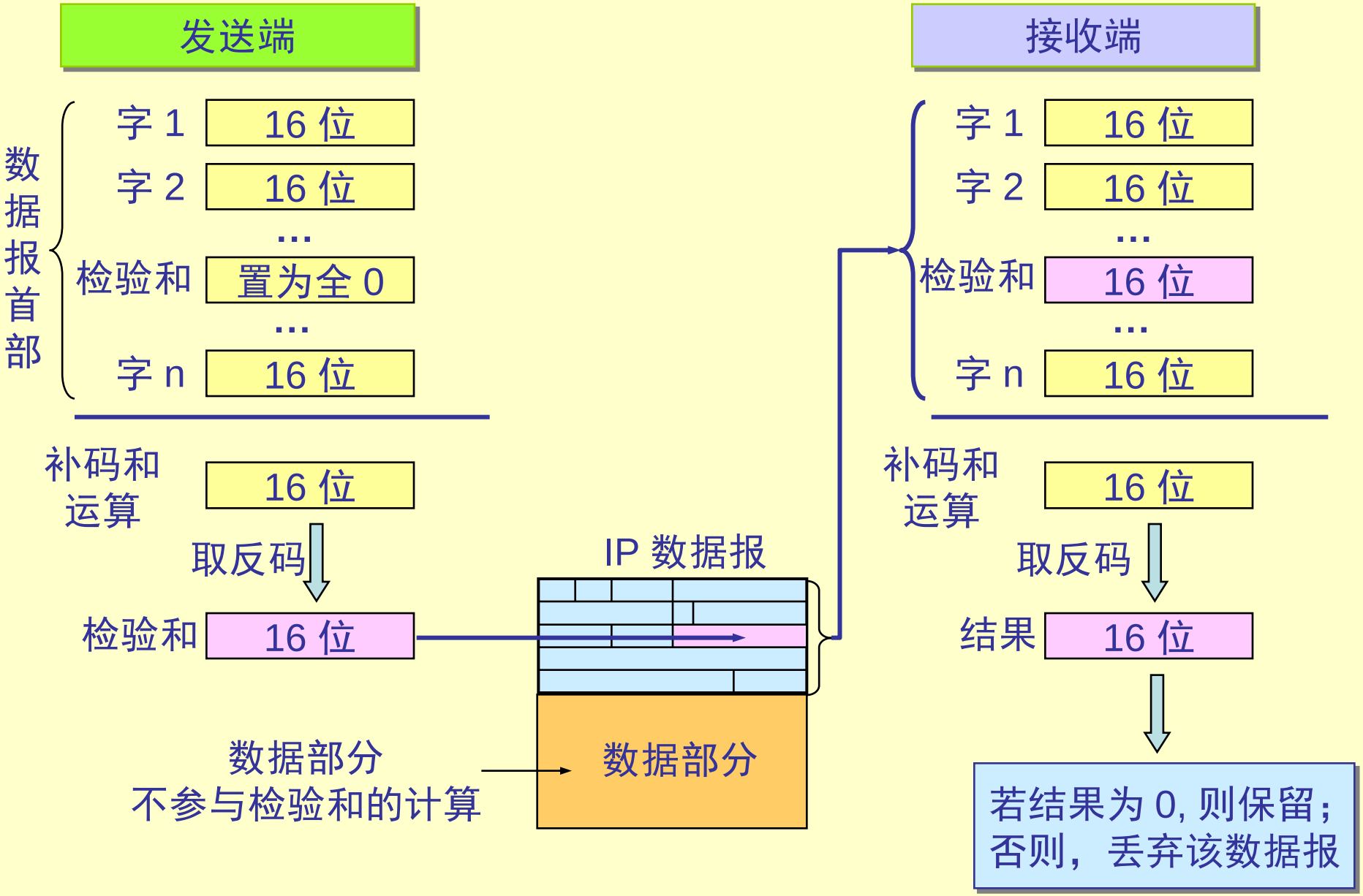
数据报在网络中可通过的路由器数的最大值。

协议(8 位)字段指出此数据报携带的数据使用何种协议以便目的主机的 IP 层将数据部分上交给哪个处理过程


运输层
TCP(6)
UDP(17)
网络层
ICMP(1)
IGMP
OSPF
数 据 部 分
IP 数据报
协议字段指出应将数据
部分交给哪一个进程

首部检验和(16 位)字段只检验数据报的首部
不检验数据部分。

这里不采用 CRC 检验码而采用简单的计算方法。



补码和运算：带循环进位的加法

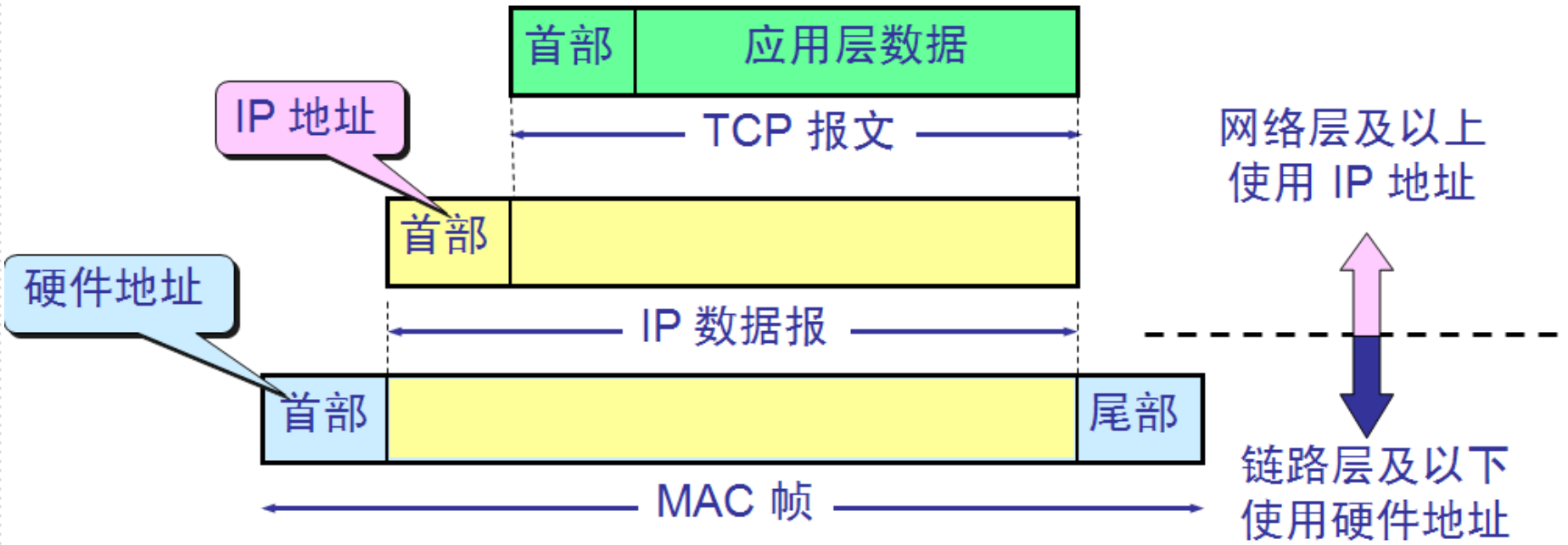
反码：二进制各位取反

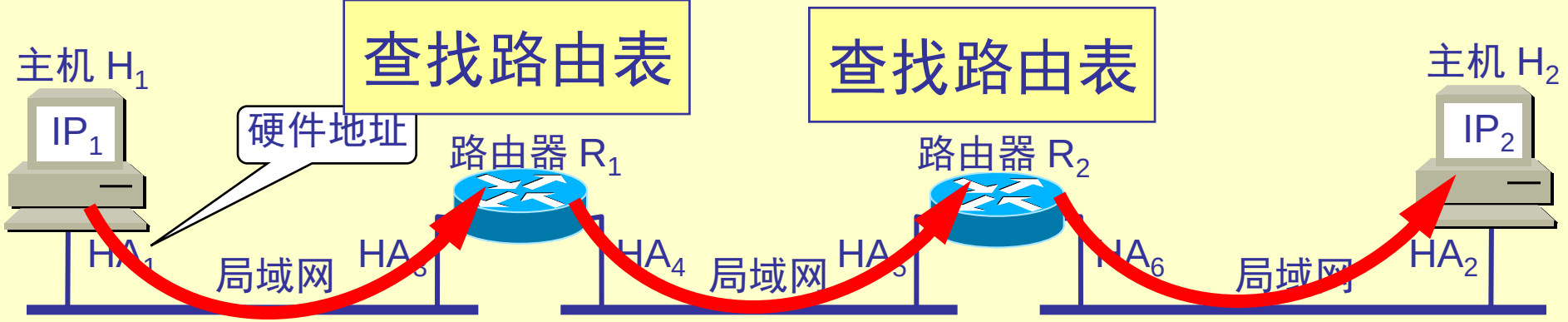
源地址和目的地址都各占 4 字节

- IP数据报首部的可变部分
  - IP首部的可变部分就是一个选项字段，用来支持排错、测量以及安全等措施，内容很丰富
  - 选项字段的长度可变，从1个字节到40个字节不等，取决于所选择的项目
  - 增加首部的可变部分是为了增加IP数据报的功能，但这同时也使得IP数据报的首部长度成为可变的。这就增加了每一个路由器处理数据报的开销
  - 实际上这些选项很少被使用

IP数据报传输过程

- IP地址与硬件地址

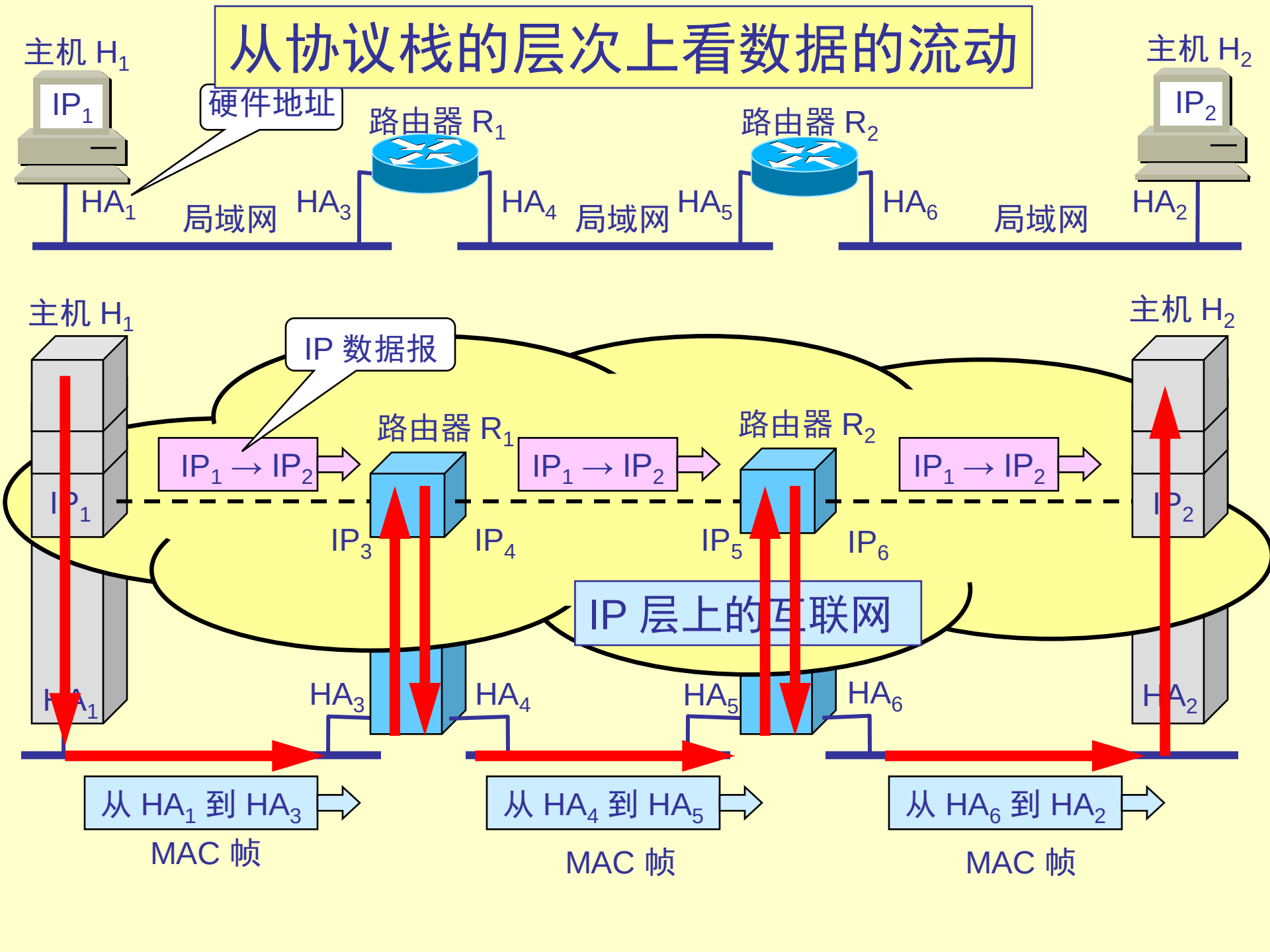


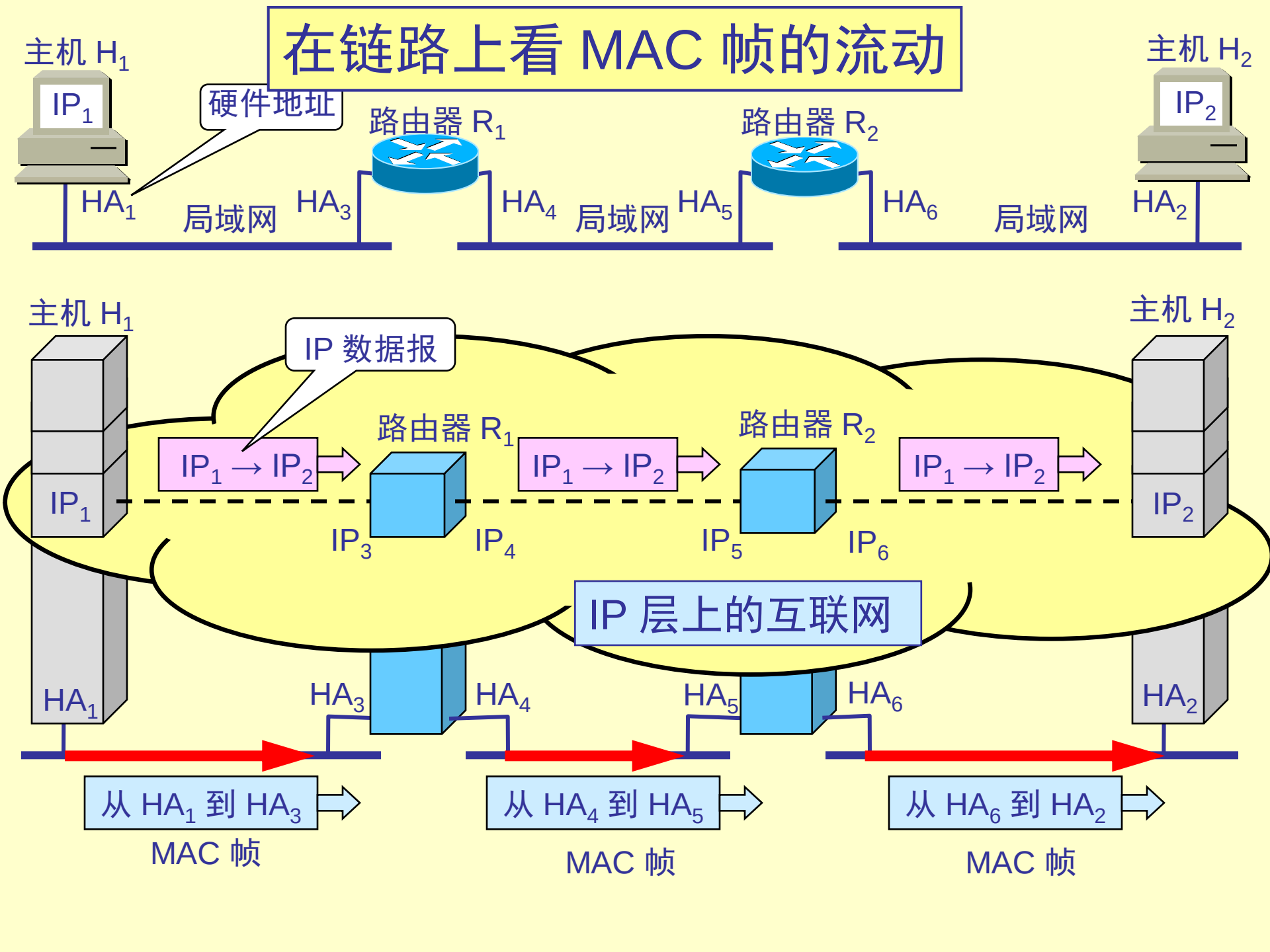


通信的路径

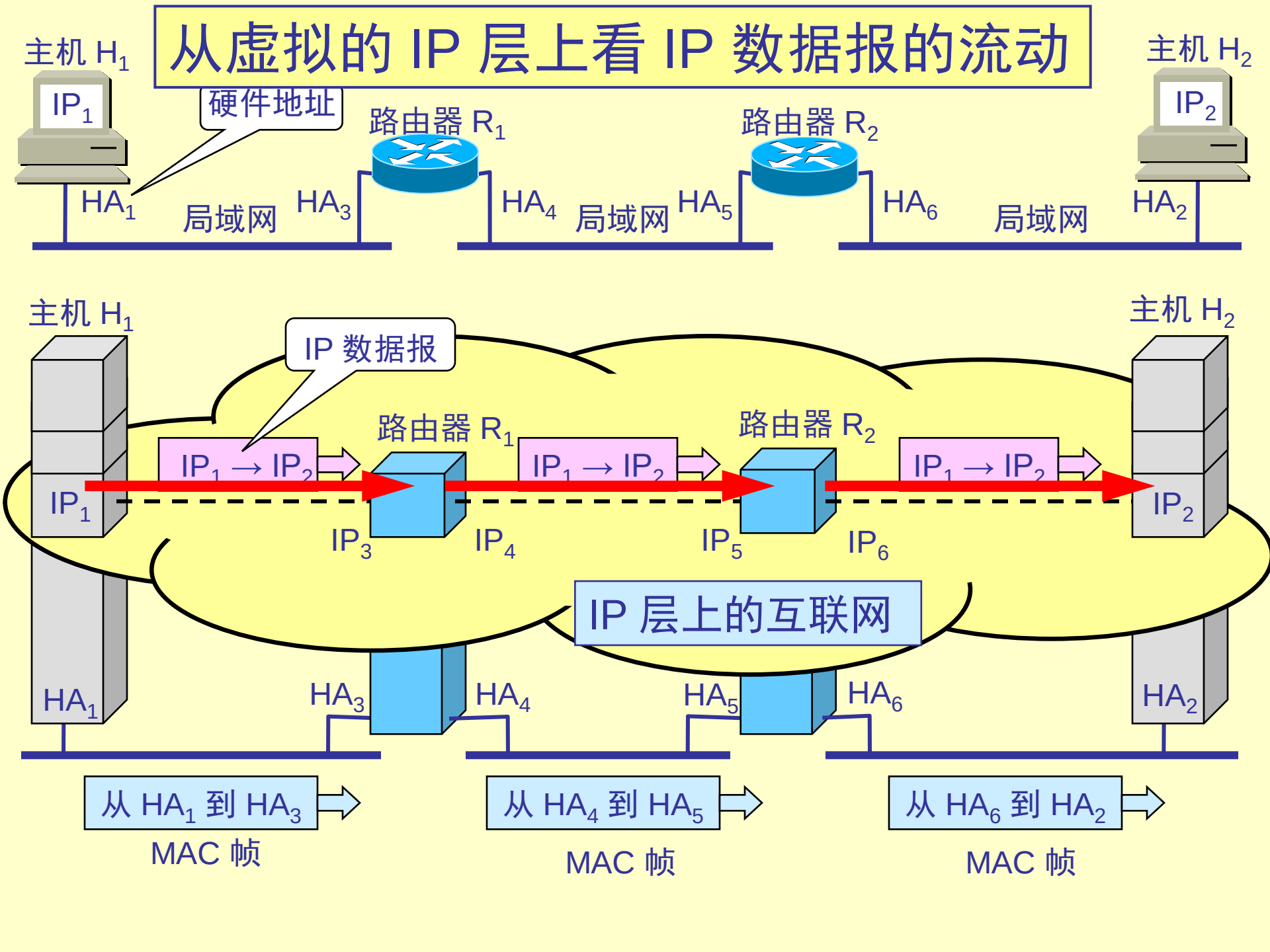
H_1→经过 R_1 转发→再经过 R_2 转发→H_2

从协议栈的层次上看数据的流动




在链路上看 MAC 帧的流动
主机 H1
IP1
硬件地址
HA1
局域网
路由器 R1
HA3
HA4
局域网
路由器 R2
HA5
HA6
局域网
主机 H2
IP2
HA2
主机 H1
IP 数据报
IP1 → IP2
路由器 R1
IP1 → IP2
路由器 R2
IP1 → IP2
主机 H2
IP1
IP3
IP4
IP5
IP6
IP2
IP 层上的互联网
HA1
HA3
HA4
HA5
HA6
HA2
从 HA1 到 HA3
MAC 帧
从 HA4 到 HA5
MAC 帧
从 HA6 到 HA2
MAC 帧

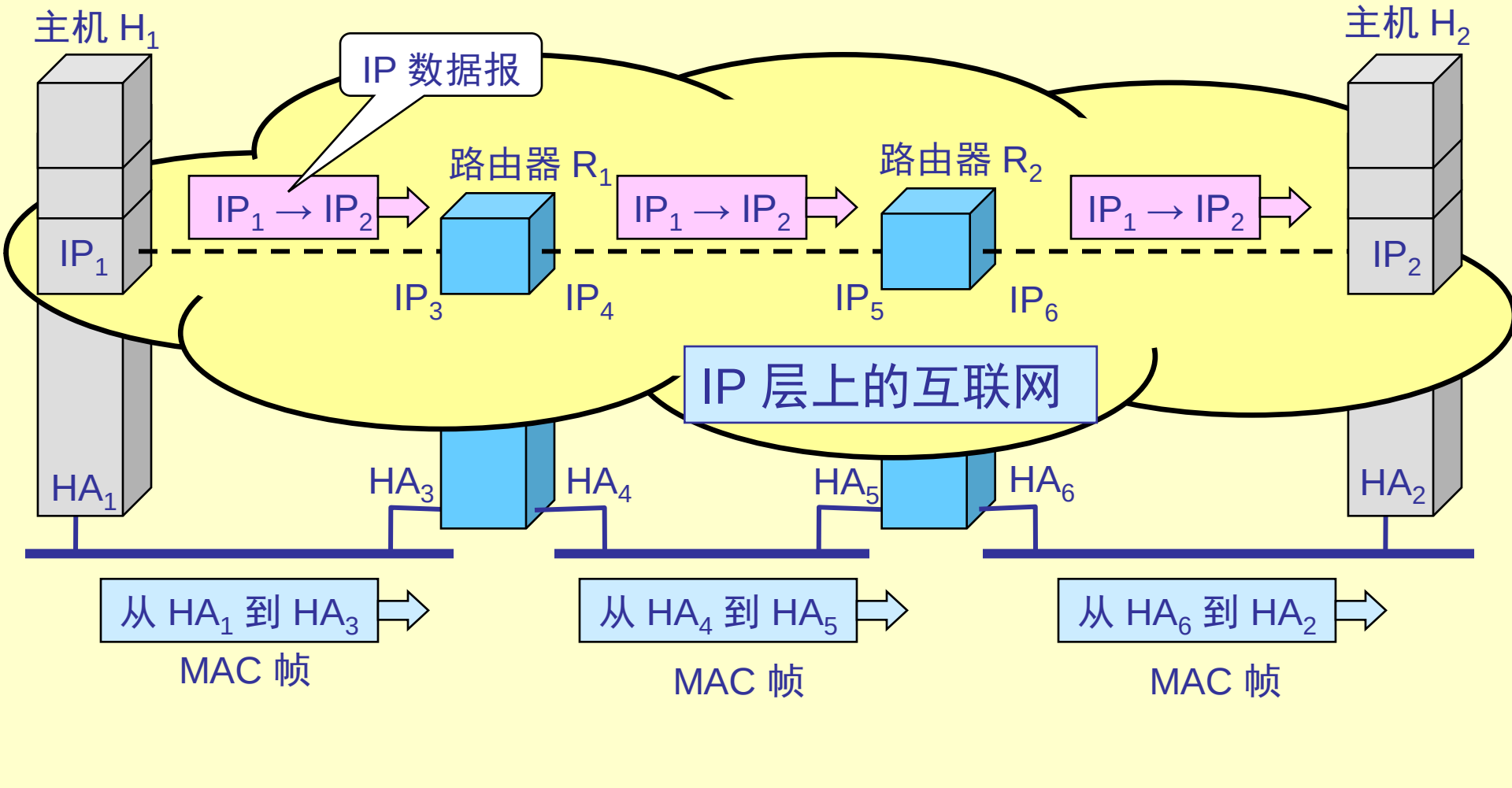
从虚拟的 IP 层上看 IP 数据报的流动



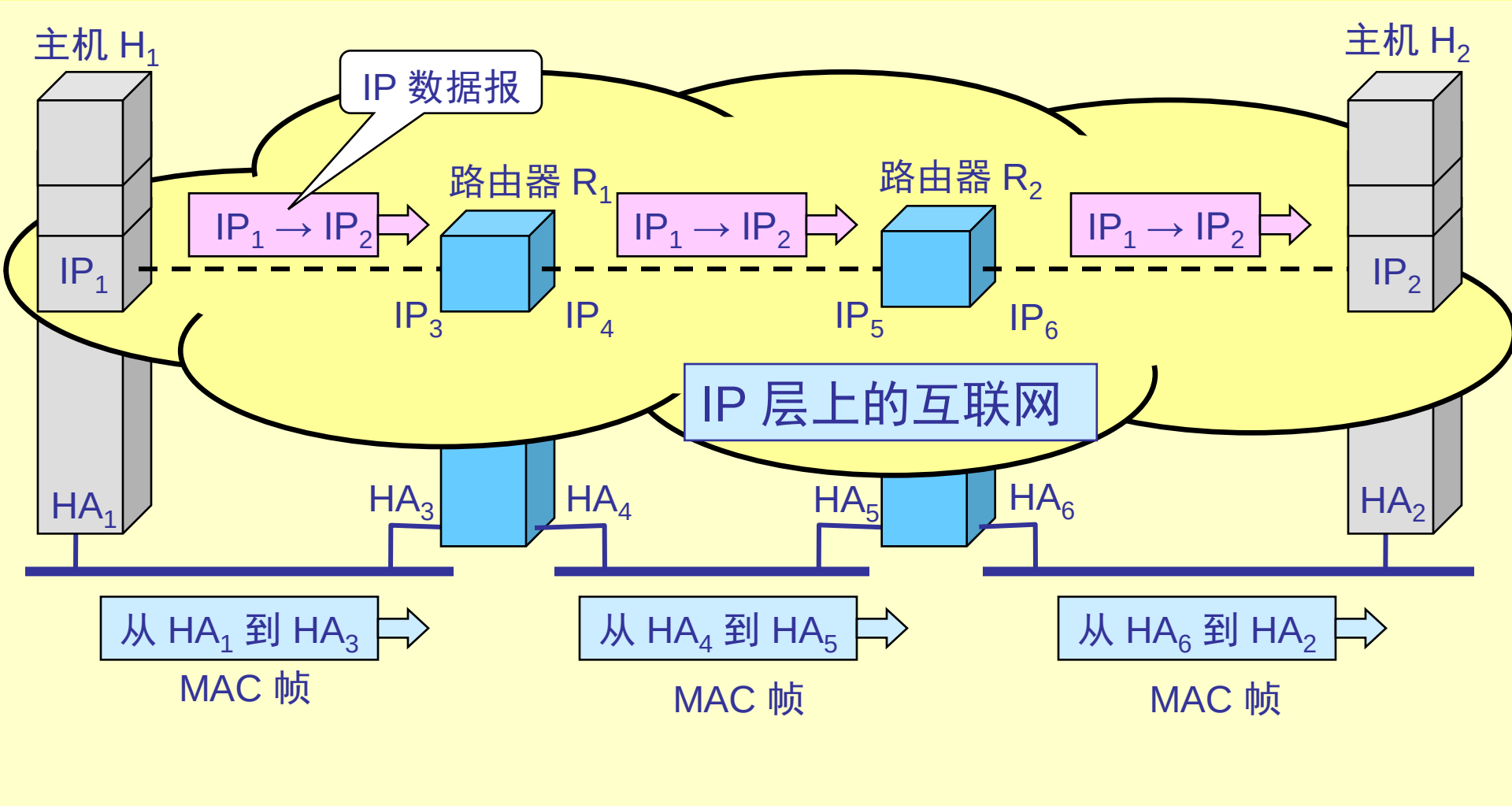
在 IP 层抽象的互联网上只能看到 IP 数据报

图中的 $IP_1 \rightarrow IP_2$ 表示从源地址 IP_1 到目的地址 IP_2

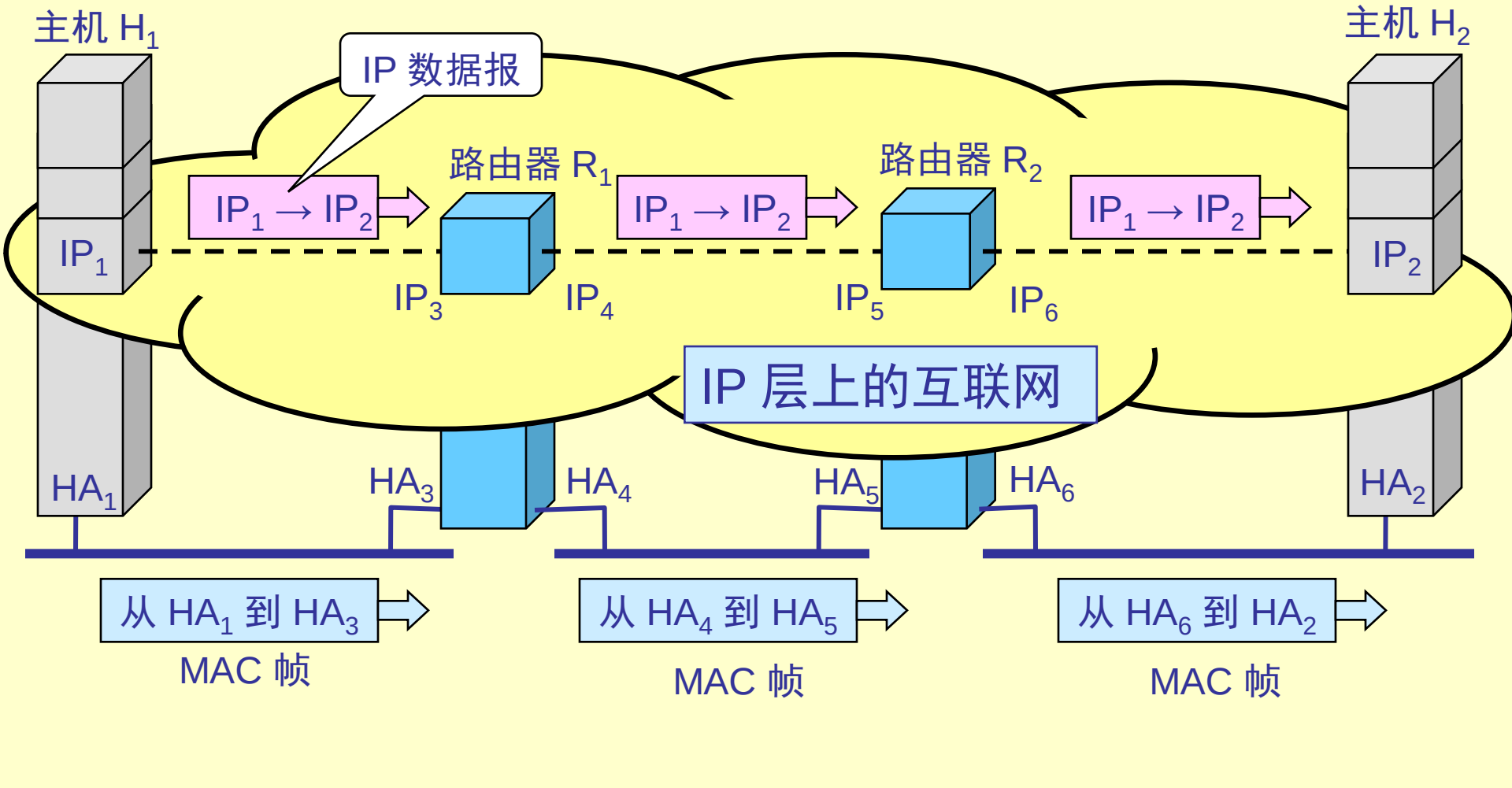
两个路由器的 IP 地址并不出现在 IP 数据报的首部中



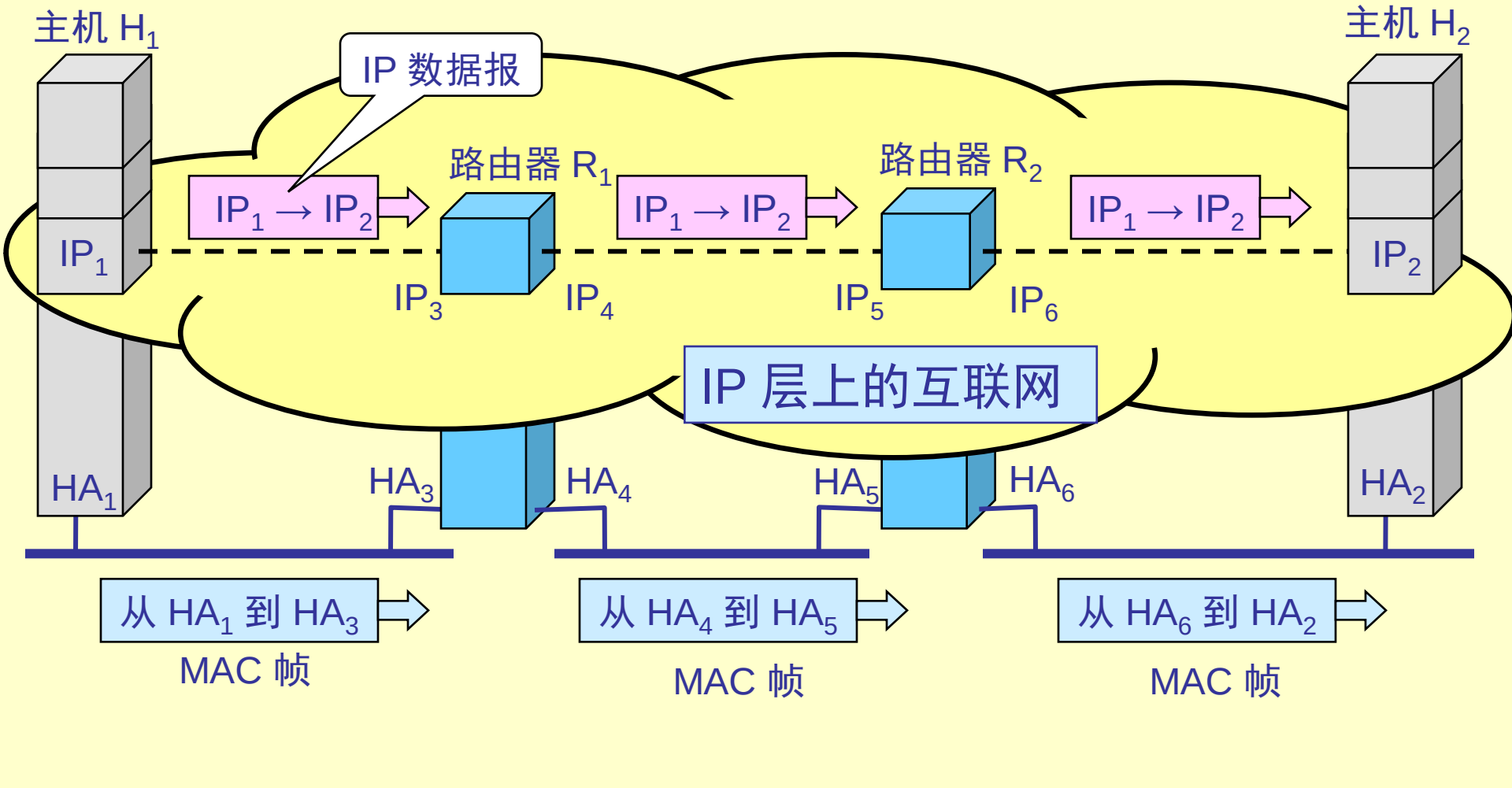
路由器只根据目的站的 IP 地址的网络号进行路由选择



在具体的物理网络的链路层
只能看见 MAC 帧而看不见 IP 数据报
会出现途经的主机及路由器的MAC地址

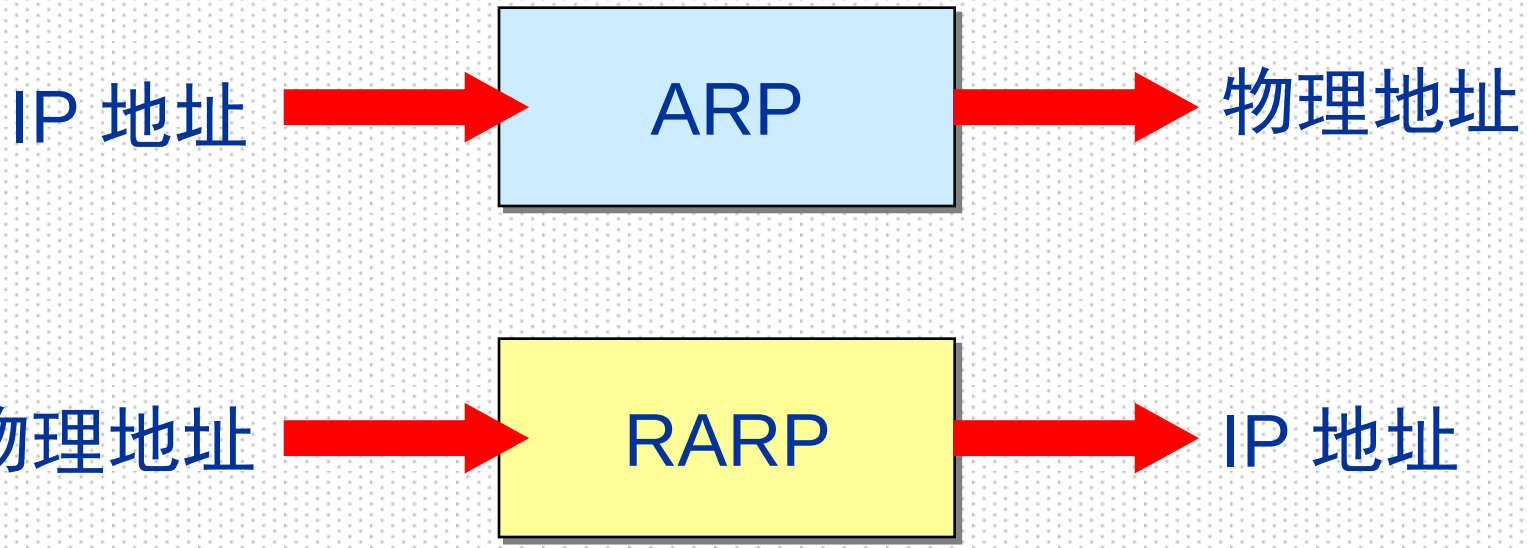


IP层抽象的互联网屏蔽了下层很复杂的细节
在抽象的网络层上讨论问题，就能够使用
统一的、抽象的 IP 地址
研究主机和主机或主机和路由器之间的通信



5.4.3 地址解析协议(ARP)

- 地址解析协议ARP和逆地址解析协议RARP

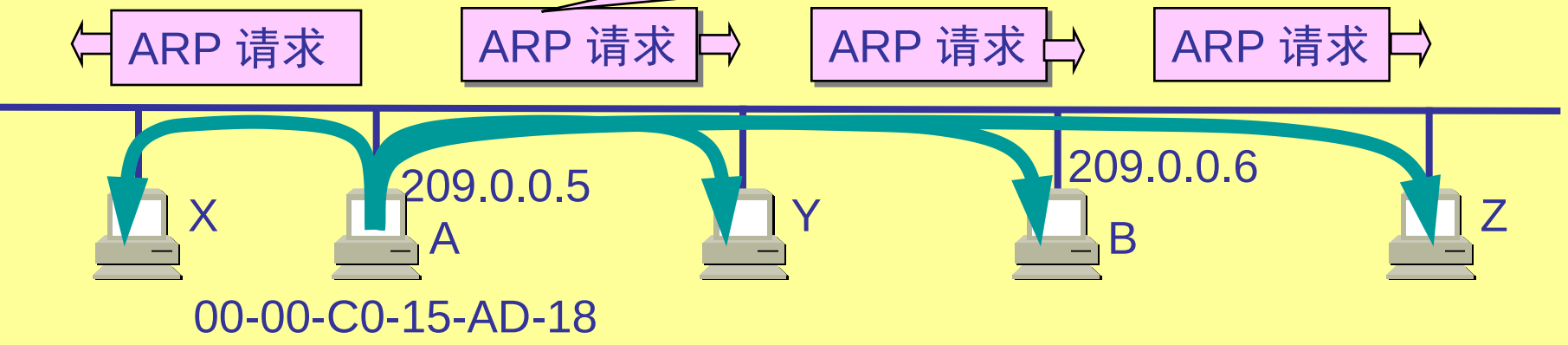


5.4.3 地址解析协议(ARP)

- **地址解析协议ARP**
  - 不管网络层使用什么协议，在实际网络链路上传送数据帧时，最终还是必须使用硬件地址
  - 每一个主机都设有一个ARP高速缓存，里面有所在局域网上各主机和路由器的 IP 地址到硬件地址的映射表
  - 当主机A欲向本局域网上某个主机B发送IP数据报时，就先在其ARP高速缓存中查看有无主机B的IP地址。如有，就可查出其对应的硬件地址，再将此地址写入MAC帧，然后通过局域网将该MAC帧发往此硬件地址

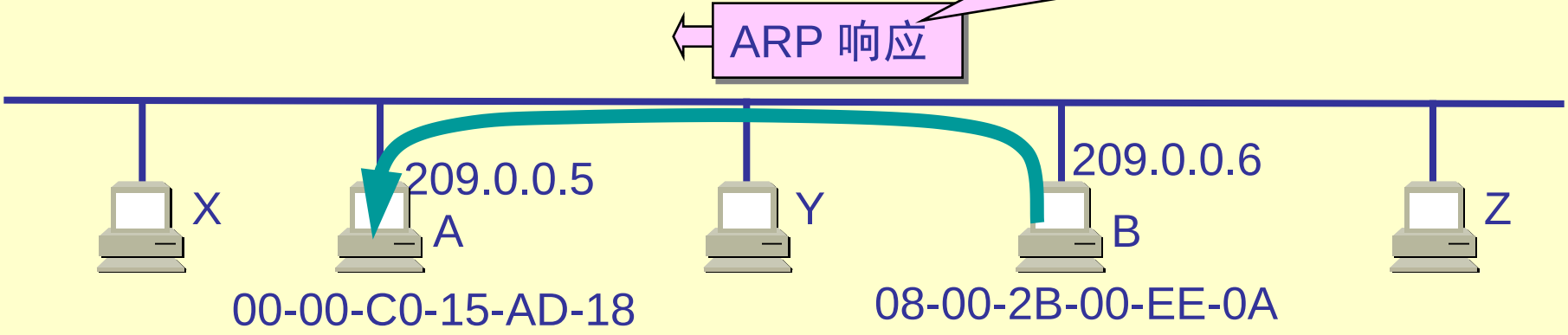
主机 A 广播发送
ARP 请求分组

我是 209.0.0.5，硬件地址是 00-00-C0-15-AD-18
我想知道主机 209.0.0.6 的硬件地址



主机 B 向 A 发送
ARP 响应分组

我是 209.0.0.6
硬件地址是 08-00-2B-00-EE-0A



5.4.3 地址解析协议(ARP)

- ARP高速缓存的作用
  - 为了减少网络上的通信量，主机A在发送其ARP请求分组时，就将自己的IP地址到硬件地址的映射写入ARP请求分组
  - 当主机B收到A的ARP请求分组时，就将主机A的这一地址映射写入主机B自己的ARP高速缓存中。这对主机B以后向A发送数据报提供了方便

5.4.3 地址解析协议(ARP)

- 应当注意的问题
  - ARP是解决同一个局域网上的主机或路由器的IP地址和硬件地址的映射问题
  - 如果所要找的主机和源主机不在同一个局域网上，那么就要通过ARP找到一个位于本局域网上的某个路由器的硬件地址，然后把分组发送给这个路由器，让这个路由器把分组转发给下一个网络。剩下的工作就由下一个网络来做

5.4.3 地址解析协议(ARP)

- 应当注意的问题
  - 从IP地址到硬件地址的解析是自动进行的，主机的用户对这种地址解析过程是不知道的
  - 只要主机或路由器要和本网络上的另一个已知IP地址的主机或路由器进行通信，ARP协议就会自动地将该IP地址解析为链路层所需要的硬件地址
  - 屏蔽了复杂的硬件地址转换，给用户带来方便

使用ARP的四种典型情况

- 发送方是主机，要把IP数据报发送到本网络上的另一个主机。这时用ARP找到目的主机的硬件地址
- 发送方是主机，要把IP数据报发送到另一个网络上的一个主机。这时用ARP找到本网络上的一个路由器的硬件地址，剩下的工作由这个路由器来完成
- 发送方是路由器，要把IP数据报转发到本网络上的一个主机。这时用ARP找到目的主机的硬件地址
- 发送方是路由器，要把IP数据报转发到另一个网络上的一个主机。这时用ARP找到距目的网络更近的一个路由器的硬件地址。剩下的工作由这个路由器来完成

5.4.4 逆地址解析协议(RARP)

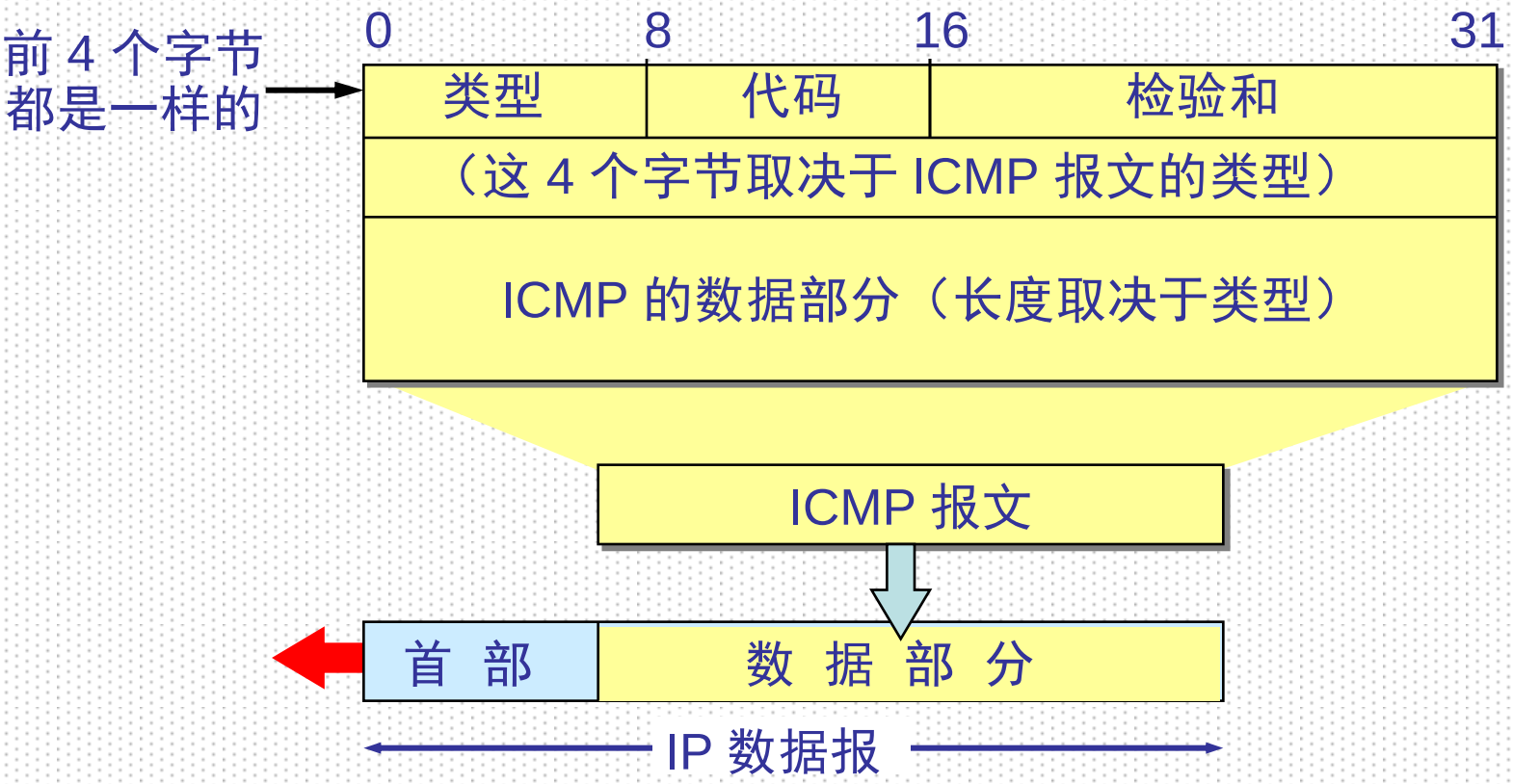
- 逆地址解析协议RARP使只知道自己硬件地址的主机能够知道其IP地址
- 这种主机往往是无盘工作站
- 因此RARP协议目前已很少使用

5.4.5 因特网控制报文协议(ICMP)

- 网际控制报文协议ICMP
  - IP是一种不可靠的传输协议，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

5.4.5 因特网控制报文协议(ICMP)

- ICMP报文的格式



5.4.5 因特网控制报文协议(ICMP)

- ICMP报文的种类
  - ICMP报文的前4个字节是统一的格式，共有三个字段：即类型、代码和检验和。接着的4个字节的内容与ICMP的类型有关

5.4.5 因特网控制报文协议(ICMP)

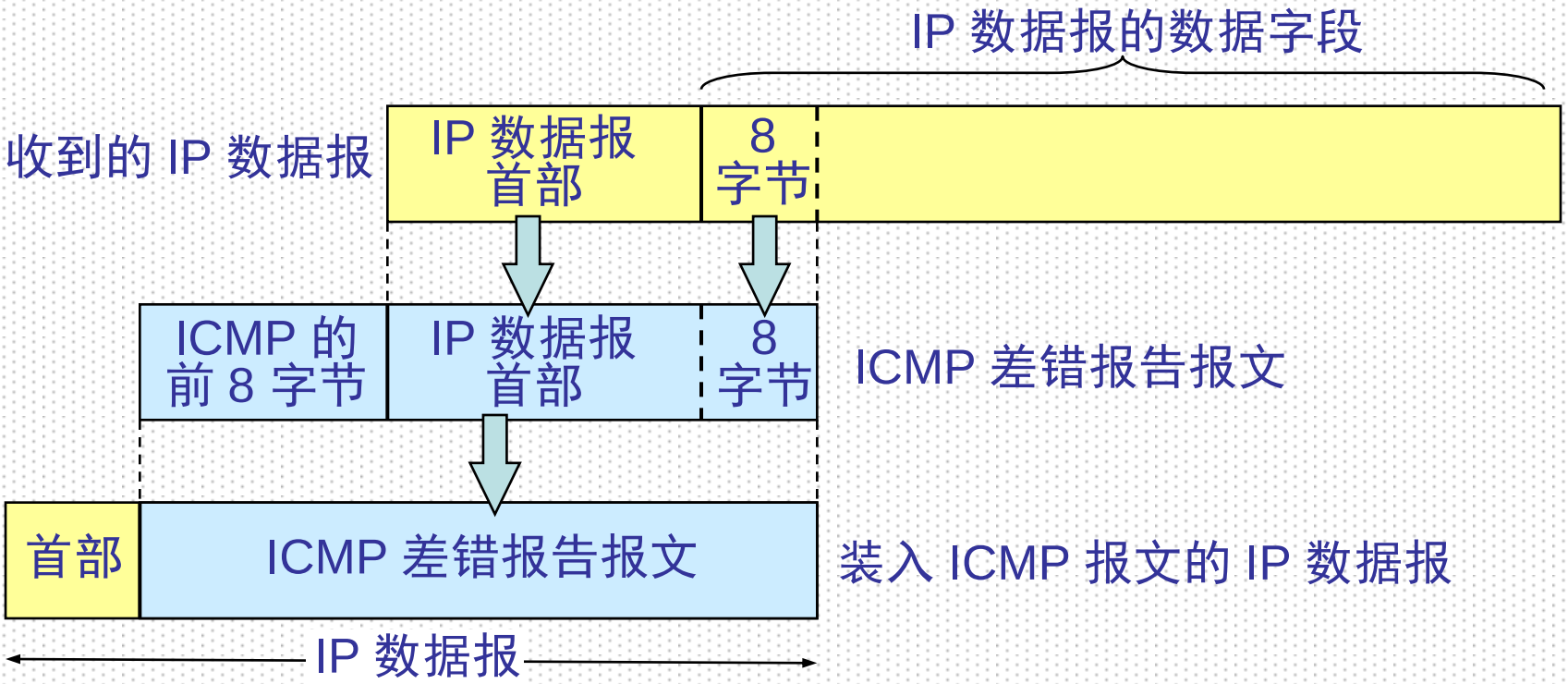
- ICMP差错报告报文
  - 终点不可达：目的站或路由器不能交付数据报时向源站发送终点不可达报文
  - 时间超过：目的站或路由器收到生存时间TTL=0或分片重组超时的数据报时向源站发送时间超过报文
  - 参数出错：目的站或路由器收到的数据报首部中有字段的值不正确时向源站发送参数出错报文

仅提供差错报告，伴随抛弃出错数据报

目的站或路由器→源站，无法解决中间路由器引起的错误

5.4.5 因特网控制报文协议(ICMP)

- ICMP差错报告报文的数据字段的内容



5.4.5 因特网控制报文协议(ICMP)

- ICMP控制报文
  - 源点抑制(Source quench)：路由器或主机由于拥塞而丢弃数据报时，向源点发送源点抑制报文，使源点知道应放慢数据报发送速率
  - 改变路由（重定向）(Redirect) ：路由器向主机发送改变路由报文，使主机得知下次应将数据报发送给另外的路由（更优路径）

5.4.5 因特网控制报文协议(ICMP)

- ICMP询问报文有两种
  - 回应请求/应答报文：用于测试目的站的可达性
  - 时间戳请求/应答报文：用于时钟同步
- 下面的几种ICMP报文不再使用
  - 掩码地址请求/应答报文
  - 信息请求/应答报文
  - 路由器询问和通告报文

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回应请求与回应应答报文
  - 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

5.4.5 因特网控制报文协议(ICMP)

- PING的应用举例

```
C:\Documents and Settings\XXR>ping mail.sina.com.cn

Pinging mail.sina.com.cn [202.108.43.230] with 32 bytes of data:

Reply from 202.108.43.230: bytes=32 time=368ms TTL=242
Reply from 202.108.43.230: bytes=32 time=374ms TTL=242
Request timed out.
Reply from 202.108.43.230: bytes=32 time=374ms TTL=242

Ping statistics for 202.108.43.230:
    Packets: Sent = 4, Received = 3, Lost = 1 (25% loss),
Approximate round trip times in milli-seconds:
    Minimum = 368ms, Maximum = 374ms, Average = 372ms
```

5.4.6 路由选择协议(RP)

- 路由选择协议的作用
  - 交换路由信息、执行路由选择算法、更新路由表
- 理想的路由算法
  - 算法必须是正确的和完整的
  - 算法在计算上应简单
  - 算法应能适应通信量和网络拓扑的变化，有自适应性
  - 算法应具有稳定性
  - 算法应是公平的
  - 算法应是最佳的

5.4.6 路由选择协议(RP)

- 路由度量指标
  - 距离：路由的长度
  - 跳数：路由所经过的路由器数目
  - 时延：分组由源站到达目的站所花费的时间
  - 费用：借助电信等部门的通信线路需交纳的费用
  - 可靠性：链路的误码率

5.4.6 路由选择协议(RP)

- 关于“最佳路由”
  - 不存在一种绝对的最佳路由算法
  - 所谓“最佳”只能是相对于某一种特定要求下得出的较为合理的选择而已
  - 实际的路由选择算法，应尽可能接近于理想的算法
  - 路由选择是个非常复杂的问题
    - 它是网络中所有结点共同协调工作的结果
    - 路由选择的环境往往是不断变化的，而这种变化有时无法事先知道

5.4.6 路由选择协议(RP)

- 分层次的路由选择协议
  - 因特网采用分层次的路由选择协议
  - 因特网的规模非常大。如果让所有的路由器知道所有的网络应怎样到达，则这种路由表将非常大，处理起来也太花时间，所有这些路由器之间交换路由信息所需的带宽就会使因特网的通信链路饱和
  - 许多单位不愿意外界了解自己单位网络的布局细节和本部门所采用的路由选择协议（这属于本部门内部的事情），但同时还希望连接到因特网上

5.4.6 路由选择协议(RP)

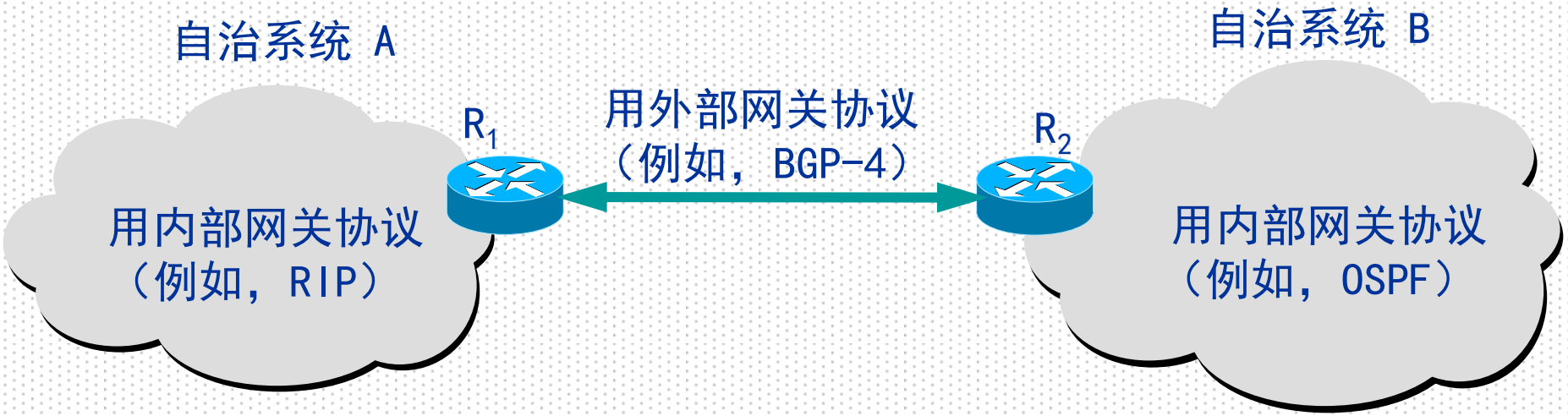
- **自治系统**AS(Autonomous System)
  - 自治系统定义：在单一的技术管理下的一组路由器，而这些路由器使用一种AS内部的路由选择协议和共同的度量以确定分组在该AS内的路由，同时还使用一种AS之间的路由选择协议用以确定分组在AS之间的路由
  - 现在对自治系统AS的定义是强调下面的事实：尽管一个AS使用了多种内部路由选择协议和度量，但重要的是一个AS对其他AS表现出的是一个单一的和一致的路由选择策略

5.4.6 路由选择协议(RP)

- 两类路由选择协议
  - 内部网关协议IGP (Interior Gateway Protocol)：即在一个自治系统内部使用的路由选择协议。目前这类路由选择协议使用得最多，如RIP和OSPF协议
  - 外部网关协议EGP (External Gateway Protocol)：若源站和目的站处在不同的自治系统中，当数据报传到一个自治系统的边界时，就需要使用一种协议将路由选择信息传递到另一个自治系统中。这样的协议就是外部网关协议EGP，目前使用最多的是BGP-4

5.4.6 路由选择协议(RP)

- 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)

5.4.7 内部网关协议(RIP)

- 路由信息协议RIP
  - 各路由器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，如每隔30秒
  - 使用距离矢量算法优化更新路由，处理的过程是一个分布式处理过程

5.4.7 内部网关协议(RIP)

- 路由表的建立
  - 路由器在刚刚开始工作时，只知道到直接连接的网络的距离（此距离定义为1）
  - 以后，每一个路由器也只和数目非常有限的相邻路由器交换并更新路由信息
  - 经过若干次更新后，所有路由器最终都会知道到达本自治系统中任何一个网络最短距离和下一跳路由器的地址
  - RIP协议的收敛(convergence)过程较快，即在自治系统中所有的结点都得到正确的路由选择信息的过程

距离矢量算法

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

临站(X)给出了本站不知道的路由

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。

经临站(X)去往某个目的网络的距离发生变化

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。

临站(X)知道去某个目的网络距离更短的路由

(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）。

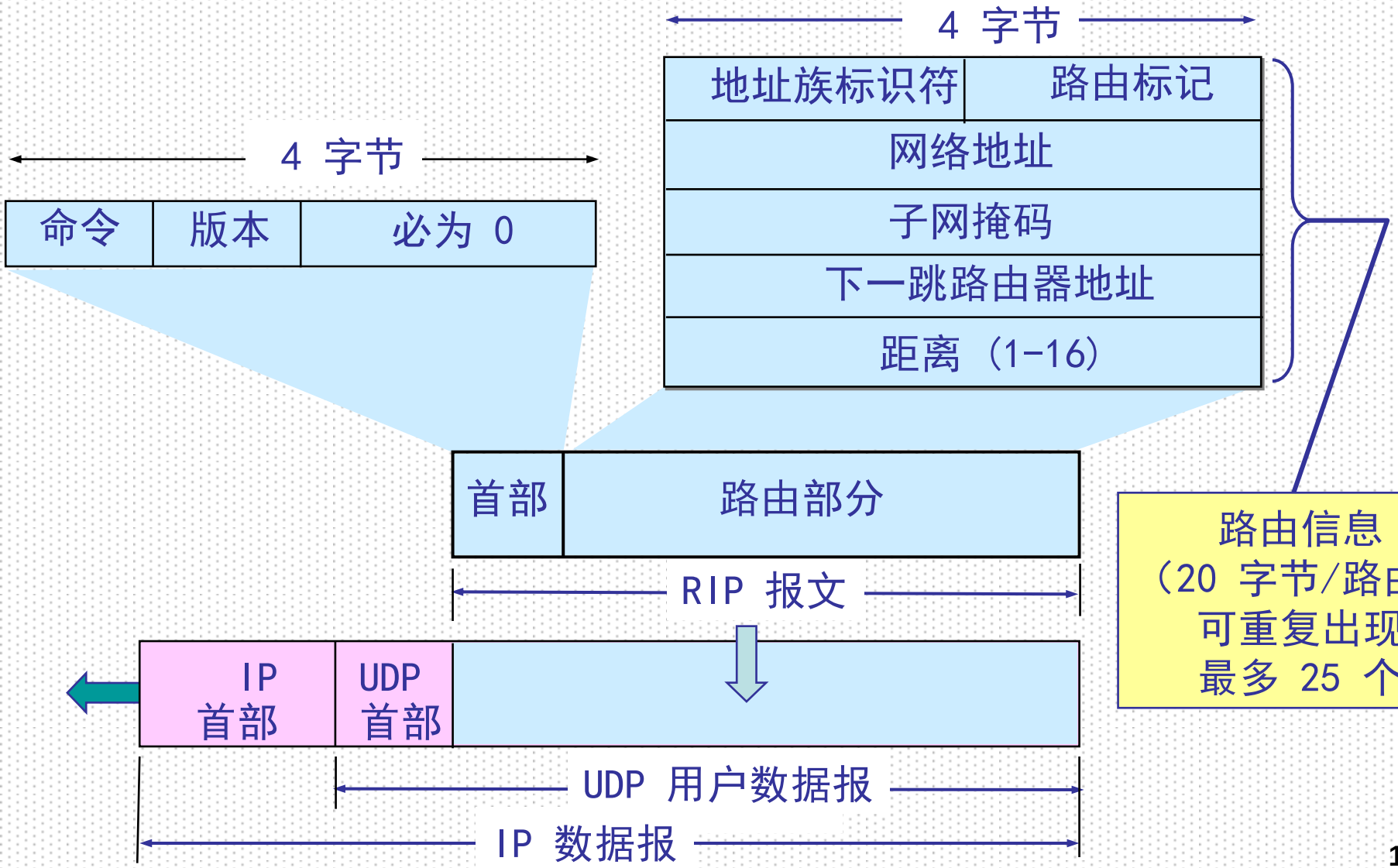
(4) 返回。

规定时间内未收到临站(X)的路由报文

5.4.7 内部网关协议(RIP)

- 路由器之间交换信息
  - RIP协议让互联网中的所有路由器都和自己的相邻路由器不断交换路由信息，并不断更新其路由表，使得从每一个路由器到每一个目的网络的路由都是最短的（即跳数最少）
  - 虽然所有的路由器最终都拥有了整个自治系统的全局路由信息，但由于每一个路由器的位置不同，它们的路由表当然也应当是不同的

RIP2协议的报文格式

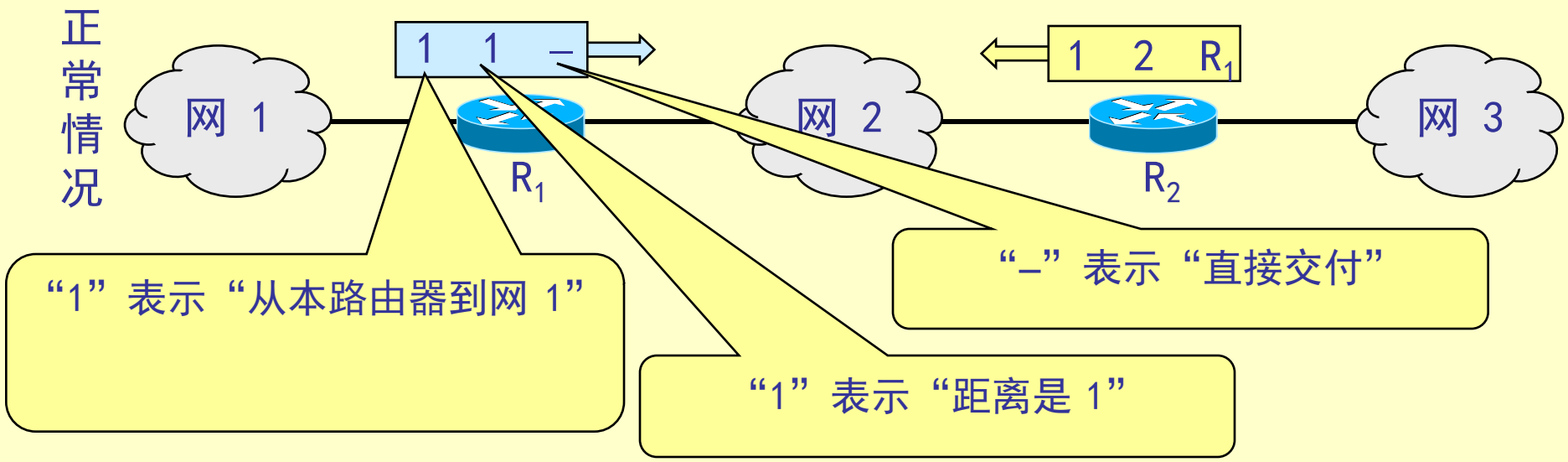


5.4.7 内部网关协议(RIP)

- RIP2的报文由首部和路由部分组成
  - RIP2报文中的路由部分由若干个路由信息组成。每个路由信息需要用20个字节。地址族标识符（又称为地址类别）字段用来标志所使用的地址协议
  - 路由标记填入自治系统的号码，这是考虑使RIP有可能收到本自治系统以外的路由选择信息。再后面指出某个网络地址、该网络的子网掩码、下一跳路由器地址以及到此网络的距离

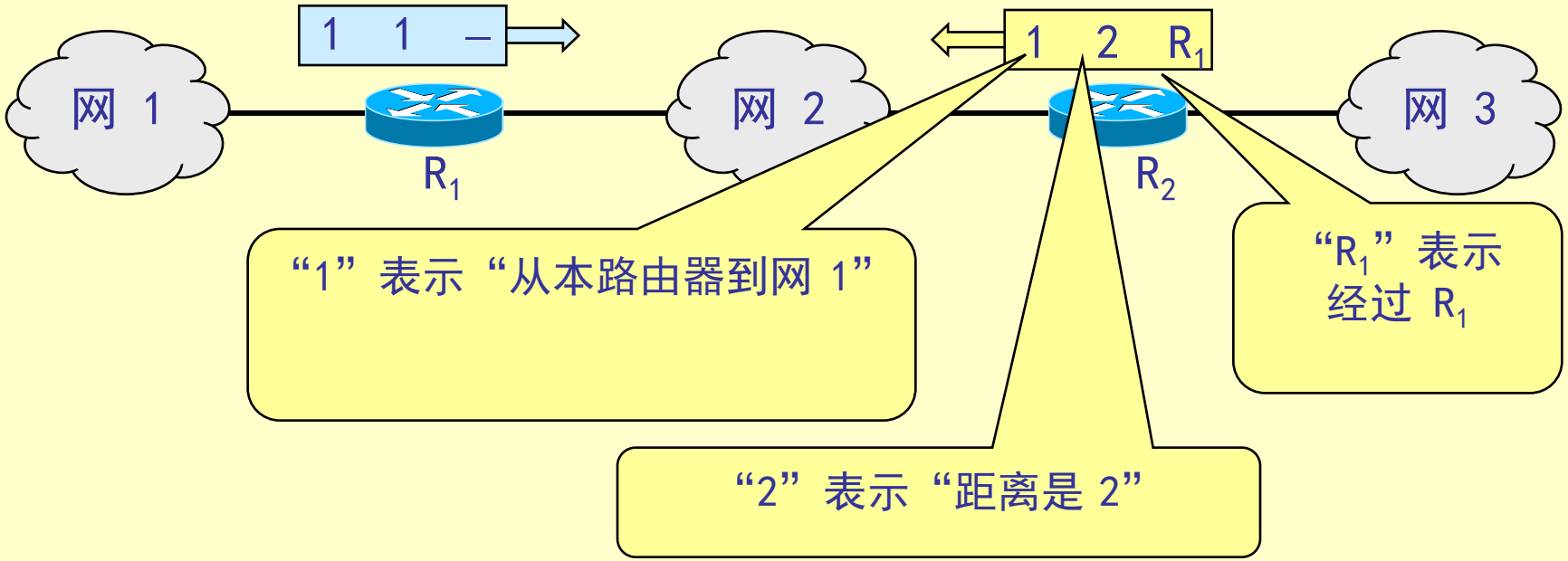
5.4.7 内部网关协议(RIP)

- RIP协议的优缺点
  - RIP协议最大的优点就是实现简单，开销较小
  - RIP存在的一个问题是当网络出现故障时，要经过比较长的时间才能将此信息传送到所有的路由器
  - RIP限制了网络的规模，它能使用的最大距离为15（16表示不可达）
  - 路由器之间交换的路由信息是路由器中的完整路由表，因而随着网络规模的扩大，开销也就增加



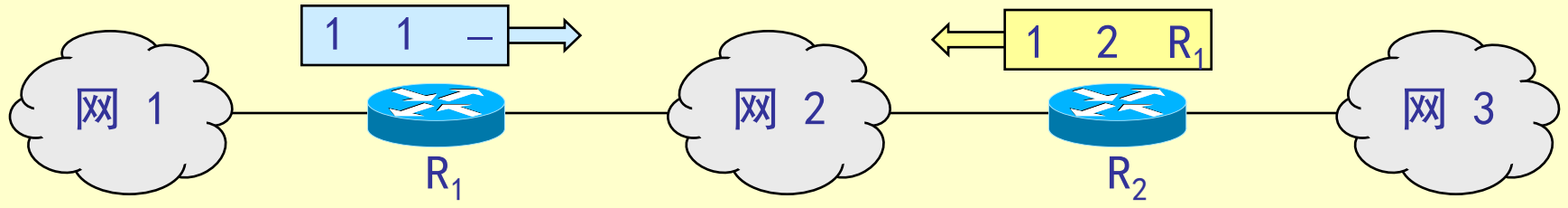
R_1 说：“我到网 1 的距离是 1，是直接交付。”

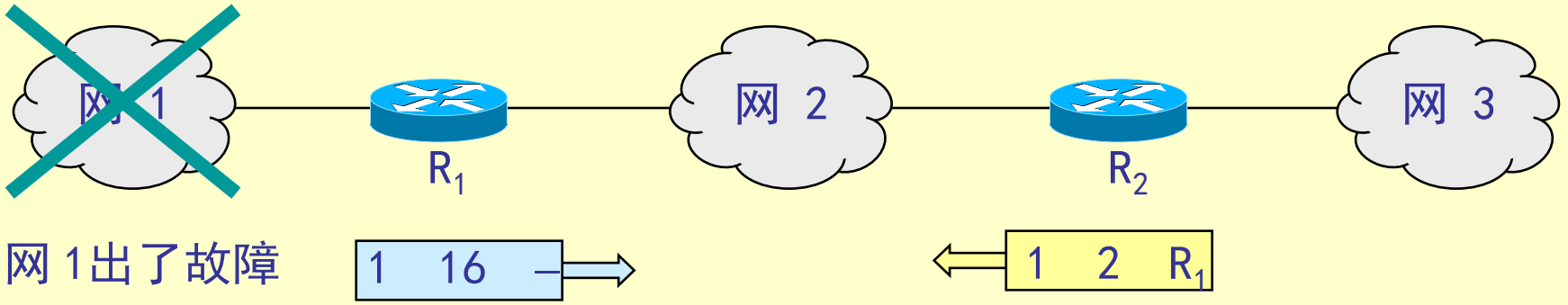
正常情况



R_2 说：“我到网 1 的距离是 2，是经过 R_1。”

正常情况





网 1出了故障

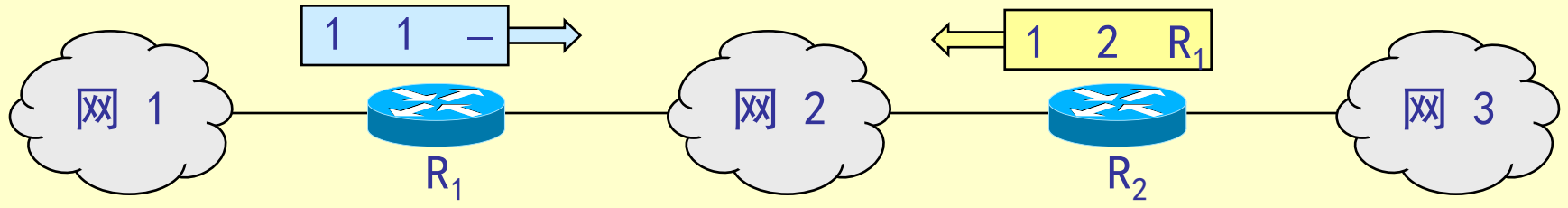
1 16 –

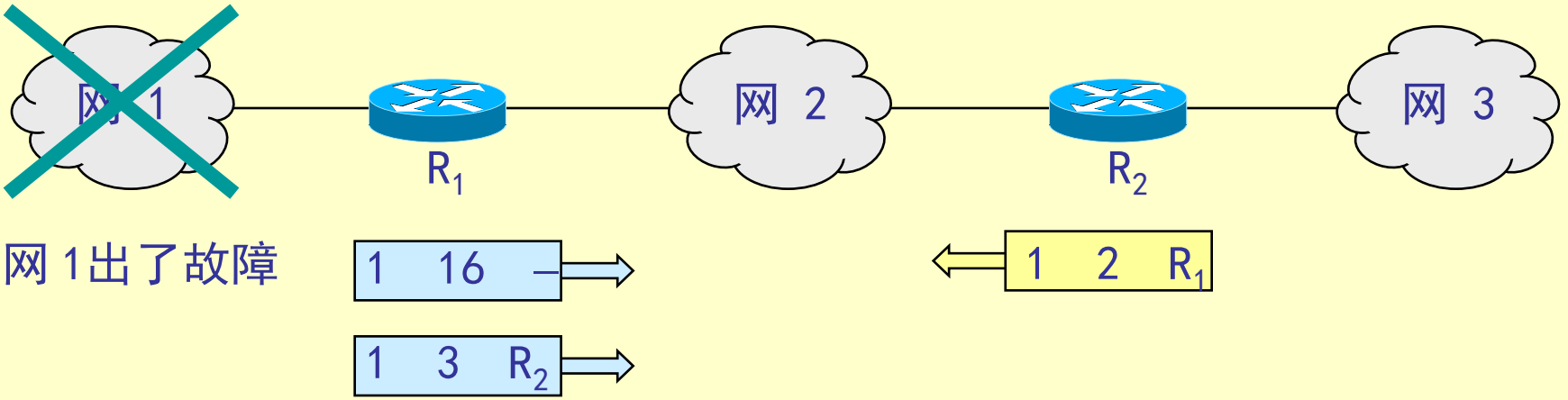
1 2 R1

R_1 说：“我到网 1 的距离是 16 （表示无法到达），是直接交付。”

但 R_2 在收到 R_1 的更新报文之前，还发送原来的报文，因为这时 R_2 并不知道 R_1 出了故障。

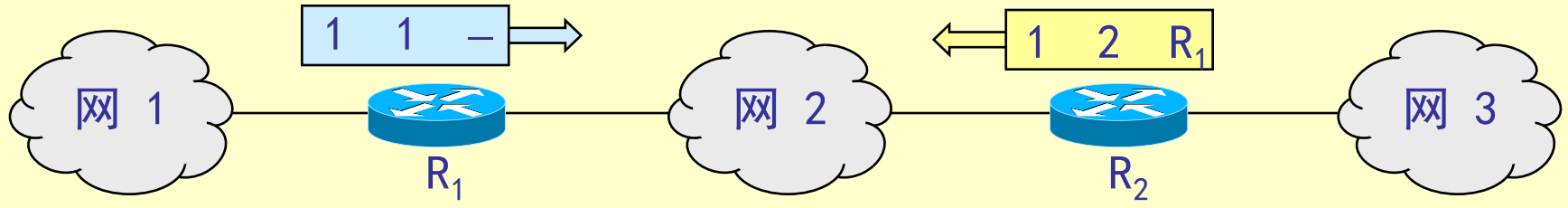
正常情况

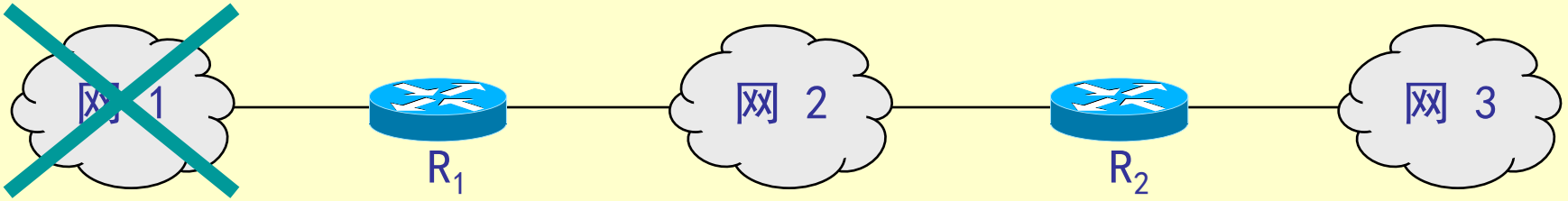




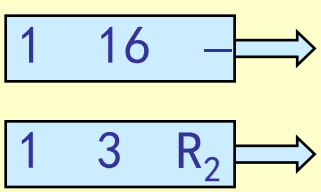
R_1 收到 R_2 的更新报文后，误认为可经过 R_2 到达网1，于是更新自己的路由表，说：“我到网 1 的距离是 3，下一跳经过 R_2”。然后将此更新信息发送给 R_2。

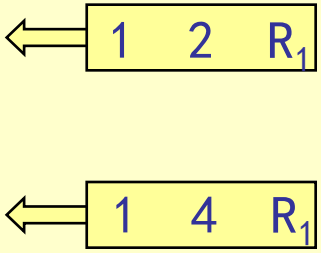
正常情况





网 1出了故障

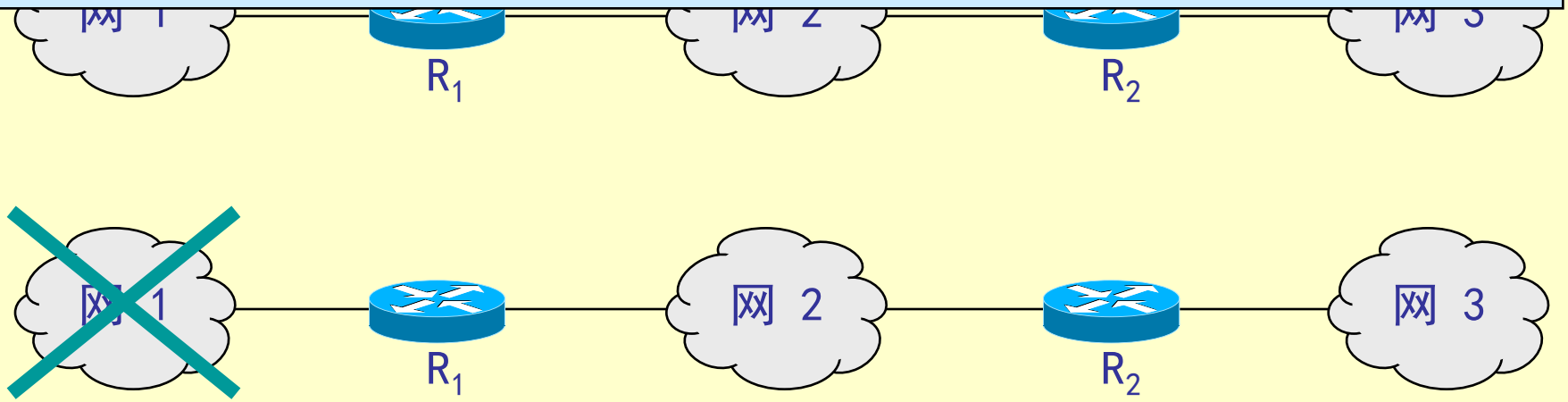




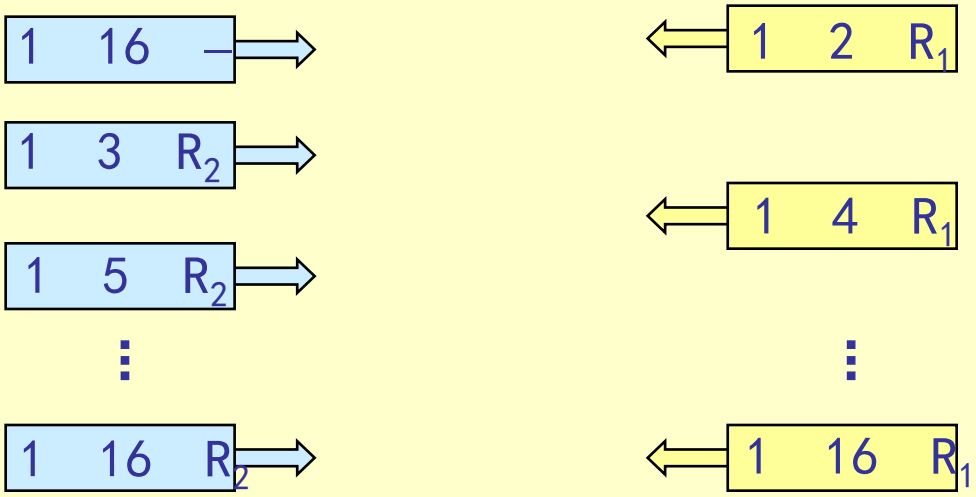
R_2 以后又更新自己的路由表为“1，4，R_1”，表明 “我到网 1 距离是 4，下一跳经过 R_1”。

这就是好消息传播得快，而坏消息传播得慢。网络出故障的传播时间往往需要较长的时间(例如数分钟)。这是 RIP 的一个主要缺点。

情况



网 1出了故障



这样不断更新下去，直到 R_1 和 R_2 到网 1 的距离都增大到 16 时，R_1 和 R_2 才知道网 1 是不可达的。

5.4.8 内部网关协议(OSPF)

- 开放最短路径优先协议OSPF基本特点
  - “开放”表明OSPF协议不受某一家厂商控制，而是公开发表的
  - “最短路径优先”是因为使用了Dijkstra提出的最短路径优先算法SPF（Shortest Path First）
  - OSPF只是一个协议的名字，它并不表示其他的路由选择协议不是“最短路径优先”
  - 是一种分布式的链路状态协议

5.4.8 内部网关协议(OSPF)

- 链路状态数据库
  - 与距离矢量算法不同，SPF算法的特点是每个路由器要知道全部的网络拓扑结构信息
  - SPF各路由器之间周期性交换链路状态信息，说明该路由器与那些路由器邻接，以及连接链路的度量
  - 因此所有的路由器最终都能建立一个链路状态数据库（LSDB），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

5.4.8 内部网关协议(OSPF)

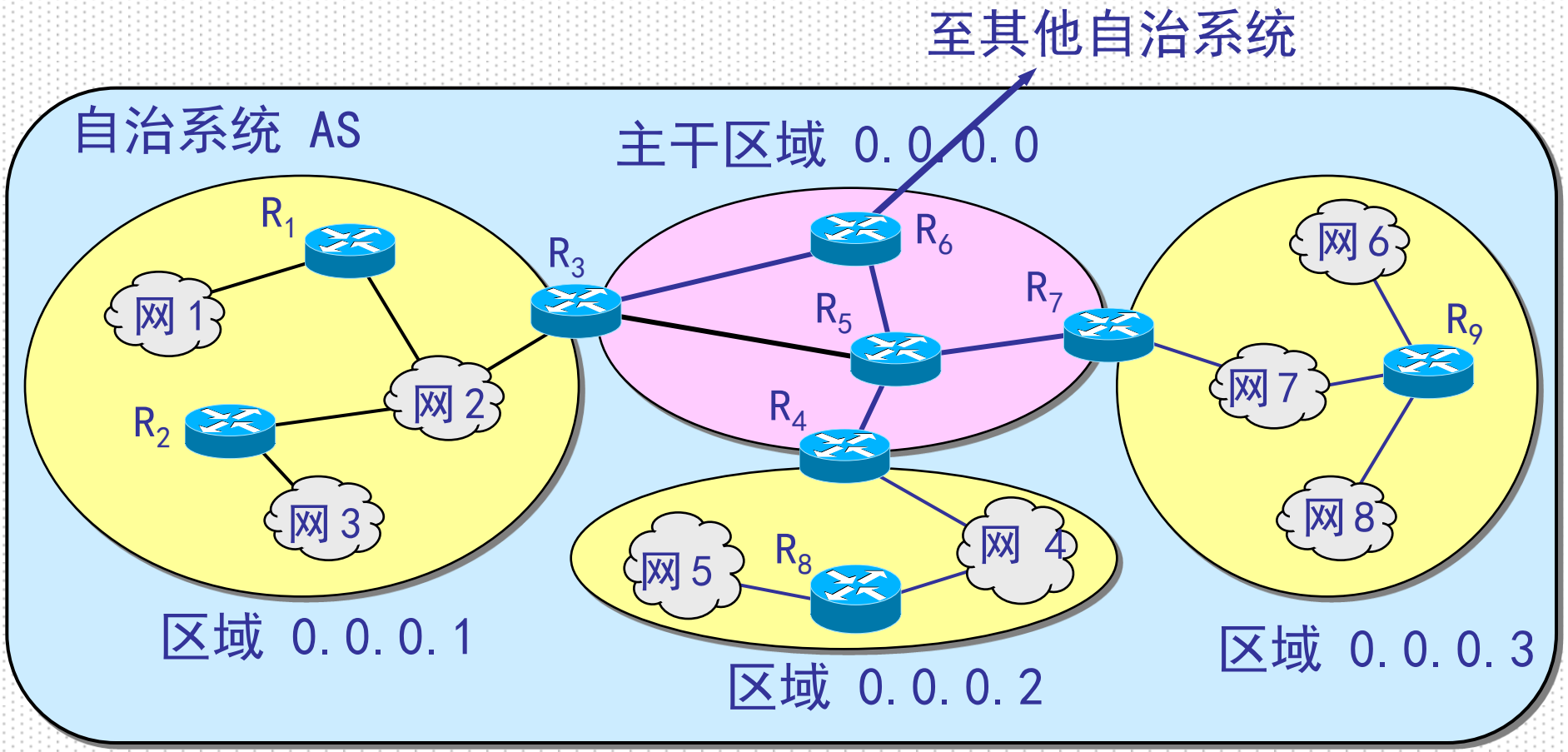
- OSPF工作流程及特点
  - 各路由器使用链路状态信息更新自己的网络拓扑图
  - 依据链路状态的新数据，各路由器使用Dijkstra算法（从单个源点开始计算到其它所有目的结点的最短路径），对网络拓扑图求最短路径
  - 两个显著特点
    - 每个路由器使用同样原始数据，独立进行最短路由计算，不依赖于中间路由器计算结果，可保证收敛性
    - 链路状态报文仅携带与单个路由器直接相连的链路信息，报文长短与网络数无关，更适合大规模的网络

5.4.8 内部网关协议(OSPF)

- OSPF的区域(area)
  - 为了使OSPF能够用于规模很大的网络，OSPF将一个自治系统再划分为若干个更小的范围，叫作区域
  - 每一个区域都有一个32位的区域标识符（用点分十进制表示）
  - 区域也不能太大，在一个区域内的路由器最好不超过200个

5.4.8 内部网关协议(OSPF)

- OSPF划分为两种不同的区域

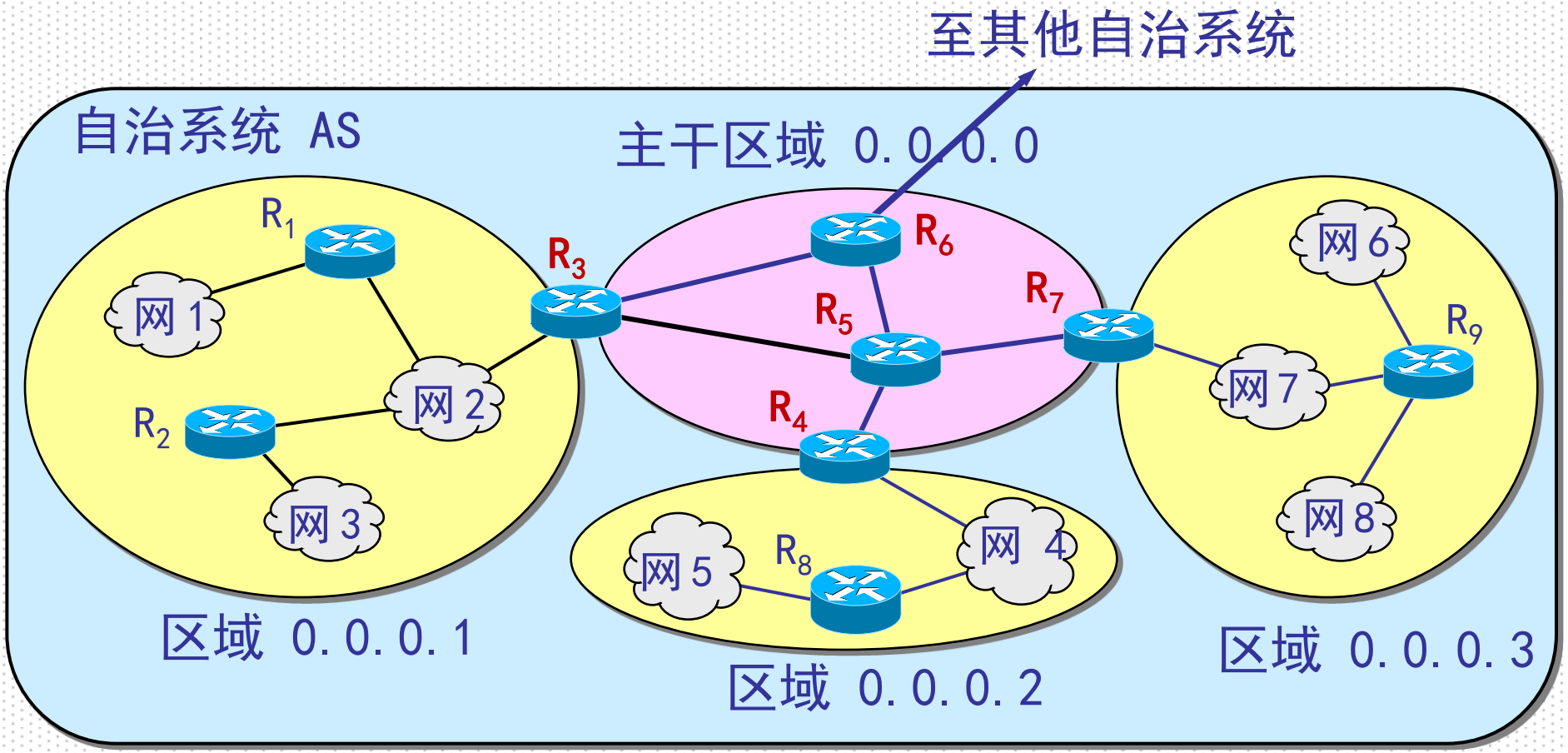


5.4.8 内部网关协议(OSPF)

- 划分区域
  - 划分区域的好处就是将利用洪泛法交换链路状态信息范围局限于每一个区域而不是整个自治系统，这就减少了整个网络上的通信量
  - 在一个区域内部的路由器只知道本区域完整网络拓扑，而不知道其他区域的网络拓扑的情况
  - OSPF使用层次结构的区域划分。上层的区域叫作主干区域（backbone area）。主干区域的标识符规定为0.0.0.0。主干区域的作用是用来连通其他在下层的区域

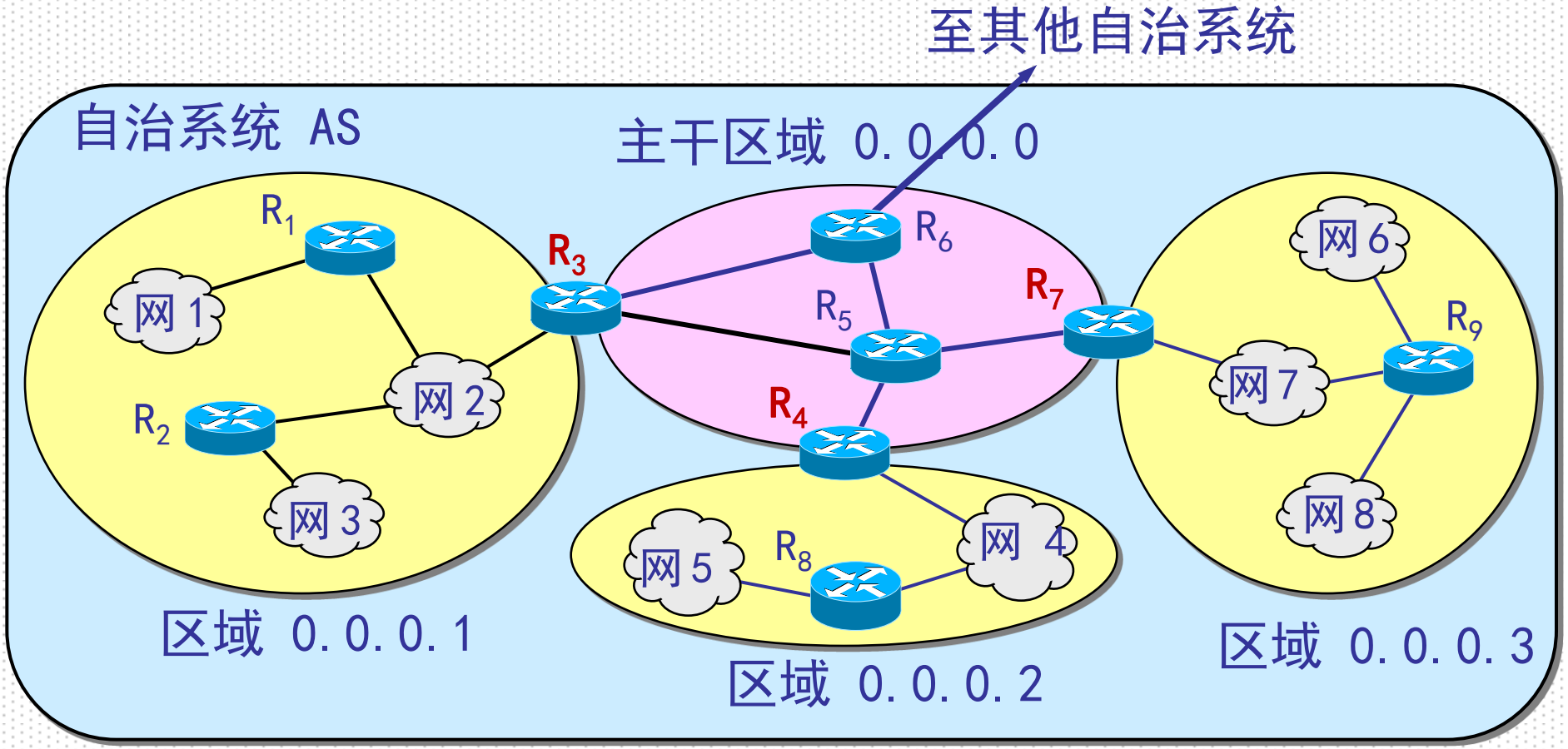
5.4.8 内部网关协议(OSPF)

- 主干路由器



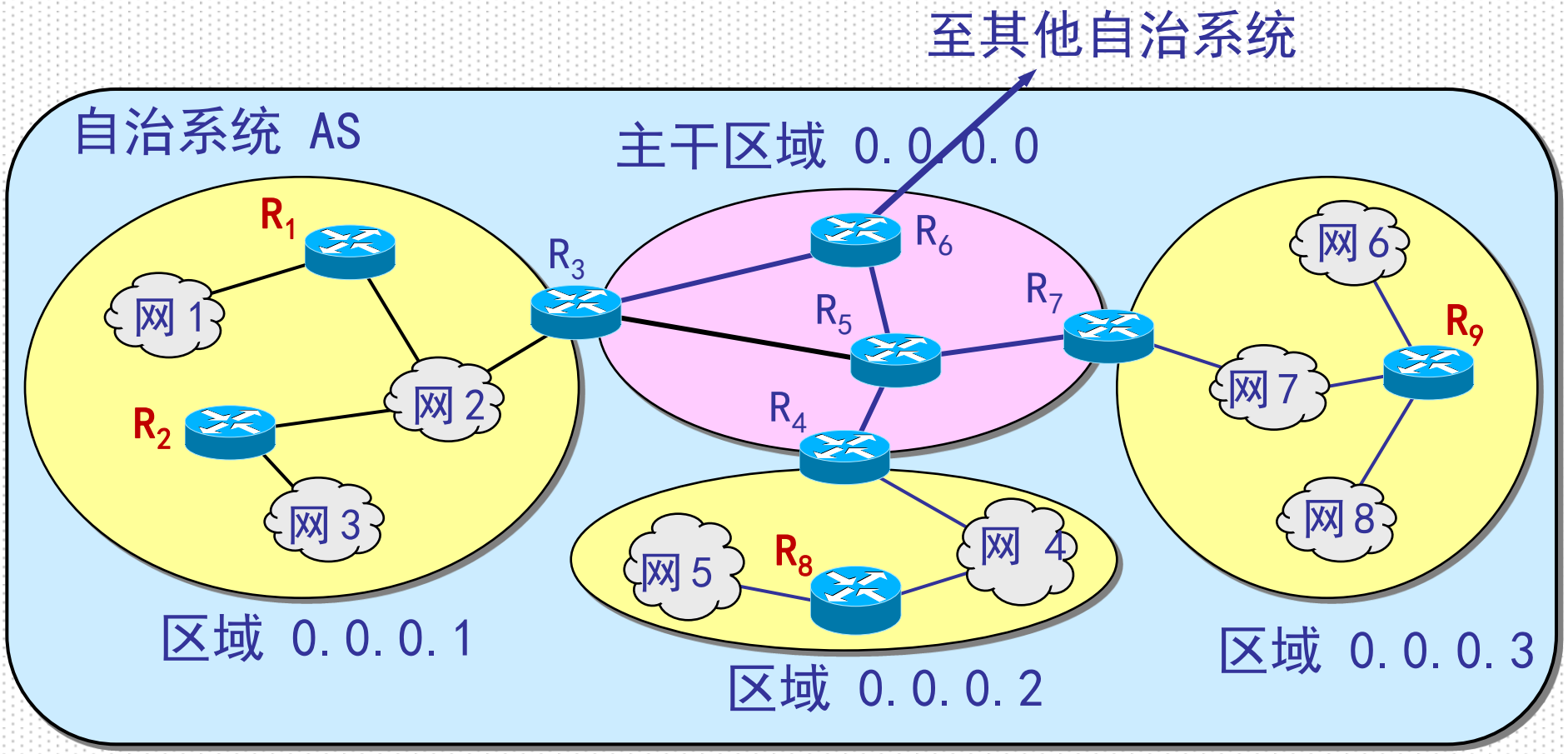
5.4.8 内部网关协议(OSPF)

- 区域边界路由器：重要的桥梁作用



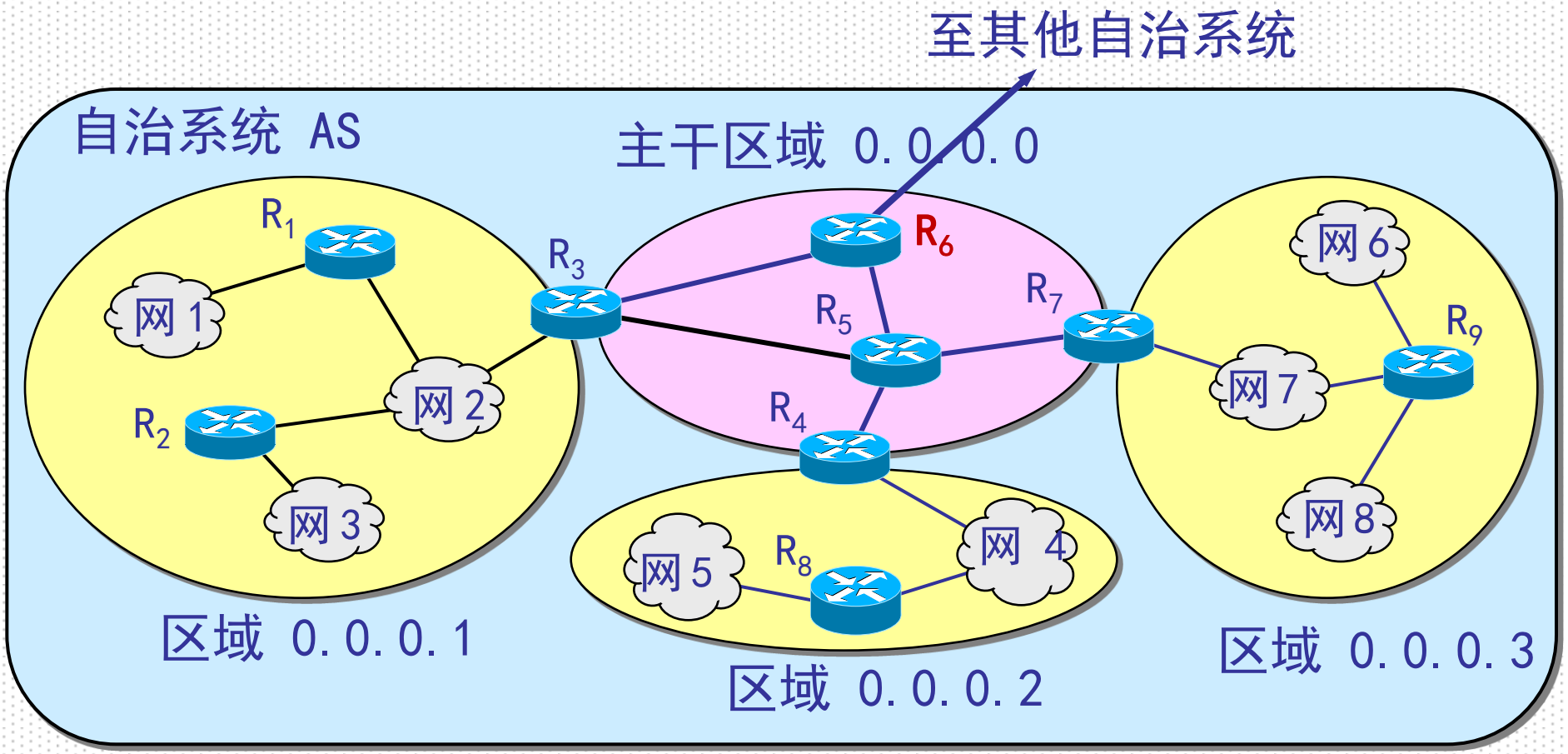
5.4.8 内部网关协议(OSPF)

- 内部路由器



5.4.8 内部网关协议(OSPF)

- 自治系统边界路由器



5.4.9 外部网关协议(BGP)

- BGP是不同AS的路由器之间交换路由信息的协议
- 因特网规模太大，使得AS之间路由选择非常困难。对于AS之间路由选择，寻找最佳路由是很不现实的
  - 当一条路径通过几个不同AS时，要想对这样的路径计算出有意义的代价是不太可能的
  - 比较合理的做法是在AS之间交换“可达性”信息
- AS之间的路由选择必须考虑有关策略
  - 跨越不同国家和大洲，要考虑费用、安全乃至政治因素
- 因此，边界网关协议BGP只能是力求寻找一条能够到达目的网络且比较好的路由（不能兜圈子），而并非要寻找一条最佳路由

5.4.9 外部网关协议(BGP)

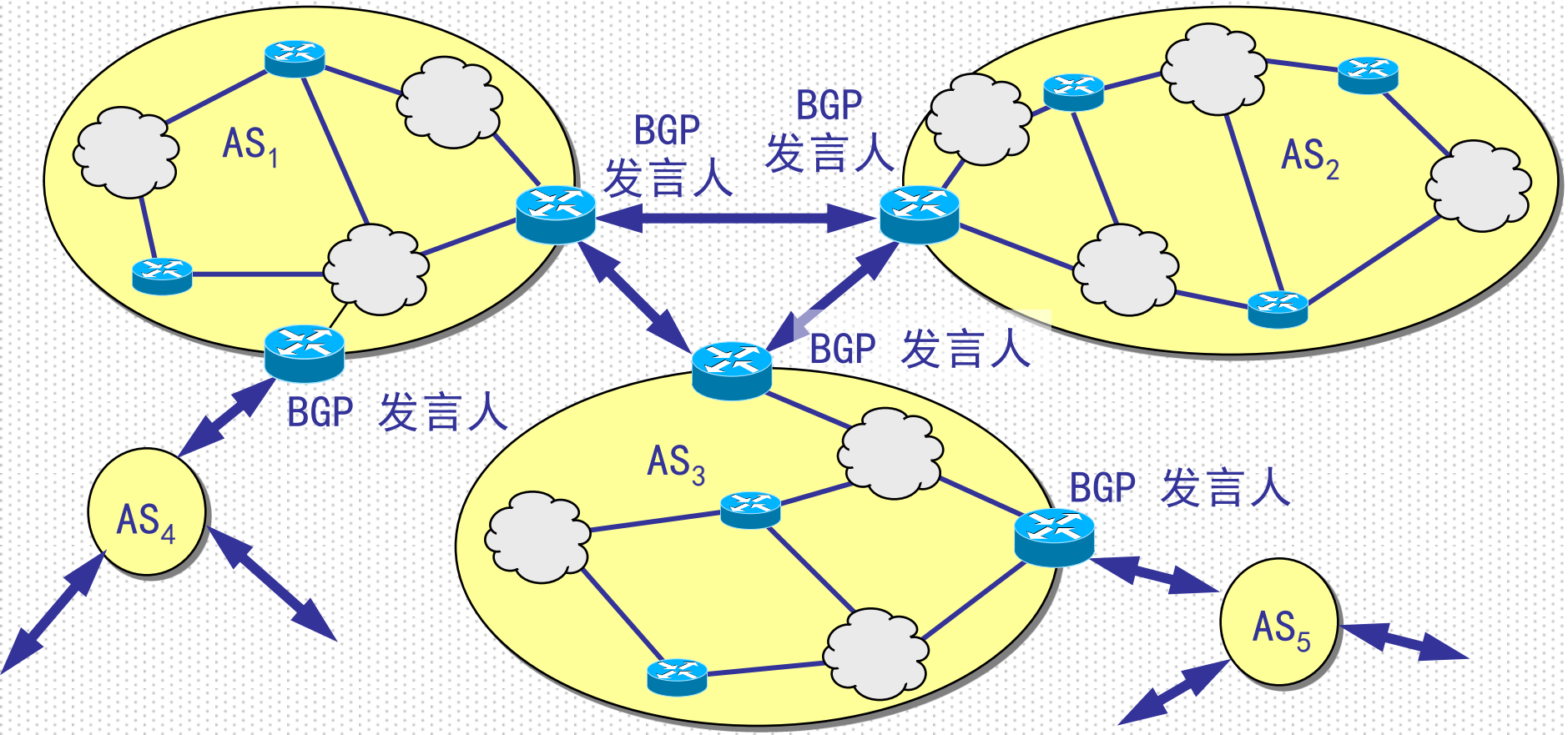
- 外部网关协议BGP
  - BGP较新版本是2006年1月发表的BGP-4（BGP第4个版本，可简写为BGP），即RFC 4271~4278
- BGP发言人
  - 每一个自治系统的管理员要选择至少一个路由器作为该自治系统的“BGP发言人”
  - 两个BGP发言人常通过一个共享网络连接在一起，BGP发言人往往就是BGP边界路由器，但也可以不是BGP边界路由器

5.4.9 外部网关协议(BGP)

- BGP交换路由信息
  - 一个BGP发言人与其他自治系统中的BGP发言人要交换路由信息，就要先建立TCP连接，然后在此连接上交换BGP报文以建立BGP会话（session），利用BGP会话交换路由信息
  - 使用TCP连接能提供可靠服务，也简化了路由选择协议
  - 使用TCP连接交换路由信息的两个BGP发言人，彼此成为对方的邻站或对等站
  - 参与交换信息的节点数是AS数的量级，简化了协议

5.4.9 外部网关协议(BGP)

- BGP发言人和自治系统AS的关系

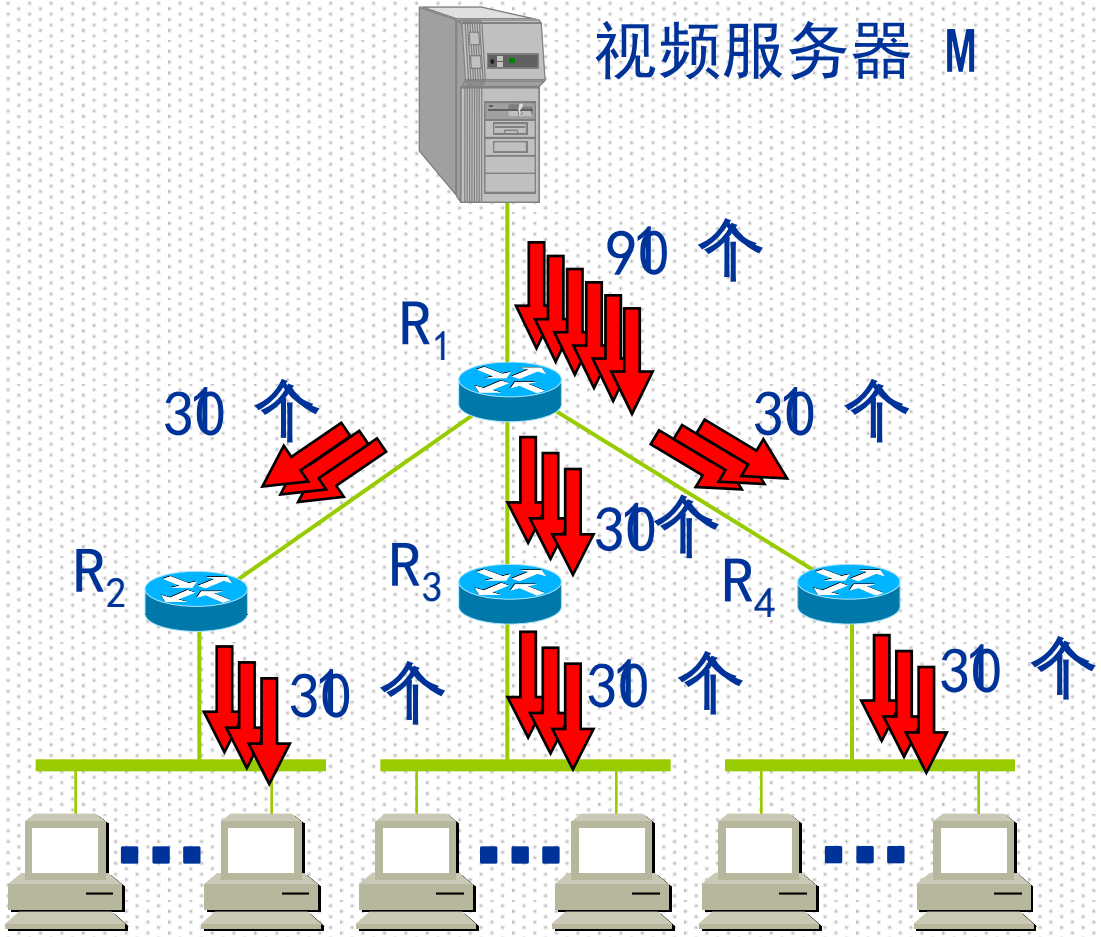


5.4.9 外部网关协议(BGP)

- BGP协议的特点
  - BGP支持无类别域间路由选择CIDR，因此BGP的路由表也就应当包括目的网络前缀、下一跳路由器，以及到达该目的网络所要经过的各个自治系统序列
  - 在BGP刚刚运行时，BGP的邻站是交换整个的BGP路由表。但以后只需要在发生变化时更新有变化的部分。这样做对节省网络带宽和减少路由器的处理开销方面都有好处

5.4.10 IP多播

- IP多播的基本概念



不使用多播时需要发送 90 次单播，即向每台主机各发送1次IP数据报

使用多播时仅发送1次IP数据报

5.4.10 IP多播

- IP 多播的一些特点
  - 多播使用组地址——IP使用D类地址支持多播。多播地址只能用于目的地址，而不能用于源地址
  - 临时多播地址和永久多播地址。永久组地址由因特网号码指派管理局IANA负责指派
    - 224.0.0.1　本地网络上所有的系统
    - 224.0.0.2　本地网络上所有的路由器

    } 永久组地址示例
  - 动态的组成员，主机可任意加入/退出多播组
  - 使用硬件进行多播，需底层网路/路由器支持

5.4.11 因特网管理协议(IGMP)

- 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

5.4.11 因特网管理协议(IGMP)

- IGMP协议
  - 1989年公布的RFC 1112（IGMPv1）早已成为了因特网的标准协议
  - 1997年公布的RFC 2236（IGMPv2，建议标准）对IGMPv1进行了更新
  - 2002年 公布了RFC 3376（IGMPv3，建议标准），宣布RFC 2236（IGMPv2）是陈旧的

5.4.11 因特网管理协议(IGMP)

- 几种多播路由选择协议
  - 距离向量多播路由选择协议DVMRP（Distance Vector Multicast Routing Protocol）
  - 基于核心的转发树CBT（Core Based Tree）
  - 开放最短通路优先的多播扩展MOSPF（Multicast Extensions to OSPF）
  - 协议无关多播-稀疏方式PIM-SM（Protocol Independent Multicast-Sparse Mode）
  - 协议无关多播-密集方式PIM-DM（Protocol Independent Multicast-Dense Mode）

5.4.12 最新网际协议(IPv6)

- IPv4技术缺陷 —先天不足
  - 地址危机
  - 配置复杂
  - 路由表的膨胀
  - ……

5.4.12 最新网际协议(IPv6)

- 技术的原因
  - 32bit地址空间
- 分配策略的原因
  - 理论： 32-bit space: 4 billion devices
  - 实际： 32-bit space: 250 million devices
- 历史的原因
  - 互联网发展初期地址分配策略不合理
  - 全球互联网起步时间不同，区域性地址分配不均（中国全部地址量约318B，美国仅MIT就有256B）

5.4.12 最新网际协议(IPv6)

- IPv6优势
  - 地址容量巨大
    - 采用128bit地址空间，理论上有2^{128}个地址，足够为地球上每一粒沙子分配一个独立地址
    - 比IPv4地址空间大7.9×10^{28}倍，每个人拥有6×10^{28}个IP地址
  - 简化处理
    - 新的简化的报头格式，更有利于中间路由器的处理，将报头开销减少到最小程度
    - 新的路由算法，减小了路由表的大小，提高了路由选择效率
  - 对流的支持、安全功能、即插即用功能

5.4.12 最新网际协议(IPv6)

- IPv6地址表示

冒分十六进制记法

  - IPv4以8位作为分界，IPv6以16位做为分界，IPv6标准地址格式为X:X:X:X:X:X:X:X（X为16位/十六进制值表示）
  - 例如21DA:00D3:0000:2F3B:02AA:00FF:FE28:9C5A
  - 零压缩：如果地址出现多个0，则可以用一对冒号“::”进行压缩。如FE80:0:0:0:2AA:FF:FE9A:4CA2可以用FE80::2AA:FF:FE9A:4CA2来表示。但需要注意的是，零压缩只能在给定的地址中使用1次
  - 三种地址：单播，多播，任播
  - IPv4与IPv6的过渡：双协议栈技术、隧道技术

本章知识点小结

- 网络层提供的服务
- 数据交换方式
- IP协议-作用、地址、两级寻址、子网掩码
- ARP地址解析
- 路由选择协议RP
- IPv6协议
- 思考题
  - P197：5.1、5.2、5.3、5.5、5.7、5.10、5.14、5.19、5.37、5.38