第九节 ARP、ICMP 和分组转发过程

一、课程目标

了解网络层相关概念，掌握 ARP、ICMP 和分组转发过程。

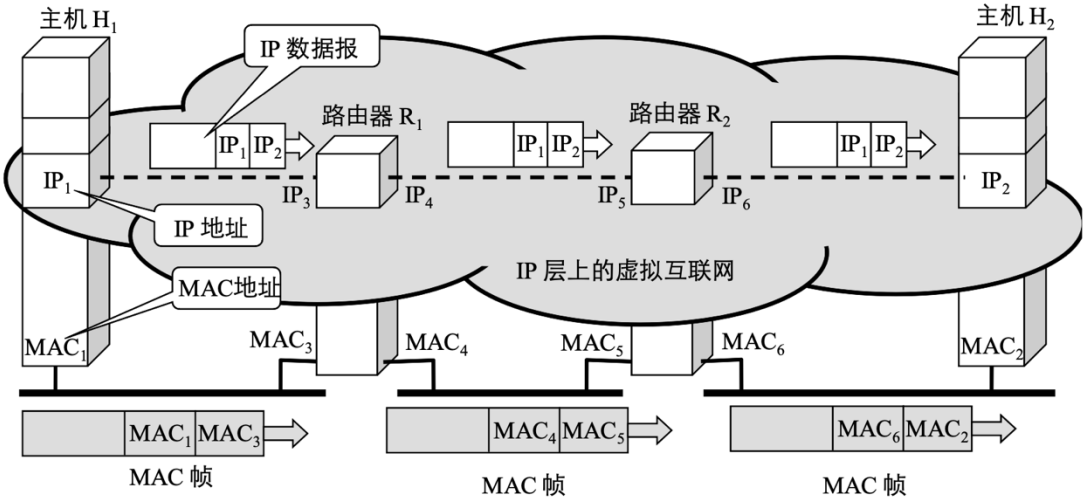
二、课程内容

1、从网络层看数据包的传输过程：

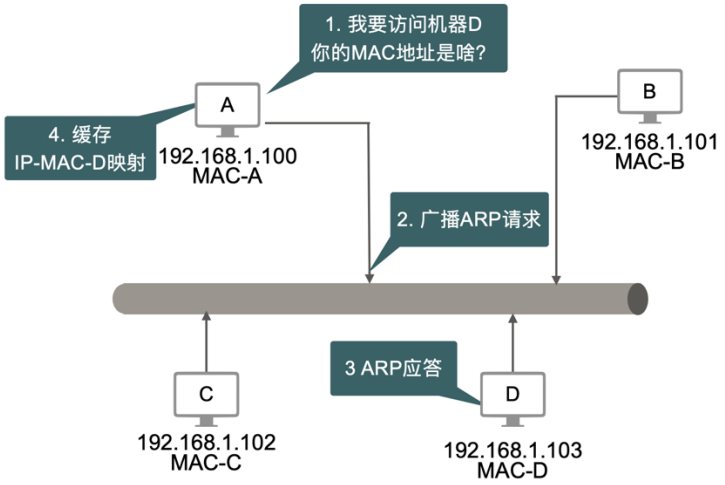
（1）MAC 地址是物理地址，IP 地址是逻辑地址。

（2）有的局域网是异构的，并没有采用 MAC 地址机制，需要对 MAC 地址做地址转换。引入 IP 地址屏蔽数据链路层差异。

（3）硬件地址与物理位置无关，IP 地址与物理地址有关。

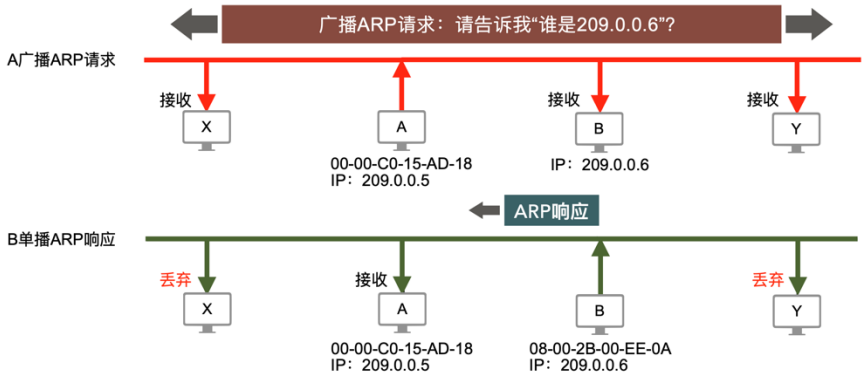


2、**ARP 协议**【RFC 826】：获取同一个局域网上已知 IP 地址的主机（适配器）或路由器的 MAC 地址，**解决 IP 地址和 MAC 地址的映射问题**。ARP 协议不能穿透路由器。



3、ARP 请求以广播帧形式发送，所有主机接收该帧；ARP 响应以单播帧形

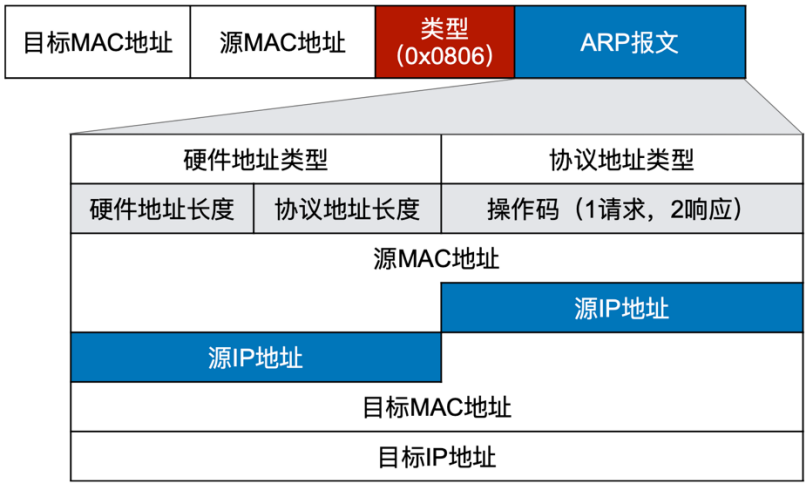
式发送，只有目标主机接收。



4、ARP 缓存：存放最近获得的 IP 地址到 MAC 地址的绑定，以减少 ARP 广播的数量。

```
Mac-mini-2:~ li$ arp -a
? (192.168.1.1) at d4:41:65:ee:5c:c0 on en0 ifscope [ethernet]
? (192.168.1.4) at 80:d6:5:16:c5:7a on en0 ifscope [ethernet]
? (192.168.1.8) at 8c:fe:57:39:8b:1b on en0 ifscope [ethernet]
? (224.0.0.251) at 1:0:5e:0:0:fb on en0 ifscope permanent [ethernet]
```

5、ARP 报文格式




```
Frame 136: 42 bytes on wire (336 bits), 42 bytes captured (336 bits) on inter
Ethernet II, Src: Elitegro_4f:47:d2 (74:27:ea:4f:47:d2), Dst: Broadcast (ff:f
Address Resolution Protocol (request)                          MAC广播帧
  Hardware type: Ethernet (1)  ← 硬件地址类型：以太网
  Protocol type: IPv4 (0x0800) ← 协议地址类型：IP地址
  Hardware size: 6             ← 硬件地址长度：6字节
  Protocol size: 4             ← 协议地址长度：4字节
  Opcode: request (1)          ← 操作码：1表示请求
  Sender MAC address: Elitegro_4f:47:d2 (74:27:ea:4f:47:d2) ← 发送方硬件地址
  Sender IP address: 172.20.28.40 ← 发送方IP地址
  Target MAC address: 00:00:00_00:00:00 (00:00:00:00:00:00) ← 目标硬件地址：未知
  Target IP address: 172.20.31.254 ← 目标IP地址
```

```
Frame 137: 60 bytes on wire (480 bits), 60 bytes captured (480 bits) on inter
Ethernet II, Src: Hangzhou_e1:c9:06 (00:0f:e2:e1:c9:06), Dst: Elitegro_4f:47:
Address Resolution Protocol (reply)                          MAC单播帧
  Hardware type: Ethernet (1)
  Protocol type: IPv4 (0x0800)
  Hardware size: 6
  Protocol size: 4
  Opcode: reply (2)  ← 操作码：2表示响应
  Sender MAC address: Hangzhou_e1:c9:06 (00:0f:e2:e1:c9:06)
  Sender IP address: 172.20.31.254
  Target MAC address: Elitegro_4f:47:d2 (74:27:ea:4f:47:d2)
  Target IP address: 172.20.28.40
```

6、ARP 协议流程

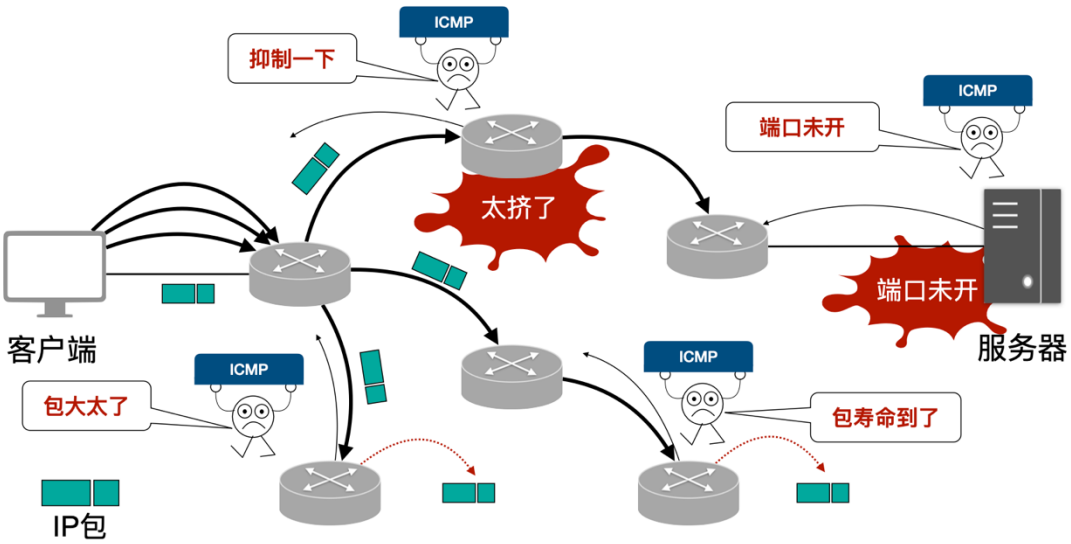
（1）当主机 A 欲向本局域网上的某个主机 B 发送 IP 数据报时，就先在其 ARP 高速缓存中查看有无主机 B 的 IP 地址。

（2）如有，就可查出其对应的硬件地址，再将此硬件地址写入 MAC 帧，然后通过局域网将该 MAC 帧发往此硬件。

（3）如没有，ARP 进程在本局域网上广播发送一个 ARP 请求分组，收到 AQP 响应分组后，将得到的 IP 地址到硬件地址的映射写入 ARP 高速缓存。

7、ICMP 协议：为了更有效地转发 IP 数据报和提高交付成功的机会，在网际层使用网际控制报文协议 ICMP，是一种特殊的 IP 协议（不是更高层协议）。

主要功能：主机或路由器报告差错情况和提供有关异常情况的报告



ICMP 报文分为两种：差错报告报文和询问应答报文。

常见差错报告报文包括（1）终点不可达、（2）超时、（3）参数错误、（4）改变路由(重定向，Redirect) 。

常见询问应答报文包括（1）回送请求和回答报文、（2）时间戳请求和回答报文。

Type	Code	Description	Query	Error
0	0	Echo Reply：回送回答（ping应答）	√	
3	0	Network Unreachable：网络不可达		√
3	1	Host Unreachable：主机不可达		√
3	2	Protocol Unreachable：协议不可达		√
3	3	Port Unreachable：端口不可达		√
3	5	Source routing failed：源站选路失败		√
3	6	Destination network unknown：目的网络未知		√
3	7	Destination host unknown：目的主机未知		√
5	1	Redirect for host：主机重定向		√
8	0	Echo request：回送请求（ping请求）	√	
11	0	TTL equals 0 during transit：传输期间生存时间为0		√
12	0	IP header bad (catchall error)：坏的IP首部（包括各种差错）		√
17	0	Address mask request：地址掩码请求	√	
18	0	Address mask reply：地址掩码应答	√	

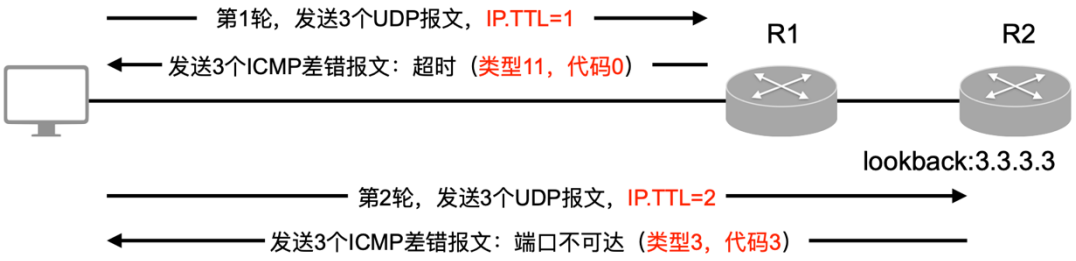
8、ICMP 协议应用：

（1）PING (Packet InterNet Groper) ：连通性测试

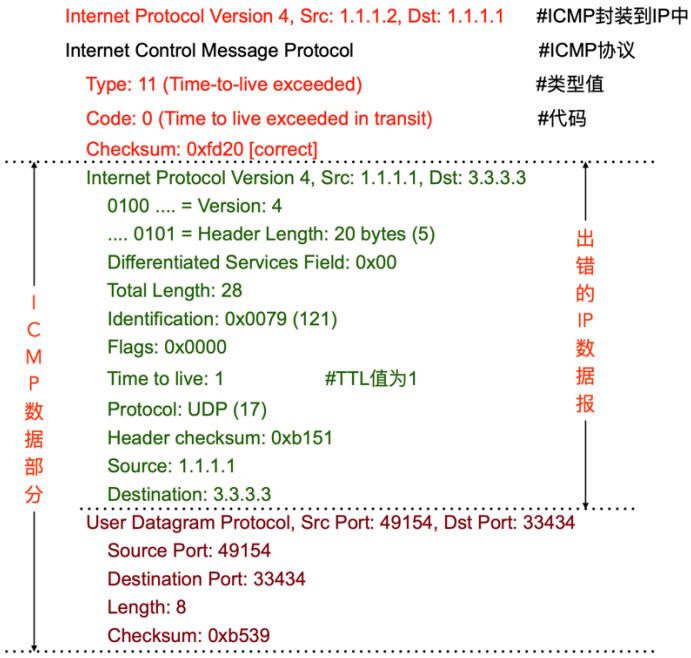
```
li@ubuntu1604:~$ ping -c 3 www.baidu.com
PING www.wshifen.com (103.235.46.39) 56(84) bytes of data.
64 bytes from 103.235.46.39: icmp_seq=1 ttl=38 time=387 ms
64 bytes from 103.235.46.39: icmp_seq=2 ttl=38 time=470 ms
64 bytes from 103.235.46.39: icmp_seq=3 ttl=38 time=472 ms

--- www.wshifen.com ping statistics ---
3 packets transmitted, 3 received, 0% packet loss, time 2000ms
rtt min/avg/max/mdev = 387.418/443.393/472.366/39.596 ms
li@ubuntu1604:~$
```

（2）Traceroute：分组路径跟踪测试

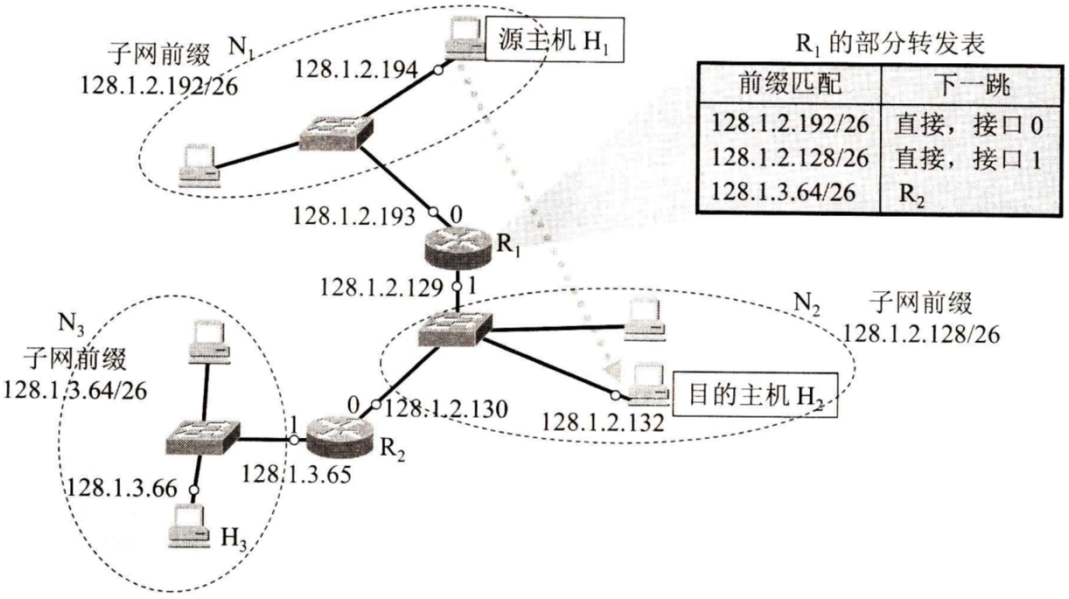


超时差错报告报文：



9、分组转发过程

目前，互联网采用**基于终点（目的地址）的转发**。路由器执行**查表转发**（逐行寻找前缀匹配）操作。



10、**最长前缀匹配**：在采用 CIDR 编址时，如果一个分组在转发表中可以找到多个匹配的前缀，那么就应当选择前缀最长的一个作为匹配的前缀。这个原则称为最长前缀匹配。网络前缀越长，其地址块就越小，因而路由就越具体。

11、两种特殊路由：**主机路由**和**默认路由**。

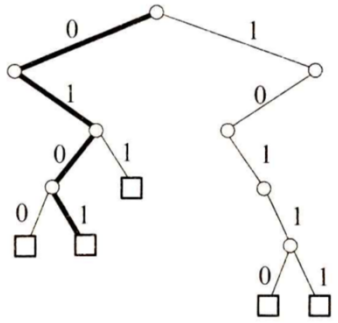
主机路由(host route)：又叫作特定主机路由，这是对特定目的主机的 IP 地址专门指明的一个路由。采用特定主机路由可使网络管理人员更方便地控制网络和测试网络，同时也可在需要考虑某种安全问题时采用这种特定主机路由。在对网络的连接或转发表进行排错时，指明到某一台主机的特殊路由就十分有用。假定这个特定主机的点分十进制 IP 地址是 a.b.c.d，那么在转发表中对应于主机路由

的网络前缀就是 a.b.c.d/32。实际的网络不可能使用 32 位的前缀，因为没有主机号的 IP 地址是没有实际意义的。但这个特殊的前缀却可以用在转发表中。不难看出，32 个 1 的子网掩码和 IP 地址 a.b.c.d 按位进行 AND 运算后，得出的结果必定是 a.b.c.d，也就是说，找到了匹配。这时就把收到的分组转发到转发表所指出的下一跳。**主机路由在转发表中都放在最前面**。

默认路由(default route)：这就是不管分组的最终目的网络在哪里，都由指定的路由器 R 来处理。这在网络只有很少的对外连接时非常有用。在实际的转发表中，用一个特殊前缀 0.0.0.0/0 来表示默认路由。这个前缀的掩码是全 0（/0 表示网络前缀是 0 位，因此掩码是 32 个 0）。用全 0 的掩码和任何目的地址进行按位 AND 运算，结果一定是全 0，即必然是和转发表中的 0.0.0.0/0 相匹配的。这时就按照转发表的指示，把分组送交下一跳路由器 R 来处理（即间接交付）。

12、**二叉线索查找（高效查找）**：IP 地址中从左到右的比特值决定从根节点逐层向下层延伸的路径，二叉线索中的各个路径代表转发表中存放的各个地址。

32 位的 IP 地址	唯一前缀
01000110 00000000 00000000 00000000	0100
01010110 00000000 00000000 00000000	0101
01100001 00000000 00000000 00000000	011
10110000 00000010 00000000 00000000	10110
10111011 00001010 00000000 00000000	10111



三、重点习题

P202：全部