Laura Chappell

Fragmentation

Good, Bad, and Downright Ugly

Editor's Note: For more information about network analysis techniques and uses, you can attend Session 164 "Introduction to Network Analysis" by Laura Chappell at BrainShare 2000 in Salt Lake City. You may also want to visit http://www.netanalysis.org to download PINGFRAG.CAP—a trace file that contains fragmented Internet Control Message Protocol (ICMP) packets.

Unlike IPX, IP supports fragmentation over smaller Maximum Transmission Unit (MTU) links. On an IPX network, devices must negotiate the lowest common packet size on startup and pray that this packet size is sufficient. If many packets are lost, you may begin to think that your company's network resembles a tiny garden hose connecting fire hydrants together. On an IP network, a larger packet (such as a Token Ring 4,096-byte packet) is automatically fragmented into smaller packets to cross a link (such as an Ethernet link) that supports a smaller MTU.

Although you may think fragmentation sounds like a nice feature, beware: IP fragmentation is not all that it's cracked up to be (I couldn't resist).

HOW FRAGMENTATION WORKS

The fragmentation and reassembly process is performed at the IP layer of the TCP/IP stack. If a router cannot forward an oversized packet, that router fragments the packet into smaller, more acceptable-sized packets. For example, if a device on a Token Ring network sends a 4,096-byte Token Ring packet to an Ethernet network, the router between the two networks splits the packet into three smaller Ethernet packets and sends each of these packets separately.

The router must perform the following tasks to properly fragment the packet:

- The router places the value of the original packet's IP header Identification field in each fragment.
- The router decrements the original packet's Time to Live (TTL) value by one and places the new TTL value in each fragment.
- The router calculates the relative location of the fragmented data and includes that value in the Fragment Offset field of each fragment.
- The router sends each fragment as a separate packet with a separate Media Access Control (MAC) header and checksum calculation.

```
DLC:  ----- DLC Header -----
DLC:
DLC:  Frame 2 arrived at  13:57:45.3559; frame size is 1514 (05EA hex)
DLC:  Destination = Station 00104B30C44A
DLC:  Source      = Station 00A0CC30C8DB
DLC:  Ethertype   = 0800 (IP)
DLC:
IP:  ----- IP Header -----
IP:
IP:  Version = 4, header length = 20 bytes
IP:  Type of service = 00
IP:        000. ....   = routine
IP:        ...0 ....   = normal delay
IP:        .... 0...   = normal throughput
IP:        .... .0..   = normal reliability
IP:  Total length    = 1500 bytes
IP:  Identification  = 19485
IP:  Flags           = 2X
IP:        .0.. ....   = may fragment
IP:        ..1. ....   = more fragments
IP:  Fragment offset = 1480 bytes
IP:  Time to live    = 32 seconds/hops
```

Figure 1. *To obtain this trace, I forced a device to fragment a large ICMP echo packet by using the following command: ping -l 4096 10.0.0.1.*

FRAGMENTED PACKETS

Figure 1 shows a fragmented packet. The Fragment Offset value of 1480 indicates that this fragment is not the first of the set. This fragment is not the last in the set either, as denoted by the More Fragments bit setting.

When the fragments arrive at the destination device, this device uses the fragment offset value, which is contained in the IP header, to put the fragments back in order. Elegant, eh? Unfortunately, fragmentation also has an ugly side.

THE UGLY SIDE OF FRAGMENTATION

Fragmentation has some ugly characteristics that make it undesirable traffic on a network. First, the fragmentation procedure takes processing time at the router or device that is fragmenting the packet.

Second, all fragments must arrive before the expiration of the first-received fragment's TTL. If one of the fragments does not arrive in time, the destination device sends an ICMP message type 11 (Time Exceeded) with code 22 (Fragmentation Reassembly Time Exceeded). In this case, the sending device resends the original packet, which must be fragmented again. On a network that has limited available bandwidth, the fragment retransmission process causes more traffic on the wire.

Laura Chappell writes technical training books for podbooks.com and is a senior protocol analyst at NetAnalysis Institute.