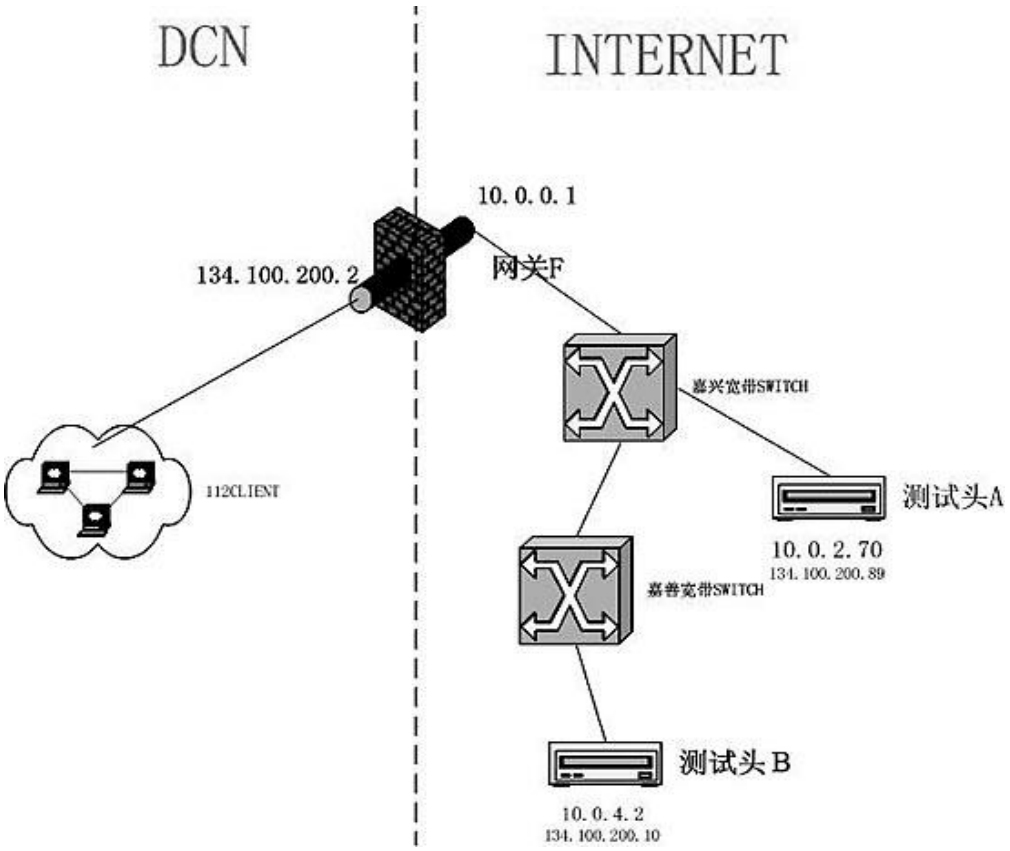
使用 Sniffer 分析 ARP 网络故障

电信网络内部一套 112 测试系统，涉及到一系列服务器和测试头(具有 TCP/IP 三层功能的终端)，原有的拓扑在电信内网(DCN)中。由于测试范围的扩大，有些机房没有内网接入点，变通的方案是在城域网上建立一个 VPN，将那些没有 DCN 接入点的测试头设备接在此 VPN 上 ,然后此 VPN 通过一个防火墙(PIX)与 DCN 做接口。可以将这些测试头看作一些提供测试服务的服务器，使用 NAT 静态转换将这些测试头映射为 DCN 内网网段上的 IP 地址 ,内网的 一些客户端使用这些映射后的地址访问测试头。

方案实施后，用 DCN 内网设备访问有些测试头，时通时不通，对这些局点的 112 测试工作带来了极大的困扰。通过使用 Sniffer 抓包工具，结合对 ARP 协议的理解，逐步分析出了故障的真正原因，解决了问题。

故障现象说明

112 系统的部分网络拓扑图如图 1 所示。



故障现象

1.DCN 中的 112CLIENT 有时访问不到测试头 A。112CLIENT ping 不通测试头 A，网关 F 上也 ping 不通测试头 A。

2. F 上始终有 ARP 记录：例如嘉兴某 NPORT 测试头 A

Internet 10.0.2.70 118 0090.e809.b82f ARPA FastEthernet0/1

3. 如果 F 上 clear arp，则 112CLIENT 再 ping，可以 ping 通。

4. 如果不采取步骤 3，用 DCN 内机器 telnet 134.100.200.10(测试头 B)，再用 B 来 ping 10.0.2.70(测试头 A)，能 ping 通。再用 112CLIENT ping A，能 ping 通。

5. 将测试头换下，换上同 IP 地址笔记本电脑，没有任何问题。

对问题的预先判断中，有两种倾向性猜测，如下：

◆ A：NPORT 测试头的 TCP/IP 实现不规范。测试头是厂家应局方要求加工组装的，其 TCP/IP 协议簇的实现是建立在 NPORT MOXA 卡上的，主要是为了实现 TCP/IP 与 SERIAL 协议之间的转换。而这种实现的可靠性并没有 100%的把握。如果是这个原因，需厂家解决。

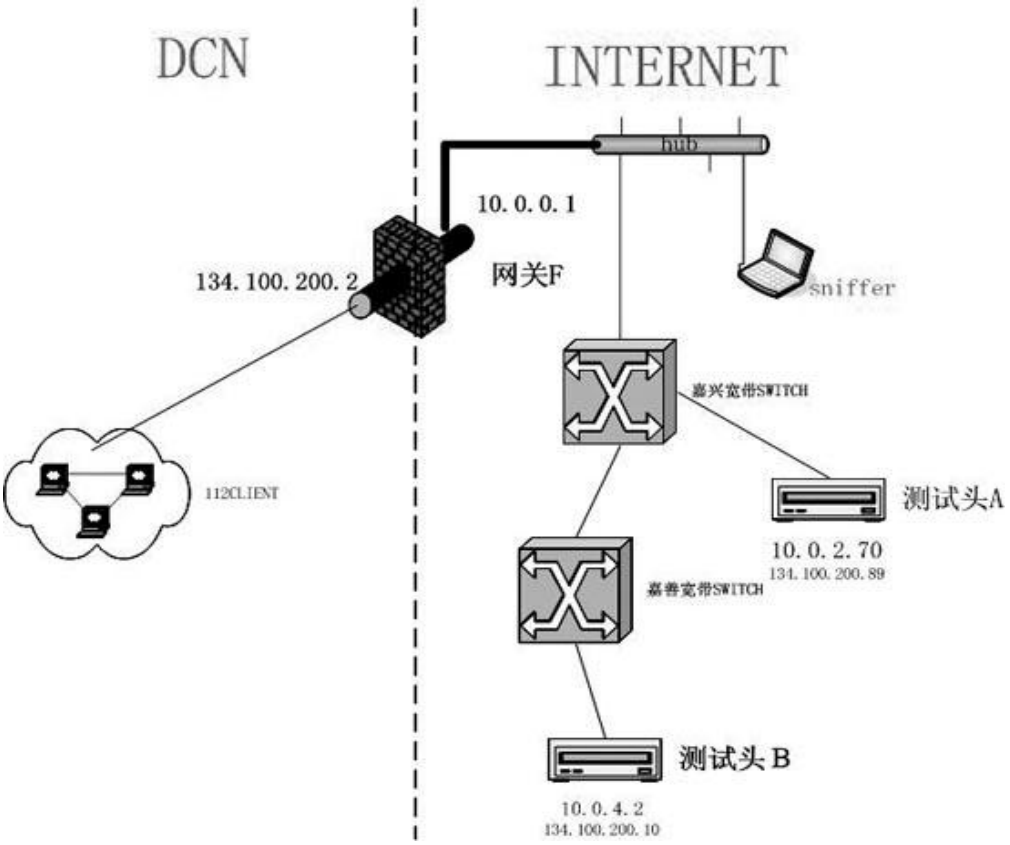
◆ B：宽带交换机的设置不科学。交换机的 ARP 条目失效时间对其 ARP 对照表有很大影响，设的太短，很快就失效，包过来后就会不知道流向哪个端口，会被交换机丢弃。宽带交换机属于数据部门维护，一般情况下不会提供给我们口令，没有确实的判断，他们一般不愿意改交换机设置。

所以确实的定位问题的所在，是我们解决故障的先决条件。

查找故障源

在不能确定故障源的情况下，我们同时从以上两种倾向性猜测的角度出发，力图从两个方向做出解释，最后找出符合实际的故障点。

首先，改变拓扑结构如图 2 所示，网关接口之一连接一台共享带宽的 HUB，HUB 上的两个端口分别连接宽带部分和一台运行 Sniffer 的电脑。这样，Sniffer 能“抓”到所有宽带与网关 F 之间的包。



针对现象一：IDSCLIENT ping 不通测试头 A

测试动作一：

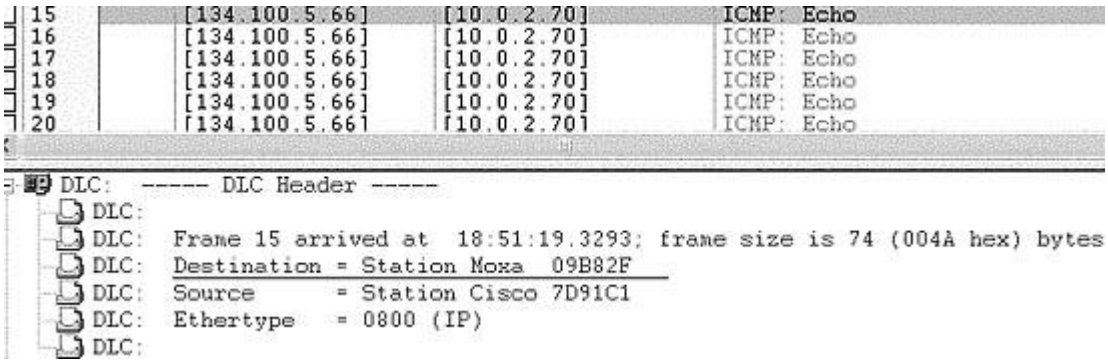
1)网关 F 上有 A 的 ARP 记录。

112_edge#sh arp | include 10.0.2.70

Internet 10.0.2.70 3 0090.e809.b82f ARPA FastEthernet0/1

2)用内网的 IDSCLIENT 来 ping A，结果 ping 不通。

用 Sniffer 抓包，从图 3 中可以清楚地看出，ICMP 探测包从网关 F 准确地向目的 A 10.0.2.70(09B82F)发送,但 A 没有回响应包。所以结果为 ping 不通。



基于两种猜测，故障的原因可能解释有：

解释 A：应该为 A 的 ARP 缓存中没有网关 F 的 ARP 记录，所以 A 找不到网关的 MAC 地址，而且它对这种“找不到网关的 MAC 地址”不作为(NPORT 测试头对 ARP 的实现不完善)。

解释 B：连接测试头 A 的宽带交换机中的 MAC 对端口的对应记录过期，在 MAC 地址表中目的 MAC 地址无对应端口，交换机丢掉此包。

针对现象二：将测试头换下，换上同 IP 地址笔记本电脑，没有任何问题。

测试动作二：

1)A 的位置换上一台电脑 hongjing(IP 与 A 一致),且让网关 F 有 hongjing 的 ARP 记录。

以下是引用片段：

112_edge#sh arp | include 10.0.2.70

Internet 10.0.2.70 3 000b.dbe0.1de9 ARPA FastEthernet0/1

2)IDSCLIENT2(134.100.5.52) ping 10.0.2.70(HONGJING),能 ping 通。

基于两种猜测，故障的原因的解释有：

解释 A：包从网关 F 中发过来，ICMP 探测包准确的发送到目的 A 10.0.2.70,hongjing 同样由于本机 ARP 缓存中没有网关 F 的记录，不能立即发送 ICMP 回应包。但 hongjing 没有“不作为”，而是根据 ICMP 包的源 IP 地址跟自己的掩码判断此 ICMP 查询包发自广播域外，所以 hongjing 当机立断，向本广播域发起 ARP 查询 ,要查出网关 10.0.0.1 的 MAC 地址 ,查到后 ,将 ICMP 回应包发送到 10.0.0.1,所以网络能通。

对比动作一，动作二的网络包分析，不难发现问题所在。相同的条件与情况下，产生“通”与“不通”的两种结果，关键在于测试头(A)与电脑(hongjing)对 ICMP 查询包的“态度”不一样所致。电脑 hongjing 的态度“积极”，当没有该包的传递者 F 的 MAC 地址时 ,会想方设法找到“回答”的路径 ,并“回答”。而测试头 A 的态度“消极”,收到询问包时 ,发现自己没有该包传递者 F 的 MAC 地址时，没有采取任何措施，保持“沉默”，所以没回答。

解释 B：笔记本电脑 hongjing 一接上交换机后立刻发出广播包，通知局域网内其他机器，hongjing 的 MAC 地址是多少。此时，交换机记下 hongjing-MAC 与端口的映射。所以包从网关 F 过来后，能到达测试头 A。

针对现象三：“如果 F 上 clear arp，则 112CLIENT 再 ping ，可以 ping 通”。

测试动作三：

登录网关 F,执行 clear arp 命令，然后在内网中，用 IDSCLIENT ping A，结果可以 ping 通。

基于两种猜测的原因解释：

解释 A :本来由于测试头的“消极”,是不通的。但网关 F 上执行了 clear arp 命令后，网关 F 由于 ARP 地址影射清空，F 不知网关的 MAC，会向广播域发送 ARP 包，该包中包含了自己的 MAC 地址。根据 RFC826，虽然广播域中的机器不会回应此包，但会将 F 的 MAC 地址记录到 ARP 缓存中，所以能使得本不通的 112CLIENT pingA 能 ping 通。

解释 B :网关 F 上执行了 clear arp 命令后 ,网关 F 由于 ARP 地址映射清空，F 不知网关的 MAC ,会向广播域发送 ARP 包 ,该包中包含了自己的 MAC 地址。测试头 A 上连的交换机会将 F 的 MAC 地址和相关端口绑定;A 回应此 ARP 请求时 ,交换机又会将 NPORT 测试头 A 的 MAC 地址与相关端口绑定。所以后续的连接能通。

针对现象四：“用 DCN 内机器 telnet 134.100.200.10(测试头 B)，再用 B 来 ping 10.0.2.70(测试头 A)，能 ping 通。再用 112CLIENT ping A，能 ping 通。”

测试动作四：

用内网机器 IDSCLIENT telnet 到 134.100.5.66，然后从 134.100.5.66 上 ping 测试头 B，结果本来 ping 不通的，现在可以 ping 通了。

基于两种猜测的原因解释：

解释 A：此现象用猜测 A 解释不了。

解释 B：测试头 B 向测试头 A ping 时，先会发 ARP 广播，测试头 B 回应此 ARP 请求。这个过程中，A 上连的交换机会将 A<->相应端口，B<->相应端口的记录记在地址端口映射表。

所以 F 到 A 的包就能通了。

至此，可以排除猜测 A。同时，由于同一批次的 NPORT 测试头在其他地区及内网用的比较正常，所以，倾向于猜测 B。为进一步证实猜测 B，进一步做了以下测试。

做动作一的时候，在交换机与 A 间抓包。看是否有源地址为 F 的物理地址，目的地址为 A 的物理地址的包从交换机端口出来，结果确实无包被监听到，所以，从理论上得出，猜测 B 是正确的。从理论上定位出正确的故障原因后，我们理直气壮的联系数据部门，请他们修改了部分交换机的 ARP 失效时间。经过一段时间的检验，系统运行良好，原有故障消失