Subject: Bridging veth dev creates massive delay
Posted by MeMu on Thu, 13 Sep 2007 08:05:55 GMT
View Forum Message <> Reply to Message

Hello!
I'm fighting with this problem for some time now without any positiv results.
I tried to ping from an external host to a VZ node using wlan. But the ping rtt
are dramatically high and this occures only when pinging the VZ node.
Pinging a dummy device which is also connected to the bridge works sometimes
perfect but the next day its also delayed.

Short Hardware Summary:
OpenVZ Host:
Dual Xeon 2.80GHz with 2GB RAM,
DLink (Atheros) WLAN
OpenVZ 2.6.18-028stab039 and 2.6.22-ovz003.1 tested

Additional External Host without OpenVZ:
Asus R2H, 900MHz Celeron M, 768MB RAM,
Linux 2.6.22.1 with modified debiankernel config
Fritz WLAN USB Stick
(Also tested with differend external hardware nodes to eliminate
problems on this side; same results)

ASUS(wlan0, 10.0.0.23) <--> VZHost(Bridge vzbr0 [ath0, dummy0, veth111.0])

My VZ node uses the IP 10.0.0.111
Here are some details from the host system:

```
# brctl show
bridge name     bridge id               STP enabled     interfaces
vzbr0           8000.001346e9df09       no              dummy0
                                                        veth111.0
                                                        ath0

# route -n
Kernel IP Routentabelle
Ziel            Router          Genmask         Flags Metric Ref    Use Iface
10.0.0.0        0.0.0.0         255.255.255.0   U     0      0        0 vzbr0

# ifconfig
ath0      Protokoll:Ethernet  Hardware Adresse 00:13:46:E9:DF:09
          inet6 Adresse: fe80::213:46ff:fee9:df09/64 Gültigkeitsbereich:Verbindung
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:3728 errors:0 dropped:0 overruns:0 frame:0
          TX packets:5396 errors:0 dropped:0 overruns:0 carrier:0
          Kollisionen:0 Sendewarteschlangenlänge:0
```

RX bytes:364504 (355.9 KiB) TX bytes:634091 (619.2 KiB)

dummy0   Protokoll:Ethernet Hardware Adresse 76:29:A2:99:3B:BE
inet Adresse:10.0.0.77 Bcast:10.0.0.255 Maske:255.255.255.0
inet6 Adresse: fe80::7429:a2ff:fe99:3bbe/64 Gültigkeitsbereich:Verbindung
UP BROADCAST RUNNING NOARP MTU:1500 Metric:1
RX packets:0 errors:0 dropped:0 overruns:0 frame:0
TX packets:533 errors:0 dropped:0 overruns:0 carrier:0
Kollisionen:0 Sendewarteschlangenlänge:0
RX bytes:0 (0.0 b) TX bytes:89933 (87.8 KiB)

veth111.0 Protokoll:Ethernet Hardware Adresse 00:12:34:56:78:11
inet6 Adresse: fe80::212:34ff:fe56:7811/64 Gültigkeitsbereich:Verbindung
UP BROADCAST RUNNING MULTICAST MTU:1500 Metric:1
RX packets:285 errors:0 dropped:0 overruns:0 frame:0
TX packets:663 errors:0 dropped:0 overruns:0 carrier:0
Kollisionen:0 Sendewarteschlangenlänge:0
RX bytes:29448 (28.7 KiB) TX bytes:85244 (83.2 KiB)

vzbr0   Protokoll:Ethernet Hardware Adresse 00:13:46:E9:DF:09
inet Adresse:10.0.0.110 Bcast:10.0.0.255 Maske:255.255.255.0
inet6 Adresse: fe80::212:34ff:fe56:7811/64 Gültigkeitsbereich:Verbindung
UP BROADCAST RUNNING MULTICAST MTU:1500 Metric:1
RX packets:18306 errors:0 dropped:0 overruns:0 frame:0
TX packets:3906 errors:0 dropped:0 overruns:0 carrier:0
Kollisionen:0 Sendewarteschlangenlänge:0
RX bytes:2384712 (2.2 MiB) TX bytes:356868 (348.5 KiB)

wifi0   Protokoll:UNSPEC Hardware Adresse
00-13-46-E9-DF-09-30-3A-00-00-00-00-00-00-00-00
UP BROADCAST RUNNING MULTICAST MTU:1500 Metric:1
RX packets:19226 errors:0 dropped:7326 overruns:0 frame:120
TX packets:5448 errors:529 dropped:0 overruns:0 carrier:0
Kollisionen:0 Sendewarteschlangenlänge:199
RX bytes:1405104 (1.3 MiB) TX bytes:800827 (782.0 KiB)
Interrupt:16

Everything looks good and works until you increase the "traffic"/packets per second.
Here are my strange ping results:
ping vom external device via WLAN to dummy device in bridge
ping 10.0.0.77 -i 0.01 -c 100 -q
PING 10.0.0.77 (10.0.0.77) 56(84) bytes of data.

--- 10.0.0.77 ping statistics ---
100 packets transmitted, 84 received, 16% packet loss, time 1239ms
rtt min/avg/max/mdev = 9.484/924.517/1752.211/508.240 ms, pipe 64

ping vom external device via WLAN to VZ node connected to the bridge
ping 10.0.0.111 -i 0.01 -c 100 -q
PING 10.0.0.111 (10.0.0.111) 56(84) bytes of data.

--- 10.0.0.111 ping statistics ---
100 packets transmitted, 82 received, 18% packet loss, time 1264ms
rtt min/avg/max/mdev = 5.986/883.845/1724.167/502.285 ms, pipe 63

ping vom VZ node to dummy device
vn11:~# ping 10.0.0.77 -i 0.01 -c 100 -q
PING 10.0.0.77 (10.0.0.77) 56(84) bytes of data.

--- 10.0.0.77 ping statistics ---
100 packets transmitted, 100 received, 0% packet loss, time 890ms
rtt min/avg/max/mdev = 0.016/0.027/1.129/0.110 ms

Sometimes the ping packets to the dummy device are perfect but never the connection between ext.hardware node to the vz node.
Its also strange, that the RTT increses after every new ping request packet... Maybe some queue is full - some overflows or whatever?
ping vom external device via WLAN to VZ node connected to the bridge
ping 10.0.0.111 -i 0.01
PING 10.0.0.111 (10.0.0.111) 56(84) bytes of data.
64 bytes from 10.0.0.111: icmp_seq=1 ttl=64 time=5.43 ms
64 bytes from 10.0.0.111: icmp_seq=2 ttl=64 time=62.4 ms
64 bytes from 10.0.0.111: icmp_seq=3 ttl=64 time=70.0 ms
64 bytes from 10.0.0.111: icmp_seq=4 ttl=64 time=110 ms
64 bytes from 10.0.0.111: icmp_seq=5 ttl=64 time=147 ms
64 bytes from 10.0.0.111: icmp_seq=6 ttl=64 time=151 ms
64 bytes from 10.0.0.111: icmp_seq=7 ttl=64 time=216 ms
64 bytes from 10.0.0.111: icmp_seq=8 ttl=64 time=223 ms
64 bytes from 10.0.0.111: icmp_seq=9 ttl=64 time=229 ms
64 bytes from 10.0.0.111: icmp_seq=10 ttl=64 time=237 ms
64 bytes from 10.0.0.111: icmp_seq=11 ttl=64 time=274 ms
64 bytes from 10.0.0.111: icmp_seq=12 ttl=64 time=278 ms
64 bytes from 10.0.0.111: icmp_seq=13 ttl=64 time=280 ms
64 bytes from 10.0.0.111: icmp_seq=14 ttl=64 time=320 ms
64 bytes from 10.0.0.111: icmp_seq=15 ttl=64 time=332 ms
64 bytes from 10.0.0.111: icmp_seq=16 ttl=64 time=336 ms
64 bytes from 10.0.0.111: icmp_seq=17 ttl=64 time=399 ms
64 bytes from 10.0.0.111: icmp_seq=18 ttl=64 time=396 ms
64 bytes from 10.0.0.111: icmp_seq=19 ttl=64 time=391 ms
64 bytes from 10.0.0.111: icmp_seq=20 ttl=64 time=414 ms

64 bytes from 10.0.0.111: icmp_seq=21 ttl=64 time=460 ms
64 bytes from 10.0.0.111: icmp_seq=22 ttl=64 time=459 ms
64 bytes from 10.0.0.111: icmp_seq=23 ttl=64 time=536 ms
64 bytes from 10.0.0.111: icmp_seq=24 ttl=64 time=533 ms
64 bytes from 10.0.0.111: icmp_seq=25 ttl=64 time=527 ms
64 bytes from 10.0.0.111: icmp_seq=26 ttl=64 time=523 ms
64 bytes from 10.0.0.111: icmp_seq=27 ttl=64 time=512 ms
64 bytes from 10.0.0.111: icmp_seq=28 ttl=64 time=551 ms
[.....]

Are there maybe some settings in the vz node? or bridge? Maybe it is a strange race condition?
Hope someone has some hints/advices for me. Thanks in advance!
bye MeMu

ps: ping from ext node to open vz host without the bridge interface works fine (openvz ath0: 10.0.0.8)
ping 10.0.0.8 -i 0.01 -c 100 -q
PING 10.0.0.8 (10.0.0.8) 56(84) bytes of data.

--- 10.0.0.8 ping statistics ---
100 packets transmitted, 100 received, 0% packet loss, time 891ms
rtt min/avg/max/mdev = 1.148/1.496/2.071/0.210 ms

Subject: Re: Bridging veth dev creates massive delay
Posted by MeMu on Wed, 10 Oct 2007 14:24:09 GMT
View Forum Message <> Reply to Message

Hi there...
Anybody has an idea how to solve this issue?

Here another example: ping from ext. node to dummy device in bridge on the vz host

vz modules unloaded
100 packets transmitted, 100 received, 0% packet loss, time 1982ms
rtt min/avg/max/mdev = 1.825/2.678/12.824/1.105 ms

vz modules loaded
100 packets transmitted, 100 received, 0% packet loss, time 1983ms
rtt min/avg/max/mdev = 2.196/4.290/43.214/6.691 ms, pipe 3

vz node started and member of the bridge
100 packets transmitted, 99 received, 1% packet loss, time 2177ms
rtt min/avg/max/mdev = 72.001/508.651/918.070/248.995 ms, pipe 32

If there is an vz node in the bridge, the connection is... well... not as good as it should be

cu
MeMu