



第5章 网络层_1

杜广龙（参考：袁华）

华南理工大学计算机科学与工程学院

广东省计算机网络重点实验室



主要内容

■ 通信网络及提供的服务 (5.1)

- 无连接的服务
- 面向连接的服务

■ IPv4协议 (5.2)

■ 与IP有关的其它协议或技术 (5.3)



网络层





网络层的主要功能和重要地位

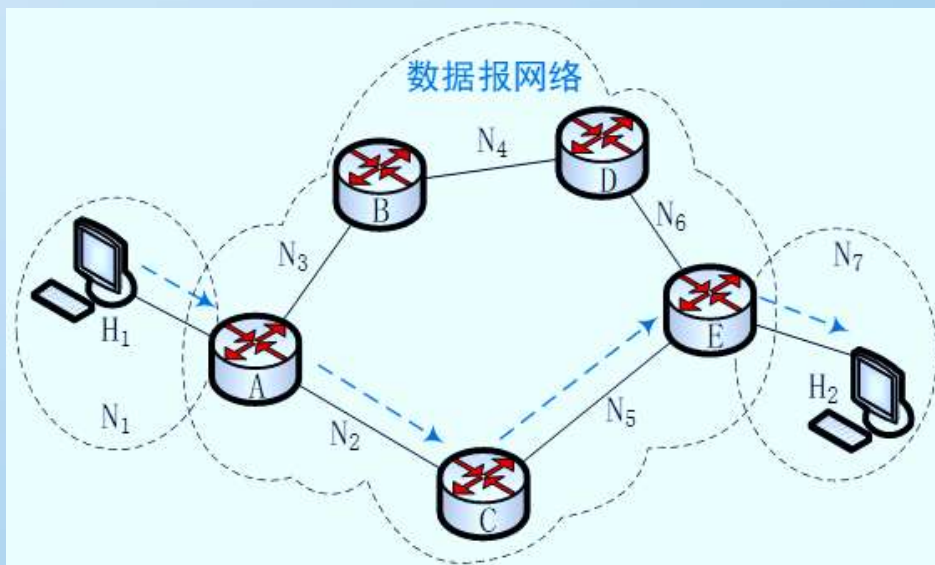
- 核心层之一
- 网络层的功能：将分组从源机一路送到目的机。
- 寻址：分组从源机找到目的机的过程叫寻址。
- 通信网络：源机和目的机可能位于同一个局域网，更大的可能是源机和目的机位于完全不同的局域网中；源机和目的机之间隔着一个中间网络。
 - 主要由中间**交换设备**和连接交换设备的**链路**构成



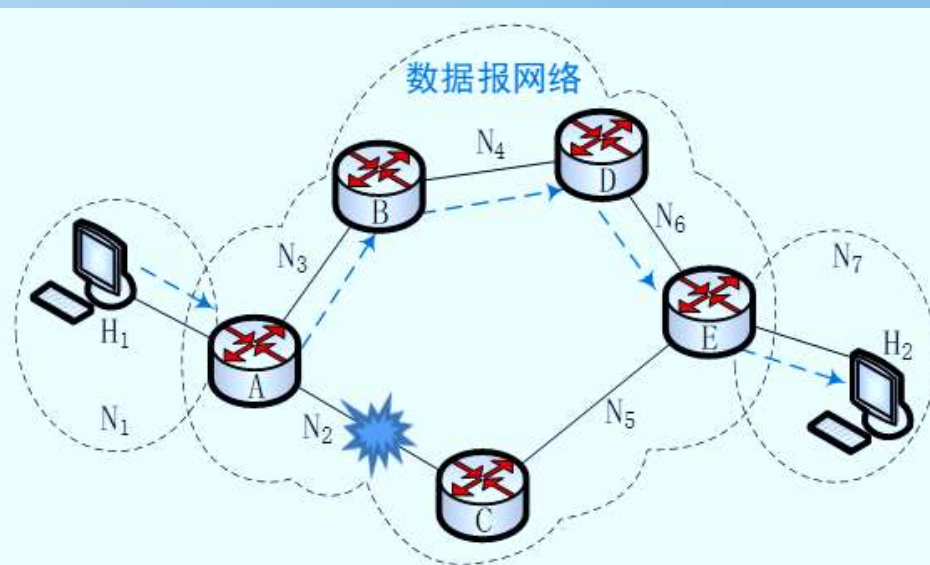


数据报网络：提供无连接的服务

- 路由器：主要交换设备
 - 维护路由表
 - 每个分组携带目的地址
 - 路由器根据路由表和目的地址，将分组推送到源机所在的网络



(a) 网络正常时 $H_1 \rightarrow H_2$ 的穿越路径



(b) A 和 C 之间的网络故障时





无连接服务的主要技术特征

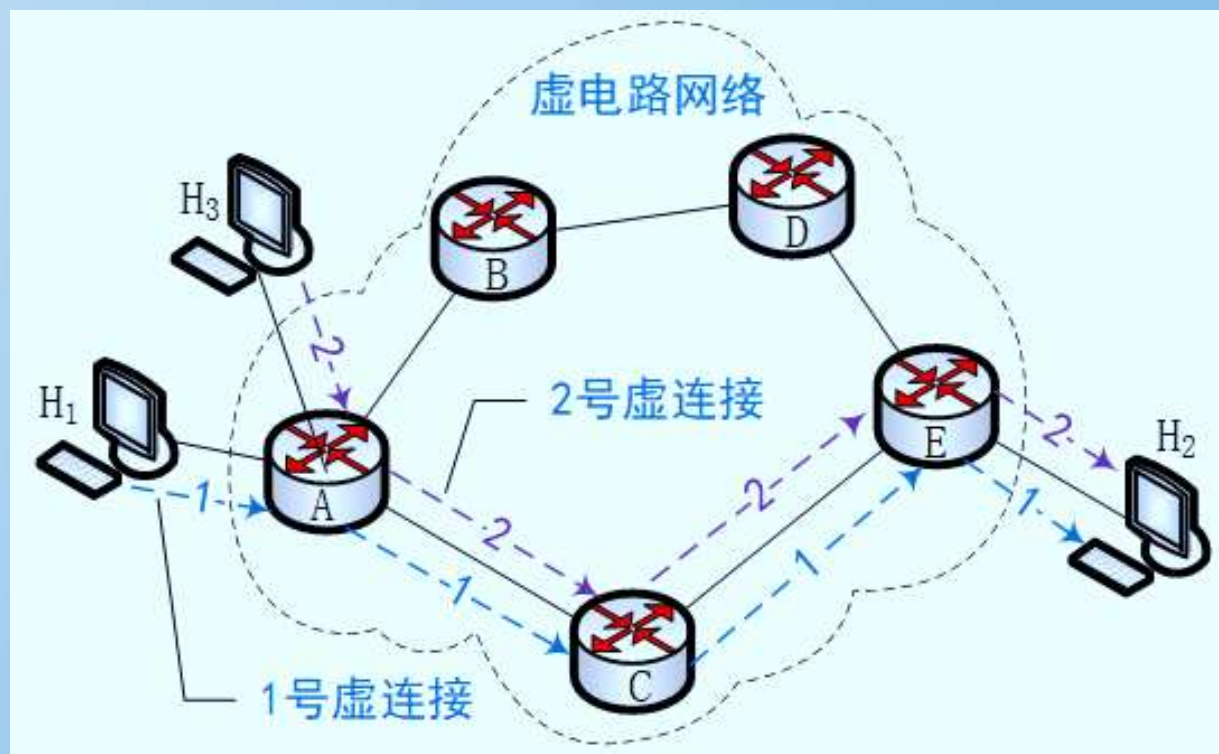
- **无连接**：数据分组传输无需建立连接
- **独立寻径**：每个分组都携带目的地址，可以独立寻址
- **乱序到达**：先发的分组可能因为重传或走了一条慢的路径而后到，后发的分组也可能因为走了捷径而先到
- **重组**：目的机需要对收到的分组重新排序
- **无状态**：路由器无需保留过往分组的状态，只是尽力而为地转发
- **抗毁性**：单点故障，甚至多点故障，都不会让网络瘫痪，分组可以在尚通达的网络中自由寻路。
- **灵活分片**：路由器可能承担分片的任务
- **服务质量难以保证**





虚电路网络：提供面向连接的服务

- 面向连接的服务：在分组传输之前，搭建一条连接，从源机到目的机的分组只需要沿着这条连接就可以穿越整个虚电路网络，到达目的机。
- 连接号：区分不同的连接





面向连接服务的特点

- 按路径传输：数据分组传输之前，需要建立连接
- 无需地址寻径：每个分组携带连接号，而不需要携带目的机地址
- 按序到达：先发的先到，后发的后到
- 连接损坏：分组传送中的连接如果发生故障，连接中断，需要重新搭建虚连接
- 有状态：路由器工作较为复杂，要保存和维护所有虚连接的状态信息
- 服务质量保证：可以容易地进行拥塞控制和服务质量保证





主要内容

- 通信网络及提供的服务 (5.1)
- IPv4协议 (5.2)
 - IP分组和IP地址
 - 子网划分
 - IP寻址和MAC寻址
 - 无类域间路由CIDR
- 与IP有关的其它协议或技术 (5.3)





IP是什么？

- 互联网协议：Internet Protocol
- IP为路由提供路由所需要的信息，比如IP地址、数据、数据尺寸等，并把这些信息封装在分组/报包（**Packet**）中。
- 主要包括两方面的内容
 - IP分组格式
 - IP地址编址

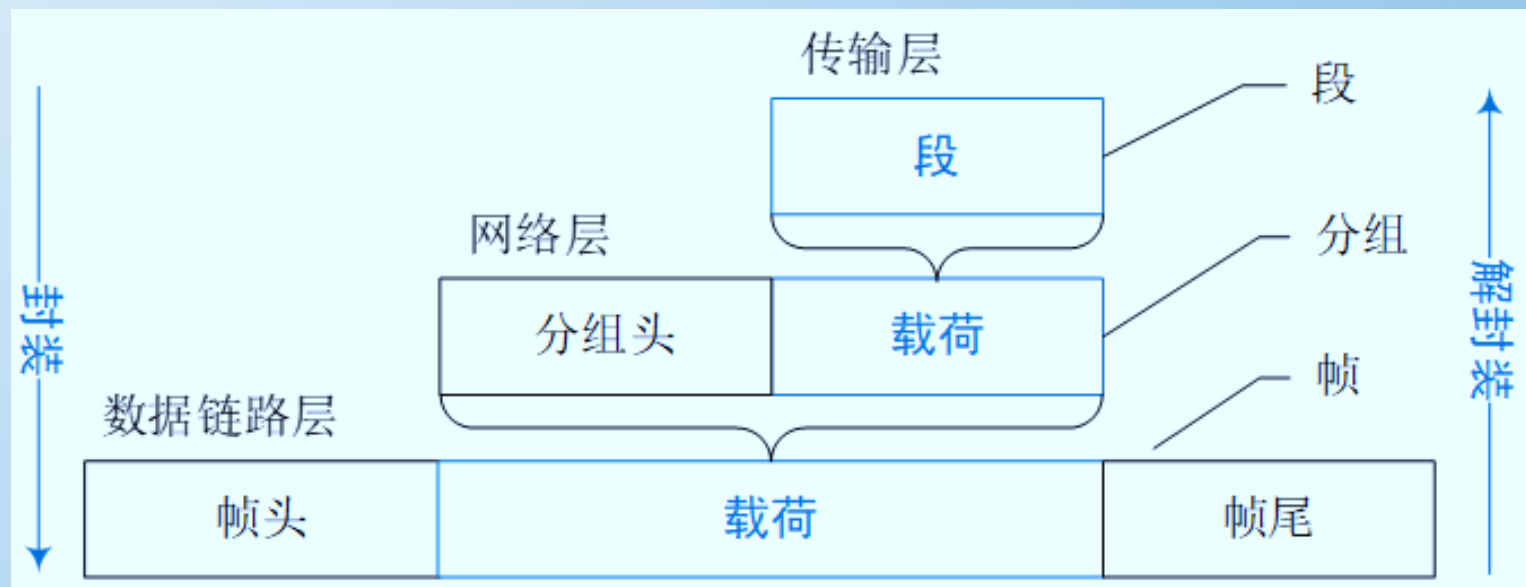




IP分组、段和帧的关系

■ 网络层和传输层及数据链路层的关系

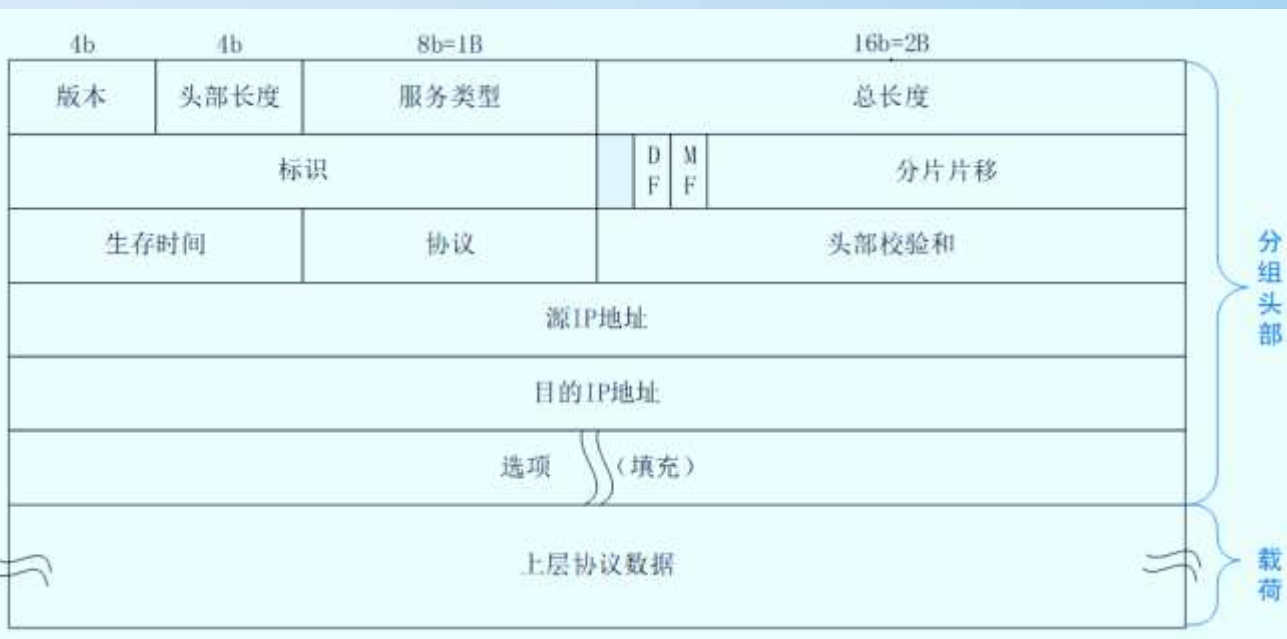
□ 3层对应的PDU：分组、段、帧





IP分组的构成

- 分组头部和载荷
- 分组头部由14个（或15/16）字段构成





构成IP分组头部的字段_1

- **协议字段**: 0100
- **头部长度**: 0101-1111 , 单位4B, 对应十进制: 20B-60B
- **服务类型**: 表示 IP 分组的优先级别或重要程度, 为差分服务提供被怎样区别处理的依据
- **总长度**: 长 16 位, 表示整个 IP 分组 (头部和数据载荷) 的总长度, 单位是字节, 分组长度的上限是65535 字节
- **标识**: 长 16 位, 用来标识分组的序号。由发送方维护, 每发出一个数据分组, 标识增加 1





构成IP分组头部的字段_2

- **标记位**：第⑥、⑦和⑧这 3 个字段合起来称为标记共长 3 位
 - ⑥标记位未使用，须为 0
 - ⑦不要分片DF标记位，取值 0 表明允许分片
 - ⑧更多分片MF标记位叫，取值 0 表明这是最后一个分片
- **片偏移**：第⑨个字段，长 13 位，单位是 8 字节。
- **生存时间**：第⑩个字段，常称为 TTL (Time To Live) ， 每经过一个路由器，其值减1





构成IP分组头部的字段_3

■ **协议**：第⑪个字段，长 8 位，表示网络层之上的传输层使用的协议

协议字段值	上层协议	备注
1	ICMP	互联网控制消息协议，用以报告 IP 分组传输的错误和测试网络。参考 5.3.2 节
2	IGMP	组成员管理协议，用于组播成员管理。参考 5.5.5 节
6	TCP	传输控制协议，用以提供可靠的数据段传输。参考 6.3 节
9	IGP	内部网关协议，用以支持私人定制的内部网关协议。参考 5.5.1 节
17	UDP	用户数据报协议，用以提供简洁高效的数据段传输。参考 6.2 节
41	IPv6	应用于 IPv4 向 IPv6 的过渡技术，表示载荷搭载的是 IPv6 分组。参考 5.4 节
89	OSPF	开放最短路径优先协议，最流行的内部网关协议。参考 5.5.3 节





构成IP分组头部的字段_4

- **校验和**：头部校验和，长 16 位，用以检查 IP 分组头部有无发生传输错误。采用分组头部的数据计算互联网校验和
- **源 IP 地址**：长 32 位，标识当前分组的发出接口，即从哪儿来
- **目的 IP 地址**：长 32 位，标识当前分组的目的接口，即到哪儿去
- **选项**：其长度必须是 4 字节的整数倍，如果不是，需要用 0 填充；选项的长度上限是 40 字节。
 - 安全、松散\严格源路由、时间戳、流标识





关于分片

- 每个网络都有自己的承载能力限制
- 最大传输单元MTU：数据帧中载荷的最大长度。即分组的总长度的最大值



网络名	MTU（字节）	备注
以太网 II	1500	如果是 Jumbo 帧，此值可以达到 9000 字节，甚至更高。
PPPoEv2	1492	用 PPP 携带 Ethernet II 数据时，扣除 PPP 的 8 字节头部。
802.11	2304	2304 是 MAC 服务单元数据（MAC），如果加上 WEP（早期认证方法）的 8 字节，应为 2312 字节，如果加上 WPA2 的 16 字节，应为 2320 字节。
FDDI	4352	曾经非常引人注目的一种网络，可靠的快速（100M）传输
令牌环	4464	采用确定性介质访问控制，无冲突，现在已经很少见
X.25	576	一种古老的网络，提供面向连接（虚电路）的传输





分片条件和方法

- 路由器收到分组后，如果分片，需要满足下面两个条件：

- 分组的长度大于转出网络的 MTU,即： $L > M$ 。
- 分组中的标记位 DF 没有被置位，即： $DF = 0$

- 分片方法

$$F_i = \frac{d}{8} \times (i - 1), \text{ 其中 } 1 \leq i \leq n$$

- 假设需要分片的总片数为 n ，分组头部长度是 20 字节，分片载荷的长度上限是 d ，则 $d = \lfloor \frac{M-20}{8} \rfloor \times 8$ $n = \lceil \frac{L-20}{d} \rceil$

- 除了最后一个分片，其余的分片必须满载

$$F_i = \frac{d}{8} \times (i - 1), \text{ 其中 } 1 \leq i \leq n$$

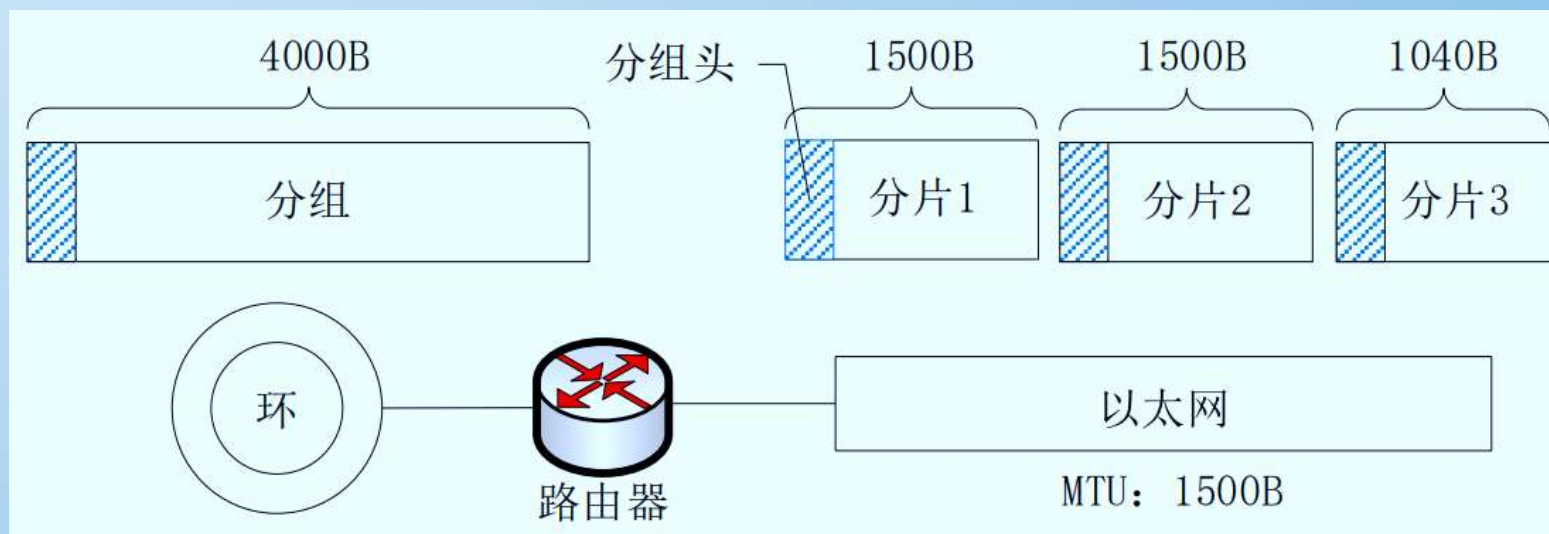
$$L_i = \begin{cases} d + 20, & 1 \leq i < n \\ L - d \times (n - 1), & i = n \end{cases}$$





例5-1

- 一个分组长 4000 字节，其中整个头部长度为 20 字节，头部标识字段值为 8580，标记位 DF=0，当分组到达路由器时，发现转出网络是以太网，其 MTU 是 1500 字节。路由器是否需要分片？如果进行分片，应分为多少片？每个分片的总长度、头部标识、标记位 DF 和 MF 以及片偏移的值分别是多少





例5-1_解答

■ 因为 $4000\text{B} > 1500\text{B}$, $\text{DF}=0$, 满足条件

■ 分片的最大载荷 $d = \lfloor (1500 - 20) / 8 \times 8 = 1480\text{B}$, 则分片总片数为

$$n = \left\lfloor \frac{L-20}{d} \right\rfloor = \left\lfloor \frac{4000-20}{1480} \right\rfloor = 3$$

字段名 分 组	总长度(字节)	标识	DF	MF	片偏移
原分组	4000	8580	0	0	0
分片 1	1500	8580	0	1	0
分片 2	1500	8580	0	1	185
分片 3	1040	8580	0	0	370

■ 前两个分片满载, 第三个分片的总长度应该是 $4000\text{B} - 1480\text{B} \times 2 = 1040\text{B}$, 其中的净载荷为 1020B





练习

假设网络报头大小为 20 字节。一个路由器的一个接口接收MTU为2000字节 链路的数据报，并输出到一个MTU为500字节的链路。如果路由器接收一个输入数据报，大小为 2000字节，路由器需要生成多少个分片，才能转出到MTU为500字节 的链路进行传输？

- A. 5
- B. 6
- C. 3
- D. 4

分片数= $\lceil (2000-20) / (500-20) \rceil = \lceil 4.125 \rceil = 5$ ，即需要5个分片，第5个分片不是满负荷。





练习

一个IP分组头部中的字段**MF=1**，分片偏移值是“0”，是判断这个数据分组是哪一个分片？

- A. 第一个分片
- B. 中间的分片
- C. 最后一个分片
- D. 无法判断

MF=1，意味着，后面还有分片，这不是最后一个分片，可能是第一个或中间的分片；分片偏移值等于零，意味着，这是第一个分片。



课后练习



节点 A 到 B 通过路由器 R1 和 R2 路由。通过网络发送的 IP 数据报有 20 字节长的报头。A-R1 链路的 MTU 是 1800 字节、R1 — R2 链路的 MTU 是 1200 字节，R2 — B 链路的 MTU 是 600 字节。如果 A 想要发送长度为 2800 的消息，分别发送了两个大小为 1800 字节和 1040 字节数据报才发完了这个消息。（注意第一个数据报携带 1780 字节的有效载荷，第二个数据报携带 1020 字节的有效载荷。）问：B 接收的数据报的总个数是多少？（假定源数据报中的 DF=0）。题意所表达的拓扑示意如下：

- A. 9
- B. 7
- C. 6
- D. 5

解析：各链路能够承载的最大数据分别是：A-R1： $\lfloor (1800-20)/8 \rfloor * 8 = 1776B$ ；R1-R2： $\lfloor (1200-20)/8 \rfloor * 8 = 1176B$ ；R2-B： $\lfloor (600-20)/8 \rfloor * 8 = 576B$ 。源 A 根据所在网络的 MTU，将 2800 字节的消息封装成两个数据报：1796 字节（载荷 1776 字节）和 1044 字节（载荷 1024 字节）。R1 收到这两个报，不得对 1796 字节的数据报进行分片，分成 1196 字节（载荷 1176B）和 620 字节（载荷 600B）。R2 收到了 1196B、620B 和 1044B 的三个数据报，转出到 R2-B 链路时，不得不进行再次分片：1196B = 596B + 596B + 44B（3 片）；620B = 596B + 44B（2 片）；1044B = 596B + 468B（2 片）。最后 B 收到的数据报个数为：3 + 2 + 2 = 7。





为什么需要IP地址？

■ 为了将分组送达，必须知道分组要去哪里！

□ IP地址

□ IP分组中的第11、12个字段（源、目的）

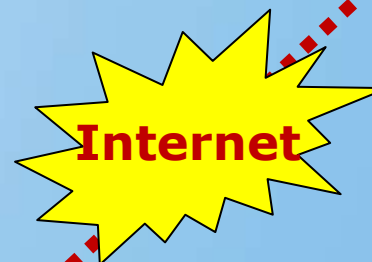
■ 谁需要IP地址？

□ 主机、路由器、其它设备

□ 确切地说：是接口！



网卡



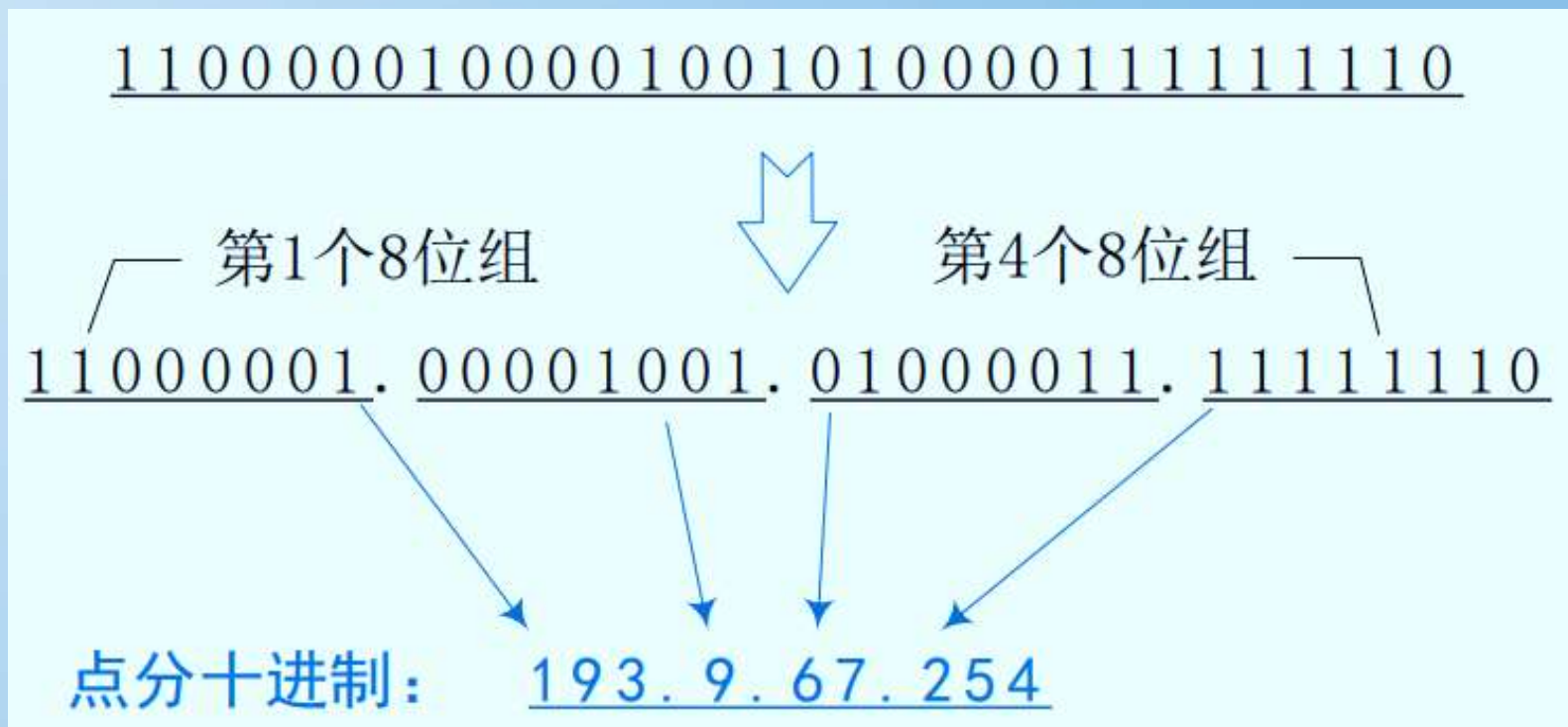
网卡





IP地址的表示

- 32位二进制位
- 点分十进制

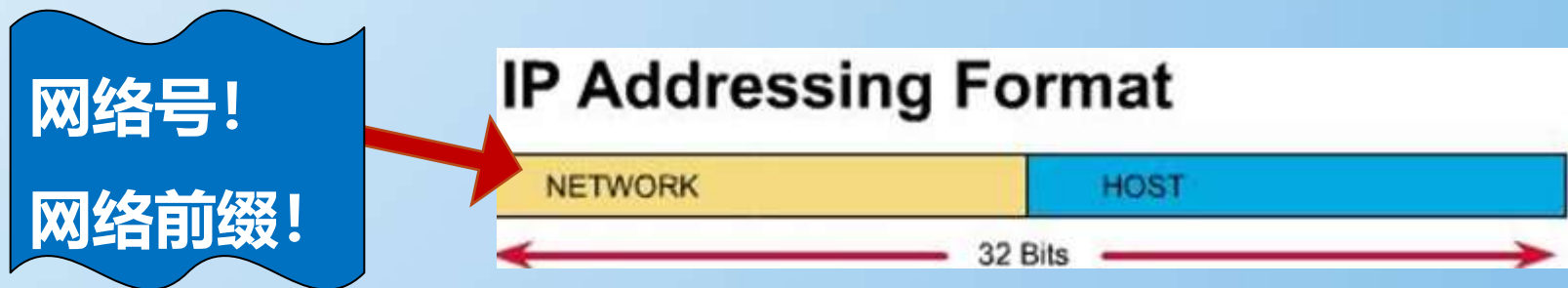




IP地址的本质属性

■ IP地址的内涵：天生2层结构

□ 网络号：所属的网络



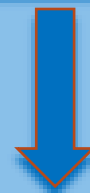
■ IP地址定义了一台设备的网络位置，而不仅仅是一个名字！

□ 全网唯一

□ 网络位置改变了，IP地址必须相应改变！（搬家了）

□ 一台设备可以有多个IP地址（Multihomed）

广州市天河区五山路



广州市番禺区广州大学城

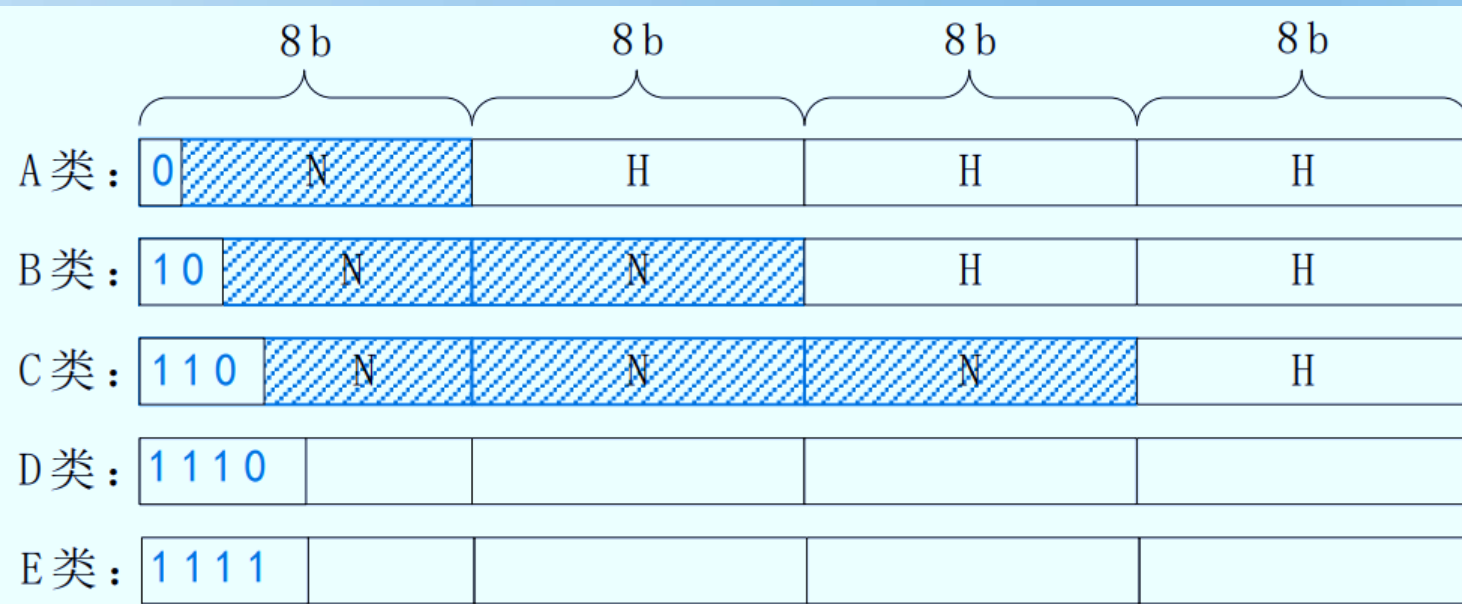




IP地址的分类

■ 全球地址总数约43亿 $2^{32}=4,294,967,296$

■ 为了便于管理，分为5类

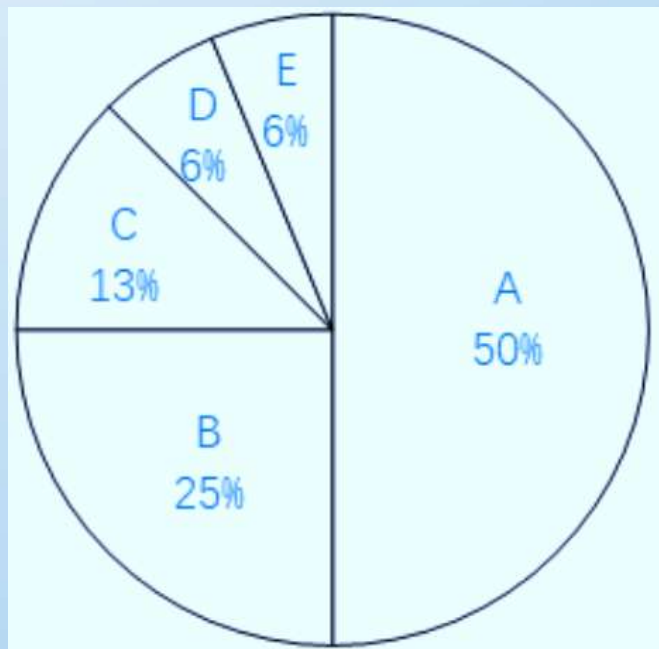


■ 其中ABC 3类常用

类别	最高固定位	第 1 个 8 位组的值	网络位数/个数		主机位数/个数	
A	0	0-127	8b	128	24b	约 1600 万
B	10	128-191	16b	约 1.6 万	16b	约 6.6 万
C	110	192-223	24b	约 210 万	8b	256

各类IP地址的占比

- A类网络地址数最少 (128个) , 但总数最多, $126 \times (2^{24}-2)$



- 分类导致极大的浪费

American Registry for Internet Numbers		
000/8	IANA - Local Identification	
001/8	APNIC	
002/8	RIPE NCC	
003/8	Administered by ARIN	
004/8	Level 3 Parent, LLC	1992-12
005/8	RIPE NCC	2010-11
006/8	Army Information Systems Center	1994-02
007/8	Administered by ARIN	1995-04
008/8	Administered by ARIN	1992-12
009/8	Administered by ARIN	1992-08
010/8	IANA - Private Use	1995-06
011/8	DoD Intel Information Systems	1993-05
012/8	AT&T Bell Laboratories	1995-06
013/8	Administered by ARIN	1991-09
014/8	APNIC	2010-04
015/8	Administered by ARIN	1994-07
016/8	Administered by ARIN	1994-11
017/8	Apple Computer Inc.	1992-07
018/8	Administered by ARIN	1994-01

施乐

HP

DEC

MIT



中国IPv4地址数排名：第二

■ 人均约0.2个，而美国4、5个人拥有一个IP

項次	國家	202102	202101	202012	202011	202010	202009
1	美國	1615279712	1615464288	1614306912	1610092544	1610265856	1610075392
2	中國大陸	344414464	344413952	344408576	344374016	344896768	340703744
3	日本	190005504	190004224	189992448	190119168	190086656	190087424
4	德國	123839104	123833472	123787392	124521856	124344960	124254336
5	英國	113375096	113357176	114504824	114552088	114624024	114640408
6	南韓	112475136	112474624	112473088	112483584	112483584	112482560
7	巴西	87182592	87112960	87115008	87115776	87115520	87113216
8	法國	82413584	82404368	82368528	82709520	82725648	82709520
9	加拿大	69713664	69847040	69753856	69510144	69509376	69494528
10	義大利	54989632	54986560	54950976	55045952	55028032	55014208





可是，中国网民人数第一！

**2022年6月，中国网民人数10.5亿！
地址缺口巨大！**

#	Country or Region	Internet Users 2020 Q1	Internet Users 2000 Q4	Population, 2020 Est.	Population 2000 Est.	Internet Growth 2000 - 2020
1	China	854,000,000	22,500,000	1,439,062,022	1,283,198,970	3,796 %
2	India	560,000,000	5,000,000	1,368,737,513	1,053,050,912	11,200 %
3	United States	313,322,868	95,354,000	331,002,651	281,982,778	328 %
4	Indonesia	171,260,000	2,000,000	273,523,615	211,540,429	8,560 %
5	Brazil	149,057,635	5,000,000	212,392,717	175,287,587	2,980 %
6	Nigeria	126,078,999	200,000	206,139,589	123,486,615	63,000 %
7	Japan	118,626,672	47,080,000	126,854,745	127,533,934	252 %
8	Russia	116,353,942	3,100,000	145,934,462	146,396,514	3,751 %
9	Bangladesh	94,199,000	100,000	164,689,383	131,581,243	94,199 %
10	Mexico	88,000,000	2,712,400	132,328,035	2,712,400	3,144 %





小结：IP地址的特点

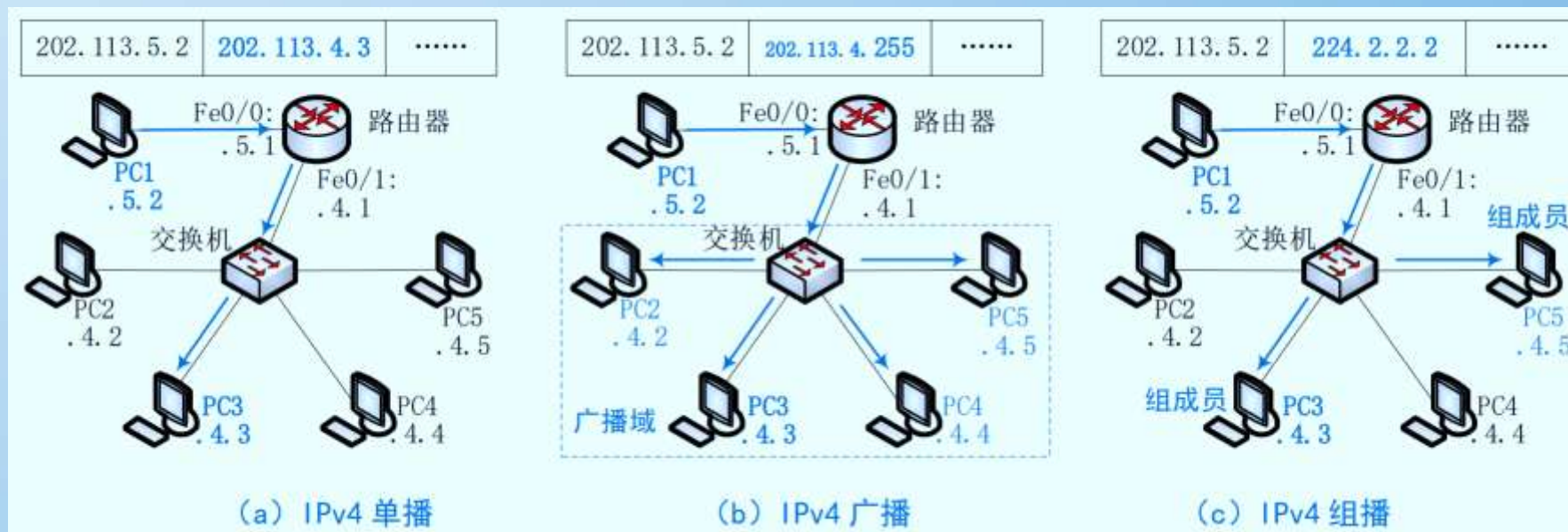
- 全球唯一性：唯一性是IP寻址的依据（私人地址除外）
- 层次性：天生二层，实际上更多层
- 位置相关性：一个IP地址所属的网路携带了它的物理位置
- 多归属性：IP地址绑定在接口，一台主机可以有多个接口，具有多个IP地址
- 同一性：同属于一个网络的所有主机（接口）的IP地址的网络号相同





IP地址的单播、广播和组播

- IPv4 单播：分组从一台主机向另外一台主机发送分组，其目的地址是单播地址，代表某一台主机。是一对一的分组传输
- IPv4 广播：分组从一台主机向所有其它主机发送，其目的地址是广播地址，可以是受限广播地址，也可以是定向广播地址
- IPv4 组播：分组从一台主机向某个组的成员发送，其目的地址是 IP 组播地址，即 D 类地址





特殊的IP地址_1

- 受限广播地址**：用受限广播地址作为目的地址的 IP 分组，不会被路由器转发
- 未指定地址**：一台不知道自己 IP 地址的主机，在需要发送一个 IP 分组的时候，其源地址可用这个未指定地址。

名称	IP 地址表示	用途
受限广播地址	255.255.255.255	只能用作 IP 分组中的目的地址，向全网所有主机广播，实际上是本地广播。
未指定地址	0.0.0.0	用作特殊情况下 IP 分组的源地址。
环回地址	127.*.*.*	只能用作 IP 分组中的目的地址，用于测试本机 IPv4 软件。
链路本地地址	169.254.*.*	仅可以链路本地通信。
直接广播地址	Net-ID.255	只能用作分组中的目的地址，向 Net-ID 所在网络的所有主机广播。
网络地址	Net-ID.0	不能用作分组中的源或目的地址，代表 Net-ID 网络本身。





特殊的IP地址_2

- **环回地址**：只能用作 IP 分组的目的地址，这样的分组不会流出本机，通常用来测试本机的 IP 软件。
- **链路本地地址**：路由器不能转发源自链路本地地址的 IP 分组，所以，使用了此类地址的主机只能够跟直接相连的设备通信，
- **定向\直接广播地址**：代表网络号所在的网络的所有主机，只能用作分组的目的地址
- **网络地址**：代表网络号所在的网络本身，既不能用作分组的源地址，也不能用作目的地址。



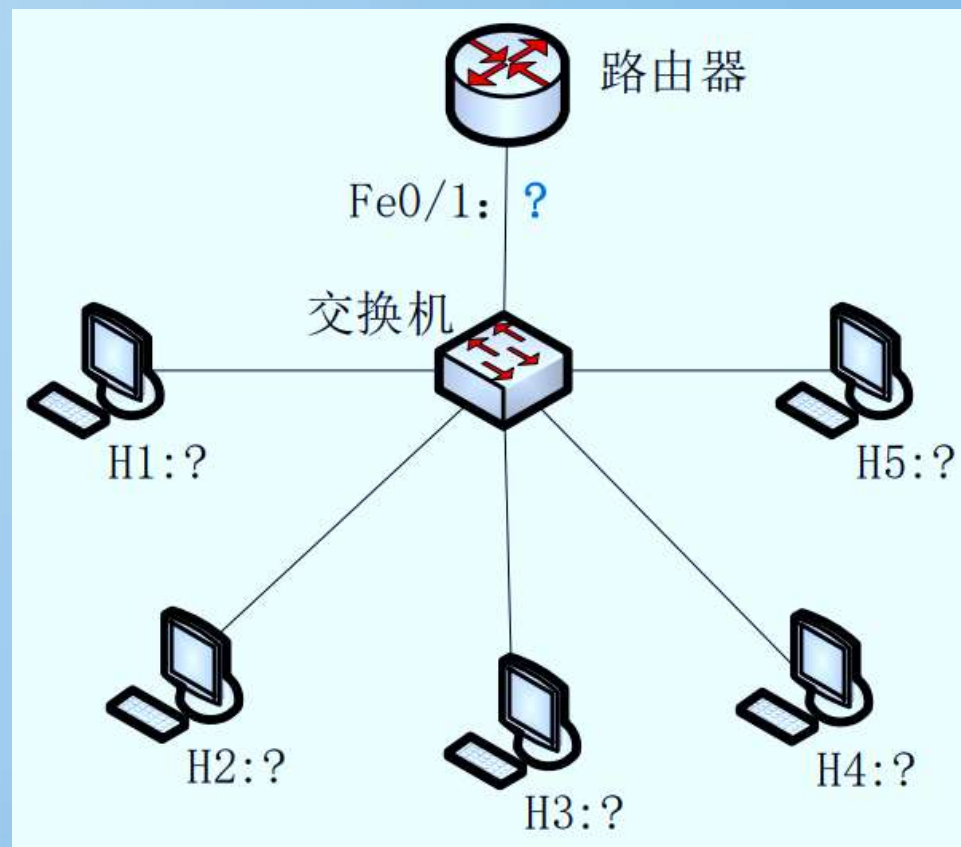
哪有有什么用呢？





例 5-2：IP地址的使用

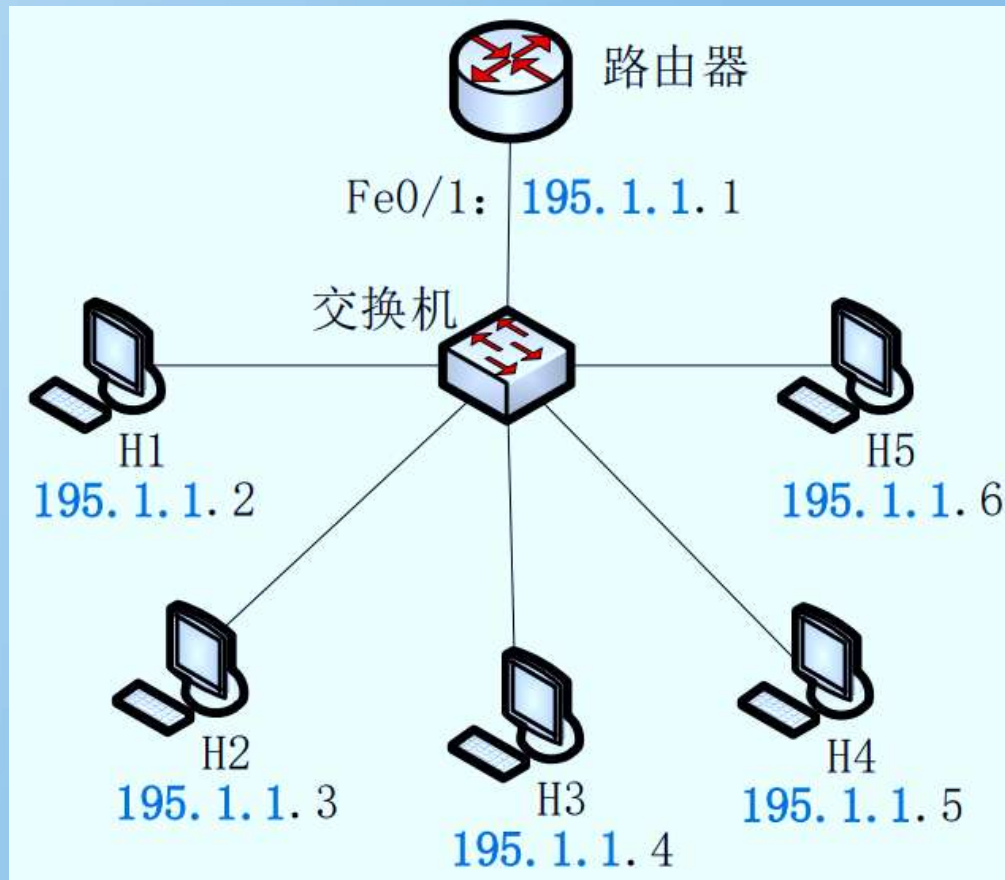
- 一个 C 类网络号 195.1.1 分配给如图 5-13 所示的网络，网络中有 H1、H2、H3、H4、H5 共 5 台主机，这是接入以太网常见的连接方式，请为所有的主机分配一个可使用的合法的 IP 地址。





例 5-2_解答

- C 类网络地址是 195.1.1.0，其中包含 254 个可以使用的合法 IP 地址，即 195.1.1.1~195.1.1.254。
- 路由器的接口 Fe0/1、交换机和 5 台主机共同构成了一个广播域，是一个物理的局域网，习惯称为子网
- 惯例：第一个地址 *.1 分给默认网关 fe0/1 接口





练习

多选题 1、下面是一些主机的IP地址，哪些属于私人地址？

- A. 192.0.2.3
- B. 172.15.1.9
- C. 172.17.9.6
- D. 10.10.10.10
- E. 192.168.5.1

解析：三段私人地址空间是10.0.0.0~10.255.255.255、172.16.0.0~172.31.255.255、192.168.0.0~192.168.255.255。192.0.2.3和172.15.1.9不在这三段地址中，不是私人地址。





练习

2、IP地址4.5.6.7是哪一种类型？

- A. A类
- B. B类
- C. C类
- D. D类

解析：第一个8位组是4，正好落在1-126之间，是A类地址。





练习

3、IP地址229.1.2.3是哪一种类型？

- A. A类
- B. B类
- C. C类
- D. D类

解析：第一个8位组是229，正好落在224-239之间，是D类地址，即组播地址。





练习

4、IP地址191.1.2.3是哪一种类型？

- A. A类
- B. B类
- C. C类
- D. D类

解析：第一个8位组是191，正好落在128-191之间，是B类地址。





练习

5、IP地址169.5.0.0是哪一种类型？

- A. 主机IP
- B. 定向广播地址
- C. 受限广播地址
- D. 网络地址

解析：第一个8位组是169，表明这是一个B类地址，后面2个8位组是主机位，全部为“0”，正是网络地址。





练习

6、IP地址169.5.1.1是哪一种类型？

- A. 主机IP
- B. 定向广播地址
- C. 受限广播地址
- D. 网络地址

解析：第一个8位组是169，表明这是一个B类地址，后面2个8位组是主机位，分别是“00000001”，是一个普通的主机地址。





练习

7、IP地址169.5.255.255是哪一种类型？

- A. 主机IP
- B. 定向广播地址
- C. 受限广播地址
- D. 网络地址

解析：第一个8位组是169，表明这是一个B类地址，后面2个8位组是主机位，全部为“1”，正是定向广播地址。





练习

8、关于IP地址241.1.2.3，哪一个说法是正确的？

- A. 网络部分是241（NetID）
- B. E类地址
- C. 主机部分是1.2.3（HostID）
- D. 网络部分是11110

解析：第一个8位组是241，落在范围240-255内，是E类地址。





练习

9、一台设备有两个IP地址，这台设备可能是以下哪个？

- A. 计算机
- B. 路由器
- C. 网关
- D. 其余选项都是

解析：一台计算机可以安装双网卡，配置2个IP地址；一台路由器通常连接了多个网络，对应的口上（**interface**）都有IP地址，有两个的可能；网关的概念很广，在某种意义上讲，网关也可以是一台路由器，也可以配置两个甚至多个IP地址。所以，都可能。





练习

10、下面哪种地址可以作为源IP地址？

- A. 0.0.0.0
- B. 255.255.255.255
- C. 环回地址127.*.*.*
- D. 0.0.0.64

解析：受限广播地址255.255.255.255只能用作目的地址；环回地址用来测试，也只能用作目的地址；0.0.0.64可能是一个 这个网络的特定主机（**specific host on this network**）地址（即网络部分全为0的地址），被一个网络的主机用于给同一个网络的某个主机发分组，只能用作目的地址。只有0.0.0.0表示这个网络的这个主机，仅可用作源IP地址，在刚启动且还未获得IP地址时使用。





练习

11、一台拥有142.5.0.1地址的主机，想要测试其内部软件，发送的分组的地址可以是下面哪个？（多选）

- A. 127.0.0.0
- B. 127.0.0.1
- C. 127.127.127.127
- D. 以上都可以

解析：环回地址127.*.*.*都可用于测试，但是127.0.0.0是网络地址，既不能用作源地址，也不能用作目的地址。所以，选“127.0.0.1”和“127.127.127.127”





主要内容

- 通信网络及提供的服务 (5.1)
- IPv4协议 (5.2)
 - IP分组和IP地址
 - 子网划分
 - IP寻址和MAC寻址
 - 无类域间路由CIDR
- 与IP有关的其它协议或技术 (5.3)





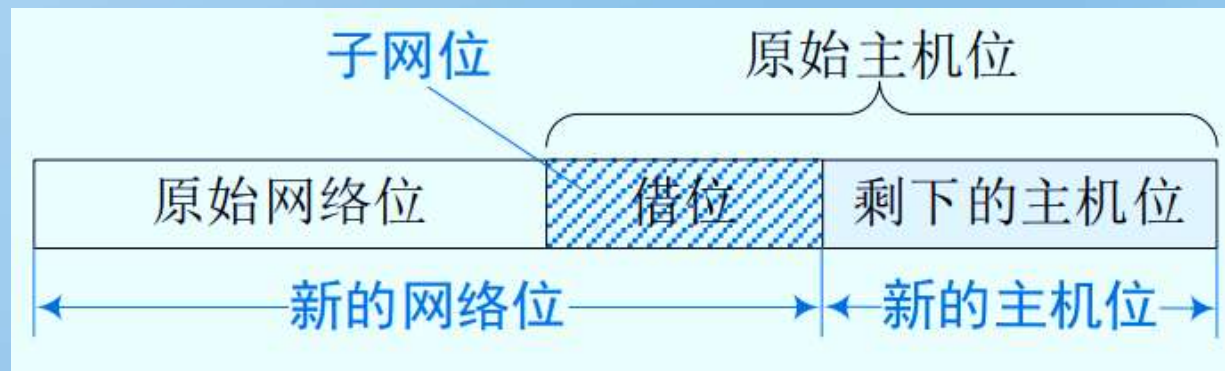
子网规划





子网划分（子网规划）

- 子网划分：把原有 IP 地址块划分为跟物理网络（或 VLAN）大小相适应的 IP 地址块。
- 划分方法：借位
- 借位原则
 - 从主机域的高位开始借位。从主机域的高位开始借位，可以保证所借的位紧紧挨着原网络位，剩下的主机位仍然在是一整块、连续的主机号
 - 主机域至少要保留 2 位





子网掩码_1

- 子网掩码：指示了对应的 IP 地址所在的网络以及这个网络的规模。
- 子网掩码也用 32 位表示
 - 数字“1”表示 IP 地址的对应位是一个网络位，而“0”表示对应位是一个主机位

	25位网络位	7位主机位		25位网络位	7位主机位
IP 地址：	125.216.4.0	0000101	网络地址：	125.216.4.0	0000000=125.216.4.0
子网掩码：	255.255.255.1	0000000	广播地址：	125.216.4.0	1111111=125.216.4.127





子网掩码_2

■ 3类IP地址缺省的子网掩码

	点分十进制表示法	前缀表示法
A 类	255.0.0.0	/8
B 类	255.255.0.0	/16
C 类	255.255.255.0	/24

■ 类C (C-Like) 地址

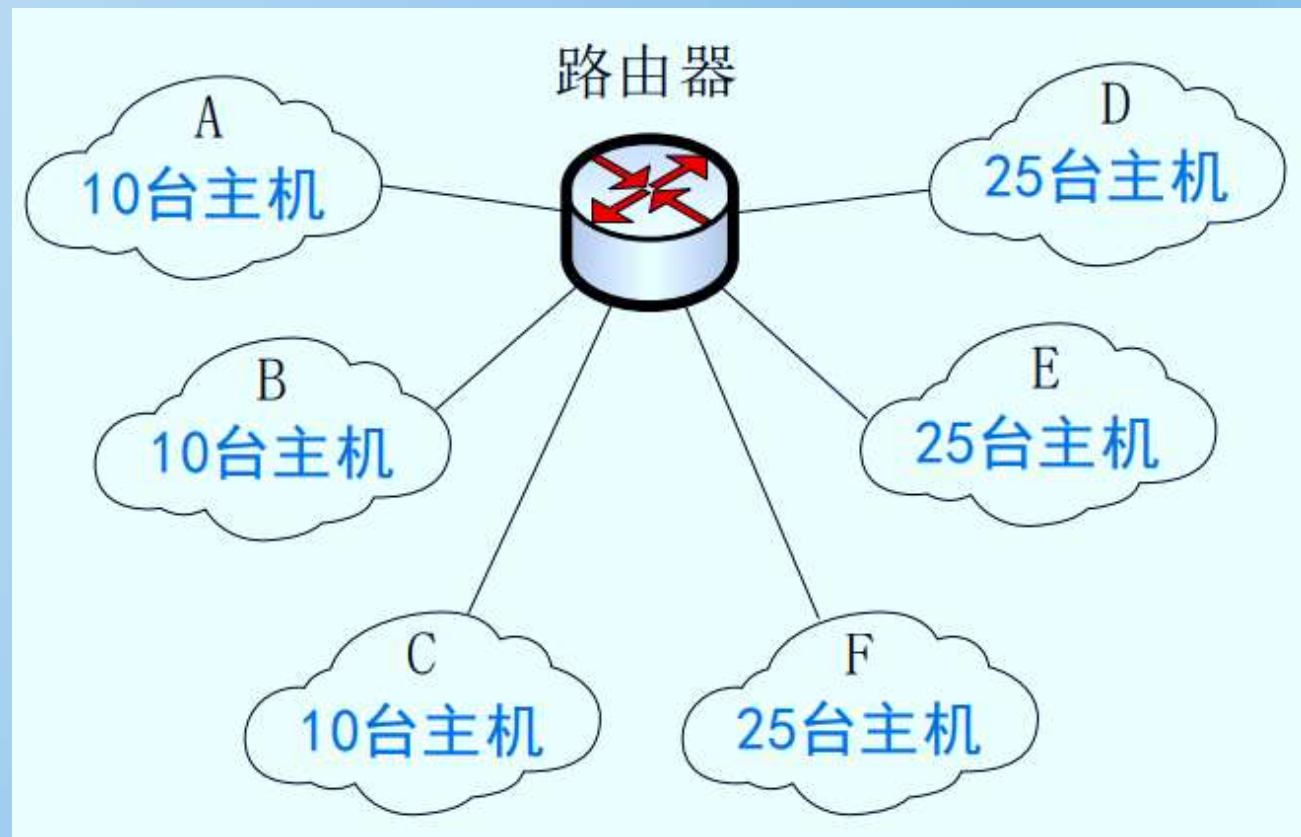
- A或B 类地址通过借位，形成255.255.255.0的子网掩码，C类地址的缺省子网掩码相同
- 比如，138.6.33.0/24 就是一个类 C 地址， 虽然它本是一个 B 类地址





例 5-3

- 一个机构部署了一个网络，包含 6 个物理网络，即 6 个子网，其拓扑和各子网挂载的主机数，如图所示，有一个网络地址块 192.168.1.0/24 可用于该网络，应该如何规划这个地址块？





例 5-3_解答

- 需求分析：6个子网，最多主机数25
- 借3 位才能满足子网数量的需求；剩余 5 位主机位，可以提供30个可用 IP 地址

子网序号	子网掩码	子网络地址	子网络广播地址	可用的地址范围	分配
1	255. 255. 255. 224 (或/27)	192. 168. 1. 0	192. 168. 1. 31	192. 168. 1. 1~30	A
2		192. 168. 1. 32	192. 168. 1. 63	192. 168. 1. 33~62	B
3		192. 168. 1. 64	192. 168. 1. 95	192. 168. 1. 65~94	C
4		192. 168. 1. 96	192. 168. 1. 127	192. 168. 1. 97~126	D
5		192. 168. 1. 128	192. 168. 1. 159	192. 168. 1. 129~158	E
6		192. 168. 1. 160	192. 168. 1. 191	192. 168. 1. 161~190	F
7		192. 168. 1. 192	192. 168. 1. 223	192. 168. 1. 193~222	备用
8		192. 168. 1. 224	192. 168. 1. 255	192. 168. 1. 226~254	备用





注意

- 例5-3解答中，每个子网都剩余多多少少的IP地址，造成浪费
- IP地址的划分，并没有唯一正确的答案，可以很多种划分方法
- 从表格中注意到了一些规律：
 - 每个子网的主机容量是32，即 2^5
 - 每个子网的网络地址，是32的整数（0-7）倍，即192.168.1.0、32、64.....、224
 - 每个子网的广播地址，是本子网的最后一个地址，也是下一个子网的网络地址的前一个地址
 - 每个子网的可用地址范围：本子网的第一个地址和最后一个地址之间，即网络地址和广播地址之间。





子网规划技巧_1

例：一个主机的IP地址是202.112.14.37，掩码是255.255.255.240，要求计算这个主机所在网络的网络地址和广播地址。

解答：

- 1) 从掩码推算子网可容纳的IP地址数量：容纳的IP地址有 $256 - 240 = 16$ 个（包括网络地址和广播地址）；
- 2) 子网网络地址是可容纳IP数量的整数倍，如202.112.14.0、16、32、48.....
- 3) 网络地址 < 子网IP地址 < 广播地址，而广播地址是下一个网络地址减1，如 $32 < 37 < 47$ (48-1)，所以202.112.14.37所在的网络地址和广播地址分别是202.112.14.32和202.112.14.47。





子网规划技巧_2

例：一个C类地址，需要具有10台主机的子网，请进行子网地址的规划和计算子网掩码。

解答：

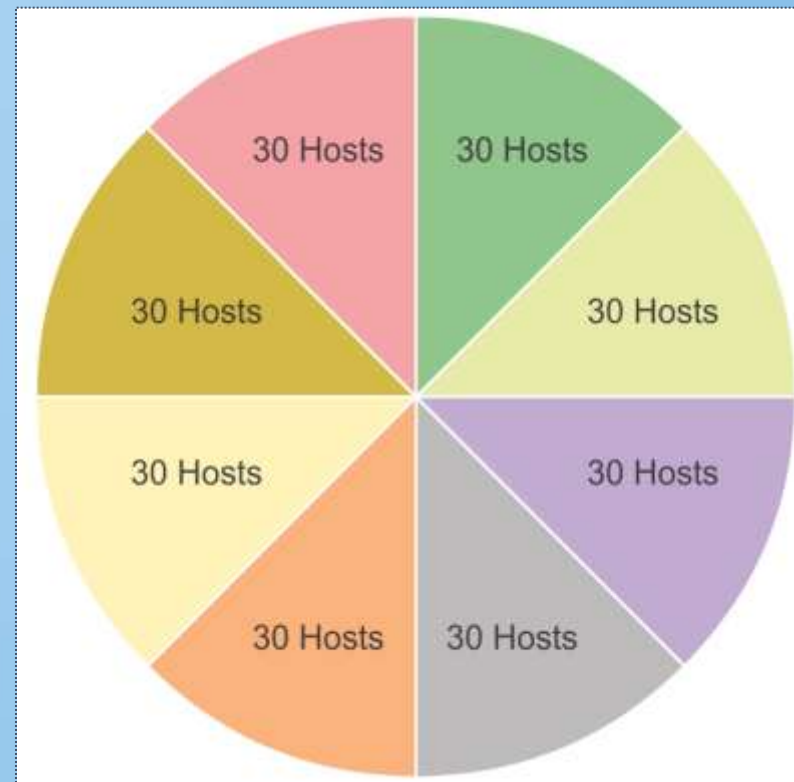
- 1) 计算子网需要的IP地址数量，如 $10+1+1=12$;
- 2) 找到满足大于所需IP数量的最小 2^n ，如 $12 < 16 = 2^4$;
- 3) n 就是主机位数，确定借位，如4位主机位，借位 $8-4=4$;
- 4) 写出子网掩码，如**255.255.255.240**，其中 $240=256-16$ ；或 $8*3+4=28$ ，**"/28"**





传统子网划分 浪费了地址

- **传统子网划分**——为每个子网分配**相同数量**的地址。
- 需要较少地址的子网中存在未使用（**浪费**）的地址。例如，**链路只需要2个地址**。
- **可变长子网掩码（VLSM）**或**细分子网**可以提供更有效的地址使用。





可变长子网子网掩码的子网划分

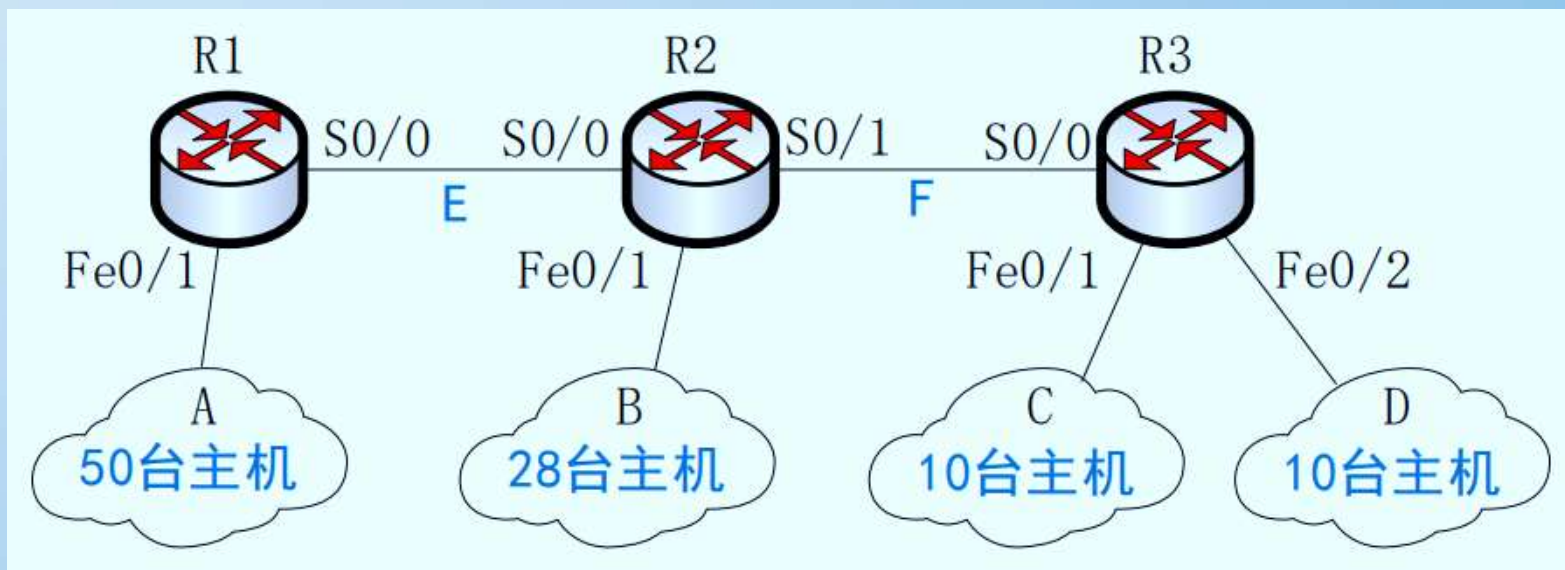
- 可变长子网掩码的子网划分：可以对子网进一步划分，继续借位，创建更小的子网，精打细算，尽量不浪费地址；这样子网位数不再是定长的，而是变化的
 - 子网的规模不同
- 可变长子网掩码子网划分方法
 - 满足子网个数的需求
 - 先满足大规模子网的需求
 - 最后满足小规模子网的需求





例 5-4

- 一个机构部署了一个企业网络，由三台路由器连接了 A、B、C、D、E、F 共 6 个物理网络，每个子网的规模其拓扑如图 5-17 所示；有一个网络地址块 192.168.6.0/24 可用于该网络，应该如何规划这个地址块？



定长子网掩码无法
满足需求！





例 5-5解答

- 需求分析：网络 A、B、C 和 D 分别需要 51 个、29 个、11 个和 11 个可用的 IP 地址
- 最大的子网A需要主机位6位才能满足需求，所以，首先借2位，创建4个子网，第1个子网分配给A
- 将第2个子网继续划分，借1位，生成2个子网，其中一个分配给B
 - 剩下的那个子网继续划分，再借1位，生成2个子网，正好满足C和D的需求
- 将第3个子网192.168.6.128/26继续划分.....



192.168.6.0/24

.XX000000

192.168.6.0/26

192.168.6.64/26

192.168.6.128/26

192.168.6.192/26

.XXX00000

192.168.6.64/27

192.168.15.96/27

.XXXX0000

192.168.15.96/28

192.168.15.112/28

子网 A 50
台主机

子网 B
28 台主机

子网 C
10 台主机

子网 D
10 台主机

(续)





IP地址的分配

■ 互联网号码分配机构IANA

□ 总地址池的分配

■ 5大地区网络信息中心

□ AFRINIC、APNIC、ARIN、LACNIC、RIPE NCC

■ 中国互联网络信息中心CNNIC



REGISTRY	AREA COVERED
AFRINIC	Africa Region
APNIC	Asia/Pacific Region
ARIN	Canada, USA, and some Caribbean Islands
LACNIC	Latin America and some Caribbean Islands
RIPE NCC	Europe, the Middle East, and Central Asia





一台主机/设备如何获取IP地址

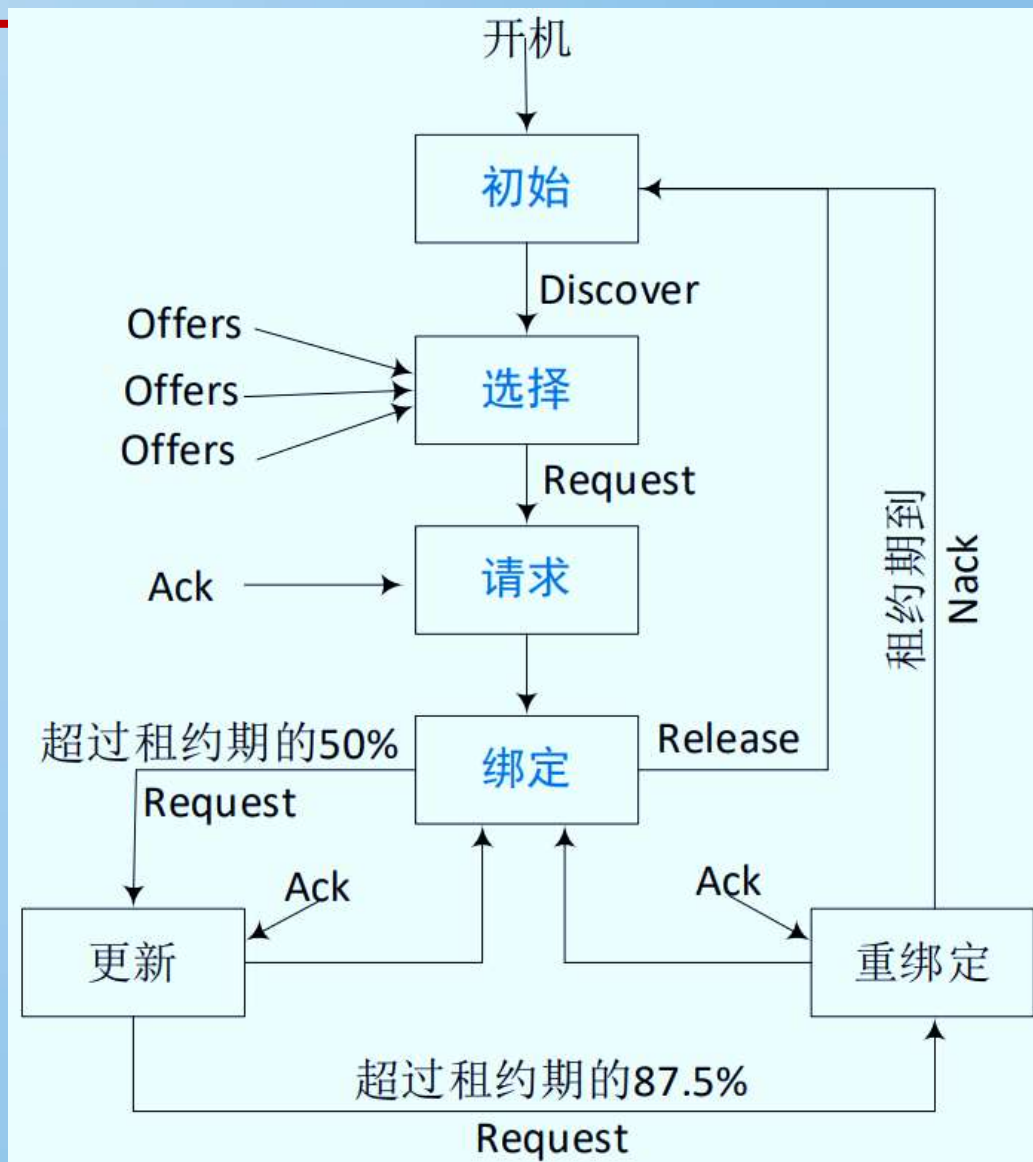
- 逆向地址解析协议RARP，主机从 RARP 服务器处获得一个与自己 MAC 地址对应的 IP 地址
 - 早期的无盘工作站使用，已淘汰
- 引导程序协议BOOTP
 - 个客户/服务器协议，服务器可放在网络的任何地方；同时，服务器提供包含 IP 地址在内的上网所需要的主要信息
- 动态主机配置协议
 - BOOTP的改进版本，目前主要使用的方法





动态主机配置协议

- 初始状态：开机启动，发送Discover消息
- 选择状态：选择Offer消息，并发出Request消息
- 请求状态：接收ACK
- 绑定状态：使用服务器提供的IP地址等上网资源
- 更新状态、重绑定状态





IP寻址





IP寻址和MAC寻址

- IP 寻址指的是根据 IP 地址定位到目的机的过程，IP 寻址类似于快递投递过程
- MAC 寻址是根据 MAC 地址定位到目的机的过程，MAC 寻址发生在较小的网络，通常是广播式网络
- IP寻址和MAC两种寻址方式，不是对立的，而是相辅相成的
 - IP 寻址将数据分组推向目的机所在的网络
 - MAC寻址，将数据分组推向目的网络中的目的主机





IP寻址和MAC寻址的区别_1

- 适用的网络范围不同
 - IP 寻址适用于整个互联网络；而 MAC 寻址就发在子网中
- 两种地址在数据报文中的位置不同
 - MAC 寻址所依赖的目的 MAC 地址在帧头部；IP 寻址所依赖的目的 IP 地址在分组的头部
- 寻址的地址依据不同
 - IP 寻址依据的是 IP 地址不可或缺的特性；而 MAC 寻址依据的是 MAC 地址，它是平面地址，在小网络范围内广播式找寻、定位目的机。





IP寻址和MAC寻址的区别_2

■完成寻址的交换设备不同

□进行 IP 寻址的主要是路由器，而进行MAC寻址的主要是交换机

□路由器进行IP寻址的过程

■接收分组，提取目的IP地址，确定目的网络

■查找路由表

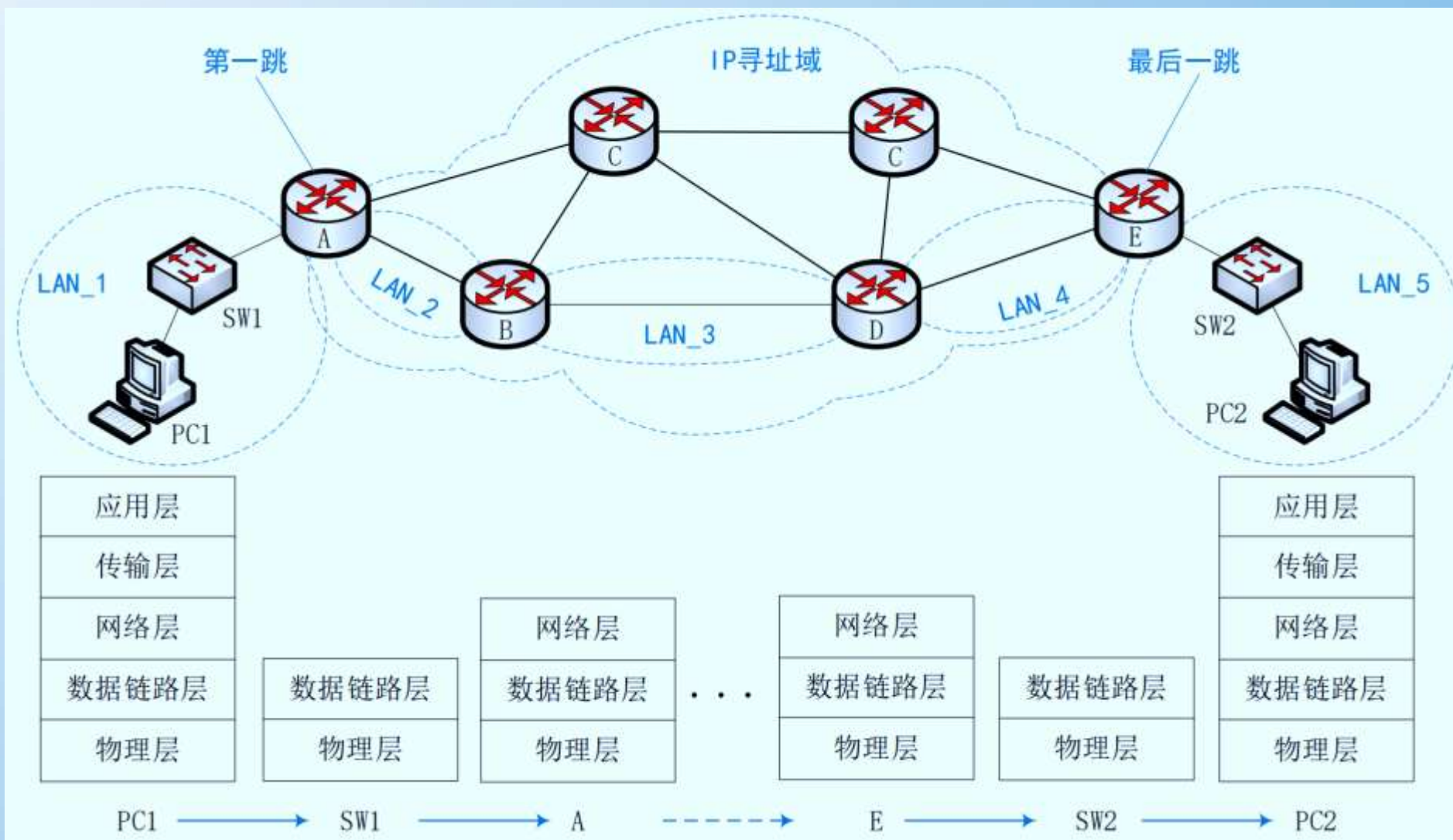
■重新封装、转发

目的IP地址:	01111101.11011000.00110001.00100011
子网掩码: &	11111111.11111111.11111111.00000000
<hr/>	
“与”结果:	01111101.11011000.00110001.00000000





IP寻址和MAC寻址共同完成数据的交付





无类域间路由的提出

- 1990s年代，互联网面临的问题
 - B 类地址空间耗尽
 - 互联网路由器的路由表规模增长，超出了当时的软件、硬件和有效管理能力
 - 32 位的 IPv4 地址空间也终将全部耗尽（实际上，2011 年 2 月，总地址池已经枯竭）
- 无类域间路由 CIDR 并不试图解决第三个问题，而是通过提供一种机制去降低 IPv4 地址的耗尽速度以及路由表的增长速度





CIDR提出按需分配IP地址

- 按需分配地址块：不再按照A\B\C地址类别划分
- 【例 5-5】一个 1000 人的公司，申请 1000 个主机 IP 地址。按照 CIDR 的无类编址机制为该公司提供地址分配
 - 解答：首先确定主机位、找到大于 1000 且最靠近 1000 的 2^n 。
此时 $n=10$ ，表示用 10 位主机位的 IP 地址，即可满足该公司的地址数需求，只有少量余量，可供后续扩展之用。
 - 网络位数=总位数-主机位数=32-10=22，所以只需要分配“/22”地址块，就满足了公司的 IP 地址数需求。





例 5-6

- 一个互联网服务提供商 ISP_1 获得了一个 IP 地址块 **202.112.0.0/20**，现在有 4 所学校 A、B、C 和 D 向 ISP_1 分别申请 IP 地址数 1000、1000、500 和 500，ISP_1 为 4 所学校分配满足需求的 IP 地址。





例 5-6_解答

- 地址块 202.112.0.0/20 的网络位是 20 位，含 4096 个 IP 地址
- A 学校分配了网络位为 22 的 IP 地址块 202.112.0.0/22，其中，第一个 IP 地址是网络地址 202.112.0.0，最后一个 IP 地址是广播地址 202.112.3.255，共提供 1024 个 IP 地址，满足需求。
- B 学校分配了网络位为 22 的 IP 地址块 202.112.4.0/22
- C 学校分配了网络位为 23 的 IP 地址块 202.112.8.0/23
- D 学校分配了网络位为 23 的 IP 地址块 202.112.10.0/23
- 余下的地址块为 202.112.12.0/22，备用。





常见的CIDR地址块

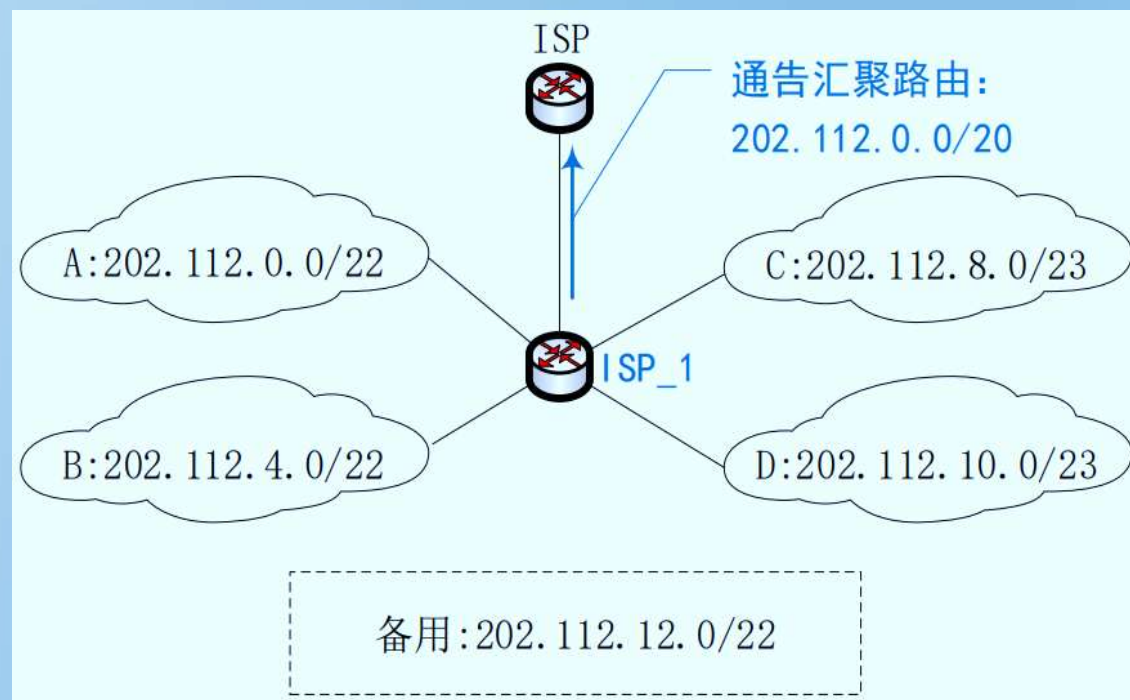
CIDR 地址块	子网掩码	IP 地址数	主机数	备注
/13	255. 248. 0. 0	524288	524286	2048 个 C
/14	255, 252. 0. 0	262144	262142	1024 个 C
/15	255. 254. 0. 0	131072	131070	512 个 C
/16	255. 255. 0. 0	65536	65534	256 个 C
/17	255. 255. 128. 0	32768	32766	128 个 C
/18	255. 255. 192. 0	16384	16382	64 个 C
/19	255. 255. 224. 0	8192	8190	32 个 C
/20	255. 255. 240. 0	4096	4094	16 个 C
/21	255. 255. 248. 0	2048	2046	8 个 C
/22	255. 255. 252. 0	1024	1022	4 个 C
/23	255. 255. 254. 0	512	510	2 个 C





CIDR：路由汇聚

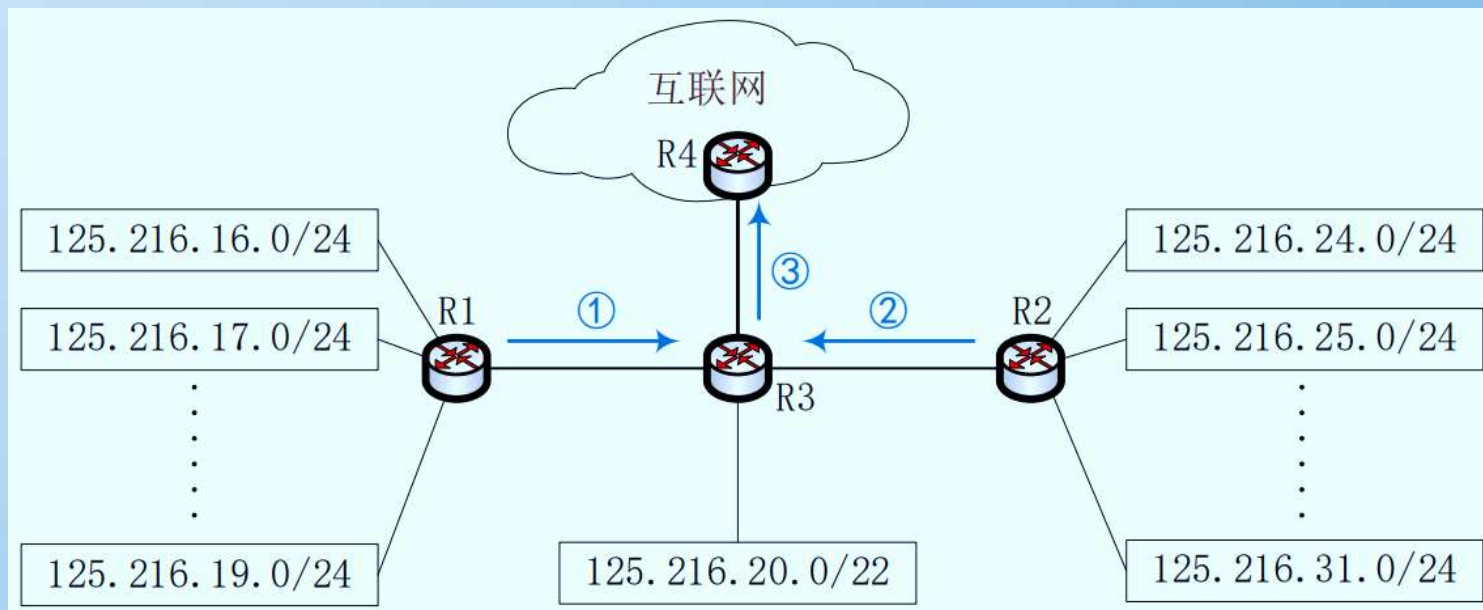
- 路由汇聚/汇总
- 汇聚结果：超网，采用“基地址/网络位数”表示
 - 其中的基地址，指的是形成的超网地址空间中的第一个地址
- 汇聚方法
 - 网络位：地址空间中不变的位





例 5-7

- 一个网络的拓扑如图所示，路由器 R1 直接连接了 4 个 “/24” 的网络，路由器 R2 直接连接了 8 个 “/24” 的网络，路由器 R3 连接了 1 个 “/22” 的直连网络，还连接了路由器 R1 和 R2；R1 向 R3 通告时发生了路由汇聚，用①表示；R2 向 R3 通告时发生了路由汇聚，用②表示；R3 向 R4 通告时，也发生了路由汇聚，用③表示。请分别写出汇聚①、②、③发生后产生的超网。





例 5-7_解答

- ①4个子网汇聚：25.216.16.0/22

不变的位		
16:	000100	00
17:	000100	01
18:	000100	10
19:	000100	11

- ②8个子网汇聚：125.216.24.0/21

不变的位		
24:	00011	000
25:	00011	001
:	:	
:	:	
31:	00011	111

- ③2个超网和1个直连子网汇聚
: 125.216.16.0/20

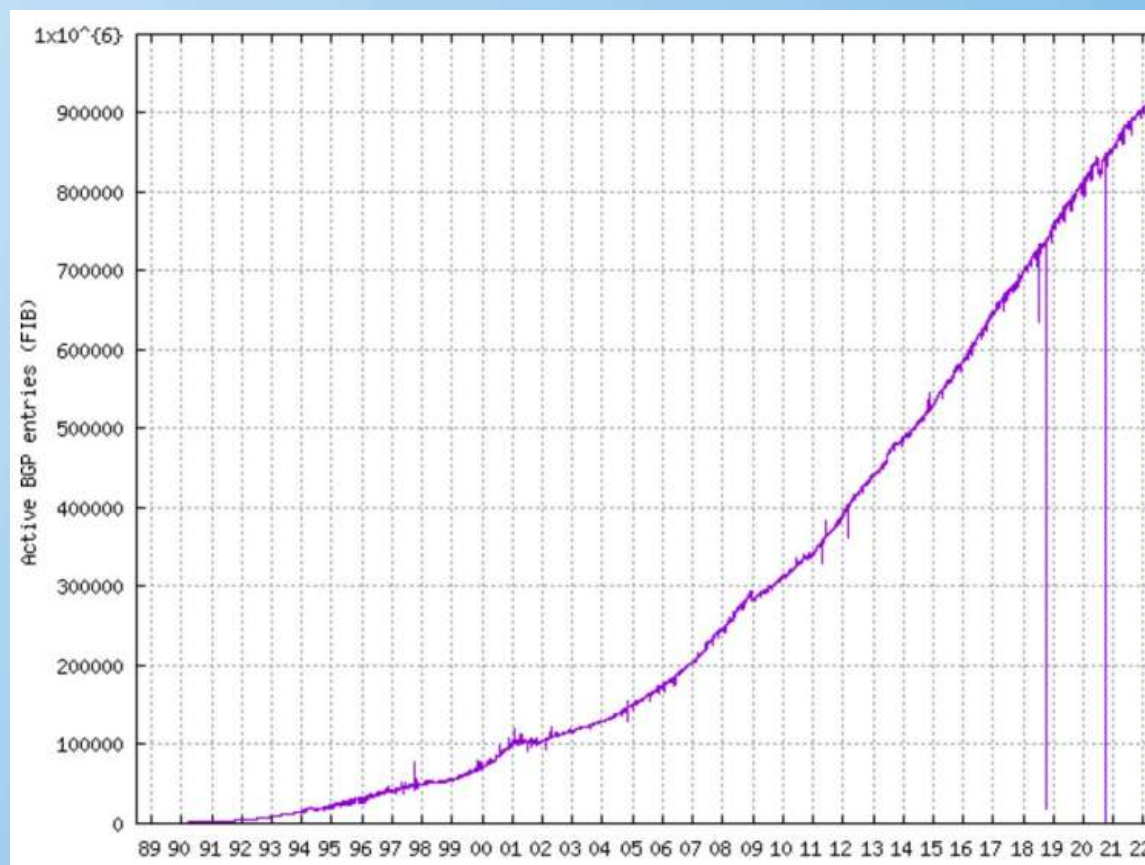
不变的位		
16:	0001	0000
17:	0001	0001
:	:	
20:	0001	0100
:	:	
:	:	
31:	0001	1111





CIDR的部署及效果

- 采用 CIDR 地址分配机制之后，B 类地址枯竭的趋势得到了遏制（虽然没有改变地址耗尽的命运），同时，CIDR 也改写了路由表的指数增长趋势





最长地址前缀匹配

■ 路由器收到了一个分组，其目的 IP 地址是 202.112.12.9

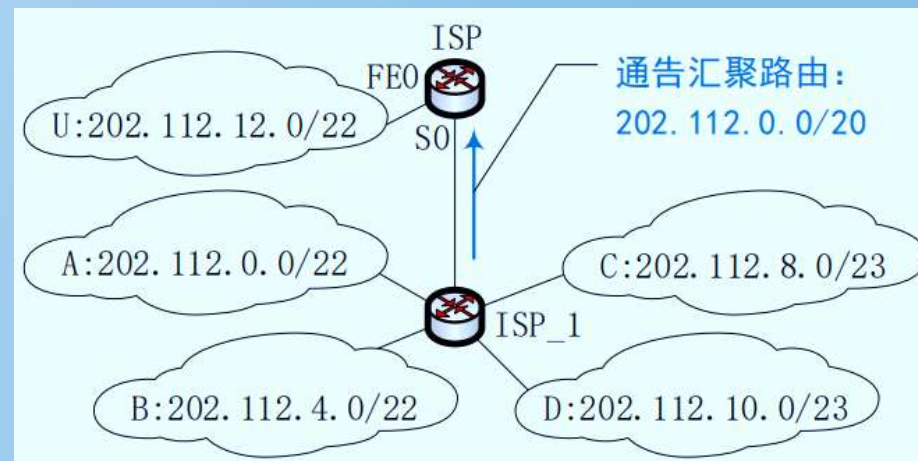
■ 匹配两个目的网络

目的 IP: 202.112.00001100.00001001
子网掩码: & 255.255.11111100.00000000

“与”结果: 202.112.00001100.00000000
目的网络地址: 202.112.12.0

目的 IP: 202.112.00001100.00001001
子网掩码: & 255.255.11110000.00000000

“与”结果: 202.112.00000000.00000000
目的网络地址: 202.112.0.0



目的网络	子网掩码	转出接口
202.112.12.0	/22 (255.255.252.0)	FE0
202.112.0.0	/20 (255.255.240.0)	S0
.....		





练习

1、下面列出的前缀中，哪一个是匹配 136.127.4.1 的最长前缀网络地址？

- A. 136.127.4.0/24
- B. 136.127.4.128/25
- C. 136.127.0.0/16
- D. 136.0.0.0/8

1、对于一个特定的IP地址，它一定属于某个小网络，小网络又属于大些的中网络，中网络属于更大些的网络.....，换言之，这个IP属于所有这些网络。最长前缀的网络是这些网络中最小的那个。136.127.4.1同时在136.0.0.0/8、136.127.0.0/16和136.127.4.0三个网络中，但是最长前缀是“24”。136.127.4.128/25的前缀“25”比“24”更长，但是它包含的全部IP范围是136.127.4.128~255，不包含136.127.4.1。





练习

2、下面列出的前缀中，哪一个是匹配 136.127.4.1 的最长前缀？

- A. 136.127.0.0/21
- B. 136.127.4.0/21
- C. 136.127.0.0/22
- D. 136.127.0.0/20

2、136.127.00000100.00000001/20 所属网络的网络地址是：136.127.0.0/20，但 $20 < 21$ ；
136.127.00000100.00000001/22 所属网络的网络地址是：136.127.4.0/22，答案 136.127.0.0/22 不正确；
136.127.00000100.00000001/21 所属的网络地址是 136.127.0.0/21（将 11 位主机位置为 0），且 $21 > 20$ 。





练习

3、如果有一个子网的网络前缀是 192.168.176.0/20， 那么该子网中 IP 地址的最大数目是多少？

- A. 12^2
- B. 2^{12}
- C. 2^{16}
- D. 2^{20}

3、前缀 20 表明网络位有20位，那么主机位有 $32-20=12$ 位，能够提供的IP地址数是 2^{12} 。





练习

4、某单位申请到了一个网络地址：12.24.0.0/17，需要划分为4个相同大小的子网。每一个子网的前缀是什么？

- A. 12.24.0.0/18, 12.24.64.0/18, 12.24.128.0/18, and 12.24.192.0/18
- B. 12.24.0.0/18, 12.24.32.0/18, 12.24.64.0/18, and 12.24.96.0/18
- C. 12.24.0.0/19, 12.24.32.0/19, 12.24.64.0/19, and 12.24.96.0/19
- D. 12.24.0.0/19, 12.24.64.0/19, 12.24.128.0/19, and 12.24.192.0/19

4、4个子网，需要的子网位数是2，所以，从12.24.0.0/17中的主机位借位，总网络位变为 $17+2=19$ 。从高位数起，第18、19位有4种组合00、01、10、11，分别对应于4个子网，即：12.24.0.0/19, 12.24.32.0/19, 12.24.64.0/19, and 12.24.96.0/19。





练习

5、如果一台路由器内部的路由表中，有如下的表项。问：一个目的IP地址为135.46.70.10的数据分组到达了该路由器，应从哪里转发该分组？

- A. Interface C
- B. Interface D
- C. Interface A
- D. Interface B

Address/mask	Next hop
135.46.64.0/21	Interface A
135.46.68.0/23	Interface B
135.46.64.0/20	Interface C
135.46.70.0/23	Interface D

解析：目的IP地址135.46.70.10同时属于这几个网络：135.46.64.0/20、135.46.64.0/21和135.46.70.0/23，但是最长前缀是“23”，所以在路由表中选择135.46.70.0/23网络对应的接口进行转发。





练习

6、一个带有子网掩码为255.255.252.0的子网172.16.128.0，该子网中有多少个可用的（可分配给接口的）IP地址？（From CCNA）

- A. 510
- B. 512
- C. 1022
- D. 1024
- E. 2046
- F. 2044

解析：子网掩码255.255.252.0表示网络位数是 $8*2+6=22$ ，对应的主机位是 $32-22=10$ ，所以，一个这样的网络提供的IP数量是 $2^{10}=1024$ ，扣除广播地址和网络地址，剩下的1022个地址是合法的可分配给主机使用的。





练习

7、有一个网络，拓扑示意图如下所示，现有一个IP网络地址128.107.0.0/16可用，每个子网容量相同，有足够的地址可用，且需要保留有最多的子网数，那做了子网划分后，子网掩码应该是多少？(From CCNA)

- A. 255.255.255.0
- B. 255.255.255.128
- C. 255.255.255.192
- D. 255.255.255.224
- E. 255.255.255.240

解析：图中，最大的子网里有100台主机，所以，主机位保留7位即可满足所有子网对IP地址的需求，每个子网的容量是 $2^7=128$ ，可用IP地址数是 $126>100$ ，子网个数 $2^{(16-7)}=512$ ；子网掩码255.255.255.0 中有8位主机位，也可以满足子网的主机地址需求，但是提供的子网数是 $2^{(16-8)}=256$ ，不满足保留最大子网数目的需求。





练习

7、有一个网络，拓扑示意图如下所示，现有一个IP网络地址128.107.0.0/16可用，每个子网容量相同，有足够的地址可用，且需要保留有最多的子网数，那做了子网划分后，子网掩码应该是多少？(From CCNA)

- A. 255.255.255.0
- B. 255.255.255.128
- C. 255.255.255.192
- D. 255.255.255.224
- E. 255.255.255.240

解析：图中，最大的子网里有100台主机，所以，主机位保留7位即可满足所有子网对IP地址的需求，每个子网的容量是 $2^7=128$ ，可用IP地址数是 $126>100$ ，子网个数 $2^{(16-7)}=512$ ；子网掩码255.255.255.0 中有8位主机位，也可以满足子网的主机地址需求，但是提供的子网数是 $2^{(16-8)}=256$ ，不满足保留最大子网数目的需求。





练习

8、一个公司有一个网络地址192.168.1.64，对应的子网掩码是255.255.255.192，这个公司想创建2个子网，1个可容纳10台主机，另外一个可容纳18台主机，下面哪两个网络地址可用于这两个子网？（选其中两项）（From CCNA）

- A. 192.168.1.16/28
- B. 192.168.1.64/27
- C. 192.168.1.128/27
- D. 192.168.1.96/28
- E. 192.168.1.192/28

解析：采用可变长子网掩码（VLSM）。子网掩码255.255.255.192对应的网络位数是 $8*3+2=26$ ，如果采用定长子网掩码，只需要借用1位，创建两个子网192.168.1.64/27和192.168.1.128/27；可以满足需求，但是没有192.168.1.96/27这个选项。且192.168.1.96/27提供的子网数是32，大大超过了需求，所以对起进行进一步规划，划分为192.168.1.96/28和192.168.1.112/28两个子网，其中之一分配使用，另外一个备用。





练习

9、在下面这些点分十进制表示的IPv4地址中，哪一个子网172.25.0.64/26中的最后一个可用地址？

- A. 172.25.0.94
- B. 172.25.0.95
- C. 172.25.0.126
- D. 172.25.0.127

解析：网络地址 172.25.0.64 的最后1个8位组 的二进制表示是：01000000，其低6位是主机位，可从6个0变化到6个1，最后一个可用地址的主机部分应该是111110，那最后一个可用地址的最后1个8位组的二进制是01111110，对应的十进制数是126，所以，172.25.0.126是个网络地址空间中的最后一个可用IP地址。





练习

10、某主机的IP地址是180.80.77.55，子网掩码是255.255.252.0，若该主机向其所在子网发送广播分组，则目的地址可以是什么？（2012考研真题）

- A. 180.80.76.0
- B. 180.80.76.255
- C. 180.80.77.255
- D. 180.80.79.255

解析：从子网掩码 255.255.252.0，我们知道主机位是10位，所以，180.80.77.55是网络180.80.76.0/22中的一个，它的广播地址是：180.80.79.255。这个题的答案还可以是255.255.255.255这个受限广播地址，只是这里没有这个选项。





练习

11、将255和15按位“与”（AND）之后的结果是哪个？

- A. 255
- B. 15
- C. 0
- D. 不确定

解析：255=11111111，任何数与“1”进行按位“与”，结果就是那个数。





练习

12、将0和15按位“与”（AND）之后的结果是哪个？

- A. 255
- B. 15
- C. 0
- D. 不确定

解析：任何数与“0”进行按位“与”，结果都是0。





练习

13、将192和65按位“与”（AND）之后的结果是哪个？

- A. 192
- B. 65
- C. 0
- D. 64

解析：192=11000000，65=01000001，按位“与”后得到01000000=64





练习

14、下面哪个不是有效的子网掩码？

- A. 255.255.255.254
- B. 255.255.224.0
- C. 255.148.0.0
- D. 255.248.0.0

解析：148=10010100，“1”不连续，子网掩码是由连续的“1”紧跟连续的“0”而成。





练习

15、下面哪个不可能是A类地址块下的子网掩码？

- A. 255.255.255.192
- B. 255.248.0.0
- C. 248.0.0.0
- D. 255.255.128.0

解析：A类地址的缺省子网掩码是255.0.0.0，即“/8”，其下的子网，网络位必然大于8（因为要从主机位借位而增大网络位的数目），其它三个子网掩码的网络位都大于“8”，只有“248.0.0.0”等同于“/5”，网络位小于5，不可能是A类地址块下的子网。





练习

16、下面哪个不可能是B类地址块下的子网掩码？

- A. 255.255.255.192
- B. 255.248.0.0
- C. 255.255.224.0
- D. 255.255.255.129

解析：B类地址的缺省子网掩码是255.255.0.0，即“/16”，其下的子网，网络位必然大于16（因为要从主机位借位而增大网络位的数目），其它三个子网掩码的网络位都大于“16”，只有“255.248.0.0”等同于“/13”，网络位小于16，不可能是B类地址块下的子网。





练习

15、下面哪个不可能是A类地址块下的子网掩码？

- A. 255.255.255.192
- B. 255.248.0.0
- C. 248.0.0.0
- D. 255.255.128.0

解析：A类地址的缺省子网掩码是255.0.0.0，即“/8”，其下的子网，网络位必然大于8（因为要从主机位借位而增大网络位的数目），其它三个子网掩码的网络位都大于“8”，只有“248.0.0.0”等同于“/5”，网络位小于5，不可能是A类地址块下的子网。





练习

16、下面哪个不可能是C类地址块下的子网掩码？

- A. 255.255.255.192
- B. 255.255.224.0
- C. 255.0.0.0
- D. 128.0.0.0

解析：C类地址的缺省子网掩码是255.255.255.0，即“/24”，其下的子网，网络位必然大于24（因为要从主机位借位而增大网络位的数目），子网掩码“255.255.224.0”、“255.0.0.0”、“128.0.0.0”对应的网络前缀分别是“/19”、“/8”、“/1”，都小于“24”，不可能是C类地址块下的子网。只有“255.255.255.192”等同于“/26”，大于“24”，是C类地址块下的子网。





练习

17、下图中，为了让两台主机PC A和PC B能够访问互联网，并且让路由器的CPU开销和网络带宽利用最少，应该做什么样的路由配置方案？（Edge 和 R1 是两台路由的名字）

- A. 配置一条从R1到Edge的静态路由和一条从Edge到R1的动态路由。
- B. 配置一条从R1到Edge的默认路由、一条从Edge到Internet的默认路由以及一条从Edge到R1的静态路由。
- C. 配置一条从R1到Edge的动态路由和一条从Edge到R1的动态路由。
- D. 在R1和Edge之间配置一个动态路由选择协议，并通告所有的路由。

解析：不选择有动态路由的两个选项，静态路由的运行需要消耗带宽、CPU等资源，而静态路由的开销要小得多，默认路由是静态中的一种，常用于末节网络。此网络拓扑中，为了达到目的，必须创建两条路由：一条从R1到Edge的缺省路由，一条Edge到R1的静态路由作为返回路径，这是对类似PC A和PC B所处的末节/存根（stub）网络最好的解决方法。





练习

18、如果将18.31.128.0/17和18.31.0.0/17聚合成一个单一前缀，这个前缀应是什么？

- A. 18.31.0.0/18
- B. 18.31.0.0/17
- C. 18.31.0.0/16
- D. 18.31.128.0/16

解析：聚合的关键是要确定连续的地址空间中的不变的位，即 网络位。18.31.0.0/17包含了从18.31.0.0到18.31.127.255的地址，而18.31.128.0/17包含了从18.31.128.0到18.31.255.255的地址。合起来正好是一大块连续的地址空间18.31.0.0~18.31.255.255。显然，前两个8位组18.31在整个空间中没有发生任何变化，是固定的，是网络位，有8+8=16位。而后两个8位组可从全零变化为全1，是主机位。所以，聚合后的地址“基地址/网络位数”表示为“18.31.0.0/16”。





练习

19、有3个网络57.6.128.0/21、57.6.136.0/21和57.6.144.0/20，汇聚为一个网络，汇聚后的超网可以表示为下面哪一个？

- A. 57.6.128.0/19
- B. 57.6.128.0/20
- C. 57.6.0.0/19
- D. 57.6.128.0/21

解析：汇聚的关键是找到汇聚后连续地址空间中的网络位数和第一个IP地址，因为汇聚后的超网表示为“基地址/网络位数”。前两个8位组是确定无疑的网络位，最后一个8位组也是确定无疑的主机位，关键是第3个8位组。128=10000000，136=10001000，144=10010000，只有高3位是固定的“100”，所以，汇聚后的超网网络位是16+3=19，且第一个地址是57.6.128.0，所以表示为：57.6.128.0/19。





主要内容

- 通信网络及提供的服务 (5.1)
- IPv4协议 (5.2)
- 与IP有关的其它协议或技术 (5.3)
 - 地址解析协议ARP
 - 互联网消息控制协议ICMP
 - 网络地址转换NAT





地址解析协议ARP

- ARP: Address Resolution Protocol, ARP
- ARP是一个将 IP 地址映射为 MAC 地址的协议
- ARP 的任务是找到一个给定IP地址 (Who?) 所对应的MAC地址
- ARP定义于 RFC 826。

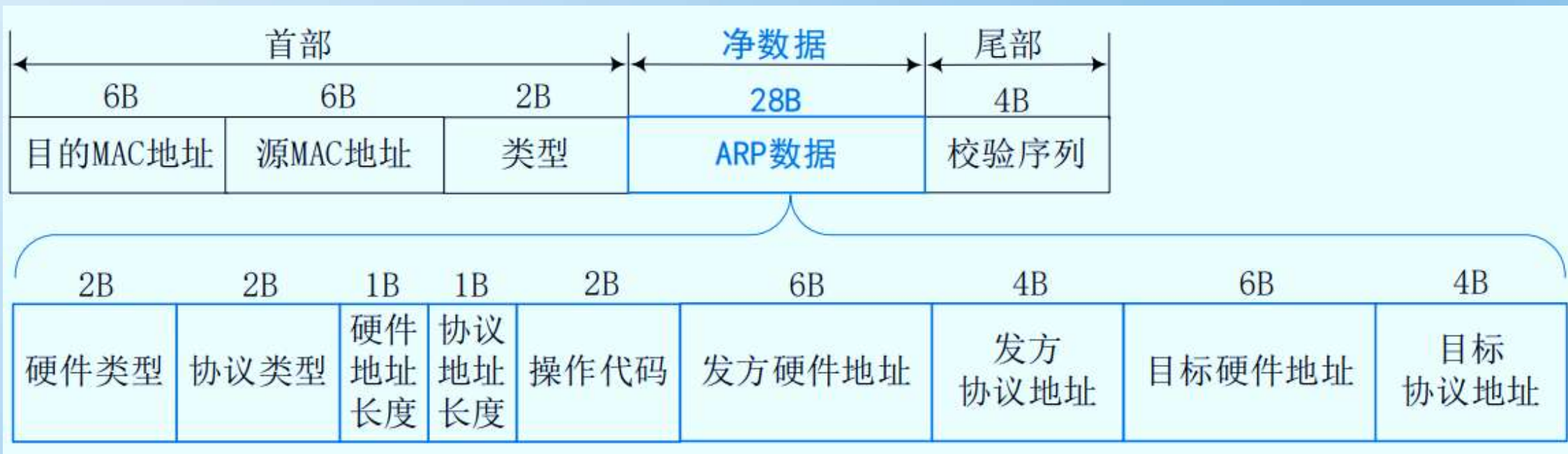
问题：为什么需要地址解析？





ARP帧

- ARP请求帧：广播帧，目的MAC地址是FFFF-FFFF-FFFF
- ARP应答帧：单播帧

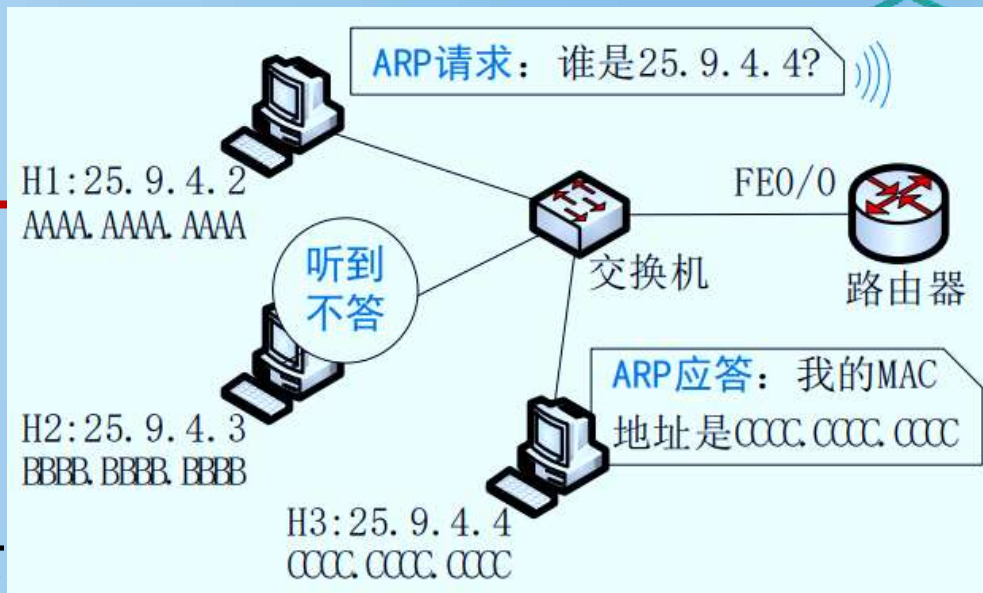


ARP基本工作原理

■基本原理：请求-应答

■H1请求H3的过程

- H1 封装 ARP 请求帧，并将其广播出去
- 所有跟 H1 位于同一子网的主机都会收到这个 ARP 请求
- 除了 H3，其它非目标机收到不理睬
- H3 发出 ARP 应答并单播给 H1
- H1 收到应答并获得目标机的 MAC 地址





ARP表

- 如每次发信息都发起请求-应答过程，不仅效率低，也浪费资源
- ARP表：存储着 IP地址和它的 MAC 地址这一对映射
- ARP表的建立更新和维护
 - 一个发出 ARP 请求的节点通过提取收到的 ARP 应答中的发方 IP/MAC 地址映射对，并打上使时戳
 - 一个收到 ARP 请求的目标节点通过提取 ARP 请求中的发方 IP/MAC 地址映射对，并打上使时戳
 - 一个收到广播 ARP 请求的非目标节点，可通过提取发方的 IP/MAC 地址映射对，并打上使时戳





免费ARP (Gratuitous ARP P364)

■免费ARP

- 当一台主机启动时，发送一个免费ARP请求，**请求自己**IP的MAC地址（如果意外收到一个应答，即是IP地址发生了**冲突**）
- 当一个接口（interface）的配置发生了改变，会发送一个免费ARP

■例子：一台主机 (172.16.1.1, 0002 4A87 0D92) 发送的免费ARP





一个免费ARP帧的样例

- A host (172.16.1.1, 0002 4A87 0D92)

Ethernet II

0	4	8	14	19	Bytes
PREAMBLE: 101010...1011		DEST MAC: FFFF.FFFF.FFFF		SRC MAC: 0002.4A87.0D92	
TYPE: 0x806		DATA (VARIABLE LENGTH)		FCS: 0x0	

ARP

0	8	16	31	Bits
HARDWARE TYPE: 0x1		PROTOCOL TYPE:		
HLEN: 0x6		PLEN: 0x4		OPCODE: 0x1
SOURCE MAC: 0002.4A87.0D92 (48 bits)				SOURCE IP (32 bits)
172.16.1.1				
TARGET MAC: 0000.0000.0000 (48 bits)				
TARGET IP: 172.16.1.1 (32 bits)				

**Source IP
==
target IP**





免费ARP的一个功能

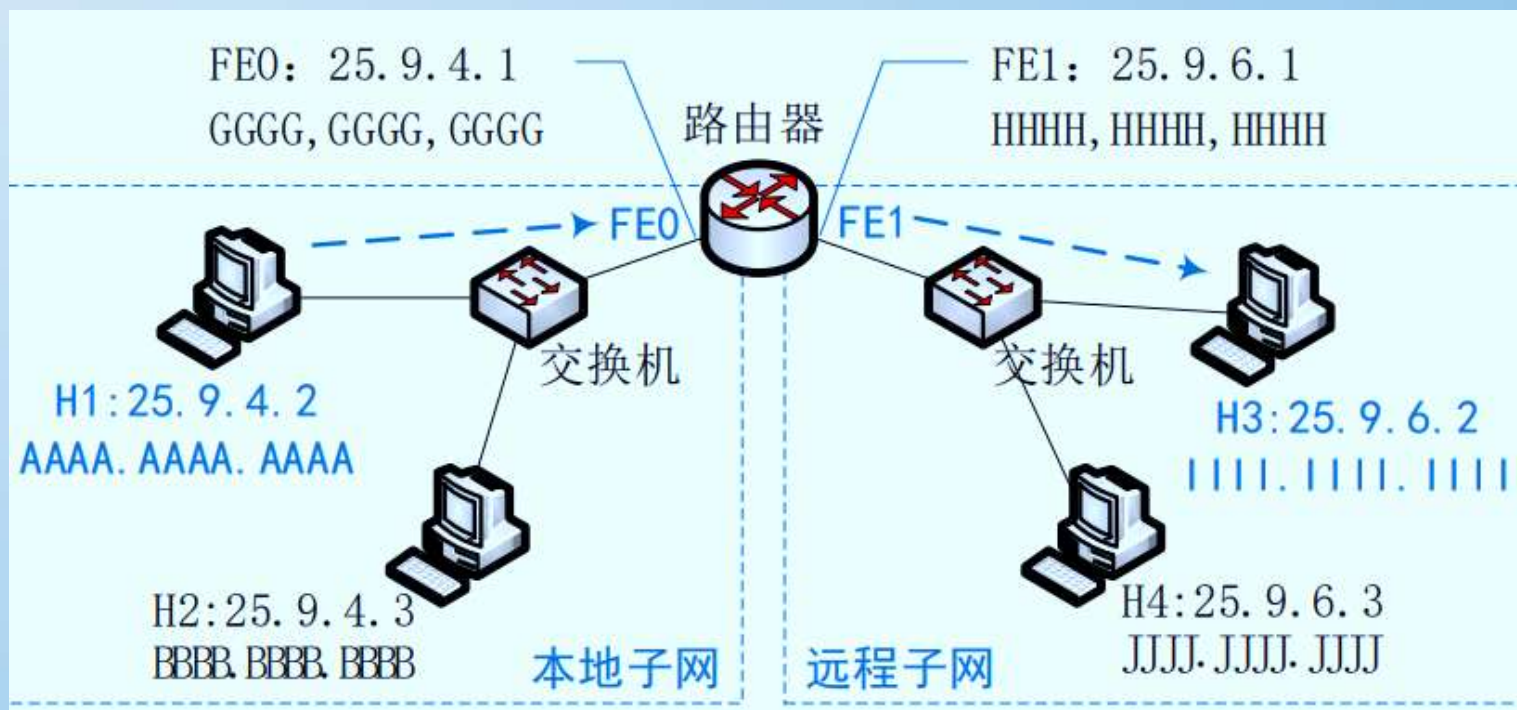
- RFC5227 (2008 年) : 检测子网中是否存在 IP 地址冲突
- 当一台主机因启动或重新配置而广播了免费 ARP 信息, 其目标地址就是它自己, 它本不期望收到一个应答
 - 但是如果意外地收到了一个 ARP 应答, 那一定是有别的主机已经使用了它的地址, 意味着 IP 地址冲突, 就只能重新配置一个了





向远程主机发送数据

- ARP的请求-应答过程，只能发生在源机和目的机位于同1个子网
- 如果目的机在远程子网中，ARP请求无法送达，如果不知对方MAC地址，怎么办？
 - 比如：H1向H3发送数据，必须默认网关的参与





与远程主机通信

- 默认网关：与一台主机相连接的路由器的接口，即是这台主机的默认网关

- 默认网关是相对的

- 上图H1发送给H3的数据，分成两段

- 每一段MAC寻址

- 全局：IP寻址，路由器参与

- ✓ 解封装，提取目的IP地址：
- ✓ 处理：确定目的网络，查找路由表。
- ✓ 重新封装
- ✓ 转发：路由表中找到了目的网络的路径记录，则按照指示，从相应的接口转发出去。

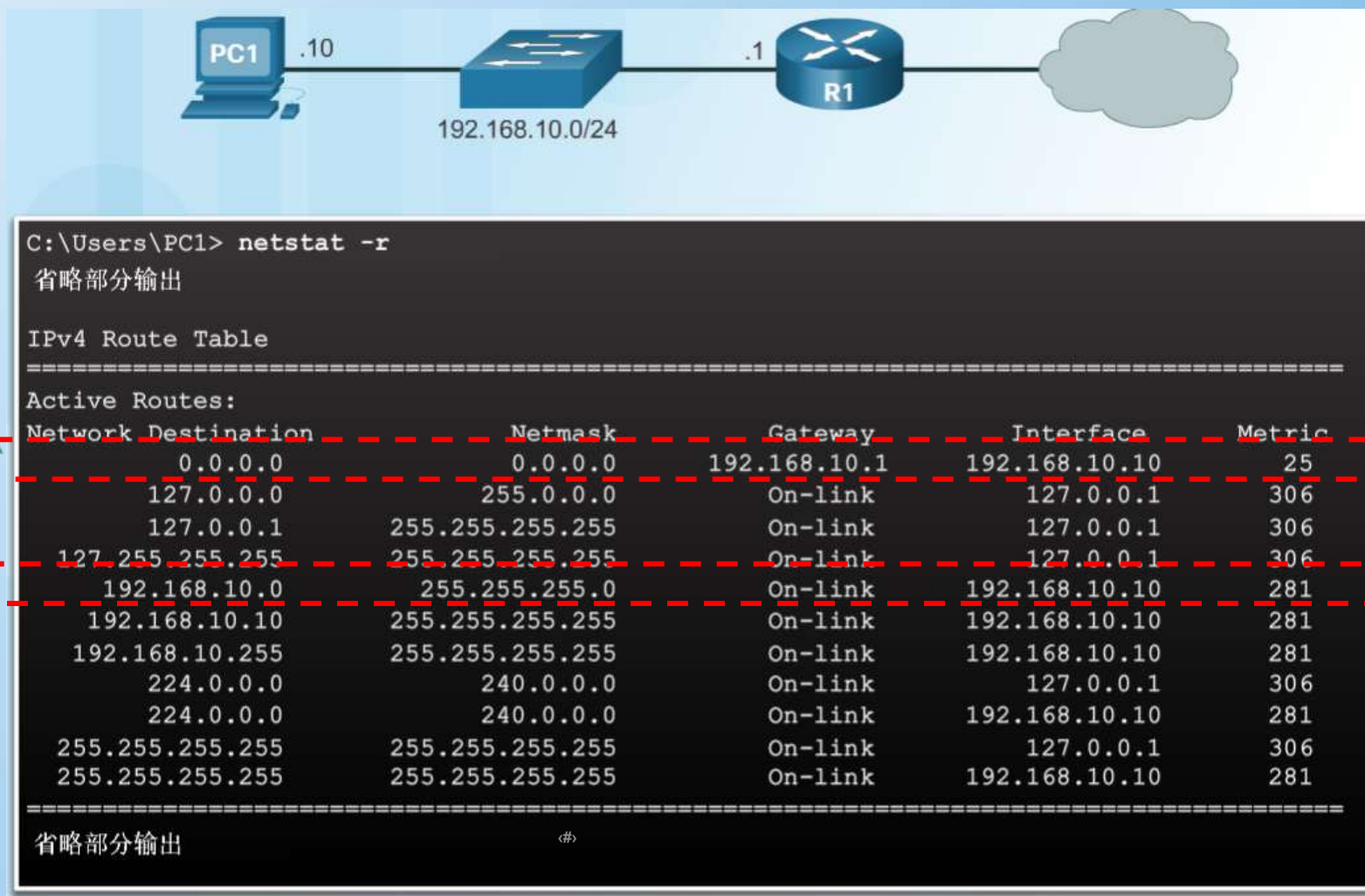
	分组头部		帧头部	
	源 IP 地址	目的 IP 地址	源 MAC 地址	目的 MAC 地址
H1→路由器 FE0	25.9.4.2	25.9.6.2	AAAA.AAAA.AAAA	GGGG.GGGG.GGGG
路由器 FE1→H3	25.9.4.2	25.9.6.2	HHHH.HHHH.HHHH	IIII.IIII.IIII





怎么判断目的机在远程？ 主机中的路由表

- 可使用命令route print 或 netstat - r ， 观察这张路由表



远程
网络

本地
网络





练习

1、ARP的功能是（ ）。（2012考研真题）

- A. 通过IP地址找到 MAC地址
- B. 通过IP地址找到 IP地址
- C. 通过域名找到 IP地址
- D. 通过IP地址找到域名

解析：地址解析ARP的主要功能是找到某台具有某个IP地址的目标主机对应的MAC地址。





练习

- 2、下列网络设备中，能够抑制广播风暴的是（ ）。（2010考研真题）
- I. Repeater（中继器） II. Hub（集线器）
III. Bridge（网桥） IV. Router（路由器）
- A. 仅 I 和 II
B. 仅 III
C. 仅 III 和 IV
D. 仅 IV

解析：一层设备只会增大广播域，扩散广播；二层设备可以控制冲突规模，分隔冲突域，但不能分隔广播域；只有路由器这种3层的设备，可以分隔广播域，控制广播。





练习

3、下列关于IP路由器功能的描述中，正确的是（ ）。（2012考研真题）
I.运行路由选择协议，及设置路由表；
II.当拥塞发生的时候，合理丢包；
III.检查分组头部的错误，所以可以确保没有分组丢失；
IV.根据目的IP地址，将分组转发到合适的线路上去；

- A. 仅 III 和 IV
- B. 仅 I ， II 和 III
- C. 仅 I ， II 和 IV
- D. 仅 I ， II ， III 和 IV

解析：III 是错的，差到错的分组会被丢弃，并不能保证不丢包





互联网控制消息ICMP

- ICMP是为了反馈或调整 IP 分组传输过程中遇到的问题而设计的，它并不保证 IP 分组传输的可靠，但它会报告其传输过程中遇到的问题。
- ICMP是 IP 协议的一部分，被称为 IP 协议的姊妹协议
- ICMP消息
 - ❑ 差错报告
 - ❑ 查询消息





ICMP消息格式

- 类型：具体类型
- 校验和：互联网校验和
- 代码：细分类别
- 数据：长度和内容不定

	类型	代码	消息名称	解释
差错报告	3	1	主机不可达	到达了目的主机所在网络，却无法找到目的
		3	端口不可达	到达了目的主机，但是找不到对应的端口
		4	需要分片但不允许分片	路由器需要将入境分组分片，才能适应出境接口的 MTU，但是分组中的控制位 DF=1
	11	0	超时	TTL 字段值已减为 0，分组仍然未到目的
		1	重组超时	目的机收到的分片不完全，无法重组，但分片定时器超期了
查询信息	8	0	回声请求	向目标机发出的探测分组
	0	0	回声应答	目标机发回的对回声请求的应答





ICMP应用1： ping

- 回声请求和应答
- **标识**：并未被严格定义，可以用“0”，也可以使用传输层的端口号，用以匹配请求和应答。
- **序号**：长 2 字节，由发起方产生和递增。
- **可选数据**：不定长，由发起方产生。

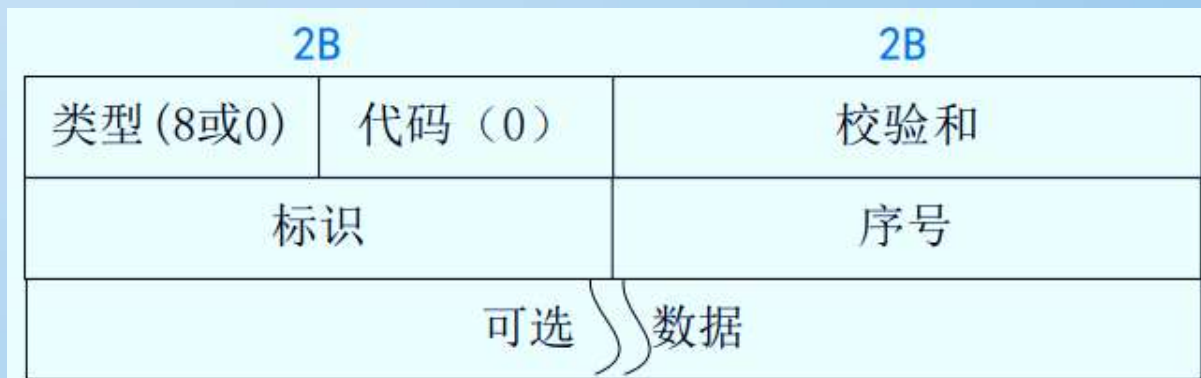


图 5-32 回声请求和回声应答





ping常用场景

- ping 127.0.0.1: 可以测试本机的 TCP/IP 协议栈是否工作正常
- ping 本机 ip 地址: 测试本机接口 (适配器) 设置是否正常
- ping 默认网关: 测试本机的对外 (远程子网) 通道是否通畅
- ping 域名: 测试本地的域名解析是否正常, 以及与域名所代表的主机的通达性
- ping 目标机:
测试本机和目标机
之间的通达性

```
C:\WINDOWS\system32>Ping 192.168.1.1

正在 Ping 192.168.1.1 具有 32 字节的数据:
来自 192.168.1.1 的回复: 字节=32 时间=7ms TTL=64
来自 192.168.1.1 的回复: 字节=32 时间=11ms TTL=64
来自 192.168.1.1 的回复: 字节=32 时间=15ms TTL=64
来自 192.168.1.1 的回复: 字节=32 时间=20ms TTL=64

192.168.1.1 的 Ping 统计信息:
    数据包: 已发送 = 4, 已接收 = 4, 丢失 = 0 (0% 丢失),
    往返行程的估计时间(以毫秒为单位):
        最短 = 7ms, 最长 = 20ms, 平均 = 13ms
```





ICMP应用2: traceroute

■路由追踪工具

■基本原理

- 构造并发送TTL=1的分组开始，未达目的的，路由器回发超时消息
- 到达目的机，回发ICMP端口不可达消息，完成追踪

2B		2B
类型(11)	代码(0)	校验和
保留未使用(0)		
IP分组头部+分组数据的前8个字节		

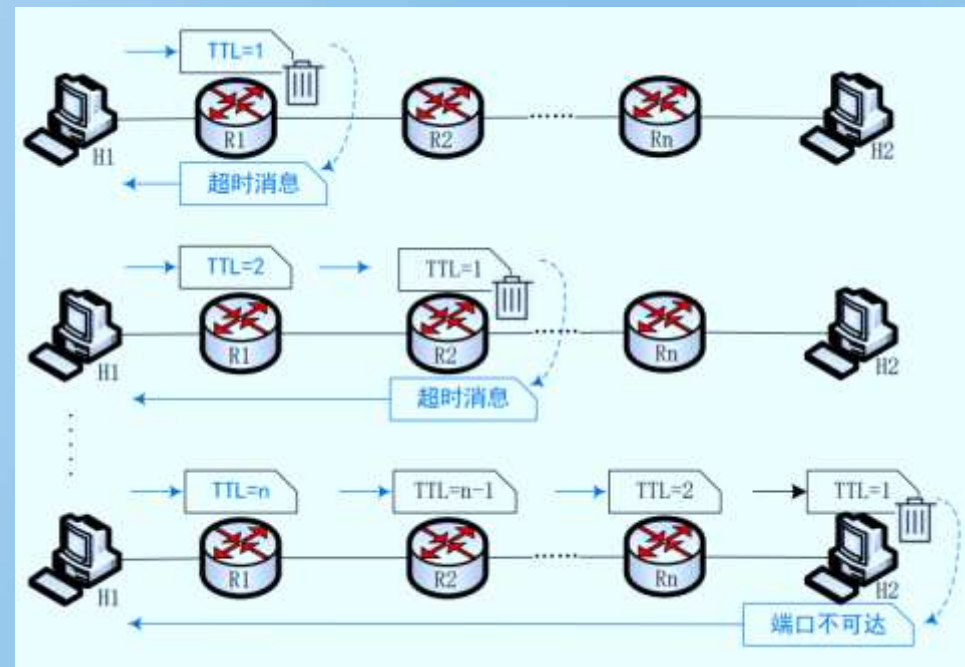
2B		2B
类型(3)	代码(3)	校验和
保留未使用(0)		
IP分组头部+分组数据的前8个字节		





H1追踪到H2中间经过的路由器（跳）

- 封装并发送第 1 个分组。TTL=1、目的 IP 地址是 H2;
 - 第一跳
- 封装并发送第 2 个分组：封装 TTL=2、目的 IP 地址是 H2
 - 第二跳
- 封装并发送第 3,, n 个分组：每多发一个分组，TTL 值增加 1，重复上述步骤，
- 完成追踪：H1收到端口不可达消息





一个路由追踪的例子

■ 注意

- 每个分组是独立寻径的，所以，从源机到目标机的分组走的路径不一定一样，所以，每次追踪路由的结果不一定相同。

```
C:\Users\HFF>tracert www.baidu.com
```

```
通过最多 30 个跃点跟踪  
到 www.a.shifen.com [182.61.200.7] 的路由:
```

1	7 ms	4 ms	4 ms	gw.gcableconfig.local [192.168.1.1]
2	18 ms	14 ms	15 ms	10.131.0.1
3	*	*	*	请求超时。
4	*	*	*	请求超时。
5	*	*	*	请求超时。
6	*	*	*	请求超时。
7	*	*	*	请求超时。
8	*	*	*	请求超时。
9	*	*	*	请求超时。
10	*	*	2015 ms	172.31.237.57
11	23 ms	16 ms	14 ms	103.27.24.98
12	34 ms	18 ms	16 ms	182.61.255.240
13	55 ms	56 ms	61 ms	182.61.255.110
14	*	63 ms	53 ms	182.61.255.45
15	*	*	*	请求超时。
16	*	*	*	请求超时。
17	*	*	*	请求超时。
18	72 ms	58 ms	53 ms	182.61.200.7

```
跟踪完成。
```





ICMP应用3：PMTU

- 路径 MTU：PMTU：是从源到目的的路径上的所有网络的 MTU 的最小值
- PMTU 发现 (PMTUD)：一种探测技术，用以发现源机到目的机路径上的所有网络的最小 MTU
- ICMP 差错报告消息
 - 类型为 3，代码为 4

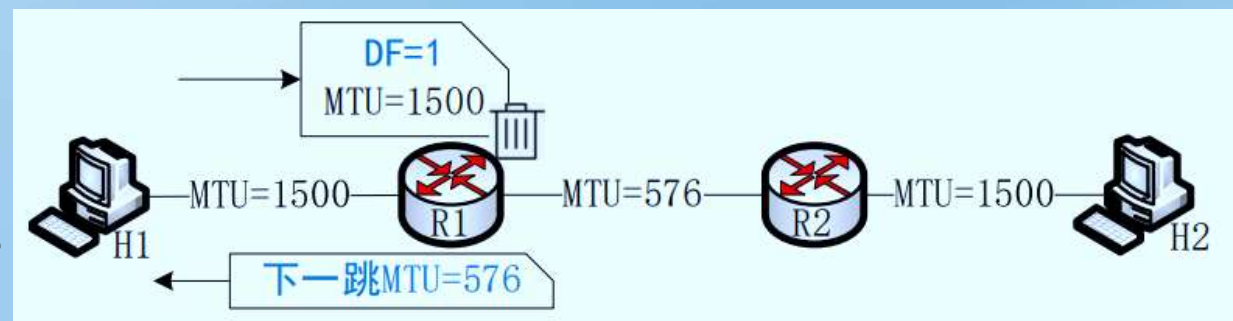
2B		2B
类型 (3)	代码 (4)	校验和
保留未使用 (0)		下一跳的MTU
IP分组头部+分组数据的前8个字节		





PMTU的工作原理

- 第一个分组：源机按照第一跳（源机的默认网关）MTU 值作为 PMTU 的初值，并按照此值分割和封装数据，且将分组头部的 DF 置为 “1”，表示不允许中途路由器进行分片
- 如果收到差错报告，提取出下一跳MTU，并按照此值封装
- 如果再次收到差错报告，重复上述过程
- 直到分组可以顺利到达目的地
 - 最后一次的MTU作为PMTU值





练习

- 1、回声请求（echo request）和回声应答（echo reply）两个互联网控制ICMP消息的主要功能是什么？
- A. 报告错误
 - B. 检查节点（node）之间的通达性
 - C. 检查IP分组的生存时间
 - D. 发现IP地址

解析：回声请求和回声应答不用于差错报告，而是用于诊断网络问题，通过这两个消息的交互以发现两个节点之间是否通达。所以，“报告错误”选项是错误的；“检查IP分组的生存时间”选项也是错误的；“发现IP地址”是早期功能，已经废止；只有“检查节点（node）之间的通达性”选项是正确的。





练习

2、封装了ICMP差错报告消息的IP分组，向下面哪一个发送？

- A. 发生差错的IP分组的源机
- B. 发生差错的IP分组的目的地
- C. 发生差错的路由器
- D. 上面都没有

解析：在IP报文头部只有源机IP和目的地IP可用，ICMP差错报告按规定应向发生了差错的IP分组的源机报告，这种规定是有道理的。因为发生的差错有很多种情况，比如可能有目的不可达、拥塞等等，如果把差错报告向目的地发送，有可能差错报告本身也会遭遇同样的差错而无法到达，且目的地也许根本不关心那个发生了差错的报文。





练习

3、以下关于**ICMP**报文结构的描述中，错误的是_____。

- A. IP分组头协议字段值为**1**，表示数据搭载的是**ICMP**报文
- B. **ICMP**报文的前**4B**的格式是统一的
- C. 第一个字段的**1B**是代码
- D. **4B**类型相关的头部之后是数据字段

解析：**1**在IP分组头中，协议字段值为**1**表示IP分组的数据部分是**ICMP**报文。“IP分组头协议字段值为**1**，表示数据搭载的是**ICMP**报文”选项是正确的。**ICMP**报文的前**4B**的格式是统一的，即第一个字段（**1B**）是类型，第二个字段（**1B**）是代码，第三个字段（**2B**）是校验和。第四个字段（**4B**）的内容与消息类型相关。因此，“**ICMP**报文的前**4B**的格式是统一的”是正确的，“第一个字段的**1B**是代码”选项是错误的。在这四个字段之后是数据字段。“**4B**类型相关的头部之后是数据字段”是正确的。





练习

4、当一个路由器收到一个IP分组，其中的生存时间TTL=1，此时路由器会发送一个互联网控制ICMP消息给源，这个ICMP消息的类型是什么？

- A. 目标不可达（destination-unreachable）
- B. 超时（time-exceeded）
- C. 超时（time-exceeded）
- D. 重定向（redirection）

解析：路由器将 $TTL-1=0$ ，丢弃这个分组，并向源报告一个超时消息。





练习

5、什么时候路由器会丢掉一个tracert分组？ (From CCNA)

- A. 当路由器接收到ICMP?超时消息时
- B. 当RTT值等于0时
- C. 当主机用ICMP回声应答消息回应时
- D. 当TTL值等于0时

解析：当路由器接收到 tracert 分组， $TTL-1$ ，当 $TTL-1=0$ 时，路由器丢掉分组，且向源发回一个超时消息。





练习

6、从29.18.0.0 到29.18.128.255 的一组IP 地址已经被聚合到29.18.0.0/17 。然而，这里有一个空闲地址块，即从29.18.60.0 到29.18.63.255 之间的1024 个地址还没有被分配。现在这块空闲地址突然要被分配给一台使用不同出境线路的主机。试问是否有必要将聚合地址分割成几块，然后把新的地址块加入到路由表中，再来看是否可以重新聚合？如果没有必要这样做，请问该怎么办呢？

解答：

没有必要。只要在路由表中添加一项：29.18.60.0/22 就可以了。

当有一个分组到来时，如果它既匹配29.18.0.0/17，又匹配29.18.60.0/22，那么它将被发送到掩码位数较大的目标地址，即29.18.60.7/22（最长地址前缀匹配）。这样做的好处是使得一个大段的地址能够被指定到一个目标，但又允许其中少量的地址出现例外的情况。





网络地址转换

■网络地址转换（NAT）：一个快速的地址短缺修补方案

■私人地址（RFC1918）

□可反复重用，不具有全球唯一性

私人地址块	第一个地址	最后一个地址	备注
10.0.0.0/8	10.0.0.0	10.255.255.255	1 个 A 类地址
172.16.0.0/12	172.16.0.0	172.16.31.255	16 个 B 类地址
192.168.0.0/16	192.168.0.0	192.168.255.255	256 个 C 类地址

■私网\内网：使用私人地址的网络

□如果要与外网通信，需要使用NAT转换



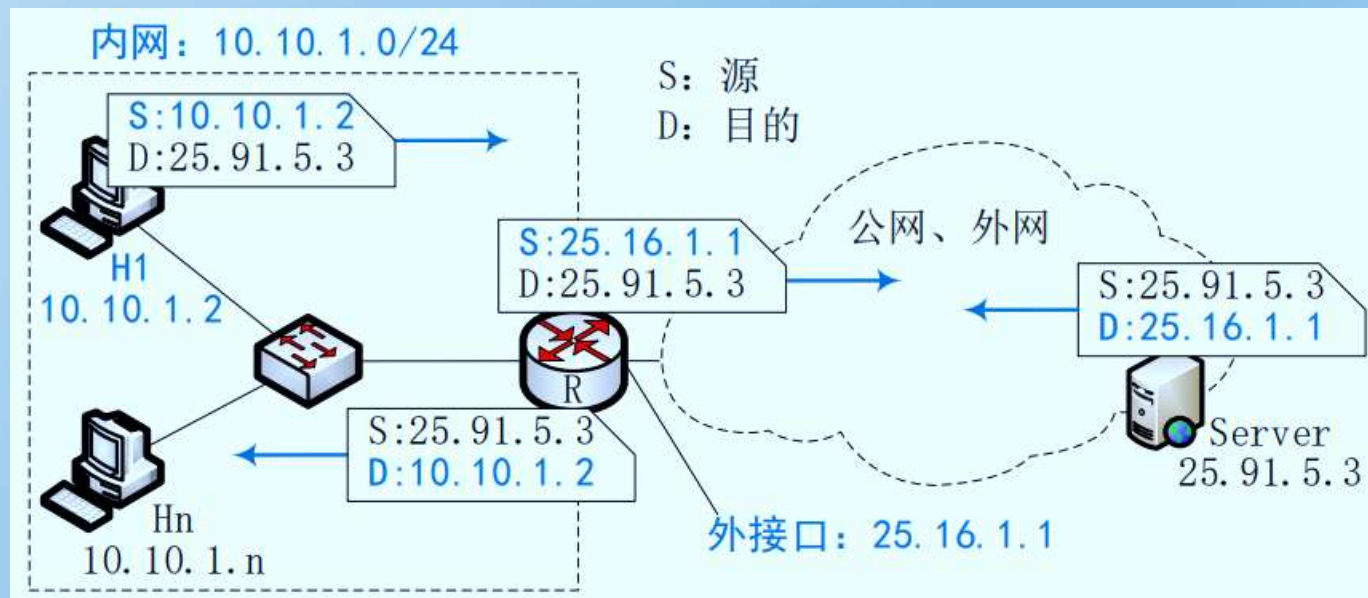


网络地址转换NAT的工作原理

- NAT：是把一些私有 IP 地址映射为另一些公有 IP 地址的方法
- NPAT：是把私有 IP 地址和传输协议端口号转换成一个公有 IP 地址和不同传输协议端口号的方法

□ 超载

- NAT转换器：完成转换





NAT转换过程_1

- (1) 主机 H1 首先封装自己的数据，源地址是 10.10.1.2
- (2) 数据分组到达边界路由器 R 时，其源地址和源端口被置换为 25.16.1.1 和 2001，转换的信息被记录到了转换表中，，转换后的端口号可用作转换表的索引。

序号	转换后		转换前	
	端口号	公网地址	端口号	私人地址
1	2001	25.16.1.1	6000	10.10.1.2
2	2002	25.16.1.1	8000	10.10.1.3
.....				





NAT转换过程_2

- (3) 转换后的数据分组中已经没有任何私有 IP 地址的痕迹, ; 应答分组的目的 IP 地址 和端口号分别是 25.16.1.1 和 2001
- (4) 应答分组一路寻址, 最终会到达目的机 25.16.1.1 (路由器 R), 路由器 R 解封装并按照目的端口号2001去查找转换表.将该应答分组的目的地址和目的端口号置换为表项指示的10.10.1.2 和6000。
- (5) 转换后的分组被转发到私网。出去的分组和回来的分组头部的 IP 地址和端口号在 NAT 转换器这里都发生了变化, 如





NAT转换

- NAT转换器：可以是路由器，也可以是独立的服务器
- 由内网主机发起
- 对源机和目的机完全透明

转换方向↵	数据头部信息变化↵			
内网→公网（出去）↵	内网↵		公网↵	
	源 IP 地址↵	源端口号↵	源 IP 地址↵	源端口号↵
	10. 10. 1. 2↵	6000↵	25. 16. 1. 1↵	2001↵
内网←公网（返回）↵	公网↵		内网↵	
	目的 IP 地址↵	目的端口号↵	目的 IP 地址↵	目的端口号↵
	25. 16. 1. 1↵	2001↵	10. 10. 1. 2↵	6000↵





NAT转换的弊端

- 端到端的传输被打断：NAT 的采用，使 IP 分组的传输变成了“源→NAT 转换器→目的机”的有“连接”传输
- NAT 转换器成为了一个瓶颈：内网的所有外访主机都必须经过 NAT 转换器
- NAT 的转换是有方向的，内外网互访并不对等，从内网向外网发起主动访问是可行的，反过来却不行
- 部分应用会失效：NAT 仅处理分组的头部和段的头部，如果某个应用所需的 IP 地址在载荷数据中，将导致应用失效
- 非TCP\UDP的应用失效：NAT 的工作过程使用了传输层的端口号，如果应用进程并没有使用 TCP 或UDP，将导致应用失效。





本章小结_1

- 两种通信网络：数据报网络（无连接）和虚电路网络（面向连接）
- IP的功能：将分组从源机一路送达目的机
- IP分组格式（12、13或14个字段），不含选项的头部长20字节
 - 头部长、TTL、用户协议、总长度、分片、源地址、目的地址.....
- IP地址的点分十进制表示及特点
- IP地址分类以及特殊IP地址
- 子网划分方法
 - 子网掩码





本章小结_2

- IP地址的分配和获取
 - IANA及5的地区互联网信息中心
 - 动态主机配置协议
- IP寻址和MAC寻址，共同完成数据的交付
- 无类域间路由CIDR缓解了地址枯竭趋势，缩小了路由表规模
- 地址解析协议：子网内、远程主机通信
- 互联网控制消息协议ICMP：ping、traceroute、PMTU
- 网络地址转换NAT：原理、过程和弊端





本章对应的中英文术语_1

英文名称	缩写	对应的中文名称
Virtual-circuit Subnet		虚电路子网
Datagram Subnet		数据报子网
Routing Table	RT	路由表
Internet Protocol Version 4	IPv4	互联网协议第 4 版
Internet Protocol Version 6	IPv6	互联网协议第 6 版
Simple Internet Protocol Plus	SIPP	增强的简单因特网协议
IP Packet		IP 分组
Time to Live (IPv6: hop limit)	TTL	生存时间 (以跳数为单位)
Flow Label		流标签
Next Header		下一个头
Dotted Decimal Notation		点分十进制
Colon Hex Notation		冒分十六进制
Link- Local Address		链路本地地址





本章对应的中英文术语_2

英文名称	缩写	对应的中文名称
Subnet		子网（规划）
Subnet Mask	SM	子网掩码
Congestion Control		拥塞控制
Flow Control		流量控制
Load Shedding		载荷脱落
Random Early Detection		随机早期检测
Leaky Bucket Algorithm		漏桶算法
Token Bucket Algorithm		令牌桶算法
Internet Control Message Protocol	ICMP	互联网控制消息协议
ARP(Address Resolution Protocol)	ARP	地址解析协议
BOOTP		自举协议
Dynamic Host Configuration Protocol	DHCP	动态主机配置协议
Classless InterDomain Routing	CIDR	无类别域间路由选择
Network Address Translation	NAT	网络地址转换





致谢和声明

- 本课程课件中的绝大部分插图来自于
 - 袁华，王昊翔，黄敏主编《深入理解计算机网络》，清华大学出版社
- 其余的部分素材来自于
 - 华为智能基座课程
 - 思科网络技术学院教程
 - 网络上搜到的其它公开资料
- 特别对资料的各提供机构和个人表示诚挚的感谢！
- 引用的素材，仅用于课程学习，如果有任何问题，请与我们联系！

