



第五单元 网络层

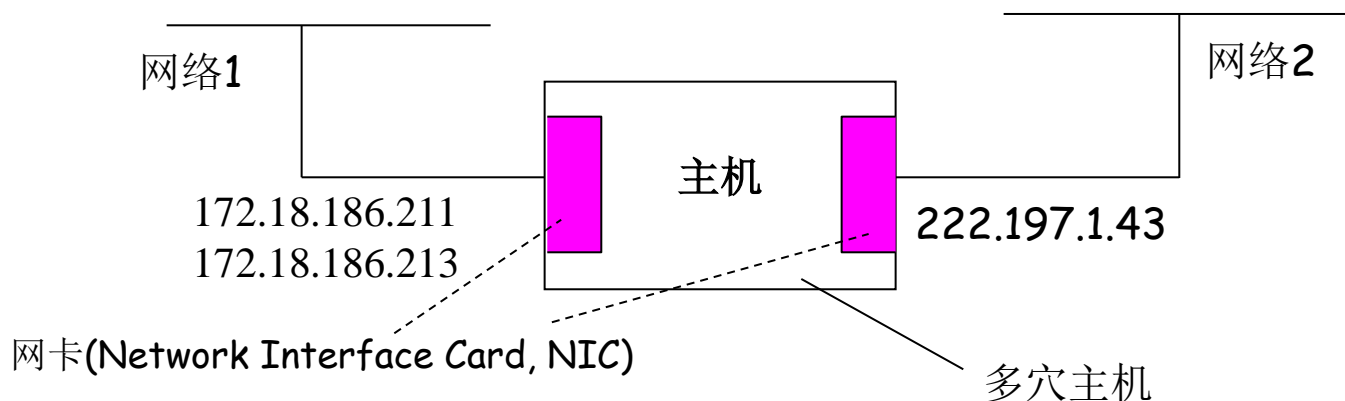
-IP地址

- IP地址空间
- IP地址结构
- 有类网和子网划分(VLSM、CIDR)
- 特殊的IP地址
- 私有IP地址和NAT
- 多播IP地址
- ARP协议
- DHCP协议
- ICMP协议



IP地址空间

- ❑ 48位的**MAC**地址和32位的**IP**地址都是全局的(全球分配), 但是**IP**地址空间是分层的, 是可路由的(**routable**)。
- ❑ **IP**地址由**ICANN**统一负责并逐级分配。亚洲由**APNIC**负责, 中国由**CNNIC**负责。
- ❑ **IP**地址属于接口(网卡)。主机或路由器的每个接口可以配置一个或多个**IP**地址。



ICANN --The Internet Corporation for Assigned Names and Numbers

IP地址结构



- ❑ 一个IP地址可以划分为两个部分: 网络号(**network numbers**)和主机号(**host identifier**)。
- ❑ 网络号也称为网络前缀(**network prefix**)、网络标识 (**network ID**) 。它是用来确定拥有该IP地址的主机位于哪个网络，而主机号用于确定属于该网络的哪台主机。

有类网

点分十进制(dotted decimal)

所属类	IP地址的格式					地址范围	每个网络的地址数	
	31	23	15	7	0			
A	0	网络号		主机号		0.0.0.0 ~127.255.255.255 (2 ⁸ =128 个)	2 ²⁴ = 16,777,216	
B	10	网络号		主机号		128.0.0.0 ~191.255.255.255 (2 ¹⁴ =16384个)	2 ¹⁶ = 65536	
C	110	网络号		主机号		192.0.0.0 ~223.255.255.255 (2 ²¹ = 2,097,152个)	2 ⁸ = 256	
D	1110	多播地址					224.0.0.0 ~239.255.255.255	
E	1111	保留					240.0.0.0 ~255.255.255.255	

* 实际上，在有类网模型下，可用的A类网个数要减2，因为网络号全0和全1不可用。B类网和C类网个数也要减2。

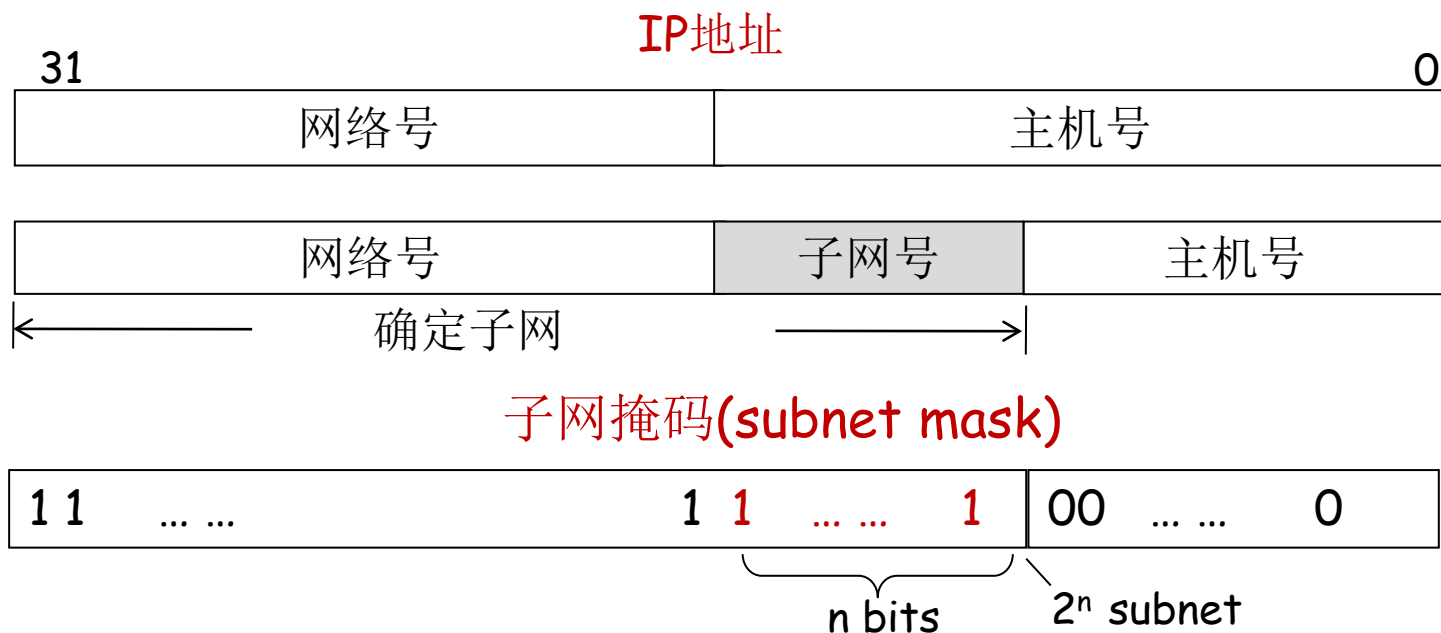
IPv4地址耗尽问题

[**CIDR Address Strategy** , Fuller, Li, Yu & Varadhan , **1993**]

- As the Internet has evolved and grown over in recent years, it has become evident that it is soon to face several serious scaling problems. These include:
 1. **Exhaustion of the class B network address space.** One fundamental cause of this problem is the lack of a network class of a size which is appropriate for mid-sized organization; class C, with a maximum of 254 host addresses, is too small, while class B, which allows up to 65534 addresses, is too large for most organizations.
 2. **Growth of routing tables** in Internet routers beyond the ability of current software, hardware, and people to effectively manage.
 3. **Eventual exhaustion of the 32-bit IP address space.** It has become clear that the first two of these problems are likely to become critical within the next one to three years.

子网划分

- 一个有类网可以划分为多个相同大小的子网(subnet):



子网掩码也可以用点分十进制表示：**C类网192.168.1.0**划分为四个子网的子网掩码为**255.255.255.192**，子网号分别为**00、01、10、11**。

- * **主机号为全1或者全0的地址被保留，不能使用。**
- * **子网号为全0或全1的子网现在都可以使用（以前规定不能使用）。**

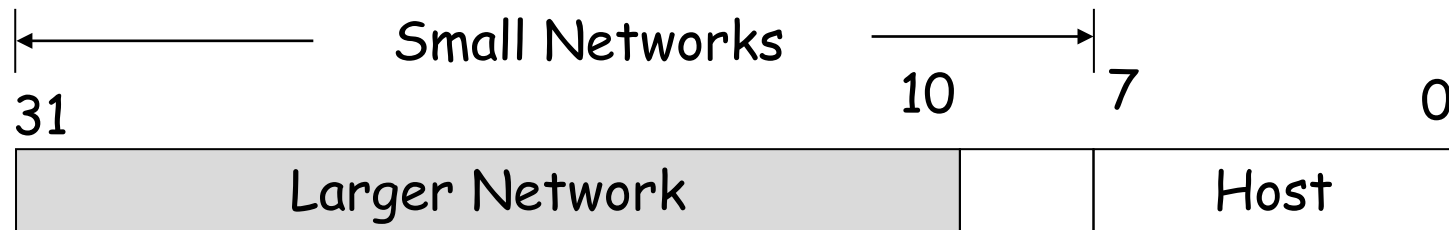
变长子网掩码

- ❑ 变长子网掩码(Variable-Length Subnet Mask, VLSM) 允许把一个有类网划分为多个不同大小的子网。
- ❑ 例如，给定一个有类网**199.1.12.0**，如何把它们划分为四个子网，使它们可以分别容纳 **100**，**60**，**25**和**20**台主机？

子网1:	199.1.12.0XXXXXXXX/25	(100台主机)
子网2:	199.1.12.10XXXXXXXX/26	(60台主机)
子网3:	199.1.12.110XXXXXX/27	(25台主机)
子网4:	199.1.12.111XXXXXX/27	(10台主机)

用长度来表示子网掩码：/26表示**255.255.255.192**

无类域间路由选择协议



- ❑ 无类域间路由选择协议(Classless Inter-Domain Routing, CIDR)允许把多个有类网合并为一个更大的网络,称为超网(supernet)。
- ❑ 例如,把有类网192.24.8.0~192.24.15.0合并为网络号为192.24.8.0、子网掩码为255.255.248.0的超网。
- ❑ CIDR可以显著减少路由表中路由的数量,例如,上例就把八个路由减少路由减少为一个路由,称为路由聚合(route aggregation)。
- ❑ 通过引入CIDR,加上子网掩码,现在的网络号(可能包含子网号)可以看成是没有边界,即是无类的。

特殊的IP地址

(1)

0 0 0	0
-------	--------	---

未知或秘密IP地址，只用作源地址

(2)

0 0 0	0	Host
-------	--------	---	------

同一子网的主机，只用作源地址

(3)

1 1 1	1
-------	--------	---

有限广播，对于一个直连物理网络的广播

(4)

network	1 1	1
---------	-----	--------	---

对于一个远程网络的广播

(5)

network	0 0	0
---------	-----	--------	---

用32比特表示的网络号(含子网号)

(6)

0 1 1 1 1 1 1 1	any value
-----------------	-----------

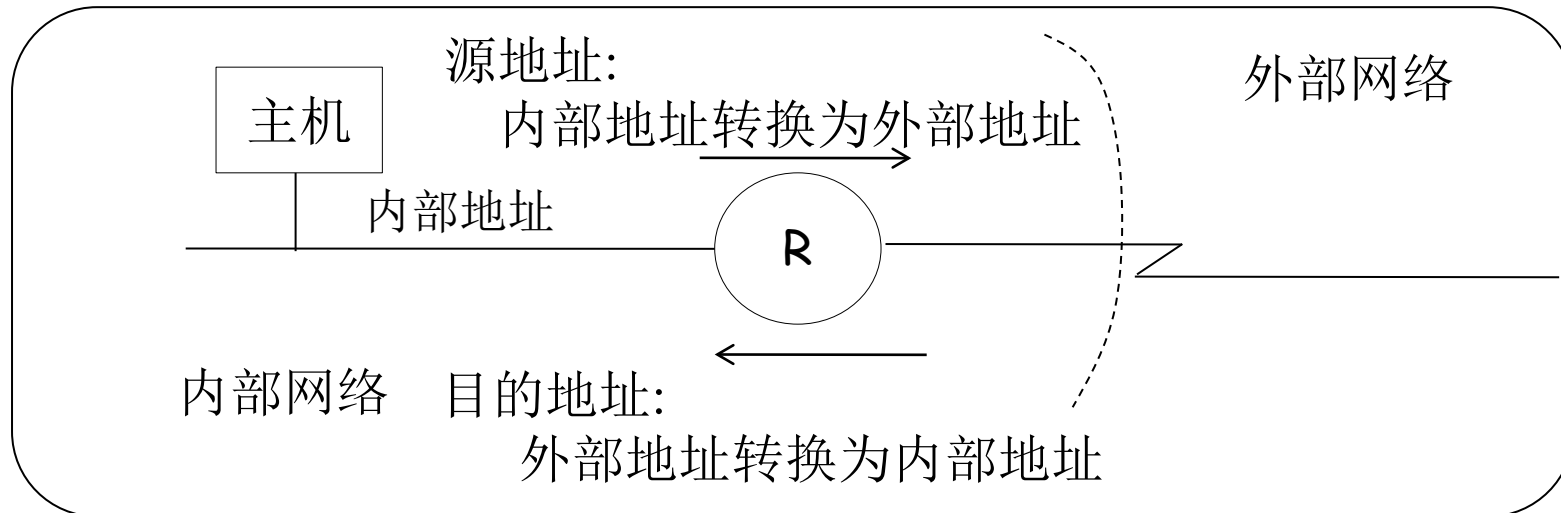
环回地址(loopback) - 本机 127.0.0.1 - 本地地址(localhost)

私有IP地址

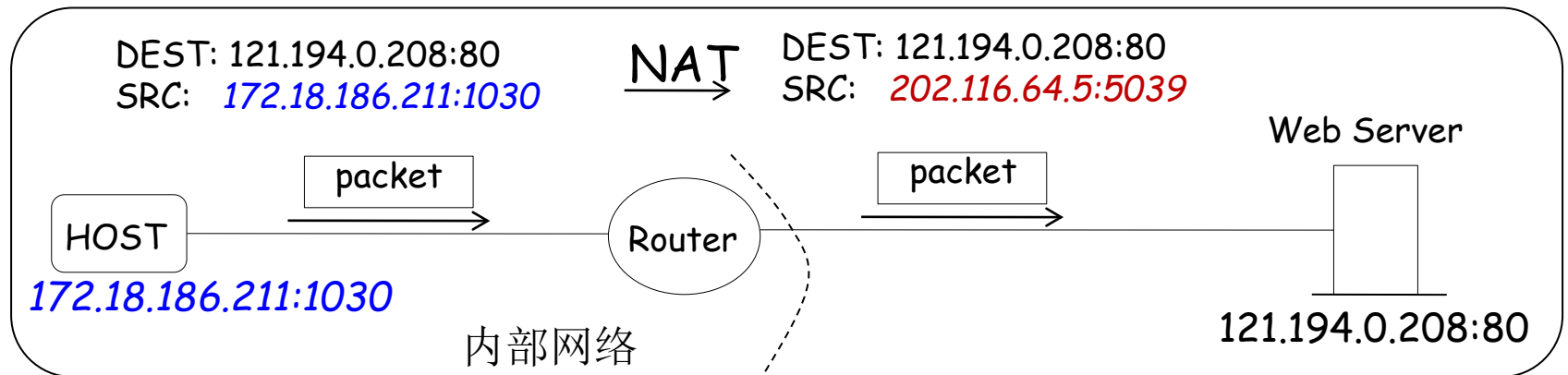
- 私有IP地址就是无需IANA分配、任何人都可以使用的IP地址：
 - (1) **10.0.0.0 ~ 10.255.255.255**
 - (2) **172.16.0.0 ~ 172.31.255.255**
 - (3) **192.168.0.0 ~ 192.168.255.255**
- 私有地址只能用于内部网络。主干网上的路由器会过滤掉目的地址为私有地址的IP数据报。因此，离开内部网络的IP数据报必须使用由IANA分配的全局地址作为目的地址。

网络地址转换

- ❑ 网络地址转换(**Network Address Translation, NAT**)是一种把内部地址映射为外部地址的技术。例如, 把私有地址映射为全局地址。
- ❑ 出口路由器在内网数据报发往外网时自动把内网地址映射为外网地址的方法称为动态**NAT**。每个动态映射都关联一个**TTL**。如果在**TTL**时间内没有使用一个映射, 该映射将被出口路由器删除。直接由管理员加入映射的方法称为静态**NAT**。静态**NAT**加入的映射不会被自动删除。



- ❑ NAT(Network Address Port Translation) 把端口号也加入到NAT的映射中，也称为PAT(Port Address Translation)或带有 过载 NAT(NAT with overload)。
- ❑ 下面的图显示了从私有网络的主机发包给Web服务器的地址转换方法：



source address	destination address	global address	src port	dest. port	global port	protocol	connection released	timer
172.18.186.211	121.194.0.208	202.116.64.5	1030	80	5309	TCP	no	2.0

* The timer expires in 2 minutes

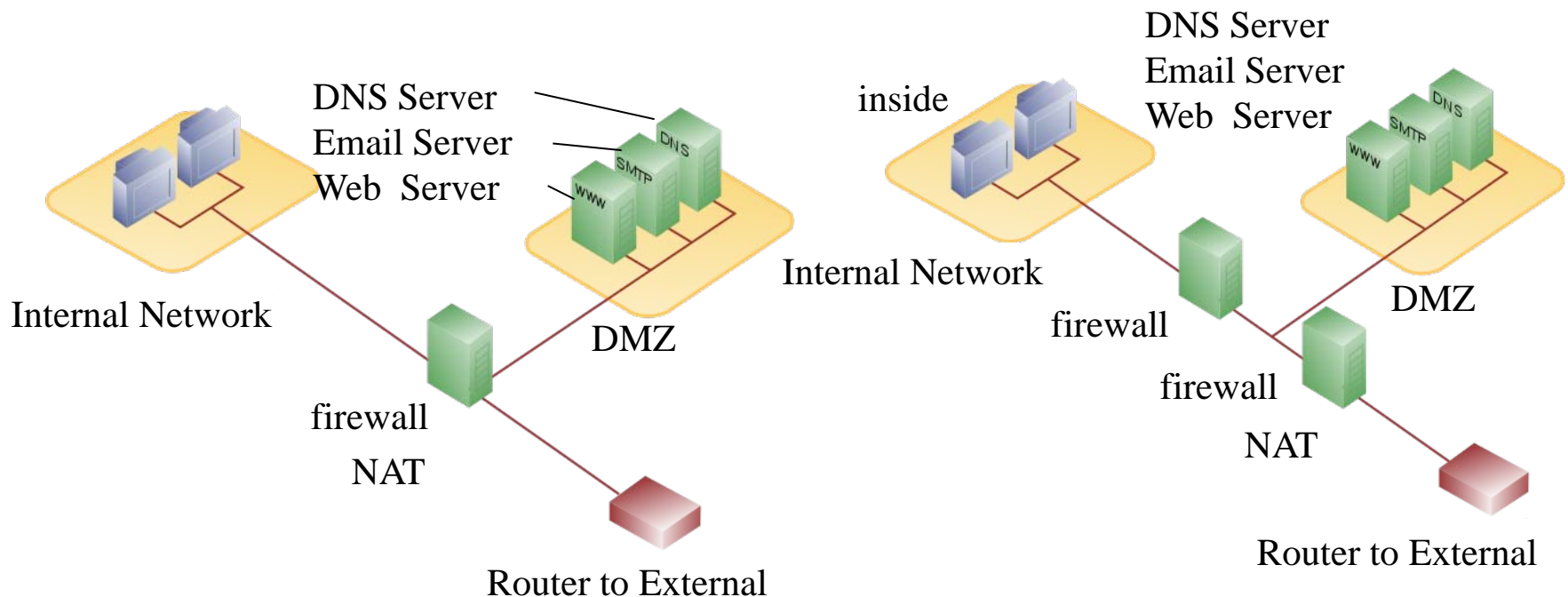
<http://tools.ietf.org/html/rfc4966>

Windows NAT表

SIST-STU1 - 网络地址转换会话映射表格								
通讯协议	方向	专用地址	专用端口	公用地址	公用端口	远程地址	远程端口	空闲时间
TCP	出站	10.1.100.9	2,130	172.18.186.211	62,447	210.28.176.12	80	8
TCP	出站	10.1.100.9	2,131	172.18.186.211	62,448	210.28.176.12	80	11
TCP	出站	10.1.100.9	2,132	172.18.186.211	62,449	210.28.176.12	80	11
TCP	出站	10.1.100.9	2,133	172.18.186.211	62,450	64.233.181.154	80	11
TCP	出站	10.1.100.9	2,134	172.18.186.211	62,451	210.28.176.12	80	11
TCP	出站	10.1.100.9	2,135	172.18.186.211	62,452	210.28.176.12	80	11
TCP	出站	10.1.100.9	2,136	172.18.186.211	62,453	210.28.176.12	80	11
TCP	出站	10.1.100.9	2,137	172.18.186.211	62,454	64.233.181.154	80	11
TCP	出站	10.1.100.9	2,138	172.18.186.211	62,455	119.42.233.243	80	11
TCP	出站	10.1.100.9	2,139	172.18.186.211	62,456	210.28.176.12	80	11
TCP	出站	10.1.100.9	2,140	172.18.186.211	62,457	210.28.176.12	80	11
TCP	出站	10.1.100.9	2,141	172.18.186.211	62,458	210.28.176.12	80	11
TCP	出站	10.1.100.9	2,142	172.18.186.211	62,459	64.233.181.154	80	11
TCP	出站	10.1.100.9	2,143	172.18.186.211	62,460	64.233.181.154	80	11
TCP	出站	10.1.100.9	2,159	172.18.186.211	62,476	210.28.176.12	80	10
TCP	出站	10.1.100.9	2,160	172.18.186.211	62,477	210.28.176.12	80	10
TCP	出站	10.1.100.9	2,161	172.18.186.211	62,478	210.28.176.12	80	10
TCP	出站	10.1.100.9	2,162	172.18.186.211	62,479	210.28.176.12	80	10

非军事化区*

- ❑ 非军事化区(**Demilitarized Zone, DMZ**)是位于内部网络和外部网络之间并为双方提供因特网服务的区域。
- ❑ 内网主机可以访问内网主机、**DMZ**和因特网。内网主机可以使用内部地址或全局地址访问**DMZ**的服务器。外部主机只能通过全局地址访问**DMZ**的服务器，不能访问内网主机。



多播IP地址*

多播地址范围	用法
224.0.0.0~239.255.255.255	IPv4的多播地址空间
224.0.0.0~224.0.0.255	由IANA分配的永久地址。路由器不转发目的地址为这些地址的IP数据包
224.0.1.0~224.0.1.255	由IANA分配的永久地址。路由器转发目的地址为这些地址的IP数据包
232.0.0.0~232.255.255.255	用于指定源的多播应用
233.0.0.0~233.255.255.255	由AS分配的全局多播地址
239.0.0.0~239.255.255.255	私有多播地址
其它地址	临时多播地址(transient address)

知名多播地址*

224.0.0.0	base address (reserved)		RFC 1222
224.0.0.1	All System on this subset	本网中的所有节点	RFC 1222
224.0.0.2	All routers on this subset	本网中的所有路由器	RFC 1222
224.0.0.3	Unassigned		
224.0.0.4	DVMRP		RFC 1075
224.0.0.5	OSPF-IGP-all routers	所有OSPF路由器	RFC 1583
224.0.0.6	OSPF-IGP-designated routers	所有OSPF指定路由器	RFC 1583
224.0.0.7	ST routers		RFC 1190
224.0.0.8	ST hosts		RFC 1190
224.0.0.9	RIP2	所有RIPv2路由器	RFC 1723
224.0.0.10	IGRP routers	所有IGRP路由器	Cisco
224.0.0.11	Mobile-agents		
224.0.0.12	DHCP server/relay agent	所有DHCP路由器	RFC 1884
224.0.0.13	PIM	所有PIM路由器	RFC 1884
224.0.0.14-224.0.0.255	unassigned		

IP地址与MAC地址(以太网)

IP 单播地址 $\xrightarrow{\text{ARP}}$ MAC单播地址

IP 广播地址 \longrightarrow MAC广播地址
111.....1(32 bits) \longrightarrow 111111.....1(48 bits)

IP 多播地址(1110开头) \longrightarrow MAC多播地址(第1字节最后1位为1)

举例: 224.1.2.3

E0010203(十六进制)

\longrightarrow 0x01-00-5E-01-02-03

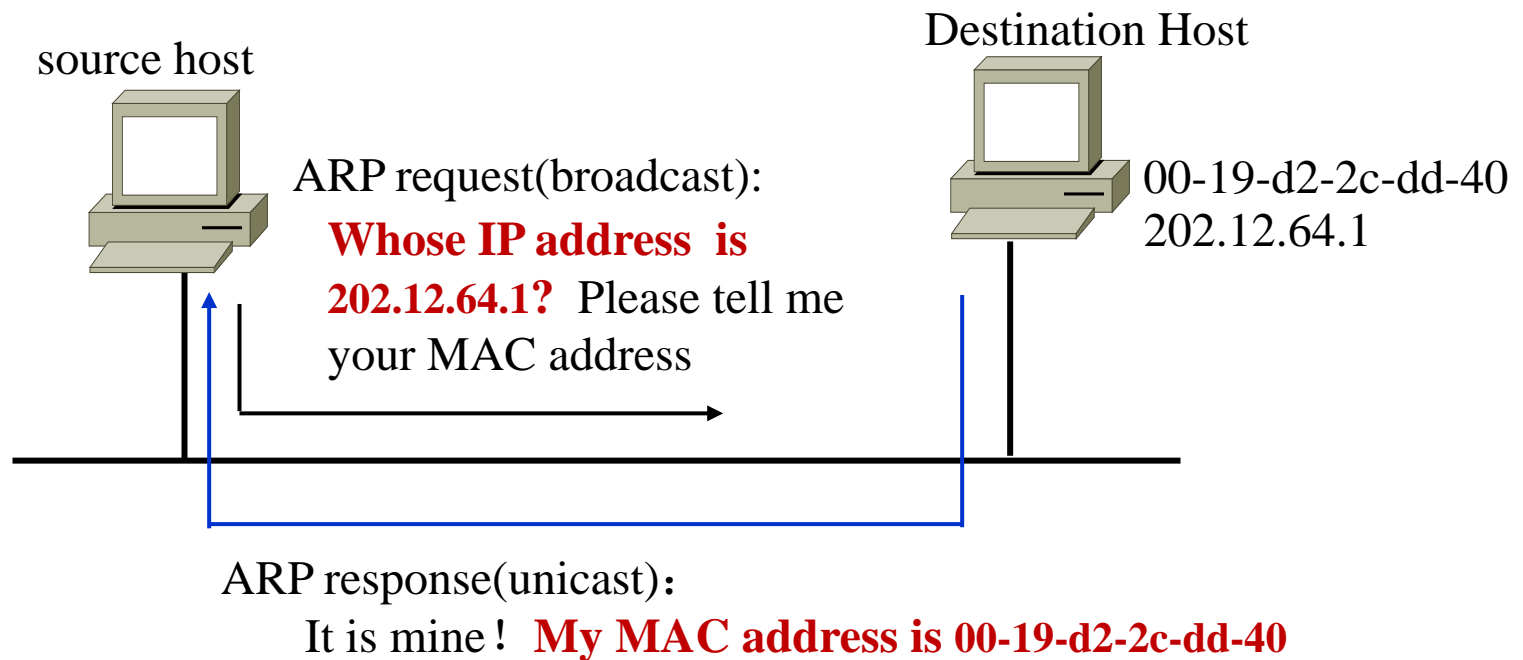


替换低23位

0x01-00-5E-00-00-00

ARP协议 (1)

- ❑ 地址解析协议 (Address Resolution Protocol)可以把IP地址映射为MAC地址。



ARP协议 (2)

- ❑ ARP协议没有超时重传机制。超时没有收到响应，则丢弃引发ARP查询的IP分组。
- ❑ 源主机获得的映射结果缓存在ARP表中<IP address, MAC address, TTL>。其中TTL用于超时删除ARP表项，其取值由系统确定，一般为2~20分钟。
- ❑ 当收到ARP请求，目的主机会缓存源主机的映射，其它主机如果已缓存该映射，则会重置TTL。
- ❑ 也可以把映射直接加入ARP缓存，称为静态ARP映射。静态ARP映射不会因超时而被删除。
- ❑ Windows中的ARP表：

```
管理员: 命令提示符
C:\Users\Administrator>arp -a

接口: 192.168.1.4 --- 0xc
Internet 地址      物理地址      类型
192.168.1.1        00-26-5a-c6-70-7f 动态
192.168.1.255      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
239.255.255.250    01-00-5e-7f-ff-fa 静态

C:\Users\Administrator>
```

ARP协议 (3)

□ ARP包格式

这里采用以太网(硬件)和IP协议举例



RARP(Reverse ARP)用于无盘工作站把MAC地址映射为IP地址。InARP(Inverse ARP)用于NBMA(Non-Broadcast Multiple Access)把VCI映射为IP地址。

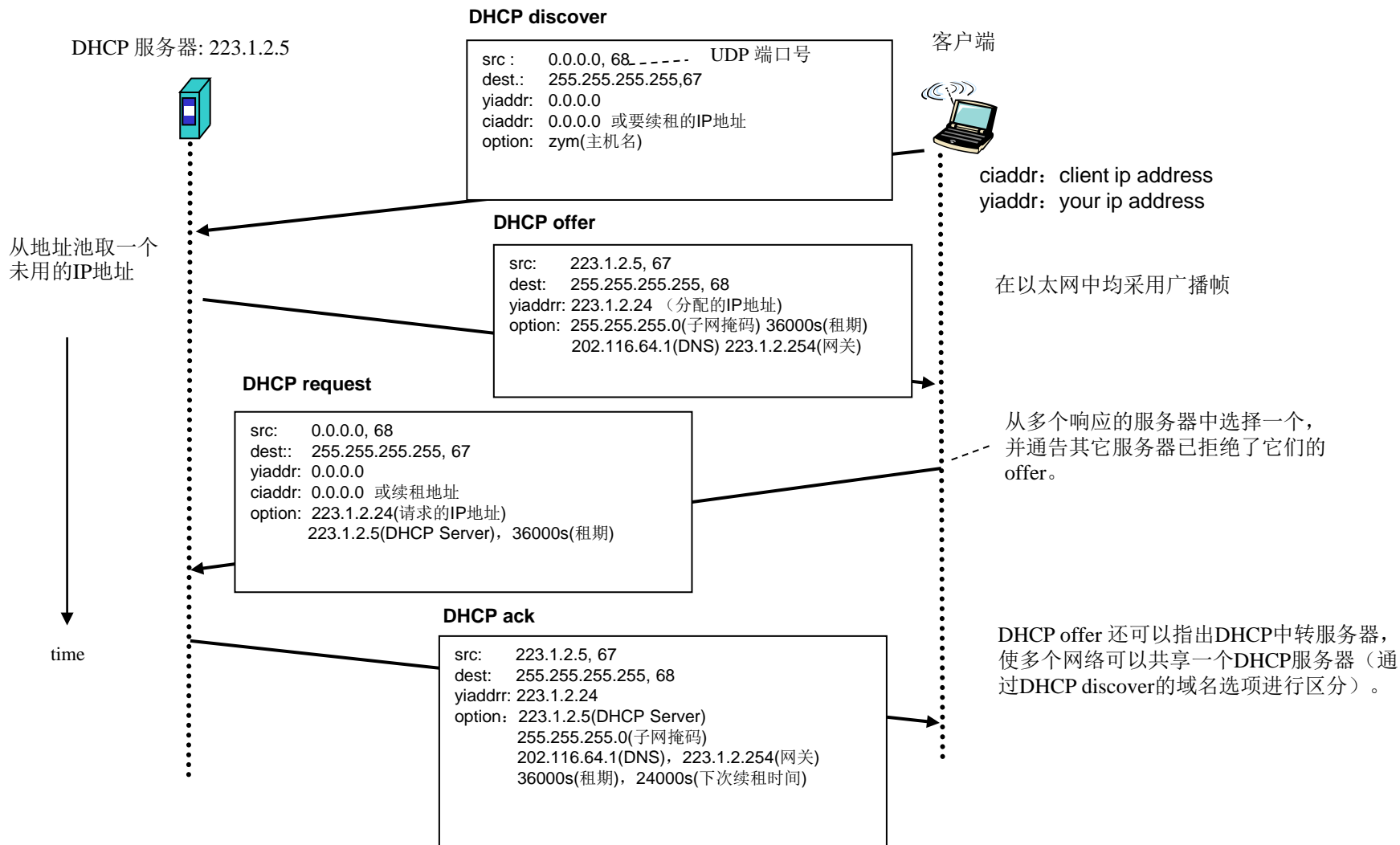
带有ARP包的以太网帧:

Preamble	Dest. Addr.	Src. Addr.	type=0x0806	ARP分组	CRC
----------	-------------	------------	-------------	-------	-----

DHCP协议

<http://www.ietf.org/rfc/rfc2131.txt>

DHCP协议(Dynamic Host Configuration Protocol)用于主机在加入网络时动态租用IP地址。



DHCP数据报

31 25 16 8 0

Operation	HType	HLen	Hops
Xid			
Secs		Flags	
ciaddr			
yiaddr			
siaddr			
giaddr			
chaddr(16字节)			
sname(64字节)			
file(128字节)			
options			

- ✓ Operation: 1-boot request(discover,request)
2-boot reply(offer,ack)
- ✓ HType: 硬件类型, 以太网为1
- ✓ HLen: 硬件地址长度, 以太网为6
- ✓ Xid: 表示一次DHCP会话。
- ✓ Hops: 初始为0, 经过一个路由器加1, 为3时表示循环。
- ✓ Secs: 由客户设置, 表示启动引导进程以来经过的秒数。
- ✓ ciaddr: 客户IP地址。0.0.0.0或要续租的IP地址。
- ✓ yiaddr: 您的IP地址。由服务器分配的IP地址。
- ✓ siaddr: 服务器IP地址。

- ✓ giaddr: 网关路由器IP地址。由中转路由器设置。
- ✓ chaddr: 客户机硬件地址。
- ✓ sname: 服务器名, 以0x00结尾。
- ✓ file: 自举文件名, 包含BOOTP客户端所需的启动映像。
- ✓ optional: 选项, 例如: 最大租期、子网掩码、默认网关、DNS。

- DHCP数据报采用UDP分组进行传送, DHCP Server和DHCP Client的端口号分别为67和68。
- DHCP服务器可以给主机自动分配一个有租期的或者永久使用的IP地址。
- DHCP Message Type(Options): 1-discover,2-offer,3-request,5-ack

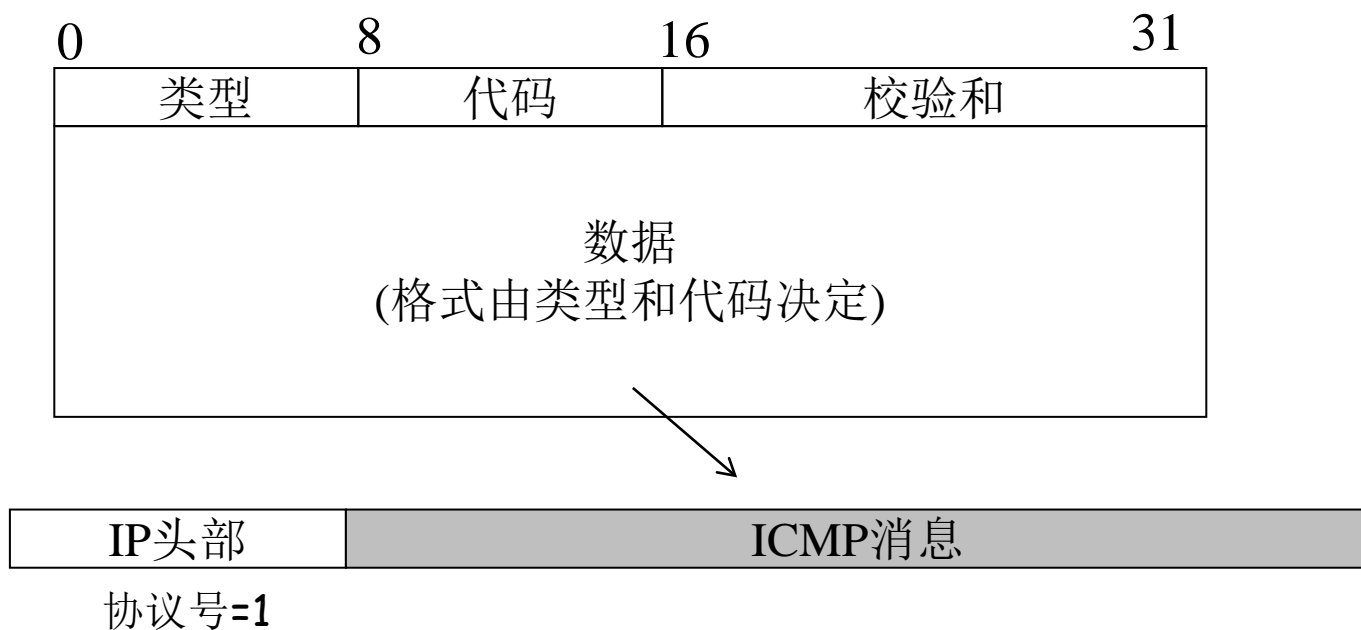
ICMP协议

- ❑ 因特网控制消息协议(Internet Control Message Protocol)用于主机或路由器发布网络级别的控制消息。 <http://tools.ietf.org/html/rfc792>
- ❑ ICMP消息的常见类型:

类型	代码	描述	查询	差错
0	0	回显应答(Ping应答)	√	
8	0	请求回显(Ping请求)	√	
3		目标不可达		
	0	网络不可达		√
	1	主机不可达		√
	2	协议不可达		√
	3	端口不可达		√
	4	需要分段但不可分段(DF=1)		√
	5	源站选路失败(IP Options)		√
	11	由于TOS, 网络不可达		√
4	0	源端抑制 (控制源主机发送速度)		√

类型	代码	描述	查询	差错
5		重定向		
	1	对主机重定向		√
11		超时		
	0	传输期间TTL=0		√
	1	数据报重组超时		√
12		参数问题		
	0	坏的IP头部(各种错误)		√
	1	缺少必要的选项		√
13	0	时间戳请求	√	
14	0	时间戳应答	√	

❑ ICMP消息的一般格式



ICMP不可达消息

8b	8b	16b
类型=3	代码=0~15	校验和
未用(必须填为0)		
原IP头部+原IP数据部份的头64比特		

<u>代码</u>	<u>含义</u>	<u>说明</u>
0	网络不可达	目的地址为私有地址，路由表出错等
1	主机不可达	不能找到到目的网络的路由
2	协议不可达	上层协议不存在
3	端口不可达	UDP端口号没有绑定进程。 如果是TCP协议，则发送TCP 复位消息而不是本消息
4	分段错误	需要分段但是设置了DF
5	源路由错误	IP源路由选项出错

ICMP回响请求和答复消息(Ping)

8b	8b	16b
类型=0或8	代码=0	检验和
标识符		序号
数据		

回响(echo)请求：类型 = 8；回响答复：类型= 0

- 从回响请求收到的标识符、序号和数据将拷贝到回响答复中。
- 标识符和序号用来区分不同的响应。例如：把进程号记录为标识符，并记录发送序号。

ICMP时间超时消息

8b	8b	16b
类型=11	代码=0或1	校验和
未用(必须为0)		
原IP头部+原IP数据部份的头64比特		

代码 说明

0 TTL减为0

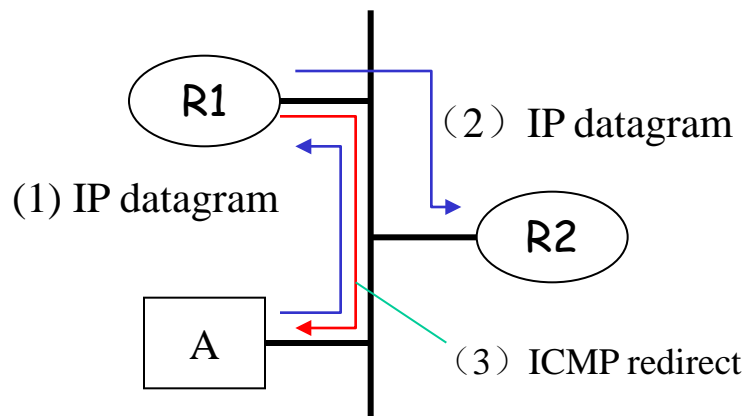
1 重组IP片段超时

ICMP重定位消息

8b	8b	16b
类型=5	代码	校验和
未用(必须为0)		
原IP头部+原IP数据部份的头64比特		

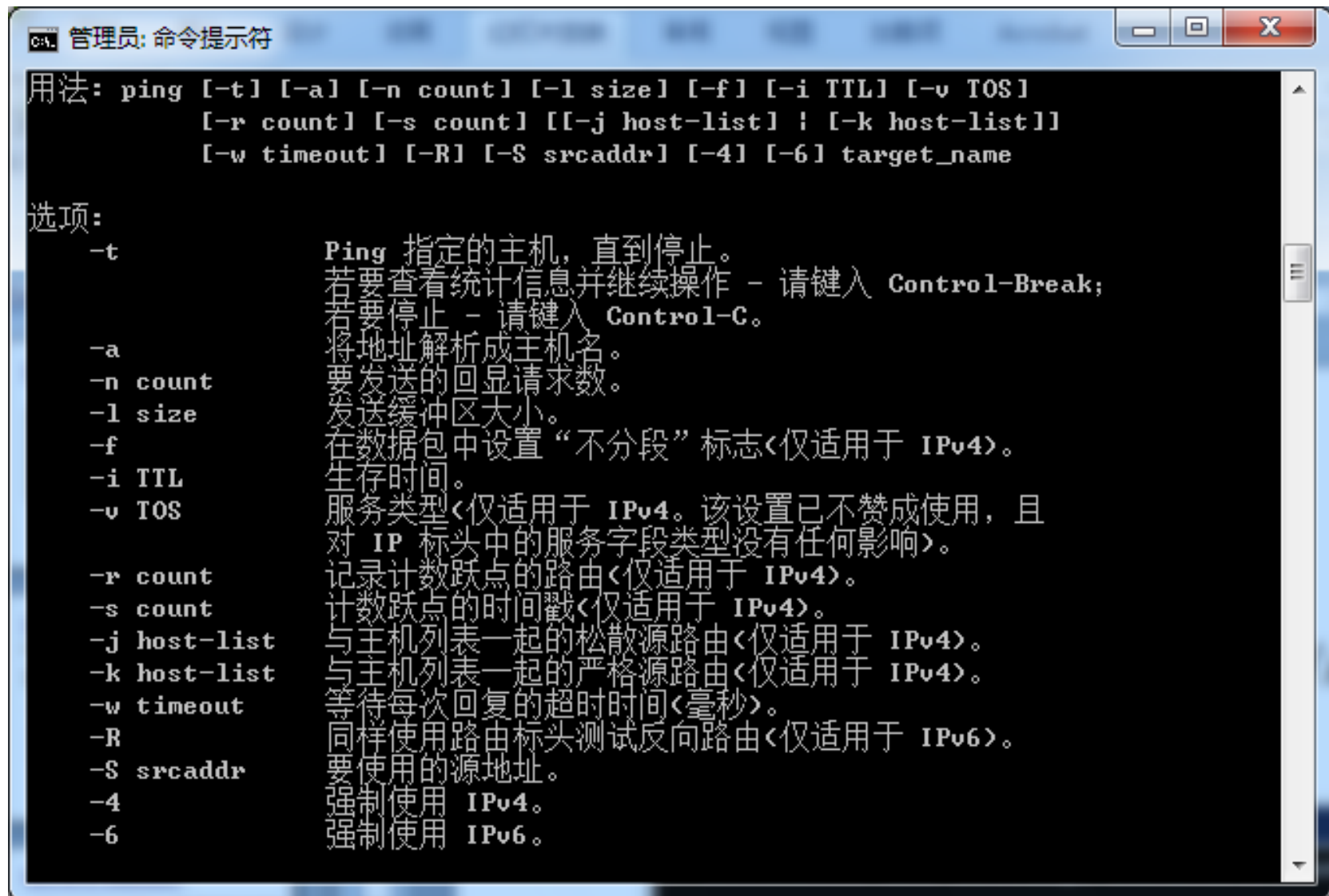
Code Description

- 0 Redirect datagrams for the Network.
- 1 Redirect datagrams for the Host.
- 2 Redirect datagrams for the Type of Service and Network.
- 3 Redirect datagrams for the Type of Service and Host.



当 R1发现把一个数据包转发给R2的接口就是其接收接口，则会把从网络重定向消息发给主机A，要主机A直接把发往这些网络的数据报直接发给R2。

Ping 和 ICMP 协议



```
管理员: 命令提示符

用法: ping [-t] [-a] [-n count] [-l size] [-f] [-i TTL] [-v TOS]
        [-r count] [-s count] [[-j host-list] : [-k host-list]]
        [-w timeout] [-R] [-S srcaddr] [-4] [-6] target_name

选项:
    -t          Ping 指定的主机，直到停止。
                若要查看统计信息并继续操作 - 请键入 Control-Break;
                若要停止 - 请键入 Control-C。
    -a          将地址解析成主机名。
    -n count    要发送的回显请求数。
    -l size     发送缓冲区大小。
    -f          在数据包中设置“不分段”标志<仅适用于 IPv4>。
    -i TTL      生存时间。
    -v TOS      服务类型<仅适用于 IPv4。该设置已不赞成使用，且
                对 IP 标头中的服务字段类型没有任何影响>。
    -r count    记录计数跃点的路由<仅适用于 IPv4>。
    -s count    计数跃点的时间戳<仅适用于 IPv4>。
    -j host-list 与主机列表一起的松散源路由<仅适用于 IPv4>。
    -k host-list 与主机列表一起的严格源路由<仅适用于 IPv4>。
    -w timeout  等待每次回复的超时时间<毫秒>。
    -R          同样使用路由标头测试反向路由<仅适用于 IPv6>。
    -S srcaddr  要使用的源地址。
    -4          强制使用 IPv4。
    -6          强制使用 IPv6。
```

C:\WINDOWS\system32\cmd.exe

C:\Documents and Settings\Administrator>ping -r 3 222.200.160.130

Pinging 222.200.160.130 with 32 bytes of data:

Reply from 222.200.160.130: bytes=32 time=11ms TTL=252

Route: 172.18.240.82 ->
222.200.160.130 ->
172.18.187.254

Reply from 222.200.160.130: bytes=32 time=9ms TTL=252

Route: 172.18.240.82 ->
222.200.160.130 ->
172.18.187.254

Reply from 222.200.160.130: bytes=32 time=14ms TTL=252

Route: 172.18.240.82 ->
222.200.160.130 ->
172.18.187.254

Reply from 222.200.160.130: bytes=32 time=7ms TTL=252

Route: 172.18.240.82 ->
222.200.160.130 ->
172.18.187.254

Ping statistics for 222.200.160.130:

Packets: Sent = 4, Received = 4, Lost = 0 (0% loss),

Approximate round trip times in milli-seconds:

Minimum = 7ms, Maximum = 14ms, Average = 10ms

C:\Documents and Settings\Administrator>_

微软拼音 半:

总结

- **IP**地址空间
- **IP**地址结构
- 有类网和子网划分(VLSM、CIDR)
- 特殊的**IP**地址
- 私有**IP**地址和**NAT**
- 多播**IP**地址
- **ARP**协议
- **DHCP**协议
- **ICMP**协议