Traceroute – The Internet's Diagnostic Tool

NTT Communications Global IP Network White Paper

By Charles Sellers, CISSP

May, 2006

May, 2006

Contents

Section 1

Executive Summary

Traceroute was originally developed with the intention of providing a quick and dirty debugging tool, which could be used to help determine which device or segment of the network may be causing network problems.

This whitepaper discusses what the traceroute program is and how to use traceroute, complete with several examples demonstrating error messages encountered. Also covered are different scenarios traceroute can be utilized in order to troubleshoot, isolate, and diagnose network problems.

Section 2

What is Traceroute?

Traceroute was originally developed with the intention of providing a quick and dirty debugging tool, which could be used to to help determine which device or segment of the network may be causing network problems. Running Traceroute can yield several pieces of information when executed. Traceroute is the program that shows you the exact route taken by data packets over the network between the source systems to the destination system, listing all the intermediate routers a connection must pass through to get to its destination for the set of probe packets sent by traceroute. This route may or may not vary, depending on what is going on between the two end systems. The route packets are sent can, and usually do, vary with time depending on many factors. It can help you determine why your connections to a given server might be poor, and can often help you figure out where exactly the problem is. It also shows you how systems are connected to each other, letting you see how your ISP connects to the Internet as well as how the target system is connected. Traceroute can also be easily used to find how exactly a particular network is organized and to determine the potential entry points.

If you execute the `traceroute` command on a source device, it sends IP packets toward the destination with Time To Live (TTL) values that increment up to the maximum specified hop count. This is 30 by default on most systems. Typically, each router in the path towards the destination decrements the TTL field by one unit while it forwards these packets. When a host in the middle of the path finds a packet with TTL = 1, it responds with an Internet Control Message Protocol (ICMP) "time exceeded" message to the source. This message lets the source know that the packet traverses that particular router as a hop

Running Traceroute

Traceroute can be run from almost any host or network system. Popular software which supports traceroute includes most Unix systems, Mac OS X, and Microsoft Windows 95, 98, 2000, and XP. If your host system does not have this capability, the source code can be downloaded from any number of places and compiled to run on your host system if there is library support for network functionality.

On Mac OS X and Unix systems, and most router platforms the command is executed with either the domain name:

```
#traceroute www.servername.com
```

or by IP address of the host:

```
#traceroute 192.168.1.1
```

On Microsoft OS products the command is executed similar to above like this:

```
C:\WINDOWS>tracert www.servername.com
```

or by IP address of the host:

```
C:\WINDOWS>tracert 192.168.1.1
```

VisualRoute (http://www.visualroute.com/), a graphical traceroute program, can be downloaded and is available for Windows, Sparc Solaris, and Linux. VisualRoute helps analyze the traceroute, and provides a world map showing where packets are going

Another option is to use traceroute portals. A couple of examples are described inn the next section.

Traceroute Portals

Traceroute.org (http://www.traceroute.org/) is a large collection of traceroute, looking glass, route servers and bgp links from which a Network Administrator can utilize to track down network issues.

This VisualRoute Server (http://www.visualware.com/) provides a graphical traceroute and ping test from this server to any other network device you choose, useful for pinpointing network connectivity problems and identifying IP address locations.

Traceroute Output

The traceroute probe will continue until it is either successful or fails, which indicates that an ICMP error message was received). The trace will stop at this point. Possible ICMP error messages are shown in Table 1.

Table 1 – Typical ICMP Error Messages

Character	Description
*	The probe timed out.
?	Unknown packet type.
!A	Administratively unreachable.
!F	Fragmentation needed. This indicates that the router is misconfigured.
!H	Host unreachable. The router has no route to the target system.
!N	Network unreachable.
!P	Protocol unreachable.
!Q	Source quench (destination too busy).
!S	Source route failed. You tried to use source routing, but the router is configured to block source-routed packets.
!T	Timeout.
!U	Unreachable
!X	Communication administratively prohibited. The network administrator has blocked traceroute at this router.

In example 1 – Traceroute to www.sec.gov observe the administratively unreachable error message on hop 11.

Note: Cisco Router ICMP Unreachables Rate Limitation

ICMP unreachables are limited to one packet per 500 ms (as a protection for Denial of Service (DoS) attacks) in a Cisco Router. From Cisco IOS Software Release 12.1 and later, this rate value is configurable.

Section 3

Using Traceroute

Traceroute can be utilized in several ways:

1. To trace the geographical location of a particular system.
2. To Get Information on Network Topography
3. Firewall Detection Purposes
4. Remote OS Detection using Traceroute
5. Latency Detection

A brief description of each use of traceroute along with an example is discussed in the following paragraphs.

Geographical Location

In Example 1 – Traceroute to www.sec.gov the approximate geographical location of the server can be discerned by observing hops 9-11. Near the destination at hop 9, the location of that hop is in Ashburn, VA. The latency numbers for hops 9-11 indicate < 4 ms from hop 9 to hop 11 indicating that the server is in or near NTT America's Ashburn, VA facility.

Example 1 – Traceroute to www.sec.gov

```
e0-0.pe-lab1#traceroute www.sec.gov
Type escape sequence to abort.
Tracing the route to www.sec.gov (162.138.185.33)
  1 129.250.33.177 4 msec 0 msec 0 msec
  2 t3-0-1-1.a00.lsanca03.us.ra.verio.net (198.173.172.169) 28 msec 32 msec 28 msec
  3 vl-4.r00.lsanca03.us.bb.gin.ntt.net (129.250.29.125) 208 msec 232 msec 232 msec
  4 xe-1-0-0.r20.lsanca03.us.bb.gin.ntt.net (129.250.5.32) 28 msec 28 msec 28 msec
  5 p64-0-3-0.r20.mlpsca01.us.bb.gin.ntt.net (129.250.4.114) 36 msec 36 msec 36 msec
  6 p64-0-0-0.r20.asbnva01.us.bb.gin.ntt.net (129.250.2.11) 96 msec 100 msec 100 msec
  7 xe-0-3-0.r21.asbnva01.us.bb.gin.ntt.net (129.250.2.17) 100 msec 100 msec 112 msec
  8 xe-1-1.r05.asbnva01.us.bb.gin.ntt.net (129.250.2.87) 100 msec 100 msec 100 msec
  9 ge-3-3.a00.asbnva01.us.ce.verio.net (168.143.105.98) 96 msec 96 msec 96 msec
 10 border1.pc1-bbnet1.wdc002.pnap.net (216.52.127.18) 100 msec 96 msec 100 msec
 11 sec-2.border1.wdc002.pnap.net (66.150.126.206) !A  *  !A
```

Network Topology Information

As shown in Example 1 – Traceroute to www.sec.gov, the probe packets traverse the NTT network until it is handed off to the pnap network in Ashburn, VA. In Example 2- Traceroute to www.google.com, the probe packets traverse NTT America's network until handed off to Level 3's network in Nos Angeles, CA (hop 5 to 6). From there the probe packets traverse the Level 3 network to Atlanta, GA, where a server farm exists for

Google, as shown with multiple hops within the data center to a Google server. Using traceroute in this manner, a network administrator can find out how a particular network is structured, the address class to which it belongs and in general and to obtain related information on topography of a remote network.

Example 2- Traceroute to www.google.com

```
e0-0.pe-lab1#traceroute www.google.com
Translating "www.google.com"...domain server (129.250.35.250) [OK]
Type escape sequence to abort.
Tracing the route to www.l.google.com (64.233.161.99)
  1 129.250.33.177 0 msec 4 msec 0 msec
  2 t3-0-1-1.a00.lsanca03.us.ra.verio.net (198.173.172.169) 28 msec 28 msec 28 msec
  3 vl-4.r01.lsanca03.us.bb.gin.ntt.net (129.250.29.141) 24 msec 28 msec 28 msec
  4 xe-0-1-0.r21.lsanca03.us.bb.gin.ntt.net (129.250.5.46) 36 msec 28 msec 28 msec
  5 p16-0.level3.lsanca03.us.bb.gin.ntt.net (129.250.9.34) 40 msec 28 msec 28 msec
  6 ae-2-56.bbr2.LosAngeles1.Level3.net (4.68.102.161) 28 msec
    ae-2-54.bbr2.LosAngeles1.Level3.net (4.68.102.97) 28 msec
    ae-2-56.bbr2.LosAngeles1.Level3.net (4.68.102.161) 28 msec
  7 as-1-0.bbr1.Atlanta1.Level3.net (209.247.9.101) 92 msec 96 msec 92 msec
  8 ae-21-54.car1.Atlanta1.Level3.net (4.68.103.98) 96 msec 92 msec
    ae-21-56.car1.Atlanta1.Level3.net (4.68.103.162) 92 msec
  9 4.78.208.114 80 msec 80 msec 80 msec
 10 66.249.95.148 96 msec 96 msec 92 msec
 11 72.14.238.96 100 msec
    72.14.238.234 96 msec 96 msec
 12 72.14.236.202 96 msec
    72.14.236.200 96 msec
    72.14.236.202 96 msec
 13 216.239.49.214 100 msec
    216.239.48.190 104 msec 100 msec
 14 www.l.google.com (64.233.161.99) 100 msec 100 msec 96 msec
```

Firewall Detection Purposes

The presence of a firewall installed near the destination host can be detected as well. When executing traceroute to www.nasdaq.com the '*' (asterisk) sign appears at hop 14 of the traceroute output. This means that the traceroute attempt has timed out. If this command is repeated to the target system several times at different times of the day and still receive the same output with the '*', then it probably means that a firewall has been installed on the target system. This is also easily verified when surfing to http://www.nasdaq.com The asterisk sign being displayed in the output of the traceroute command indicates that a firewall is installed between the source system and the destination system and is filtering out traceroute attempts. As a result, the target system does not reply to the traceroute a request which causes the probe packet to time due to the firewall. Another quick test, besides http to the destination host, is to ping the

destination host. If all three (or more) different protocol attempts to contact the destination host, fail, the destination host is most likely down or inoperative.

Example 3 – Traceroute to www.nasdaq.com

```
e0-0.pe-lab1#traceroute www.nasdaq.com
Translating "www.nasdaq.com"...domain server (129.250.35.250) [OK]
Type escape sequence to abort.
Tracing the route to www.nasdaq.com (208.249.116.71)
  1 129.250.33.177 0 msec 0 msec 0 msec
  2 t3-0-1-1.a00.lsanca03.us.ra.verio.net (198.173.172.169) 28 msec 28 msec 28 msec
  3 vl-4.r00.lsanca03.us.bb.gin.ntt.net (129.250.29.125) 28 msec 28 msec 28 msec
  4 xe-1-0-0.r20.lsanca03.us.bb.gin.ntt.net (129.250.5.32) 28 msec 28 msec 28 msec
  5 p16-0-0-0.r02.lsanca03.us.bb.gin.ntt.net (129.250.3.157) 28 msec 28 msec 28 msec
  6 POS1-2.BR3.LAX9.ALTER.NET (204.255.173.17) 32 msec 28 msec 28 msec
  7 0.so-0-3-0.XL2.LAX9.ALTER.NET (152.63.115.6) 28 msec 28 msec 52 msec
  8 0.so-0-0-0.XL2.NYC8.ALTER.NET (152.63.0.170) 104 msec 104 msec 124 msec
  9 0.so-3-0-0.XR2.NYC8.ALTER.NET (152.63.19.34) 104 msec 104 msec 104 msec
 10 182.at-6-1-0.WR2.NYC8.ALTER.NET (152.63.16.201) 104 msec 108 msec 104 msec
 11 pos6-0.ur2.ewr2.web.wcom.net (63.111.126.46) 104 msec 104 msec 104 msec
 12 63.87.224.55 108 msec 108 msec 108 msec
 13 63.87.242.37 104 msec 104 msec 104 msec
 14  *  *  *
 15  *  *  *
 16  *  *  *
 17  *  *  *
 18  *  *  *
 19  *  *  *
 20  *  *  *
 21  *  *  *
 22  *  *  *
 23  *  *  *
 24  *  *  *
 25  *  *  *
 26  *  *  *
 27  *  *  *
 28  *  *  *
 29  *  *  *
 30  *  *  *
```

Remote OS Detection

Traceroute used along with a packet sniffer can be used to detect the name and version of the operating system running on the destination system by observing the value of the TTL field of the packet. The following table is courtesy Lance Spitzner and the Honeypot Project.

Table 2 – Default Values of the TTL Field by Operating System

Operating System	Platform	TTL
Windows 9x/NT	Intel	32
Windows 9x/NT	Intel	128
Windows 2000	Intel	128
Digital UNIX 4.0	Alpha	60
Unisys x	Mainframe	64
Linux 2.2.x	Intel	64
FTX(UNIX) 3.3	STRATUS	64
SCO R5	Compaq	64
Netware 4.11	Intel	128
AIX 4.3.x	IBM/RS6000	60
AIX 4.2.x	IBM/RS6000	60
Cisco 11.2	7507	60
Cisco 12.0	2514	255
IRIX 6.x	SGI	60
FreeBSD 3.x	Intel	64
OpenBSD 2.x	Intel	64
Solaris 8	Intel/Sparc	64
SunOS 2.x	Intel/Sparc	255

Latency Detection

Traceroute sends out three probe packets by default. The three numbers given on each line of traceroute output show the round trip times (latency) in milliseconds. Smaller numbers generally mean better connections. As the latency of a connection increases, interactive response suffers. Each time listed is the total amount of time that one packet took to get from the originating machine to that specific hop and back to the originating machine. It is not the amount of time from the previous hop to the following hop. Download speed can also suffer as a result of high latency (due to TCP windowing), or as a result of whatever is actually causing that high latency

Table 3 – Connection Latency

Connection	Latency (ms)
Dial-up	120-180
ISDN	35-50
DSL	5-30
Satellite	250-1,000
Wi-Fi	2-25 (range dependent)

Causes of Latency

Latency can be caused by several different problems, a few of which are listed here.

Link Over Utilization

The most common cause of latency is over utilization of a link. When the amount of traffic passing over an IP link increases to a high percentage of the total bandwidth of that link, the latency across that link begins to increase. The closer the utilization gets to 100%, the worse the latency becomes until the link is saturate. At that point very little, if any, traffic will succeed in reaching its destination. Typically over utilization would be found near the edge of the network, either at the CPE or the link to the CPE. Link over utilization is not normally an issue with links between core routers. Links between core routers are usually monitored by service providers to ensure the link is not over utilized and for capacity planning purposes.

Link Speed

Different links will have different latency, which is normal, i.e. a trans-oceanic link will add more latency than an intra-POP or intra-City link. Different links will have different latencies as shown in Table x – Connection Latency

Router Over Utilization

When a router's processor becomes over utilized, (CPU utilization approaches 100%), it's ability to route traffic may quickly become impaired. This may be due to the router attempting to handle more traffic than it was designed for, or a shortage of router resources such as memory. Router over utilization may also be caused by malicious attacks on the router, such as denial-of-service attacks. Service Provider core routers forward traffic via specially designed Application Specific Integrated Circuits (ASICs) at wire speed and are very difficult to cause the CPU to approach 100% utilization.

Firewall Issues

Firewalls can be affected by many of the same problems as routers. These include over utilization due to lack of resources and malicious attacks.

It is important to realize latency problems may still reside elsewhere despite what the traceroute output indicates. Traceroute only gives you the forward path to the destination host, not the return path. The return path may be completely different (asymmetric routing), and there may be additional latency on that path. Asymmetric routing is not unusual on the Internet and is not in itself a problem. Asymmetric routing is typically the standard used between Service Providers. To see the return path, you will need to obtain a traceroute performed on the destination host tracing the path back to you. This can be accomplished by utilizing one of the traceroute portals mentioned earlier.

Asymmetric Routing

In Commercial Service Provider networks routing between Autonomous Systems (ASs) is achieved by passing traffic from one AS to another AS as soon as possible thereby using the other Service Provider's wide-area transit to move the packet to it's destination. This practice is what is known as hot-potato routing.

Hot-potato routing is the normal practice of most settlement-free peering agreements. It is expected that your peers at a peering point will route packets destined for your network within your AS in this manner.

Cold-potato routing, opposite of hot-potato routing, is another practice where the originating AS routes the packet internally to the AS until it is as near to it's destination as possible, then exits the AS at a peering point Cold-potato routing is more expensive to perform, but keeps the packet traffic in the originating AS as long as possible allowing the service provider of well-provisioned networks to offer a higher Quality of Service (QoS) to their customers. Cold potato routing is prone to misconfiguration as well as poor coordination between two service provider networks. In such scenarios packets often do not traverse the optimal route to the destination and are usually routed further distances. In the case of service providers who have trans-oceanic links, this can add several hundred milliseconds.

Appendix A

Microsoft Traceroute

The MS Windows tracert command uses ICMP echo request datagrams instead of UDP datagrams as probes. ICMP echo requests are launched with incrementing TTL, and the same operation as described in Cisco IOS and Linux occurs. The significance of using ICMP echo request datagrams is that the final hop does not rely on the response of an ICMP "unreachable" message from the destination host. It relies instead on an ICMP echo reply message.

The command syntax is:

```
tracert [-d] [-h maximum_hops] [-j computer-list] [-w timeout] target_name
```

This table explains the command parameters:

Parameter	Description
-d	Specifies not to resolve addresses to computer names.
-h maximum_hops	Specifies the maximum number of hops to search for a target.
-j computer-list	Specifies a loose source route along computer-list.
-w timeout	Waits the number of milliseconds specified by the timeout for each reply.
target_name	Name of the target computer.

The command syntax for IPv6 traceroute is:

```
tracert6 [-d] [-h maximum_hops] [-w timeout] [-s srcaddr] target_name
```

This table explains the command parameters:

Parameter	Description
-d	Do not resolve addresses to hostnames.
-h max_hops	Maximum number of hops to search for target.
-w timeout	Wait timeout milliseconds for each reply.
-s srcaddr	Source address to use.
-r	Use routing header to test reverse route also.

Appendix B

References

ftp://ftp.rfc-editor.org/in-notes/rfc792.txt

http://www.cisco.com/en/US/tech/tk364/technologies_tech_note09186a00801ae32a.shtml

http://www.cisco.com/en/US/products/sw/iosswrel/ps1831/products_tech_note09186a00800a6057.shtml

http://dast.nlanr.net/NPMT/

http://en.wikipedia.org/wiki/Dial-up

http://en.wikipedia.org/wiki/Integrated_Services_Digital_Network