



www.internet2.edu

What is NDT? This is the Network Diagnostic Tool, developed by Rich Carlson of the Department of Energy's Argonne National Laboratory. It can be accessed via: <http://miranda.ctd.anl.gov:7123>. For a recent presentation on this tool, see: <http://www.internet2.edu/presentations/spring03/20030410-Web100-Carlson.htm>.

What is Ethereal? Ethereal is a free network protocol analyzer for Unix and Windows that lets users examine data from 1) a live network or 2) a capture file on disk. Users can interactively browse the capture data, viewing summary and detail information for each packet. Ethereal features include a rich display filter language and the ability to view the reconstructed stream of a TCP session. For more information, see: <http://www.ethereal.com/>.

What is a broadcast storm? This is when a large number of broadcast packets happen at the same time. For example, when a power failure causes a significant number of machines to reboot at the same time, each machine broadcasts a question at the same time (such as "who has this IP address?"), causing a "storm" of packets. This can bring the network down again... and again... and again.

An alternate scenario is one involving redundant layer 2 connections between switches running a spanning tree. In normal operation, the switches select one of the connections to be the active one and disable the alternate path. When the primary path fails, the alternate path can be activated and connectivity is maintained. If the spanning tree protocol is somehow disabled, a broadcast storm results: any broadcast packet (continued)

Hey! Where Did My Performance Go?

A Case Study for E2Epi

Shawn, an assistant research scientist at the University of Michigan, was seeing problems with routine data transfers. As part of his research with the Large Hadron Collider, Shawn routinely sends large streams of data to locations across the nation and around the world. On June 11, 2003, when sending a stream, Shawn experienced 20% packet loss to locations outside of his departmental LAN. Why?

Recently, Shawn has been instrumental in developing MGRID (Michigan Grid Research and Infrastructure Development) for the University; the project is attempting to develop a scalable grid infrastructure such that the tools could be replicated on a national and international level. As part of the development effort, Shawn is documenting problems encountered, along with the solutions discovered. Unlike many users confronted with an end-to-end problem, Shawn was not only expecting trouble, but prepared to deal with it, expeditiously.

As the chair of Internet2's End-to-End Performance Initiative (E2Epi) Technical Advisory Group (TAG), Shawn is familiar with many of the diagnostic tools that network performance experts use to identify and solve performance problems. He installed NDT (Network Diagnostic Tool), developed by Rich Carlson of Argonne National Labs, on a Web100 Kernel (on a Linux 2.4.20 box) and used the tool to debug poor performance on his local network.

Using a tuned host, connecting via 100 Mbps Ethernet, a user would normally see 95 Mbps or so worth of throughput. Shawn discovered that the maximum throughput was 20-60 Mbps over many tests from a tuned client to the NDT server along a FastEthernet path. The University of Michigan has a robust networking infrastructure; normally, he would

have no problem sending a 95 Mbps TCP stream across his own campus between properly configured machines.

NDT helped Shawn localize two problems by identifying the bandwidth limitation, which wasn't present earlier, as well as indicating a significant amount of packet loss on an under-utilized network. Further research with **Ethereal**, a network packet capturing utility, showed bursts of broadcast packets during normal network operation and coincident with packet "loss" events. Testing within a local subnet didn't exhibit the problem. The problem seemed to involve the connection between the local subnet and the rest of the departmental LAN. When the department network administrator was contacted about this specific network connection, Shawn discovered that the department had established a broadcast packet rate limit of 10 packets/second to protect against a known bug that had caused a **broadcast storm**.

The reason for wanting rate limiting was in response to an earlier problem:

- The departmental network design had Ethernet switches.
- The switches had a protocol among them to form a spanning tree – a single path for broadcast packets.
- A firmware upgrade on a network switch had silently turned off this feature on that switch; now broadcast packets would come in one link and broadcast on all others.
- Unfortunately, the departmental network design also had redundant links; this meant that broadcast packets kept multiplying when they passed the redundant links, until the network was completely filled with these (now useless) broadcast packets.

received on one of the redundant connections is rebroadcast to all interfaces EXCEPT the interface on which it was received. The problem is that one of these interfaces leads back to the switch that originally sent the packet – which now loops the packet back to first switch. Very quickly, these switches amplify the original broadcast packet and spend all their time resending the same broadcast, creating a broadcast storm. Thus, to protect the network against broadcast storms, some links are rate limited.

What is the E2Eperf Interest Group?

The E2Eperf Interest Group is an email list for those interested in end-to-end issues. It was started to help form Internet2's End-to-End Performance Initiative (E2Epi) and continues today as a place for discussions of end-to-end performance problems. To subscribe to the mailing list:

- Send a message to the address: listproc@internet2.edu
- Do not include anything in the subject line
- In the body of the message include the line: `subscribe e2eperf-interest FirstName LastName [i.e., subscribe e2eperf-interest John Smith]`

Once you are subscribed, you may send a message to the list by using this address: E2Eperf-INTEREST@Internet2.edu

Archives of the E2Eperf-INTEREST list can be found at: <http://archives.internet2.edu/>.

What is E2Epi? E2Epi is Internet2's End-to-End Performance Initiative. For more information about us, see our website at: <http://e2epi.internet2.edu/>. There you'll find links to recent projects (like **E2E piPEs**) and presentations, as well as other information related to end-to-end performance issues, measurement tools, upcoming events, and related activities.

To try to protect the network against something like this in the future, some of the links were rate limited. A traffic limiting device was installed in the middle of the departmental LAN; it was designed to turn on the rate limiting when it recognized a certain level of broadcast packets, but, apparently, it limited *all traffic* on the link (not just broadcasts).

NOTE: As a result of this problem, the University of Michigan network administration team worked with the manufacturer of the network switch to ensure both that the firmware upgrade would not silently turn off a spanning tree and that the existing release notes warned users about this existing behavior.

The 10 packet/second limit, when it was originally configured, was not a problem on the subnet, which typically had around 2-3 broadcast packets/second. However some newly installed (and apparently misconfigured) software was causing bursts of broadcast packets, which intermittently caused the subnet to exceed 10 packets/second, thus causing intermittent connectivity problems due to exceeding the broadcast limit.

Resolution: When Shawn identified the problem and brought it to the attention of his network staff, they increased the rate limit by a factor of three, which decreased the probability that "normal" broadcast traffic (or at least traffic with broadcast rates significantly below a real broadcast storm) would trip the limit. In addition, Shawn notified one of the subnet user's about their misbehaving software and they correctly reconfigured it.

Recommendations

- Have a good understanding of your network topology.
- When a good connection suddenly goes bad, talk to network staff at the departmental and campus levels. Tell them about the problem encountered; ask if they have installed any upgrades that could have silently modified the paths?
- Learn about network tools available to help define and isolate the problem, such as NDT (<http://miranda.ctd.anl.gov:7123>), which is designed for novice (to expert) users, and Ethereal (www.ethereal.com) or tcpdump, which are primarily intended for expert users. Another option is to use Iperf (<http://dast.nlanr.net/Projects/Iperf/>) with a server at the edge of your campus as a testpoint.
- Ask if your department or campus has initiated rate limiting; all the bugs have not yet been worked out of traffic limiting devices and they could be causing a problem.
- More generally, even simple switch and router configurations can have unforeseen consequences, especially with regards to performance. Talk to the people who run your network; lack of communication is often the largest part of the problem.