

第六章 网络层

合肥工业大学
计算机与信息学院

□ TCP/IP产生的背景

1972-1980

ARPnet、ALOHAnet、SNA（IBM）、Telenet（商用分组交换网）

各种局域网（以太网、令牌环网、令牌总线网....）

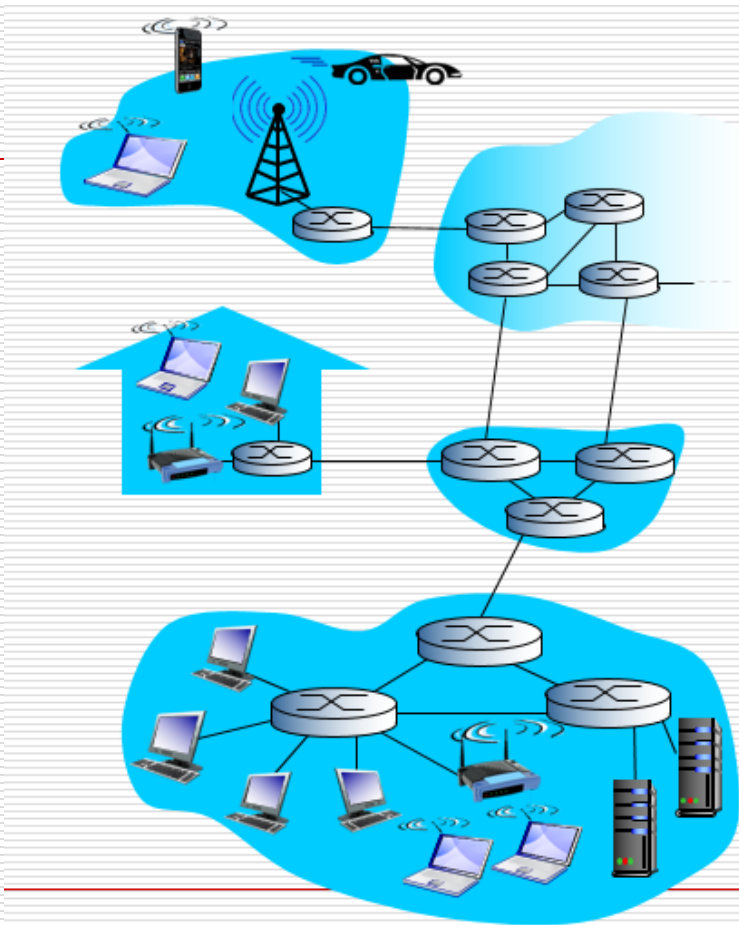
网络互连

□ 异构网络互连：网络的网络

➤ TCP

➤ UDP

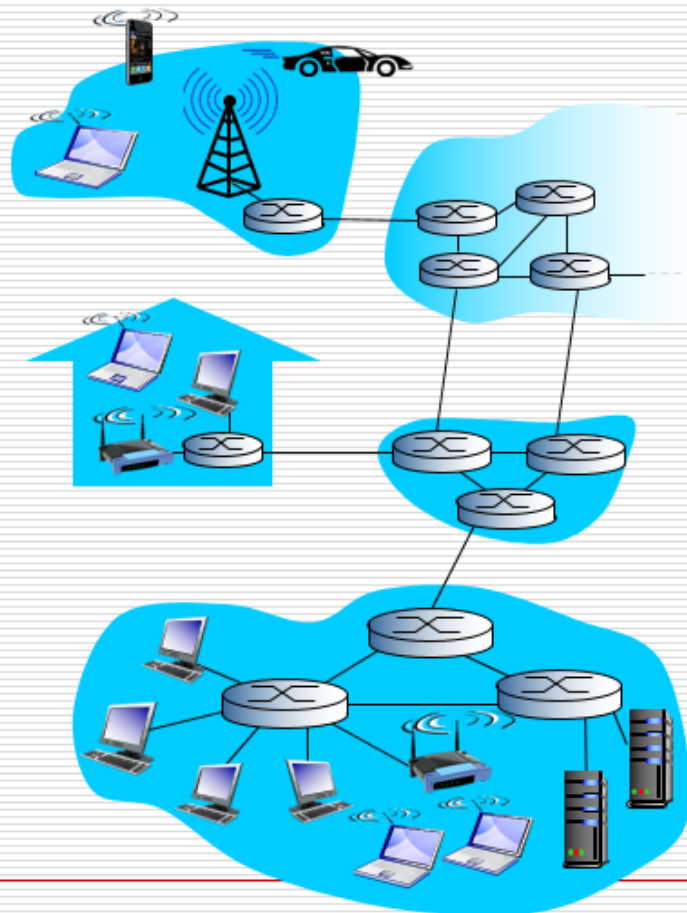
➤ IP



网络互连

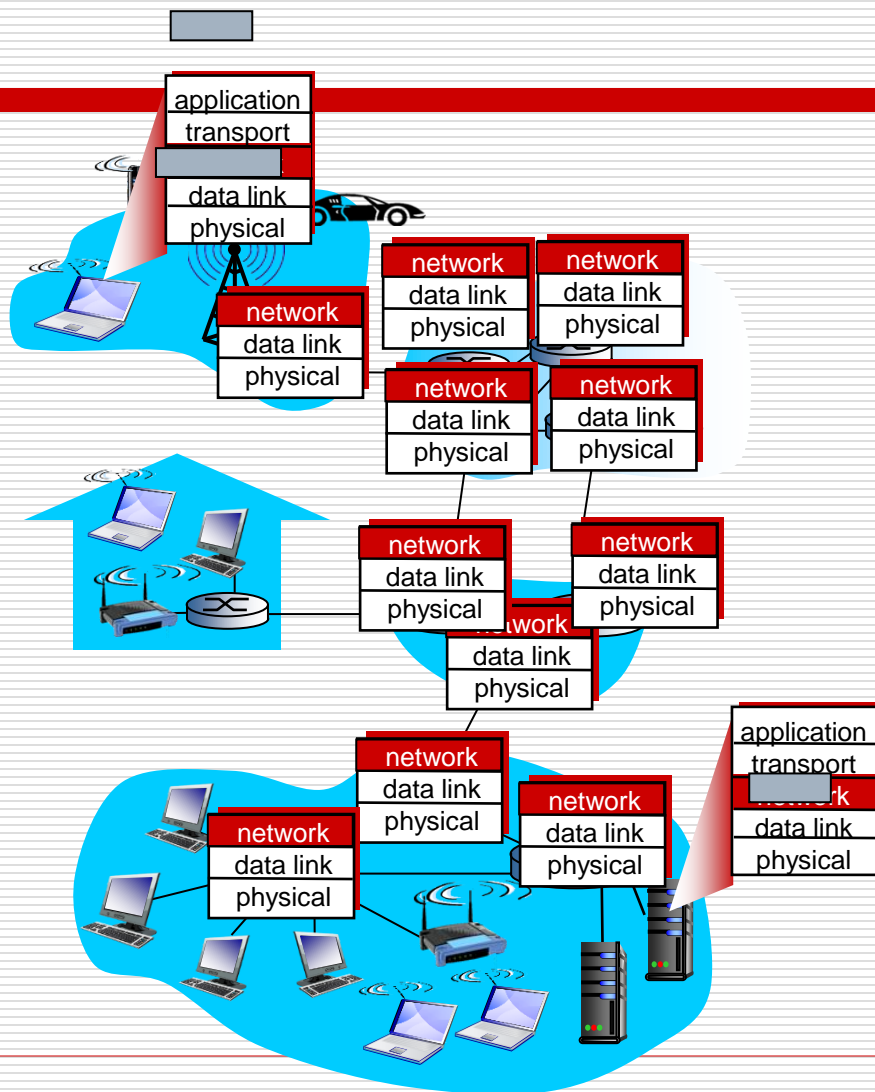
□ 1983.1.1: TCP/IP作为ARPAnet新的标准主机协议，正式部署，替代了NCP协议

□ TCP功能扩展，DNS出现



网络层服务

- 核心：IP协议
 - 实现主机到主机的通信
 - 网络层互连设备：路由器
-

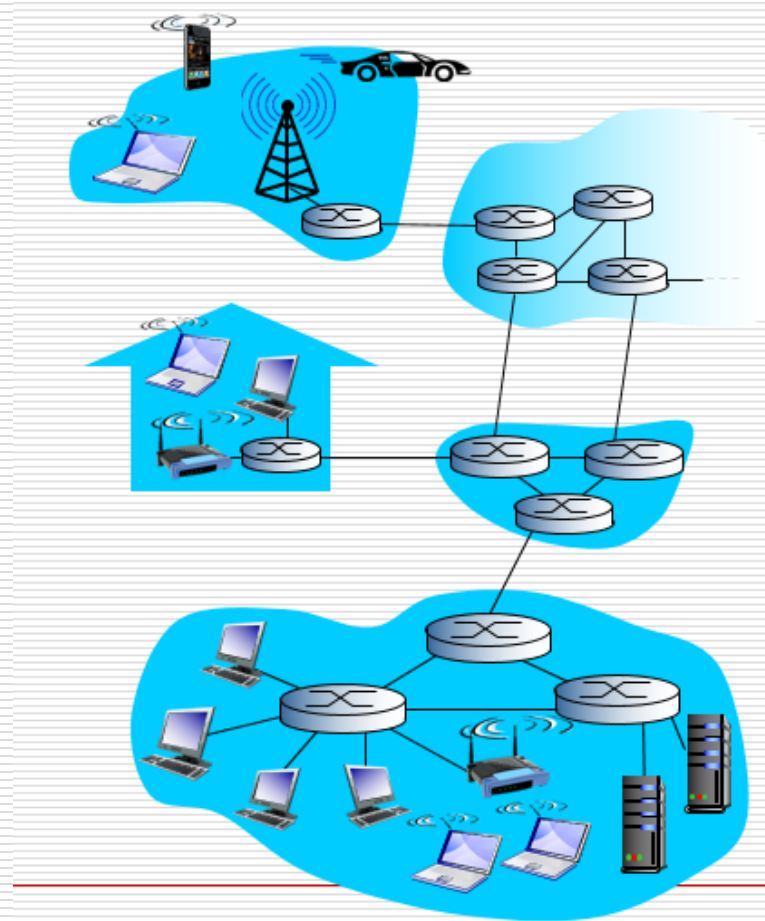


- 在发送主机和接收主机对之间传送段（segment）
- 在发送端将段封装到数据报中
- 在接收端，将段上交给传输层实体
- 网络层协议存在于每一个主机和路由器
- 路由器检查每一个经过它的IP数据报的头部

网络层：数据平面和控制平面

□ 数据平面

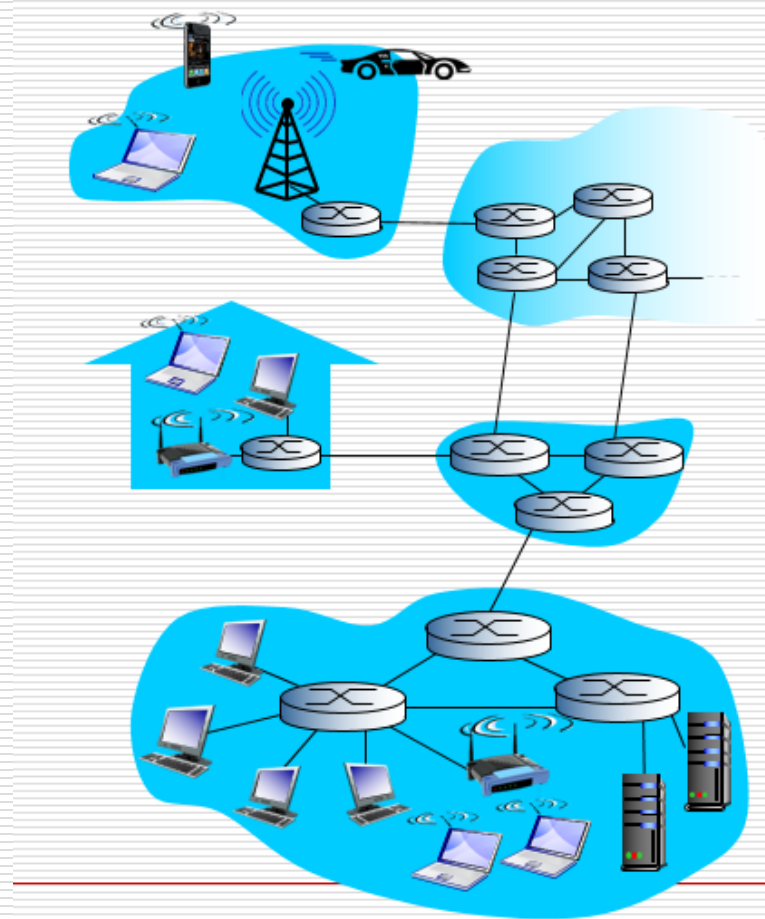
- 局部：每个路由器功能
- 决定从路由器输入端口到达的分组如何转发到输出端口
- 转发表

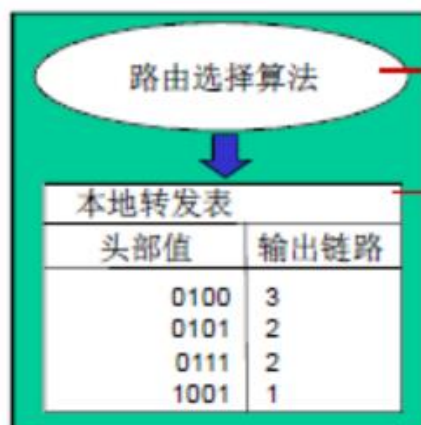


网络层：数据平面和控制平面

□ 控制平面

- 全局：多个路由器
- 决定数据报如何在路由器之间路由，决定数据报从源到目标主机之间的端到端路径

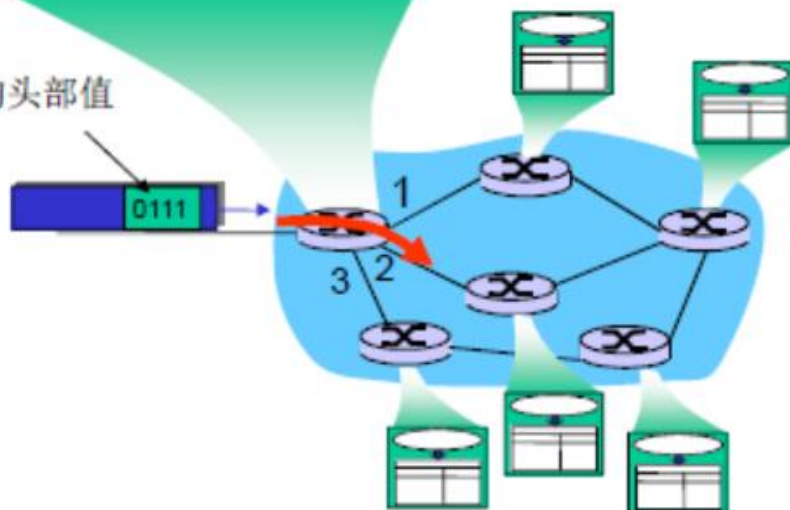




控制平面：路由算法决定
端到端路径

数据平面：IP协议根据转发表
决定了IP数据报在此路由器上的
局部转发

到达分组的头部值



网络服务模型

□ 从发送方主机到接收方主机传输数据报，网络层提供什么样的服务模型？

✓ 可靠交付

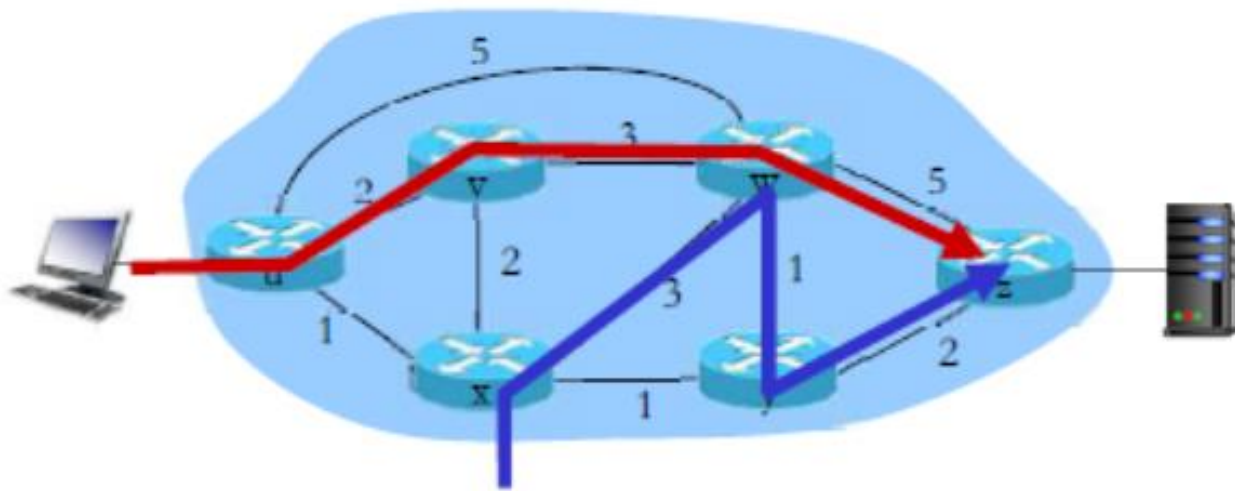
✓ 时延

✓ 有序

✓ 安全

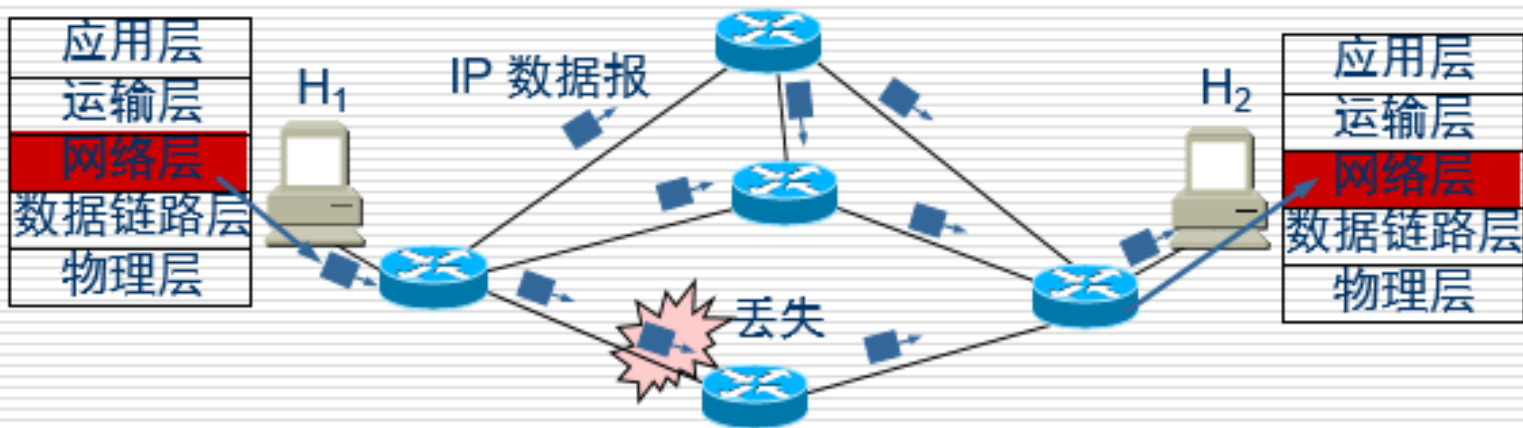
✓

虚电路（网络层建立连接）

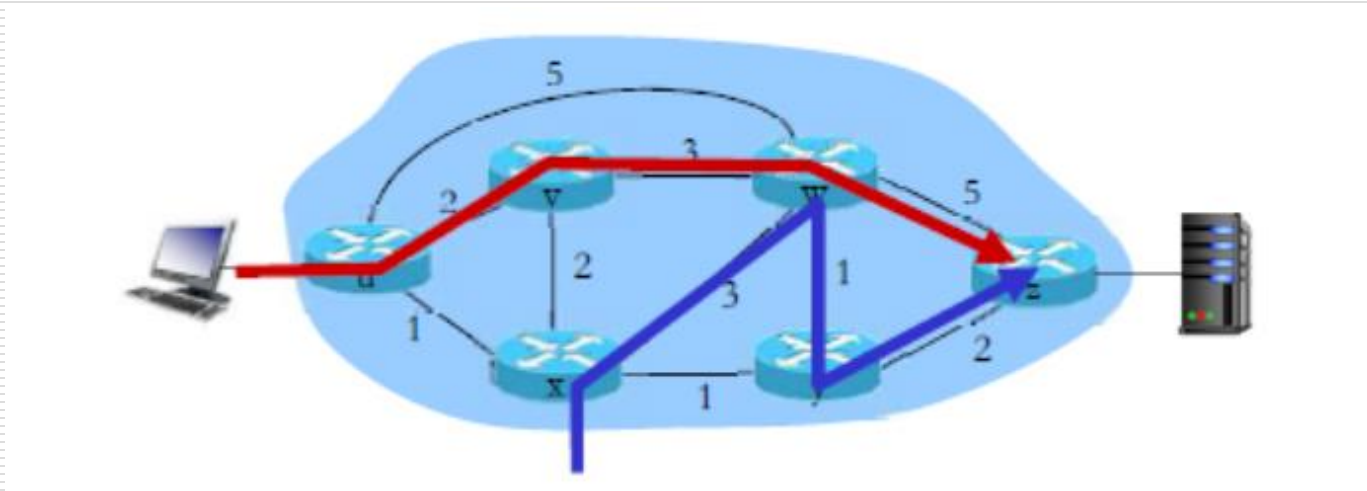


在分组传输之前，在两个主机之间，通过一些路由器所构成的路径上建立一个网络层连接（涉及到路由器）

数据报（网络层无连接）

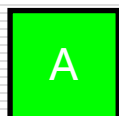


-
- ✓ 虚电路和数据报都属于分组交换方式
 - ✓ 虚电路和电路交换的区别



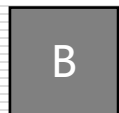
	数据报	虚电路
延时	分组传输延时	虚电路建立延时，分组传输延时
路由选择	每个分组单独选择路由	建立虚电路时选择路由，以后所有分组都使用该路由
状态信息	子网无需保存状态信息	每个结点要保存一张虚电路表
地址	每个分组携带完整的源/目的地址	每个分组分配一个较短的虚电路号
节点失败的影响	除了崩溃时正在该节点处理的分组都丢失外，无其他影响	所有经过失效节点的虚电路都要被终止
拥塞控制	难	容易

() 能保证分组的按序到达?



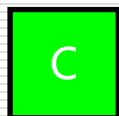
A

虚电路



B

数据报



C

电路交换

提交

□ Internet的网络层采用了数据报方式

网络架构	服务模型	保证？				拥塞反馈
		带宽	丢失	保序	延迟	
Internet	best effort	none	no	no	no	no (inferred via loss)

□ 网络层

1) 数据平面（转发）

- ✓ IP、转发算法
- ✓ ARP、ICMP、NAT

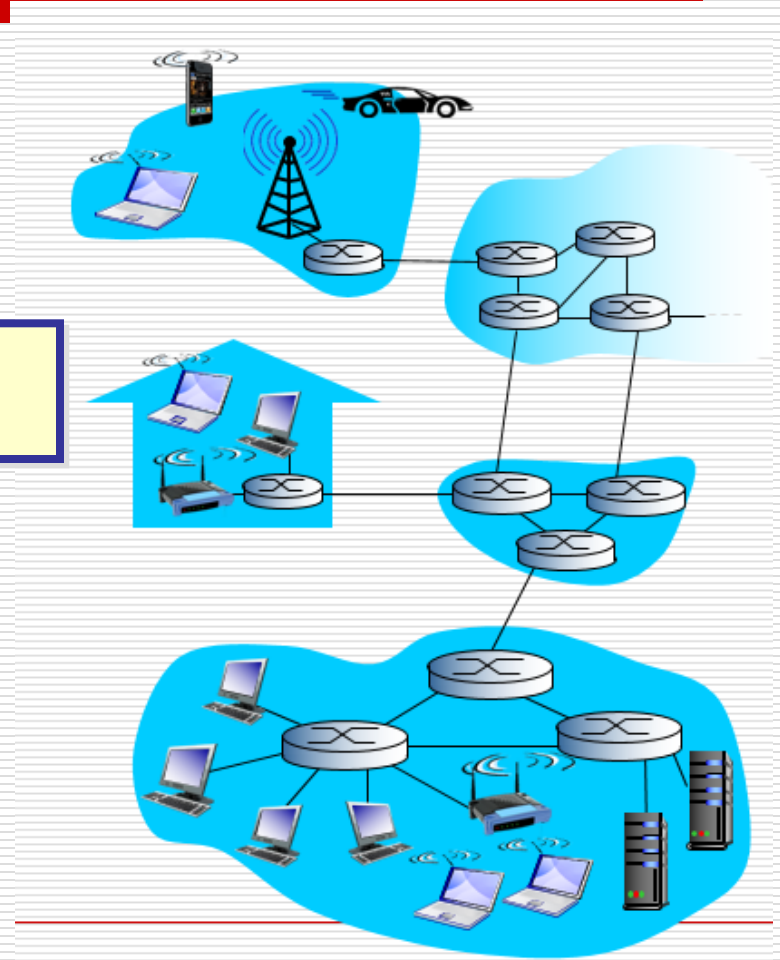
2) 控制平面（路由）

- ✓ 路由协议（RIP、OSPF、BGP）
-

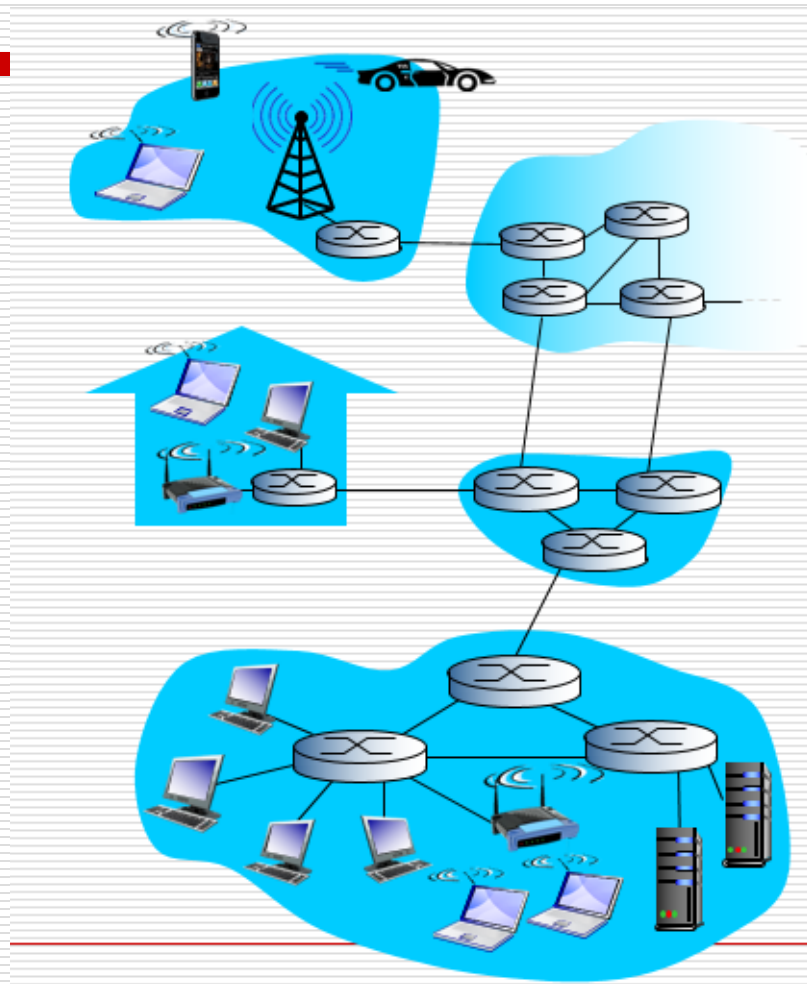
IP: Internet Protocol

□ IP地址: 32bit

IP 地址 ::= { <网络号>, <主机号> }

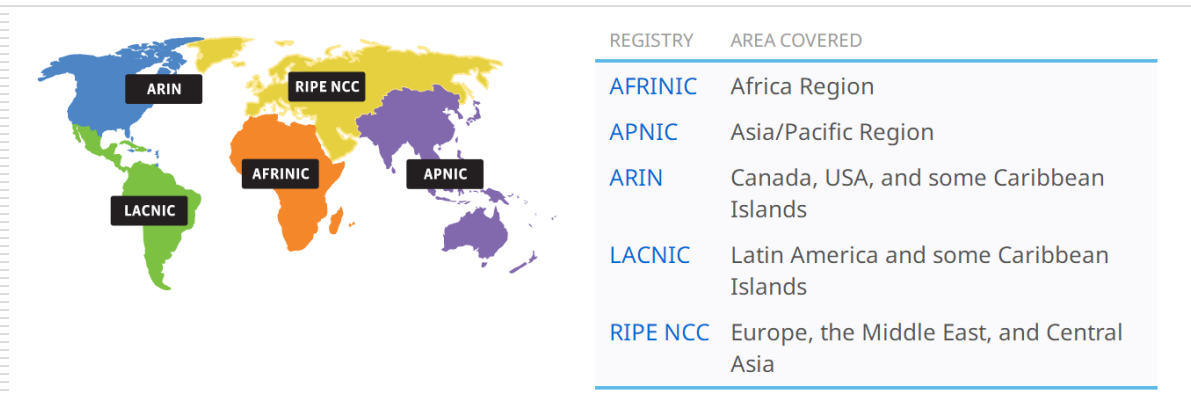


- ✓ IP地址：对主机或路由器的接口进行编址
- ✓ 接口：主机/路由器和物理链路的连接处



如何分配IP地址

□ 网络号，由国际互联网信息中心(InterNIC)分配



□ 主机号：系统管理员分配

IP 地址 ::= { <网络号>, <主机号> }

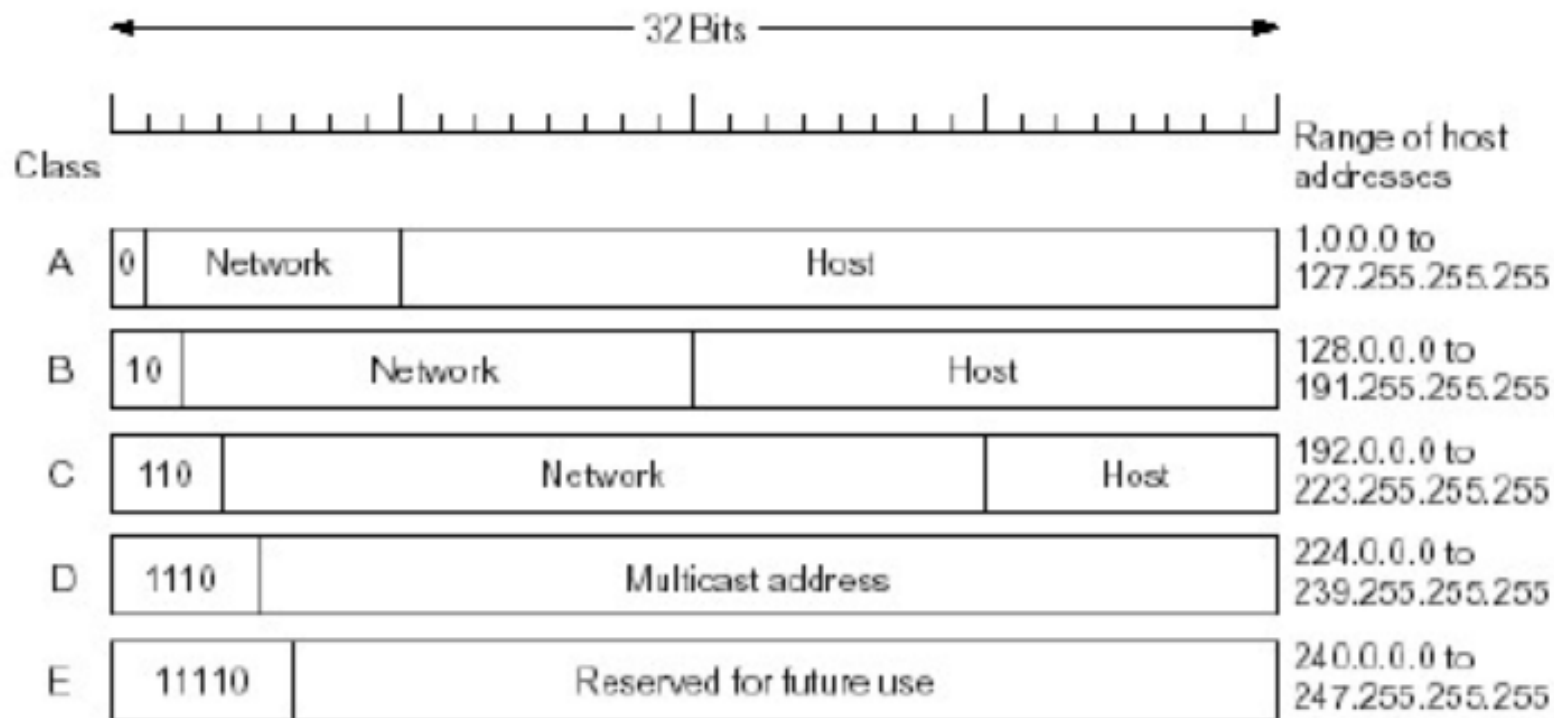
□ 网络的规模（网络容纳的主机）不同

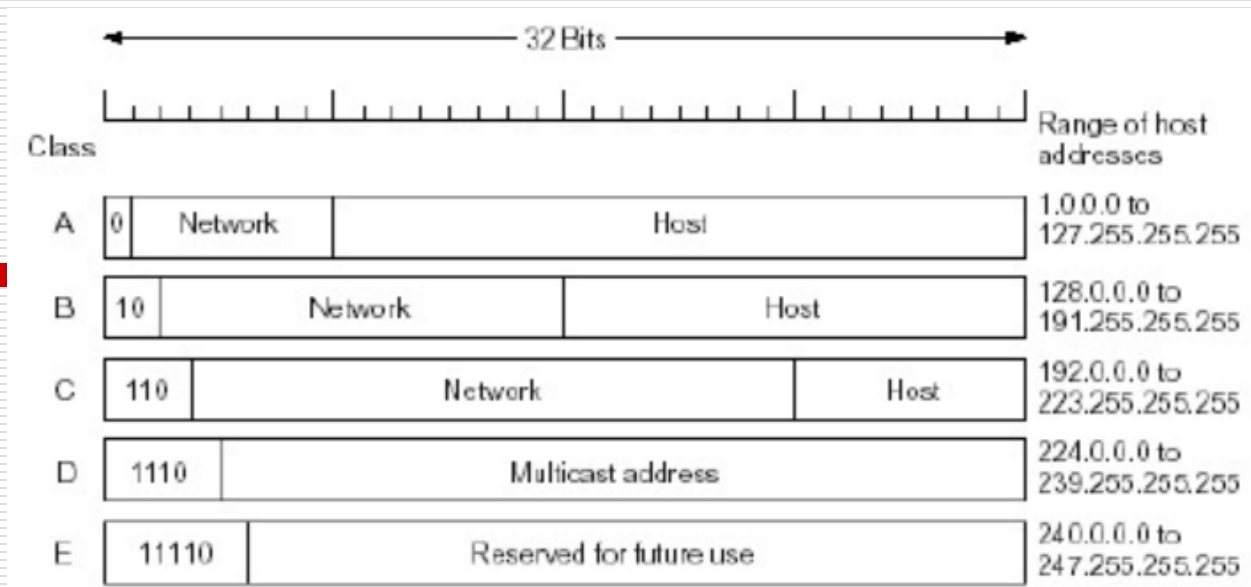
■ 网络号多长？

IP 地址 ::= { <网络号>, <主机号> }

■ 分类：A(8)、B(16)、C(24)三类

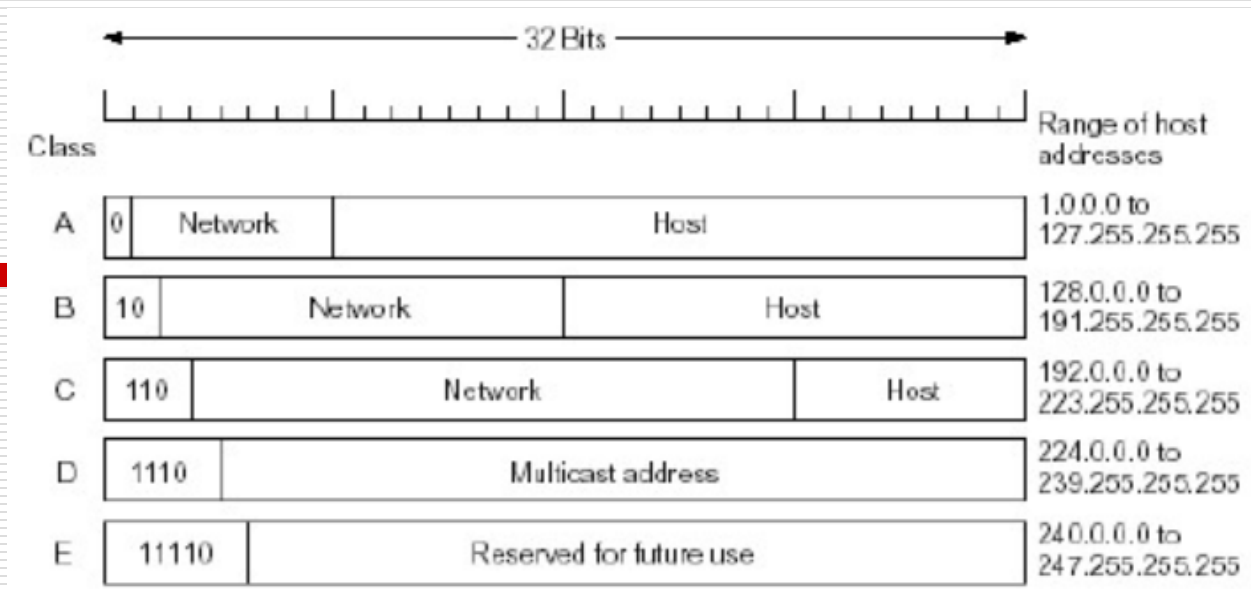
分类IP地址



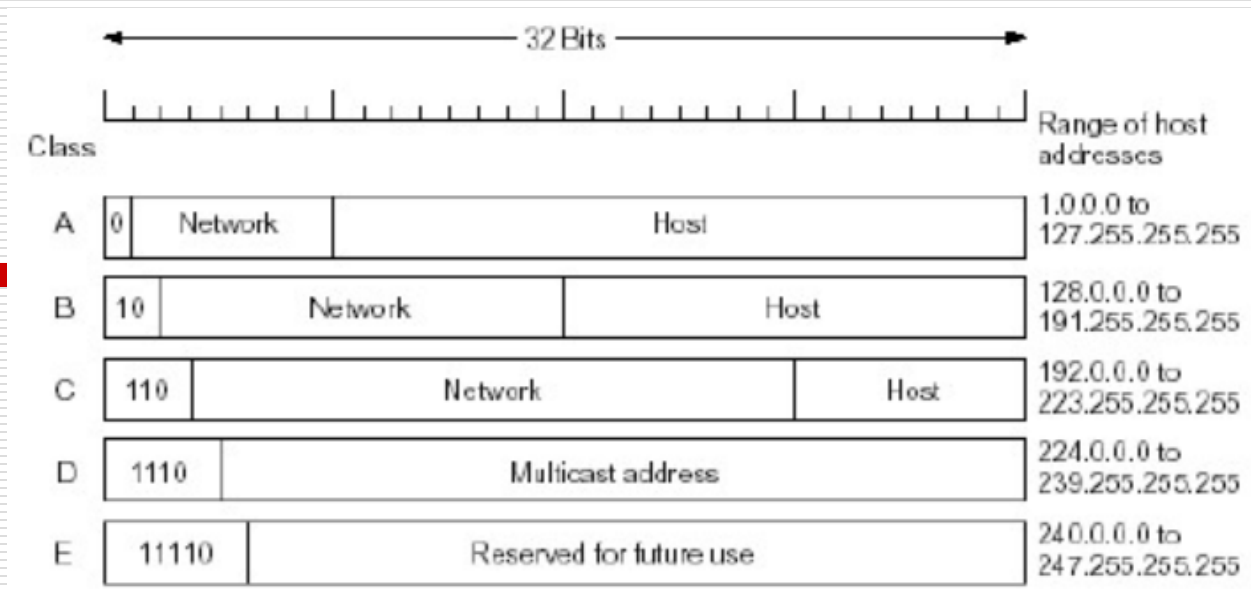


本网广播地址

- ✓ 32bit 全1: 255.255.255.255
- ✓ 主机号全1: ***.***. 255.255



- ✓ 主机号全0：为网络号，不作为主机IP地址分配
- ✓ 网络号全0：本网络的某个主机



● 127.*.*.*: 回环地址

- ✓ 允许运行在同一台主机上的客户程序和服务器程序通过 TCP/IP 进行通信
- ✓ 网络层检测到目的地址为回环地址，则不发送到链路层，而是放入 IP 回环接口，传给本机的上层 TCP

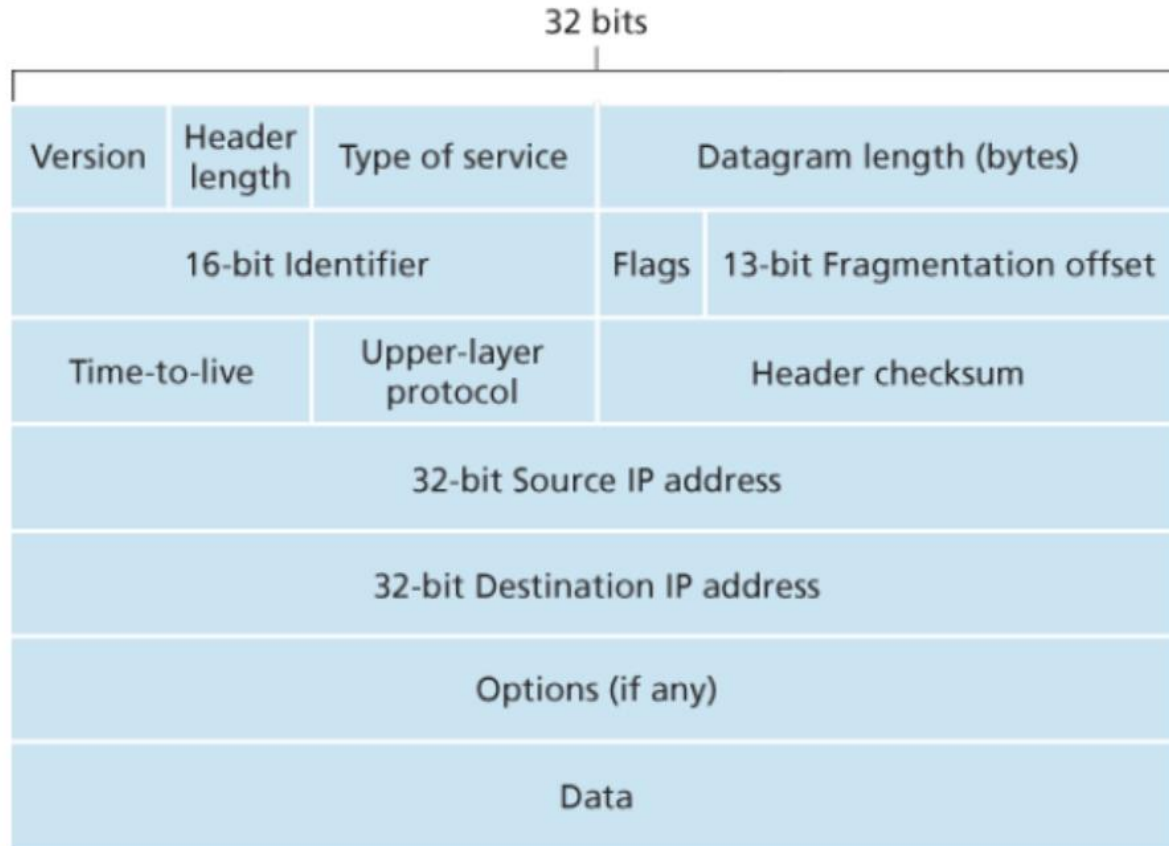
某主机IP地址：180.172.32.12

- 1) 所在网络的分类[填空1]
- 2) 该网络的网络号 [填空2]
- 3) 该网络的广播地址 [填空3]
- 4) 可分配的IP地址范围 [填空4] [填空5]

作答

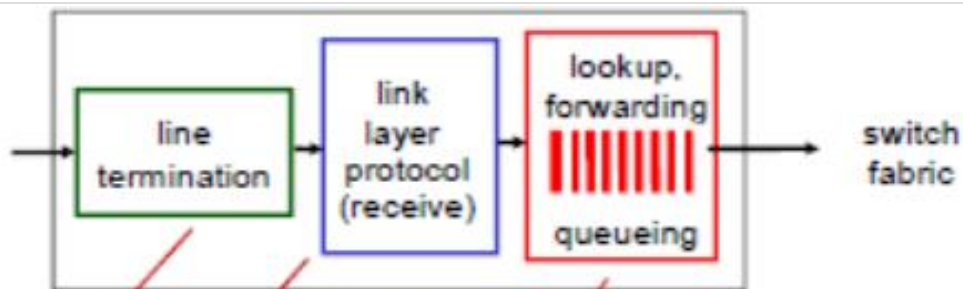
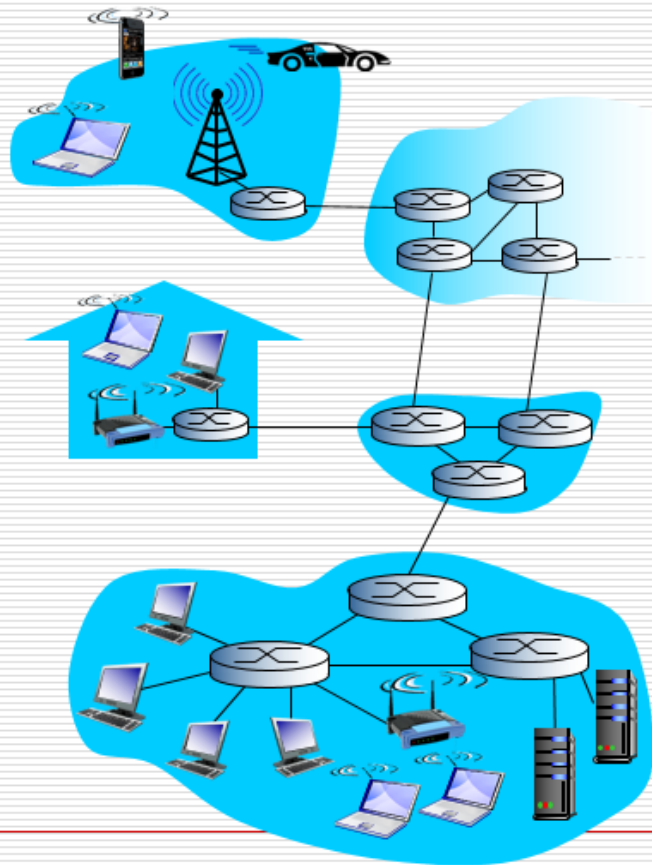
IP: Internet Protocol

□ IP分组/IP数据报



IP: Internet Protocol

□ 路由器转发



物理层:
Bit级的接收

数据链路层

operates in millisecond
time frame

□ 转发表

网络号	下一跳地址	接口
.....
.....
.....

- ✓ **网络号**：路由器为**每一个网络**指定路由，而不是为**每个主机**指定路由，极大缩小了转发表的规模
- ✓ **下一跳**：**相邻路由器的IP地址**（相邻：两个路由器直接相连或连在同一个物理网络中）
- ✓ **接口**：路由器的端口标识

转发

□ 主机

- 目的主机与源主机在一个物理网络中（如以太网），直接交到链路层，
- 否则，发给默认路由器，由路由器转发

□ 路由器

- 根据IP数据报中的目的地址，提取网络号，查找转发表，向对应接口转发
-

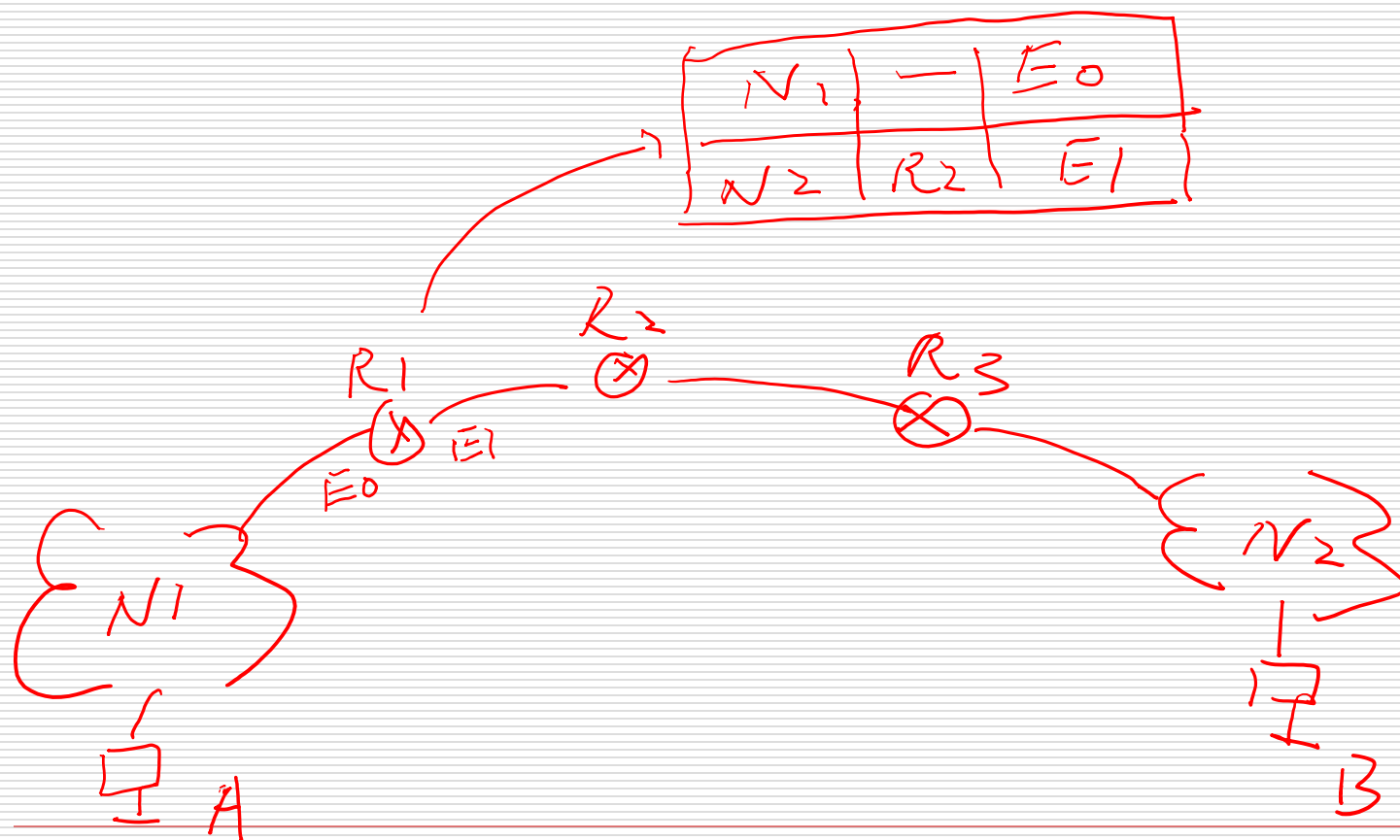
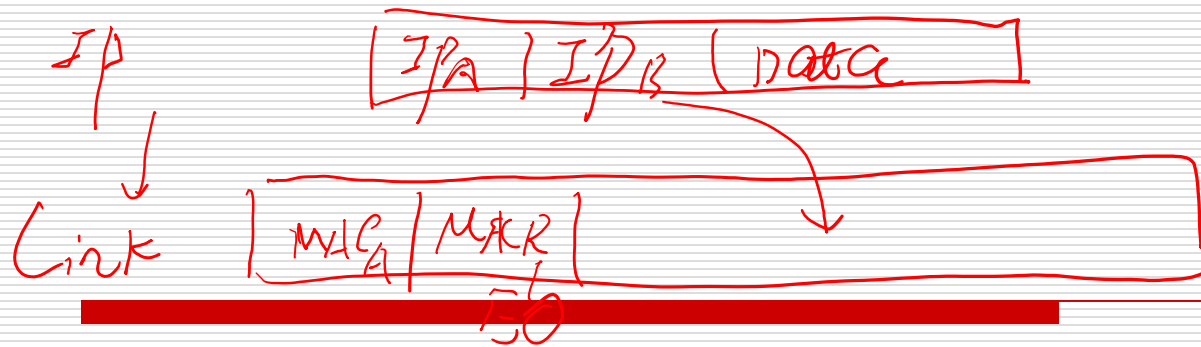
□ 网络层：转发

■ IP协议（IP地址、IP数据报首部格式）

- ✓ 分层的IP地址（32bit）
- ✓ 源IP，目的IP，TTL，首部长度的，

■ 路由器：转发表

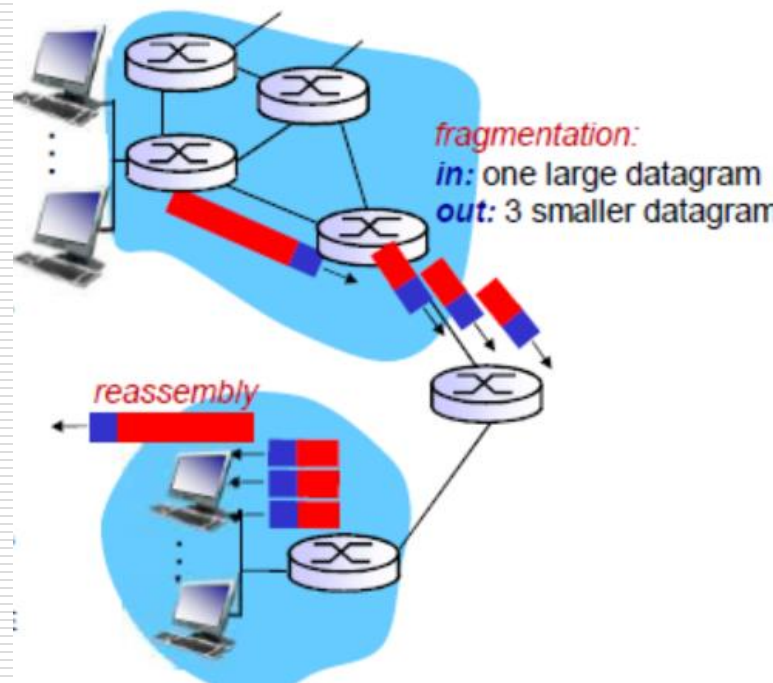
网络号、下一跳、端口号



IP: Internet Protocol

□ 分片与重组

- 不同网络，链路层的**MTU**（最大传输单元）不同
- **分片**：一个IP数据报被分成若干小的数据报
- **重组**：在目的主机进行

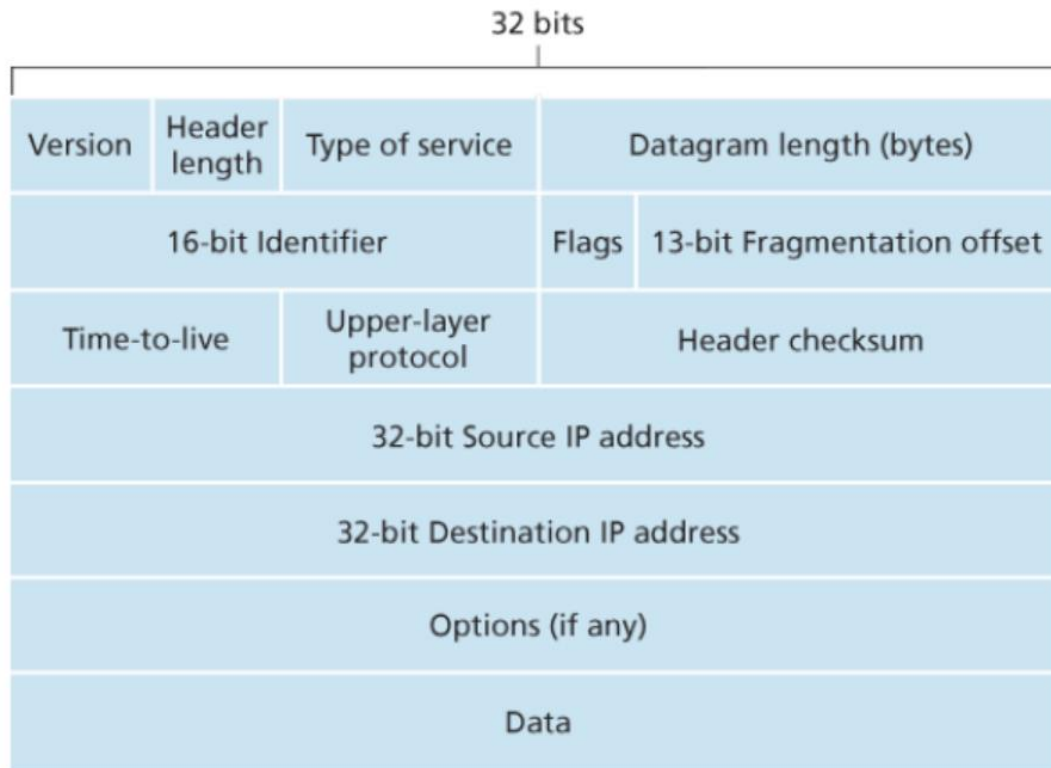


□ 标志位

3位，DF (Don't Fragment)：“0”能分片；“1”不能分片

MF (More Fragment)：“1”后面还有分片；“0”最后一个分片

□ 偏移量 (13bit)，单位：8bytes



IP分片与重组

例

Length: 2400↵

ID=X↵

DF:0

MF:0↵

Offset: 0↵

□ 2400字节数据报

■ 固定首部：20字节

■ 数据长度：2380字节

□ MTU：1000字节

1) 20字节+976字节

Length: 996↵

ID=X↵

DF:0

MF:1↵

Offset: 0↵

2) 20字节+976字节

Length: 996↵

ID=X↵

DF:0

MF:1↵

Offset: 122↵

3) 20字节+428字节

Length: 448↵

ID=X↵

DF:0

MF:0↵

Offset: 244↵

路由器数据平面（转发）工作在（）

- ☐ A 物理层
- ☐ B 链路层
- ☒ C 网络层
- ☐ D 应用层

提交

一个C类网能容纳的主机数最多为（）

- ☐ A 65535
- ☐ B 256
- ☐ C 65534
- ☒ D 254

提交

路由器根据（）进行IP分组的转发

- ☐ A 源主机的IP地址
- ☐ B 源主机所在的网络号
- ☐ C 目的主机的IP地址
- ☒ D 目的主机所在的网络号

提交

IP: Internet Protocol

□ 子网 (subnet)

■ 分类IP地址存在的问题

B类网: 128.10.0.0

将一个网络划分若干地址空间不重复的网络——
subnet

□ 子网划分

- 从主机号取若干位，与网络号一起作为子网号

B类网：128.10.0.0

划分成2个子网

例： 128.10.0.0

4个子网

8个子网

□ 划分子网后带来的问题

B类网：128.10.0.0

划分成2个子网

□ 子网掩码 (subnet mask)

32位：对应网络号的位置1，对应主机号的位置0

网络号=IP地址 \cap 子网掩码

B类网: 128.10.0.0

划分成2个子网

子网掩码:

A类、B类、C类

——255.0.0.0

——255.255.0.0

——255.255.255.0

网络号↵	子网掩码↵	下一跳地址↵	接口↵
↵	↵	↵	↵
↵	↵	↵	↵

□ 固定长度子网

- ✓ 子网号长度固定，子网掩码相同
- ✓ 每个子网的主机数相同

□ 变长子网

P184

□ 202.120.224.0, 划分成5个子网, 3个子网容纳主机数50台, 2个子网容纳主机数30台

子网号能否为全0和全1?

□ 路由器转发表示例

网络号↵	子网掩码↵	下一跳地址↵	接口↵
特定主机↵	255.255.255.255↵	R1↵	E1↵
223.1.1.0↵	255.255.255.0↵	直连↵	E0↵
223.1.2.0↵	255.255.255.0↵	R1↵	E1↵
默认路由↵	0.0.0.0↵	R2↵	E2↵

转发算法

1. 路由器从IP数据报的首部提取**目的IP地址**，与**子网掩码**“与”运算，获得网络号N
2. 判断是否可**直接交付**
3. 判断是否为**特定主机路由**
4. 判断是否为**间接路由**
5. **默认路由/报告错误**

网络号 [⌘]	子网掩码 [⌘]	下一跳地址 [⌘]	接口 [⌘]
特定主机 [⌘]	255.255.255.255 [⌘]	R1 [⌘]	E1 [⌘]
223.1.1.0 [⌘]	255.255.255.0 [⌘]	直连 [⌘]	E0 [⌘]
223.1.2.0 [⌘]	255.255.255.0 [⌘]	R1 [⌘]	E1 [⌘]
默认路由 [⌘]	0.0.0.0 [⌘]	R2 [⌘]	E2 [⌘]

□ 分类IP地址

- 子网划分，缓解地址空间的分配

- A类和B类越来越少

- C类网络增多，路由表规模不断增长

□ 无类别域间路由选择（CIDR: Classes InterDomain Routing）

——取消分类IP地址的网络号长度规定，网络号可以为任意长度

CIDR

- **斜线记法:**在 IP 地址面加上一个斜线 “/”，写上网络号（网络前缀）所占的位数

128.14.32.0/20

子网掩码： 11111111 11111111 11110000 00000000

根据需要分配适当大小的 CIDR 地址块,分配到一个**CIDR地址块**的组织，仍可以根据需要划分子网

CIDR 前缀长度	点分十进制	包含的地址数	相当于包含分类的网络数
/13	255.248.0.0	512 K	8 个 B类或 2048 个 C 类
/14	255.252.0.0	256 K	4 个 B 类或1024 个 C 类
/15	255.254.0.0	128 K	2 个 B 类或512 个 C 类
/16	255.255.0.0	64 K	1 个 B 类或256 个 C 类
/17	255.255.128.0	32 K	128 个 C 类
/18	255.255.192.0	16 K	64 个 C 类
/19	255.255.224.0	8 K	32 个 C 类
/20	255.255.240.0	4 K	16 个 C 类
/21	255.255.248.0	2 K	8 个 C 类
/22	255.255.252.0	1 K	4 个 C 类
/23	255.255.254.0	512	2 个 C 类
/24	255.255.255.0	256	1 个 C 类
/25	255.255.255.128	128	1/4 个 C 类
/26	255.255.255.192	64	1/4 个 C 类
/27	255.255.255.224	32	1/8 个 C 类

□ 分类IP地址

- 子网划分，缓解地址空间的分配

- A类和B类越来越少 -----CIDR

- C类网络增多，路由表规模不断增长

-
- CIDR支持路由聚合(地址聚会)，减小路由表的规模

目的网络	掩码	下一跳
192.60.128.0/24	255.255.255.0	R1
192.60.129.0/24	255.255.255.0	R1
192.60.130.0/24	255.255.255.0	R1
192.60.131.0/24	255.255.255.0	R1

-
- 路由聚合带来的问题：一个IP数据报可能会匹配到多个选项

目的网络	掩码	下一跳
192.60.128.0/24	255.255.255.0	R1
192.60.129.0/24	255.255.255.0	R1
192.60.130.0/24	255.255.255.0	R1
192.60.131.0/24	255.255.255.0	R1

最长网络前缀匹配

某路由表中的4条路由选项，下一跳相同，

35.230.32.0/21

35.230.40.0/21

35.230.48.0/21

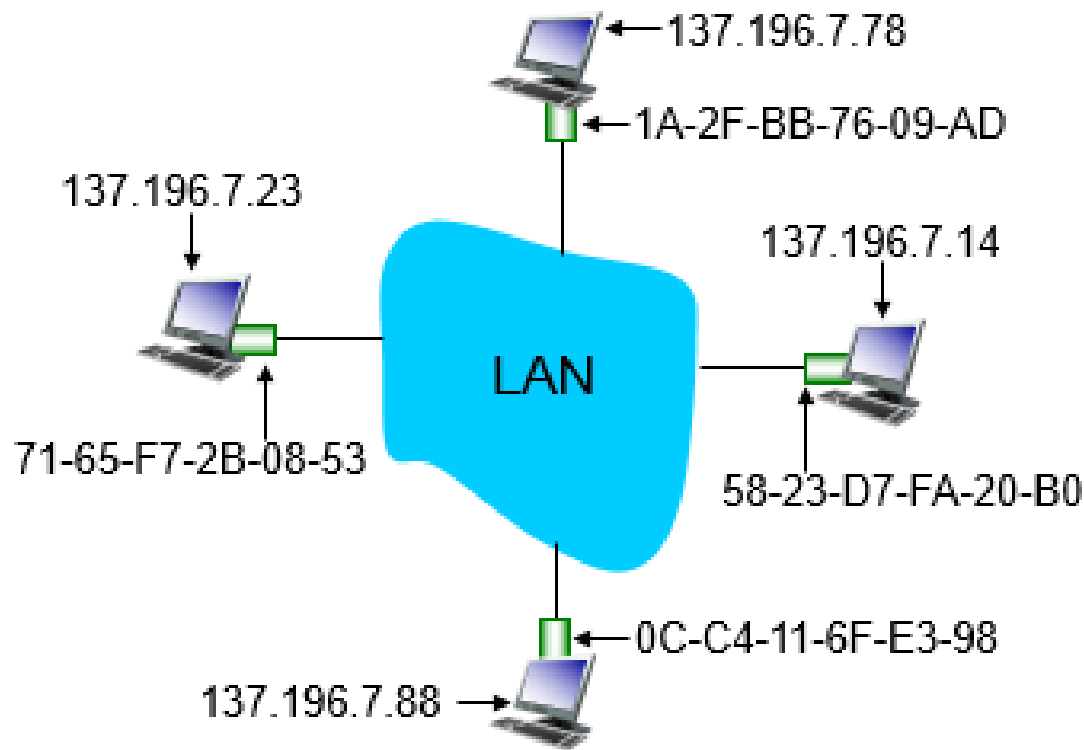
35.230.56.0/21

聚合后的路由： [填空1] ， Mask [填空2]

正常使用填空题需3.0以上版本雨课堂

作答

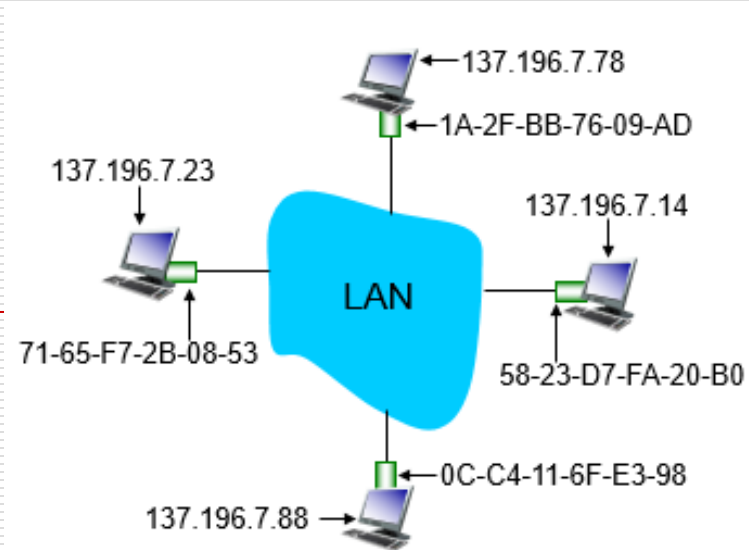
ARP: Address Resolution Protocol



□ 每台主机或路由器在内存中维持一个ARP表

< IP address; MAC address; TTL >

A向B发送IP数据报，B的MAC地址不在ARP表中？



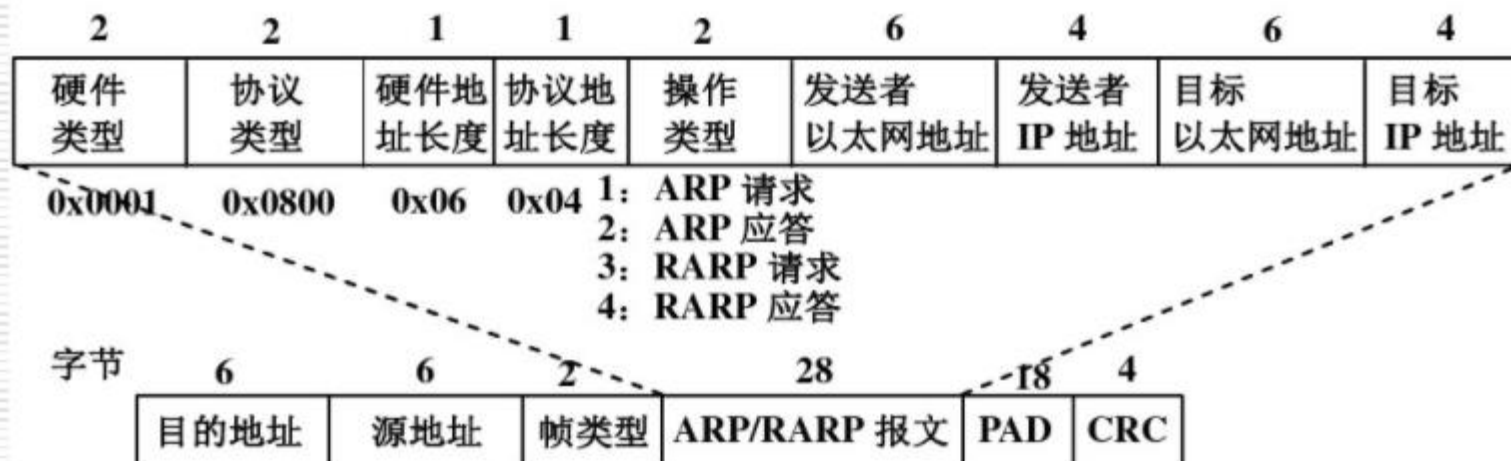
□ 则A构造ARP查询分组

- ✓ 查询分组：A的IP地址、A的MAC地址、B的IP地址、B的MAC地址（全0）
- ✓ 封装成帧（目的MAC地址：FF-FF-FF-FF-FF-FF）

-
- B收到ARP查询分组，回复ARP响应分组（单播）

< IP address; MAC address; TTL >

ARP分组



ARP: 0x0806

RARP: 0x8035

硬件类型：值1（以太网）

协议类型：值0800（IPv4）

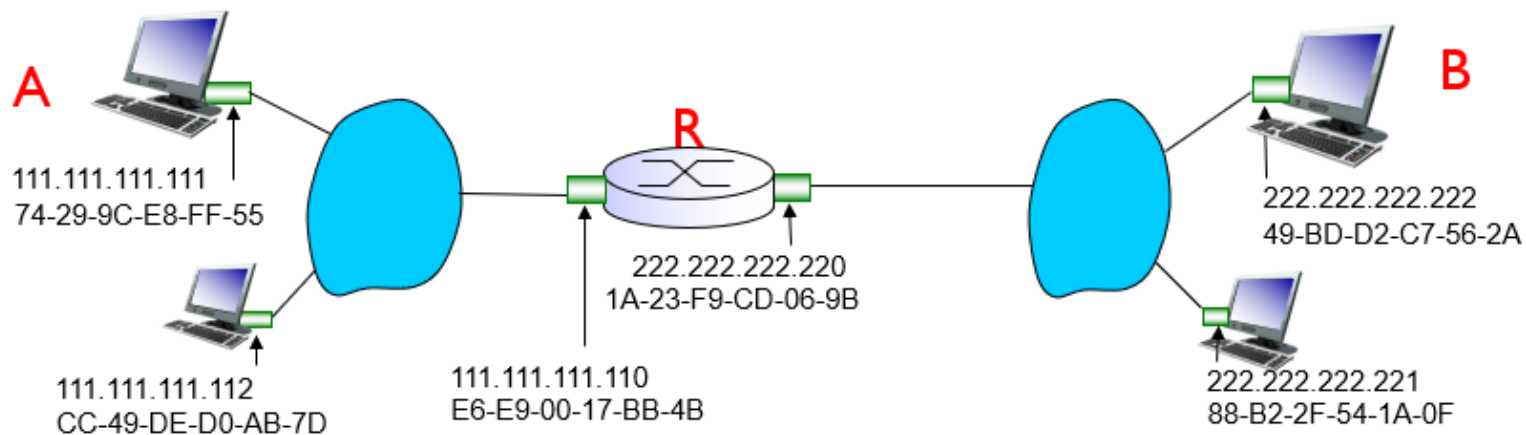
硬件地址长度：值6

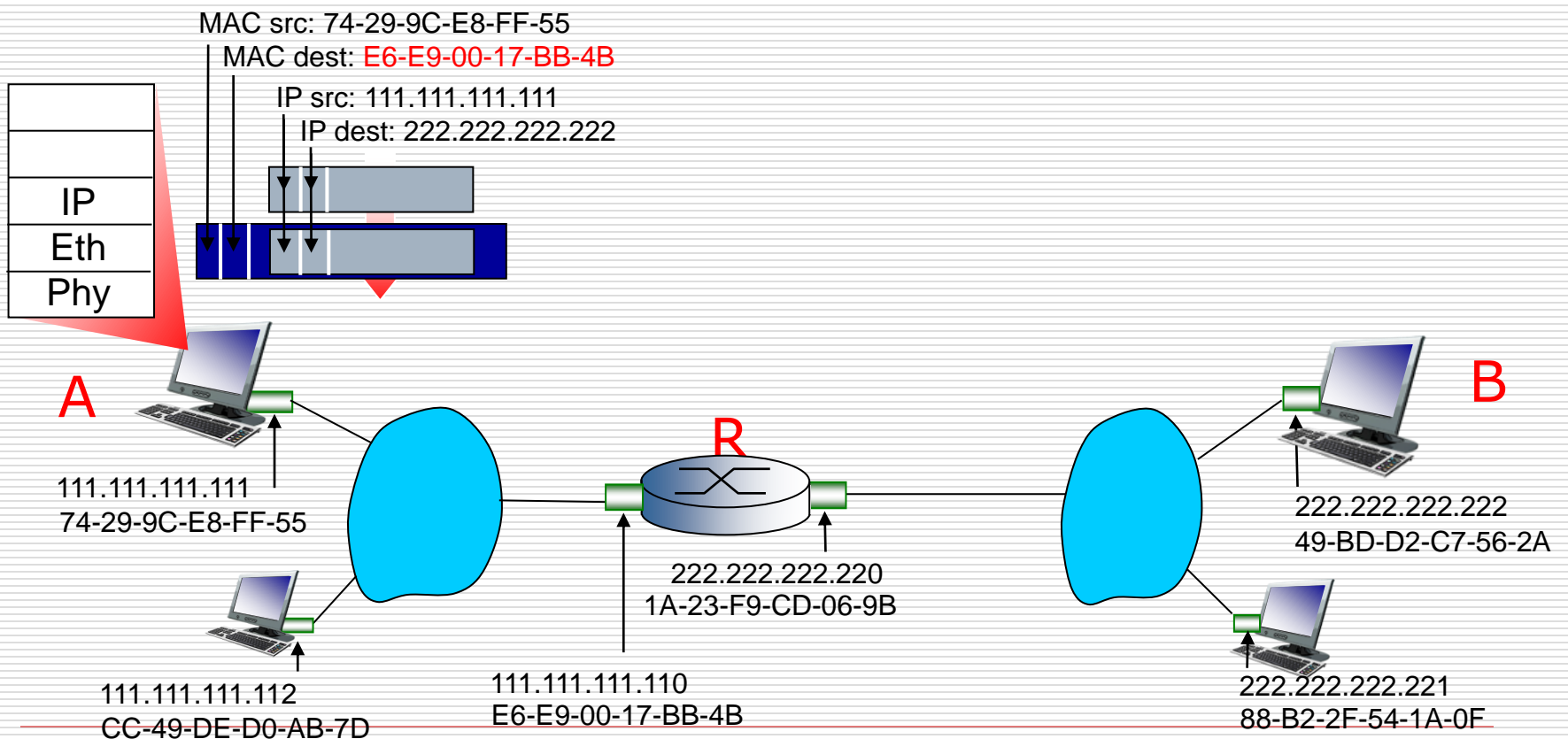
协议地址长度：值4

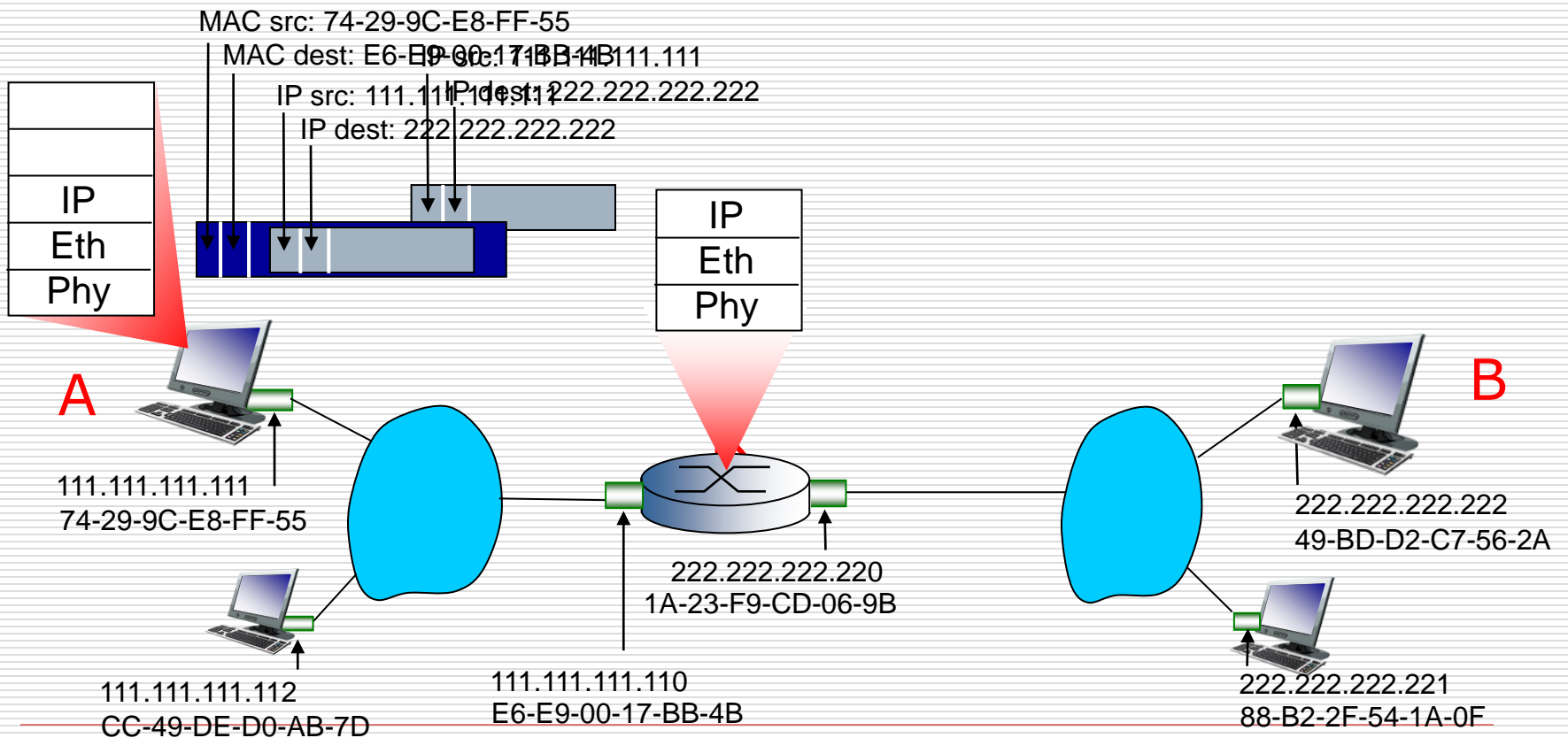
操作：请求1，响应2

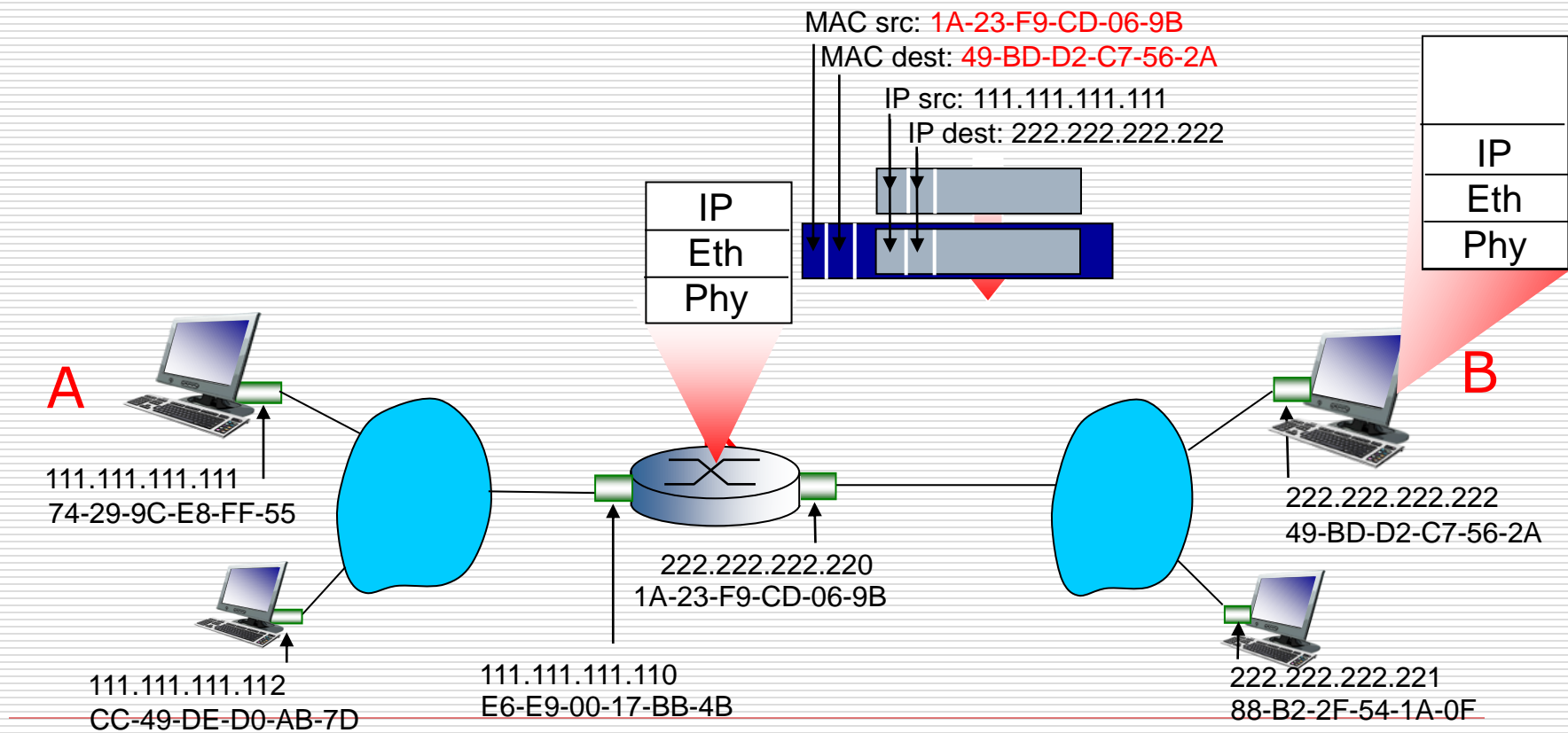
- 发送方硬件地址（以太网：6个字节）
- 发送方协议地址（IP：4个字节）
- 目标硬件地址（以太网：6个字节）
- 目标协议地址（IP：4个字节）

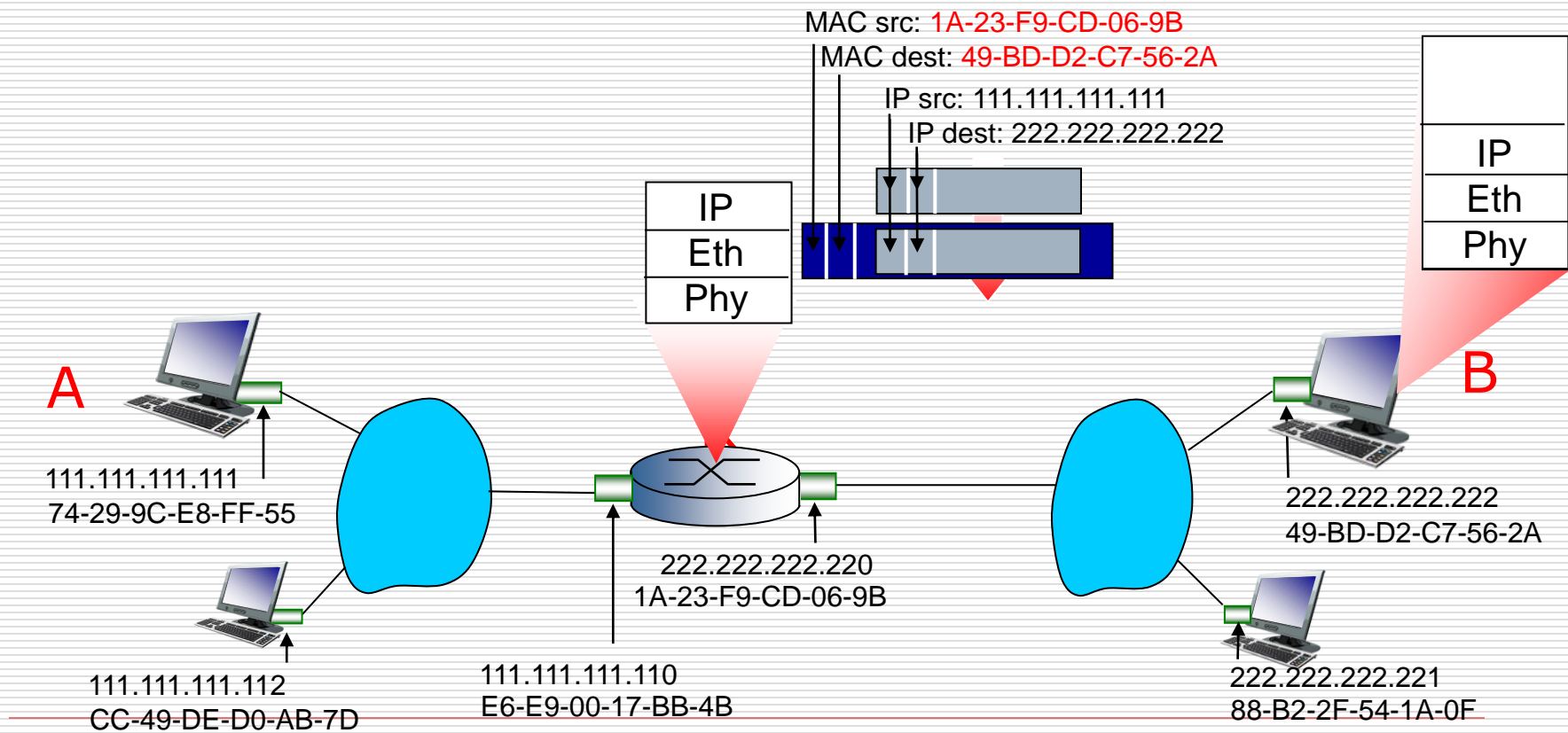
不在一个LAN?

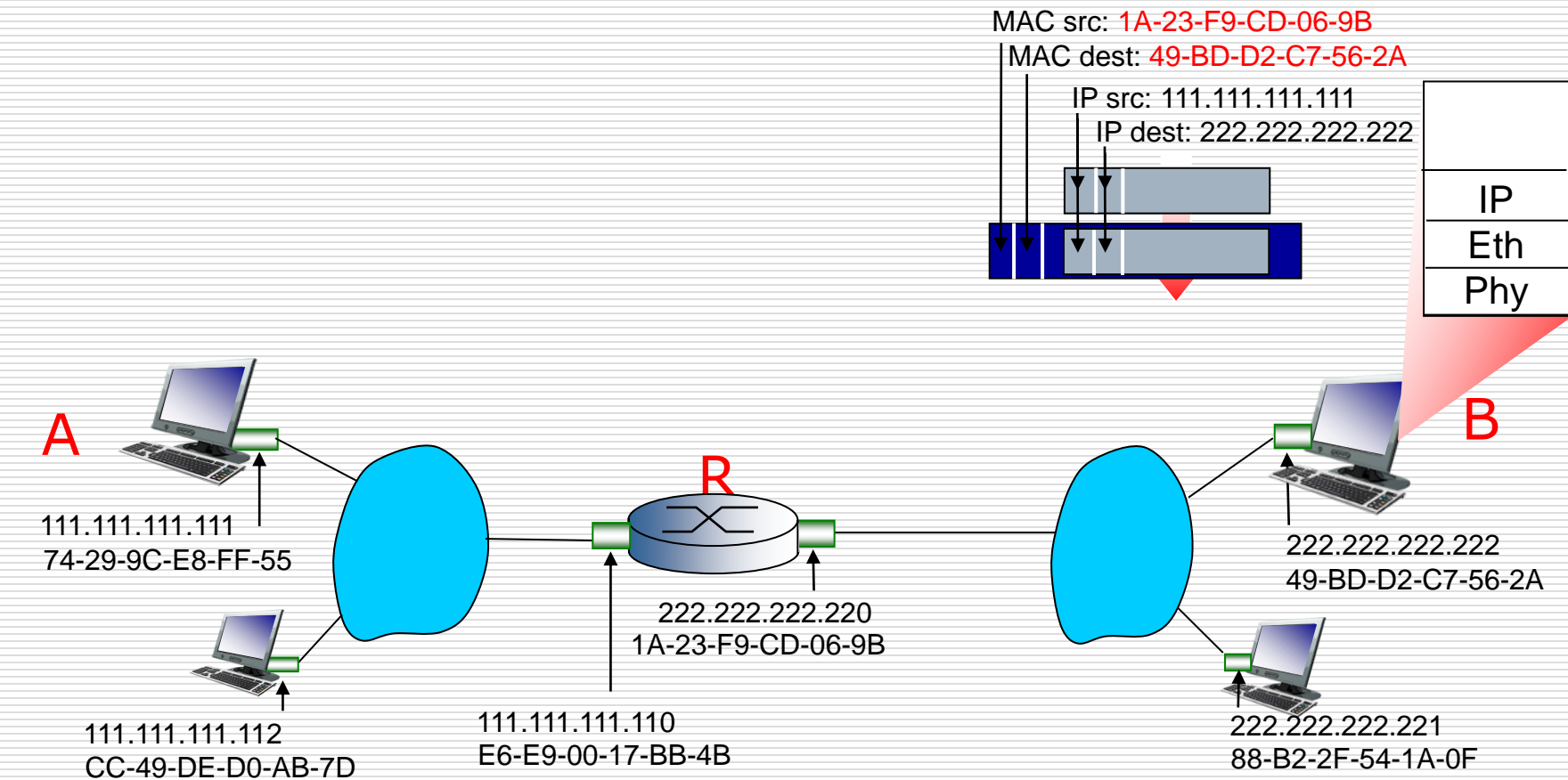










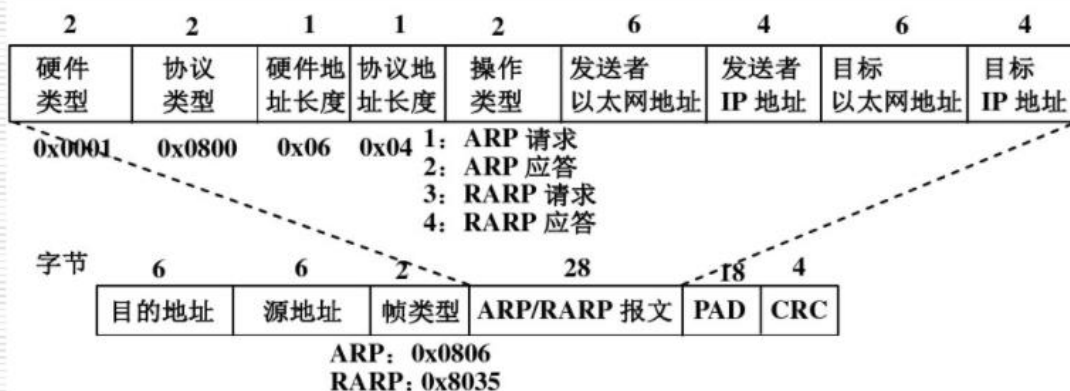


ARP

□ 网络层协议？

□ 链路层协议？

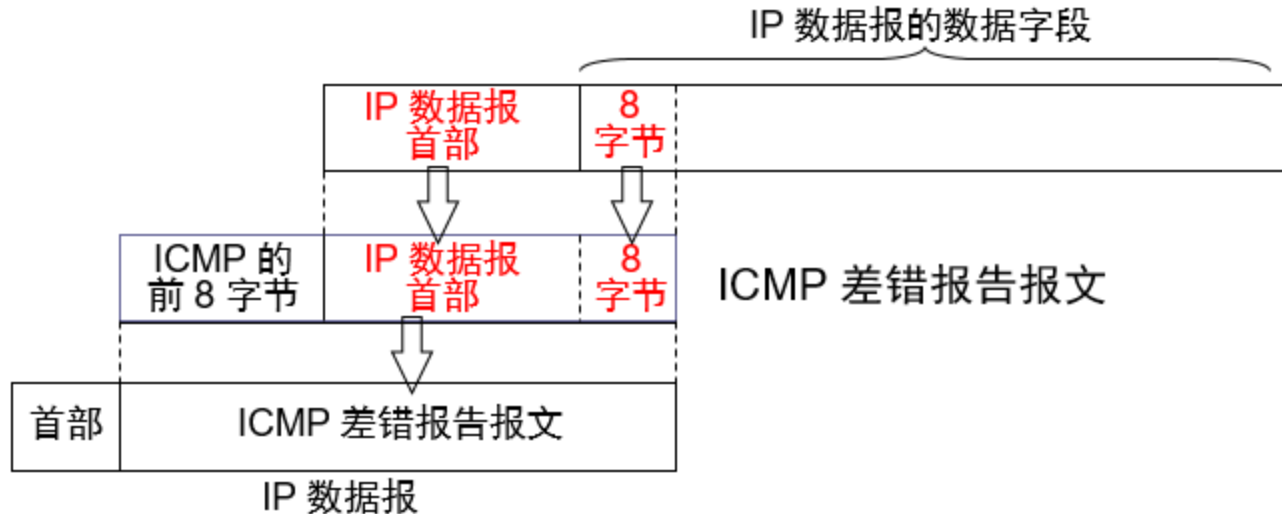
——跨链路层和网络层的协议



ICMP: Internet Control Message

□ Internet 控制报文协议

- ✓ 路由器转发或交付出错，向源主机发送差错报告
 - ✓ 传递网络层信息
-



❑ 差错报告

- ✓ 不可达（网络/主机/协议/端口）
- ✓ 超时（TTL为0）
- ✓ 路由重定向
- ✓

□ ICMP查询报文

——测试主机或路由器在网络层是否可达

Echo请求报文

Echo响应报文

□ Ping: ICMP 的Echo请求和响应报文

□ Traceroute/Tracert

- ✓ ICMP差错报文（TTL超时）

- ✓ ICMP差错报文（端口不可达）

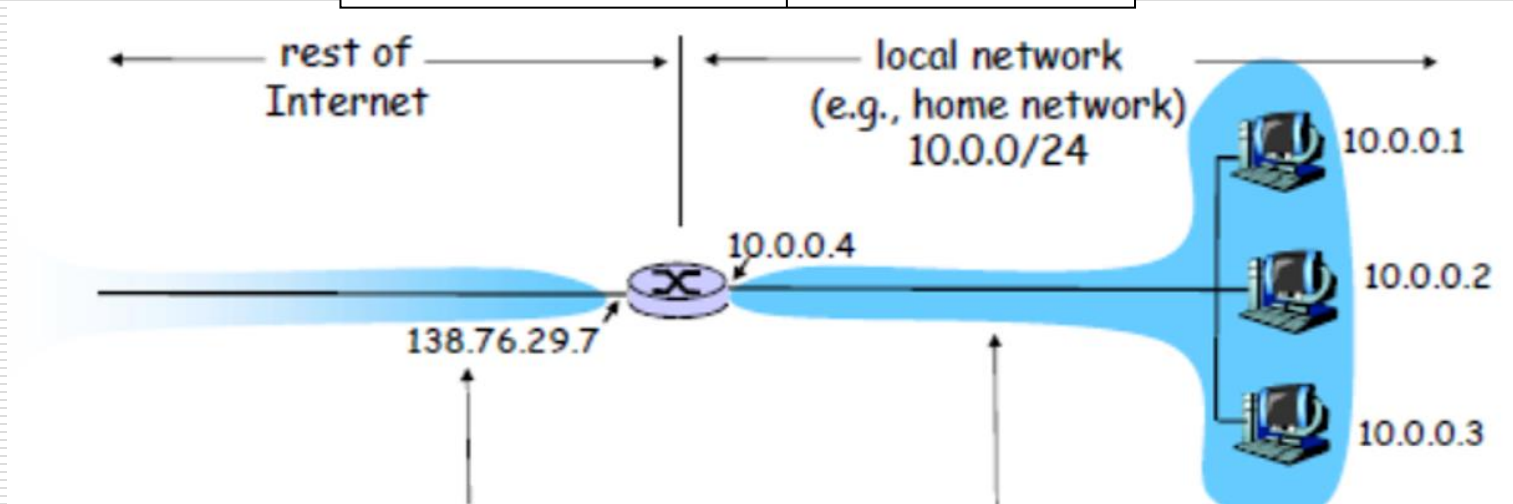
NAT: Internet Control Message

□ 内部网络（private: 专有）地址

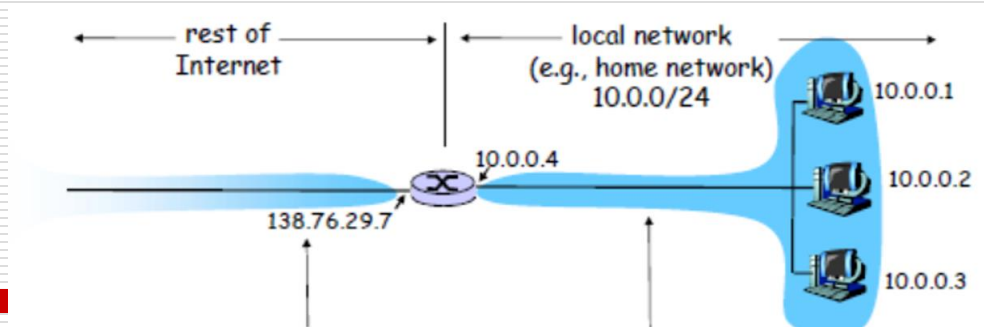
- ✓ 仅在**机构内部使用的 IP 地址**，可以由本机构自行分配，不需要向互联网管理机构申请（仅在机构内网中有意义，区分不同的设备）

- | | |
|---------------------------------------|--------------------|
| ● Class A 10.0.0.0-10.255.255.255 | MASK 255.0.0.0 |
| ● Class B 172.16.0.0-172.31.255.255 | MASK 255.255.0.0 |
| ● Class C 192.168.0.0-192.168.255.255 | MASK 255.255.255.0 |
-

NAT 转换表	
WAN side addr	LAN side addr
138.76.29.7	10.0.0.1
.....



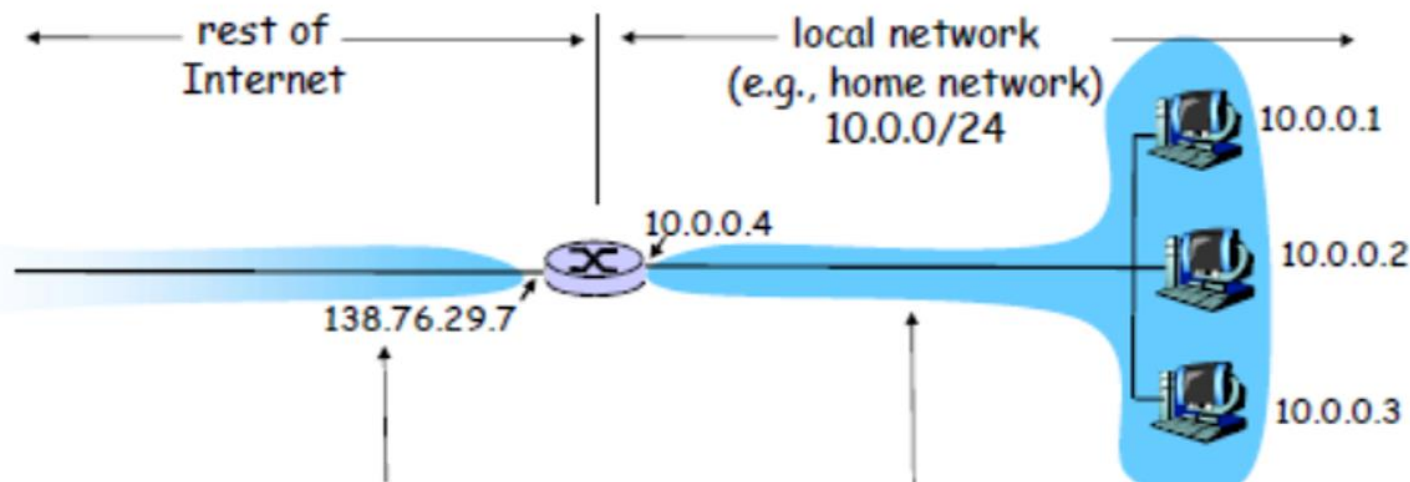
NAT: 网络地址转换



□ NAT的动机：本地网络只有1个或几个外部IP地址

- 节省地址空间（省钱）：不需要从ISP申请一个地址块，将1个或几个IP地址用于局域网所有设备
- 安全：局域网内部设备对外不可见
- 互不影响：内部地址和外部ISP地址更换，不会互相影响

NAT 转换表	
WAN side addr	LAN side addr
138.76.29.7	10.0.0.1
.....



NAT translation table	
WAN side addr	LAN side addr
138.76.29.7, 5001	10.0.0.1, 3345
.....

□ 对NAT（NAPT）的争议

- ✓ NAPT：路由器转发工作在网络层，使用了传输层的端口号
- ✓ 外网到内网的访问（P2P）——NAT穿越

地址空间不足：IPv6

作业

P249: 5.1

P250: 5.8, 5.11, 5.12, 5.13
