Subject: Re: [patch 2/6] [Network namespace] Network device sharing by view
Posted by ebiederm on Wed, 28 Jun 2006 14:39:38 GMT
View Forum Message <> Reply to Message

jamal <hadi@cyberus.ca> writes:

> On Wed, 2006-28-06 at 15:36 +0200, Herbert Poetzl wrote:
>
>> note: personally I'm absolutely not against virtualizing
>> the device names so that each guest can have a separate
>> name space for devices, but there should be a way to
>> 'see' _and_ 'identify' the interfaces from outside
>> (i.e. host or spectator context)
>>
>
> Makes sense for the host side to have naming convention tied
> to the guest. Example as a prefix: guest0-eth0. Would it not
> be interesting to have the host also manage these interfaces
> via standard tools like ip or ifconfig etc? i.e if i admin up
> guest0-eth0, then the user in guest0 will see its eth0 going
> up.
>
> Anyways, interesting discussion.

Please no.

We really want the fundamental rule that a network device
is tied to a single namespace, and that a socket is tied
to a single namespace. If those two conditions are met
we don't have to tag packets with a namespace identifier.

We only have to modify hash table lookups in the networking
code to look at a namespace tag in addition to the rest because
that is less expensive than allocating new hash tables.

Currently with a network device only being usable in one
network namespace we have the situation where we can
fairly safely give a guest CAP_NET_ADMIN without problems.

In addition currently nothing in the implementation knows about
the hierarchical structure of how the network namespace will be
used. To allow ifconfig guest0-eth0 to work would require
understanding the hierarchical structure and places serious questions
on how safe we can make CAP_NET_ADMIN.

Now I am open to radically different designs if they allow the
implementation cost to be lower and they have clean semantics,
and don't wind up being an ugly unmaintainable wart on the linux

networking stack. The only route I could imagine such a thing coming
from is something like tagging flows, in some netfiler like way.
Which might allow ifconfig guest-eth0 from the host without problems.
But I have not seen such a design.

Eric