Subject: Path MTU Discovery doesn't work properly
Posted by uDgCHciH on Thu, 03 Mar 2011 19:44:15 GMT
View Forum Message <> Reply to Message

Hi,

all my containers have set a MTU size of 1500 Bytes. If someone connects to it through a VPN he or she won't get all TCP packets.

A TCP dump shows that the router next to the containers sends an ICMP Type 3 Code 4 (fragmentation needed) message with a recommended MTU size of 1419 Bytes back to the container.
The container then retransmits the criticized TCP packet with a MTU of 1419 Bytes. The retransmitted packet reaches its destination and gets acknowledged.
But all other transmissions are still send up to a MTU of 1500 Bytes.

The mentioned router sends not more than one ICMP messages per second back in case of too big TCP packets (to save it self). So the container doesn't know that the messages in between were dropped and waits till the retransmission time out occurs. Because every second one messages gets through, the receiving system realizes that it gets packets out of order and sends back a double ack messages.
The container then starts to resend any packet beginning of the sequence within the double ack message. But then again the first messages gets dropped by the router with a proper ICMP message and only the first message gets resend again with the correct MTU size.
It seems to me that after receiving the 'fragmentation needed' message a rule with the correct MTU size is added to the routing cache but gets removed / flushed as soon the criticized message is resend.

Currently the only way to work around this issue is to reduce the default MTU of the Container. But that can not be a final solution.

I haven't found any similar issue in the bug tracker of Open VZ or Red Hat.

First I thought I've found someone with the same problem.
See Linux Kernel Mailing List for:
[135/156] ipv4: Dont drop redirected route cache entry unless PTMU actually expired

But that occurs only in case of an ICMP-Redirect messages which weren't existent in my case.

Does someone ever has seen the same behavior or has any idea how to solve it?

Kernel: ovzkernel-2.6.18-194.26.1.el5.028stab079.1 (RHEL based)
OS Hardware Node: CentOS release 5.5
OS Containers: RHEL 4.8 or CentOS 5.5

Many thanks in advance
Steve

Subject: Re: Path MTU Discovery doesn't work properly
Posted by kir on Fri, 04 Mar 2011 03:06:31 GMT

http://bugzilla.openvz.org/show_bug.cgi?id=1798