

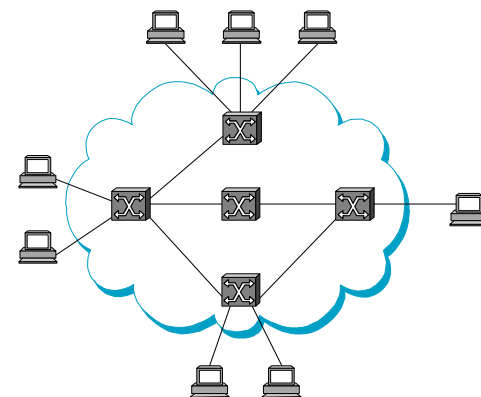
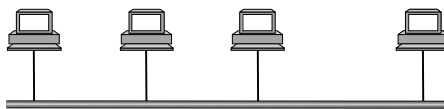
## 计算机网络 Computer Networks

# 网络互联：IP协议

# 单一网络的局限

## □ 单一类型的网络

- 互联的主机采用相同的寻址机制, 介质访问控制协议, 及服务模型, 等等



## □ 局限性

### ➤ 异构性

- 当某一类型网络中的用户期望与另一类型网络中的用户进行通信, 如何提供跨网络的主机之间的服务?

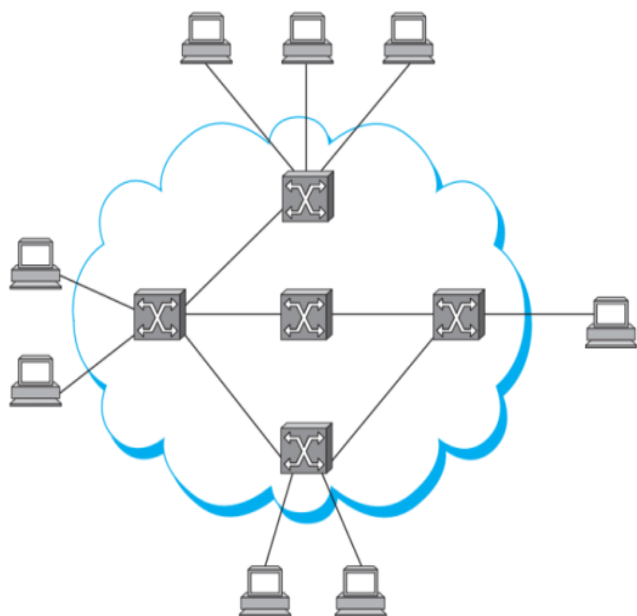
### ➤ 扩展性

- 当网络规模增大, 如何发现有效的路径?
- 当网络规模增大, 如何对网络节点进行有效的标识?

# 网络互联

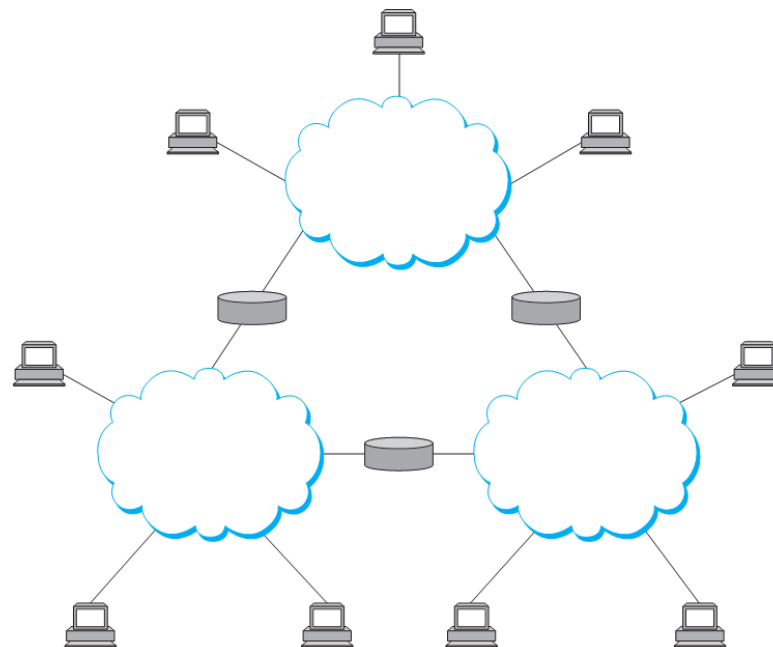
## □ 单一网络面临的问题

➤ 网络异构和扩展的局限性

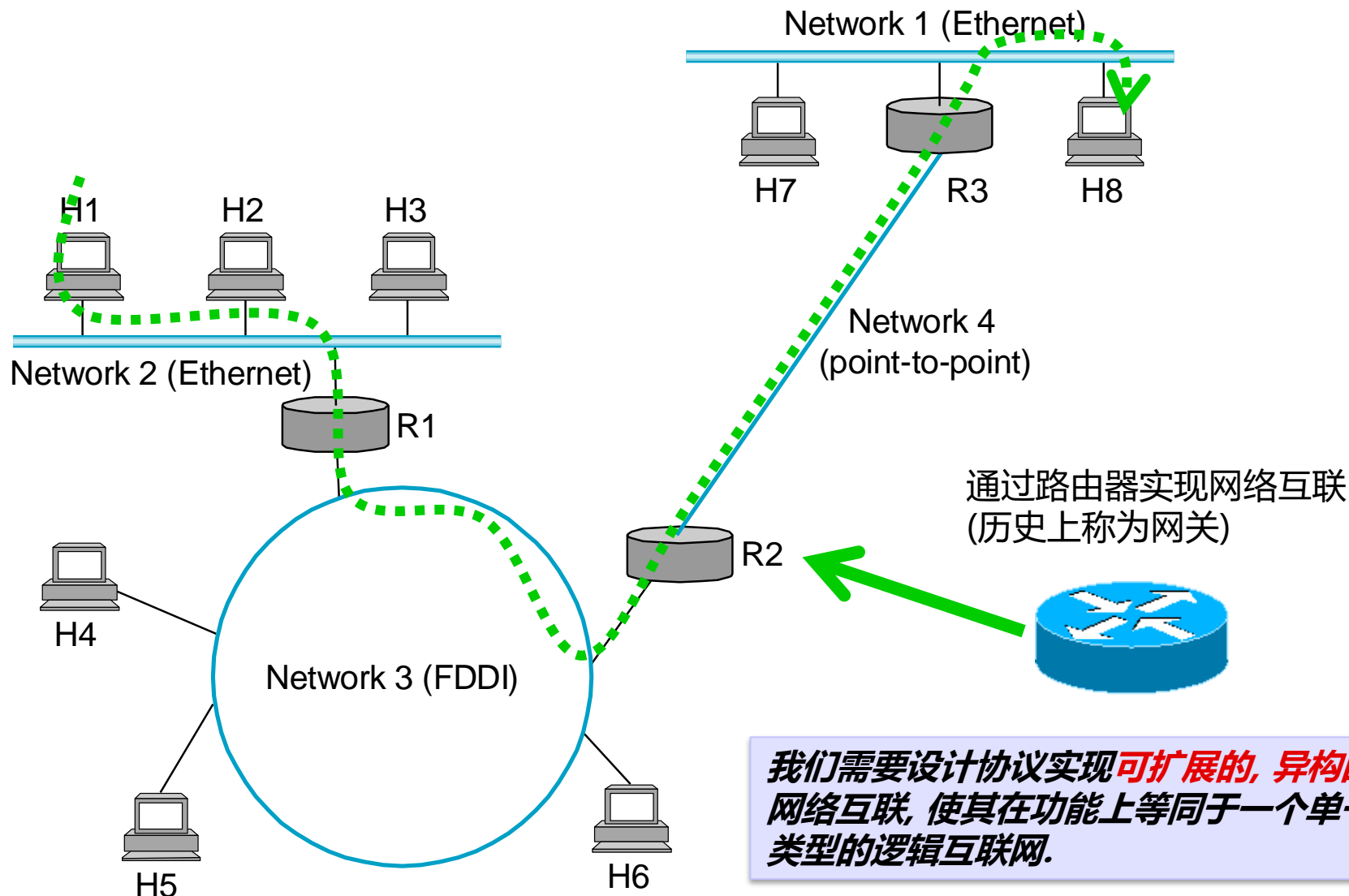


## □ 解决方案

➤ 网络互联



# 网络互联: 示例



## □ 问题: 网络如何互联

- 网络互联与服务模型
- IP协议
- IP地址
  - 标准分类IP地址
  - 子网划分
  - 无类别域间路由CIDR
  - 网络地址转换NAT
- IP附属协议
  - 网络控制管理协议ICMP
  - 地址解析协议ARP
- 分组交付与路由选择

## □ 问题: 网络如何互联

- 网络互联与服务模型
- IP协议
- IP地址
  - 标准分类IP地址
  - 子网划分
  - 无类别域间路由CIDR
  - 网络地址转换NAT
- IP附属协议
  - 网络控制管理协议ICMP
  - 地址解析协议ARP
- 分组交付与路由选择

# 什么是互联网?

## □ “单一”网络

- 采用同一种网络技术, 例如点到点链路, 共享介质网络以及链路层交换机组建的网络

## □ 互联网

- “单一”网络的互联集合
- 视为“网络的网络”: 由许多较小的网络构成
- 范例: Internet, 当今应用最为广泛的全球性互联网

## □ 另一种观点

- 互联的底层网络视为物理网络
- 互联网是由物理网络集合构成的逻辑网络

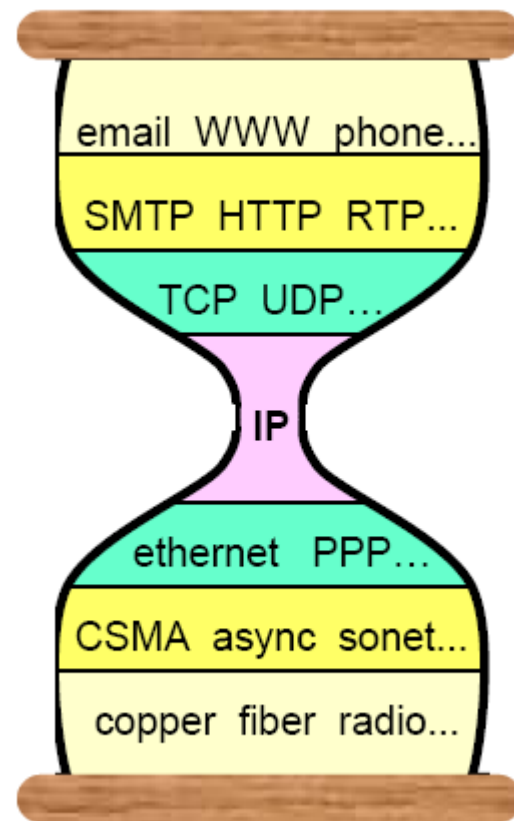
# 服务模型

## □ Internet 协议(IP)

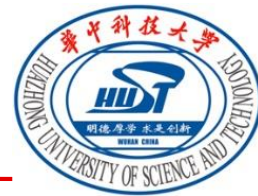
- 连接不同类型网络的网络层协议
- 最成功的网络互联协议
  - 并非唯一的网络互联协议

## □ 设计理念: 提供最基本的服务(最小服务集)

- 可以支持任意类型的网络技术(现行的, 未来兴起的)







# IP服务模型: 数据报传送

## □ 数据报传送

- 无连接方式
- 每一个分组携带转发所需的信息

## □ 尽最大努力交付: 不可靠服务

- 尽最大努力进行数据报传送, 但不对传送过程进行任何保证
- “不保证”不仅仅意味着分组丢失
  - 分组集合可能经过不同的路径乱序到达
  - 同一个分组可能多次传送
- 运行于IP之上的高层协议或应用必须能够意识到这些可能的错误模式

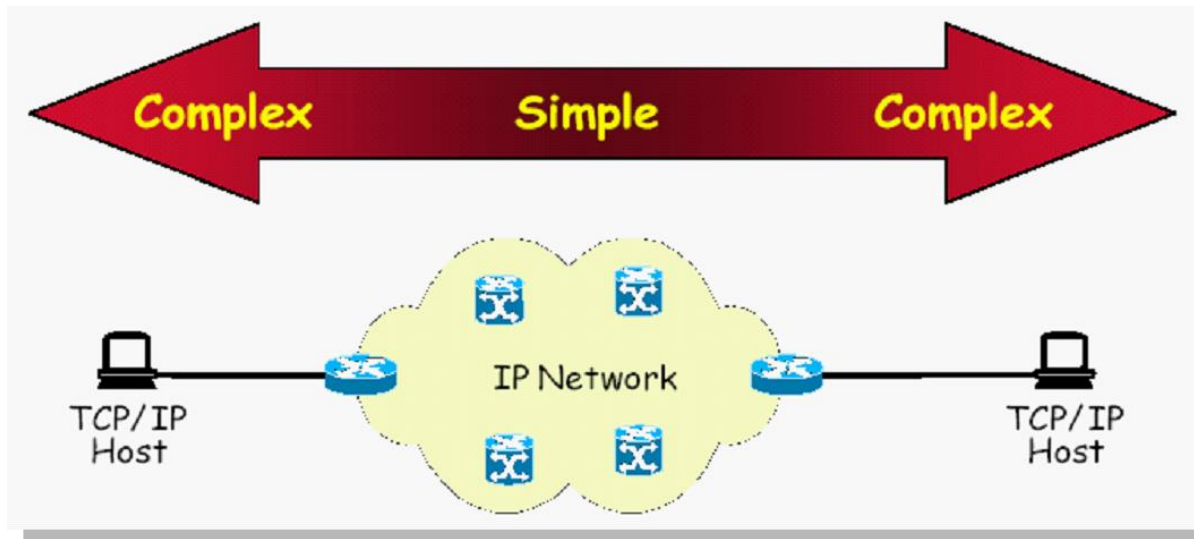
# IP服务模型: 数据报传送

## □ 优点

- IP可以“在任何技术上运行”
- 现今IP赖以运行的很多网络技术在IP被发明时还不存在

## □ 隐含的选择

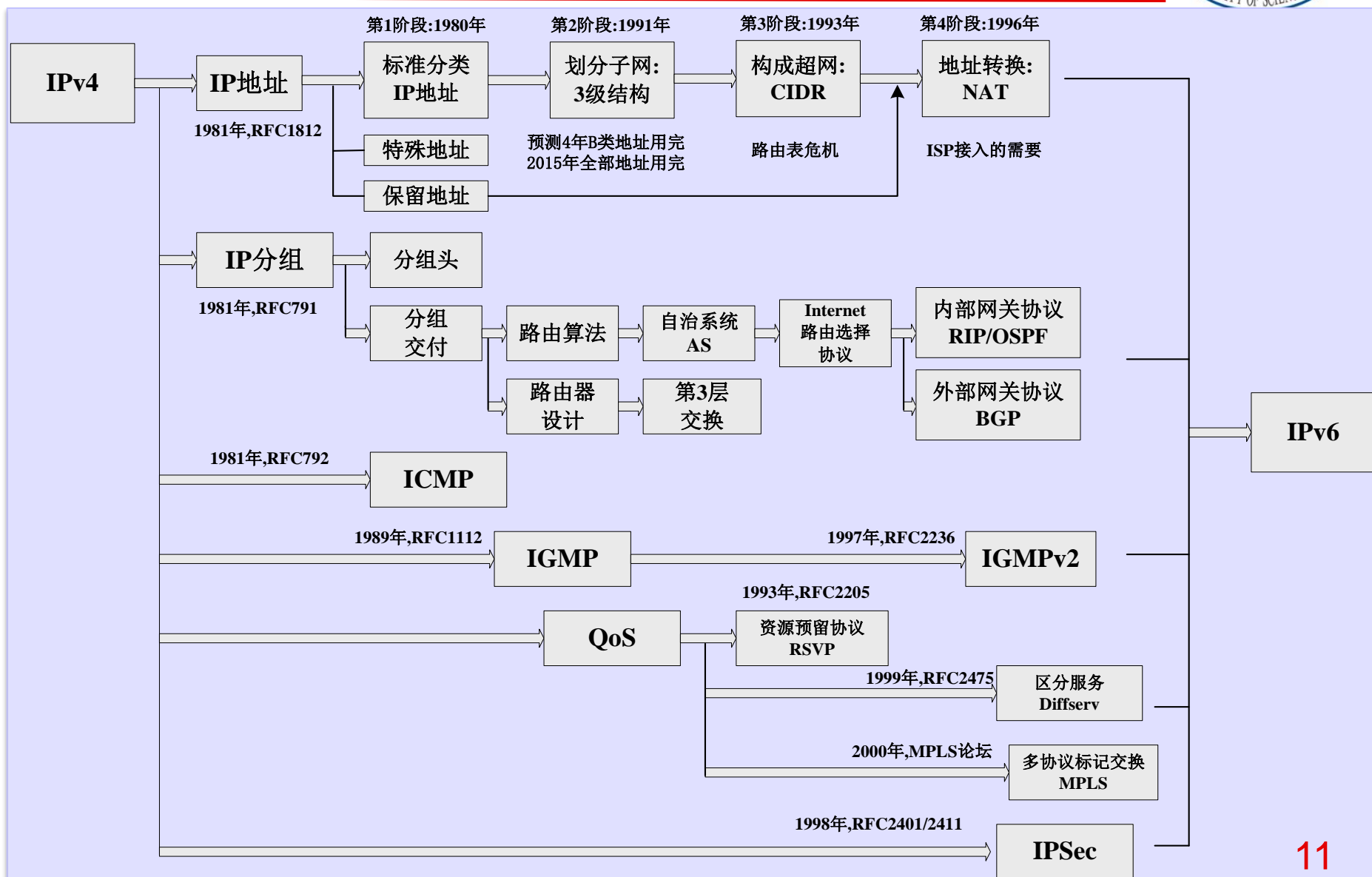
- 将复杂的事情(服务保证)交给主机完成
- 扩展: 简单的核心和复杂的边缘



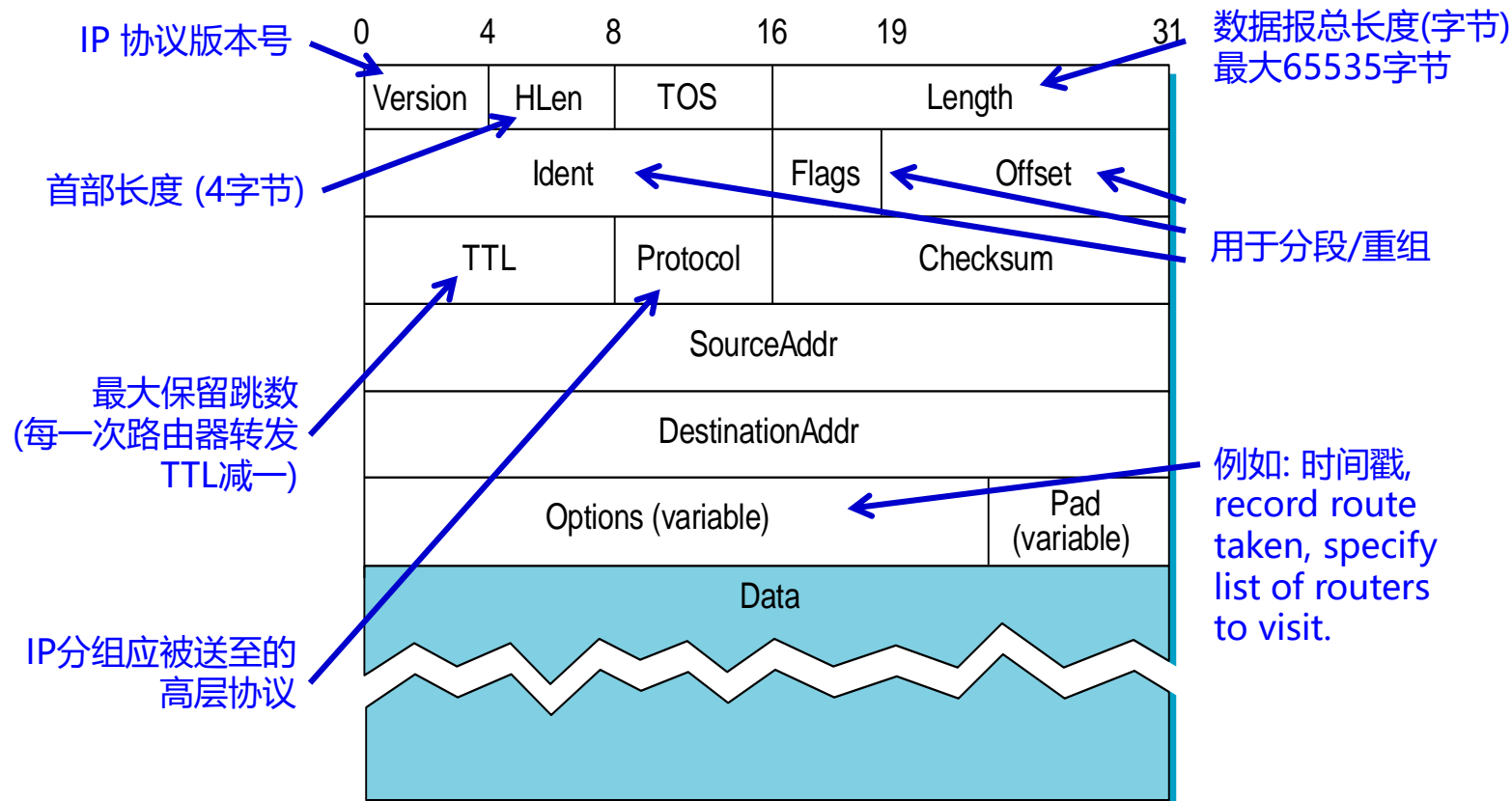
## □ 问题: 网络如何互联

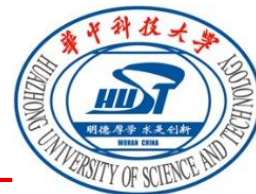
- 网络互联与服务模型
- IP协议
- IP地址
  - 标准分类IP地址
  - 子网划分
  - 无类别域间路由CIDR
  - 网络地址转换NAT
- IP附属协议
  - 网络控制管理协议ICMP
  - 地址解析协议ARP
- 分组交付与路由选择

# IP协议的演变



# IP协议: 分组格式





# IP协议: 分组格式

## □ IP首部的不同字段

### ➤ 自身的描述

- 版本(4b), 首部长度(4b), 分组长度(16b), 校验和(16b)

### ➤ 寻址

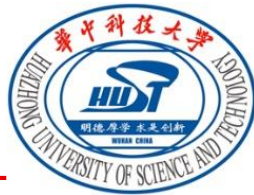
- 源地址(32b), 目的地址(32b)

### ➤ 底层的传送能力

- 标识(16b), 标志(3b), 偏移量(5b)

### ➤ 向上层提供的服务

- 服务类型(8b), 生存期(8b), 协议(8b)



# IP协议: 分段与重组

## □ 最大传输单元 (MTU)

- 网络中一帧能够承载的最大的分组
- 示例: 以太网 (1500 字节), PPP (532 字节)

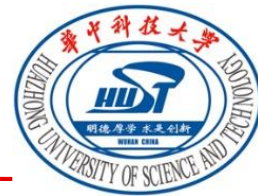
## □ 原因

- 不同的网络存在不同的MTUs

## □ 所有不同类型的网络为什么不采用最小的MTUs?

### ➤ 问题

- 新的网络技术可能出现
- 效率低: 分组首部越长则协议开销越大, 且会产生更多的待处理分组



# IP协议: 分段与重组

## □ 策略

- 主机按照与之直接相连的网络的MTU发送分组
- 需要进行分组分段( $MTU < \text{分组长度}$ )
- 分段产生独立的数据报

## □ 分段重组由目的主机完成

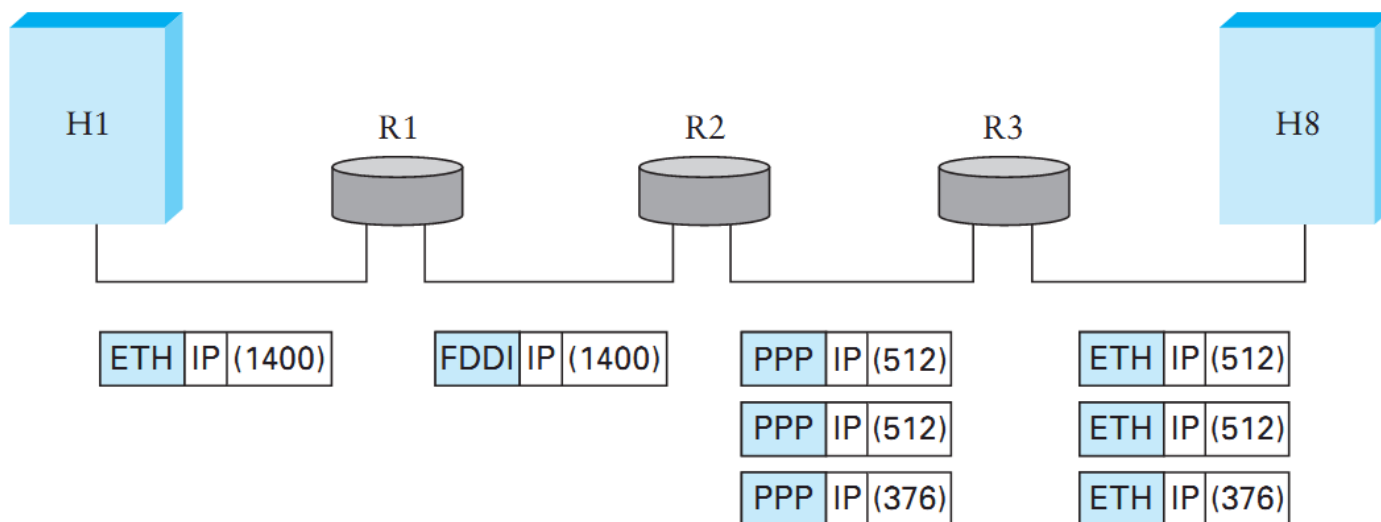
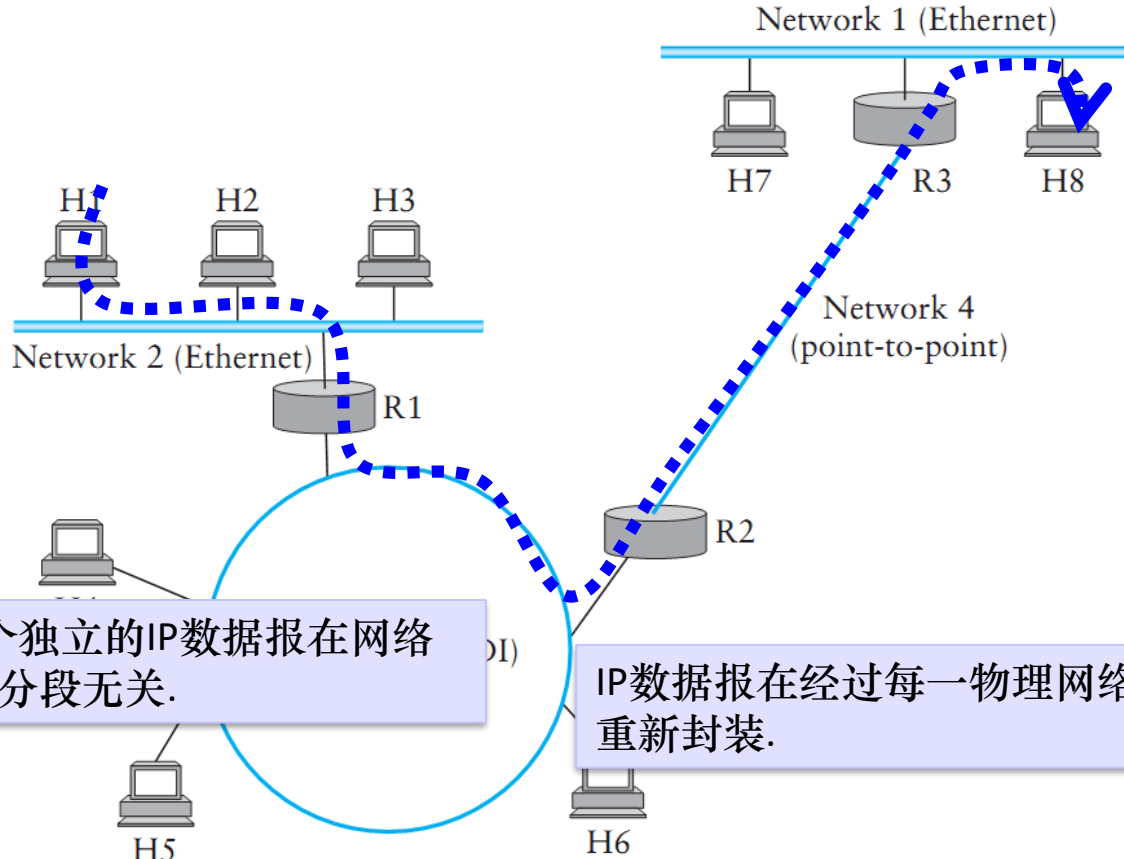
- 分段丢失无法恢复
- 迭代分段可能发生

## □ 问题

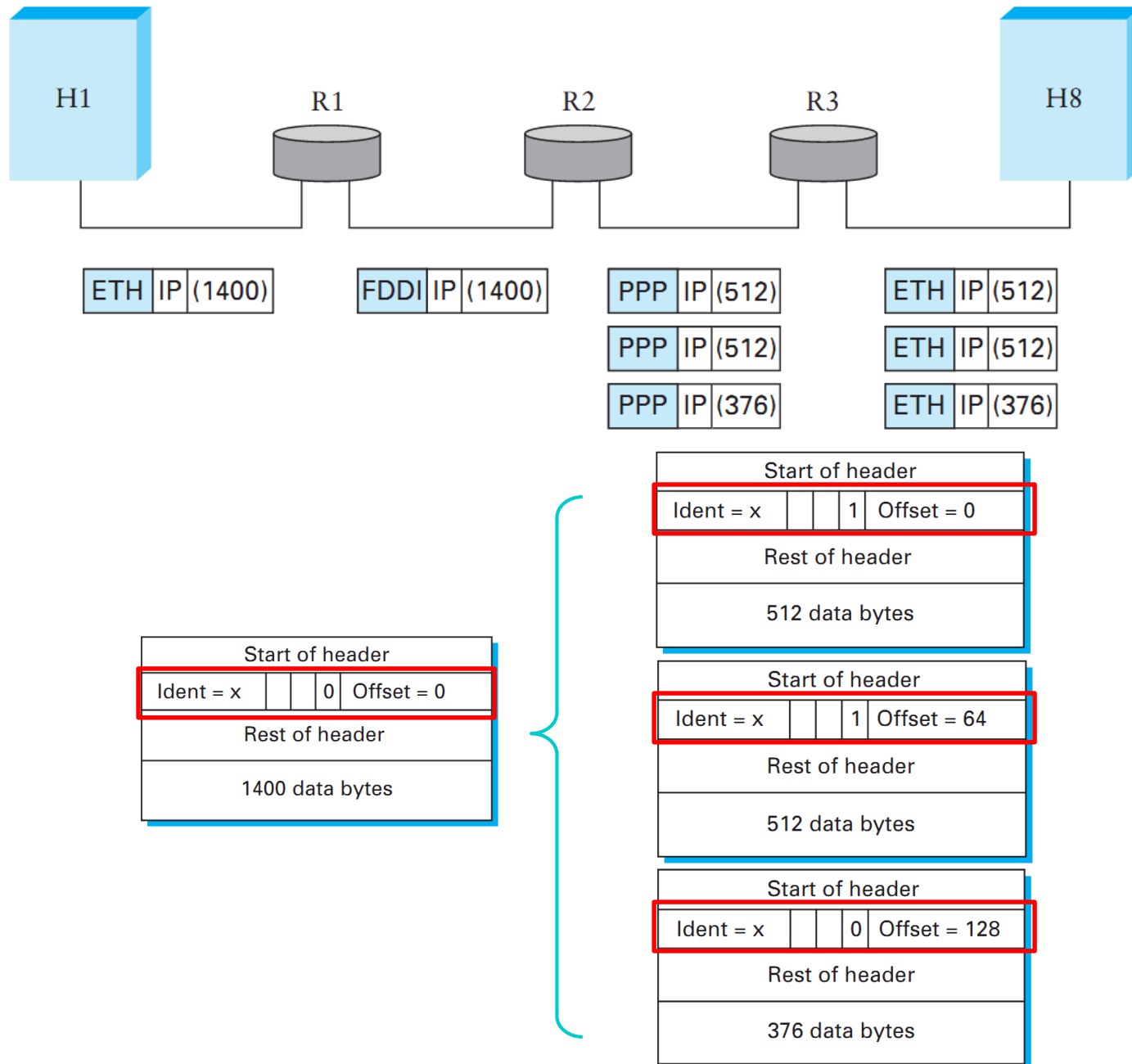
- 如果某一个分段丢失, 目的主机将放弃重组进程并丢弃已收到的分段
  - 尽量避免分段, 建议主机执行“路径MTU发现”



# 示例:

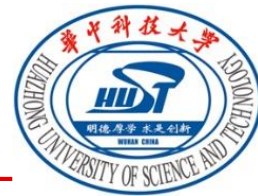


# 示例:



## □ 问题: 网络如何互联

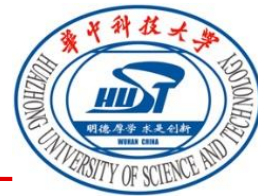
- 网络互联与服务模型
- IP协议
- IP地址
  - 标准分类IP地址
  - 子网划分
  - 无类别域间路由CIDR
  - 网络地址转换NAT
- IP附属协议
  - 网络控制管理协议ICMP
  - 地址解析协议ARP
- 分组交付与路由选择



# IP地址分配

---

- ❑ 连接到Internet的每一个主机（计算机或路由器）至少有一个32位的IP地址
- ❑ IP地址是分配给网络接口的
- ❑ 多归属主机可以有多个IP地址
- ❑ 一个网络接口也可以分配多个IP地址
  - 网桥、Ethernet交换机属于数据链路层设备，使用MAC地址，不属于网络层设备，不分配IP地址。



# IP地址的表示

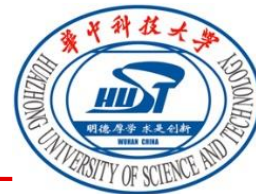
## □ IP地址采用点分十进制方式表示

- IP 地址记为四个十进制整数, 用点分离
- 其中每一个整数表示为地址中1个字节对应的十进制数值
- 例如, 本机的IP地址为 171.69.210.245

## □ 另一种节点命名方式: Internet 域名

- 域名标识为由“.”隔开的ASCII字符串, 例如 *www.edu.cn*, *www.hust.edu.cn*, *itec.hust.edu.cn*
- 应用层的DNS (域名解析服务) 完成域名和IP地址之间的映射转换

# 点分十进制法



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

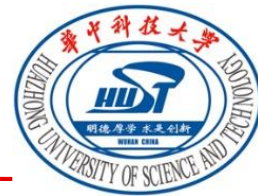
# 寻址设计基础

## □ Flat 扁平化编址方案

- 例如每个主机有 48-bit MAC 地址
- 路由器需要记录全球每个主机的每条地址
  - 路由表太大了
  - 路由表太难以维护了(主机加入、离开时)

## □ Hierarchy 层次化编址方案

- 例如邮政编码 430074, 430072
- 地址分成不同的有特定区分含义的片段
- 可以实现从大到小的不同区域的路由交付
- 如果主机接入网络发生变化, 只需要局部的更新



# 全局地址

## □ 网络互联的寻址问题

- 首先, 每一个主机拥有唯一的地址
- 其次, 在一个大规模网络中需要跟位置和拓扑相关联的层次化IP地址

## □ IP的解决方案

- 每一个地址**全局唯一**
- 层次化寻址
  - 两级地址空间 {**网络号:主机号**}

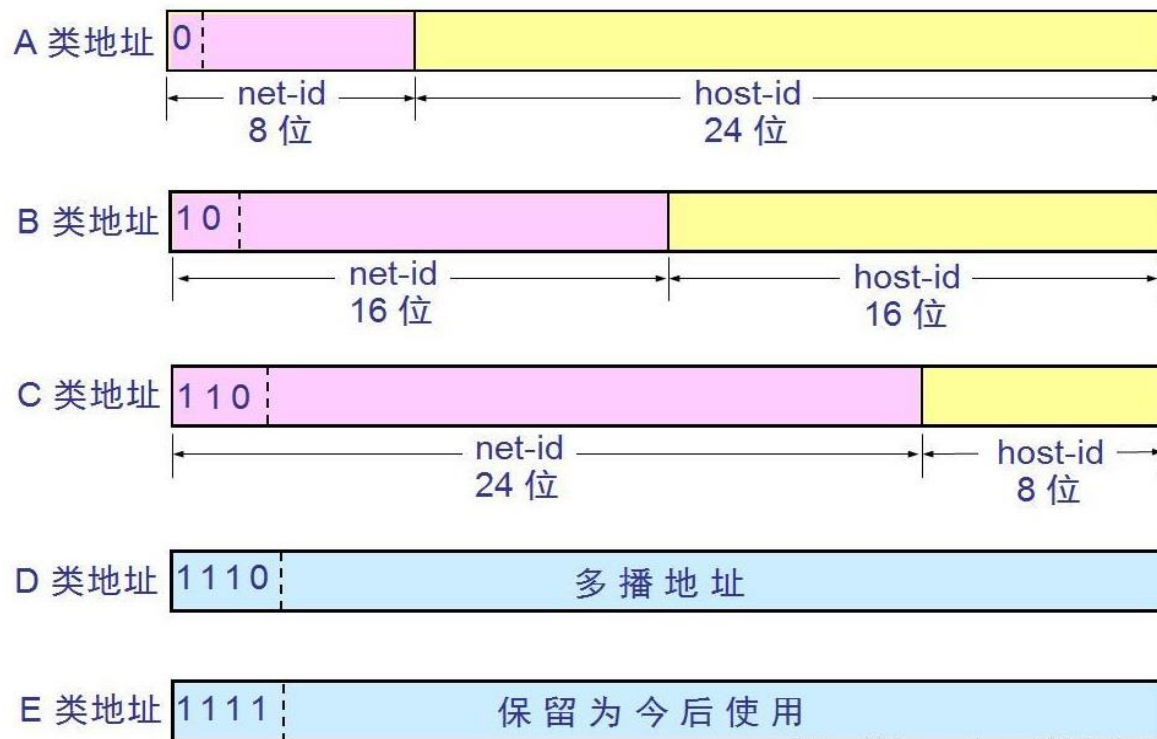


# 层次化地址空间

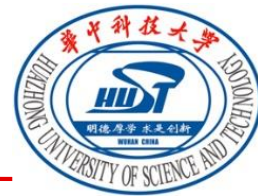
□ 每一个IP地址包括2个部分

➤ 网络号+主机号

□ IP地址分类



网络号范围	主机数量
1~127	16,777,214
128~191	65,534
192~223	254



# IP地址如何获取

## □ 公有IP地址要求全球唯一

- ICANN (Internet Corporation for Assigned Names and Numbers) 即互联网名字与编号分配机构向ISP分配
- ISP再向所属机构或组织逐级分配

## □ 静态设定

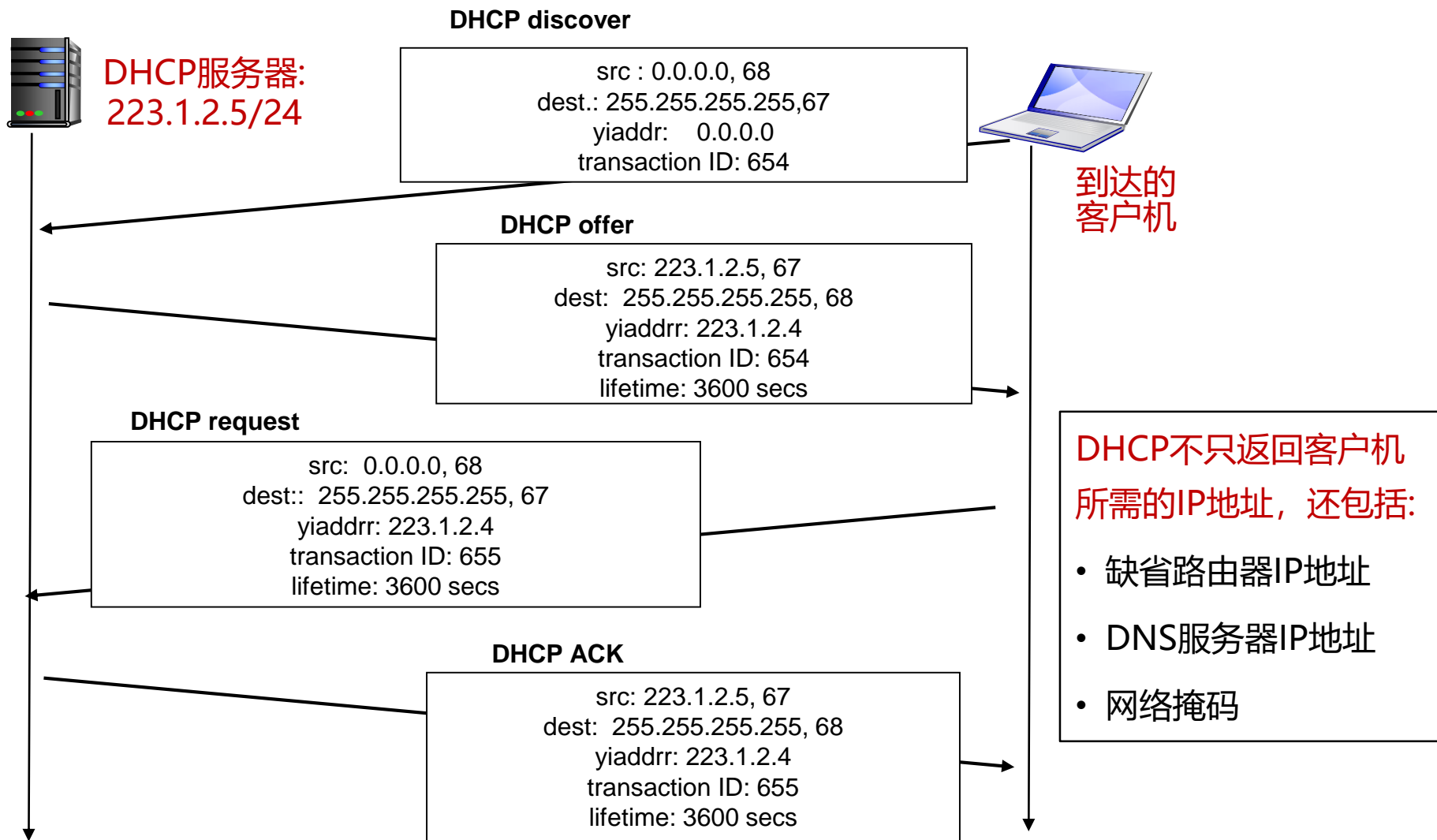
- 申请固定IP地址, 手工设定, 如路由器、服务器

## □ IP地址不仅要求在给定的互联网中唯一, 而且必须能够反映互联网中结点的连接情况

## □ 动态获取

- 使用诸如DHCP动态主机配置协议
- 当主机加入IP网络, 允许主机从DHCP服务器动态获取IP地址
- 可以有效利用IP地址, 方便移动主机的地址获取

# DHCP工作过程



# 分类寻址

## □ IP地址的网络号唯一标识一个网络

- 连接在同一个网络内的主机以及路由器接口的IP地址中包含相同的网络号
- 给定任意IP地址, 路由器可以识别该IP地址所属的网络

## □ 分类寻址

- 假设: Internet包含

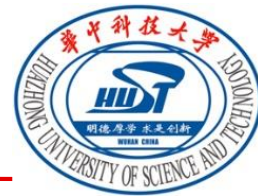
- 少量的广域网 (A类网络)
- 相当数量的站点- (校园-) 规模的网络 (B类网络)
- 大量的LANs (C类网络)

## □ 可能存在一些不公平的且未充分利用的地址分配

- Internet中采用“**无分类寻址**”方式

## □ 问题: 网络如何互联

- 网络互联与服务模型
- IP协议
- IP地址
  - 标准分类IP地址
  - 子网划分
  - 无类别域间路由CIDR
  - 网络地址转换NAT
- IP附属协议
  - 网络控制管理协议ICMP
  - 地址解析协议ARP
- 分组交付与路由选择



# 子网划分

---

## □ 问题

- IP 地址耗尽问题

## □ 动机

- 引入新的层次

## □ 解决方案

- 寻址: 子网掩码
- IP 转发

# IP 地址耗尽问题

## □ IPv4 包含32bit的地址空间

- 总共包含40亿个合法地址
- 并非所有的地址均用于主机或接口标识
- 部分地址用于多播或保留使用

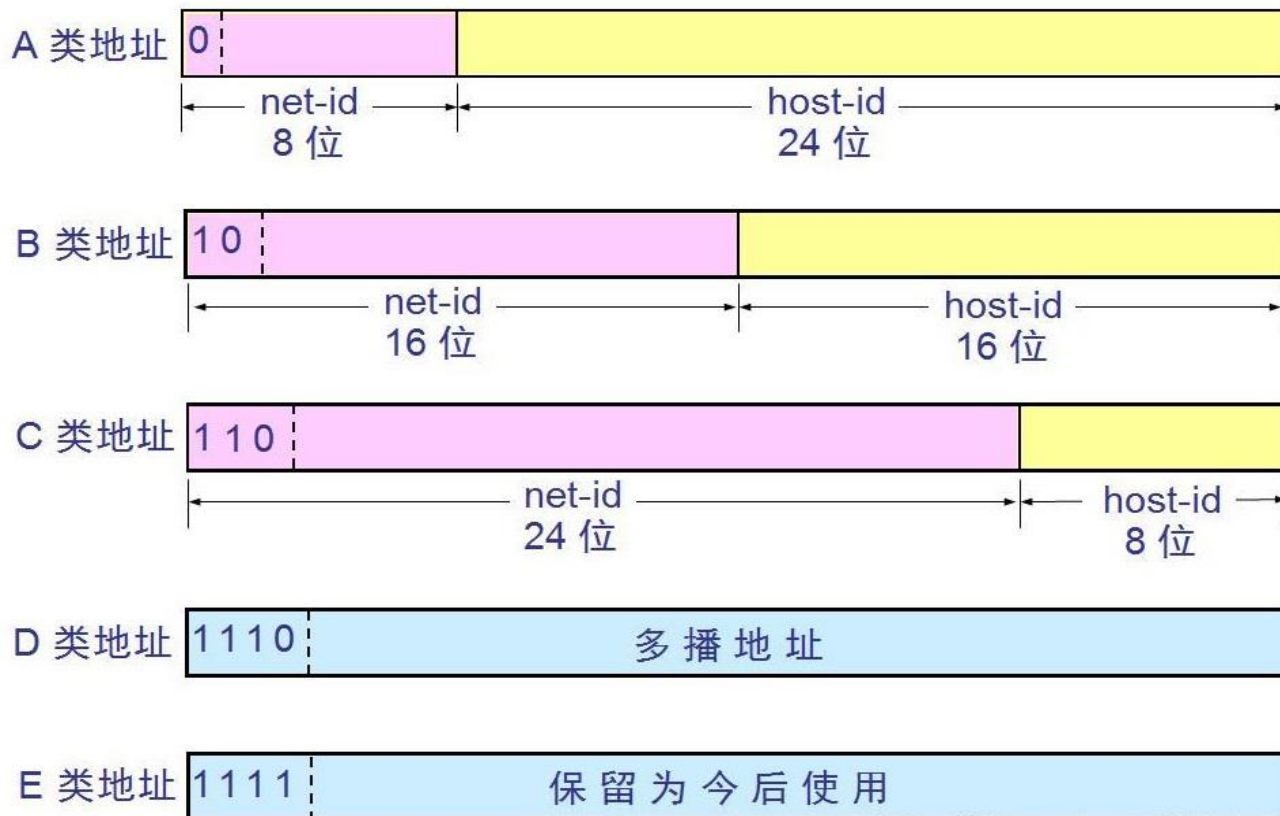
## □ 为什么IP地址会耗尽?

- Internet的快速发展
  - 特别是物联网的快速发展，万物互连的需求
- 低效的地址分配

# IP地址回顾

## □ IP 地址最初采取分类方式管理

➤ 网络被分为A, B, C三类





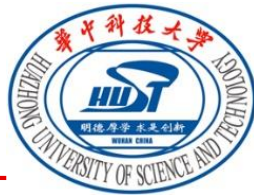
# 低效的地址分配

## □ 示例1: 一个包含10台主机的网络

- 需要一个C类网络号
- 效率:  $10/254$  约为 4%

## □ 示例2: 一个拥有300台主机的网络

- 需要一个B类网络号
- 效率:  $300/65534$  约为 0.5%



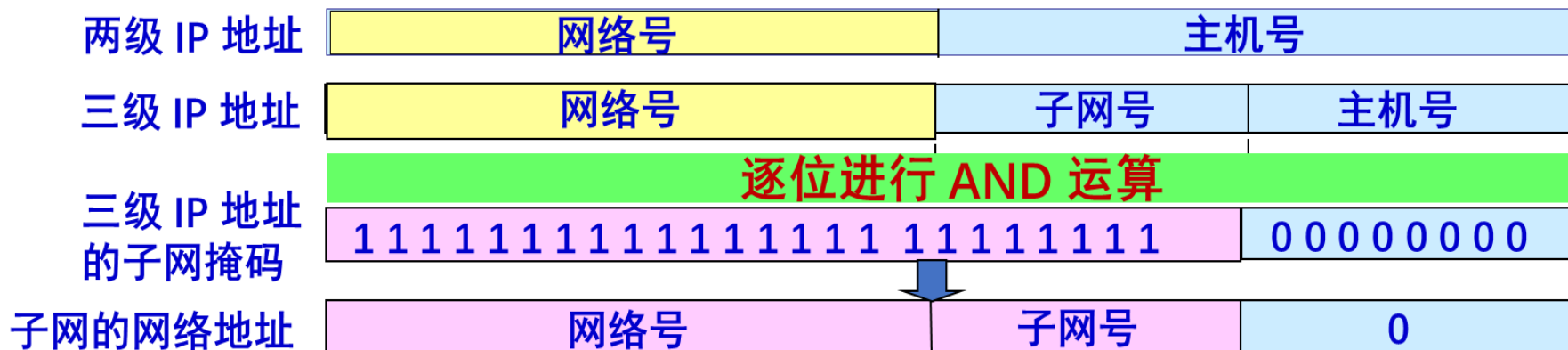
# 子网划分的动机

□ 子网在IP地址中引入新的层次

□ 子网划分将拥有同一个网络号的IP地址分配给若干个物理网络 (子网)

- 子网彼此离得很近，拥有同一个网络号
- 路由器只能选择一条路由到达任何子网
- 子网可以减少对网络号的需求
  
- 划分子网纯属一个单位内部的操作
- 单位对外仍然表现为没有划分子网的网络

# 子网划分

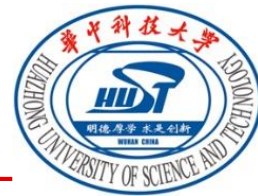


## □ 子网掩码(subnet mask)

- 从主机号部分借用若干个位作为子网号
- 与 IP 地址对应, 是32 bit 的二进制数, 置1表示网络位, 置0表示主机位

## □ 注意

- 主机号部分为全0表示这个网络
- 主机号部分为全1表示该网络中的所有主机, 即广播地址



# IP寻址示例

---

- ❑ 网络地址: 主机号部分全部为0
- ❑ 广播地址: 主机号部分全部为1
- ❑ 如果一台主机的IP地址为202.112.14.137, 子网掩码为255.255.255.224, 则网络地址和广播地址计算如下:
  - $137 = 1000\ 1001$
  - $224 = 1110\ 0000$
  - 网络地址的最后8bit为 $1000\ 0000 = 128$
  - 网络地址为202.112.14.128
  - 广播地址的最后8bit为 $1001\ 1111 = 159$
  - 广播地址为202.112.14.159

# IP寻址示例

- 在子网中, 网络地址 (全0) 和广播地址 (全1) 占用了两个额外的地址
  - 在实际的以太网中, 选择其他地址作为主机或网关的IP地址
  
- 对于包含10台主机的子网, 计算其子网掩码:
  - 所有地址  $10 + 1 + 1 = 12$ ,  $2^3 < 12 < 2^4$
  - 因此主机号部分占有4bits
  - 掩码的最后8bit为  $1111\ 0000 = 240$
  - 子网掩码为255.255.255.240

# 子网划分的效率

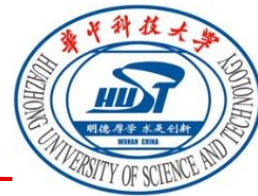
□ 示例: 3 个网络, 其中2个包含50台主机, 另1个包含100台主机

➤ 传统分配

- 需要3个C类网络号
- 效率:  $200/(3 \times 254)$  约为 26%

➤ 子网划分

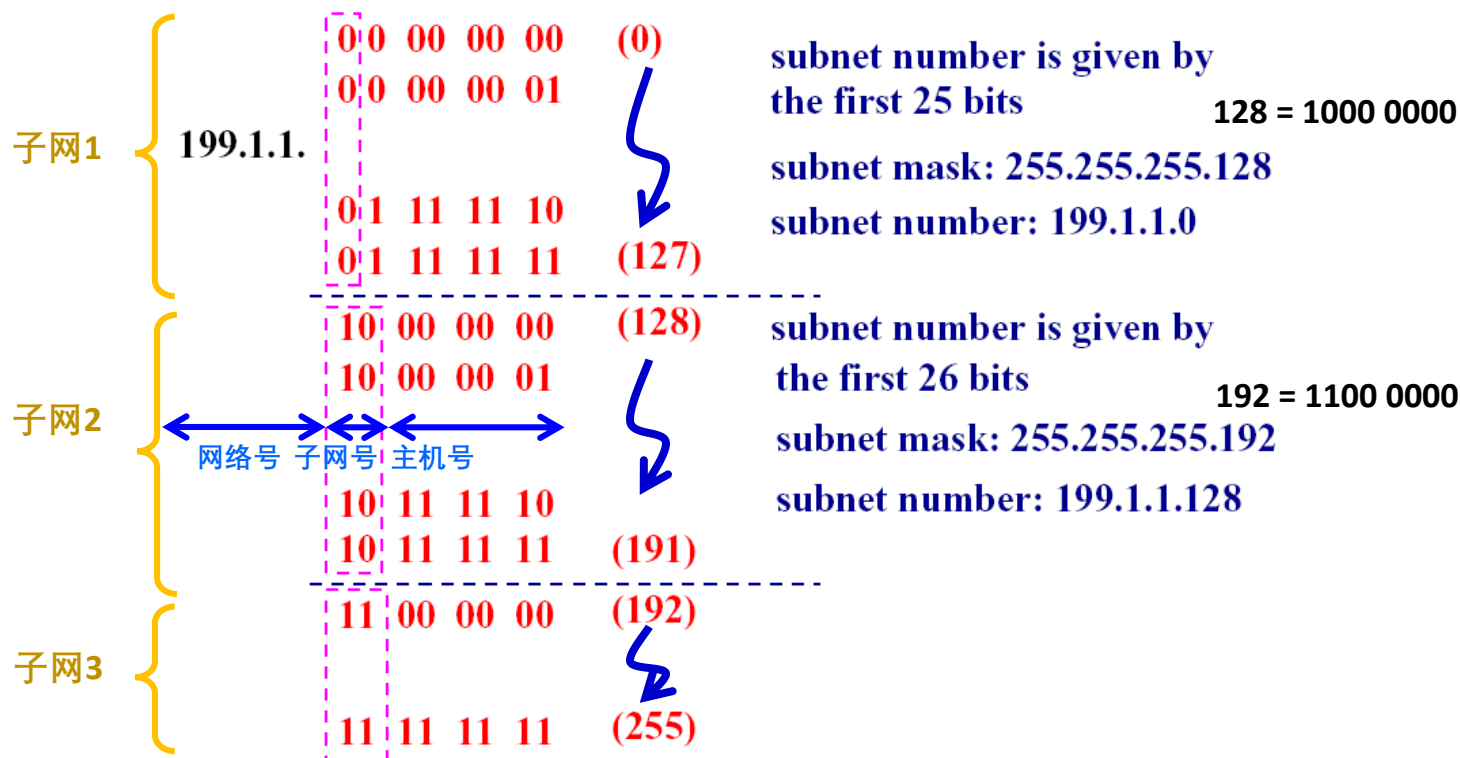
- 仅需要1个C类网络号
  - 划分3个子网: 其中2个分配64的地址, 另1个分配128个地址
- 效率:  $200/254$  约为 78%



# 如何划分子网

- 示例: 给定一个C类网络号199.1.1.0, 划分为3个子网, 其中2个分配64个地址, 另1个128个地址
  - 对于包含128个地址的子网, 地址中的主机号部分占有7bits
    - 子网掩码为255.255.255.128
  - 对于包含64个地址的子网, 地址中的主机号部分占有6bits
    - 子网掩码为255.255.255.192

# 如何划分子网

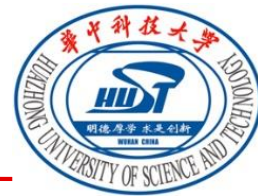


子网号	子网掩码	IP地址
192.1.1.0	255.255.255.128	192.1.1.0 – 192.1.1.127
192.1.1.128	255.255.255.192	192.1.1.128 – 192.1.1.191
192.1.1.192	255.255.255.192	192.1.1.192 – 192.1.1.255



# 如何划分子网

199.1.1.	00	00	00	00	(0)	subnet mask: 255.255.255.192 subnet number: 199.1.1.128
	00	11	11	11	(63)	
	01	00	00	00	(64)	<del>subnet mask: 255.255.255.0 subnet number: 199.1.1.0</del>
	01	00	00	01	(65)	
	10	11	11	10	(190)	
	10	11	11	11	(191)	
	11	00	00	00	(192)	优先划分较大的子网 !!
	11	11	11	11	(255)	



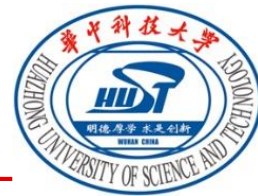
# 小结: 子网划分的优点

---

- 子网划分通过两种方式解决了网络扩展问题
- 第一, 提高了地址分配效率
  - 加入一个新的物理网络时不需要用光B类或C类网络的所有地址
- 第二, 有利于信息聚合
  - 可以减少路由器为转发数据报到这些网络所需存储的信息量

## □ 问题: 网络如何互联

- 网络互联与服务模型
- IP协议
- IP地址
  - 标准分类IP地址
  - 子网划分
  - 无类别域间路由CIDR
  - 网络地址转换NAT
- IP附属协议
  - 网络控制管理协议ICMP
  - 地址解析协议ARP
- 分组交付与路由选择



# 无分类路由选择(CIDR)

---

## □ 问题

- IP地址固定结构的低效性

## □ 动机

- IP地址不采用固定的分类结构

## □ 解决方案

- CIDR
- 汇聚路由

# 固定分类地址的问题

## □ 示例: 某ISP组建一个包含1000台主机的网络

### ➤ 如果采用分类地址结构

- 需要一个B类网络地址
- 效率:  $1000/65536$  约为0.15%

## □ 实际问题

- 网络大小的粒度划分非常粗略
- 网络大小只能为256, 65k, 17M

## □ 即使采用子网划分技术也无法避免该问题

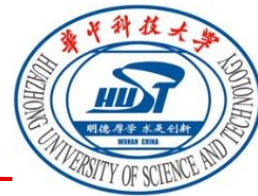
# 固定分类地址的问题

## □ 一种可能的解决方案

- 以256个地址(C类网络地址)为单位分配空间
- 然而, 这个方案会引起一个严重的问题: 对路由器超量存储的需求

## □ 需要在两者之间达到平衡

- 减少路由器的路由记录数 vs. 高效分配地址
- 汇聚路由(Route Aggregation)
  - 仅使用转发表中的一条记录来知道如何到达多个不同的网络(子网)



# 无分类路由的动机

---

## □ 寻址

- 更加合理的进行网路大小的粒度划分
- 网络可以为合适的大小: 256, 512, 1024, 2048, ...

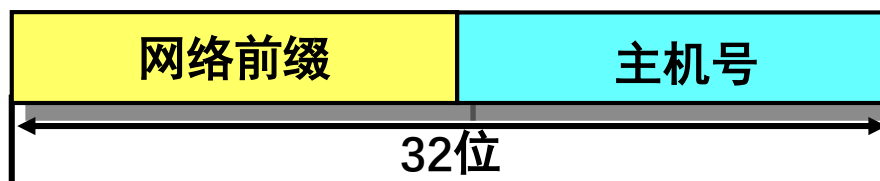
## □ 路由选择

- 分发多个具有相同前缀的连续C类地址块
- 地址块包含数目为2的幂次方的C类网络

# 解决方案

## □ CIDR: Classless InterDomain Routing 无分类域间路由选择

- IETF于1993年提出的用于替换分类IP地址的方案
- 基于变长的子网掩码允许分配任意长度前缀的子网，将地址块分为256, 512, 1024, 2048, ...
- 将32位的IP地址划分为两个部分，并采用斜线记法，即在IP地址后加上“/”，再写上网络前缀所占位数



IP地址 ::= {<网络前缀>, <主机号>}



# 示例



## □ 示例: 某单位组建一个包含3000台主机的网络

- 成功获得了16个连续的C类网络地址块
  - 假设C类地址网络号为 199.199.0-199.199.15
  - 这些C类地址网络号部分的高20bit是相同的
- 基于CIDR, 可以认为这是一个网络号占用20位的网络, 记为199.199.0.0/20
  - 这种标识表示网络号为199.199.0.0的高20位
- 思考: 如果直接给该单位分配16个C类地址块?
  - 路由表更大 (每一个C类网络占一项)
  - 路由开销更大 (主机通信因跨网络而使用路由器)

# CIDR和转发表

## □ 转发表

网络/掩码长度	下一跳
199.199.199.0/22	Router A
199.199.192.0/17	Router B
200.200.16.0/20	Router C
200.200.21.0/24	Router D

## □ 当路由器收到一个分组

➤ 查询其目的IP地址

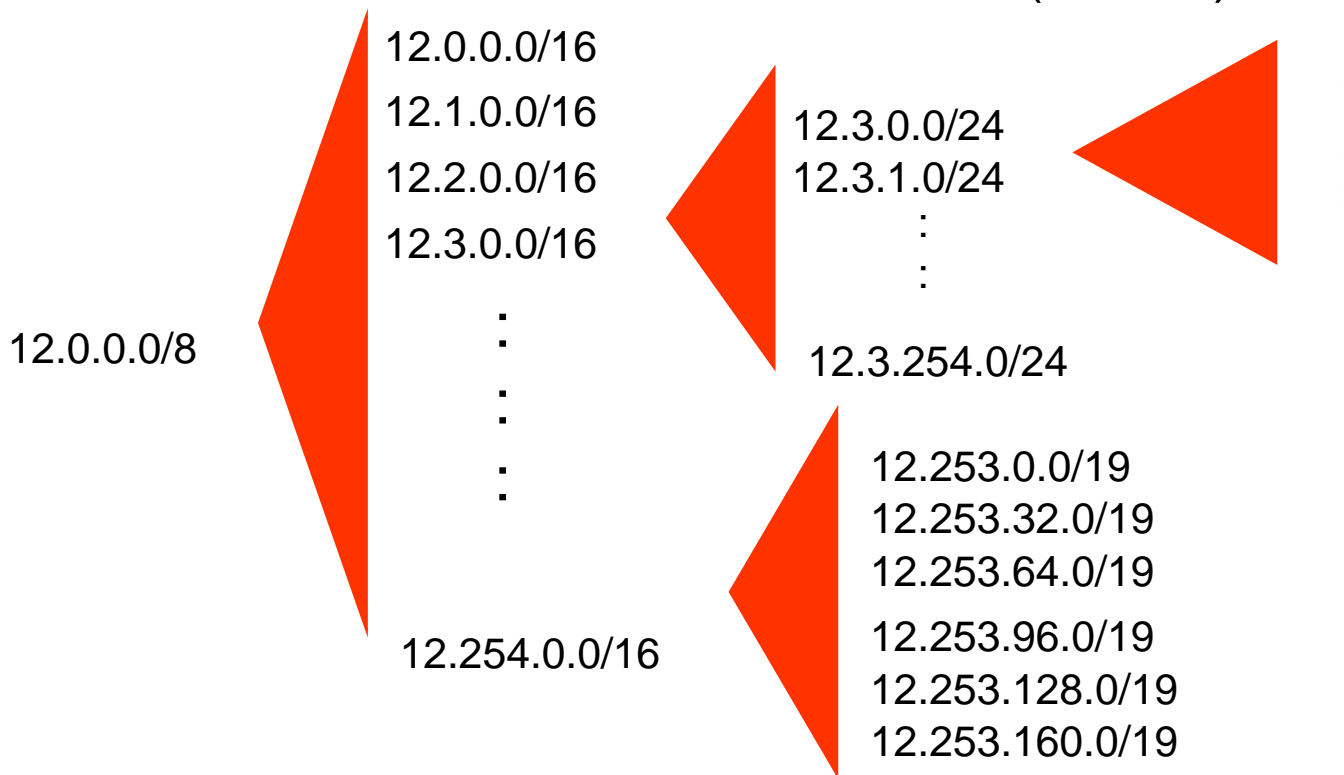
➤ 遍历转发表

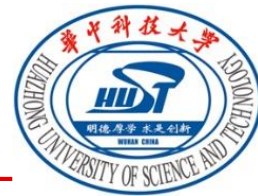
● 匹配: 掩码长度为 $n$ , 目的IP地址的高 $n$ 位 = 网络号的高 $n$ 位

# CIDR: 灵活的层次化地址分配

## □ 采用前缀进行Internet扩展

- 对连续的块进行地址分配 (前缀聚合, 形成超网)
- 基于前缀进行路由选择和分组转发
- 至今, 路由表包含约200,000 个前缀(vs. 4B)





# 最长前缀匹配

---

□ 示例: 目的IP地址为200.200.21.1

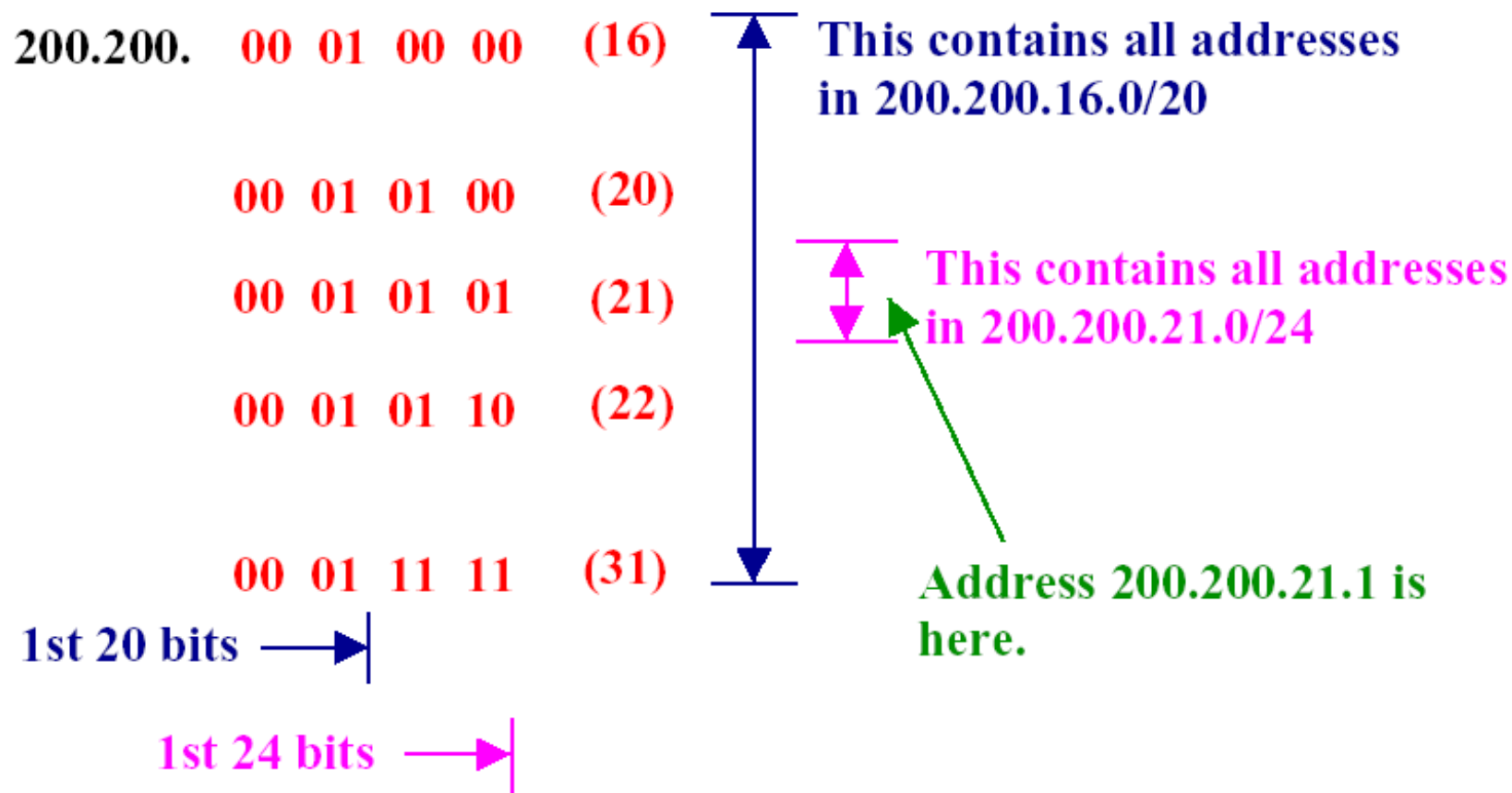
- 倒数第2条记录: 200.200.16.0/20
- 最后一条记录: 200.200.21.0/24
- 均与目的IP地址相匹配

□ 原则: 分组与最长的前缀匹配

- 本例选择结果: 200.200.21.0/24

# 最长前缀匹配: 图示说明

We ignore the last 8 bits



# 小结

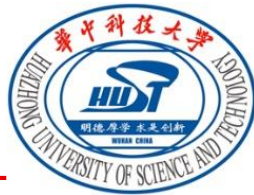
---

## □ IP的扩展

- 问题 1: 低效的地址空间使用
- 问题 2: 路由表和转发表的记录数量

## □ 两种改进方法

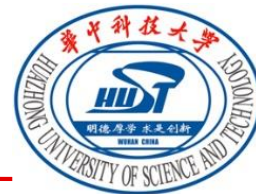
- 子网划分: 仅能解决问题1
- CIDR: 同时有效解决问题1 和 2



# 练习

- ❑ 某部门申请到一个C类IP地址，若要分成8个相同规模的子网，其掩码应为（ 255.255.255.224 ）
- ❑ 如果用户网络需要划分5个子网，每个子网最多20台主机，则适用的子网掩码是（ 255.255.255.224 ）
- ❑ 一个B类网络172.16.0.0，使用子网掩码255.255.255.192来划分子网，则理论上可以利用的网络数和每个网络中的主机数分别为（ 1024、62 ）
- ❑ IP地址为192.168.1.\*的网络中，网络掩码为255.255.255.0，则其网络地址为（ 192.168.1.0 ），广播地址为（ 192.168.1.255 ），主机可用IP地址范围为（ 192.168.1.1 — 192.168.1.254 ）
- ❑ 与10.110.12.29（掩码为255.255.255.224）属于同一网段的主机IP地址是（ B ）
  - A. 10.110.12.0
  - B. 10.110.12.30
  - C. 10.110.12.31
  - D. 10.110.12.32

# 练习

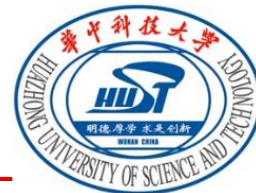


- 某公司包括A、B、C、D、E五个部门，分别有157、108、59、25、16台主机。该公司现有IP地址块10.10.10.0/23，需要划分用于五个部门网络的编址。给出一种满足上述要求的子网地址划分方案，包括网络号和网络掩码，结果直接填入下列表格中。

部门	网络号	网络掩码
A	10.10.10.0/24	255.255.255.0
B	10.10.11.0/25	255.255.255.128
C	10.10.11.128/26	255.255.255.192
D	10.10.11.192/27	255.255.255.224
E	10.10.11.224/27	255.255.255.224



# 练习



- 网络地址规划：一个单位有一个C类网络199.1.1.\*。考虑到共有四个部门，准备划分子网。这四个部门内的主机数目分别是：A — 40台，B — 85台，C — 31台，D — 17台。要求各部门的IP地址分配保持连续，给出一种划分方案，说明各子网的网络地址、子网掩码和允许的主机IP地址范围。

部门	子网编号	最大可用地址	子网号	子网掩码
A	1	64	199.1.1. <u>00000000</u>	255.255.255. <u>00000000</u>
C	2	64	199.1.1. <u>01000000</u>	255.255.255. <u>01000000</u>
B	3	64	199.1.1. <u>10000000</u>	255.255.255. <u>10000000</u>
	4	32	199.1.1. <u>11000000</u>	255.255.255. <u>11000000</u>
D	5	32	199.1.1. <u>11100000</u>	255.255.255. <u>11100000</u>

# 练习

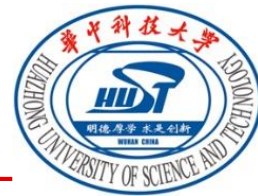


- 网络地址规划：一个单位有一个C类网络199.1.1.\*。考虑到共有四个部门，准备划分子网。这四个部门内的主机数目分别是：A — 40台，B — 85台，C — 31台，D — 17台。要求各部门的IP地址分配保持连续，给出一种划分方案，说明各子网的网络地址、子网掩码和允许的主机IP地址范围。

部门	主机数	网络地址	广播地址	子网掩码	地址范围
A	40	199.1.1.0	199.1.1.63	255.255.255.192	199.1.1.1-199.1.1.62
C	31	199.1.1.64	199.1.1.127	255.255.255.192	199.1.1.65-199.1.1.126
B	85	199.1.1.128	199.1.1.191	255.255.255.192	199.1.1.129-199.1.1.190
		199.1.1.192	199.1.1.223	255.255.255.224	199.1.1.193-199.1.1.222
D	17	199.1.1.224	199.1.1.255	255.255.255.224	199.1.1.225-199.1.1.254

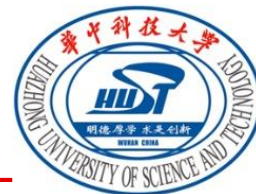
## □ 问题: 网络如何互联

- 网络互联与服务模型
- IP协议
- IP地址
  - 标准分类IP地址
  - 子网划分
  - 无类别域间路由CIDR
  - 网络地址转换NAT
- IP附属协议
  - 网络控制管理协议ICMP
  - 地址解析协议ARP
- 分组交付与路由选择



# 补充：特殊IP地址

IP 地址	说明
0.0.0.0	在缺省网关(路由器)作为所有不清楚的主机和目的网络的代表
255.255.255.255	限制广播地址，指本网段内（同一广播域）的所有主机
127.0.0.1	本机地址，主要用于测试，别名是“Localhost”
224.0.0.1	组播地址，从224.0.0.0到239.255.255.255都是这样的地址
169.254.x.x	DHCP服务器发生故障时，操作系统会自动分配该地址
10.x.x.x	私有地址，被大量用于企业内部网络中。使用私有地址的私有网络在接入Internet时，要使用地址转换(NAT)，将私有地址翻译成公用合法地址。
172.16.x.x ~ 172.31.x.x	
192.168.x.x	



# 网络地址转换的需求

## □ IP地址分配的需求

- 公网IP地址数量有限
- 实际运营中很多内部专用网络都需要IP地址

## □ 私有网络地址的提出

- RFC 1918 为私有网络预留出了三个IP地址块
- 私有网络地址不会在因特网上被分配，因此可以不必向ISP 或注册中心申请而在公司或企业内部自由使用

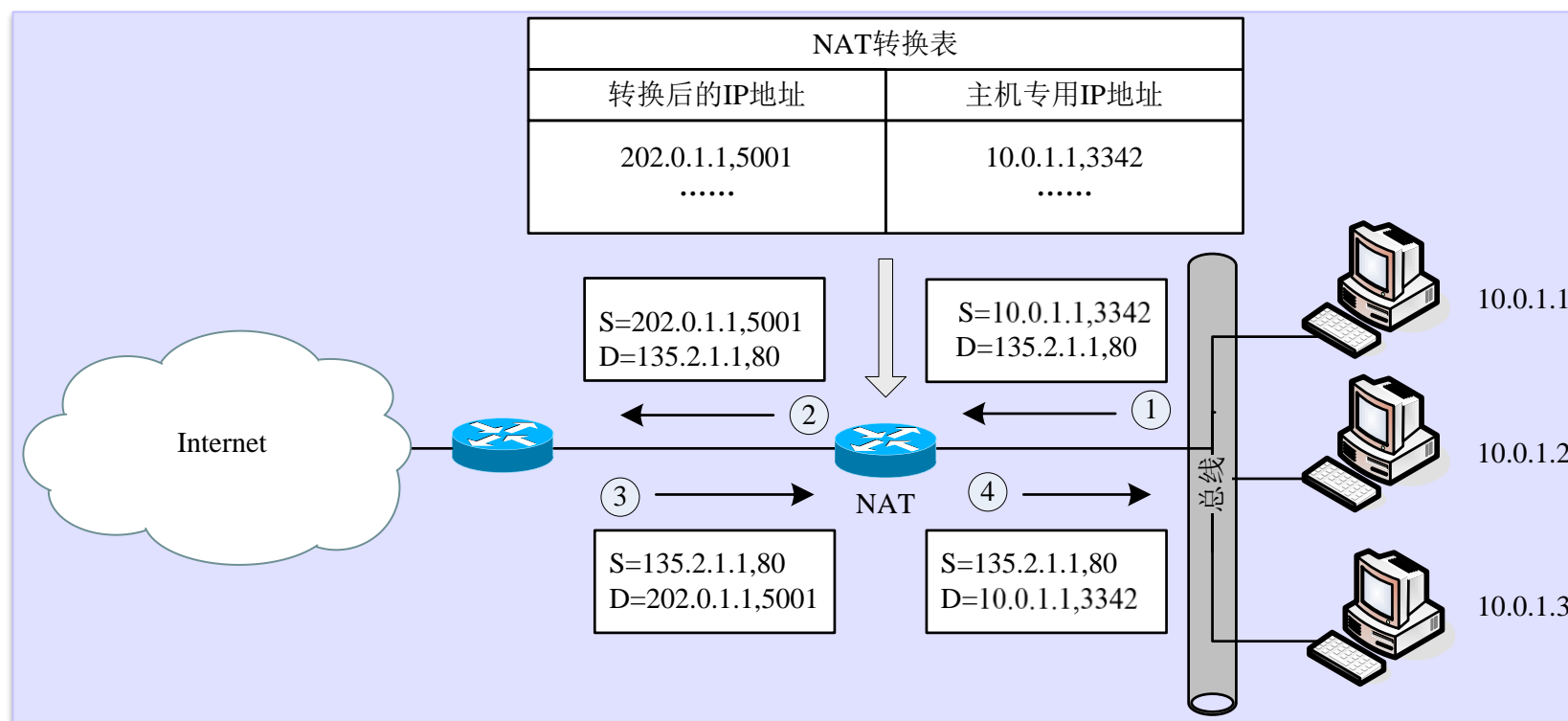
## □ 网络地址转换的需求

- 内部网络主机使用私有地址，通过网关路由器发送数据包时，私有地址被转换成合法公网的IP地址
- 一个局域网只需使用少量IP地址（甚至是1个）即可实现私有地址网络内所有计算机与Internet的通信需求

# 网络地址转换NAT

## □ NAT网络地址转换(Network Address Translation)

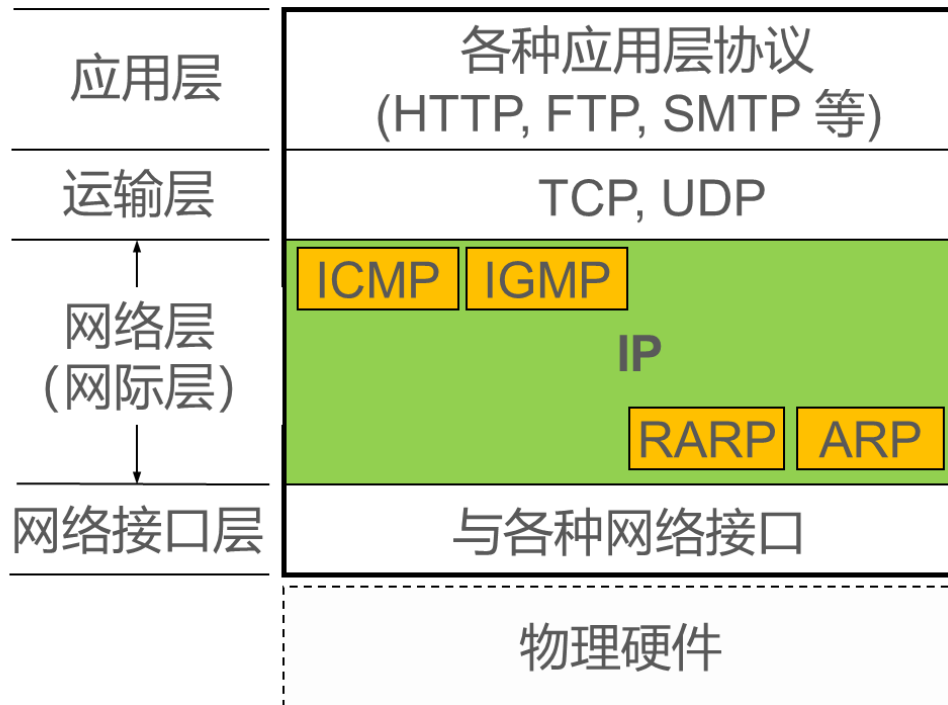
- 1994年RFC 1632, 路由器可将多个内部地址映射为一个合法公网地址, 以不同的协议端口号与不同的内部地址相对应



## □ 问题: 网络如何互联

- 网络互联与服务模型
- IP协议
- IP地址
  - 标准分类IP地址
  - 子网划分
  - 无类别域间路由CIDR
  - 网络地址转换NAT
- IP附属协议
  - 网络控制管理协议ICMP
  - 地址解析协议ARP
- 分组交付与路由选择

# 网络层IP附属协议



**ICMP** (Internet Control Message Protocol): Internet控制报文协议

**IGMP** (Internet Group Management Protocol): Internet 组管理协议

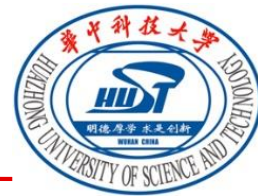
**ARP** (Address Resolution Protocol): 地址解析协议

**RARP** (Reverse Address Resolution Protocol): 反向地址转换协议



## □ 问题: 网络如何互联

- 网络互联与服务模型
- IP协议
- IP地址
  - 标准分类IP地址
  - 子网划分
  - 无类别域间路由CIDR
  - 网络地址转换NAT
- IP附属协议
  - 网络控制管理协议ICMP
  - 地址解析协议ARP
- 分组交付与路由选择



# 问题: 如何报告错误

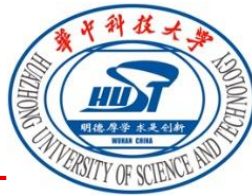
---

## □ Internet如何处理错误?

- 当IP在数据报传送受阻而要将其丢弃时
- 不能默默地归于失败

## □ Internet控制报文协议(ICMP)

- IP协议的附属协议
- 当路由器或主机无法成功处理IP数据报时向源主机发送错误报文(定义了错误报文集合)



# Internet控制报文协议

## ❑ 主机和路由之间进行网络层的通信

- 错误报告: 无法到达的主机, 网络, 端口, 协议
- echo request/reply (ping命令)

## ❑ 运行于IP之上的网络层协议

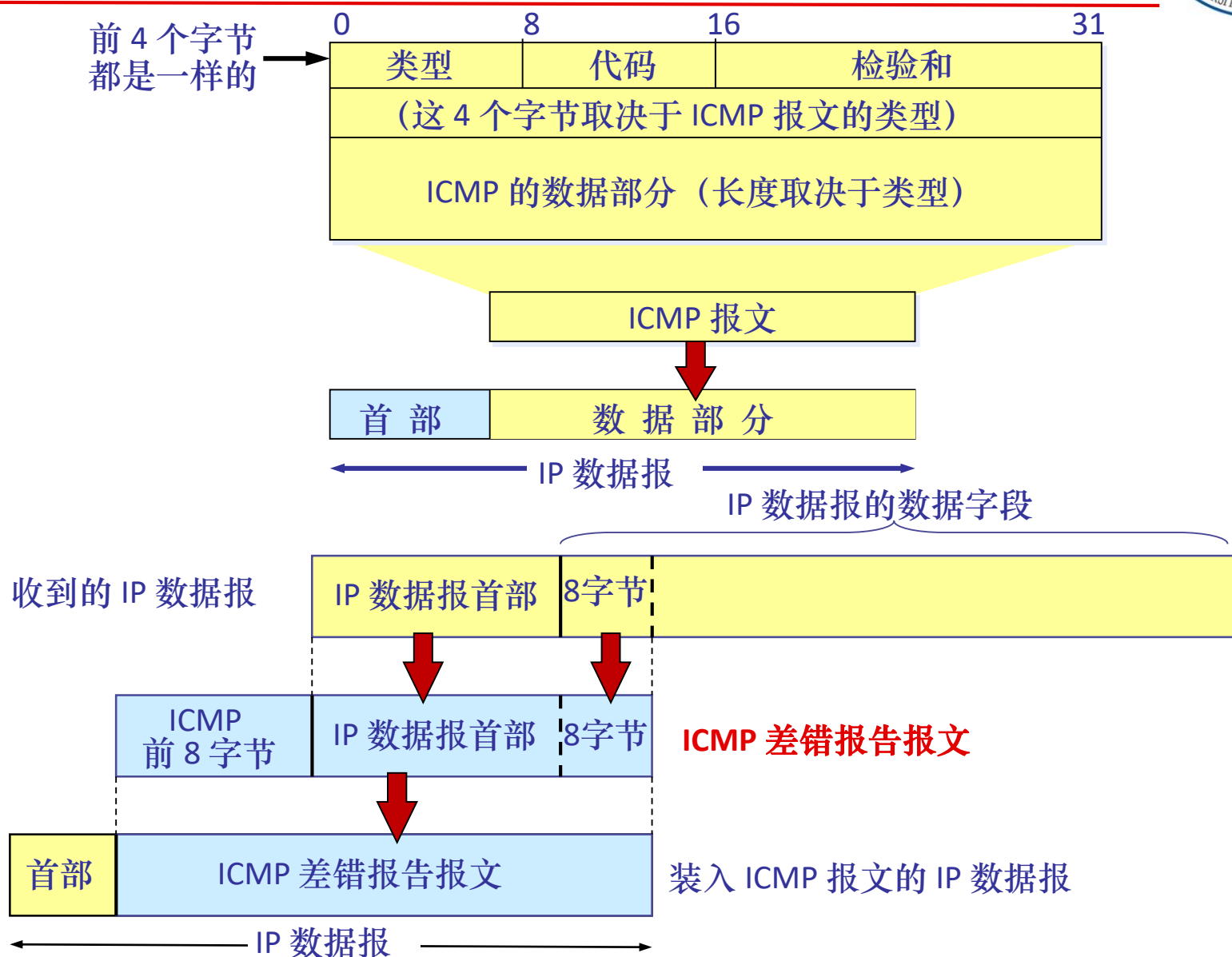
- ICMP 消息封装于IP数据报中

## ❑ ICMP 报文

- 错误原因(类型, 编码)

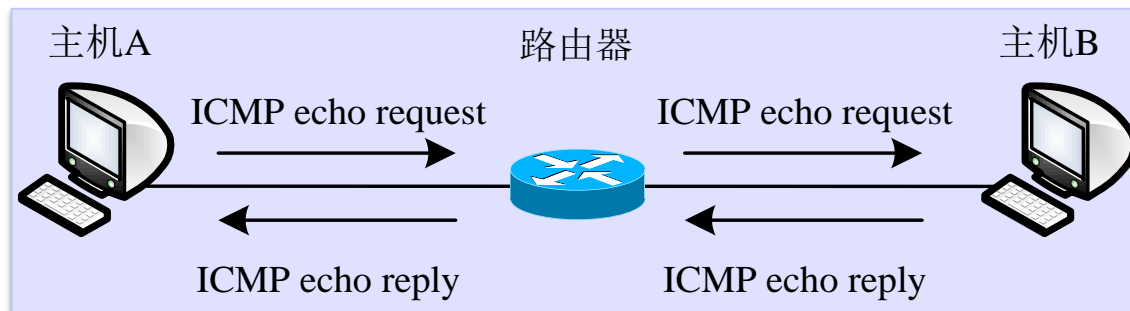
类型	编码	描述
0	0	echo reply (ping)
3	0	dest. network unreachable
3	1	dest host unreachable
3	2	dest protocol unreachable
3	3	dest port unreachable
3	6	dest network unknown
3	7	dest host unknown
4	0	source quench (congestion control - not used)
8	0	echo request (ping)
9	0	route advertisement
10	0	router discovery
11	0	TTL expired
12	0	bad IP header

# ICMP分组格式



# Ping命令

## □ 一台主机Ping另一台主机的过程



## □ Ping命令执行过程示意图

```
C:\WINDOWS\system32\cmd.exe

C:\Documents and Settings\Administrator>ping www.baidu.com

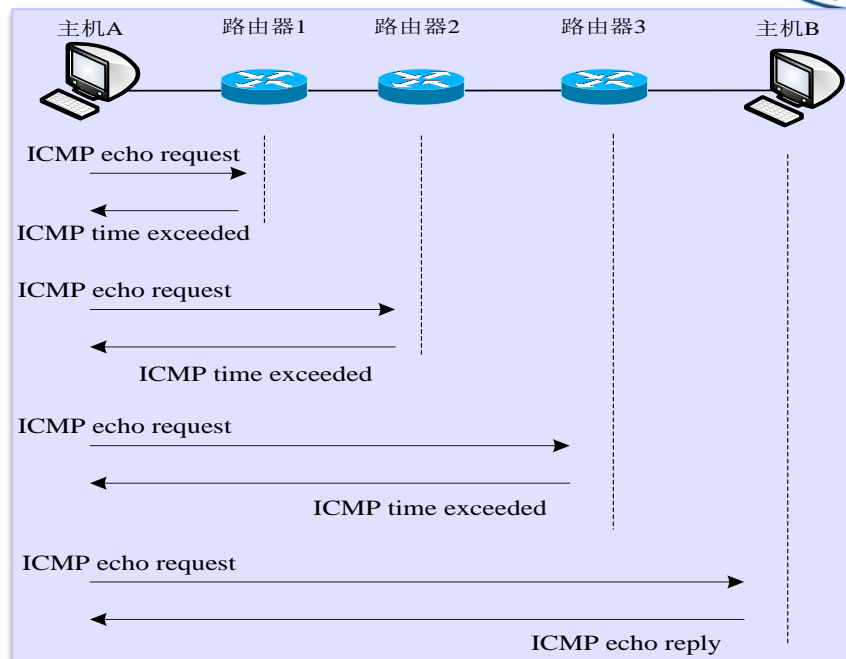
Ping www.baidu.com [119.75.217.109] with 32 bytes of data:

Reply from 119.75.217.109:bytes=32 time=5ms TTL=52
Reply from 119.75.217.109:bytes=32 time=5ms TTL=52
Reply from 119.75.217.109:bytes=32 time=5ms TTL=52
Reply from 119.75.217.109:bytes=32 time=5ms TTL=52

Ping statistics for 119.75.217.109:
    Packets: Sent=4, Received=4, Lost=0<0% loss>,
    Approximate round trip times in milli-seconds:
        Min=3ms, Max=5ms, Average=3ms
```

# Tracert命令的应用

## Tracert工作原理



## Tracert命令执行过程

```
C:\WINDOWS\system32\cmd.exe

C:\Documents and Settings\Administrator>tracert www.baidu.com

Tracing route to www.a.shien.som[119.25.217.109]
Over a maximum of 30 hops:

 1  1ms  <1ms  <1ms  192.168.1.254
 2  4ms   4ms   4ms   202.113.25.1
 3  5ms   4ms   2ms   202.113.18.129
 4  2ms   1ms   3ms   202.113.18.182
 5  5ms   4ms   2ms   202.127.5.12
 6  1ms   1ms   1ms   202.127.216.185
 7  3ms   3ms   4ms   202.112.46.89
 8  4ms   4ms   4ms   202.112.25.11
 9  4ms   5ms   3ms   202.112.6.58
10  5ms   4ms   4ms   192.168.0.5
11  5ms   5ms   4ms   10.65.190.130
12  3ms   3ms   3ms   119.75.217.109

Trace complete.
```

## □ 问题: 网络如何互联

- 网络互联与服务模型
- IP协议
- IP地址
  - 标准分类IP地址
  - 子网划分
  - 无类别域间路由CIDR
  - 网络地址转换NAT
- IP附属协议
  - 网络控制管理协议ICMP
  - 地址解析协议ARP
- 分组交付与路由选择

# 地址转换

## □ 物理地址

- 数据链路层地址 == 某一物理网络定义的硬件地址

## □ 地址转换的解决方案

- 方案1: 直接编码

- 将物理地址编码为IP地址的主机号部分
- Ethernet 48bit 地址?

- 方案2: 表

- 网络管理员在每一个主机内部维护一个IP地址-物理地址的映射表
- 如何将映射表分发给每一个节点?



# 地址转换协议(ARP)

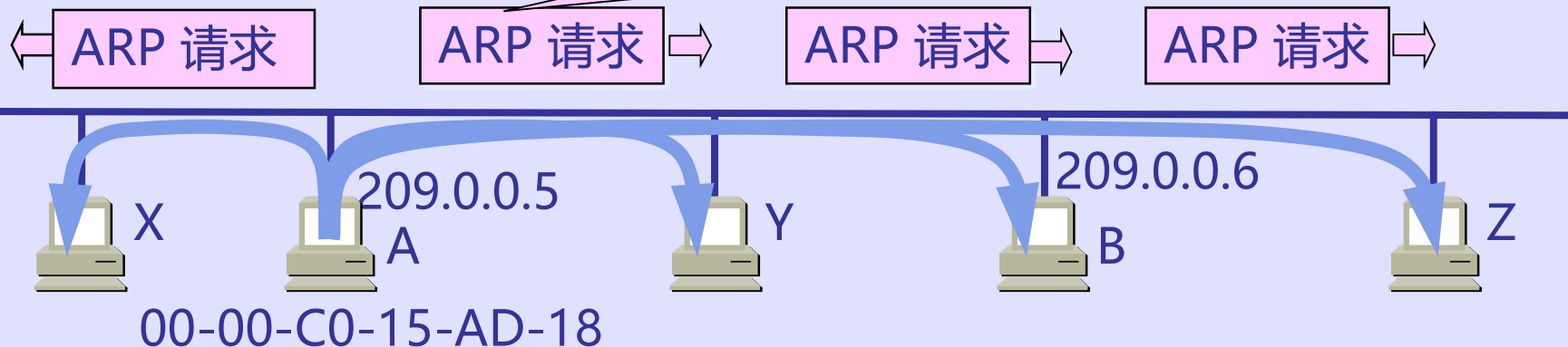
- 给定目的主机的IP地址, ARP可以找到属于同一个物理网络内部的目的主机的物理地址
  - 从IP地址到硬件地址的解析是自动进行的, 主机的用户对这种地址解析过程是不知道的。
- ARP采用广播机制
- 每一个主机维护一个ARP表: <IP 地址, 物理地址>
  - 15分钟记录未更新则超时失效
- 当ARP表中无目的IP地址的对应记录, 则ARP协议开始启动

# ARP工作原理

- ❑ 主机X广播一个ARP请求报文来探寻未知的节点D的IP地址
- ❑ 节点D向源节点反馈其物理地址
  - D节点在本地ARP表中增加或更新节点X对应的记录
- ❑ 节点X将收到的信息写入其ARP表
- ❑ 同一网络内的其他主机
  - 如果本地的ARP表中已存在节点X的IP地址对应的记录, 则更新记录 **<IP address of X, physical address of X>**
  - 否则, do nothing

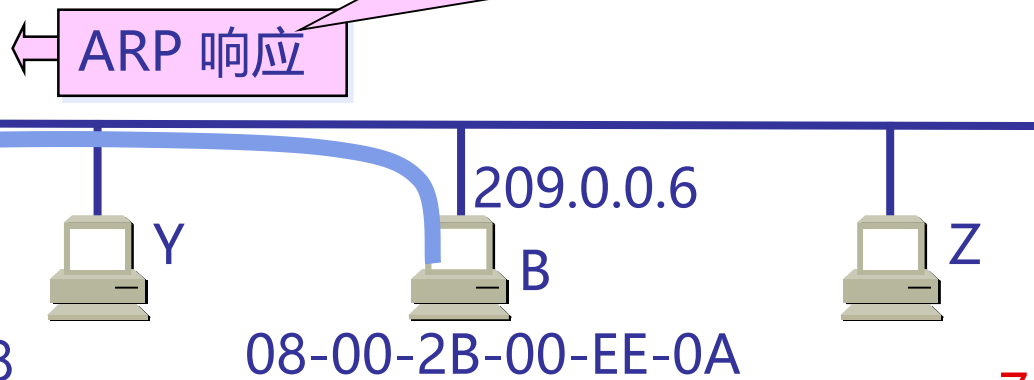
主机 A 广播发送  
ARP 请求分组

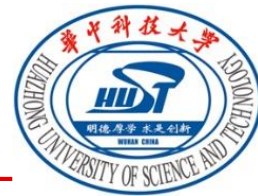
我是 209.0.0.5, 硬件地址是 00-00-C0-15-AD-18  
我想知道主机 209.0.0.6 的硬件地址



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





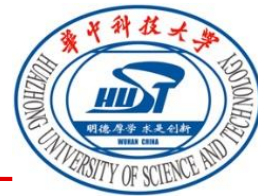
# 使用ARP的四种典型情况

## □ 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。
- 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。
- 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。



# ARP的深入讨论

---

## □ ARP机制不同的数据链路地址

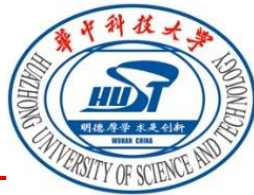
- IP地址 → 以太网MAC地址
- IP地址 → 令牌环网MAC地址
- IP地址 → ATM地址(LANE), ...

## □ RARP

- 反向地址解析协议
- 数据链路层地址 → IP地址

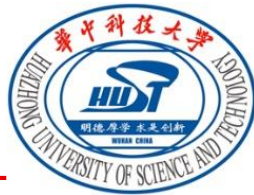
## □ ARP表和ARP广播风暴

## □ Windows 命令: arp



# 考纲要求

- ❑ 理解：IP尽力服务的服务模型
- ❑ 理解：IP数据报转发的原理
- ❑ 掌握：IPv4分段与重组的概念及其计算
- ❑ 掌握：IPv4地址设计的相关概念，包括网络地址、广播地址、掩码、子网、超网等
- ❑ 掌握：IP地址规划，能够划分子网、分配IP地址
- ❑ 掌握：CIDR的地址规划与路由聚合，能够根据最长前缀匹配原则测算路由寻址过程
- ❑ 了解：与IP地址管理与使用相关的协议，包括DHCP、NAT、ICMP、ARP等
- ❑ 了解：如何辨析IPv6地址



# 补充: IPv6

## □ 地址空间

- 128-bit 地址空间包括  $2^{128}$  (约合 $3.4 \times 10^{38}$ ) 个IP地址
- 与IPv4 (43亿) 相比, 增长 $10^{29}$  倍多
- 在地球表面平均每平方厘米 $6.7 \times 10^{19}$ 个地址

## □ 表示方式

- 冒分十六进制的方式表示为8个16-bit字段
  - 例如: 47CD:1358:39CD:A37D:1845:2A3B:6479:DAEF
- 每个x前面的0可省略, 可把连续0的x表示为 "::", 且 "::"只能出现1次
  - 简化前地址, 2001:0DA8:0000:0000:200C:0000:0000:00A5
  - 简化后地址, 2001:DA8:0000:0000:200C::A5

# 要点回顾

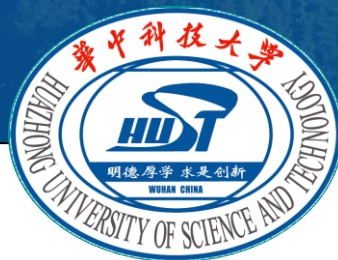
## □ IP机制解决了异构性问题

- 最大努力交付服务模型, 提供无连接、不可靠的数据报传送
- 提供一种通用的分组格式, 分段和重组机制使分组能够适应具有不同MTU的网络
- 提供识别所有主机的全局地址空间, ARP机制使其能够在具有不同物理地址编制方案的网络上运行

## □ IP机制解决了扩展性问题

- Solution1: 两级→三级 (主机字段借位), 子网掩码 (有分类限制的掩码)
- Solution2: 完全打破A,B,C分类规则, 变长掩码 (24位自适应聚合), CIDR网络前缀
- Solution3: 利用私有机制 NAT
- Solution4: IPv4→IPv6





**THANKS**  
谢谢聆听

Email: [chenwang@hust.edu.cn](mailto:chenwang@hust.edu.cn)

Website: <http://www.chenwang.net.cn>