第5章　ARP和RARP

作者：Tim Parker

本章内容包括：

- 使用地址
- 使用地址解析协议
- 使用ARP命令

IP地址是运行TCP/IP协议机器的通用标识，但是 IP地址自身不能使报文到达其目的地。网络系统自身及其行为对网络操作系统和硬件类型而言是特殊的。为了更好地理解数据从源机到目的机的路由行为，需要理解网络是如何与机器互联的。这一章首先考查典型的网络系统(特别是以太网)，看一看TCP/IP如何把IP地址转化成网络能发现的与网络相关的地址。

5.1　使用地址

使用IP地址的目的是帮助TCP/IP把报文传送到正确的目的地，通常有三个和寻址相关的术语：名字、地址和路由。

名字是一台机器、一个用户或一个应用的特殊标识，它通常是惟一的并提供了要传送报文的绝对目标。地址通常标识目标所在地。通常是它在网络内的物理或逻辑位置。路由告诉系统如何把数据报送到正确的地址。当使用术语地址时要注意，它经常和通信协议一起使用指代许多不同的内容，它可以是目的地、一个机器的端口、一个存储器位置、一个应用或其他。

接收方的登录名通常是整个传输过程的关键。一个称为名字服务器的软件包用来把用户名和机器名解析成地址和路由，对用户隐藏了 TCP/IP的路由和发送的内部机制。除了使寻址和路由对端用户透明之外，使用名字服务器有另一个好处：它给网络管理员按照需要改变网络带来了很大的自由度，而不需单独更新每个用户的机器。只要一个应用能够访问位于某处的名字服务器，应用和用户就能忽略路由的变化。

5.1.1　子网寻址

当用户把一块数据发送到另一台机器时，经常通过 IP地址来完成。虽然TCP/IP设计成围绕着IP地址工作，但是实际的网络软件和硬件却不是这样。相反，网络使用编码至网络硬件中的地址来识别每一台机器。从 IP地址得到物理地址，不是 TCP/IP协议的标准部分，所以开发了许多特殊的协议来完成这一部分任务。这些协议在下一节讨论，但是这一节首先考查网络物理地址是如何构成和如何处理的。在一个局域网中，有一些保证报文正确传输的必要信息，其中基本的信息是目的机器的物理地址和数据链路层地址。二者很重要，值得进一步讨论。

1. 物理地址

网络上的每一个设备有一个惟一的物理地址 (physical address)，有时被称为硬件地址或数据链路地址。对于网络硬件而言，地址通常编码到网络的接口卡中。物理地址有时可以通过开关或软件由用户设置。更常见的情况是，这些地址用户根本不能改变，因为一个惟一的编号已经编到可编程只读存储器 (PROM)中；生产商经常联手以保证在任一个网络上地址没有冲

突的可能性，一个地址只能出现一个。否则，名字服务器将没有办法来区分目标机。物理地址的长度依赖于网络系统。举例来说，以太网和其他一些网络使用 48位地址。为了通信需要有两个地址：一个地址标识发送设备；一个用于接收设备。

IEEE负责分给子网配置惟一的物理地址 (以前由Xerox公司完成这项工作，Xerox开发了以太网)。对每一个子网， IEEE分配一个 24位长的组织惟一标识 (organization unique identifier,OUI)，其他24位由各组织随便分配。实际上 24位OUI中的2位是控制位，所以只有 22位标识组织。OUI的格式由图 5-1示出。24位的OUI和24位局部分配位的组合称为媒体访问控制(media access control,MAC)地址。当一个数据报文要组装在 TCP/IP上发送时，要有两个MAC地址，一个来自发送机，一个标识接收机。

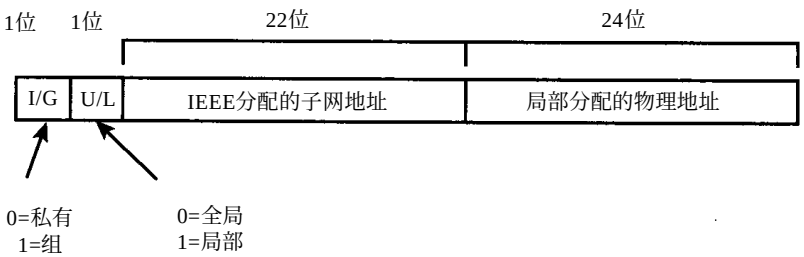

图5-1 组织惟一标识结构

地址的最低位 (结构中的最左一位)称为私有或组位。假如这一位设为 0，剩余的地址是一个私有地址，设为 1意味着剩余的地址域标识一个需要进一步解析的组地址。假如整个的 OUI设为1，则整个网络上的每个站都是目的地，这是 OUI支持的一个特殊约定。

OUI结构中的第 2位称为局部或全局位。假如第二位设为 0，则它是被全局管理团体设置的。这是分配给 IEEE OUI的。假如第二位设为 1，则OUI是局部分配的，如果按 IEEE分配的地址解码就会导致问题。通常，第 2位设为1的结构要维持在局域网或广域网内部，不能传送到按IEEE地址格式编码的网络中。

OUI结构中的剩余 22位组成子网的物理地址，由 IEEE分配。剩下的一组 24位标识局域网地址，可以局部管理。假如一个组织要用光物理地址 (24位可有1600万个地址)，IEEE能为这一组织再分配一个子网地址。

2. 链路层地址

IEEE以太网标准使用另一个称为链路层地址的地址 (常被缩写为LSAP-link service access point)。LSAP标识数据链路层中链路协议的类型。和物理地址一样，一个数据报将携带接收端和发送端的LSAP。

3. 网络帧

被传输数据报文的信息排列因所使用的网络协议不同而有所区别。然而，看一看前面提到的地址和其他相关信息是如何在数据发送到网络上之前附加在数据报上的情形将是有启发性的。我们将以与 TCP/IP一起广泛使用的以太网为例。以太网和其他的系统非常相似，虽然头结构可能不同。记住这是网络协议封装由 TCP/IP创建的头的方式，和具体的 TCP/IP没多大联系。典型的以太帧(准备好网络发送的报文)如图5-2所示。

前导	接收方地址	发送方地址	类型	数据	循环冗余校验
64位	48位	48位	16位	可变长	32位

图5-2 以太帧结构

64位的前导符用于通信过程的同步和消除发送前几位时的随机噪声。前导域的后面是一系列位称为帧定界符(start frame delimiter,SFD)，指示帧紧接在后面。

以太帧中接收方、发送方的以太网地址使用 IEEE的48位格式，其后跟着 16位的类型域，用于标识使用的协议类型。实际的数据 (由TCP/IP组装的数据报)跟在类型指示后面。对标准以太网而言，数据域在 46和1 500字节之间。如果数据小于 46字节，则补零直到46字节。以太帧的最后是循环冗余校验 (cyclic redundancy check,CRC)，用于保证帧的内容在传输过程中没有被改变。传输路径上的每一台机器计算帧的 CRC并和帧最后的 CRC比较。假如二者一样，帧就能继续沿网络传送或传到子网内部；否则，帧必定被改过，并且应该被丢弃。

一些和以太网相关的协议，如 IEEE 802.3，使用的整个帧排列是一样的，但对一些内容进行了一些更改。802.3把用于识别协议类型的 16位数替换成指示数据长度的值。同样，一个新的域附加在数据区前面。

5.1.2 IP地址

正如读者所知， TCP/IP使用32位地址用于标识网络上的机器以及和它相连的网络。 IP地址标识机器到网络的连接，而不是机器自身，这二者有重要的不同之处。但是一台机器在网络上的位置发生变化时， IP地址有时也要改变，这要依赖于网络建造的方式。 IP地址是一组许多人在他们自己的工作站或终端上看到的数，如 127.40.8.72，这个数惟一地标识了设备。IP地址由 4组8位数组成，总共 32位。 IP地址只能由网络信息中心 (Network Information Center,NIC)分配，但是如果一个网络不与 TCP/IP相连，也能决定自己的编号。 IP地址使用的十进制表示法称为点分四元表示法。

依赖网络的大小，有四种 IP地址格式， 4种格式从A类到D类，如图5-3所示。地址类别由前几个位序列标识，在此图中， A类使用1位，D类使用4位。类别可以由前3位(高位)决定。实际上，大多数情况下，前两位足够了，因为 D类网很少。

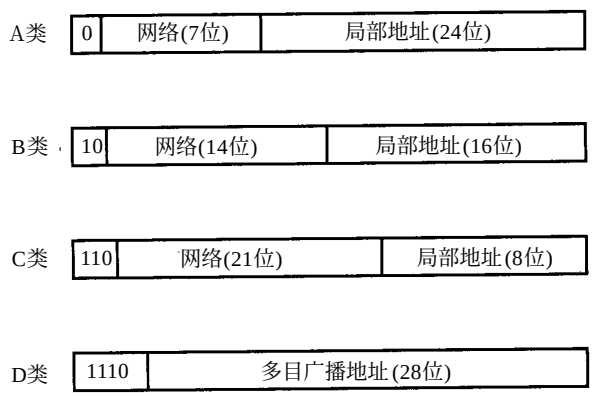

图5-3 类IP地址结构

A类地址用于有许多机器的大型网络。这种情况需要24位的局部地址(也被称为主机地址)。网络地址7位，限制了可识别的网络数量。 B类地址用作中等规模的网络， 16位的局部或主机地址， 14位的网络地址。

C类网络只有8位表示局部或主机地址，限制了设备数最多为 256个，C类使用21位表示网络地址。最后， D类网络用作组播，当需要一个以上的设备接收时，采用这种地址类型。 IP地址中每个段的长度经过了仔细挑选以提供最灵活的网络和主机地址分配。

通过IP地址，网关能够决定数据是发送到 Internet上(或其他互联网络)还是保留在局域网内部。假如网络地址和当前报文地址的网络部分相同 (路由到局域网设备，称为直接主机)，就不用把报文送到互联网络上；所有其他网络地址路由到网关，离开局域网 (称为间接主机)。

一台机器(特别是网关)如果连接了多个网络，那么这台机器可能有多个IP地址。这样的机器叫做多穴主机，因为它们与每一个相连的网络对应一个惟一地址。两个网络被一台网关相连可以有相同的网络地址，这里有一个寻址问题，因为网关必须能区别物理地址在哪个网络上。这个问题由一个专门进行地址解析的特殊协议来解决，这个协议称为地址解析协议(Address Resolution Protocol, ARP)。

5.2 使用地址解析协议

在局域网或广域网上把报文从一台机器发送到另一台机器，如果不知道目的机器的物理地址就会发生问题。需要一些方法能把IP地址(应用提供)解析为和网络相连的机器的硬件物理地址。

一种强制的方法是在每台机器上建一个用于把 IP地址转化成物理地址的解析表。于是，当一个应用程序要把数据发送到另一台机器时，软件就检查这个转换表以得到物理地址。这个方法存在一些问题，这些问题也就是几乎没有这样实现的原因。最致命的缺点是一旦发生一些变化，每台机器上的地址表就要更新。

为了解决这个问题，开发了地址解析协议 (ARP)。ARP的任务是把IP地址转化成物理地址，这样做，就消除了应用程序需要知道物理地址的必要性。用最简单的术语表述， ARP就是把IP地址转换成相应物理地址的转换表。这个表称为 ARP表。ARP表的排列如图5-4所示。ARP在存储器中维护一个 cache，这个cache称为ARP cache，通常搜索ARP cache进行匹配，假如没有匹配成功就检查ARP表。

	IF索引	物理地址	IP地址	类型
表项1				
表项2				
表项3				
表项n				

图5-4 ARP表排列，表中的每一行表示Cache中的每一个设备

5.2.1 ARP cache

ARP cache中的每一行对应一个设备，每一个设备存储以下信息：

- IF索引 —— 物理端口(接口)。
- 物理地址 —— 设备的物理地址。
- IP地址 —— 和物理地址对应的IP地址。
- 类型 —— 这一行对应的表项类型。

类型有4种可能的值，值2意味着表项是无效的，值3意味着映射是动态的(表项可能改变)，值4说明是静态项(表项不变化)，值1意味着不是上面的任何一种情况。

当ARP解析一个IP地址时，它会搜索ARP cache和ARP表作匹配。如果找到了，ARP就把物理地址返回给提供IP地址的应用，假如ARP没找到一个匹配的IP地址，它就会向网络上发布消息，这个称为ARP请求的消息，被广播到局域网上的每一个设备。

ARP请求包括接收设备的IP地址。假如一个设备认出此IP地址属于自己，就把包含自己物理地址的应答报文返回给产生ARP请求的机器，ARP请求机器会把此信息放到ARP表和ARP cache中以备将来之用。使用这种方式，ARP能决定任何IP地址对应机器的物理地址。

ARP请求和ARP应答报文的格式如图5-5所示，当一个ARP请求发出时，除了接收端硬件地址(正是请求机想知道的)之外所有域都被使用。ARP应答中，使用所有的域。

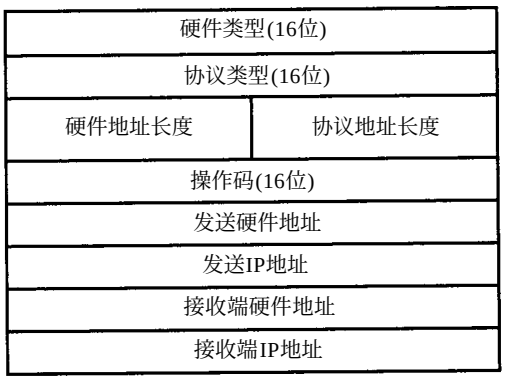

图5-5 ARP请求和应答报文格式

ARP请求和ARP应答报文中的域有几种值，本节的剩余部分会仔细解释各个域及其使用。

1. 硬件类型

硬件类型识别硬件接口类型，合法的值是：

类 型	描 述
1	以太网
2	实验以太网
3	X.25
4	Proteon ProNET(令牌环)
5	混沌网(chaos)
6	IEEE 802.X
7	ARC网络

2. 协议类型

协议类型标识发送设备所使用的协议类型， TCP/IP中，这些协议通常是EtherType，合法的值是：

十 进 制	描 述
512	XEROX PUP
513	PUP地址翻译
1536	XEROXNS IDP
2048	网际协议
2049	X.75
2050	NBS
2051	ECMA
2052	混沌网络
2053	X.25第三层
2054	地址解析协议(ARP)
2055	XEROX网络系统
4096	伯克利Trailer
21000	BBN Simnet
24577	DEC MOP (维护操作协议)Dump/Load
24578	DEC MOP (维护操作协议)远程控制
24579	DEC DECnet Phase 5
24580	DEC LAT
24582	DEC
24583	DEC
32773	HP Probe协议
32784	Excelan
32821	反向地址解析协议(RARP)
32823	AppleTalk
32824	DEC 局部网桥协议

如果协议不是EtherType，允许使用其他值。

3. 硬件地址长度

数据报中硬件地址以字节为单位的长度。

4. 协议地址长度

数据报中所用协议地址以字节为单位的长度。

5. 操作码

操作码（Opcode）指明数据报是ARP请求还是ARP应答，假如是ARP请求，此值为1；假如数据报是ARP应答，此值为2。

6. 发送方硬件地址

发送方设备的硬件地址。

7. 发送方IP地址

发送方设备的IP地址。

8. 接收方硬件地址

接收方设备的硬件地址。

9. 接收方IP地址

接收方设备的IP地址。

5.2.2 代理ARP

本章前面已提及，当两个网络通过网关联接时能够有相同的网络地址。网关必须能确定进来报文的物理地址或IP地址在哪一个网络上。网关能通过改进的 ARP —— 代理ARP来完成这个任务(有时称为杂收ARP)。

代理ARP创建一个ARP cache，其中包含这两个网络中设备的有关信息。网关必须管理穿越于两个网络的ARP请求和应答。通过把两个ARP cache组合成一个，代理ARP扩充了地址解析过程的灵活性，防止产生过多的ARP请求和ARP应答报文穿越网关。

5.2.3 反向地址解析协议

ARP协议有一个缺陷：假如一个设备不知道它自己的 IP地址，就没有办法产生 ARP请求和ARP应答。网络上的无盘工作站就是这种情况。设备知道的只是网络接口卡上的物理地址。

一个简单的解决办法是使用反向地址解析协议 (RARP)，RARP以与ARP相反的方式工作。RARP发出要反向解析的物理地址并希望返回其 IP地址，应答包括由能够提供信息的 RARP服务器发出的IP地址。虽然发送方发出的是广播信息， RARP规定只有RARP服务器能产生应答。许多网络指定多个RARP服务器，这样做既是为了平衡负载也是为了作为出现问题时的备份。

5.3 使用ARP命令

大多数TCP/IP实现(并非全部)给用户提供了检查ARP cache的方法。Unix arp命令显示出机器中所有的ARP cache表项。为了看cache内容可以使用带-a(for all)选项的命令：

```
$ arp -a
brutus <205.150.89.3> at 0:0:d2:03:08:10
```

这个例子中，叫 brutus的机器具有IP地址205.150.89.3，其MAC(媒体访问控制)地址是0:0:d2:03:08:10。

很少使用arp命令，除非网络管理员想解决 IP地址重复问题。假如两台机器具有相同地址(但有不同的MAC地址)，管理员就需要使用这个命令显示 arp cache。

5.4 小结

在这一章中，读者已经了解了地址解析协议的通常用法： ARP、代理ARP和RARP，本章首先考查了以太网帧如何对 TCP/IP数据报文进行封装。理解了这些互联网络中的复杂层次，读者应该对TCP/IP的功能和数据在网络上传送到其目的机器的过程有了更好的认识。