Van Jacobson Traceroute

Original Van Jacobson/Unix/LBL Traceroute

The well known traceroute program has been written by Van Jacobson in December 1988. In the header comment of the source, Van explains that he just implemented an idea he got from Steve Deering (and that could have come from Guy Almes or Matt Mathis). Traceroute sends "probe" packets with TTL (Time to Live) values incrementing from one, and uses ICMP "Time Exceeded" messages to detect router "hops" on the way to the specified destination. It also records "response" times for each hop, and displays losses and other types of failures in a compact way.

This traceroute is also known as "Unix traceroute" (although some Unix variants have departed from the code base, see e.g. Solaris traceroute) or "LBL traceroute" after Van Jacobson's employer at the time he wrote this.

Basic function

UDP packets are sent as probes to a high ephemeral port (usually in the range 33434–33525) with the Time-To-Live (TTL) field in the IP header increasing by one for each probe sent until the end host is reached. The originating host listens for ICMP Time Exceeded responses from each of the routers/hosts en-route. It knows that the packet's destination has been reached when it receives an ICMP Port Unreachable message; we expect a port unreachable message as no service should be listening for connections in this port range. If there is no response to the probe within a certain time period (typically 5 seconds), then a * is displayed.

Output

The output of the traceroute program shows each host that the packet passes through on its way to its destination and the RTT to each gateway en-route. Occasionally, the maximum number of hops (specified by the TTL field, which defaults to 64 hops in *NIX implementations) is exceeded before the port unreachable is received. When this happens an ! will be printed beside the RTT in the output.

Other error messages that may appear after the RTT in the output of a traceroute are:

!H : Host unreachable
!N : Network unreachable
!P : Protocol unreachable
!S : Source-route failed (that is to say, the router was not able to honour the source-route option set in an IP packet)

!X : Administratively prohibited. The gateway prohibits these packets, but sends an ICMP message back to the source of the traceroute to inform them of this.

!V : Host precedence violation
!C : Precedence cut-off in effect

! [num] : displays the ICMP unreachable code, as defined in a variety of RFC, shown at http://www.iana.org/assignments/icmp-parameters

UDP Probes

The original traceroute sends UDP packets as probes to a high ephemeral port (usually in the range 33434–33525) with the Time-To-Live (TTL) field in the IP header increasing by one for each probe sent until the end host is reached. The originating host listens for ICMP Time Exceeded responses from each of the routers/hosts en-route. It knows that the packet's destination has been reached when it receives an ICMP Port Unreachable message; we expect a port unreachable message as no service should be listening for connections in this port range. If there is no response to the probe within a certain time period (typically 5 seconds), then a * is displayed.

Why UDP?

It would seem natural to use ICMP ECHO requests as probe packets, but Van Jacobson chose UDP packets to presumably unused ports instead. It is believed that this is because at that time, some gateways (as routers were called then) refused to send ICMP (TTL exceeded) messages in response to ICMP messages, as specified in the introduction of RFC 792, "Internet Control Message Protocol". Therefore the UDP variant was more robust.

These days, all gateways (routers) send ICMP TTL Exceeded messages about ICMP ECHO request packets (as specified in RFC1122, "Requirements for Internet Hosts – Communication Layers", so more recent traceroute versions (such as Windows tracert=) do indeed use ICMP probes, and newer Unix traceroute versions allow ICMP probes to be selected with the =-I option.

Why increment the UDP port number?

Traceroute varies (increments) the UDP destination port number for each probe sent out, in order to reliably match ICMP TTL Exceeded messages to individual probes. Because the UDP ports occur right after the IP header, they can be relied on to be included in the "original packet" portion of the ICMP TTL Exceeded messages, even though the ICMP standards only mandate that the first eight octets following the IP header of the original packet be included in ICMP messages (it is allowed to send more though).

When ICMP ECHO requests are used, the probes can be disambiguated by using the sequence number field, which also happens to be located just before that 8-octet boundary.
Filters

Note that either or both of ICMP and UDP may be blocked by firewalls, so this must be taken into account when troubleshooting. As an alternative, one can often use traceroute variants such as `lft=or=tcptraceroute` to work around these filters, provided that the destination host has at least one TCP port open towards the source.

References

- [original announcement](http://end2end-interest.org) from the archives of the end2end-interest mailing list

-- Main.SimonLeinen - 26 Feb 2006