Subject: Re: [patch 3/3][netns] net: hide master/linked interface from netlink
Posted by Daniel Lezcano on Thu, 27 Sep 2007 09:15:45 GMT

Eric W. Biederman wrote:
> dlezcano@fr.ibm.com writes:
>
>> From: Daniel Lezcano <dlezcano@fr.ibm.com>
>>
>> Actually when a network device is linked to another, the name appears
>> to be @<link>. For example, if a macvlan0 is created on top of eth0,
>> the ip link show is:
>>
>> 6: macvlan0@eth0: <BROADCAST,MULTICAST> mtu 1500 qdisc noop
>>   link/ether 6a:d4:10:0d:a8:55 brd ff:ff:ff:ff:ff:ff
>>
>> But if we move macvlan0 to a network namespace, eth0 does no longer
>> exist inside it and the result will be:
>>
>> 6: macvlan0@if2: <BROADCAST,MULTICAST> mtu 1500 qdisc noop
>>   link/ether 6a:d4:10:0d:a8:55 brd ff:ff:ff:ff:ff:ff
>>
>> if2 is, I guess, some random value. That can do invalid memory
>> access or inconsistent data showing.
>>
>> The patchset will avoid such case, it checks if the linked device exist
>> into the current network namespace and if it doesn't the result will
>> be:
>>
>> 6: macvlan0@NONE: <BROADCAST,MULTICAST> mtu 1500 qdisc noop
>>   link/ether 6a:d4:10:0d:a8:55 brd ff:ff:ff:ff:ff:ff
>
> Hmm. Currently the ifindex space is global. Something that ultimately
> needs to be fixed to handle migration so we can preserve the ifindex
> on a device when we migrate it.

One interesting thing with the global ifindex, is we can keep track to
the network device. For example, from outside the namespace we create
macvlan0, push it to the netns, inside the netns, we rename it eth0 (eg.
to fit os template), the netns finishes and the netdev is pop out to the
initial network namespace. The names will conflict and the eth0 (former
macvlan0) is renamed to dev0. If we want to automate the destruction of
such interface, we should be able to identify them when the namespace
exits. The global ifindex allow that.

If the ifindex is changed to be local to the namespace, the ifindex will
be changed each time we change namespaces. So in this case, it will be
hard to track these interfaces.

This point is out of migration consideration and having a ifindex per namespace makes sense.

Should we consider an identifier for the network devices ? so we can setup this identifier each time we create/migrate an interface and find them by this identifier.

> So the @if2 piece is harmless, as it is just reporting the ifindex
> number because it can't figure out the name that corresponds to
> that network device.
>
> So the worst we get is hints that some extra data exists somewhere.
> I suspect the correct fix is to not even write the additional attribute
> in this case instead of setting the value to zero.
>
> I am not yet convinced we need to handle this case yet, at most this
> is a cosmetic issue.

_____

Containers mailing list
Containers@lists.linux-foundation.org
https://lists.linux-foundation.org/mailman/listinfo/containers