Subject: Docker daemon fails to start on host but succeeds in VZ container
Posted by abufrejoval on Tue, 06 Dec 2016 17:03:22 GMT
View Forum Message <> Reply to Message

Want to run docker images on host and inside an OpenVZ container (related to CUDA)

The docker daemon fails to start on the host after installation, it fails trying to set up routing via iptables.
Specifically it can't seem to find a 'nat' table

FATA[0001] Error starting daemon: Error initializing network controller: error obtaining controller instance: failed to create NAT chain: iptables failed: iptables --wait -t nat -N DOCKER: iptables v1.4.21: can't initialize iptables table `nat': Table does not exist (do you need to insmod?)
Perhaps iptables or your kernel needs to be upgraded.

And yes, 'cat /proc/net/ip_table_names' only lists raw, mangle and filter.

Following the hint on modprobe, I can only see that there is lots of *net* related modules loaded, but what's even more strange is: It works inside an OpenVZ container.

I used a Centos7 and a VZ7 template, set up an OpenVZ container with each, and installed Docker inside as per instructions (https://openvz.org/Docker_inside_CT).

Funny enough inside the OpenVZ containers /proc/net/ip_table_names *contains* also 'nat' and the Docker daemon has no issues setting up the network at all.

So I guess I can rule out any missing modules.

If I start the Docker daemon on the host with --iptable=false it will run, but Docker containers have no network access.

I can only guess some configuration is responsible for this odd behaviour on the host and I hope you can help me find it.

P.S.:
The issue seems independent of the Docker version or variant. I've tried 1.8.2 which comes with VZ7 and 1.12.3 from docker.com.
It's the 'iptables --wait -t nat -N DOCKER' command which fails for the missing 'nat' (or inaccessible?) table on the host.

The baseline CentOS 7 I run for comparison has nat,mangle,security,raw and filter tables in /proc/net/ip_table_names. Security seems disabled in the VZ .config file but that shouldn't be an issue.

Subject: Re: Docker daemon fails to start on host but succeeds in VZ container

Posted by khorenko on Tue, 06 Dec 2016 20:37:54 GMT
View Forum Message <> Reply to Message

Most probably you have conntracks disabled on the host, check

https://openvz.org/Using_NAT_for_container_with_private_IPs

Subject: Re: Docker daemon fails to start on host but succeeds in VZ container
Posted by abufrejoval on Wed, 07 Dec 2016 11:46:52 GMT
View Forum Message <> Reply to Message

Dead on, nailed it!

On VZ7 "options nf_conntrack ip_conntrack_disable_ve0=1" is in /etc/modprobe.d/vz.conf and the funny thing is I had found that line, 'removed' it and rebooted but that didn't seem to change anything.

Either it defaults to 1, or renaming vz.conf to vz.conf.DISABLED didn't do the job or I botched the reboot.

Setting it to 0 (and rebooting), however, fixed it.

That you soo much!

Any special way to mark this thread as closed?

BTW: Still hoping for news on /proc/driver/nvidia/params or for a hint on where to change that myself (not afraid to compile a kernel, but procfs stuff is all over the place)

Subject: Re: Docker daemon fails to start on host but succeeds in VZ container
Posted by khorenko on Wed, 07 Dec 2016 16:32:05 GMT
View Forum Message <> Reply to Message

Tried the same:
mv /etc/modprobe.d/vz.conf /etc/modprobe.d/vz.conf.DISABLED
cat /etc/modprobe.d/vz.conf.DISABLED
options vzevent reboot_event=1
#options nf_conntrack ip_conntrack_disable_ve0=0
options nf_conntrack ip_conntrack_disable_ve0=1

reboot

cat /sys/module/nf_conntrack/parameters/ip_conntrack_disable_ve0
0

So in case you face again that omitting ip_conntrack_disable_ve0=0 makes your node to disable conntracks, please let us know.

Subject: Re: Docker daemon fails to start on host but succeeds in VZ container
Posted by abufrejoval on Wed, 07 Dec 2016 20:22:52 GMT
View Forum Message <> Reply to Message

Well it's official now, I am sloppy!

I repeated the procdure on my system and it behaves as you describe: Unless specifically disabled for VE0 connection tracking will be enabled also on the host.

After careful examination I cannot find evidence on me rebooting (at least not via the command line) after renaming vz.conf to vz.conf.DISABLED (no time stamps on shell history, no auditing enabled).

I however immediately rebooted after doing the reverse operation.

I knew I'd have to reboot for those changes to be effective, but evidently I did get distracted after doing the right thing first...

Thank you for taking quality seriously and sorry for making you worry in vain!
