



# 网络层：网络互联

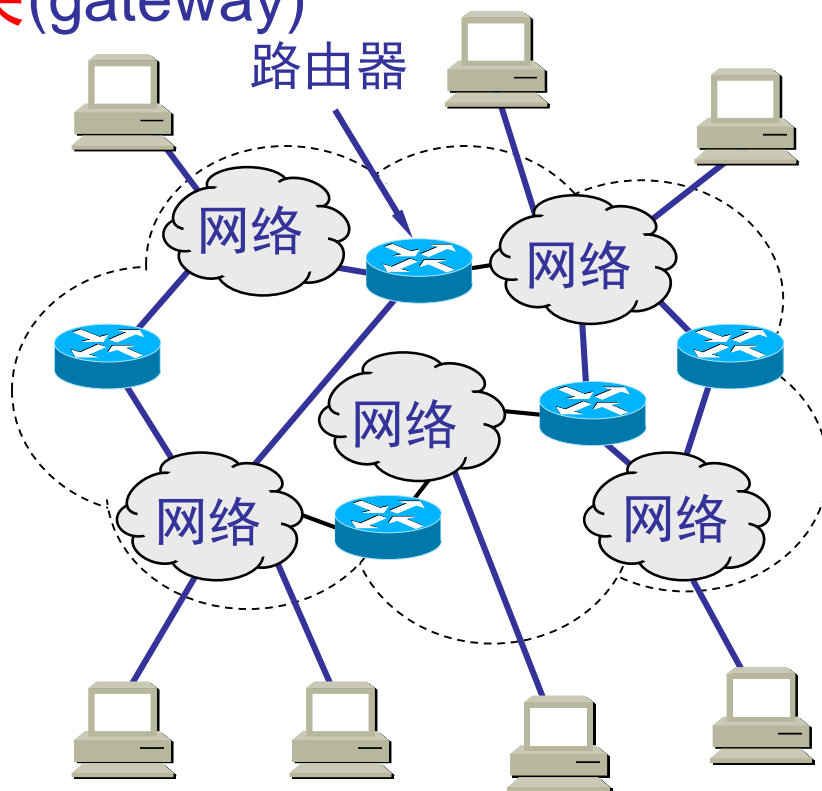
---

刘志敏

liuzm@pku.edu.cn

# Internet是一种互联网

- **互联网**：多个网络通过路由器互连
- 常见的互连设备
  - 物理层的**转发器**(repeater)；数据链路层的**网桥或桥接器**(bridge)；网络层的**路由器**(router)；网络层以上的**网关**(gateway)

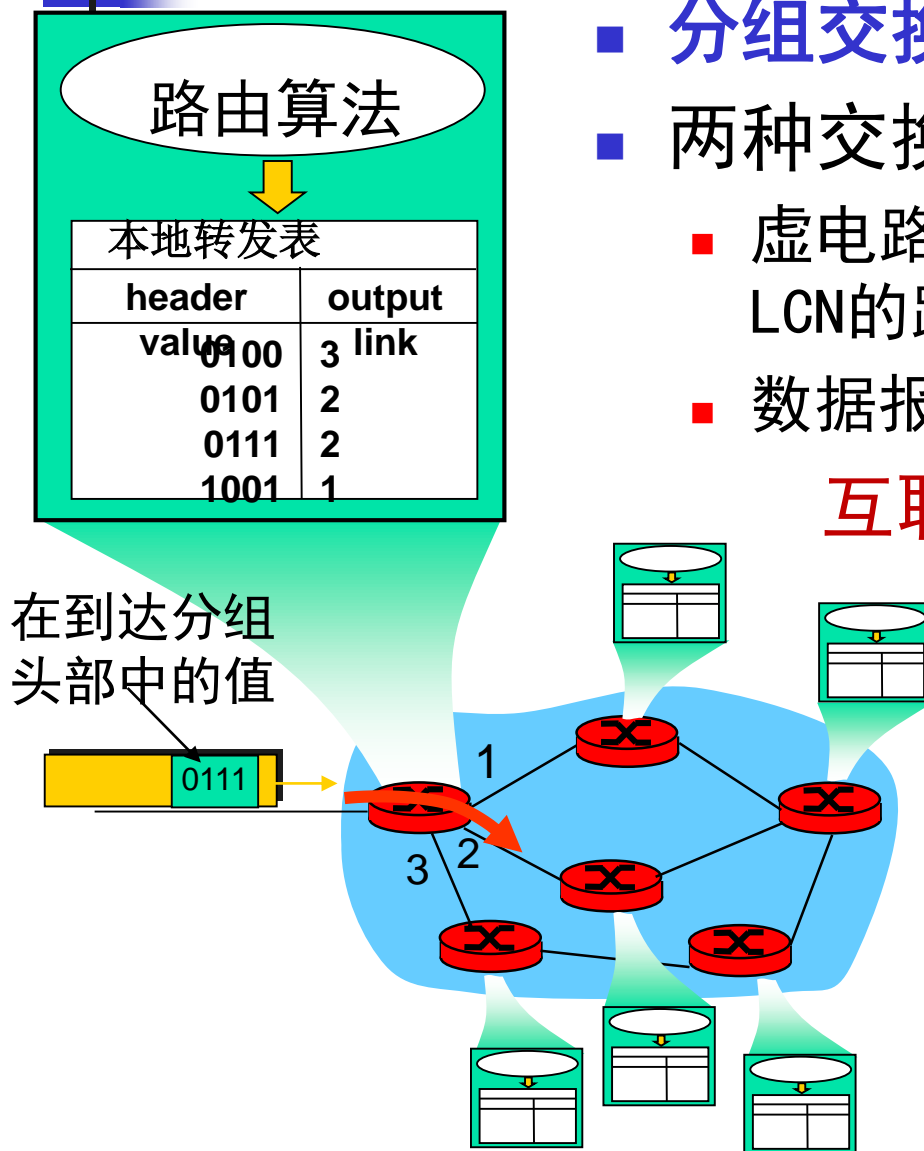


# 路由器的功能：路由与转发

- 分组交换：分组，存储转发
- 两种交换方式

- 虚电路：呼叫建立LCN，基于分组的LCN的路由
- 数据报：每个分组独立路由

互联网采用数据报方式

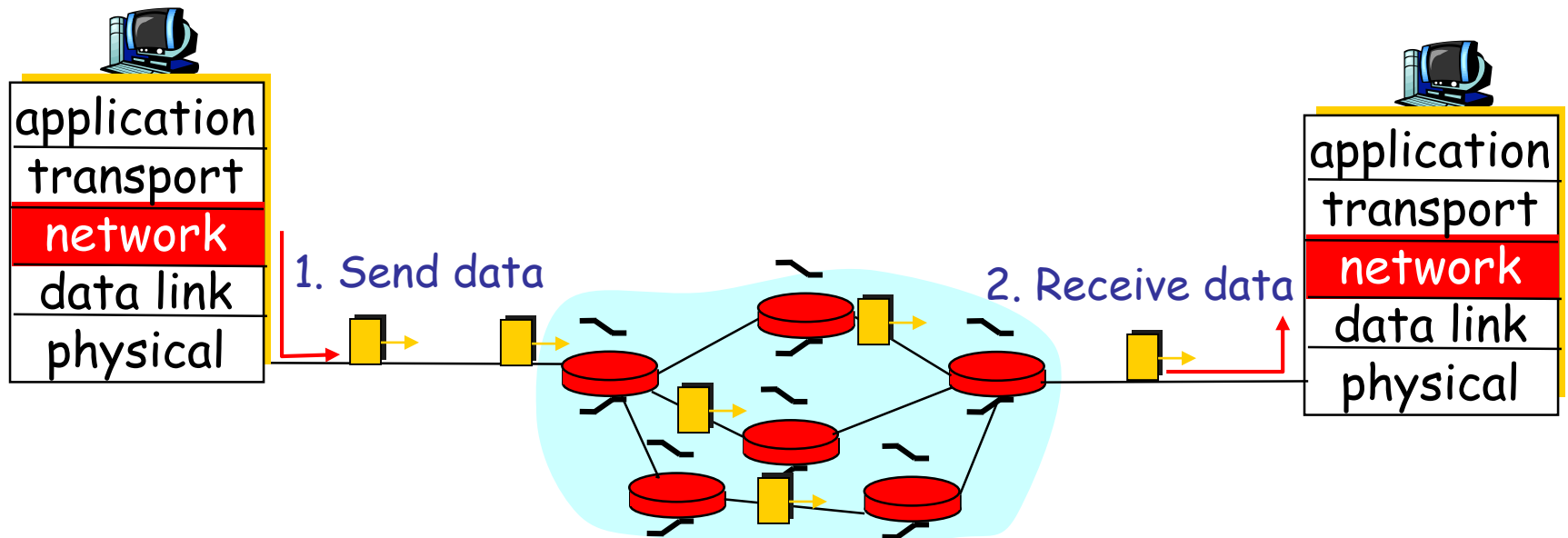


## 路由器负责分组的路由及转发

- 路由：执行路由算法，决定分组由源到目的的路径，形成路由表及转发表
- 转发：接收分组，查找转发表，将分组由输入端口移送到输出端口

# 数据报：互联网模式

- 网络层没有呼叫建立过程
- 路由器：不维护端到端的连接状态
  - 没有网络层“连接”这一概念
- 基于目的地址对分组进行路由选择
  - 一对源地址——目的主机的分组，可以选择不同路径





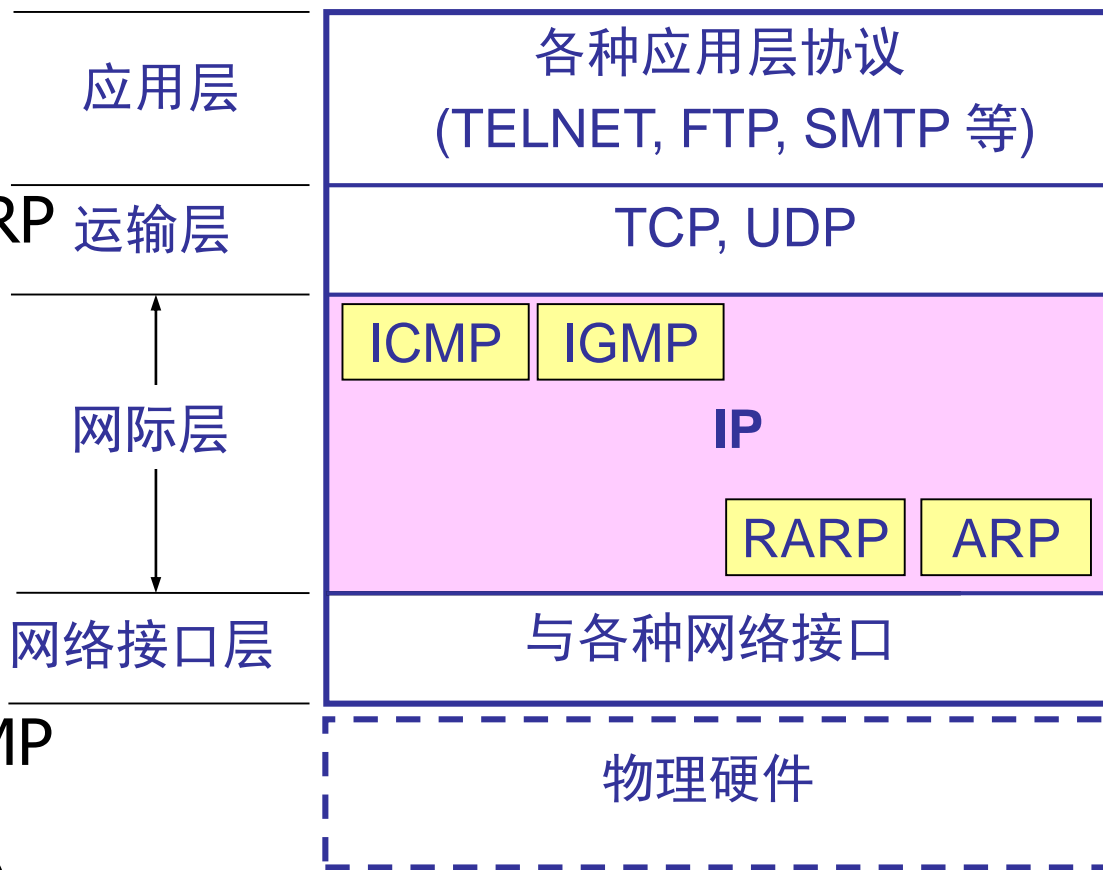
# 网络互联

---

- 需要解决的主要问题：
  - 地址分配
  - 分组传送
  - 路由与转发
  - 网络控制：超时控制、差错恢复、状态报告、拥塞检测与控制
- 互联网的核心协议是IP

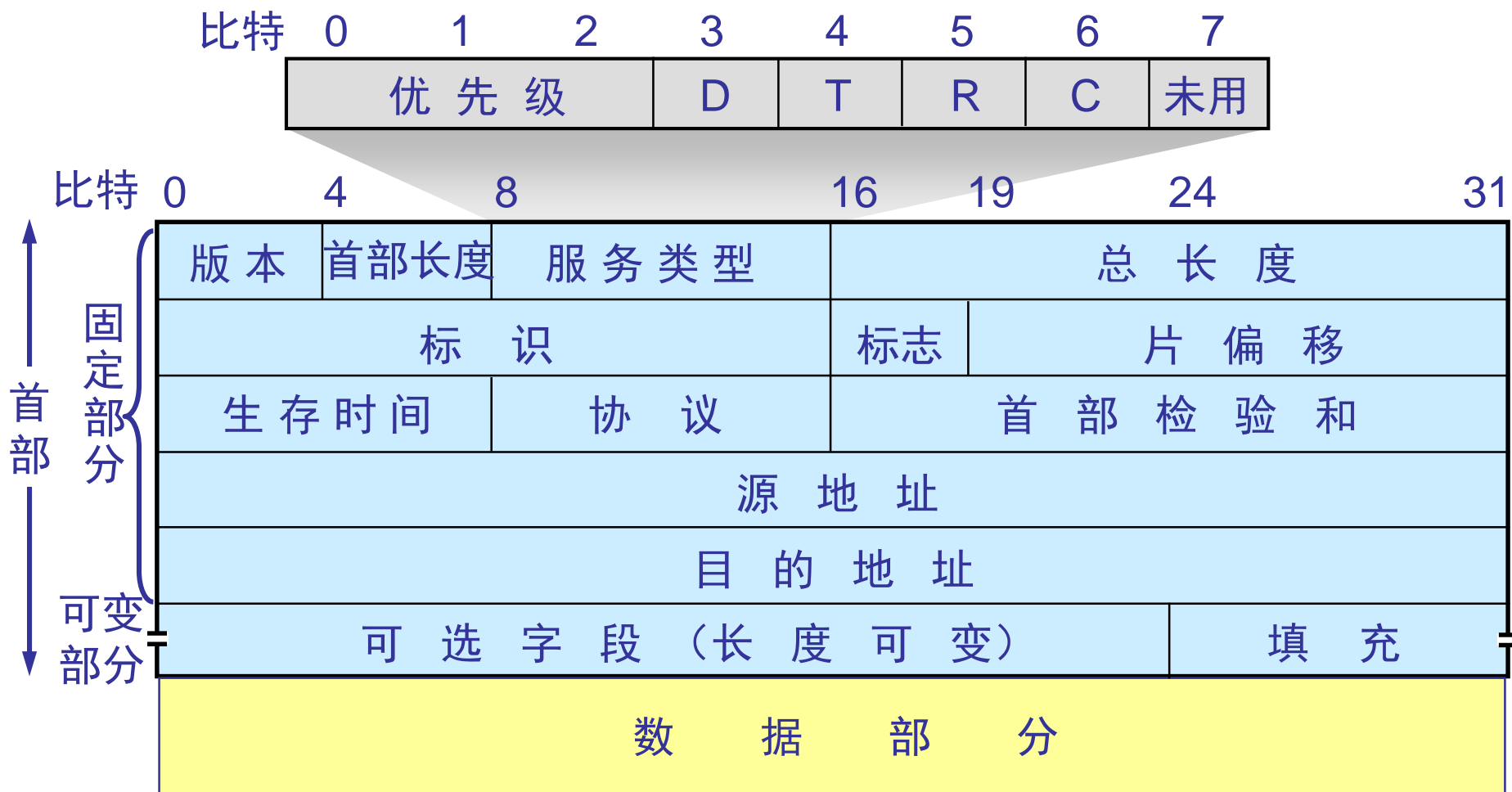
# 互联网的网络层协议

- IP
- 地址解析协议 ARP (Address Resolution Protocol)
- 反向地址解析协议 RARP (Reverse Address Resolution Protocol)
- 互联网报文控制协议 ICMP (Internet Control Message Protocol)
- 互联网组管理协议 IGMP (Internet Group Management Protocol)



# IP 分组格式

- 由首部和数据两部分组成，首部占 20 字节





# IP 地址

- 地址及标识：身份证、固定电话号码、学号等
  - 按一定的规则编码，编号唯一
- IP地址是连接在互联网上的主机（或路由器）的惟一标识；IPV4占32位，地址数为 $2^{32}$ ，IPV6占128位，地址数为 $2^{128}$ ；介绍以IPV4为主
- IP地址的编址方法
  - 分类IP地址 是最基本编址方法
  - 子网划分 是对最基本编址方法的改进
  - 构成超网 是较新的无分类编址方法
- IP地址管理：互联网域名和地址分配机构 ICANN





# 分类 IP 地址

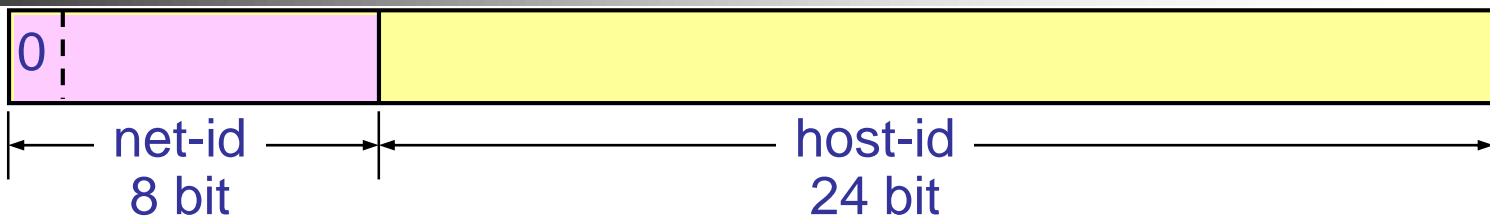
- 将IP地址分为A、B、C、D、E类，每类地址都由两个固定长度的字段组成，网络号 net-id 标志主机所连接到的网络，主机号 host-id 标志该主机。
- 两级的 IP 地址记为：

IP 地址 ::= { <网络号>, <主机号> }

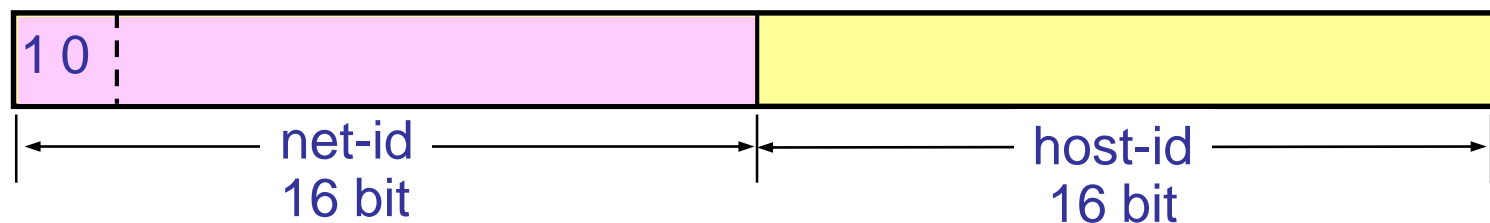
::= 代表 “定义为”

# IP 地址中的网络字段和主机字段

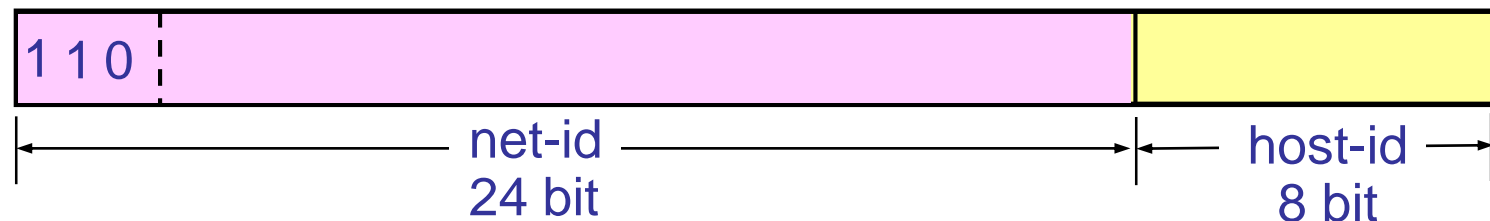
A 类地址



B 类地址



C 类地址



D 类地址



E 类地址



# 常用的三类 IP 地址

每隔 8 bit 插入一个空格  
以提高可读性

10000000 00001011 00000011 00011111

将每 8 bit 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

128.11.3.31

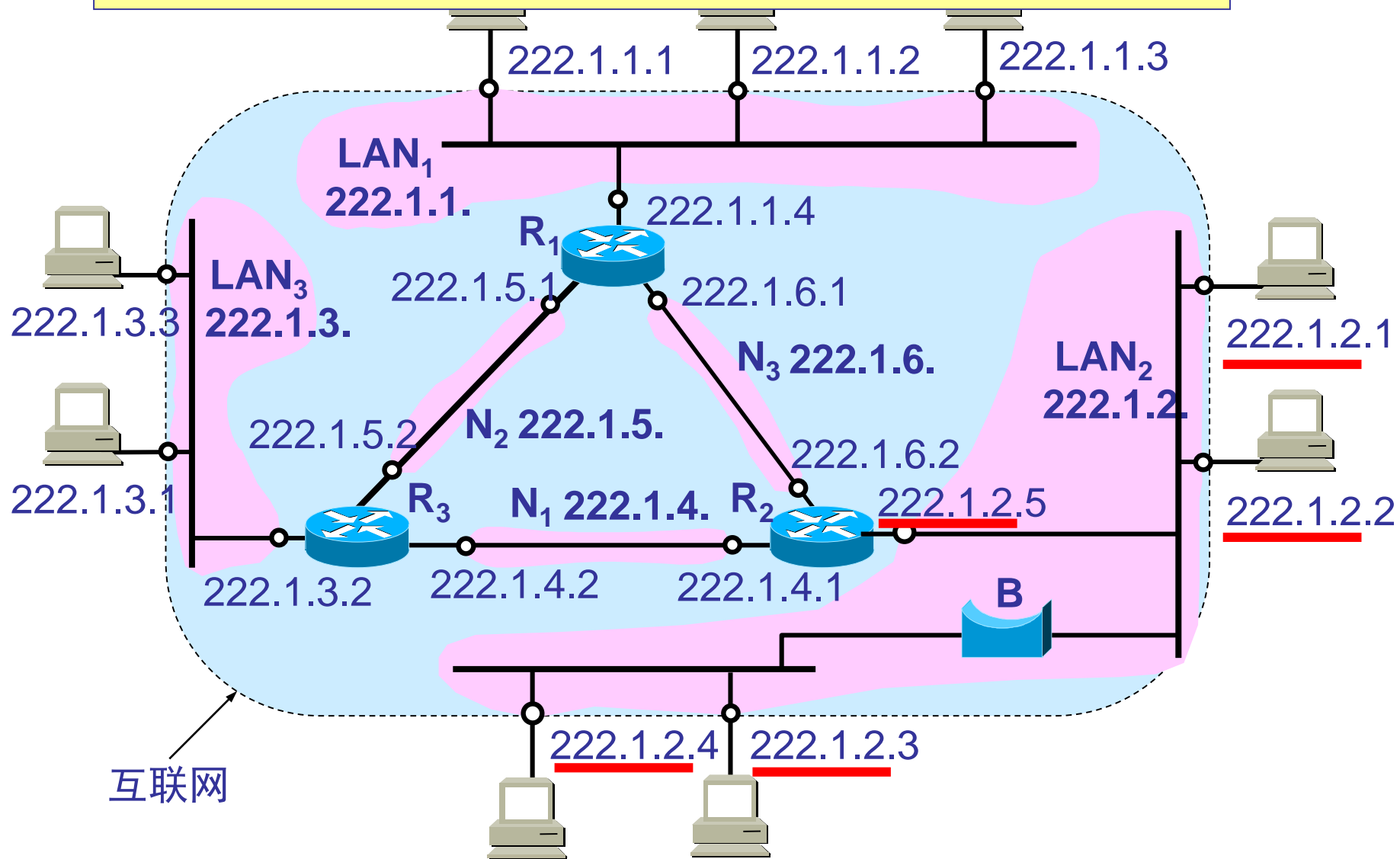
网络类别	最大网络数	第一个可用的网络号	最后一个可用的网络号	每个网络最大的主机数
A	126 ( $2^7 - 2$ )	1	126	16,777,214
B	16,384( $2^{14}-1$ )	128.1	191.255	65,534
C	2,097,152( $2^{21}-1$ )	192.0.1	223.255.255	254

# IP 地址的特点

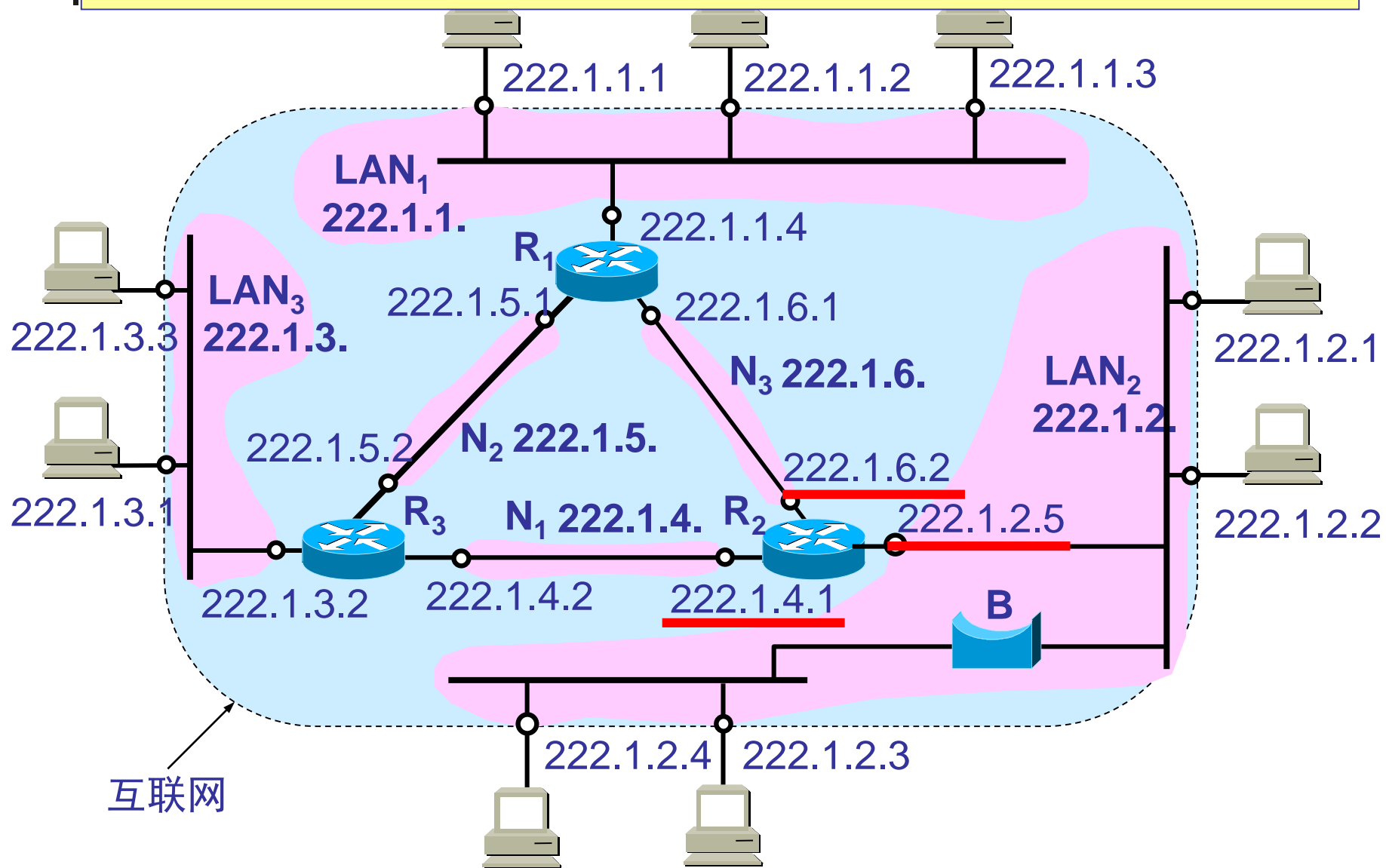
- IP地址是一种分级结构，只分配网络号，主机号由网络所属单位分配
- IP地址标志主机与链路的接口，路由器至少连接两个网络，有两个以上的IP地址
- 不使用的特殊IP

网络号	主机号	源地址	目的地址	含义
0	0	可以	不可以	本网络的本主机，用于DHCP
0	Host-id	可以	不可以	本网络的主机Host-d
全1	全1	不可以	可以	本网络上广播
Net-id	全1	不可以	可以	对Net-id的所有主机广播
127	非全0或非全1	可以	可以	用作本地软件环回测试

在同一网络上的主机或路由器，其  
IP 地址的网络号必须相同



路由器具有两个或两个以上的 IP 地址，  
每个接口的 IP 地址，其网络号不同

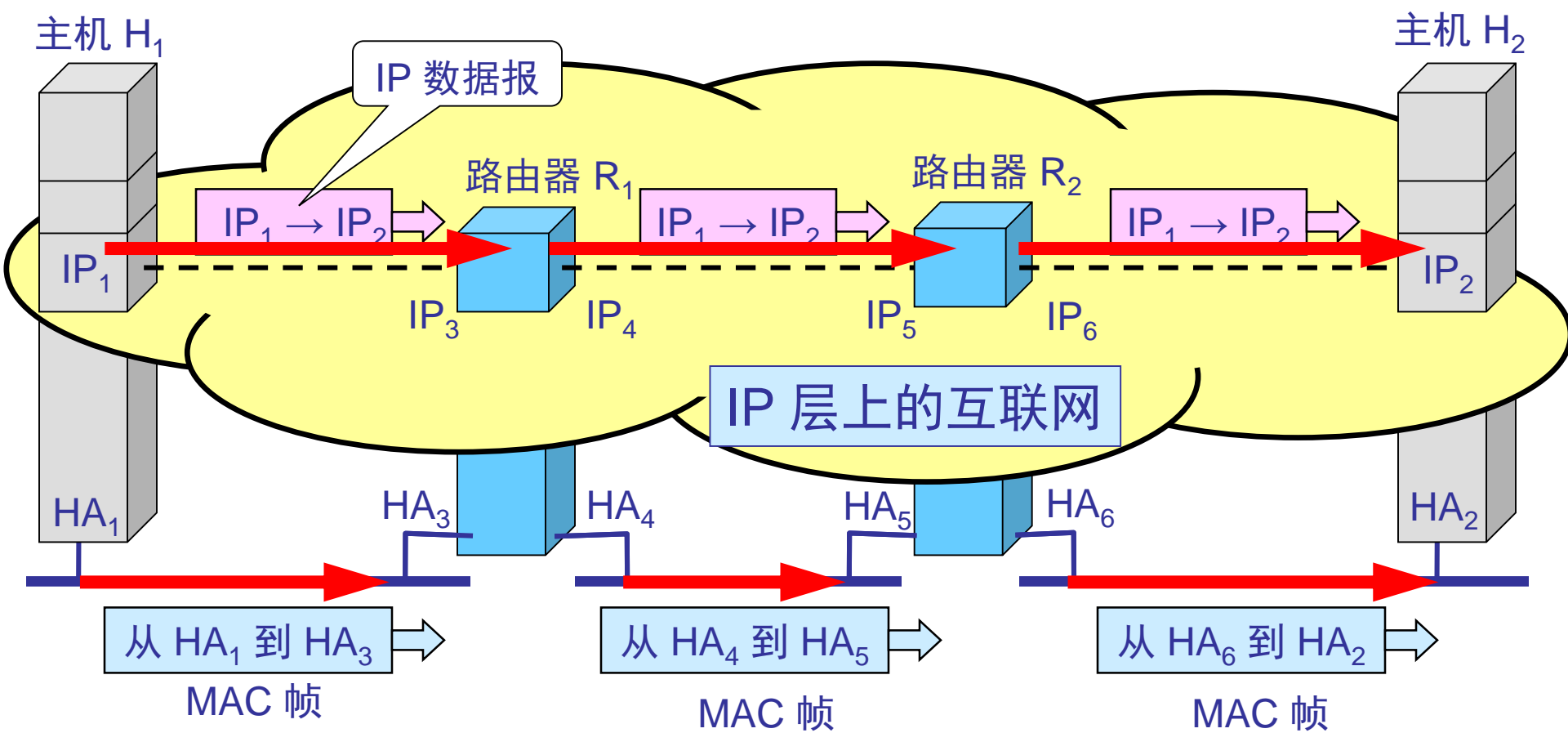
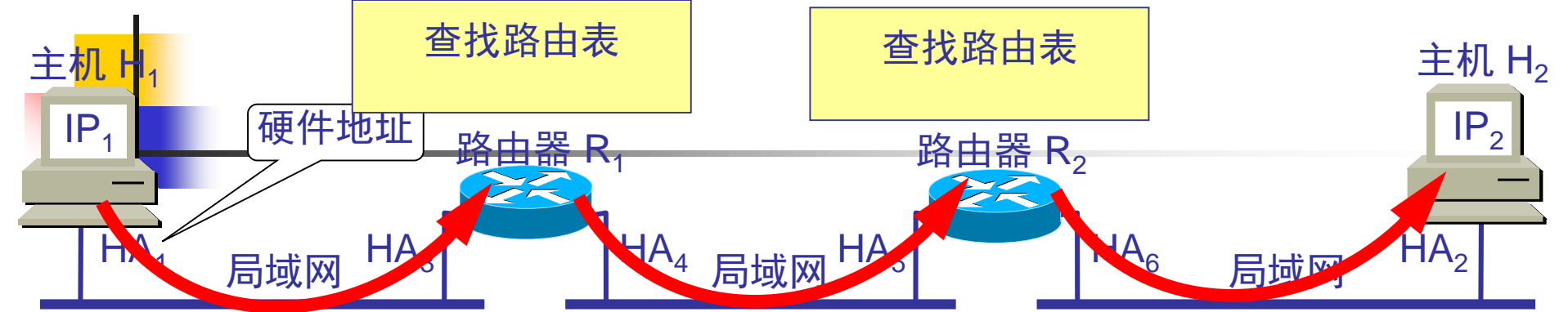




# 网络互联

---

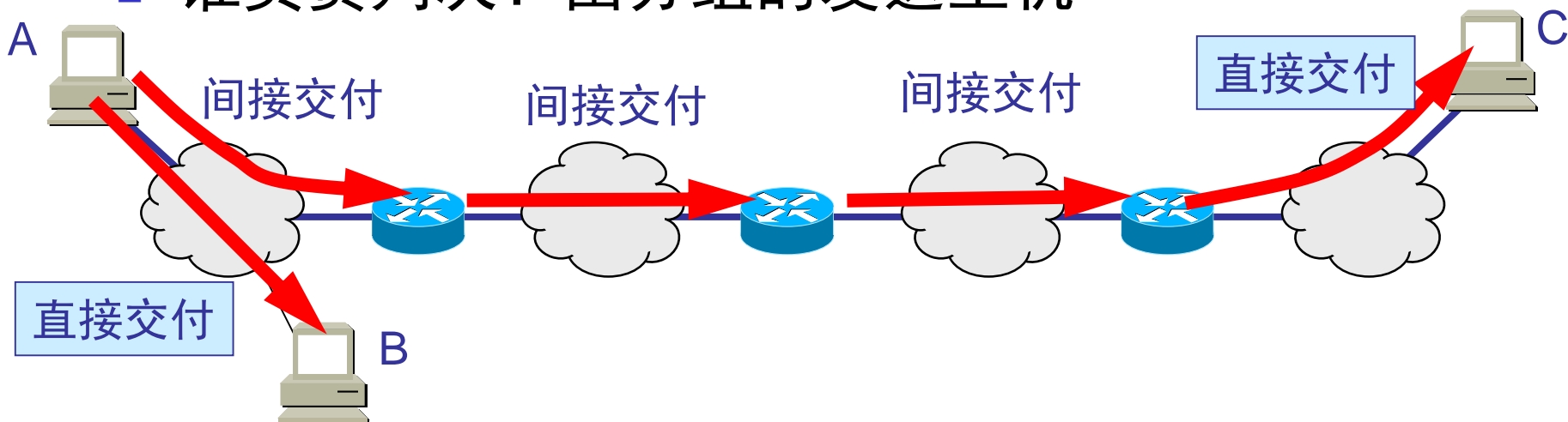
- 需要解决的主要问题：
  - 地址分配
  - 分组传送
  - 路由与转发
  - 网络控制：超时控制、差错恢复、状态报告、拥塞检测与控制
- 互联网的核心协议是IP



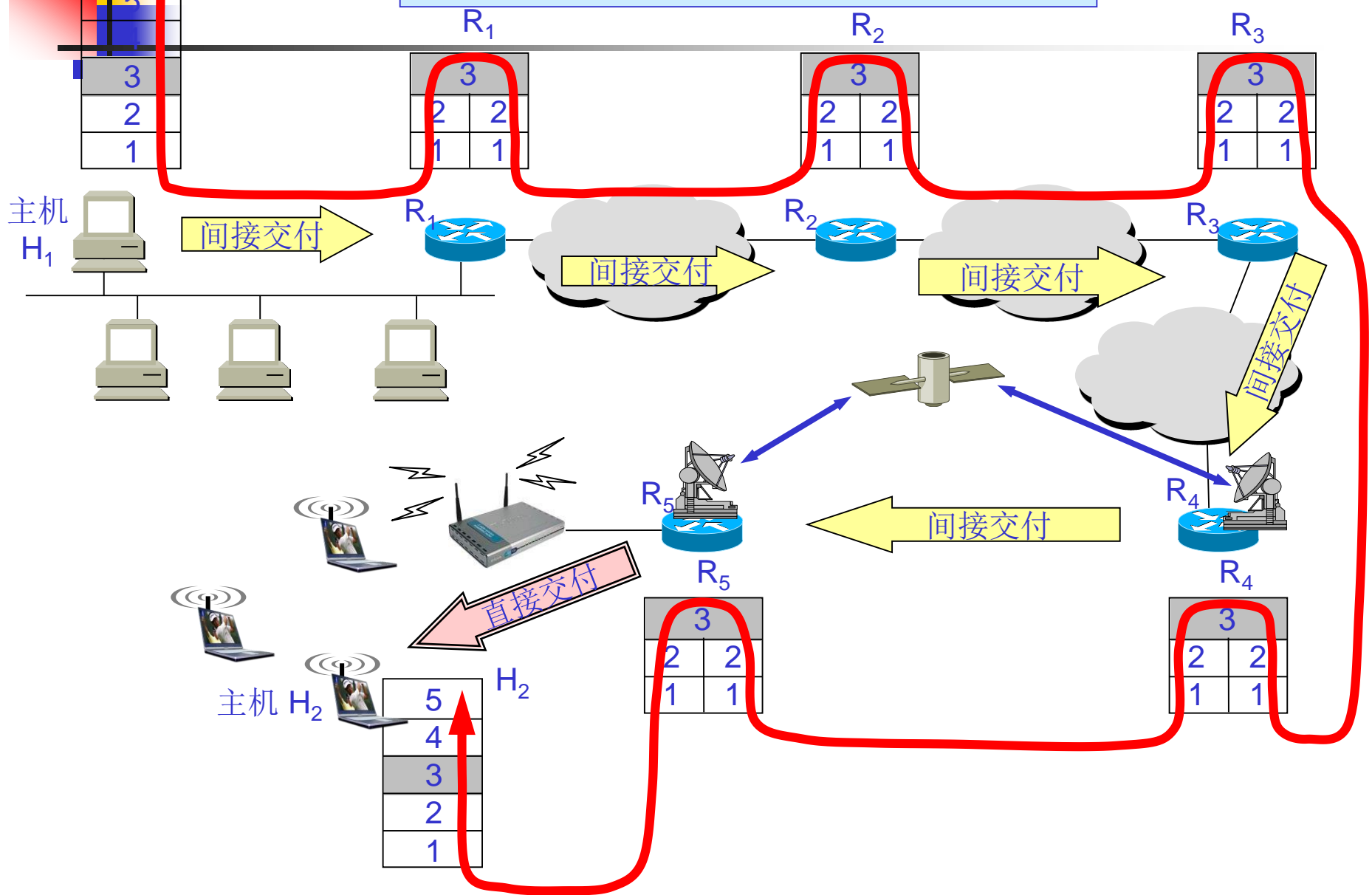


# 路由器的作用

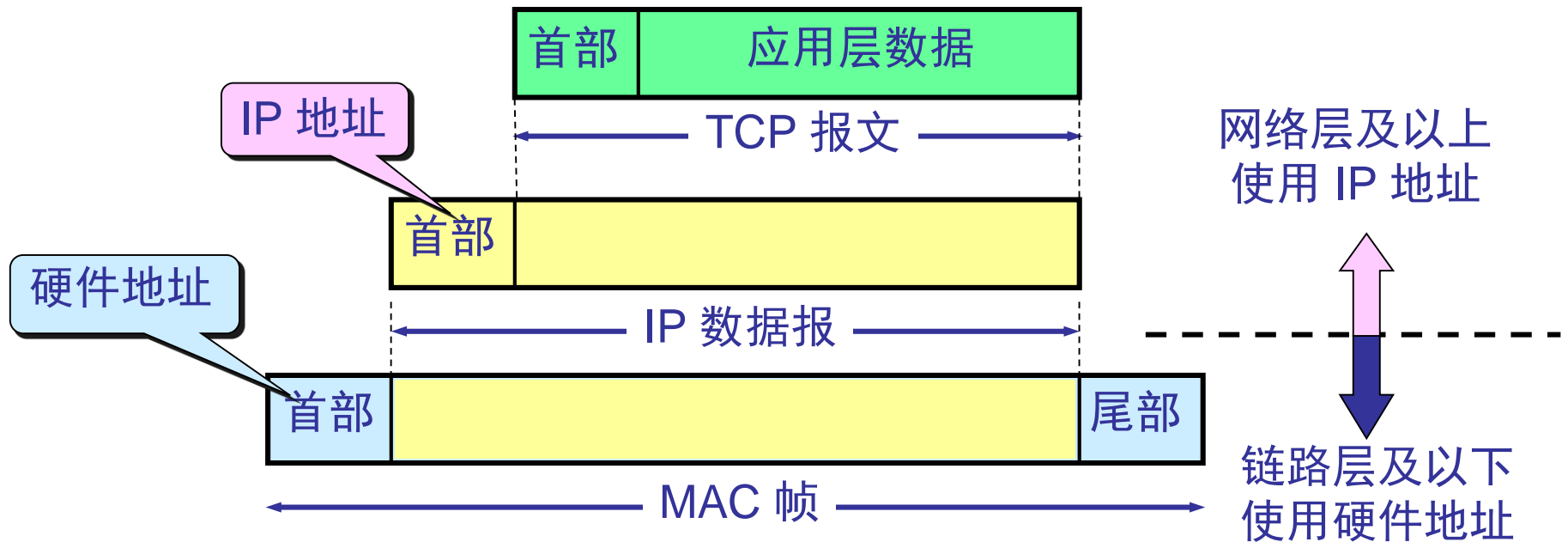
- 当主机A要向主机B发送分组时，先检查主机B是否与其在同一网络上。如果是，就**直接交付**给主机B；否则，将**间接交付**，即将分组发送给本网络上某个路由器，由路由器负责转发。
- 如何判决主机A与B是否在同一网络上？
  - 判断主机A与主机B的网络地址是否相同
- 谁负责判决？由分组的发送主机



# 分组在互联网中的传送



# IP 地址与硬件地址



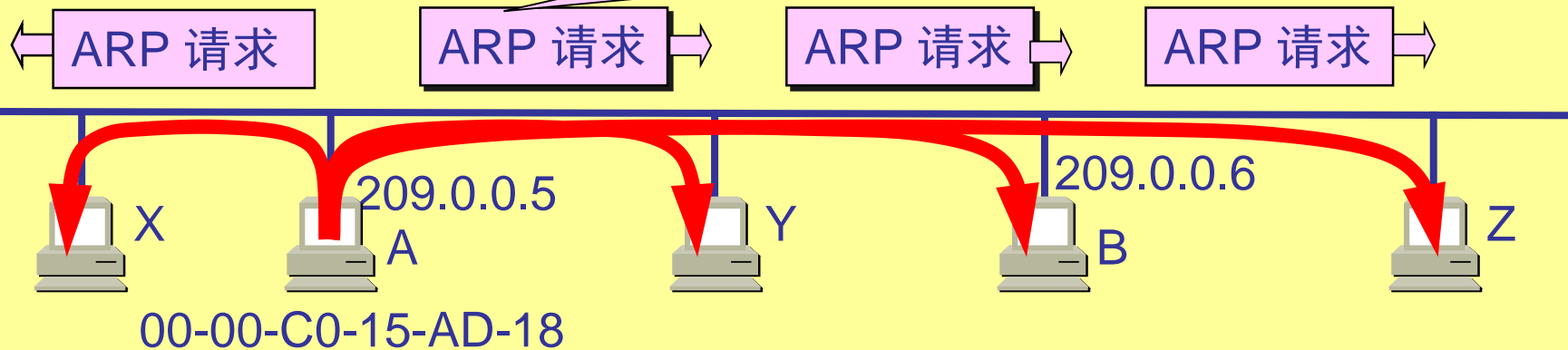


# ARP

- 在网络层上传输IP分组，用IP地址；而在链路上传送数据帧，必须使用硬件地址。
- 需要建立网络地址与硬件地址的映射关系
- 地址解析协议 ARP：解决在同一局域网（子网）上主机 IP 地址与网卡硬件地址，即MAC地址之间的映射。根据IP地址找其对应的MAC。
- 反向地址解析协议 RARP：已知主机的硬件地址，而要找到其 IP 地址。这种主机往往是无盘工作站。

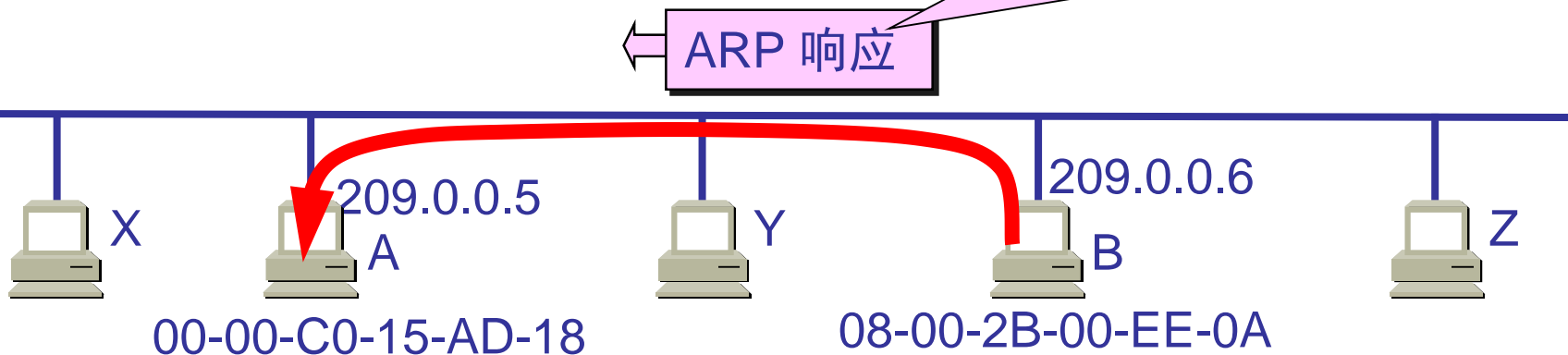
## 主机 A 广播 ARP 请求分组

我是 209.0.0.5，硬件地址是 00-00-C0-15-AD-18  
想知道 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

- ARP高速缓存：主机存储IP地址到硬件地址的映射表，减少发送ARP请求的机会。
- ARP是解决同一个网络上的主机或路由器的IP地址和硬件地址的映射问题。若主机位于不同的网络，则分组发送给其路由器，由路由器转发。
- IP地址到硬件地址的解析是自动进行的
- 为何不直接用硬件地址通信？
  - 网络种类多，硬件种类也多，转换很复杂
  - 用IP地址，用户可设置，使用方便



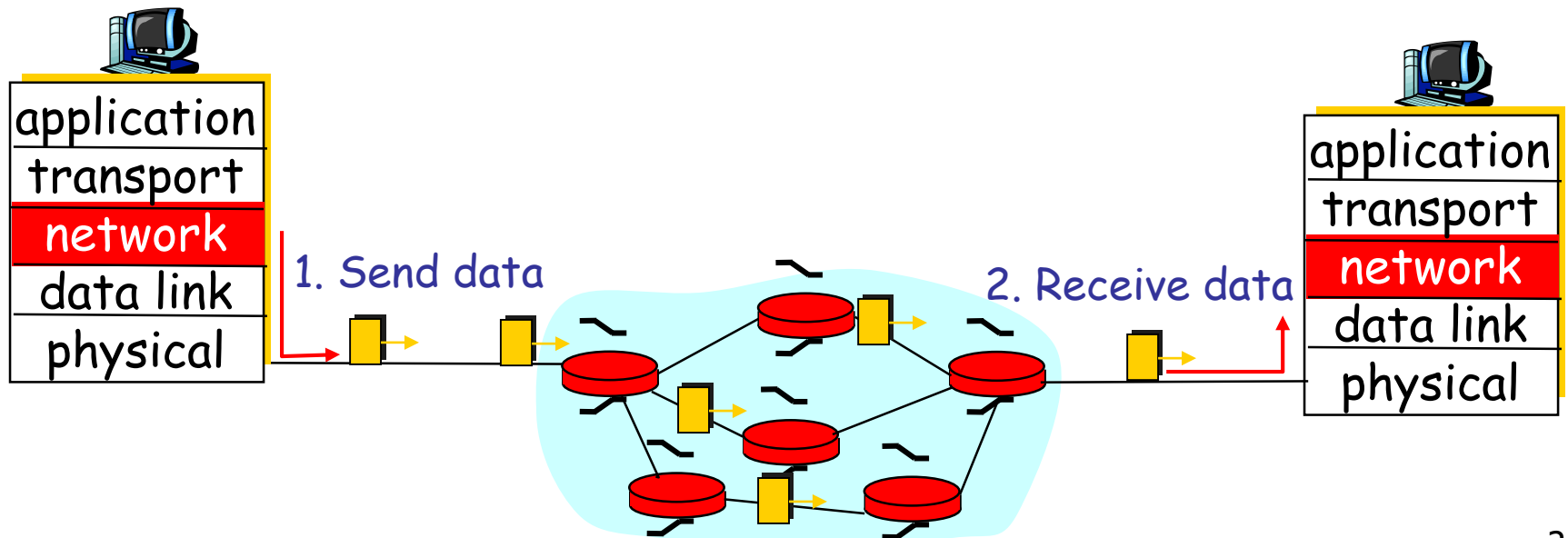
# 网络互联

---

- 需要解决的主要问题：
  - 地址分配
  - 分组传送
  - 路由与转发
  - 网络控制：超时控制、差错恢复、状态报告、拥塞检测与控制
- 互联网的核心协议是IP

# 数据报：互联网模式

- 网络层没有呼叫建立过程
- 路由器：不维护端到端的连接状态
  - 没有网络层“连接”这一概念
- 基于目的地址对分组进行路由选择
  - 一对源地址——目的主机的分组，可以选择不同路径





# 转发表（存储在路由器上）

表项可能很多

<u>Destination Address Range</u>	<u>Link Interface</u>
11001000 00010111 00010000 00000000 through 11001000 00010111 00010111 11111111	0
11001000 00010111 00011000 00000000 through 11001000 00010111 00011000 11111111	1
11001000 00010111 00011001 00000000 through 11001000 00010111 00011111 11111111	2
otherwise	3

# 更小的转发表

<u>Prefix Match</u>	<u>Link Interface</u>
11001000 00010111 00010	0
11001000 00010111 00011000	1
11001000 00010111 00011	2
otherwise	3

基于目的地址的匹配前缀（而非目的地址）查找转发端口；大大减少路由表项，加快查表速度

## Examples

DA: 11001000 00010111 00010110 10100001 Which interface?

DA: 11001000 00010111 00011000 10101010 Which interface?

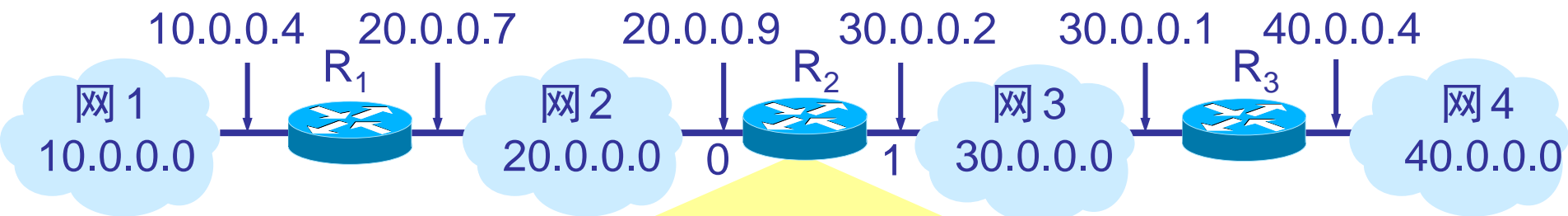
最长匹配原则：选择更长匹配前缀项



# 路由与转发

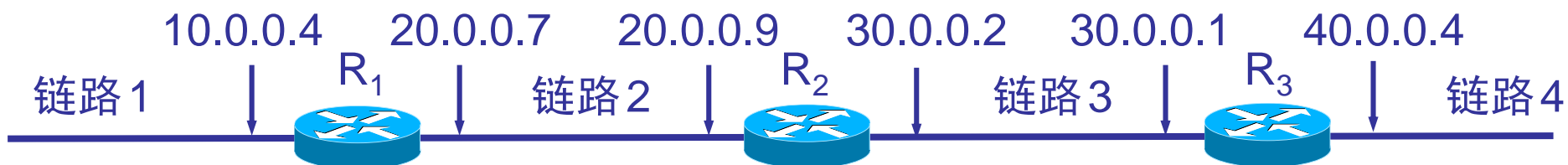
- 转发表如何产生的？
  - 路由器执行路由算法及路由协议
- 距离矢量路由，如RIP
  - 每个路由器维护一张距离矢量路由表
  - 在邻居路由器之间交换表，路由表得到更新
  - 按照距离矢量路由算法，计算最短路径路由
- 链路状态路由，如OSPF
  - 使用扩散法向所有路由器发送信息
  - 每个路由器获得完整的拓扑结构
  - 按照最短路径算法计算最短路径

# 路由器根据目的网络地址选择下一跳



路由器 R<sub>2</sub> 的路由表

目的网络	下一跳
20.0.0.0	直接交付, 接口 0
30.0.0.0	直接交付, 接口 1
10.0.0.0	20.0.0.7
40.0.0.0	30.0.0.1





# 路由器的分组转发算法

- (1) 提取分组的目的IP地址  $D$ , 得到目的网络地址  $N$
- (2) 若网络  $N$  与此路由器直连, 则直接交付; 否则, 间接交付, 执行(3)
- (3) 若在路由表中有到目的地址  $D$  的特定路由, 则将分组发送到路由表所指明的下一跳; 否则, 执行(4)
- (4) 若路由表中有到达网络  $N$  的路由, 则将分组发送到路由表所指明的下一跳; 否则, 执行(5)
- (5) 若路由表中有一个默认路由, 则将分组发送到路由表所指明的默认路由器; 否则, 执行(6)
- (6) 采用ICMP报告转发分组出错



# 再议IP地址

- IP地址是一种分级结构，只分配网络号，主机号则由网络所属单位分配
- 路由器仅根据目的主机的网络号（而非目的地址）转发分组，使路由表项数大大减少
- 降低路由表项数，提升了路由器的查表速度
- IP地址的编址方法
  - 分类IP地址 是最基本编址方法
  - 子网划分 是对最基本编址方法的改进
  - 构成超网 是较新的无分类编址方法，得到推广应用。



# 划分子网

- A类B类网络，主机数过多；存在ARP广播风暴问题
- 在IP地址中增加“子网号”，称为划分子网
- 将主机地址的若干比特作为子网号
- 划分子网是将IP地址的本地部分再划分
  - IP地址 ::= {<网络号>, <子网号>, <主机号>}
- 子网掩码：为1的部分表示子网地址，为0的部分表示主机地址
  - 例如：162.105.75.1 255.255.255.0



# 使用子网掩码的分组转发过程

- (1) 提取接收分组的首部目的 IP 地址  $D$
- (2) 用各网络的子网掩码与  $D$  相“与”，看是否与相应的网络地址匹配；若匹配，则直接交付；否则，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；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”，若结果与该行的目的网络地址匹配，则将分组传送给下一跳；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

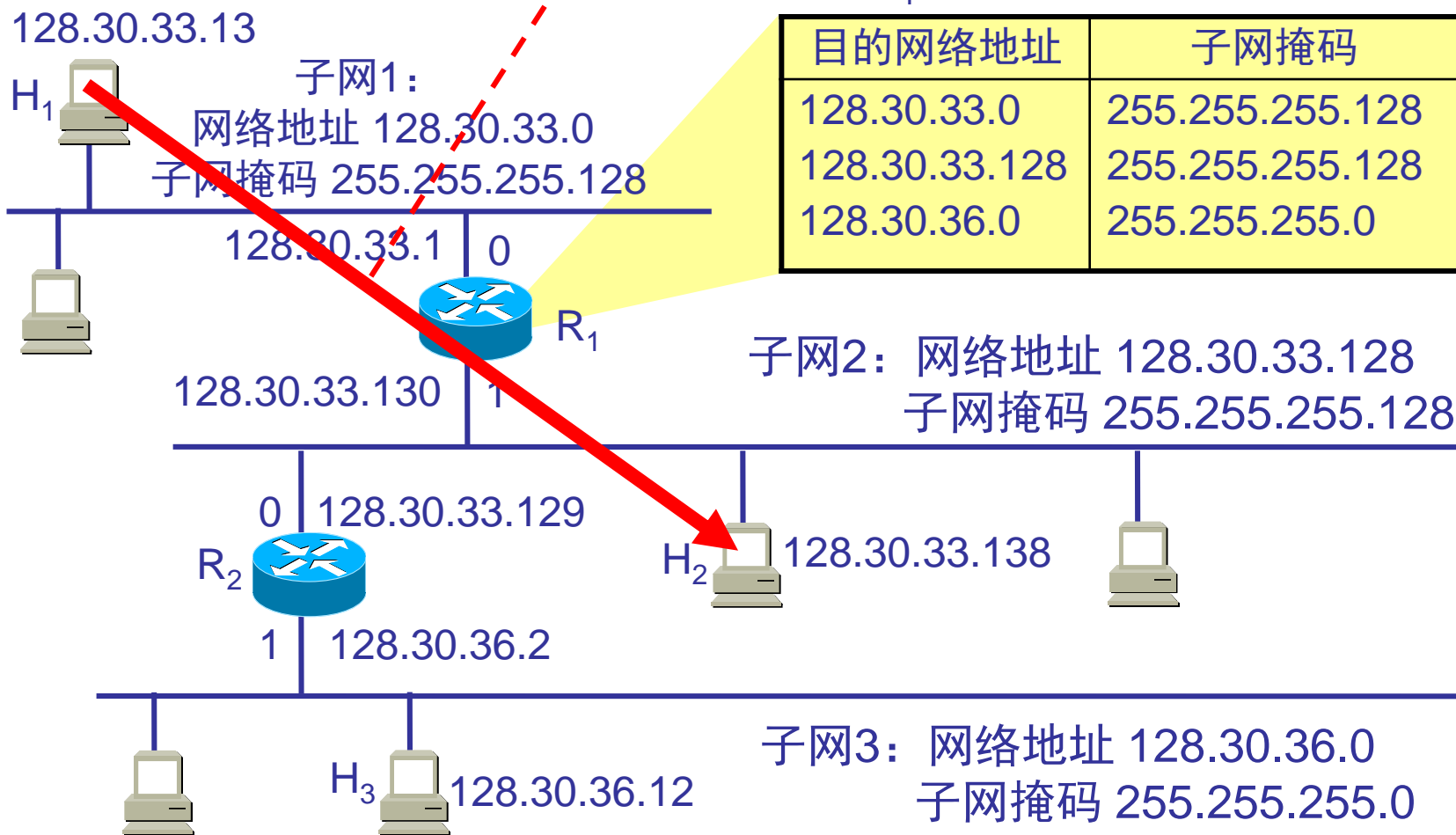


# 划分子网后分组的转发举例

发送分组的目的 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	$R_2$



# 无分类编址CIDR

- CIDR(Classless Inter-Domain Routing): 消除了A、B、C类地址及划分子网的概念
- CIDR用网络前缀代替地址中的网络号和子网号

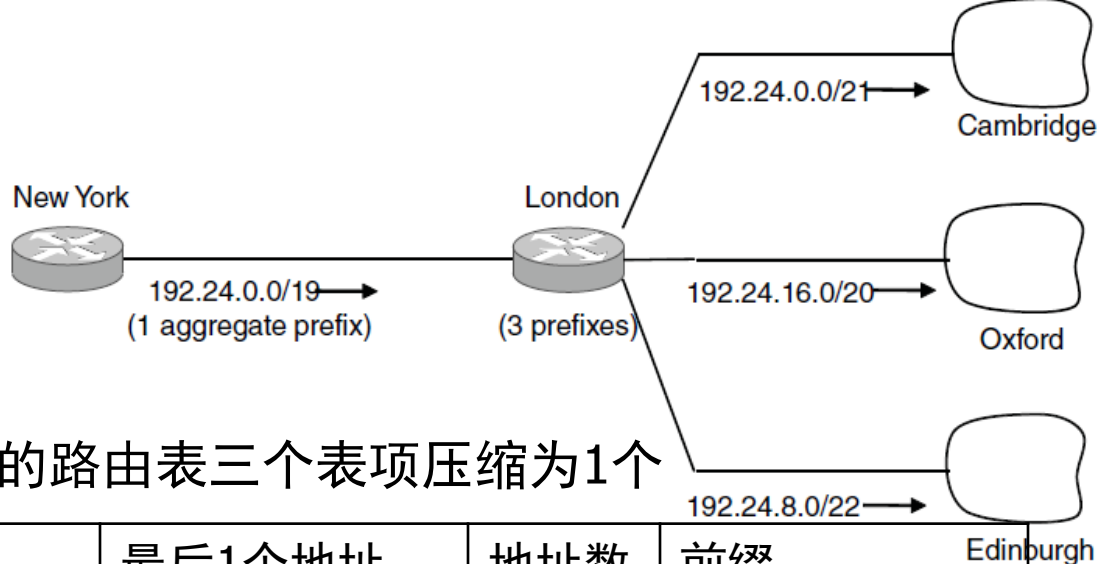
IP地址 ::= {<网络前缀>, <主机号>}

- CIDR 使用“斜线记法”，又称为CIDR记法
  - 例如162.105.75.1/16, 162.105.75.1/24
- CIDR 把网络前缀相同的连续的IP地址组成“CIDR地址块”，这种地址的聚合称为路由聚合，这样路由表中一项可表示多个传统的分类地址的路由；
- 路由聚合也称为构成超网

# CIDR举例

- 路由聚合，减少了路由表的表项

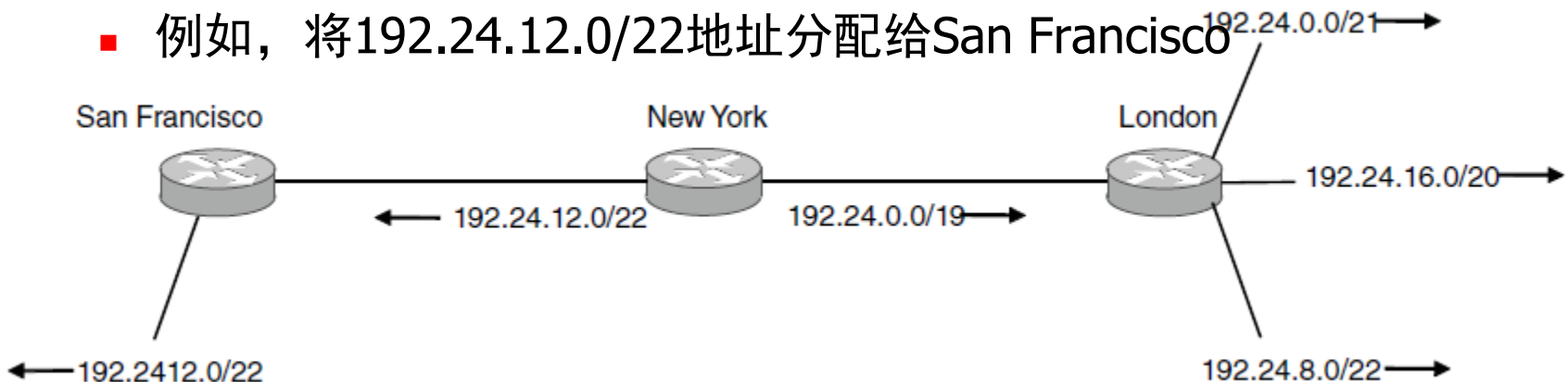
- 例如，将NewYork的路由表三个表项压缩为1个



大学	第1个地址	最后1个地址	地址数	前缀
Cambridge	192.24.0.0	192.24.7.255	2048	192.24.0.0/21
Edinburgh	192.24.8.0	192.24.11.255	1024	192.24.8.0/22
保留	192.24.12.0	192.24.15.255	1024	192.24.12.0/22
OXford	192.24.16.0	192.24.31.255	4096	192.24.16.0/20

- 最长匹配前缀，利于灵活调整地址

- 例如，将192.24.12.0/22地址分配给San Francisco



# 练习题

一个路由器的路由表如下：

地址/掩码	下一跳
135.46.56.0/22	接口0
135.46.60.0/22	接口1
192.53.40.0/23	路由器1
default	路由器2

若到达的分组，其目的地址有下述IP地址，问路由器如何处理（1）135.46.63.10（2）135.46.57.14

（3）135.46.52.2（4）192.53.40.7（5）192.53.56.7

■ 解：56=0x38, 60=0x3C, 63=0x3F 57=0x39 52=0x34

（1）接口1                      （2）接口0      （3）路由器2

（4）路由器1                  （5）路由器2



# 小结

---

- 地址分配：
  - IP地址，三种编址方式；
  - 如何分配IP地址？
  - IP地址数量不够如何解决？
- 分组传送
  - ARP：IP 地址到MAC的映射
  - 各段链路的帧长度不同，如何确定IP分组长度？
- 路由与转发：
  - RIP及距离矢量路由算法
  - 其他的路由算法及路由协议
- 网络控制：超时控制、差错恢复、状态报告、拥塞检测与控制？