
Subject: Re: L2 network namespace benchmarking
Posted by [Daniel Lezcano](#) on Wed, 28 Mar 2007 07:07:56 GMT
[View Forum Message](#) <> [Reply to Message](#)

Herbert Poetzl wrote:

> On Wed, Mar 28, 2007 at 12:16:34AM +0200, Daniel Lezcano wrote:
>> Hi,

[cut]

>> 3. General observations

>> -----

>>

>> The objective to have no performances degradations, when the network
>> namespace is off in the kernel, is reached in both solutions.

>>

>> When the network is used outside the container and the network
>> namespace are compiled in, there is no performance degradations.

>>

>> Eric's patchset allows to move network devices between namespaces and
>> this is clearly a good feature, missing in the Dmitry's patchset. This
>> feature helps us to see that the network namespace code does not add
>> overhead when using directly the physical network device into the
>> container.

>>

>> The loss of performances is very noticeable inside the container and
>> seems to be directly related to the usage of the pair device and the
>> specific network configuration needed for the container. When the
>> packets are sent by the container, the mac address is for the pair
>> device but the IP address is not owned by the host. That directly
>> implies to have the host to act as a router and the packets to be
>> forwarded. That adds a lot of overhead.

>>

>> A hack has been made in the ip_forward function to avoid useless
>> skb_cow when using the pair device/tunnel device and the overhead
>> is reduced by the half.

>

> would it be possible to do some tests regarding scalability?

>

> i.e. I would be interested how the following would look like:

>

> 10 connections on a single host (in parallel, overall performance)

> 10 connections from the same net space

> 10 connections from 10 different net spaces

> (i.e. one connection from each space)

>

> we can assume that L3 isolation will give similar results to

> the first case, but if needed, we can provide a patch to

> test this too ...

>

Ok. Assuming, Eric's and Dmitry's patchset are very similar, I will focus on the Eric's patchset because it is more mature and more easy to setup. I will have a look on the bridge optimization before doing that.

>

> PS: great work! tx!

>

Thanks.

Containers mailing list

Containers@lists.linux-foundation.org

<https://lists.linux-foundation.org/mailman/listinfo/containers>
