Subject: Re: [patch 2/6] [Network namespace] Network device sharing by view
Posted by Andrey Savochkin on Fri, 30 Jun 2006 15:01:55 GMT
View Forum Message <> Reply to Message

Jamal,

On Fri, Jun 30, 2006 at 09:50:52AM -0400, jamal wrote:
>
> BTW - I was just looking at openvz, very impressive. To the other folks,

Thanks!

> I am not putting down any of your approaches - just havent
> had time to study them. Andrey, this is the same thing you guys have
> been working on for a few years now, you just changed the name, correct?

The relations are more complicated than just the change of name,
but yes, OpenVZ represents the result of our work for a few years.

>
> Ok, since you guys are encouraging me to speak, here goes ;->
> Hopefully this addresses the other email from Herbert et al.
>
[snip]
> // create the guest
> [host-node]# vzctl create 101 --ostemplate fedora-core-5-minimal
> // create guest101::eth0, seems to only create config to boot up with
> [host-node]# vzctl create 101 --netdev eth0
> // bootup guest101
> [host-node]# vzctl start 101
>
> As soon as bootup of guest101 happens, creating guest101::eth0 should activate
> creation of the host side netdevice. This could be triggered for example by
> the netlink event message seen on host whic- which is a result of creating guest101::eth0
> Which means control sits purely in user space.

I'd like to clarify you idea: whether this host-side device is a real
device capable of receiving and transmitting packets (by moving them between
namespaces), or it's a fake device creating only a view of other namespace's
devices?

[snip]
> > However, I oppose the idea of automatic mirroring of _all_ devices appearing
> > inside some namespaces ("guests") to another namespace (the "host").
> > This clearly goes against the concept of namespaces as independent realms,
> > and creates a lot of problems with applications running in the host, hotplug
> > scripts and so on.
> >

>
> I was thinking that the host side is the master i.e you can peek at
> namespaces in the guest from the host.

"Host(master)-guest" relations is a valid and useful scheme.
However, I'm thinking about broader application of network namespaces,
when they can form an arbitrary tree and may not be in "host-guest" relations.

> Also note that having the pass through device allows for guests to be
> connected via standard linux schemes in the host side (bridge, point
> routes, tc redirect etc); so you dont need a speacial device to hook
> them together.

What do you mean under pass through device?
Do you mean using guest1-tun0 as a backdoor to talk to the guest?

>
> > > > Then the pragmatic question becomes how to correlate what you see from
> > > > `ip addr list' to guests.
> > >
> > > on the host ip addr and the one seen on the guest side are the same.
> > > Except one is seen (on the host) on guest0-eth0 and another is seen
> > > on eth0 (on guest).
> >
> > Then what to do if the host system has 10.0.0.1 as a private address on eth3,
> > and then interfaces guest1-tun0 and guest2-tun0 both get address 10.0.0.1
> > when each guest has added 10.0.0.1 to their tun0 device?
> >
>
> Yes, that would be a conflict that needs to be resolved. If you look at
> ip addresses as also belonging to namespaces, then it should work, no?
> i am assuming a tag at the ifa table level.

I'm not sure, it's complicated.
You wouldn't want automatic local routes to be added for IP addresses on
the host-side interfaces, right?
Do you expect these IP addresses to act as local addresses in other places,
like answering to arp requests about these IP on all physical devices?

But anyway, you'll have conflicts on the application level.
Many programs like ntpd, bind, and others fetch the device list using the
same ioctls as ifconfig, and make (un)intelligent decisions basing on what
they see.
Mirroring may have some advantages if I am both host and guest administrator.
But if I create a namespace for my friend Joe to play with IPv6 and sit
tunnels, why should I face inconveniences because of what he does there?

Best regards

Andrey