

第 4 章 网络层



上

第 4 章 网络层



- 4.1 网际协议 IP
- 4.2 划分子网和构造超网
- 4.3 网际控制报文协议 ICMP
- 4.4 互联网的路由选择协议

4.1 网际协议 IP



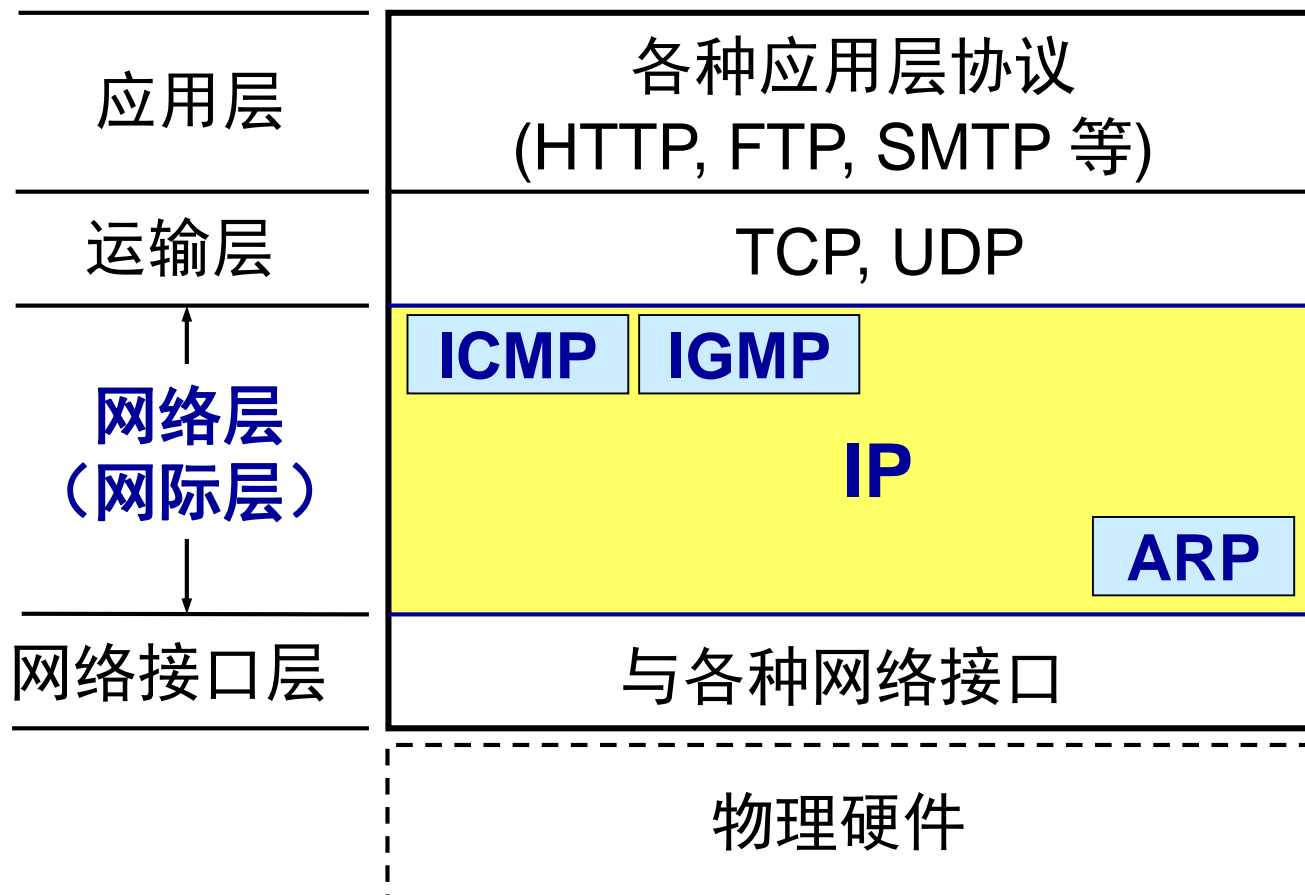
- 4.1.1 虚拟互连网络
- 4.1.2 分类的 IP 地址
- 4.1.3 IP 地址与硬件地址
- 4.1.4 地址解析协议 ARP
- 4.1.5 IP 数据报的格式
- 4.1.6 IP 层转发分组的流程

4.1 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



4.1.1 虚拟互连网络



- 将网络互连并能够互相通信，会遇到许多问题需要解决，如：
 - 不同的寻址方案
 - 不同的最大分组长度
 - 不同的网络接入机制
 - 不同的超时控制
 - 不同的差错恢复方法
 - 不同的状态报告方法
 - 不同的路由选择技术
 - 不同的用户接入控制
 - 不同的服务（面向连接服务和无连接服务）
 - 不同的管理与控制方式等

**如何将异构的网络
互相连接起来？**

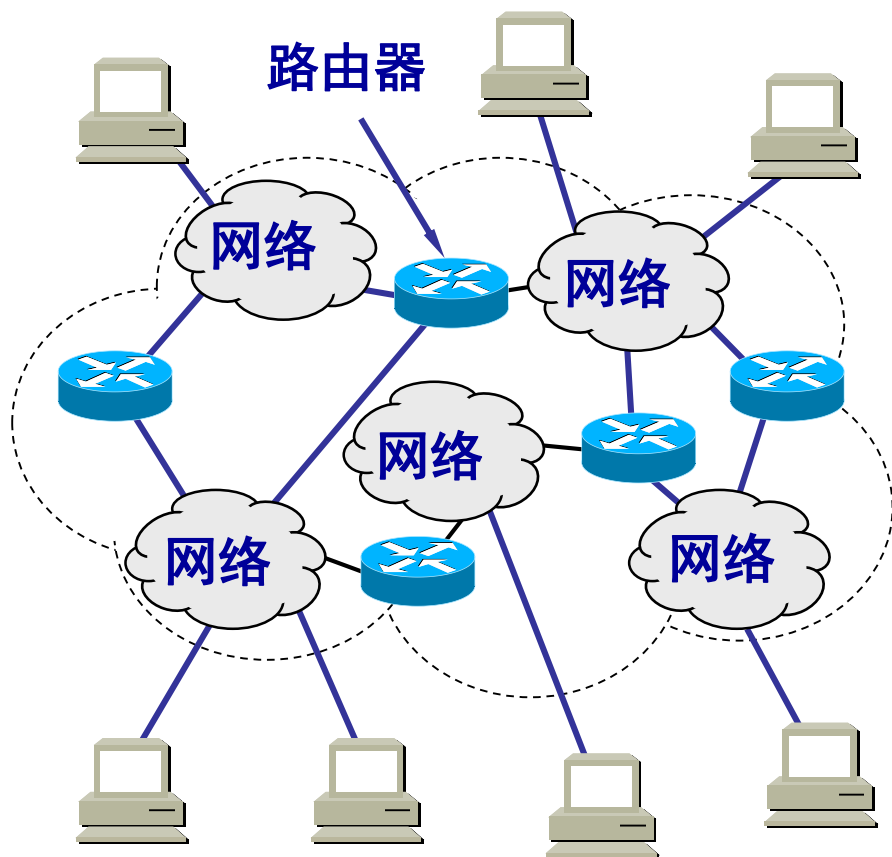
使用一些中间设备进行互连



- 将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**。
- 有以下几种不同的中间设备：
 - **物理层**中继系统：**转发器 (repeater)**。
 - **数据链路层**中继系统：**交换机(bridge)**。
 - **网络层**中继系统：**路由器 (router)**。
- 现在网络互连都是指用**路由器**进行网络互连和路由选择。



互连网络与虚拟互连网络



(a) 互连网络



(b) 虚拟互连网络

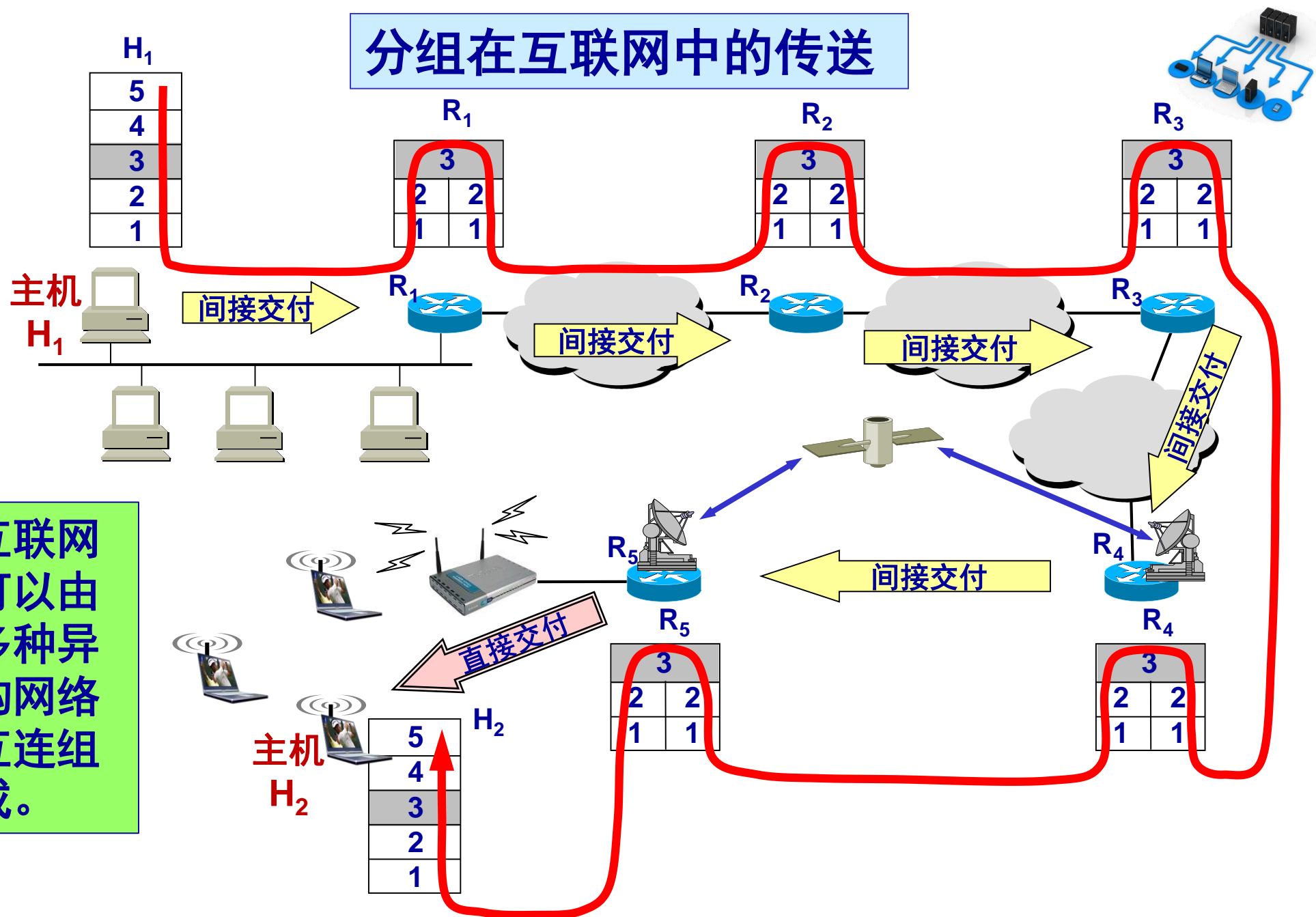
IP 网的概念

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**。
- 如果在这种覆盖全球的 **IP 网**的上层使用 **TCP 协议**，那么就是现在的互联网 (Internet)。

分组在互联网中的传送



互联网
可以由
多种异
构网络
互连组
成。

1.直接交付和间接交付



- (1) 直接交付：发送数据报接口的IP地址与目的主机IP地址在同一个网络中时，也就是发送端接口IP地址和子网掩码按位“与”操作的结果等于目的主机的IP地址和子网掩码按位“与”操作的结果。则发送端接口直接将IP数据报发送到目的主机。

1.直接交付和间接交付

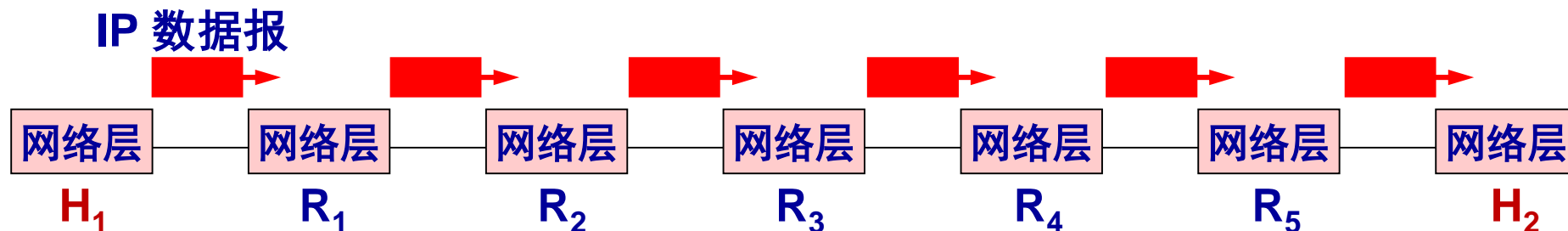


- **（2）间接交付：**发送数据报接口IP地址与目的IP地址不在同一个网络中，则发送端接口需要发送IP数据报到与下一个路由器相连的端口。在数据报间接交付过程中，在同一个网络内使用物理地址（MAC地址）转发，只有通过路由器在不同网络中转发时才使用IP地址。

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。



4.1.2 分类的 IP 地址



- 在 TCP/IP 体系中，IP 地址是一个最基本的概念。
- 本部分重点学习：
 - 1. IP 地址及其表示方法
 - 2. 常用的三类类别的 IP 地址

1. 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互联网上的主机（或路由器）分配一个在全世界范围是**唯一的 32 位的标识符**。
- IP 地址现在由**互联网名字和数字分配机构 ICANN (Internet Corporation for Assigned Names and Numbers)**进行分配。

IP 地址的编址方法



- **分类的 IP 地址。**这是**最基本的编址方法**，在 1981 年就通过了相应的标准协议。
- **子网的划分。**这是对最基本的编址方法的**改进**，其标准 [RFC 950] 在 1985 年通过。
- **构成超网。**这是比较新的**无分类编址方法**。1993 年提出后很快就得到推广应用。

分类 IP 地址

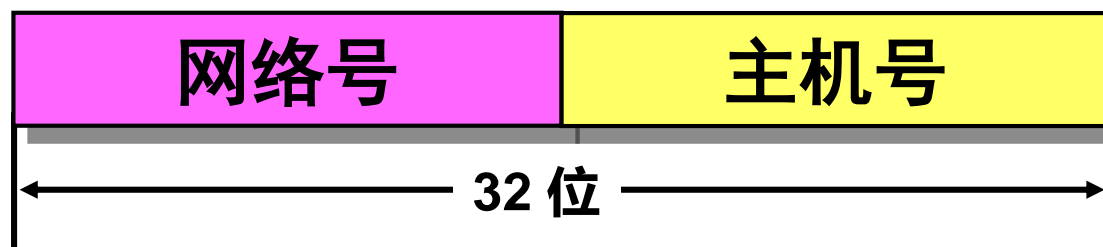


- 将IP地址划分为若干个固定类。
- 每一类地址都由两个固定长度的字段组成，其中一个字段是**网络号 net-id**，它标志主机（或路由器）所连接到的网络，而另一个字段则是**主机号 host-id**，它标志该主机（或路由器）。
- 主机号在它前面的网络号所指明的网络范围内必须是唯一的。
- 由此可见，**一个 IP 地址在整个互联网范围内是唯一的。**

分类 IP 地址



- 这种两级的 IP 地址结构如下：

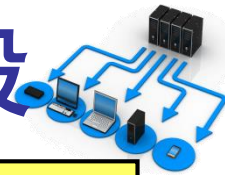


- 这种两级的 IP 地址可以记为：

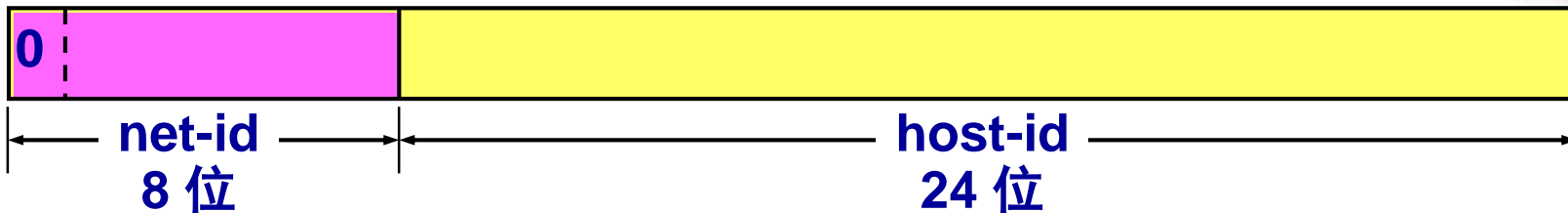
IP 地址 ::= { <网络号>, <主机号> } (4-1)

::= 代表 “定义为”

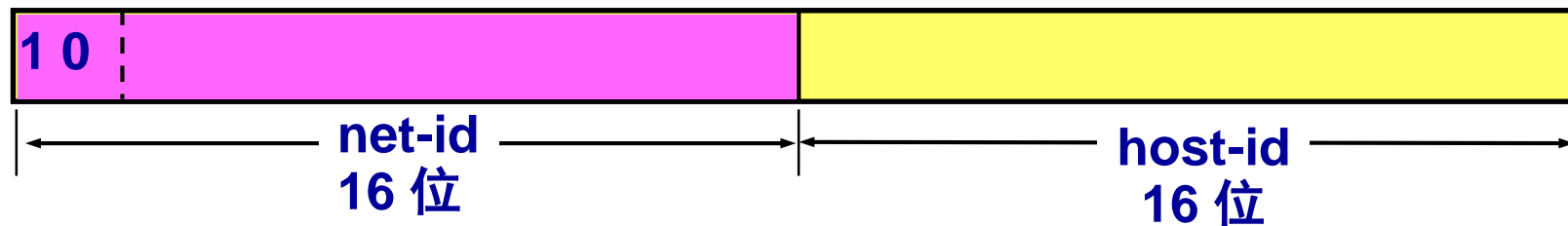
各类 IP 地址的网络号字段和主机号字段



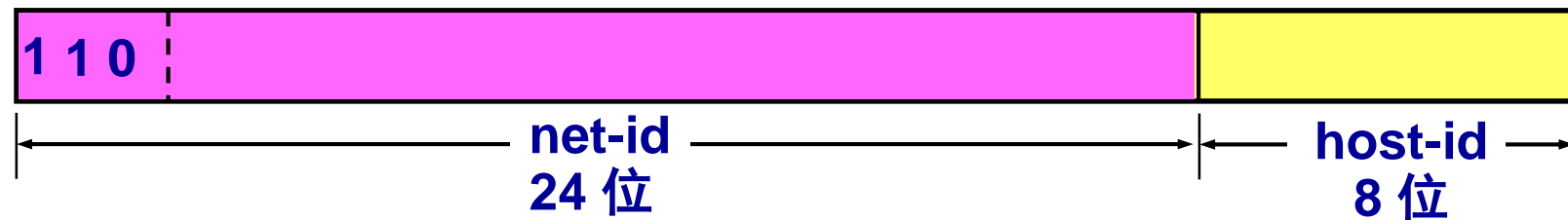
A 类地址



B 类地址



C 类地址



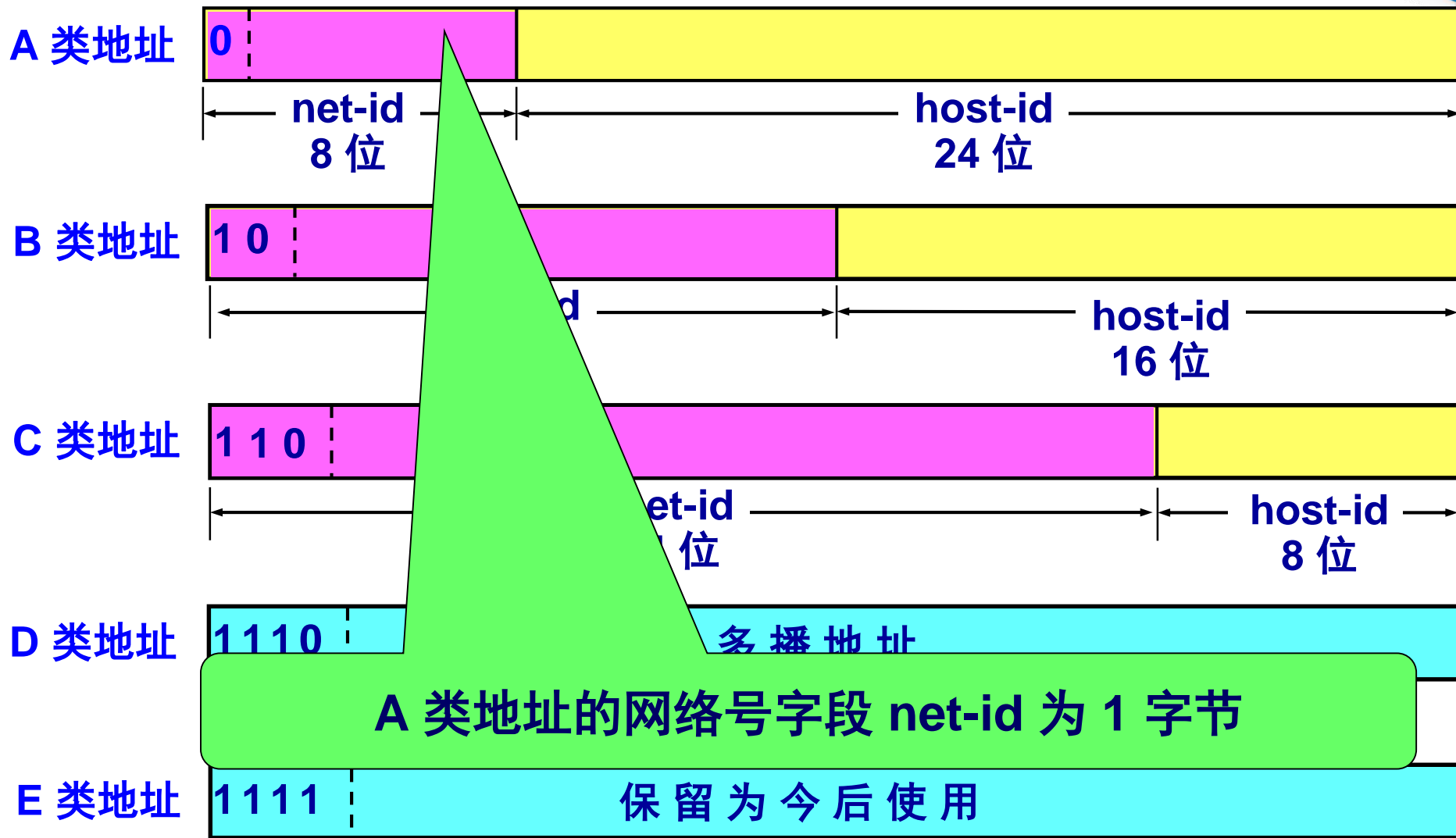
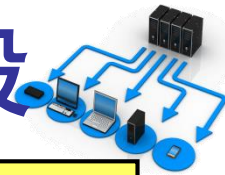
D 类地址



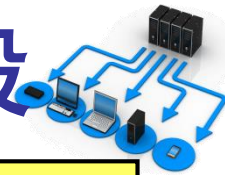
E 类地址



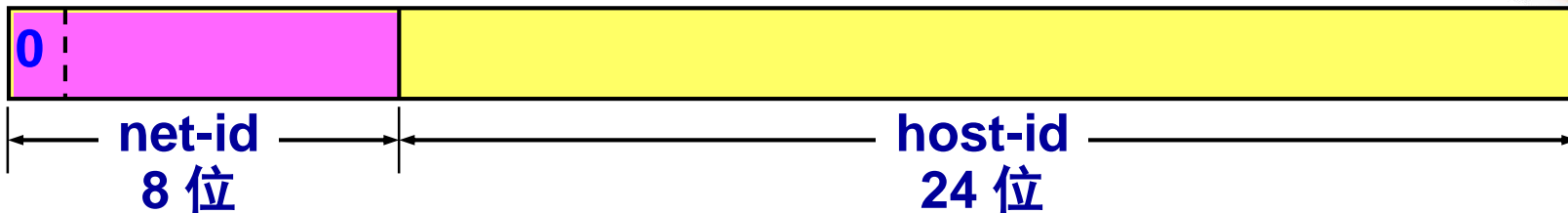
各类 IP 地址的网络号字段和主机号字段



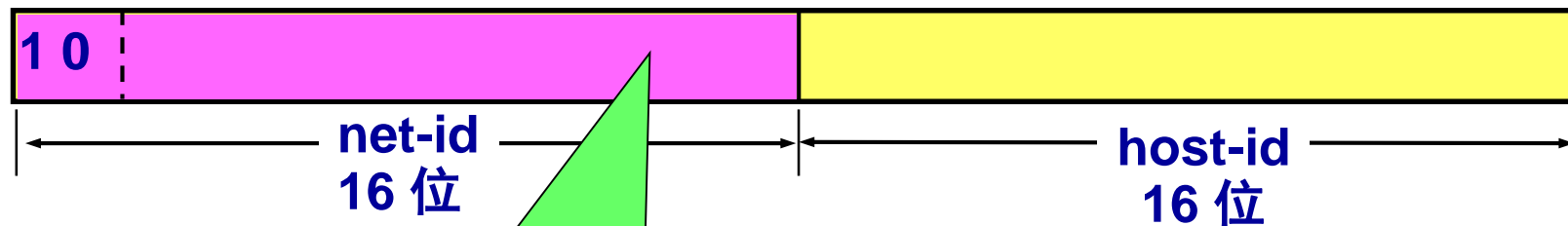
各类 IP 地址的网络号字段和主机号字段



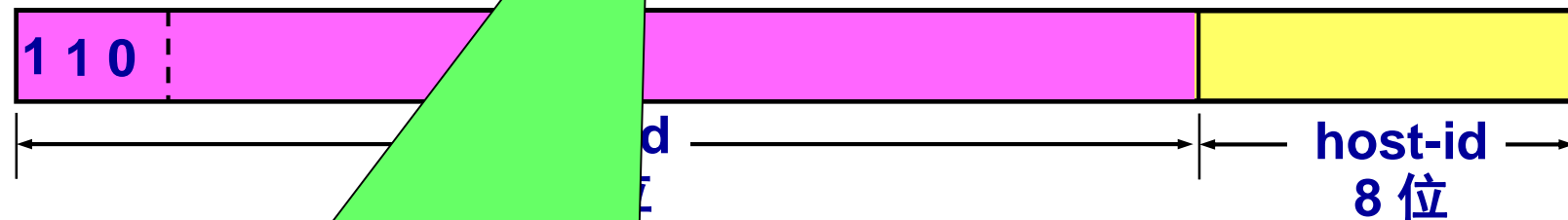
A 类地址



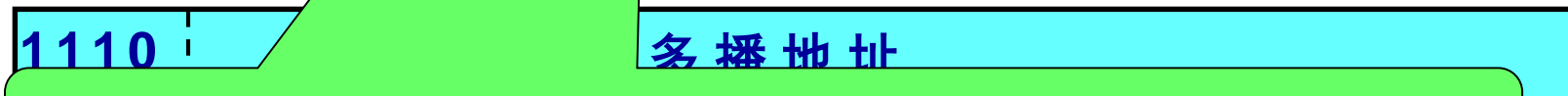
B 类地址



C 类地址



D 类地址

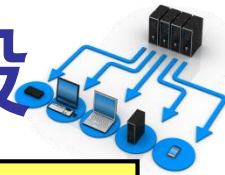


E 类地址

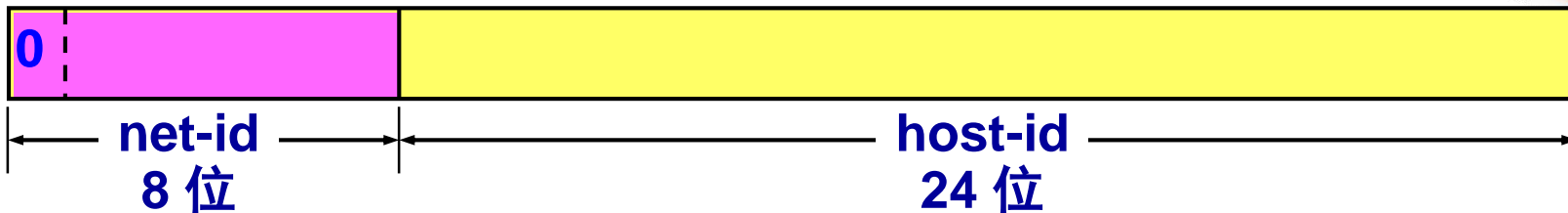


B 类地址的网络号字段 net-id 为 2 字节

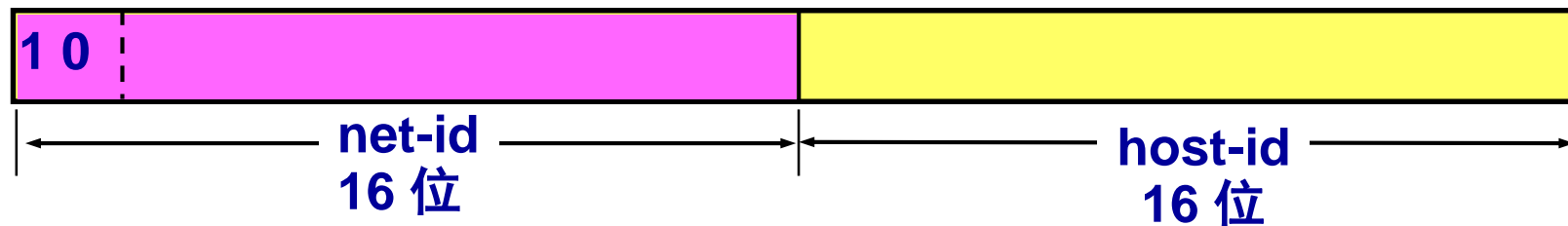
各类 IP 地址的网络号字段和主机号字段



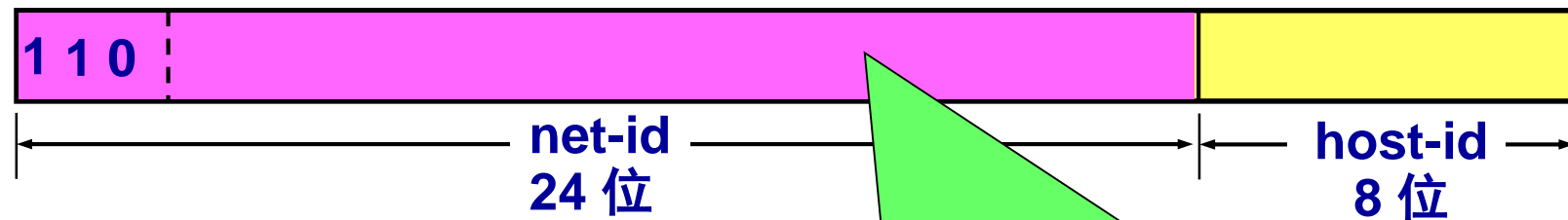
A 类地址



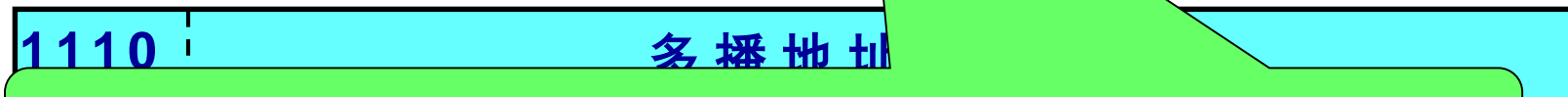
B 类地址



C 类地址



D 类地址



C 类地址的网络号字段 net-id 为 3 字节

E 类地址



点分十进制记法



机器中存放的 IP 地址
是 32 位二进制代码

100000000000010110000001100011111

每 8 位为一组

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

A类地址



- 允许的网络号为 $2^7 - 2 = 126$
- 保留地址
 - 1. 网络地址为全0的为本网路 **00000000**
 - 2. 网络地址为01111111 为环回测试
- A类地址的最大主机数为 $2^{24} - 2$
 - 1. 24位主机号全为0 表示 网络号，如10.0.0.0
 - 2. 24位主机号全为1 表示 该网络中的广播地址如10.255.255.255
- A类地址 网络地址的范围 1.0.0.0~126.0.0.0

B类地址



- 允许的网络号为 $2^{14}-1 = 16383$
- 网络地址为128.0.0.0 不指派
(10000000 .00000000.00000000.00000000)
- B类地址的最大主机数为 $2^{16}-2$
 - 1. 16位主机号全为0 表示 网络号 , 如128.1.0.0
 - 2. 16位主机号全为1 表示 该网络中的广播地址如
128.1.255.255
- B类地址 网络地址的范围 为 ? 128.1.0.0~191.255.0.0

C类地址



- 允许的网络号为 $2^{21}-1 = 2097151$
- 网络地址为192.0.0.0 不指派
(11000000 .00000000.00000000.00000000)
- C类地址的最大主机数为 $2^8-2=254$
- 1. 8位主机号全为0 表示 网络号 , 如192.0.1.0
- 2. 8位主机号全为1 表示 该网络中的广播地址如192.0.1.255
- C类地址 网络地址的范围 ?

192.0.1.0~223.255.255.0

2. 常用的三种类别的 IP 地址



IP 地址的指派范围

网络类别	最大可指派的网络数	第一个可指派的网络号	最后一个可指派的网络号	每个网络中最大主机数
A	126 ($2^7 - 2$)	1	126	16777214
B	16383 ($2^{14} - 1$)	128.1	191.255	65534
C	2097151 ($2^{21} - 1$)	192.0.1	223.255.255	254

特殊IP地址 举例



- 1.直接广播地址
- A,B,C类网络 中主机部分为全1的 为广播地址
- 如 192.168.1.255 是C类地址，这是个广播地址，向192.168.1.0 网络中所有主机广播；
- 2.受限的广播地址
- 如果目的地址为 255.255.255.255，只是向本网络内的所有主机广播。（不能到另一个网络）

特殊IP地址 举例



- 3. 环回地址
- 如果 地址为 127.0.0.1 ，表示 是环回地址，用作本地软件环回测试之用。

一般不使用的特殊的 IP 地址



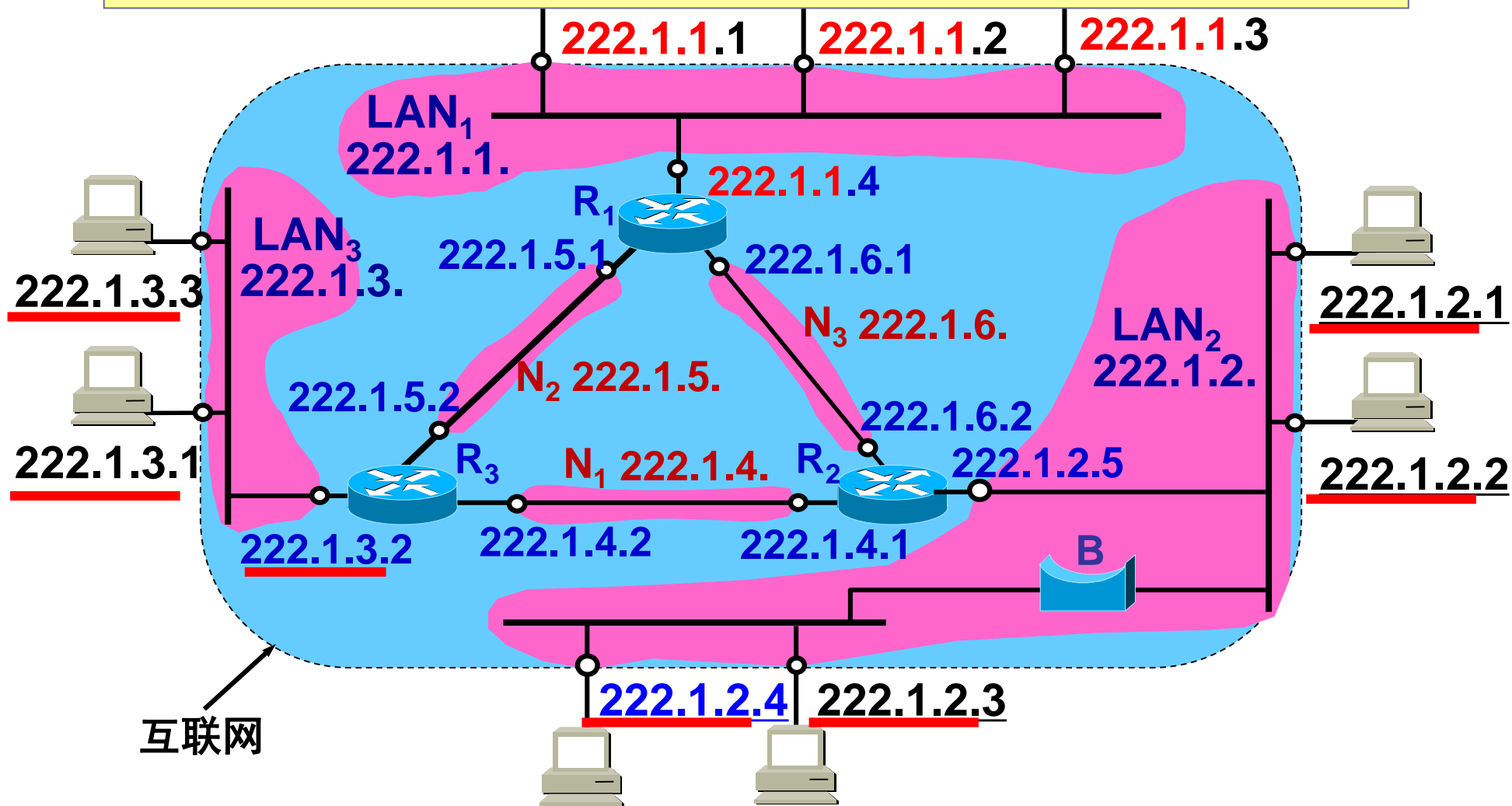
网络号	主机号	源地址使用	目的地址使用	代表的意思
0	0	可以	不可	在本网络上的本主机（见 6.6 节 DHCP 协议）
0	host-id	可以	不可	在本网络上的某台主机 host-id
全 1	全 1	不可	可以	只在本网络上进行广播（各路由器均不转发）
net-id	全 1	不可	可以	对 net-id 上的所有主机进行广播
127	非全 0 或全 1 的任何数	可以	可以	用作本地软件环回测试之用

IP 地址的一些重要特点

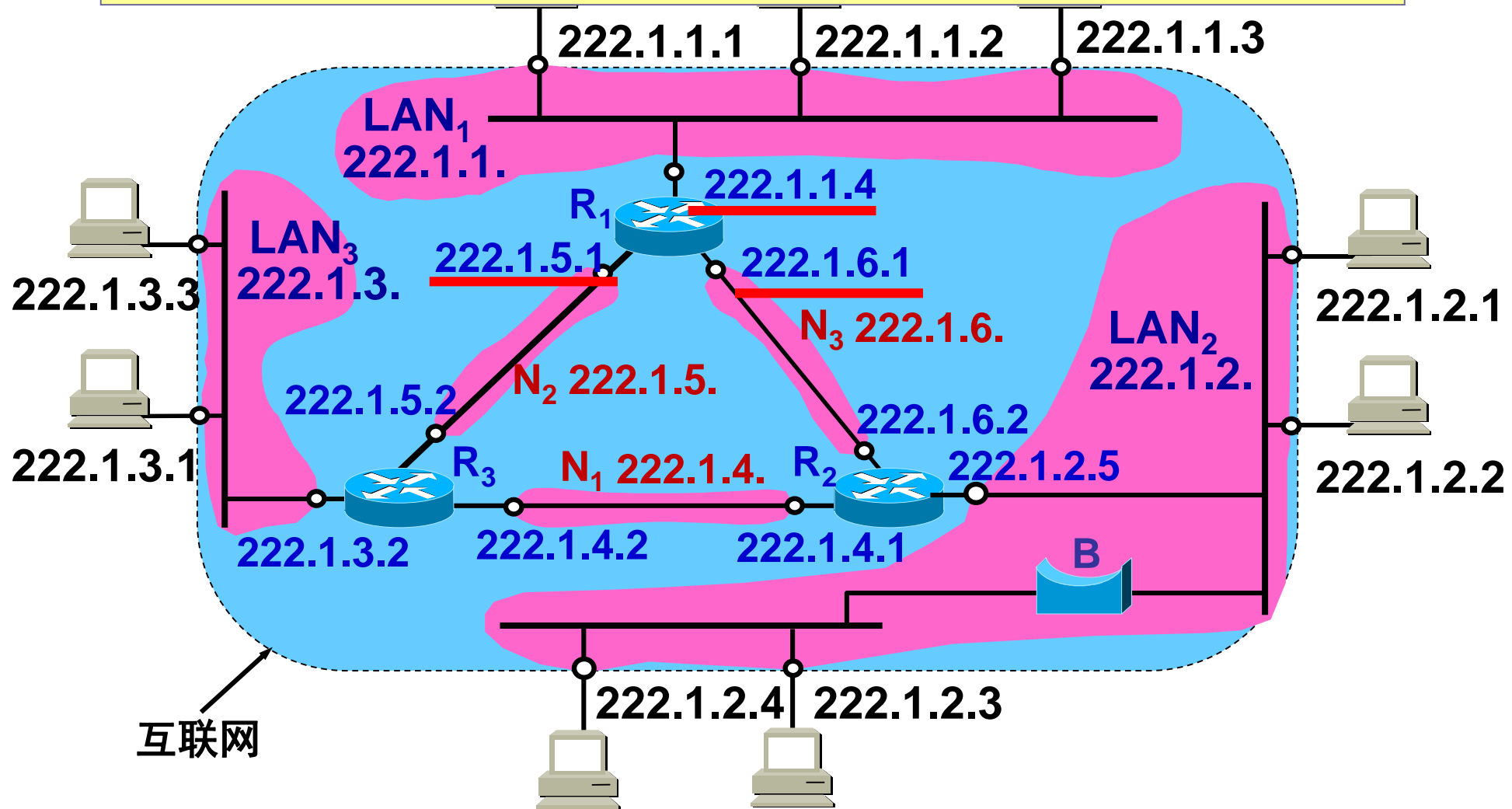


- (1) IP 地址是一种分等级的地址结构。
- (2) 实际上 IP 地址是标志一个主机（或路由器）和一条链路的连接
- (3) 用网桥、交换机连接起来的若干个局域网仍为一个网络，因此这些局域网都具有同样的网络号。
- (4) 所有分配到网络号 的网络，无论是范围很小的局域网，还是可能覆盖很大地理范围的广域网，都是平等的。

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

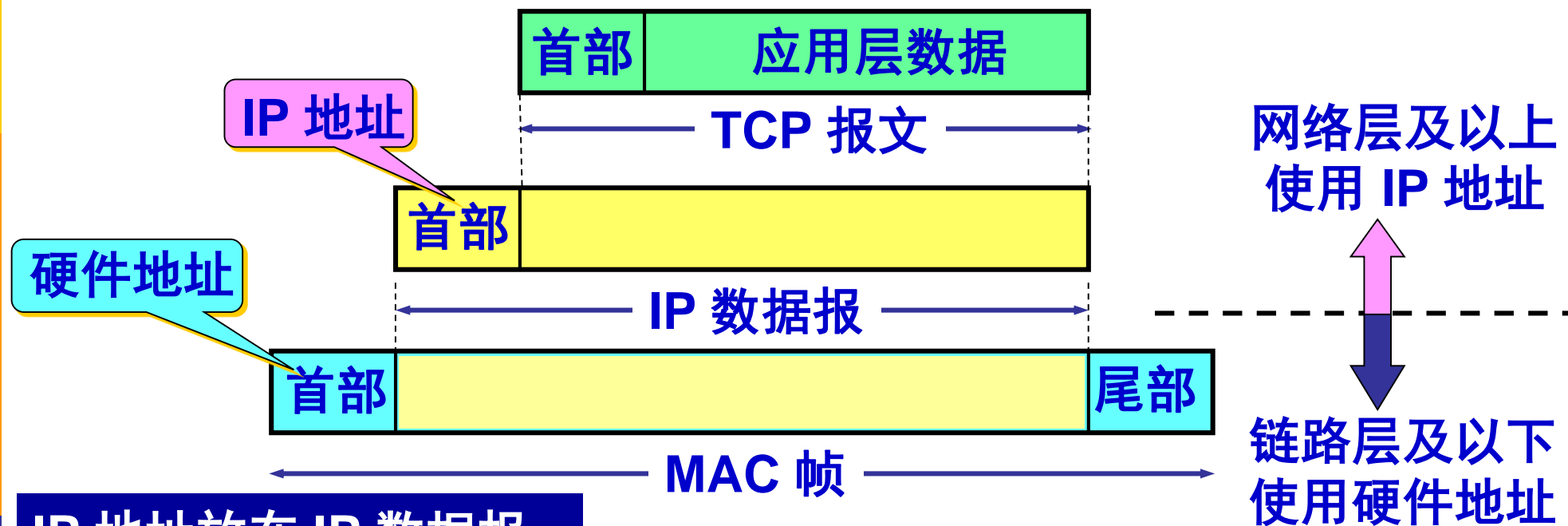


4.1.3 IP 地址与硬件地址



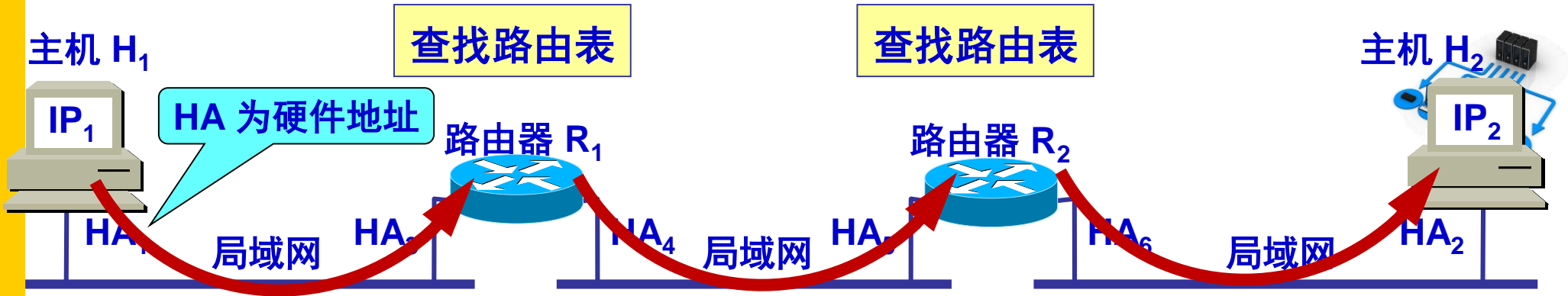
- IP 地址与硬件地址是不同的地址。
- 从层次的角度看，
 - **硬件地址（或物理地址）** 是数据链路层和物理层使用的地址。
 - **IP 地址** 是网络层和以上各层使用的地址，是一种**逻辑地址**（称 IP 地址是逻辑地址是因为 IP 地址是用软件实现的）。

4.1.3 IP 地址与硬件地址



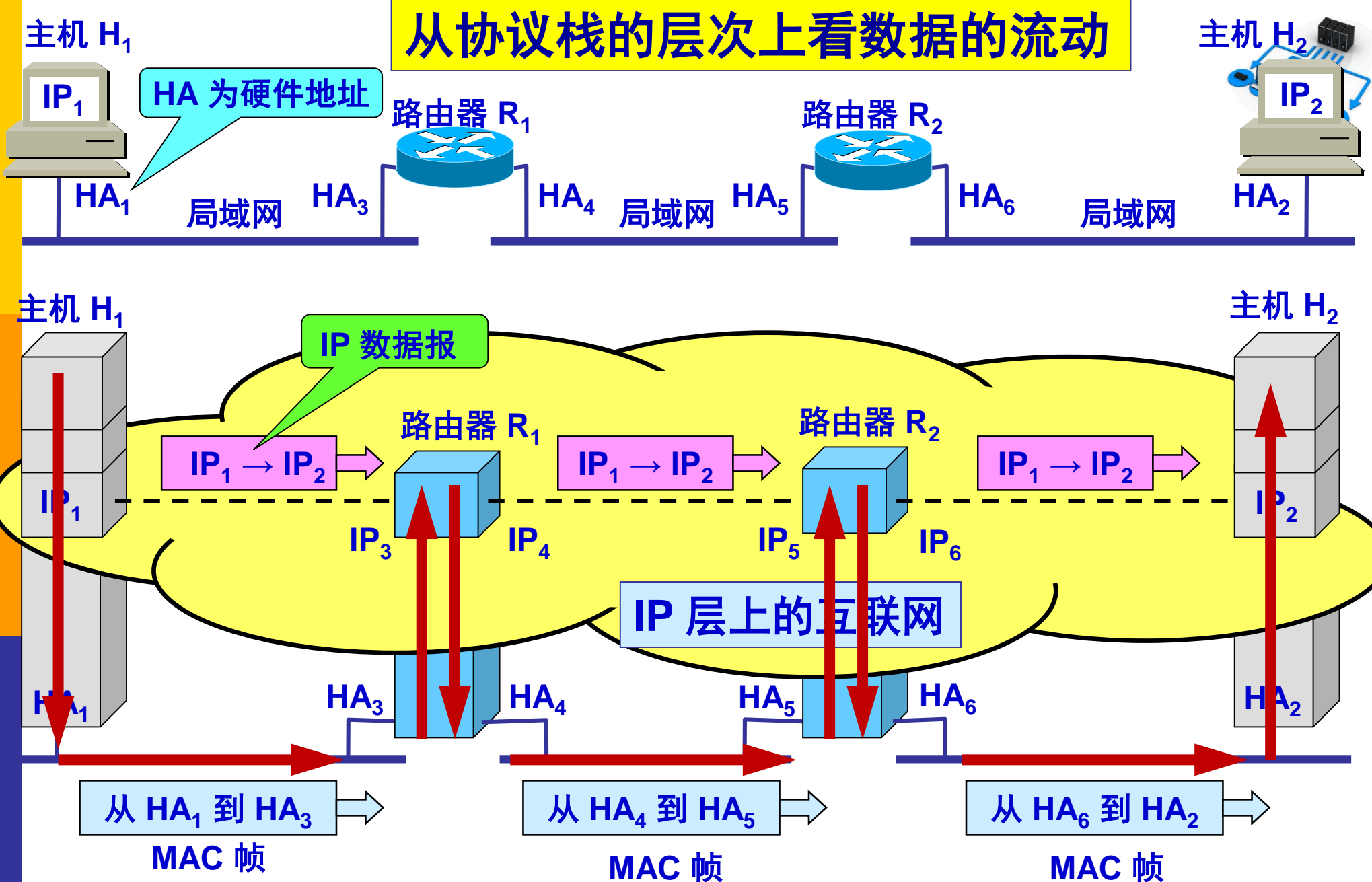
IP 地址放在 IP 数据报的首部，而硬件地址则放在 MAC 帧的首部。

IP 地址与硬件地址的区别

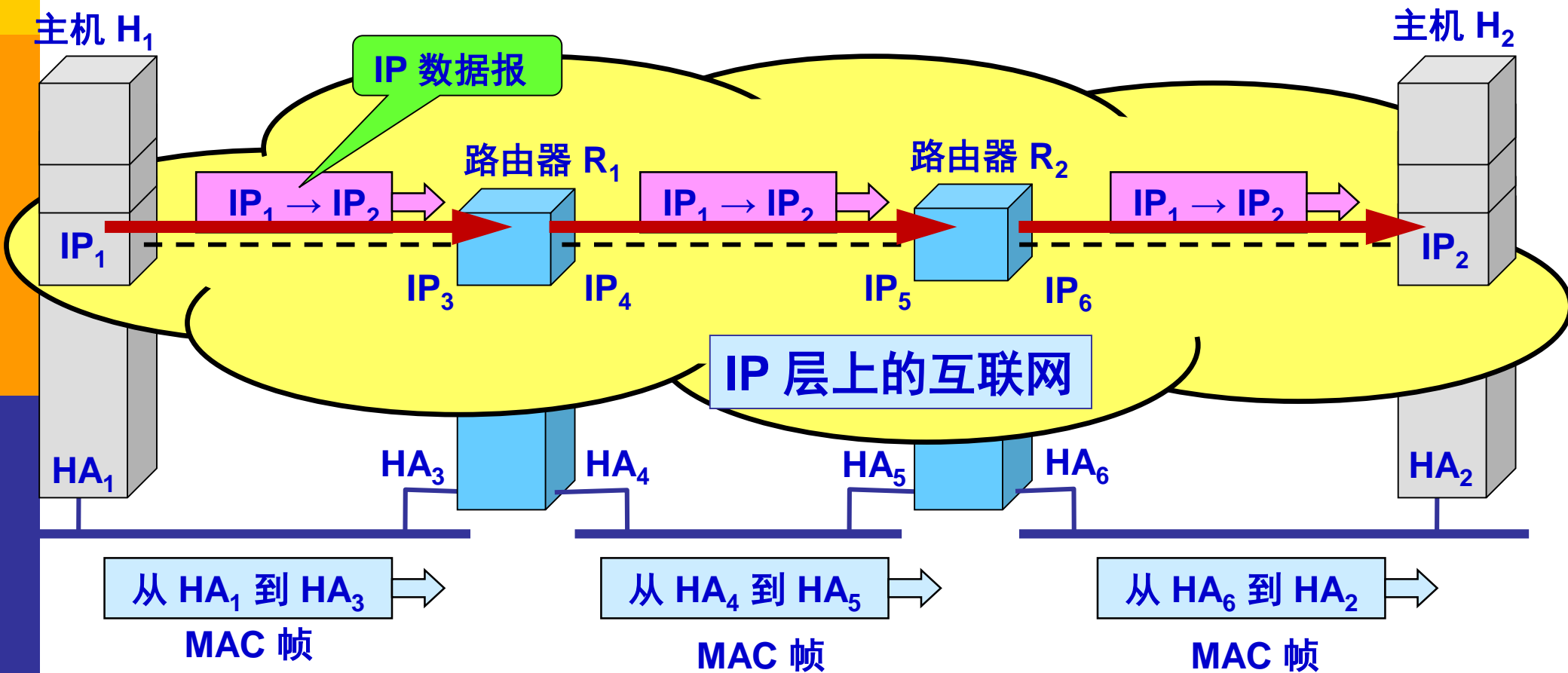


通信的路径：
 $H_1 \rightarrow$ 经过 R_1 转发 \rightarrow 再经过 R_2 转发 $\rightarrow H_2$

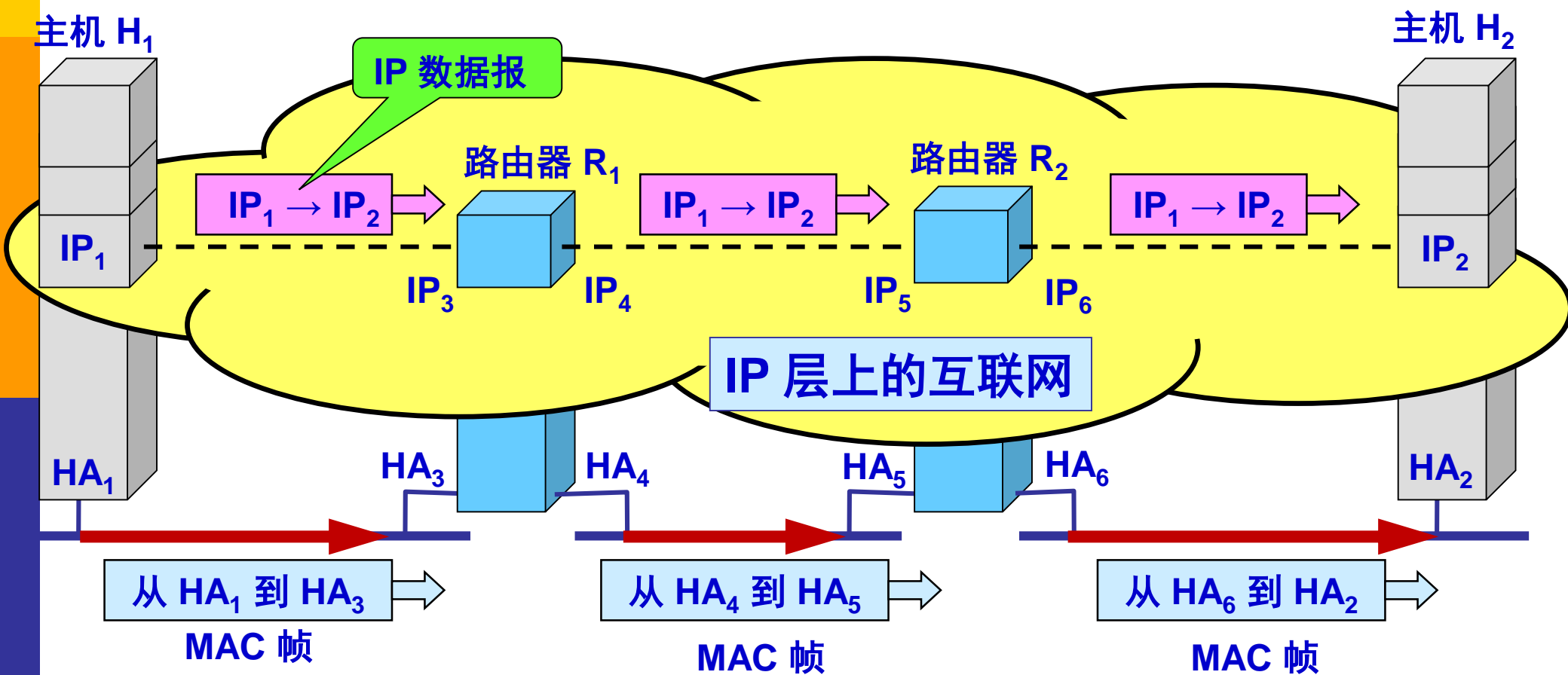
从协议栈的层次上看数据的流动



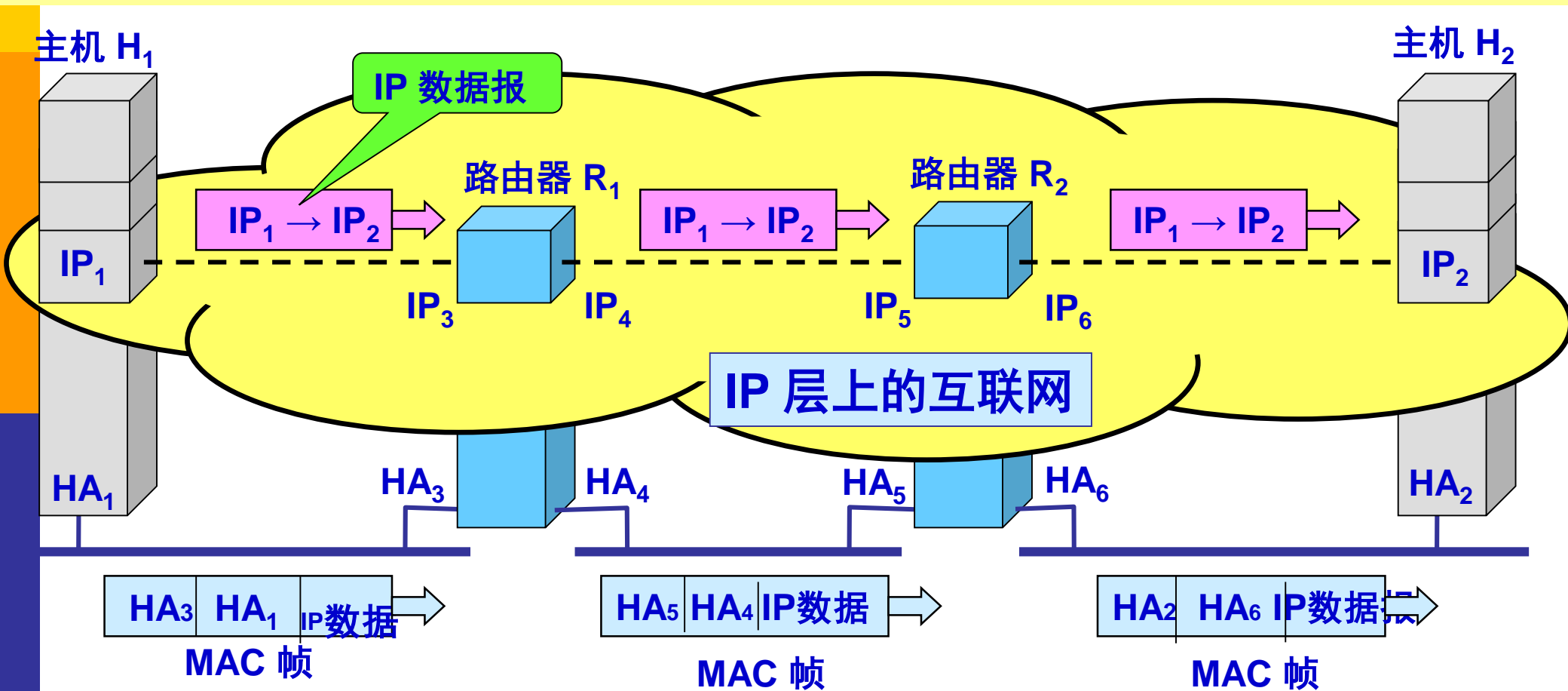
主机 H₁



在链路上看 MAC 帧的流动



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



主机 H_1

主机 H_2

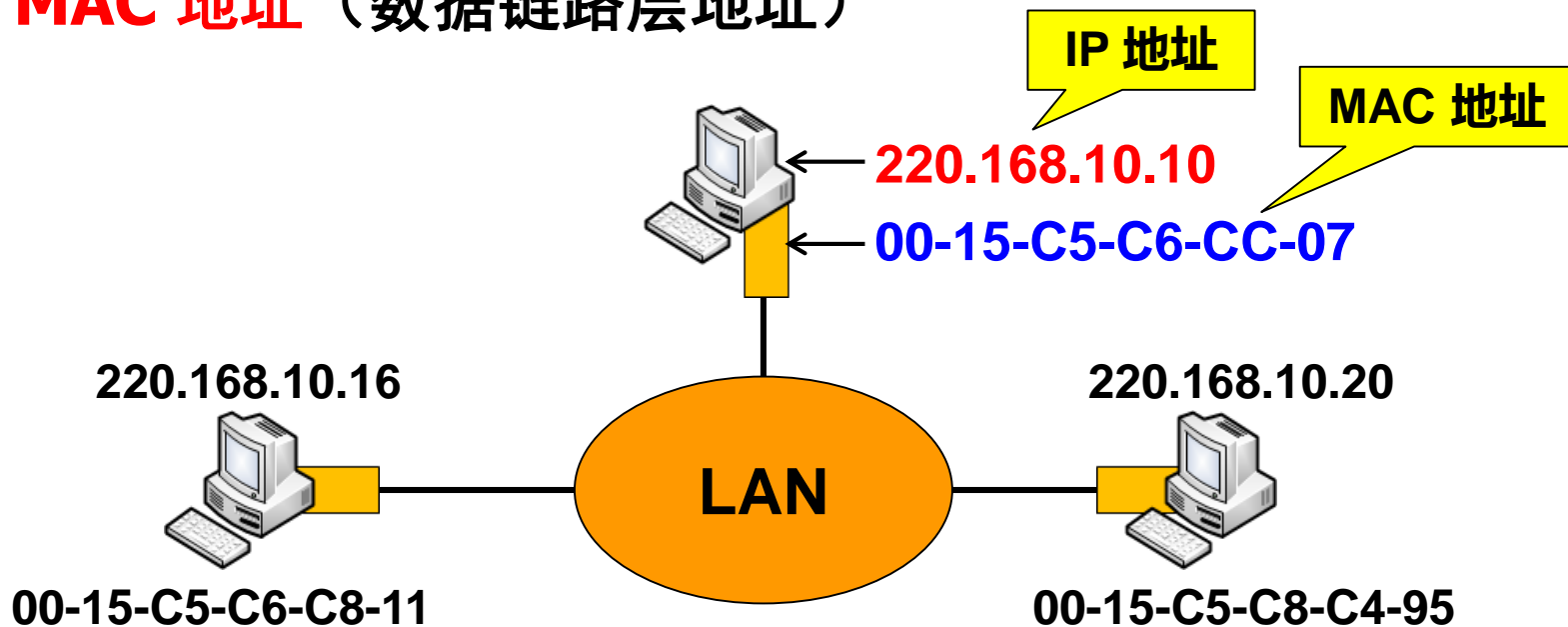


	在网络层 写入 IP 数据报首部的地址		在数据链路层 写入 MAC 帧首部的地址	
	源地址	目的地址	源mac地址	目的mac
从 H_1 到 R_1	IP_1	IP_2	?	?
从 R_1 到 R_2	IP_1	IP_2	?	?
从 R_2 到 H_2	IP_1	IP_2	?	?

4.1.4 地址解析协议 ARP



- 通信时使用了两个地址：
 - **IP 地址**（网络层地址）
 - **MAC 地址**（数据链路层地址）

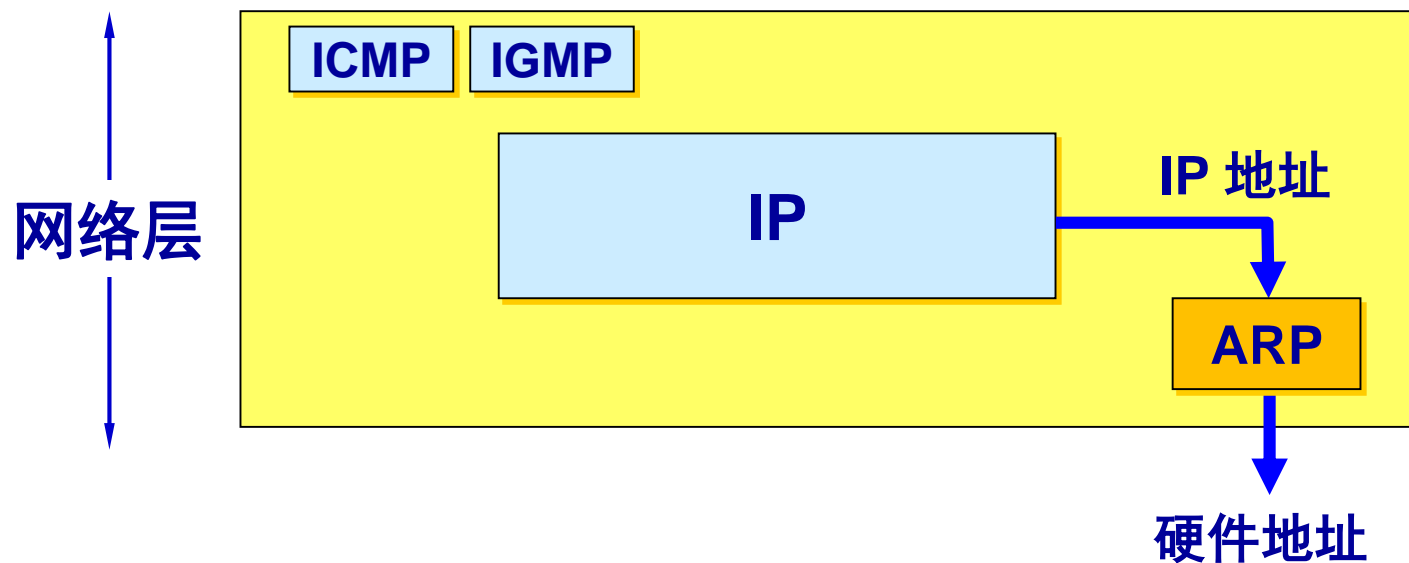


每个接口都有两个地址

地址解析协议 ARP 的作用



- 已经知道了一个机器（主机或路由器）的IP地址，如何找出其相应的硬件地址？
- 地址解析协议 ARP 就是用来解决这样的问题的。



ARP 作用：
从网络层使用的 IP 地址，
解析出在数据链路层使用的
硬件地址。

ARP 协议的作用

地址解析协议 ARP 要点



- 不管网络层使用的是什么协议，在实际网络的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** (ARP cache)，里面有所在的局域网上的各主机和路由器的 IP 地址到硬件地址的映射表。

< IP address; MAC address; TTL >

TTL (Time To Live): 地址映射有效时间。

地址解析协议 ARP 要点



- 当主机 A 欲向本局域网上的某个主机 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，再将此硬件地址写入 MAC 帧，然后通过局域网将该 MAC 帧发往此硬件地址。
 - 如没有，ARP 进程在本局域网上广播发送一个 ARP 请求分组。收到 ARP 响应分组后，将得到的 IP 地址到硬件地址的映射写入 ARP 高速缓存。

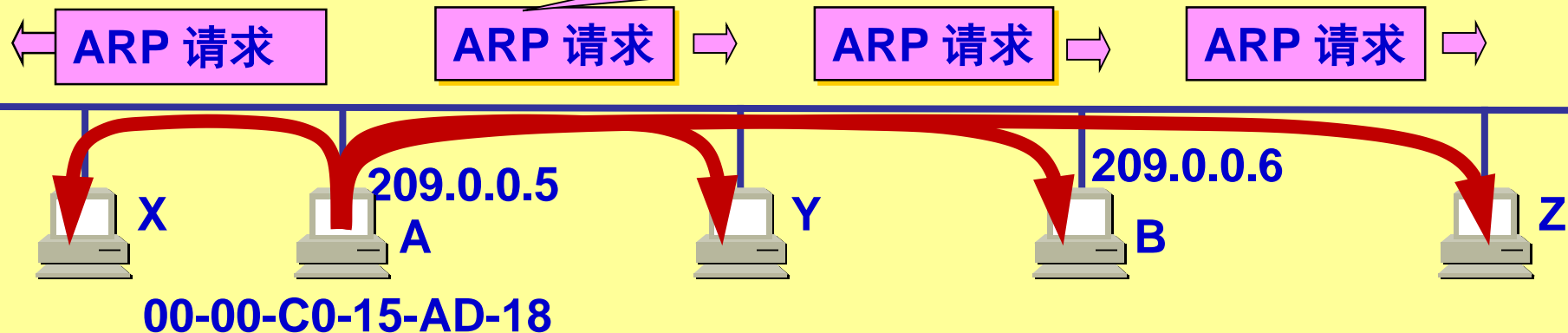
地址解析协议 ARP 要点



- **ARP请求分组**：包含发送方硬件地址 / 发送方 IP 地址 / **目标方硬件地址(未知时填 0)** / 目标方 IP 地址。
- **本地广播 ARP 请求**（路由器不转发ARP请求）。
- **ARP 响应分组**：包含发送方硬件地址 / 发送方 IP地址 / 目标方硬件地址 / 目标方 IP 地址。
- **ARP 分组封装在物理网络的帧中传输。**

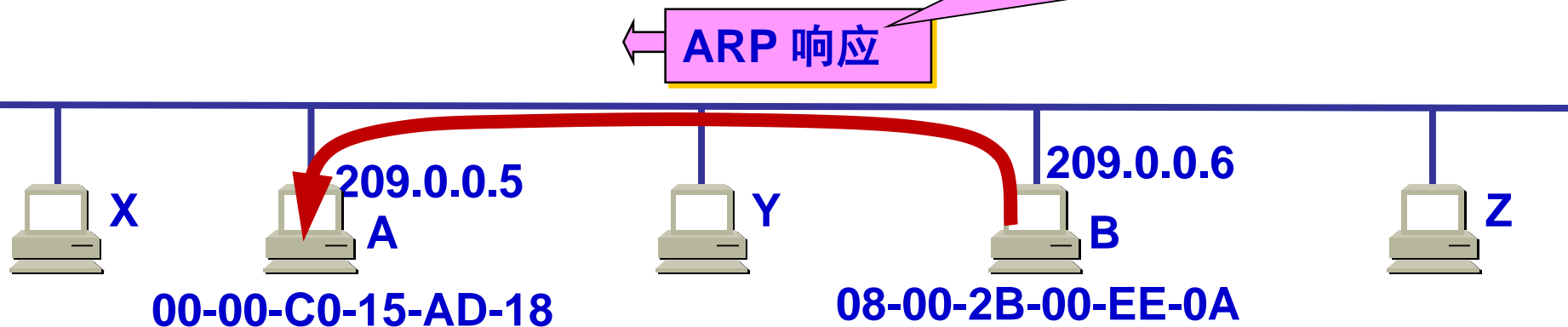
**主机 A 广播发送
ARP 请求分组**

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



ARP 高速缓存的作用



- 存放最近获得的 IP 地址到 MAC 地址的绑定，以减少 ARP 广播的数量。
- 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 发送数据报时就更方便了。

ARP缓存可以查看、添加和修改



- “arp -a”就可以查看ARP缓存中的内容

```
C:\Users\Administrator>arp -a

接口: 192.168.0.101 --- 0x3
Internet 地址      物理地址      类型
192.168.0.1        fc-75-16-ca-09-44 动态
192.168.0.255      ff-ff-ff-ff-ff-ff 静态
224.0.0.22         01-00-5e-00-00-16 静态
224.0.0.252        01-00-5e-00-00-fc 静态
224.0.1.255        01-00-5e-00-01-ff 静态
239.255.255.250    01-00-5e-7f-ff-fa 静态
255.255.255.255    ff-ff-ff-ff-ff-ff 静态

C:\Users\Administrator>
```

- 用 “arp -d”命令可以删除ARP缓存中所有的内容。
- 用 “arp -s”命令可以在ARP缓存中添加IP地址与MAC地址的映射关系，类型为static（静态）

注意的问题



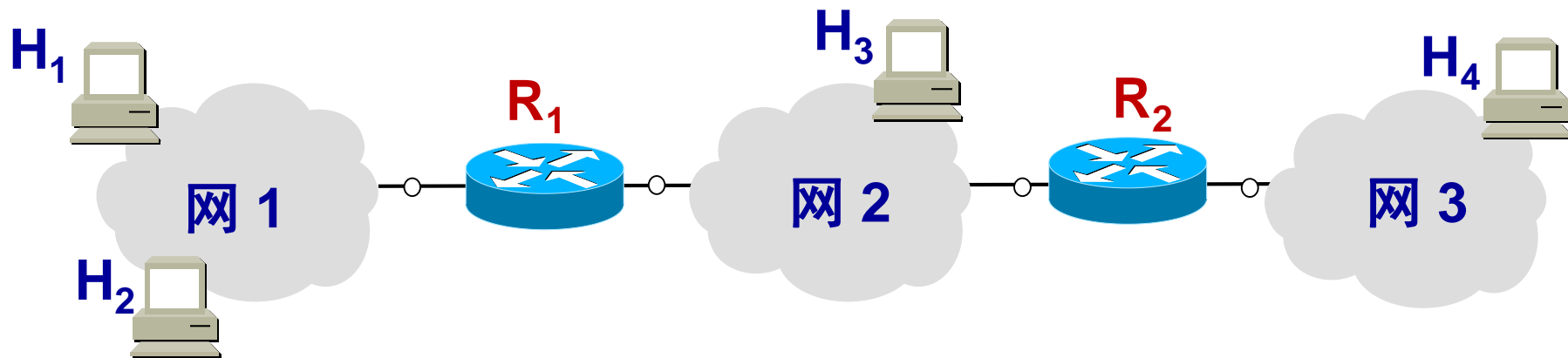
- ARP 是解决**同一个局域网**上的主机或路由器的 IP 地址和硬件地址的映射问题。
- 如果所要找的主机和源主机**不在同一个局域网**上，那么**就要通过 ARP 找到一个位于本局域网上的某个路由器的硬件地址**，然后把分组发送给这个路由器，让这个路由器把分组转发给下一个网络。剩下的工作就由下一个网络来做。

注意的问题（续）



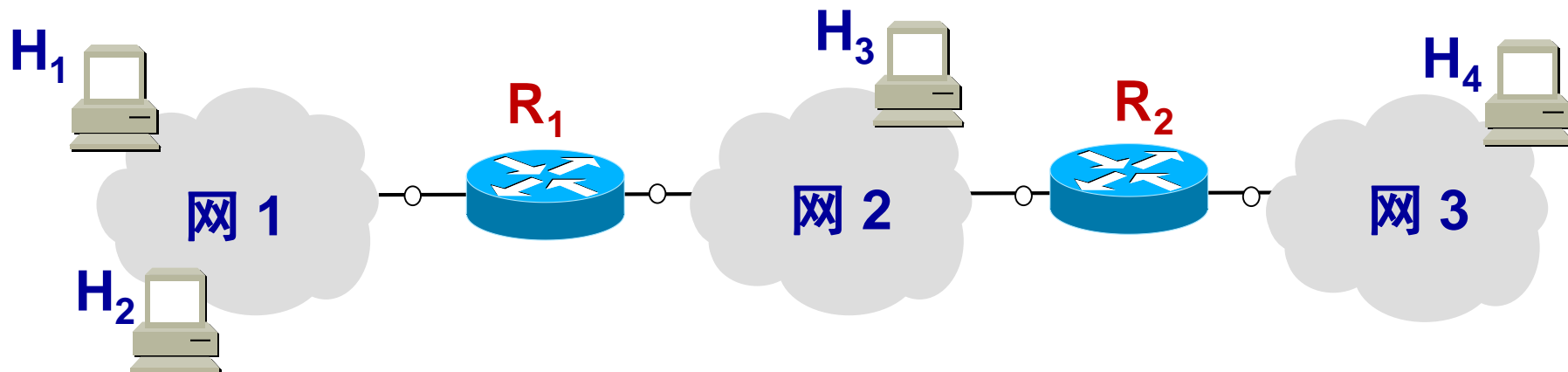
- 从 IP 地址到硬件地址的**解析是自动进行的**，主机的用户对这种地址解析过程是不知道的。
- 只要主机或路由器要和本网络上的另一个已知 IP 地址的主机或路由器进行通信，ARP 协议就会自动地将该 IP 地址解析为链路层所需要的硬件地址。

总结：使用 ARP 的四种典型情况



1. 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。（如H1→H2）
2. 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。（如H1→H3或 H1→H4）

使用 ARP 的四种典型情况



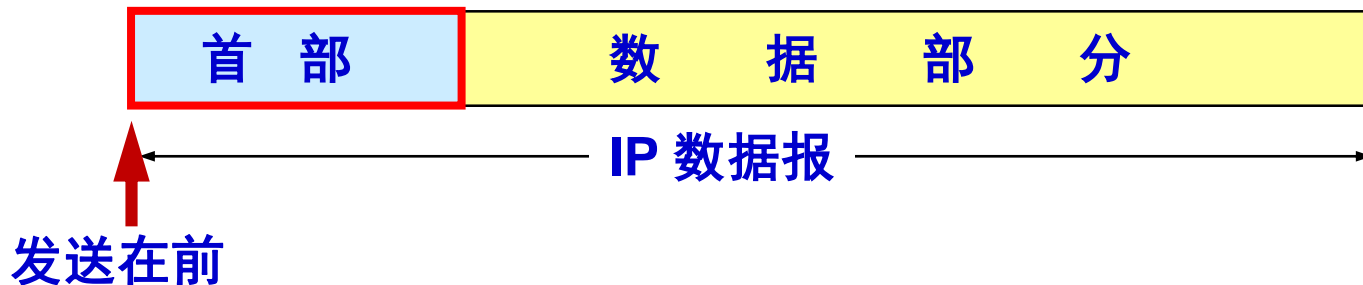
- 3. 发送方是路由器，要把 IP 数据报转发到本网络上的一个主机。这时用 ARP 找到目的主机的硬件地址。(如R1-H3)
- 4. 发送方是路由器，要把 IP 数据报转发到另一个网络上的一个主机。这时用 ARP 找到本网络上另一个路由器的硬件地址。剩下的工作由这个路由器来完成。(R1->H4)

4.1.5 IP 数据报的格式

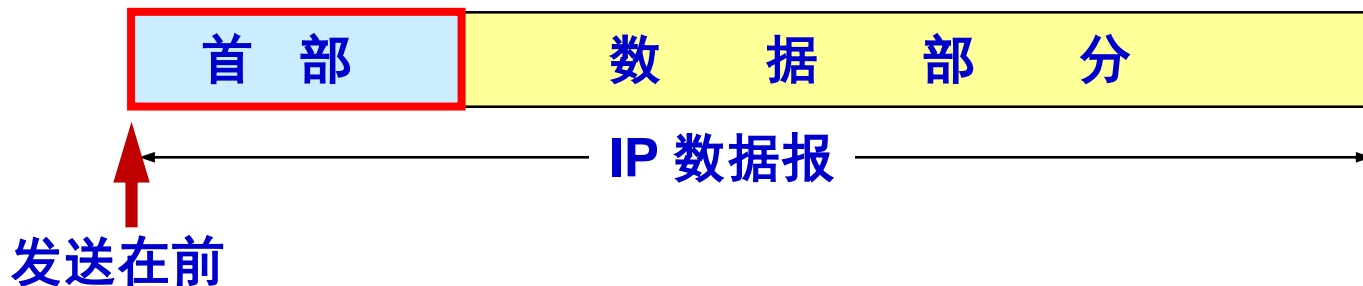
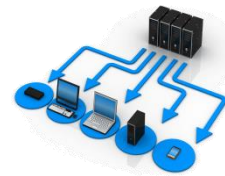


- 一个 IP 数据报由**首部**和**数据**两部分组成。
- 首部的前一部分是固定长度，共 20 字节，是所有 IP 数据报必须具有的。
- 在首部的固定部分的后面是一些可选字段，其长度是可变的。

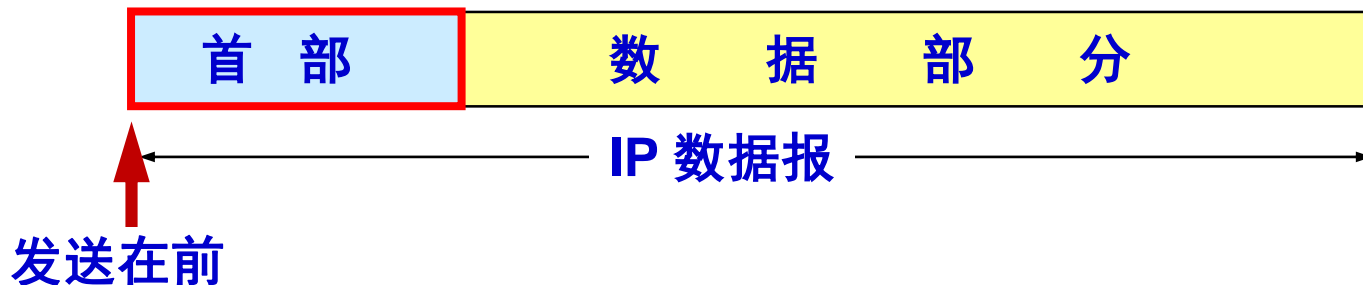
IP 数据报由首部和数据两部分组成



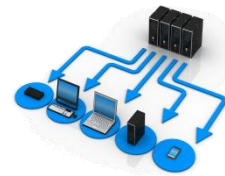
首部的前一部分是固定长度，共 20 字节，是所有 IP 数据报必须具有的。



可选字段，其长度是可变的



1. IP 数据报首部的固定部分中的各字段



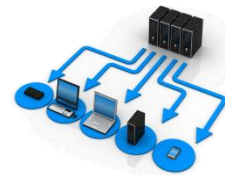
版本——占 4 位，指 IP 协议的版本。
目前的 IP 协议版本号为 4 (即 IPv4)。

1. IP 数据报首部的固定部分中的各字段



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。

1. IP 数据报首部的固定部分中的各字段



区分服务——占 8 位，用来获得更好的服务。

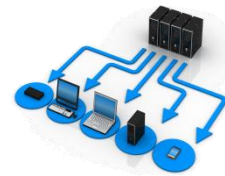
在旧标准中叫做服务类型，但实际上一直未被使用过。

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

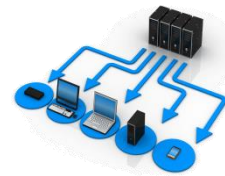
1. IP 数据报首部的固定部分中的各字段



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。

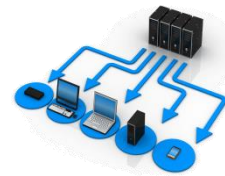
总长度必须不超过最大传送单元 MTU。

1. IP 数据报首部的固定部分中的各字段



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

1. IP 数据报首部的固定部分中的各字段



标志(flag) ——占 3 位，目前只有前两位有意义。

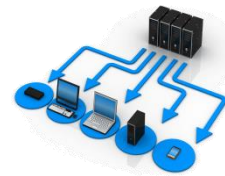
标志字段的最低位是 MF (More Fragment)。

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。

1. IP 数据报首部的固定部分中的各字段



片偏移——占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。

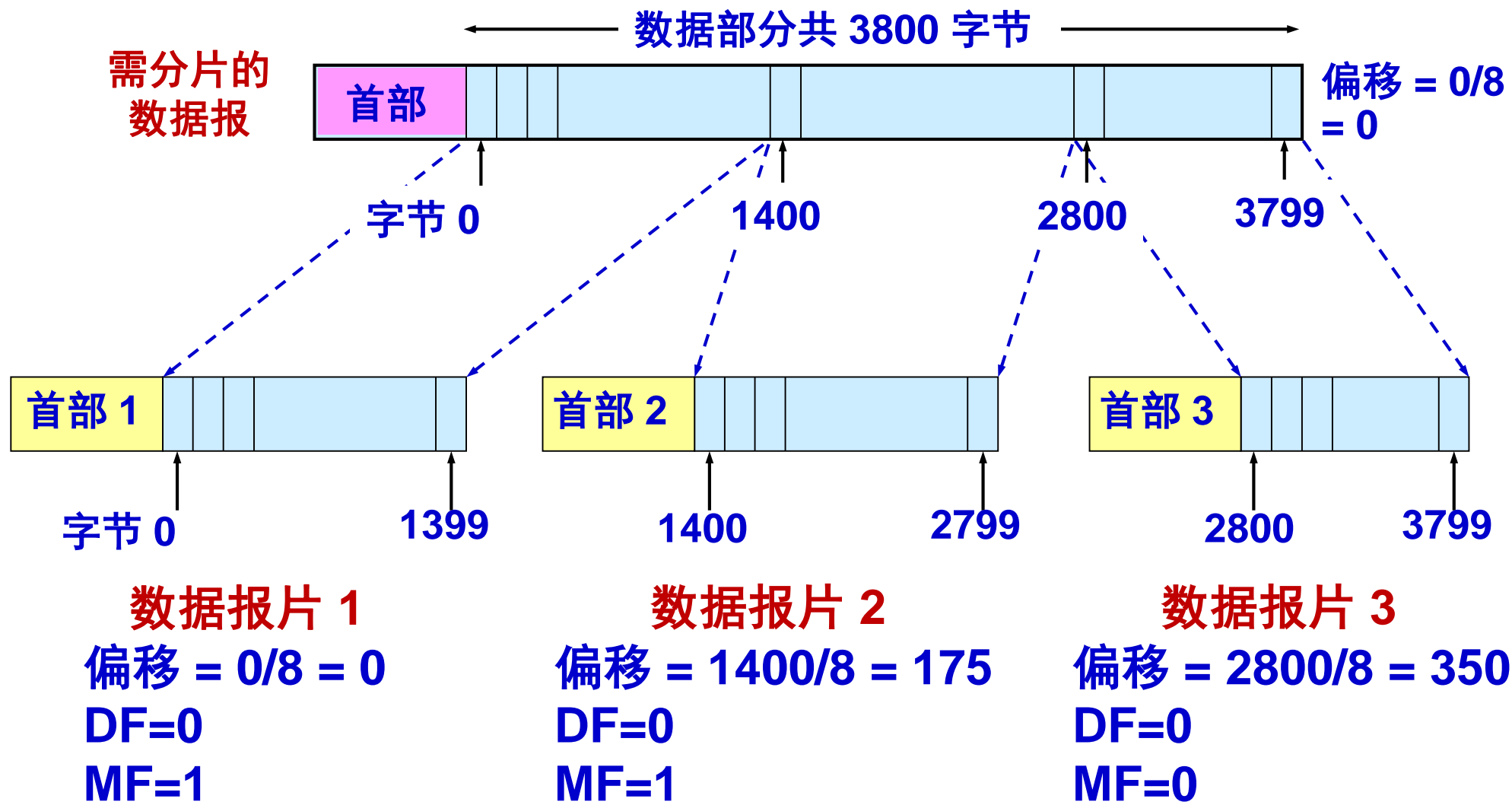
片偏移以 8 个字节为偏移单位。

【例4-1】 IP 数据报分片



- 一数据报的总长度为 3820 字节，其数据部分的长度为 3800 字节（使用固定首部），需要分片为长度不超过 1420 字节的数据报片。
- 因固定首部长度为 20 字节，因此每个数据报片的数据部分长度不能超过 1400 字节。
- 于是分为 3 个数据报片，其数据部分的长度分别为 1400、1400 和 1000 字节。
- 原始数据报首部被复制为各数据报片的首部，但必须修改有关字段的值。

【例4-1】 IP 数据报分片



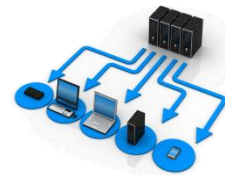
【例4-1】 IP 数据报分片



IP 数据报首部中与分片有关的字段中的数值

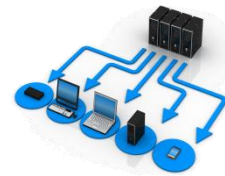
	总长度	标识	MF	DF	片偏移
原始数据报	3820	12345	0	0	0
数据报片1	1420	12345	1	0	0
数据报片2	1420	12345	1	0	175
数据报片3	1020	12345	0	0	350

1. IP 数据报首部的固定部分中的各字段



生存时间——占8位，记为 TTL (Time To Live)，
指示数据报在网络中可通过的路由器数的最大值。
最大值是多少？

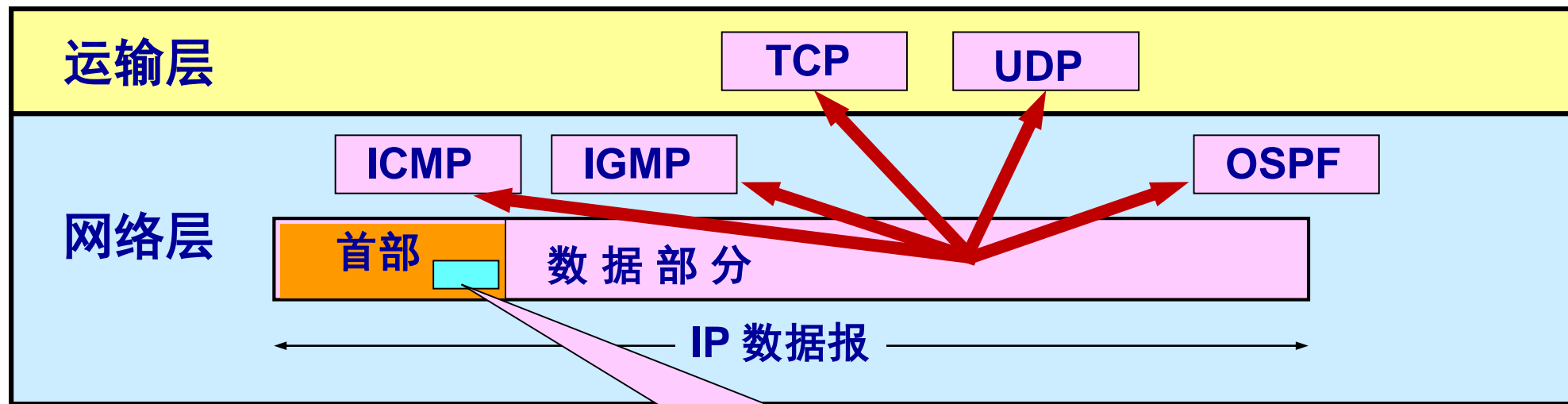
1. IP 数据报首部的固定部分中的各字段



协议——占8位，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那个处理过程



IP 协议支持多种协议， IP 数据报可以封装多种协议 PDU。

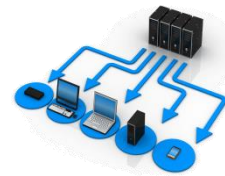


协议字段指出应将数据部分交给哪一个进程

TCP值 , UDP值

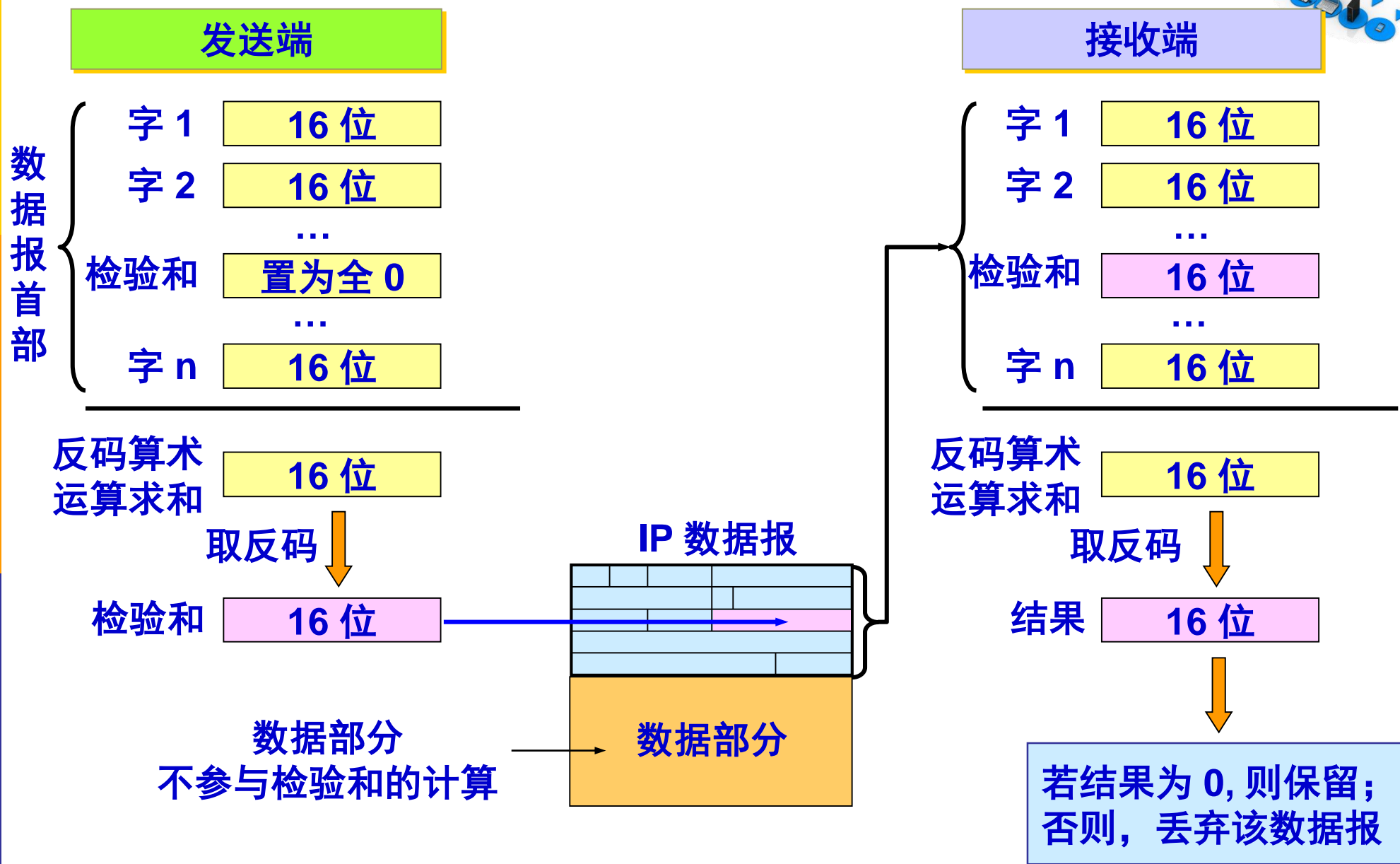
ICMP值 ?

1. IP 数据报首部的固定部分中的各字段



首部检验和——占16位，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。

IP 数据报首部检验和的计算采用 16 位二进制反码求和算法



注意：回卷



- 1010 0001 0111 1011
- 1010 1011 0000 0011
- 10000 1100 0111 1110
- 将高位的1 和16位结果 相加。
- 0000 1100 0111 1110
- 1
- 0000 1100 0111 1111

1. IP 数据报首部的固定部分中的各字段



源地址和目的地址都各占 4 字节

分析IP 数据报首部



4	5	0	28	
1			0	0
4		17	0	
10.12.14.5				
12.6.7.9				

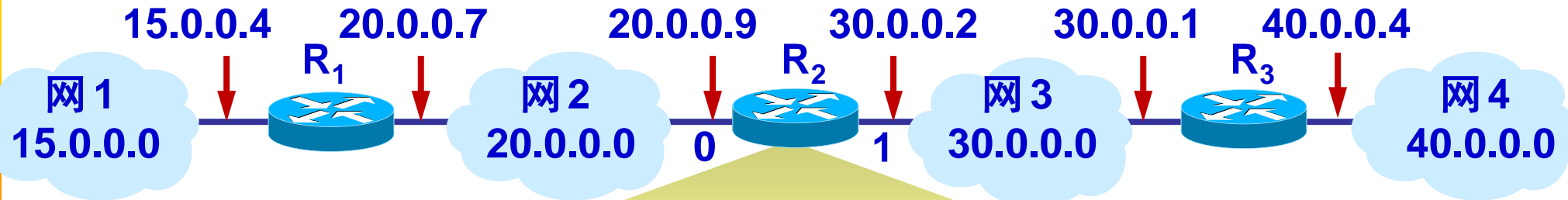
1. IP分组的总长度 为？ 数据部分 的字节数？
2. 这个IP 分组的标识是多少？
3. IP分组中封装的数据部分是哪个上层协议的数据？
4. 这个分组的MF和DF 的值 是多少？

4.1.6 IP 层转发分组的流程



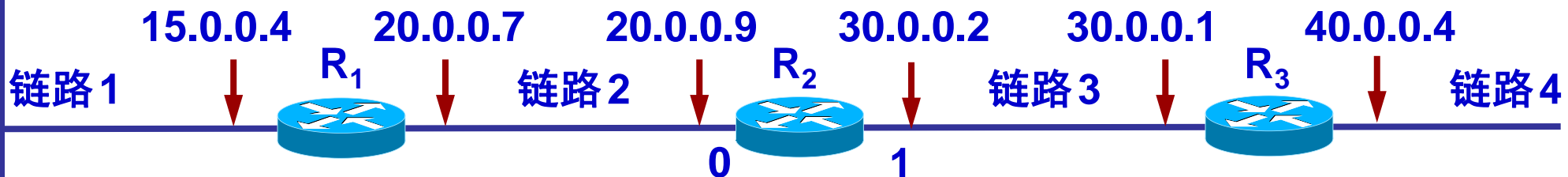
- 假设：有四个 A 类网络通过三个路由器连接在一起。每一个网络上都可能有成千上万个主机。
- 可以想像，**若按目的主机号来制作路由表**，每一个路由表就有 4 万个项目，即 4 万行（每一行对应于一台主机），则所得出的路由表就会过于庞大。
- 但**若按主机所在的网络地址来制作路由表**，那么每一个路由器中的路由表就只包含 4 个项目（每一行对应于一个网络），这样就可使路由表大大简化。

在路由表中，对每一条路由，最主要的是
(目的网络地址，下一跳地址)



路由器 R₂ 的路由表

目的主机所在的网络	下一跳地址
20.0.0.0	直接交付，接口 0
30.0.0.0	直接交付，接口 1
15.0.0.0	20.0.0.7
40.0.0.0	30.0.0.1



查找路由表



根据目的网络地址就能确定下一跳路由器，这样做的结果是：

- IP 数据报最终一定可以找到目的主机所在目的网络上的路由器（可能要通过多次的间接交付）。
- 只有到达最后一个路由器时，才试图向目的主机进行直接交付。

特定主机路由



- 虽然互联网所有的分组转发都是**基于目的主机所在的网络**，但在大多数情况下都允许有这样的特例，即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 and 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互联网连接，那么在这种情况下使用默认路由是非常合适的。

默认路由举例

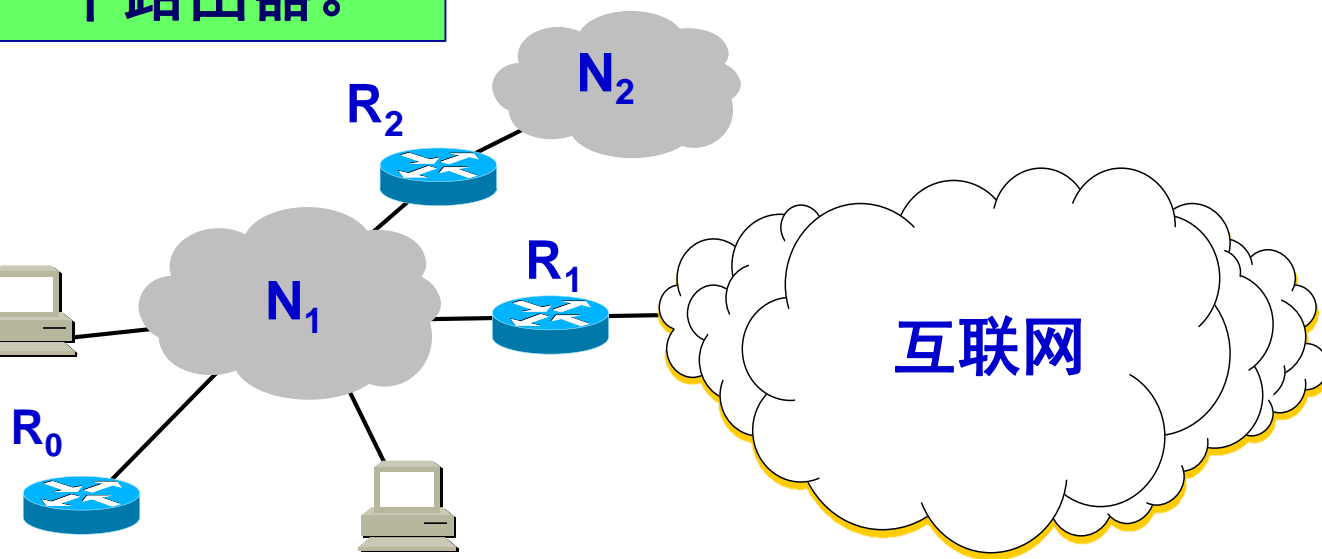


只要目的网络不是 N_1 和 N_2 ,
就一律选择默认路由,
把数据报先间接交付路由器 R_1 ,
让 R_1 再转发给下一个路由器。

R0路由表

目的网络	下一跳
N_1	直接
N_2	R_2
默认	R_1

0.0.0.0.



路由器 R_1 充当网络 N_1 的默认路由器

转发的实际过程



- IP 数据报的首部中**没有**地方可以用来指明“下一跳路由器的 IP 地址”。
- 当路由器收到待转发的数据报，不是将下一跳路由器的 IP 地址填入 IP 数据报，而是**送交下层**的网络接口软件。
- 网络接口软件**使用 ARP** 负责将下一跳路由器的 IP 地址转换成硬件地址，并将此硬件地址放在链路层的 MAC 帧的首部，然后根据这个硬件地址找到下一跳路由器。

路由器分组转发算法



- (1) 从数据报的首部提取**目的主机的 IP 地址 D** ，得出**目的网络地址为 N** 。
- (2) 若网络 N 与此路由器直接相连，则把数据报**直接交付**目的主机 D ；否则是**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**，则把数据报传送给路由表中所指明的下一跳路由器；否则，执行 (4)。
- (4) 若路由表中有**到达网络 N 的路由**，则把数据报传送给路由表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**，则把数据报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。

关于路由表



- 路由表没有给分组指明到某个网络的完整路径。
- 路由表指出，到某个网络应当先到某个路由器（即下一跳路由器）。
- 在到达下一跳路由器后，再继续查找其路由表，知道再下一步应当到哪一个路由器。
- 这样一步一步地查找下去，直到最后到达目的网络。

4.3 划分子网和构造超网



- 4.3.1 划分子网
- 4.3.2 使用子网时分组的转发
- 4.3.3 无分类编址 CIDR（构造超网）

4.3.1 划分子网



1. 从两级 IP 地址到三级 IP 地址

- 在 ARPANET 的早期，IP 地址的设计确实不够合理：
 - (1) IP 地址空间的利用率有时很低。
 - (2) 给每一个物理网络分配一个网络号会使路由表变得太大因而使网络性能变坏。
 - (3) 两级的 IP 地址不够灵活。

三级 IP 地址

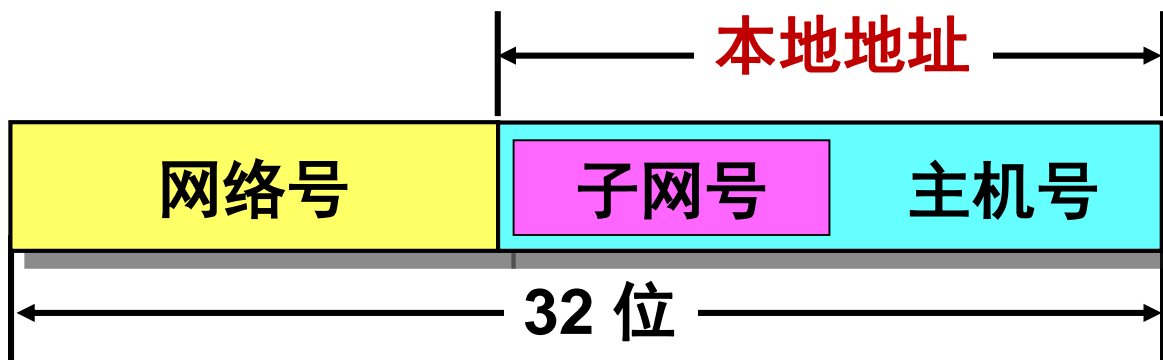


- 从 1985 年起在 IP 地址中又增加了一个“**子网号字段**”，使两级的 IP 地址变成**三级的 IP 地址**。
- 这种做法叫做**划分子网** (subnetting)。
- 划分子网已成为互联网的正式标准协议。

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。



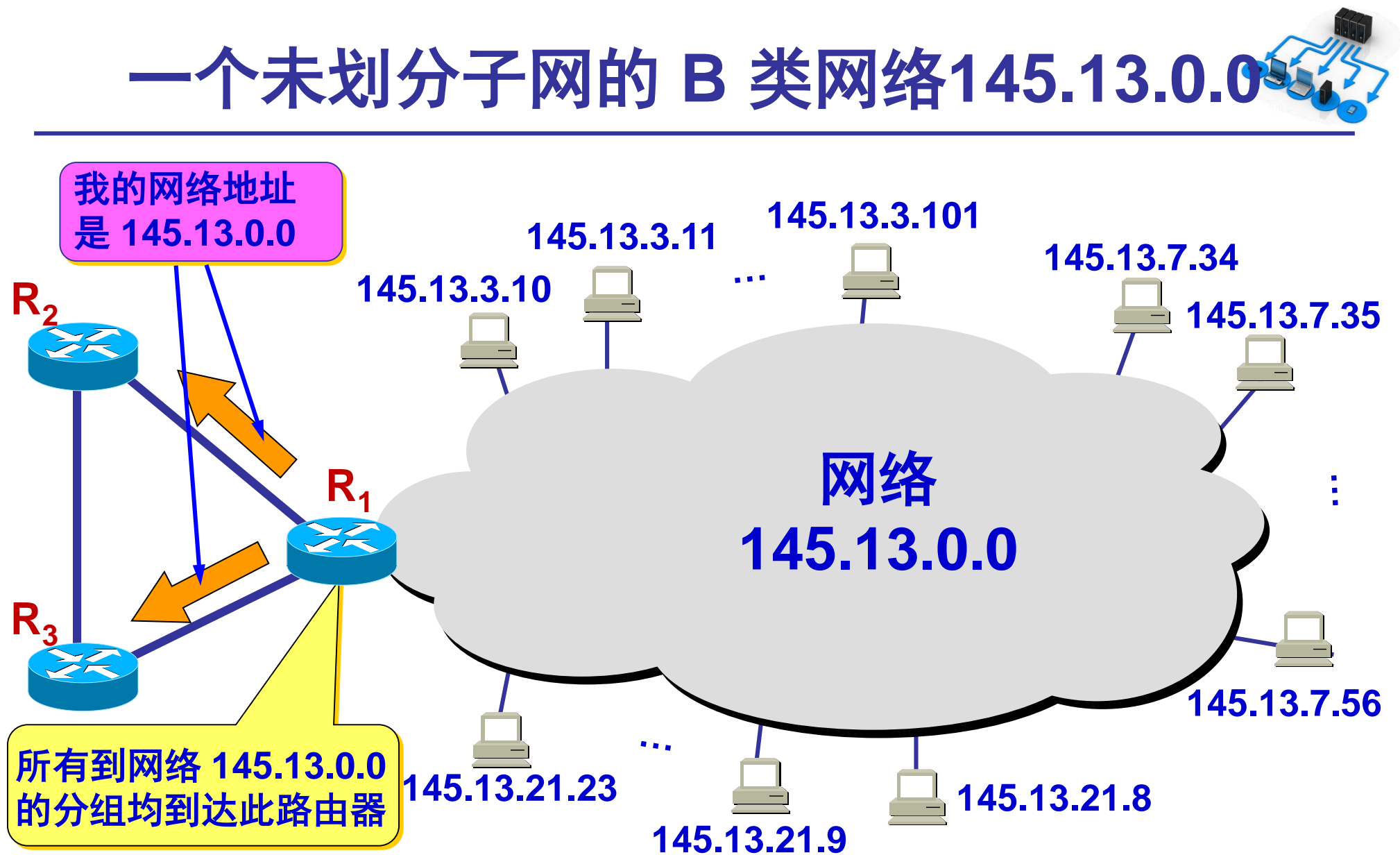
IP地址 ::= {<网络号>, <子网号>, <主机号>} (4-2)

划分子网的基本思路（续）

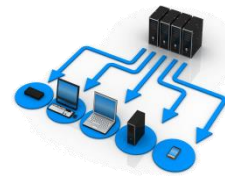


- 凡是从其他网络发送给本单位某个主机的 IP 数据报，仍然是根据 IP 数据报的**目的网络号 net-id**，先找到连接在**本单位网络上的路由器**。
- 然后**此路由器**在收到 IP 数据报后，再按**目的网络号 net-id** 和**子网号 subnet-id** 找到目的子网。
- 最后就将 IP 数据报直接交付目的主机。

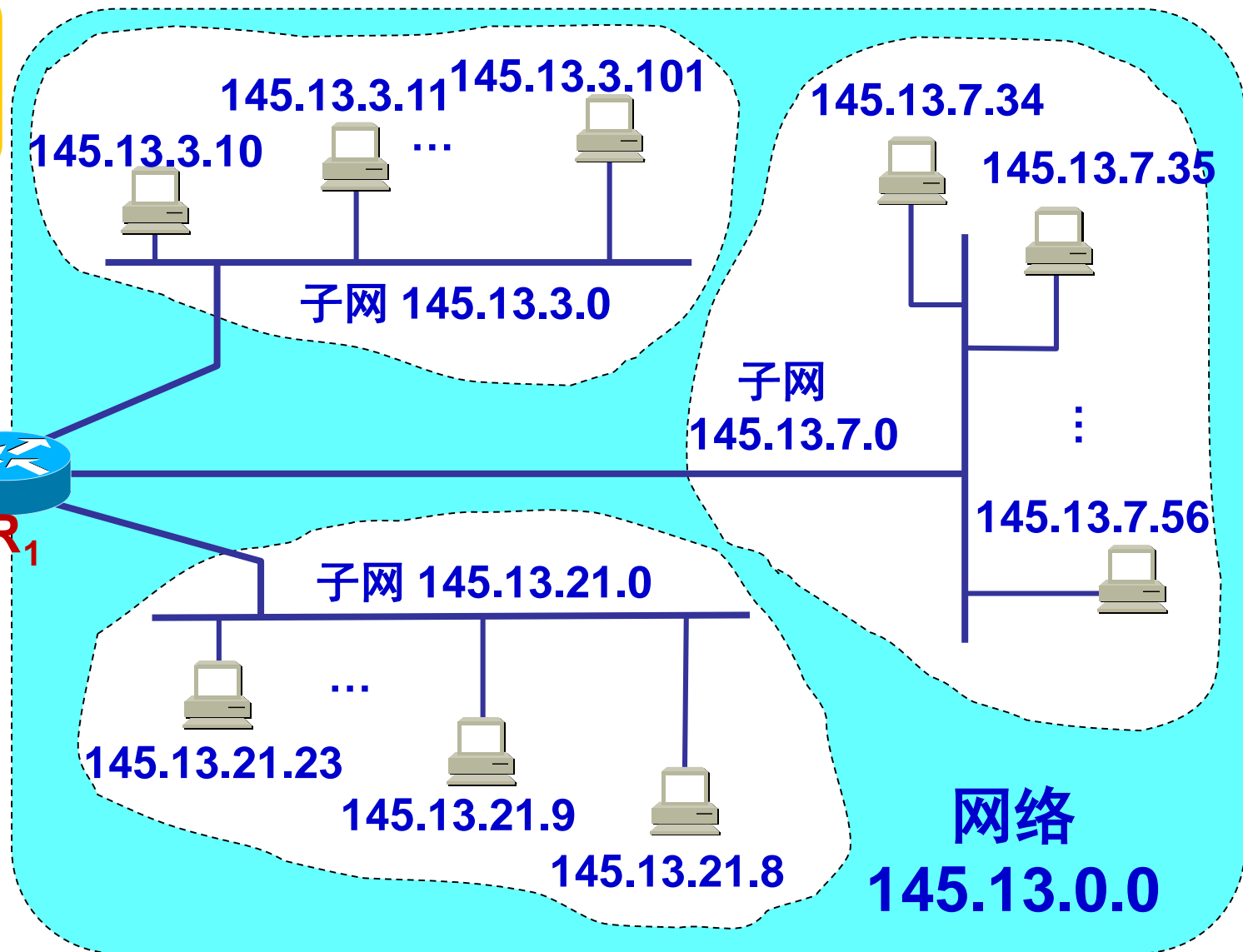
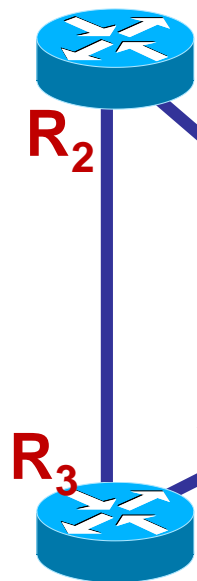
一个未划分子网的 B 类网络 145.13.0.0



划分为三个子网后对外仍是一个网络



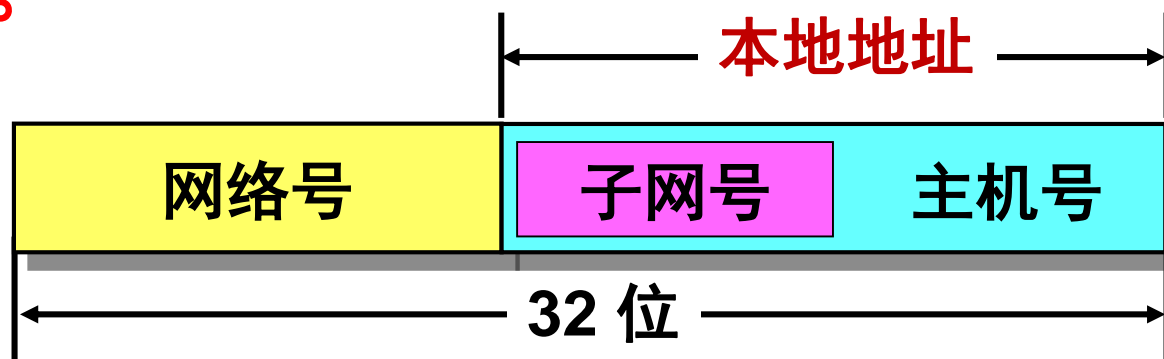
所有到达网络
145.13.0.0 的分组均
到达此路由器



划分子网后变成了三级结构



- 当没有划分子网时，IP 地址是两级结构。
- 划分子网后 IP 地址就变成了**三级结构**。
- 划分子网只是把 IP 地址的主机号 host-id 这部分进行再划分，而**不改变 IP 地址原来的网络号 net-id**。



划分子网后变成了三级结构



■ 优点

- 减少了 IP 地址的浪费
 - 使网络的组织更加灵活
 - 更便于维护和管理
- 划分子网纯属一个单位内部的事情，对外部网络透明，对外仍然表现为没有划分子网的一个网络。

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。

规则：

- 子网掩码长度 = 32 位
- **某位 = 1**：IP地址中的对应位为网络号和子网号
- **某位 = 0**：IP地址中的对应位为主机号



(IP 地址) AND (子网掩码) = 网络地址

两级 IP 地址

网络号	主机号
-----	-----

三级 IP 地址

网络号	子网号	主机号
-----	-----	-----

逐位进行 AND 运算

三级 IP 地址
的子网掩码

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
---------------------------------	-----------------	-----------------

子网的
网络地址

网络号	子网号	0
-----	-----	---

默认子网掩码



A类地址	网络地址	网络号	主机号为全 0
	默认子网掩码 255.0.0.0	1 1 1 1 1 1 1 1 0	
B类地址	网络地址	网络号	主机号为全 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	
C类地址	网络地址	网络号	主机号为全 0
	默认子网掩码 255.255.255.0	1 0 0 0 0 0 0 0 0	

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。

B 类地址的子网划分选择（使用固定长度子网）



子网号的位数	子网掩码	子网数	每个子网的主机数
2	255.255.192.0	2	16382
3	255.255.224.0	6	8190
4	255.255.240.0	14	4094
5	255.255.248.0	30	2046
6	255.255.252.0	62	1022
7	255.255.254.0	126	510
8	255.255.255.0	254	254
9	255.255.255.128	510	126
10	255.255.255.192	1022	62
11	255.255.255.224	2046	30
12	255.255.255.240	4094	14
13	255.255.255.248	8190	6
14	255.255.255.252	16382	2

表中的“子网号的位数”中没有 0, 1, 15 和 16 这四种情况，因为这没有意义。

【例4-3】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
-----	---	----	---	----	---	---

划分子网练习



- **【例3】** 如果将一个C类网络192. 168. 1. 0划分为4个子网，请问：
 - ①新的子网掩码是多少？
 - ②每个子网的主机数是多少？
 - ③每个子网的子网地址？子网广播地址？子网内主机IP地址范围？

分析：



- ①因为要划分为4个子网，而 $2^2 < (4+2) \leq 2^3$
（不使用子网号部分全0或全1的子网），因此需
要从主机号部分借用3个比特位表示相应子网，
相应的子网掩码中由24个1增加为27个1。
- 255. 255. 255. 0 → 255. 255. 255. 224

分析：



- ②由于借用3位标识子网，因此剩余的表示主机的主机号部分只有 $8-3=5$ 个比特位，所以每个子网的主机数= $2^5-2=30$ 台。

分析：



- ③由于划分子网并不改变IP地址的网络号部分，因此整个网络中的IP地址前24比特保持不变，来表示子网的3个bit 位的变化形式有4种，分别是000~111，剩余的5个比特位表示主机，变化范围是 $(00\ 000)_2 \rightarrow (11\ 111)_2$ ，因为主机号部分全为0对应的IP地址是一个子网地址，主机号部分全为1对应的IP地址是一个子网广播地址，因此每个子网内主机IP地址的主机号部分变化范围是 $(00\ 001)_2 \rightarrow (11\ 110)_2$ ，对应的地址个数是 $2^5 - 2 = 30$ 个。

练习 192.168.1.0



- 四个子网的主机号码范围是（如果子网号取001, 010, 011, 100（二进制）），写出主机地址范围。
- 子网1：192.168.1.**33**~192.168.1.**62**
- 最小的主机号：192.168.1.**001** 00001
- 最大的主机号：192.168.1.**001** 11110
- 子网2：？
- 子网3：？
- 子网4：？



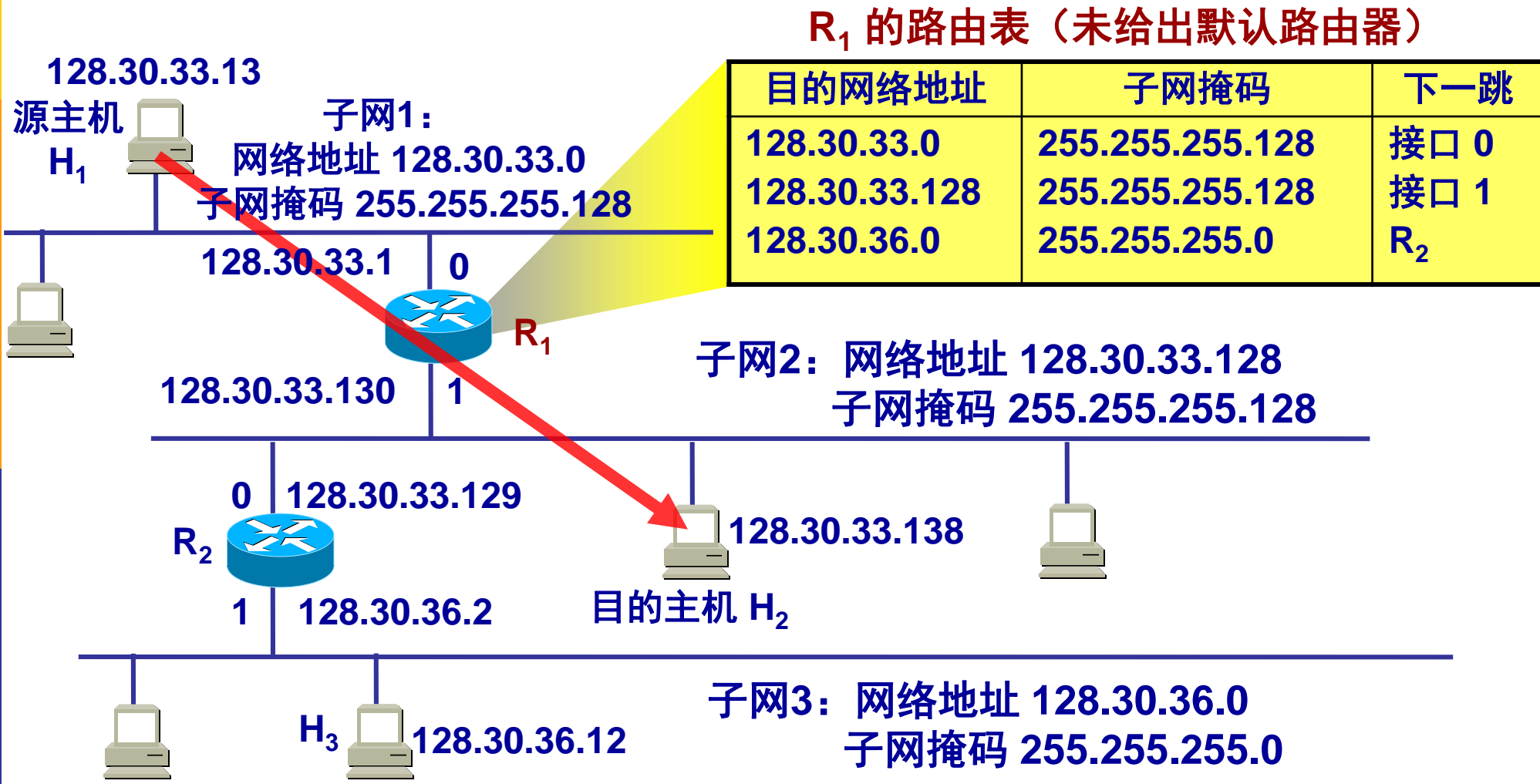
-
- 子网1: 192.168.1.33~192.168.1.62
 - 子网2: 192.168.1.65~192.168.1.94
 - 子网3: 192.168.1.97~192.168.1.126
 - 子网4: 192.168.1.129~192.168.1.158

在划分子网情况下路由器转发分组的算法



- (1) 从收到的分组的首部提取**目的 IP 地址 D** 。
- (2) 先用 本网络的**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，将**子网掩码和 D 逐位相“与”**。若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4-4】 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。



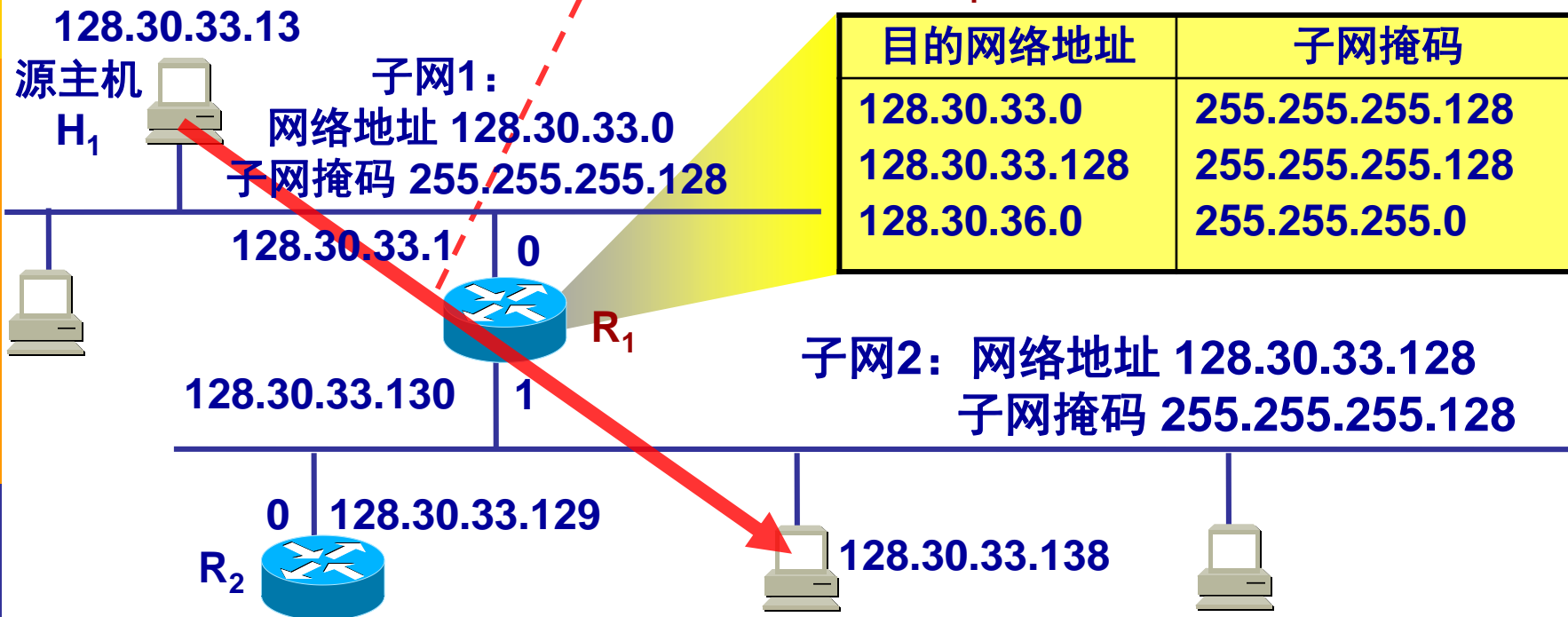
主机 H_1 要发送分组给 H_2



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。

128	→	10000000
138	→	10001010
<hr/>		
逐比特 AND 操作后		10000000 → 128

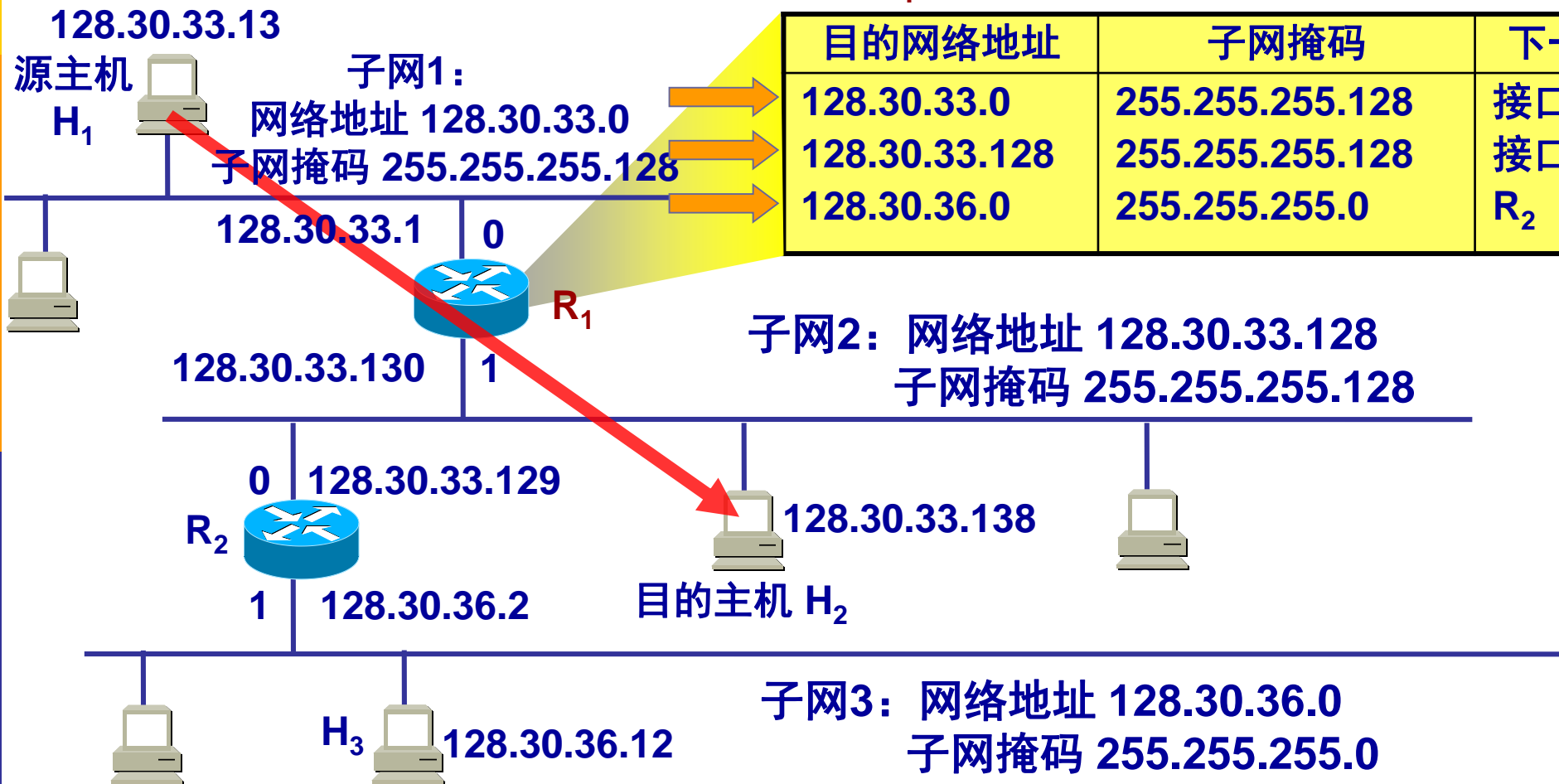
逐比特 AND 操作	255.255.255.128
	128. 30. 33.138
	<hr/>
	128. 30. 33.128

≠ H_1 的网络地址

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

R_1 的路由表（未给出默认路由器）

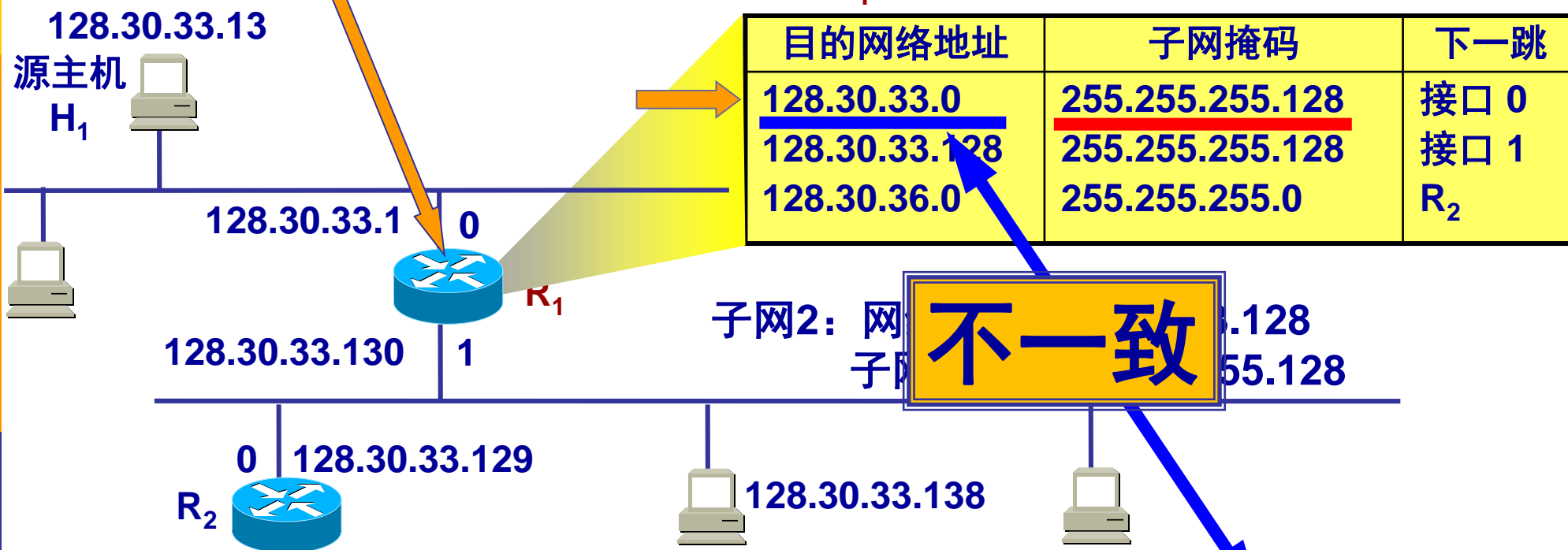
目的网络地址	子网掩码	下一跳
128.30.33.0	255.255.255.128	接口 0
128.30.33.128	255.255.255.128	接口 1
128.30.36.0	255.255.255.0	R_2



路由器 R₁ 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

R₁ 收到的分组的目的 IP 地址: 128.30.33.138

R₁ 的路由表 (未给出默认路由器)



255.255.255.128 AND 128.30.33.138 = 128.30.33.128

不匹配!

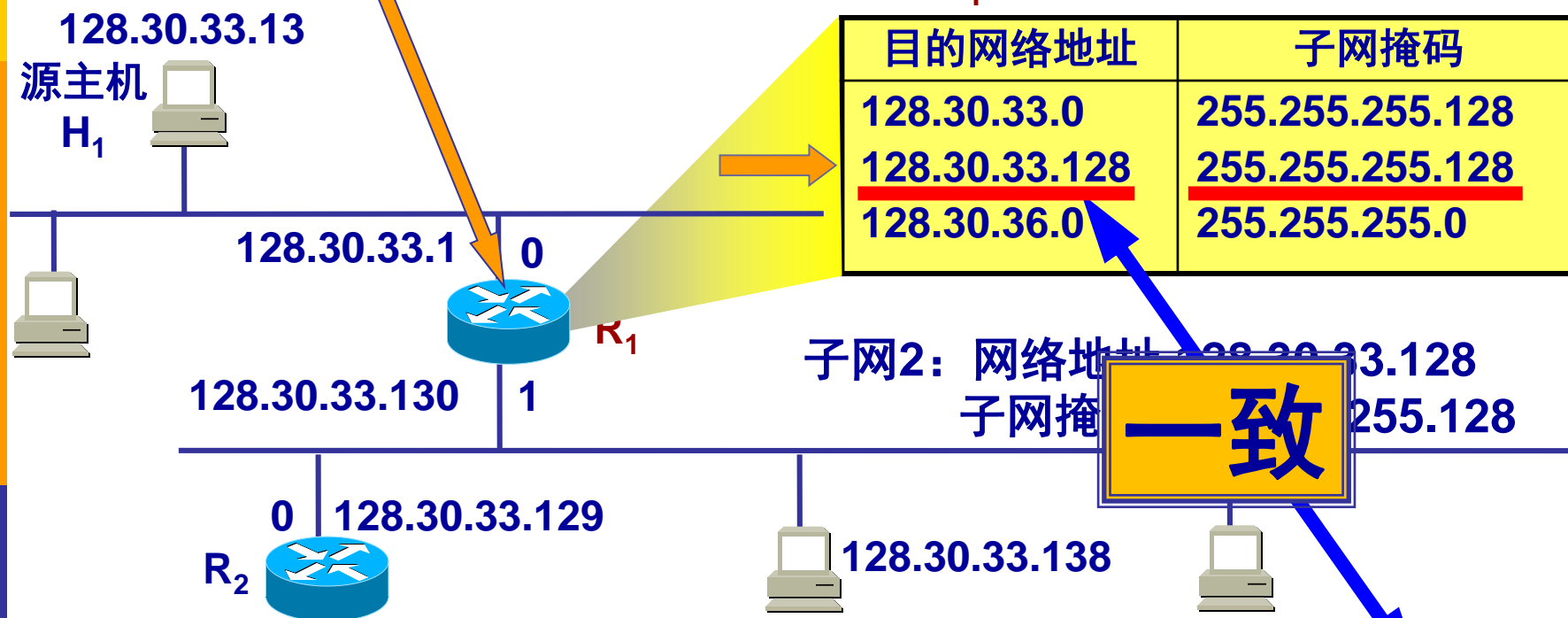
(因为128.30.33.128 与路由表中的 128.30.33.0 不一致)

路由器 R₁ 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

R₁ 收到的分组的目的 IP 地址: 128.30.33.138

R₁ 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
<u>128.30.33.128</u>	<u>255.255.255.128</u>	接口 1
128.30.36.0	255.255.255.0	R ₂



$$255.255.255.128 \text{ AND } 128.30.33.138 = \underline{128.30.33.128}$$

匹配!

这表明子网 2 就是收到的分组所要寻找的目的网络。

练习题



■ 1.一个路由器有两个IP地址，一个ip地址为192.168.11.25，另一个IP地址可能为（ ）

A.192.168.11.0 B.192.168.11.26

C.192.168.13.25 D.192.168.11.30

2. 下列地址中，属于单播地址的是（ ）

A.172.31.255.255 B.192.168.24.59

C.10.255.255.255 D.224.105.5.21