第40章　问题解决工具及要点

作者：Bernard McCargo

本章内容包括：

- 监视网络行为
- 标准应用程序
- 解决网络接口问题
- 解决网络层(IP)问题
- 解决TCP和UDP问题
- 解决应用层问题

引起大多数问题的原因非常简单，因此准确的定义它是解决问题的关键。但是事情并非总是如此，因此本章首先讨论帮助用户解决疑难问题的工具。这些工具既有价值上万美元的商业软硬件产品，也有互联网上发布的免费软件。本章主要讨论免费的诊断工具。

主要有以下几点原因：首先，商业系统的价格昂贵；其次许多管理员不能购买商业软件，但任何人都可获得免费的诊断工具。最后一点是大多数问题都可以使用免费的诊断工具解决。大型的网络可能需要商业软件如TDR，但对于系较小的网络，免费的诊断工具足够使用。

本章使用的工具在RFC1147中有详细的描述。

40.1　监视网络行为

为了发现问题，用户必须对监视网络有基本的了解。如果没有第三方或嵌入式网络监视系统，解决网络问题将花费大量的时间。

监视网络行为可使用户在问题发生前对网络进行预防性维护。当发生问题时，它可以为用户搜集问题的细节信息。因此，当报告有问题发生时，用户可以检查网络监视系统的记录，寻找哪项服务失败。远程主机名和 IP地址是什么？用户的主机名和 IP地址是什么？显示了什么错误消息？如果可能，让用户再次运行应用以确认问题，或在测试系统中复制问题。

40.2　标准应用程序

本章及全书使用的标准应用程序包括：

- ifconfig —— 提供网卡的基本配置信息。通常用来检查错误的 IP地址、不正确的子网掩码或错误的广播地址。 Unix系统均提供该应用， Windows NT操作系统的类似工具为ipconfig, Windows 95下为Winipcfg。
- arp —— 以太网地址/IP地址的转换工具。通常用来检测局域网 IP地址配置错误。
- netstat —— 提供多种信息。通常用于显示网络接口、网络socket及网络路由表的统计信息。
- ping —— 检测远程主机是否可达。该工具也可用来显示报文丢失的统计信息及报文传送时间。
- nslookup —— 提供DNS名字服务的信息。

- dig —— 提供名字服务的信息，与 nslookup类似。它可从匿名 FTP服务器venera.isi.edu的/pub/dig.2.0.tar.z获得。
- ripquery —— 提供系统接收或发送 rip更新报文的信息。它是 gated软件包的一部分。但用户不需要运行gated守护进程，当系统运行 RIP时，就可以使用此程序。
- traceroute —— 提供从本地系统到远程节点的路由信息。每一跳的信息均被显示。它可从匿名FTP服务器ftp.ee.lbl.gov的traceroute.tar.z中获得。
- etherfind —— 分析主机与网络交换的报文的信息。 etherfind实际上是TCP/IP协议分析工具，它可以检查报文的内容，包括包头。在分析协议问题时，此应用程序非常有用。在 SunOS操作系统中，该程序称为tcpdump，tcpdump可在匿名FTP站点ftp.ee.lbl.gov上获得。

本章将使用到上述所有应用程序。其 ping在所有工具中使用频率最高。

40.2.1 测试基本连接

ping命令可用于测试远程节点是否可达。这个简单的功能在测试网络连接时非常有用。ping命令使用户决定是否需要对网络作进一步测试 (低层或高层)。如果ping显示报文可到达远程主机并返回，则问题可能出在高层。如果报文不能完成往返过程，可能是低层协议出错。

通常，用户报告网络问题时，只指出他 (她)不能telnet(ftp或sendmail等)到远程主机，然后指出这些服务以前都是可用的。诸如此类的情况，首先要确定是否可以连接远程主机，此时ping是最佳工具。

使用ping测试用户提供的主机是否可达。如果你可以 ping到远程主机，让用户 ping远程主机。如果用户也可以 ping到远程主机，则需要对发生问题的应用做进一步的分析。或许是因为用户试图telnet到只提供匿名 ftp服务的主机；或许当用户登录时，主机已经关闭了。使用户再试一次，并且了解他所做的每一个细节，如果所有步骤都正确而服务依然失败，使用etherfind对应用进行详细分析，必要时联系远程服务的管理员。

如果你可以ping成功，而用户ping失败，则需要检查用户系统的配置是否正确。及用户所使用的网络路径与你使用的路径是否相同。

如果你与用户的ping都失败了，则注意显示的错误信息。 ping所显示的错误消息可以指导进一步的测试。错误消息因具体情况不同而变化。其中基本类型如下：

- 不知名主机 —— 不能将主机名解析为对应的 IP地址。名字服务器可能失效 (本地服务器或远程域名服务器)、名字不正确或用户系统与远程主机间的网络出错。如果知道远程主机的IP地址，直接ping远程主机的IP地址。如果使用 IP地址可达，则可确定问题出在名字服务器。使用nslookup或dig测试本地或远程服务器，同时注意使用的主机名的正确性。
- 网络不可达 —— 本地系统没有到达远程系统的网络路径。如果使用 IP地址也不正确，使用主机名作为ping的参数。这样可减小IP地址出错的可能性。如果使用了路由协议，确定它正确运行并且使用 netstat检查路由表。如果使用 RIP协议，运行rip query检查接收的rip更新报文的内容。如果使用缺省静态路由，重新安装它。如果没有发现问题，则需检查网关是否存在问题。
- 无应答 —— 远程系统不响应。许多网络应用都有此类错误消息。某些 ping命令还显示100%报文丢失消息。 telnet显示Connection timed out消息， sendmail返回cannotconnect错误。所有这些消息的意思相同。本地系统有路由可到达远程主机，但没有接收到远程

主机的响应。

引起这一错误的原因有多种。远程主机可能关闭，或本地主机或远程主机配置不正确，网关或本地主机与远程主机之间的线路出错，远程主机可能有路由错误等；只有进行进一步的测试方可能得到具体的原因。使用 netstat和ifconfig命令仔细检查本地主机的配置。使用traceroute检查本地系统到远程节点的路由。与远程系统的管理员联系，报告问题。

上面提到的所有工具将在本章后面讨论。下面我们讨论 ping命令的参数及其输出形式。

40.2.2 ping命令

ping命令的基本格式如下：

ping *host [packetsize] [count]*

host：为需要测试的远程主机的主机名或 IP地址。使用用户在错误报告中提供的主机名或地址进行测试。

报文长度(packetsize)定义了测试报文的字节尺寸。只有使用数量字段时才需要使用该字段。它的缺省值为56。

数量(count)指明进行测试时发送的报文数量。使用该字段时，值不要太大。否则 ping命令将一直发送报文，直到用户使用 Control-c(^C)中断为止。发送连续测试报文将消耗网络带宽和系统资源。通常5个报文比较合适。

在名为bernard的工作站上测试远程节点 uunet.uu.net是否可达，下面的命令发送5个长度为56的报文：

```
%  ping -s uunet.uu.net  56  5
PING uunet.uu.net: 56 data bytes
64 bytes from uunet.UU.NET (137.39.1.2): icmp_seq=0. Time=14. ms
64 bytes from uunet.UU.NET (137.39.1.2): icmp_seq=0. Time=14. ms
64 bytes from uunet.UU.NET (137.39.1.2): icmp_seq=0. Time=14. ms
64 bytes from uunet.UU.NET (137.39.1.2): icmp_seq=0. Time=14. ms
64 bytes from uunet.UU.NET (137.39.1.2): icmp_seq=0. Time=14. ms

----uunet.UU.NET PING Statistics----
5 packets transmitted, 5 packets received, 0% packet loss
round-trip (ms)  min/avg/max = 12/13/18
```

命令中包含-s开关是因为bernard为Sun工作站，当需要显示每个报文的情况时，需要 -s开关。否则将只显示“ uunet.uu.net is alive ”消息。其他系统上的ping不需要-s开关就可显示详细信息。

测试说明到 uunet.uu.net的广域网连接性能良好，没有报文丢失并且响应速度较快。bernard到uunet.uu.net的往返平均了仅用13毫秒。在广域网中少量的报文丢失及较大的往返时间都是正常的。ping命令的统计信息，可以显示较低层次的网络问题。关键的统计数据有：

- 报文抵达序列，由ICMP序列号(ICMP-seq)显示。
- 每个报文往返所用时间，单位为 ms(毫秒)，在time=后显示。
- 报文丢失百分比，它在ping命令输出的总结行显示。

如果报文丢失较多而响应时间低或报文乱序抵达，说明网络硬件出错。如果在广域网上进行测试，则上述现象属于正常范围。 TCP/IP适用于不可靠网络，某些广域网的报文丢失率

较高，但是若对于局域网，上述现象表明出现网络故障。

在局域网中，往返时间趋近于0，报文丢失很少或没有，并且报文按序抵达。如果出现异常，说明网络硬件故障。在以太网中，可能是线缆终结器故障，或线缆分段故障或中继器故障。首先检查线缆终结器。检查终结器非常简单，或者系统没有安装终结器，它很容易出故障，尤其是终端器放在用户可碰到的工作区中。

检查线缆硬件故障的最佳工具是时域反射表 (time domain reflectometer TDR)，TDR在线缆中发送信号并监听信号的回应。这些回应显示在测试工具前端的显示器上。如果线缆没有终结，信号将跳到显示器上顶端，正常的显示中仅有少量跳跃，使用 TDR可以很容易发现线缆问题。

ping测试的结果无论成功与否都可以帮助用户进一步查找故障。其他诊断工具用来进一步检查故障并发现较明确的起因。

40.2.3 解决网络访问故障

没有响应或不能连接等错误消息表明网络故障出现在协议的底层。如果测试指出此类网络故障则集中测试路由与网络接口。使用 ifconfig、netstat及arp命令测试网络层访问。

1. 使用ifconfig命令

ifconfig检查网络接口的配置。当用户重新配置网络接口或仅有某个用户系统不能登录远程主机而其他主机可以，则使用此命令检查用户的网络接口配置。

当ifconfig的参数只有网络接口名时，它显示分配给接口的当前值。例如检测 bernard le0接口，输出如下：

```
% ifconfig le0
le0: flags=63<UP,BROADCAST,NOTRAILERS,RUNNING>
inet 128.66.12.2 netmask ffff0000  broadcast 128.66.0.0
```

fconfig的输出仅两行，第一行显示接口名称及其特征。检查以下特征：

- UP —— 接口是否启动。如果接口没启动，让系统管理员使用 ifconfig重启接口(ifconfig le0 up)。如果仍未启动，替换接口线路并重试。如果仍然失败，检查接口硬件 (网卡)。
- RUNNING —— 表明接口可使用，如果接口没有运行，接口的驱动程序可能不正确。系统管理员需要检查安装接口的所有步骤是否有错误或遗漏。

第二行显示IP地址、子网掩码(以16进制显示)、广播地址。检查此三项以确保网络接口配置正确。

2. 使用arp命令

arp命令用于分析IP地址到以太网地址的转换问题， arp命令有三个开关可用于解决网络故障。

-a —— 显示表中所有的ARP项。

-d *hostname* —— 从ARP表中删除对应的项。

-s *hostname ether-address* —— 在表中增加新的项。

使用这三个开关，用户可以查看 ARP表的内容，删除有问题的项并添加正确的项。添加正确的项在解决问题时非常有用。

如果怀疑地址表中有不正确的项，可以使用 arp命令。arp表中的错误表现为某些命令对应

错误的IP地址(如ftp或telnet)。仅影响到特定主机的错误也可能表明arp表遭到破坏。arp表的问题通常源于两个系统使用同一个 IP地址。此类问题的间歇性源于表中的地址与最先响应的ARP请求一致，有时正确的主机响应较快。

如果用户怀疑两个系统使用相同的IP地址，可用arp -a命令显示地址解析表。示例如下：

```
% arp -a
bernard (128.66.12.2) at 8:0:20:b:4a:71
annette (128.66.12.1) at 8:0:20:e:aa:40
bernadette (128.66.12.3) at 0:0:93:e0:80:b1
```

如果每个主机的IP对应正确的以太网地址，用户很容易检查 IP地址与以太网地址的正确性。因此，当网络中添加主机时，应记录主机的 IP地址与以太网地址。如果保存了此类记录，很容易发现表中的任何异常。

如果用户没有此类记录，以太网地址的前三个字节也可以帮助解决问题。以太网地址的前三个字节标识设备的生产商。这些前缀标识可以从相应的 RFC文档中的“以太网卡厂商地址部分”找到。

表40-1列出了一些设备生产厂商及其前缀。使用这些信息，我们可以得到示例中前两项为Sun系统(8:0:20)。如果bernadette也为Sun，则0:0:93 Proteon前缀表明Proteon路由器的IP地址配置错误。

表40-1 以太网厂商前缀

前 缀	厂 商	前 缀	厂 商
00:00:0C	Cisco	08:00:0B	Unisys
00:00:0F	NeXT	08:00:10	AT&T
00:00:10	Sytek	08:00:11	Tektronix
00:00:1D	Cabletron	08:00:14	Excelan
00:00:65	Network General	08:00:1A	Data General
00:00:6B	MIPS	08:00:1B	Data General
00:00:77	MIPS	08:00:1E	Apollo
00:00:89	Cayman Systems	08:00:20	Sun
00:00:93	Proteon	08:00:25	CDC
00:00:A2	Wellfleet	08:00:2B	DEC
00:00:A7	NCD	08:00:38	Bull
00:00:A9	Network Systems	08:00:39	Spider Systems
00:00:C0	Western Digital	08:00:46	Sony
00:00:C9	Emulex	08:04:47	Sequent
00:80:2D	Xylogics Annex	08:00:5A	IBM
00:AA:00	Intel	08:00:69	Silicon Graphics
00:DD:00	Ungermann-Bass	08:00:6E	Excelan
00:DD:01	Ungermann-Bass	08:00:86	Imagen/QMS
02:07:01	MICOM/Interlan	08: 00:87	Xyplex terminal servers
02:60:8C	3Com	08:00:89	Kinetics
08:00:02	3Com (Bridge)	08:00:8B	Pyramid
08:00:03	ACC	08:00:90	Retix
08:00:05	Symbolics	AA:00:03	DEC
08:00:08	BBN	AA:00:04	DEC
08:00:09	Hewlett-Packard		

如果检查配置情况及厂商的前缀都不能帮助你标识 ARP源，则使用telnet连接ARP项的IP地址。如果设备支持telnet，则登录界面或许可以帮助你标识错误配置的主机。

3. 使用netstat检查接口

如果前面的测试说明主机到局域网的连接可能不可靠，则 netstat -i命令可以提供更有用的信息，下面为netstat -i的示例：

```
% netstat -i
Name Mtu   Net/Dest Address      Ipkts Ierrs Opkts Oerrs Collis Queue
le0 1500 family.com annette   442697  2   633524  2    50679   0
lo0 1536 loopback localhost  53040  0   53040    0    0        0
```

输出中的lo0行为本地接口可以忽略。只需注意网络接口，其中，每行的后五项对解决网络问题非常有用。

首先让我我们看最后一个字段，它说明没有一个接受查询的报文不能发送出去。如果接口启动并且运行，但是系统不能将报文发送到网络上，则可能是网络接口或线缆故障。如果更换线缆后，问题仍然存在，则需要硬件厂商维修接口硬件设备。

输入错误(Ierrs)及输出错误Oeers应当趋近于0。无论通过接口的报文有多少，若这些字段的值超过100都不正常。输出错误高说明局域网过于饱和或主机与网络的物理连接出错。输入错误高说明网络过于饱和或本地主机负载过重，也可能由于网络物理连接的问题。使用工具如ping或TDR可以确定是否出现网络物理连接故障。网络碰撞率可以判定以太网是否过于饱和。

冲突字段的值较高(Collis)是正常现象，但是如果输出报文与冲突报文的比率过高，则说明网络过于饱和。冲突率应控制在5%以内。如果网络主机的冲突率持续高于5%，则用户需要将网络划分为多个子网以减小流量负载。

冲突率为输出报文与冲突报文的比值。不要使用输入输出报文作比较，只需使用 Opkts和Collis两个字段的值之比获得冲突率。例如，上例中，发出的报文数为 633 424，冲突报文数为50 679，可得冲突率为8%，说明局域网负载过重，检查网络中的其他主机，如果其他主机的冲突率较高，则需考虑划分局域网。

4. 检查路由

网络不可达消息说明路由有问题。如果仅本地主机的路由表出错，则很容易发现并解决。首先使用netstat-nr及grep命令检查路由表中是否有到达目地主机的有效路由：例如检查到网络128.8.0.0的路由：

```
% netstat -nr ¦ grep '128\.8\.0\.0'
128.8.0.0        26.20.0.16     UG      0      37      std0
```

如果主机中没有到达网络128.8.0.0的路由，则主机无响应。例如，某用户报告网络故障，他不能telnet登录到nic.ddn.mil，并且ping该主机返回以下结果：

```
% ping -s nic.ddn.mil 56 2
PING nic.ddn.mil: 56 data bytes
sendto: Network is unreachable
ping: wrote nic.ddn.mil 64 chars, ret=-1
sendto: Network is unreachable
ping: wrote nic.ddn.mil 64 chars, ret=-1
```

```
----nic.ddn.mil PING Statistics----
2 packets transmitted, 0 packets received, 100% packet loss
```

因为出现网络不可达消息，所以需要检查用户的路由表。本例中，检查到 nil.ddn.mil的路由。假设nil.ddn.mil的IP地址为192.112.36.5，这是一个C类地址，注意路由是面向网络的。因此，我们检查到网络192.112.36.0的路由：

```
% netstat -nr ¦ grep '192\.112\.36\.0'
%
```

检查结果说明没有到网络 192.112.36.0的路由。如果找到对应的路由， grep命令将显示出在屏幕上。如果没有到指定地点的路由，需要检查缺省路由。检查缺省路由示例如下：

```
% netstat -nr ¦ grep def
default      128.66.12.1      UG      0      101277        le0
```

如果netstat命令说明有正确的路由或有缺省路由，则说明问题不在路由表。此时，使用traceroute，跟踪到达目的地的整个路径， traceroute命令将在本章后面详细介绍。

如果netstat没有返回路由信息，说明本地路由表出错。根据系统采用静态路由还是动态路由决定解决方法。如果系统采用静态路由，使用 route add命令添加所需路由，大多数系统的静态路由基本上都使用缺省路由。因此，丢失的可能是缺省路由。使系统每次启动时添加所需的路由是彻底解决问题的好方法。

如果使用动态路由，确定路由程序正确运行，例如，使用以下命令判断 gated是否运行：

```
% ps`cat /etc/gated.pid`

PID TT STAT   TIME COMMAND
27711 ?  S    304:59 gated -tep /etc/log/gated.log
```

如果没有运行正确的路由守护进程，重新启动它并指出正确的路径。路径可以使用户发现导致守护进程异常终止的原因。

5. 检查RIP更新

如果路由守护进程正确运行，并且本地系统通过 RIP路由信息协议接收路由更新技术，则使用ripquery检查由RIP提供者接收的路由更新报文。例如，检查从 annette及bernadette发送的路由更新报文， bernard的管理员可使用如下命令：

```
% ripquery -n -r annette bernadette
44 bytes from annette.family.com(128.66.12.1):
0.0.0.0, metric 3
26.0.0.0, metric 0
   264 bytes from bernadette.family.com(128.66.12.3):
128.66.5.0, metric 2
128.66.3.0, metric 2
    .
    .
    .
128.66.12.0, metric 2
128.66.13.0, metric 2
```

显示输出的第一行标识网关，其余各行显示接收的 RIP报文的内容。第一行说明 ripquery接收到来自annette的RIP报文。接下来两行说明annette发出的路由信息。 annette说明它到缺省

路由的距离为3，路由到Milnet(26.0.0.0)的距离为0。然后，ripquery显示bernadette广播的路由信息。这些路由信息路由到其他子网。

本例中使用了ripquery的两个开关：

-n —— 使ripquery显示所有输出。ripquery将把所有的IP地址解析为域名，如果不使用 -n开关。使用-n开关，可以获得更清楚的输出结果，用户不必自己将域名解析为 IP地址。

-r —— 使ripquery直接使用rip Request命令而不是RIP POLL命令查询RIP提供者。RIP DOLL没有得到广泛的支持。在ripquery命令中使用-r开关，更有可能得到需要的结果。

显示输出的路由器必须与参数中给出的一致；否则，检查 RIP提供者的配置。RIP提供者的错误配置可能使它广播不应广播的路由。而应当广播的路由信息又被忽略。用户需要根据自己对网络配置知识的了解检查这些问题。

6、跟踪路由

如果本地路由表及RIP提供者都正确，则问题有可能出在距本地机较远的主机上。远程路由问题可导致“ 无响应 ”及“ 网络不可达 ”错误。但是“ 网络不可达 ”并不说明一定是路由问题。也有可能是本地主机与远程主机间的其他错误导致网络不可达。 traceroute命令可以帮助用户定位这些错误。

traceroute命令跟踪从本地主机到远程节点的 UDP路由报文。它显示从本地主机到远程节点间的每一个网关的名字及IP地址。

traceroute使用两种技术，较小的生存期 (ttl)及无效的端口来跟踪到达目的地的报文。traceroute发送ttl值较小的UDP报文搜索网关。ttl初始值为1，每发送三个UDP报文，ttl值增1。当某个网关接收到报文，它将 ttl值减1。如果ttl的值为0，则报文不向前发送， ICMP超时消息将发送给报文的源地址。 traceroute显示发送ICMP超时报文的网关的信息，每一行显示一个网关。

当目的主机接收到来自 traceroute报文时，它返回“ ICMP端口不可达 ”消息，这是因为traceroute使用无效端口33434来人为产生该错误。如果 traceroute接收到端口不可达消息，则表明它到达目的主机，于是中止 traceroute跟踪。采用这种方法， traceroute列出一系列网关，从第一跳的网关开始，相邻网关相差一跳，直到目的主机。

以下示例说明在SURAnet网中的某台主机上运行traceroute nic.ddn.mil的情况。traceroute每发送三个报文， ttl值增1。如果没有收到报文响应，显示 *。如果收到报文响应，显示响应网关的名字及地址，报文的往返时间单位为 ms(毫秒)。

```
% traceroute nic.ddn.mil
traceroute to nic.ddn.mil (192.112.36.5), 30 hops max, 40 byte packets
1  * pgw (129.6.80.254)  4 ms  3 ms
2  129.6.1.242 (129.6.1.242)  4 ms  4 ms  3 ms
3  129.6.2.252 (129.6.2.252)  5 ms  5 ms  4 ms
4  128.167.122.1 (128.167.122.1)  50 ms  6 ms  6 ms
5  * 192.80.214.247 (192.80.214.247)  96 ms  18 ms
6  129.140.9.10 (129.140.9.10)  18 ms  25 ms  15 ms
7  nsn.sura.net (192.80.214.253)  21 ms  18 ms  23 ms
8  GSI.NSN.NASA.GOV (128.161.252.2)  22 ms  34 ms  27 ms
9  NIC.DDN.MIL (192.112.36.5)  37 ms  29 ms  34 ms
```

路径显示8个网关，每个网关都有响应报文，报文从本地主机到 nic.ddn.mil的往返时间平

均为33ms。

ICMP在不同类型网关上具体实现时的 bug及通过网络的报文路径的随机性可能会导致一些奇怪的显示输出。因此，用户不需仔细察看 traceroute的输出。traceroute输出中重要的部分为：

- 报文是否到达目的主机？
- 如果没有，在哪儿停止？

下面列出了到达nic.ddn.mil的另一条路径，示例中，报文并没有到达目的主机。

```
% traceroute nic.ddn.mil
traceroute to nic.ddn.mil (192.112.36.5), 30 hops max, 40 byte packets
1  * pgw (129.6.80.254)  3 ms  3 ms
2  129.6.1.242 (129.6.1.242)  4 ms  4 ms  4 ms
3  129.6.2.252 (129.6.2.252)  5 ms  5 ms  4 ms
4  128.167.122.1 (128.167.122.1)  6 ms  6 ms  10 ms
5  enss.sura.net (192.80.214.248)  9 ms  6 ms  8 ms
6  t3-1.cnss58.t3.nsf.net (140.222.58.2)  10 ms  15 ms  13 ms
7  t3-0.enss142.t3.nsf.net (140.222.142.1)  13 ms  12 ms  12 ms
8  GSI.NSN.NASA.GOV (128.161.252.2)  22 ms  26 ms  21 ms
9  * * *
10* * *
        .
        .
        .
29* * *
30* * *
```

当traceroute没有获得远程节点的响应，路径跟踪失败，它将连续显示三颗 *直到30。如果碰到此类情况，需要与远程主机的管理员路径中最后到达的网关的管理员联系。向他们说明碰到的问题。本例中，最后到达的网关为 GSI.NSN.NASA.GOV。我们需要与该系统的管理员及nic.ddn.mil的管理员联系，以获得帮助。

7. 检查名字服务

当用户应用返回“ unknown host(不可知主机) ”消息，说明存在名字服务器问题。名字服务器故障通常使用nslookup或dig等工具诊断。dig的功能与nslookup类似。在使用dig之前，让我们首先学习nslookup并了解如何使用nslookup解决名字服务问题。

nslookup三个重要特性可以帮助我们解决远程域名服务器故障。这三个功能特性为：

- 使用NS查询定位远程域的主服务器。
- 使用ANY查询获取远程主机的所有记录。
- 使用nslookup的ls及view命令浏览远程域的所有项。

当解决远程服务器故障时，直接使用 NS查询主服务器，不要查询非主服务器。如果问题有间歇性，查询所有的主域名服务器并比较所得结果。对于同一个查询返回不同的结果会引起间歇式名字服务器问题。

ANY查询返回主机的所有记录。仅知道简单信息就可以解决许多问题。例如：如果查询返回MX记录但没有A记录，则很容易理解用户为什么不能 telnet到远程主机。大多数主机可用mail访问，但不支持其他服务。许多用户不了解这一点，而使用不恰当的方式访问远程主机。

如果管理员不能获得用户提供的主机名的任何信息，很有可能主机名出错。如果是主机

名的问题，纠正它和大海捞针一样困难。然而 nslookup可以解决这一问题。使用 nslookup的ls命令显示远程域的文件，并将它重定向到某个文件中。然后使用 nslookup的view命令浏览该文件，寻找与用户提供的主机名类似的域名。许多问题都是由错误的主机名引起的。

dig是与nslookup类似的工具。dig通常为单行命令，而nslookup为交互行会话应用。但是dig可完成与nslookup相同的功能。具体选择哪一个工具依个人喜好而定。

例如，我们使用dig向根服务器aggie.nca&t.edu查询jhu.edu的NS记录，命令如下：

```
% dig @ aggie.nca&t.edu jhu.edu ns
```

在本例中，@ aggie.nca&t.edu为被查询的服务器。可采用名字或IP地址标识服务器。

如果要解决远程域的问题，需指定对应域的主服务器。本例中查询顶级域 (jhu.edu)的服务器名，因此需要询问根服务器。

40.3 解决网络接口层问题

虽然在这一层很简单，但底层的栈也可能发生故障。在这一层的问题可分为以下四类：

- 物理连接问题。与所有网络一样， TCP/IP必须互连才能正常工作。所有的网络都要求系统连入网络，因此，首先需要检查物理连接。
- DHCP客户端没有分配IP地址。因为可能返回给用户大量的消息所以比较明显。用户可以通过两种方式获得IP地址。第一种是静态分配 IP地址，另一种是通过 DHCP(动态主机配置协议)动态获取IP地址。在Unix中，用户需要使用 Ifconfig确信获得了IP地址，在Windows 95下可用WinIPCFG，Windows NT下可用IPCONFIG命令完成相同的功能。
- ARP问题。如果地址解析协议工作不正常，用户就不能将 IP地址解析为MAC地址。ARP应用，可以帮助用户检查地址解析是否正确。
- 网络中IP地址冲突。用户可能碰到的另一个问题是两个系统使用相同的 IP地址。如果发生此类情况，系统在把IP地址解析为MAC地址时发生错误。

如果采用静态地址解析，可以避免此类问题。但是当系统更换网络适配器时，必须要修改地址解析表。用户可以使用 ARP命令解决此类问题。

40.4 解决网络层问题

网络层负责报文的路由。因此，在网络层，我们需要仔细检查 IP地址、子网掩码及缺省网关是否正确。除了配置外，问题还可能出在路由表或本地主机远程节点间的路由器上。

40.4.1 TCP/IP配置参数

TCP/IP配置主要包括三个参数： IP地址、子网掩码及缺省网关。缺省网关通常是用户主机所在网段的路由器地址。在 Windows系统中，可以在“网络”对话框的“协议”标签页中配置，在Unix系统中，需要使用 ifconfig命令在命令行方式下进行配置。虽然 IP地址可通过DHCP服务器动态配置，但本节主要讨论静态 IP地址的配置。

TCP/IP的三个参数必须配置正确，否则将不能进入 TCP/IP网络。不正确的配置可能由于拓扑图错误造成。如果敲入错误的 IP地址、子网掩码或默认网关，用户可能不能正确连入网络或根本连接不上。

如果TCP/IP配置参数的错误是由于拓扑或错误的数字，则不正确的配置将影响通信。配置错误的类型不同，可能导致不同的故障。后续章节将详细讨论问题的类型及解决方法。

40.4.2 IP地址配置问题

错误的IP地址配置可能引起许多问题。如果 IP地址处在正确的子网且没有发生冲突，但使用了错误的主机ID，客户通信将不能很好地进行。如果与主机名对应的正确的 IP地址存放在静态文本或数据库中，正如LMHOSTS文件或DNS数据库文件中，将发生通信故障。因此，不正确的IP地址将引发若干问题。

TCP/IP参数的配置问题可以引起多种不同的症状。下面几小节将分别讨论它们对 IP通信的影响。

1. IP地址

IP地址拥有两个或三个部分惟一标识地址分配到的主机。 IP地址分为两部分，一部分标识网络地址；另一部分标识主机地址。同时，如果用户划分子网， IP地址的第三部分还可标识主机所在的子网地址。

图40-1显示不正确的网络地址的情况。在示例中，客户方的 TCP/IP地址设置不正确。其他地址设置为143.168.3.9，正确的地址应为133.168.3.9。因此，不正确的IP地址所对应的网络ID为143.168.x.x，而正确的IP所对应的网络ID为133.168.x.x。

使用不正确的地址(143.168.3.9)，客户端不能与其他任何TCP/IP主机进行通信。因为网络地址不正确，客户方发送的任何报文都将路由到错误的地址。

如果错误配置的主机(143.168.3.9)向本地客户(133.168.3.20)发送消息，发送主机TCP/IP配置表明它是一台远程主机，因为其网络地址与接收消息的主机的网络地址不同。报文不能到达本地客户端，因为接收消息的主机认为消息来自远程地址。

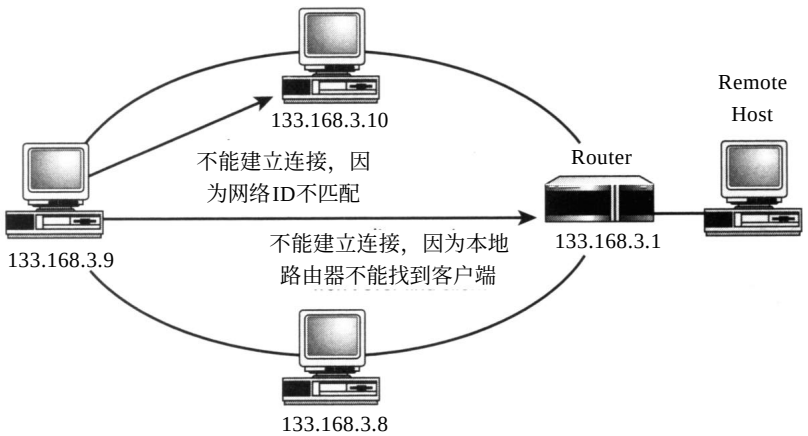

图40-1 错误IP地址示例

如果本地主机(133.168.3.6)向不正确的主机(143.168.3.9)发送消息，消息将不能到达目的主机。消息将路由到远程域 (如果使用IP地址发送)或留在子网内(使用名字发送，其对应的 IP为133.168.3.9)。如果消息被路由，错误主机将不能接收到消息，因为它位于本地子网内。如果消息没有被路由，消息仍然不能到达错误主机，因为目的主机的 IP地址为133.168.3.9与错误配置的主机IP地址143.168.3.9不匹配。

图40-2给出了另外一个错误配置IP地址的例子。在本例中，使用 A类地址(33.x.x.x)，子网掩码(255.255.0.0)表明后第二个字节用于创建子网。本示例中，即使客户端与其他主机有相同

的网络地址，但由于子网地址不同仍然会产生错误。

本例中，错误的原因是指定了不正确的子网 ID。主机33.5.8.4在子网5而其他主机的子网ID为4。因此，如果主机33.5.8.4试图与同一个子网的其他主机通信，消息将被路由，因为子网ID与源主机子网ID不匹配。如果主机33.5.8.4试图向远程主机发送消息，消息将被路由。但是，返回给主机的消息不能到达主机，因为路由器不能处理到达子网 5的消息，而只处理到达子网4的报文。

如果本地客户向主机33.5.8.4发送消息，消息将不能到达主机。如果本地客户使用 IP地址，消息将被路由，从而不可能到达对应主机，因为它在子网之内。如果本地客户使用了正确的IP地址，报文仍然不能被正确接收，因为 IP地址不匹配。

IP地址的最后一个组成部分为主机地址，如果主机地址配置错误也可能会导致通信错误。不正确的主机地址不总是引起错误。如图 40-3所示，本地客户使用了错误的IP地址，但仅只有主机地址不正确，网络地址及子网地址与其他客户相同。

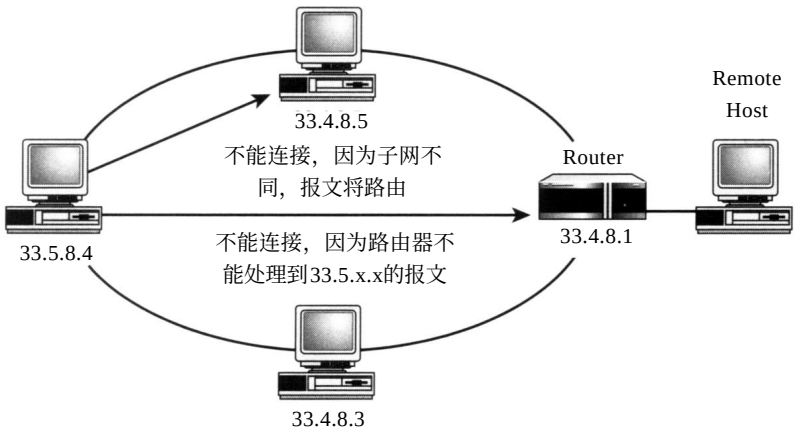

图40-2 不正确的子网地址

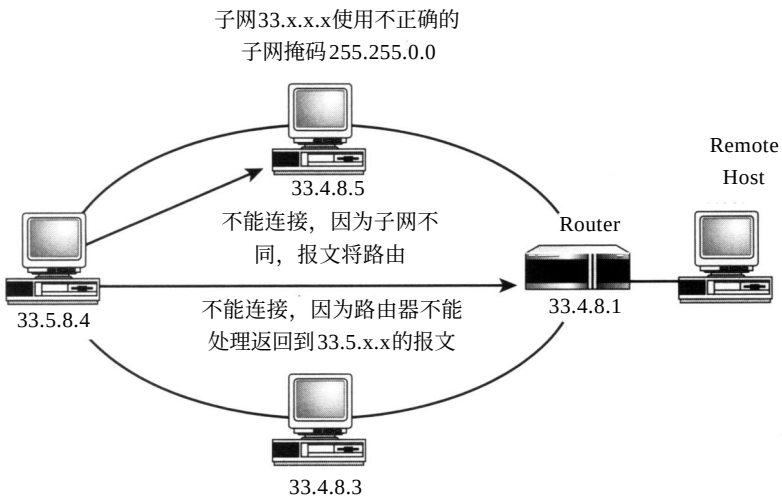

图40-3 不正确的主机ID

在本例中，如果客户向 IP地址不正确的主机发送消息，消息可以到达主机。然而如果某

用户使用正确的IP地址进行通信，将不能建立连接。实际上，他将与另一台主机建立连接，并认为此主机为他试图建立连接的主机。

如果两台主机的IP地址相同，首先启动的主机将正确运行，而后启动的主机将显示地址冲突，并且不加载TCP/IP协议栈。此时要启动的主机不能使用TCP/IP建立通信。

此外，当正确的主机地址被记录在静态文本中时，将引发其他问题。如LMHOSTS文件或DNS数据库文件中。此时，使用名字不能与正确的主机建立连接，因为服务器总是返回正确的IP地址。

通常，主机地址不正确引发的问题具有间歇性。然而，当主机是 WINS客户端时(使用Windows操作系统)，主机名将随IP地址一起注册。其他WINS客户端试图与配置不正确的主机连接时，总可以获得主机名与IP地址的准确映射。

2. 子网掩码

子网掩码表明IP地址中哪一部分表明网络地址。哪一部分表明主机地址。同时，子网掩码可以将主机地址分成多个子网。如果子网掩码设置不正确，客户主机将不能通信，或会引发通信故障。

图40-4显示使用B类网络地址138.13.x.x的TCP/IP网络的子网、IP地址的第三个字节用于划分子网，图中所有客户主机的子网ID为4，表明其子网地址为138.13.4。

不幸的是，其中一个主机的子网掩码设置为255.255.0.0。当该主机试图与同一子网其他主机建立通信时，因为子网掩码表明它们在同一子网，所以可以建立连接。然而，当该主机试图与其他子网的主机建立连接时，如138.13.3.x，连接失败。

在这种情况下，子网掩码仍认为目的主机与本地主机在同一子网下，消息将不被路由。因为目的主机在别的子网，消息将永远到达不了目的主机。子网掩码只决定出去的报文是否需要路由。因此，子网掩码配置错误的主机仍可以接收消息，然而当它发送返回消息时，消息将只在本子网内，而不会通过路由器到达其他子网。

所以，客户端只能建立单向通信。与其他网络的通信仍可正常进行，因为这些报文都将路由。

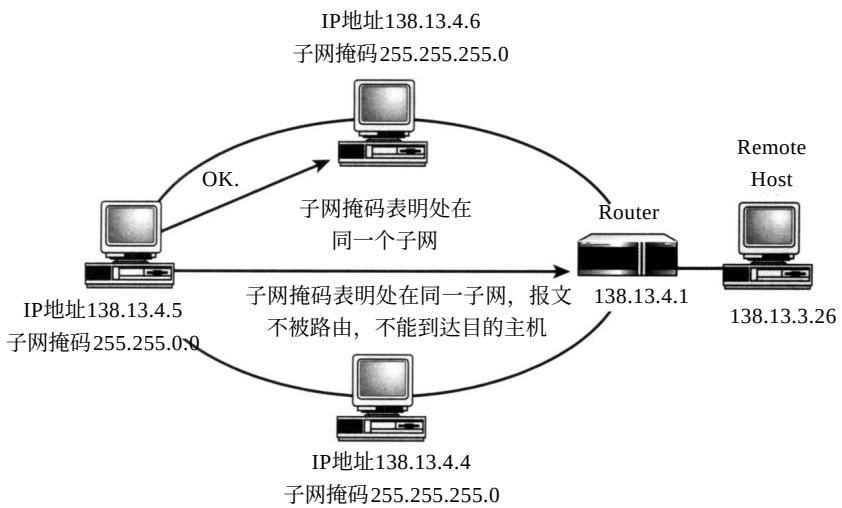

图40-4 即使IP地址正确，错误的子网ID仍会引发问题

图40-5显示子网掩码位太多的错误。一般情况下，子网掩码为 255.255.255.0，然而网络设计者试图使用子网掩码255.255.240.0，第三个字节的前四位用于子网，后四位用于主机地址。

如果正确配置的客户端向本地主机发送消息，并且 IP地址第三个字节相同，消息将不被路由，可正确建立连接。然而如果本地客户端的地址第三个字节的后四位不同，消息将被路由，从而无法到达目的主机。如果错误配置的主机试图与同子网内的其他主机建立联系，消息将路由，因为第三个字节不同。

由于子网掩码错误导致的问题有间歇性，有时连接工作正常，有时发生错误。当目的主机的IP地址引起该路由的报文未路由或不该路由的包却路由了，将会引发通信故障。

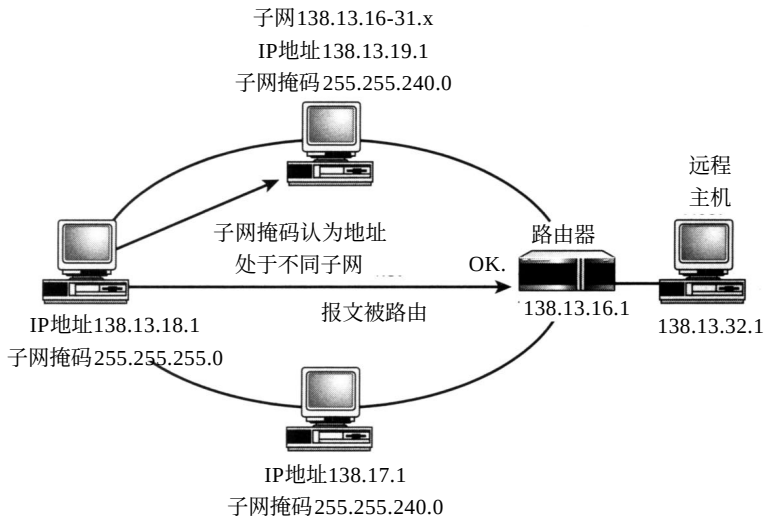

图40-5 子网掩码位太多而导致的通信故障

3. 默认网关

默认网关的地址为路由器的地址 —— 处在子网交界处的网关。如果客户主机的默认网关配置错误，它只可与本地主机通信，而不能连接远程节点。当默认网关配置错误时，主机可以接收来自子网以外的报文，因为默认网关只影响报文的发送。然而，当主机试图响应远程节点的报文时，由于网关配置错误，将不能正确发送响应报文，导致网络连接错。

40.5 解决TCP和UDP问题

传输层很少出现问题。如果用户可以使用 ping命令，说明传输层工作正确。在这一层惟一可能发现的问题与第9章中讨论的TCP窗口尺寸有关。

该问题将导致通信速度缓慢 (与70年代到80年代使用的110波特类似)。如果用户确认无其他问题，在Windows NT中检查TCP窗口尺寸只需打开注册表查找 Windows Size项即可。如果发现该项，记住其值，并将该项删除，系统将自动重设窗口尺寸。

1. Socket问题

如果用户正确配置了 IP地址及协议，并且到远程节点的路由正确，仍需要在正确的端口运行需要的服务。如果想要提供服务或连接网络上提供的服务，用户必须了解服务使用的端口号。Internet编号管理局(IANA)分配了通用端口号(socket)。然而在某些情况下，服务需要使用不同端口。下一节描述的文件列出服务所对应的端口，如果用户确认端口正确，可以使

用netstat命令检查端口接收的数据。

2. Service文件

Windows NT中system32\drivers\etc目录下的service文件中列出了大部分服务使用的端口号，以供系统初始时使用。下面是该服务文件的一小部分：

```
#   Copyright     1993-1995  Microsoft Corp.
#

#   This file contains port numbers for well-known services as
# defined by RFC 1060 (Assigned Numbers).

#


#   Format:

#


#   <service name>   <port number>/<protocol> [aliases...]   [#<comment>]
#
echo          7/tcp

echo          7/udp

discard          9/tcp        sink null
discard          9/udp        sink null
systat        11/tcp

systat        11/udp        users

daytime        13/tcp

daytime        13/udp

netstat        15/tcp

qotd        17/tcp        quote

qotd        17/udp        quote

chargen        19/tcp        ttytst source

chargen        19/udp        ttytst source
ftp-data        20/tcp

ftp        21/tcp

telnet        23/tcp
smtp        25/tcp        mail
```

```
time        37/tcp        timeserver

time        37/udp        timeserver

rlp         39/udp        resource            #  resource location
name        42/tcp        nameserver
name        42/udp                    nameserver
whois       43/tcp        nicname             #  usually to sri-nic
domain      53/tcp        nameserver          #  name-domain server
domain      53/udp        nameserver
nameserver  53/tcp        domain              #  name-domain server
nameserver  53/udp        domain
mtp         57/tcp                    #  deprecated          bootp
67/udp                    # boot program server        tftp        69/udp
```

如果某个服务有问题，用户应首先检查此文件中服务是否使用正确的端口。如果服务未列出，用户必须手工加入，使系统知道使用哪个端口提供此项服务 (service文件是普通的文本文件，可使用Edit或写字板查看和修改)。

40.6 解决应用层问题

在应用层中，用户可能会碰到两类问题： NetBIOS问题和名字解析问题。其中最常见的是名字解析问题，它影响Socket应用及NetBIOS应用。在40.6.1节，我们将讨论名字解析问题。

名字解析问题

若用户正确配置TCP/IP并且协议安装正确且可正常工作，如果连接中仍有问题则很可能是名字解析问题。当管理员使用TCP/IP地址测试连接时，可能只测试了底层连接。

当用户想要连接网络资源 —— 如映射某服务器的网络驱动器 —— 他们通常使用域名而不是IP地址。实际上，用户也不知道服务器的 IP地址。但是，当使用域名建立连接时，必须首先将域名解析为IP地址。

当检测完IP地址后，下一步是检查名字解析的IP地址是否正确。如果名字不能解析为IP地址或解析不正确，用户就不能使用名字建立连接，除非他们使用 IP地址。

用于建立连接的主机名有两类：分配给 NetBIOS主机的NetBIOS名字，如Windows NT或Windows 95系统；分配给非 NetBIOS主机的名字，如 Unix服务器。总的来说，当使用Microsoft网络连接网络共享、文件打印或应用服务器时，使用 NETBIOS名。当执行TCP/IP应用时，如使用FTP或Web浏览器时，使用主机名。

40.7 小结

由TCP/IP配置引起的问题远远超过 TCP/IP协议实现时引起的问题。用户碰到的大多数问题都可以使用本章讲解的工具解决。但是，在某些情况下，用户需要分析两个系统协议交互的情况；甚至有时需逐位分析报文中的数据。