Subject: slow network performance
Posted by John Paul Walters on Wed, 28 Mar 2007 20:42:15 GMT
View Forum Message <> Reply to Message

I'm using the faster venet device for several OpenVZ nodes and have been evaluating the network performance using netperf. I'm using kernel 2.6.16-1.2133_FC5.026test015smp with gigabit ethernet on all nodes. What I've found is that the host machines are able to get near-native throughput (around 940 Mbit/s); however, VPS nodes max out around 300 Mbit/s. Xen, on the other hand, is able to achieve network performance similar to native speeds. Is there a configurable option that I'm missing that can improve network performance? I'm getting no failures reported within /proc/user_beancounters.

Any thoughts?

best regards,
JP

Subject: Re: slow network performance
Posted by Vasily Tarasov on Thu, 29 Mar 2007 07:28:16 GMT
View Forum Message <> Reply to Message

Thank you for the info, such tests are very important for us.

I want to ask you, can you conduct this test on the last 2.6.9 OpenVZ kenel:

- in VE0
- in VE

It's also interesting to mesuare throughput from VE to VE0.

Then, please, post the results here. If they will be still bad, we'll try to find out where is the problem.

You help is very appreciated,
Thank you.

Subject: Re: slow network performance
Posted by John Paul Walters on Thu, 29 Mar 2007 17:52:32 GMT
View Forum Message <> Reply to Message

I can't conduct the tests on the original hardware because we are running FC5. I get some package conflicts if I try to install a 2.6.9 series kernel. However, I've conducted the tests again on two RHEL4 based machines that DO support the 2.6.9 series kernel and I've posted some results below.

From VPS->VPS (with kernel 2.6.9-023stab040.1-smp):

```
./netperf -H 10.1.1.209 -l 100
TCP STREAM TEST from 0.0.0.0 (0.0.0.0) port 0 AF_INET to 10.1.1.209 (10.1.1.209) port 0
AF_INET
Recv   Send    Send
Socket Socket  Message  Elapsed
Size   Size    Size     Time     Throughput
bytes  bytes   bytes    secs.    10^6bits/sec

 87380  16384  16384    99.00     461.66
```

From VE0->VE0 (with kernel 2.6.9-023stab040.1-smp):

```
./netperf -H pdcl09 -l 100
TCP STREAM TEST from 0.0.0.0 (0.0.0.0) port 0 AF_INET to pdcl09 (10.1.1.109) port 0 AF_INET
Recv   Send    Send
Socket Socket  Message  Elapsed
Size   Size    Size     Time     Throughput
bytes  bytes   bytes    secs.    10^6bits/sec

 87380  16384  16384    100.21    724.75
```

And here's the output from VPS->VE0:

```
./netperf -H 10.1.1.109 -l 100
TCP STREAM TEST from 0.0.0.0 (0.0.0.0) port 0 AF_INET to 10.1.1.109 (10.1.1.109) port 0
AF_INET
Recv   Send    Send
Socket Socket  Message  Elapsed
Size   Size    Size     Time     Throughput
bytes  bytes   bytes    secs.    10^6bits/sec

 87380  16384  16384    100.00    521.19
```

Just as a comparison point, here's the netperf output from a more recent testing kernel. So matters did improve a bit at on the VE0 side.

From VE0->VE0 (with kernel 2.6.18-ovz028test010.1-smp):

```
./netperf -H pdcl08 -l 100
TCP STREAM TEST from 0.0.0.0 (0.0.0.0) port 0 AF_INET to pdcl08 (10.1.1.108) port 0 AF_INET
Recv   Send    Send
Socket Socket  Message  Elapsed
Size   Size    Size     Time     Throughput
bytes  bytes   bytes    secs.    10^6bits/sec
```

87380 16384 16384 100.00 813.20

But on the VPS->VPS side, things are again pretty poor:

./netperf -H 10.1.1.209
TCP STREAM TEST from 0.0.0.0 (0.0.0.0) port 0 AF_INET to 10.1.1.209 (10.1.1.209) port 0 AF_INET
Recv Send Send
Socket Socket Message Elapsed
Size Size Size Time Throughput
bytes bytes bytes secs. 10^6bits/sec

87380 16384 16384 10.00 476.44

Any thoughts?
JP

Subject: Re: slow network performance
Posted by John Paul Walters on Thu, 29 Mar 2007 22:39:38 GMT
View Forum Message <> Reply to Message

I should also mention that CPU utilization is very high for these tests at approximately 25-35%. Tests on the host node show a much lower CPU load.

Subject: Re: slow network performance
Posted by Vasily Tarasov on Fri, 30 Mar 2007 08:18:11 GMT
View Forum Message <> Reply to Message

Thank you for the detailed information, we'll try to setup testing environment today and if the downgrade of throughput will reproduce - we'll make all efforts to find out where is the stopper ASAP. If it won't reproduce, we'll request you for further information about your environment.

Thank you,
Vasily.

Subject: Re: slow network performance
Posted by Vasily Tarasov on Fri, 30 Mar 2007 18:07:10 GMT
View Forum Message <> Reply to Message

Hello, after our testing we have the following results:

VE0 <-> VE0: 878.95
VE <-> VE: 793.68

VE0 (cl) <-> VE (serv): 872
VE0 (serv) <-> VE (cl): 800

(VE0s and VEs are on different nodes, server was always ran on one node and client on another)

So we had an overhead of 10%

After we swiched clien and server, we received the following results:

VE0 <-> VE0: 920.84
VE <-> VE: 910.84

Overhaed about 1%

So we want to know your environment more precisely. Note that for our testing iptables modules were unloaded, routing was very easy (default), subnet contained only two nodes. Also inform us about your CPU, network cards, etc.

Thanks.

Subject: Re: slow network performance
Posted by John Paul Walters on Fri, 30 Mar 2007 18:47:08 GMT
View Forum Message <> Reply to Message

I tried disabling iptables and rebooting and have confirmed that all iptables modules have been removed from the system. The performance improved a little - VPS <=> VPS got up to 545.68 Mbit/s - but not much.

The networking is default. These nodes are two of an 8 node cluster. Each node contains a dual port ethernet card, of which one port is in use. According to lspci, we've got a Broadcom NetXtreme BCM5721 Gigabit Ethernet PCI Express (rev 11). Each machine has 2 GB of memory. The CPUs are dual core AMD Opteron 175.

Let me know if you need anything further.

JP

Subject: Re: slow network performance
Posted by Vasily Tarasov on Mon, 02 Apr 2007 08:15:17 GMT
View Forum Message <> Reply to Message

Hello,

What Linux distribution is used on Hardware Nodes, and in VEs?

Thank you.

Subject: Re: slow network performance
Posted by John Paul Walters on Mon, 02 Apr 2007 15:13:26 GMT
View Forum Message <> Reply to Message

Thanks for your response. It doesn't seem to make much of a difference in the host distribution. I've seen the same results with either FC5 x86 host and RHEL4 x86_64 hosts. For the tests that I've detailed earlier, we're using centos-4-x86_64-default on RHEL4 x86_64 hosts. However, I've seen similar results when using FC5 hosts and FC5 guests (x86 only).

Also, the CPU utilization seems quite high within the VPS. I wouldn't expect it to reach nearly 30% (the -c -C netperf options).

JP

Subject: Re: slow network performance
Posted by Vasily Tarasov on Mon, 02 Apr 2007 15:28:57 GMT
View Forum Message <> Reply to Message

We're aware about high CPU consuption and we are working on improvement of venet algorithm. But the decrease of throughput is totally unexpected and isn't reproducable on our nodes... Can you give us an uccess to your nodes?

Thanks,
Vasily.

Subject: Re: slow network performance
Posted by John Paul Walters on Mon, 02 Apr 2007 15:40:27 GMT
View Forum Message <> Reply to Message

I may be able to arrange access to the nodes. Do you need system-level access or just a standard user account? What about to the host nodes?

JP

Subject: Re: slow network performance
Posted by Vasily Tarasov on Wed, 04 Apr 2007 12:17:45 GMT

View Forum Message <> Reply to Message

Actually we need a root access to host node, is it possible?

Thanks,
Vasily.

Subject: Re: slow network performance
Posted by John Paul Walters on Thu, 05 Apr 2007 22:30:10 GMT
View Forum Message <> Reply to Message

Vasily,

I'm sorry, I've been unable to get arrange outside access to the nodes. Is there any way that I can help you guys? If there's anything I can do, please let me know.

regards,
JP

Subject: Re: slow network performance
Posted by Vasily Tarasov on Fri, 06 Apr 2007 08:30:17 GMT
View Forum Message <> Reply to Message

It is ok.

Well, let's do the following: next week I'll write a small "specification" of test: how to setup environment, what pay attention on, etc. And you'll try to do it on your nodes. Is it possible? And then step by step, we'll try to find out there is the problem. Do you agree about such model of interaction?

Thanks,
Vasily

Subject: Re: slow network performance
Posted by John Paul Walters on Fri, 06 Apr 2007 14:07:27 GMT
View Forum Message <> Reply to Message

Yes, that'll work.