

Re: Diagnose co-location networking problem

Source: <http://unix.derkeiler.com/Mailing-Lists/FreeBSD/net/2006-12/msg00259.html>

- *From:* Matthew Hudson <fbsd@xxxxxxxxxxxxx>
 - *Date:* Thu, 28 Dec 2006 12:31:00 -0800
-

On Wed, Dec 27, 2006 at 10:08:25PM -0800, Stephan Wehner wrote:

The server FreeBSD kernel doesn't support tcpdump. I should recompile it then, but not now.

Ok, that explains the private 192.168 IP address I saw in your earlier dumps, it was from the client (a detail mentioned but that I overlooked).

So I ran the netstat tests, seeing no other suggestion. Below is the output before and after "failed" accesses. If I understand, there seems no indication of lost packets.

Actually there's significant indication of lost packets and clues that point to the location of the problem. I'll explain.

At least the problem is rather reproducible: run 'lynx -dump <http://stbgo.org> > /dev/null' in a loop, 15 times and a failure occurs. I also thought maybe the ssh session might be interfering, rather than showing a live connection; but without it the same occurs.

Generally two TCP connections on different sockets will never interfere with each other, except in extreme examples of congestion or pathologically configured address-translating gateways.

```
# Both on client and server:
$ netstat -i > /tmp/before
$ netstat -s | grep -i ret >> /tmp/before
... run test .... recognize failure ...
$ netstat -i > /tmp/after
$ netstat -s | grep -i ret >> /tmp/after
```

Client first.

Re: Diagnose co-location networking problem

```
$ cat /tmp/before
Kernel Interface table
Iface MTU Met RX-OK RX-ERR RX-DRP RX-OVR TX-OK TX-ERR TX-DRP
TX-OVR
Flg
eth0 1500 012471498 0 0 0 8604916 36 0 1
BMRU
eth0: 1500 0 – no statistics available – BMRU
eth0: 1500 0 – no statistics available – BMRU
lo 16436 0 429696 0 0 0 429696 0 0 0
LRU
66656 segments retransmited
TCPLostRetransmit: 0
TCPFastRetrans: 1233
TCPForwardRetrans: 18
TCPSlowStartRetrans: 476
$ cat /tmp/after
Kernel Interface table
Iface MTU Met RX-OK RX-ERR RX-DRP RX-OVR TX-OK TX-ERR TX-DRP
TX-OVR
Flg
eth0 1500 012471903 0 0 0 8605107 36 0 1
BMRU
eth0: 1500 0 – no statistics available – BMRU
eth0: 1500 0 – no statistics available – BMRU
lo 16436 0 429786 0 0 0 429786 0 0 0
LRU
66665 segments retransmited
TCPLostRetransmit: 0
TCPFastRetrans: 1233
TCPForwardRetrans: 18
TCPSlowStartRetrans: 476
```

So we're looking at the client here and there are few things of note:

1. No significant interface errors are being recorded so it's not a layer-2 (ethernet) issue.
2. The retransmit count went up by 9 while the overall transmit count went up by 191 packets, suggesting an approximate transient packetloss rate of 4.7% (9/191, fuzzy math) during the test which is significantly greater than the system-wide average of 0.8% (66665/8605107). Thus this possibly suggests that the client saw an abnormal packetloss rate during the test. It may be the case that all of the successful connections experienced no packet loss and only the failed connect generated the retransmits. I'm not sure if initial SYN retransmits get counted in this column or not but I believe this still may be significant. (The assumptions made in these calculations are so grossly oversimplified that the evidence derived from them is weak at best).
3. The loopback saw 90 packets of activity. I don't know how

Re: Diagnose co-location networking problem

long this test ran but that could be considered a little chatty.
As a longshot, I'd run a tcpdump on loopback and run the test again, simply to make sure that no traffic is unintentionally getting diverted over the loopback interface (unlikely but I've actually seen bugs/bad firewall configs do this).

```
Now server
$ cat /tmp/before_server
Name Mtu Network Address Ipkts Ierrs Opkts Oerrs
Coll
bge0 1500 <Link#1> 00:0b:cd:4e:40:00 156342 3 146739 0
599 bge0 1500 65.110.18.136 df1 144448 - 145988
-- bge1* 1500 <Link#2> 00:0b:cd:4e:41:11 0 0 0
0 0 pflog 33208 <Link#3> 0 0 0
0 0 lo0 16384 <Link#4> 34545 0 34545
0 0 lo0 16384 your-net localhost.stephan 34454 - 34454
-- pfsyn 2020 <Link#5> 0 0 0
0 0 565 data packets (372083 bytes) retransmitted
38 data packets unnecessarily retransmitted
540 retransmit timeouts
156 retransmitted
0 invalid return addresses
0 no return routes
$ cat /tmp/after_server
Name Mtu Network Address Ipkts Ierrs Opkts Oerrs
Coll
bge0 1500 <Link#1> 00:0b:cd:4e:40:00 156579 3 146957 0
599 bge0 1500 65.110.18.136 df1 144671 - 146206
-- bge1* 1500 <Link#2> 00:0b:cd:4e:41:11 0 0 0
0 0 pflog 33208 <Link#3> 0 0 0
0 0 lo0 16384 <Link#4> 34685 0 34685
0 0 lo0 16384 your-net localhost.stephan 34594 - 34594
-- pfsyn 2020 <Link#5> 0 0 0
0 0 565 data packets (372083 bytes) retransmitted
38 data packets unnecessarily retransmitted
540 retransmit timeouts
156 retransmitted
0 invalid return addresses
0 no return routes
```

And here are the server stats which seem to show very little but in fact are quite informative.

1. No significant interface errors, again ruling out layer-2.
2. pflog and pfsyn devices are registered in the kernel, suggesting PF firewalling has been compiled in. It doesn't seem that pflog is being used at all but this does beg the question, are you using any packetfiltering on the server? If so, I'd suggest disabling the packetfilter entirely and retesting to see if the issue is reproducible.

Re: Diagnose co-location networking problem

3. The retransmit count has gone up by zero, suggesting the server never sent a packet that it later had to retransmit.

This strongly suggests to me that the nature of the connection problems is that the server never sees the client's SYN packets. This is fairly strong evidence pointing to an intelligent filtering device / proxy in the middle of the connection. (or even a firewall configuration on the server itself).

Offhand, here's another test you can run: try and determine if this connection failure behavior is specific to HTTP or general to all TCP services. So far you've mentioned no troubles with SSH, I think you should test that further. Set up a similar test to your HTTP test but with SSH... I'd probably set up public-key authentication on a account on the server so that I could log in without a password and then run simple remote commands over ssh on the server:

```
ssh myserver echo boink
```

over and over again to see if any of those connections fail with a frequency similar to the HTTP test. If you're unable to reproduce the same failure behavior with a test like this then that suggests that the problem is only specific to HTTP which is practically a smoking gun that this is a firewall/loadbalancer/middlebox issue. You need some smarts in the middle to selectively interfere with one type of TCP traffic and not another.. there's no way that a routing problem could be so selective. It's also still possible that this could be a kernel issue since you've clearly tweaked your configuration (compiled out bpf, compiled in PF).. if you compile a GENERIC kernel and run it, can the test be reproduced? This is a more costly test but one to consider if all else fails.

Also, there's another possibility. I noticed in your earlier messages that the IP address of the server is 65.110.18.138 which in-addr.arpa maps to VPS-18-138.virtualprivateservers.ca. Looking at virtualprivateservers.ca's website it seems that they specialize in virtualized servers, begging the question: is your server running in a virtual server (xen, whatnot)? If so then that opens up a slew of other possible issues and is important information to know.

Oh, also, going back to the 198.168 address seen in the client dumps, it's clear that you're going through a NAT firewall or VPN or something on the way to your server. Thus are you able to reproduce this problem from a different external network?

Actually, I just realized that you've provided enough information for me to run this test myself which I've now done. I ran the following test;

```
i=0; while true; do ((i++)); echo $i; curl http://stbgo.org > /dev/null; done
```

I was able to make over 64 consecutive connections without a single failure before I stopped the test (didn't want to spam your site). How sure

Re: Diagnose co-location networking problem

Re: Diagnose co-location networking problem

are you that this isn't a client-side problem?

cheers.

--

Matthew Hudson

freebsd-net@xxxxxxxxxxx mailing list
<http://lists.freebsd.org/mailman/listinfo/freebsd-net>
To unsubscribe, send any mail to "freebsd-net-unsubscribe@xxxxxxxxxxx"

freebsd-net@xxxxxxxxxxx mailing list
<http://lists.freebsd.org/mailman/listinfo/freebsd-net>
To unsubscribe, send any mail to "freebsd-net-unsubscribe@xxxxxxxxxxx"