
Subject: Re: L2 network namespace benchmarking
Posted by [ebiederm](#) on Wed, 28 Mar 2007 12:06:19 GMT
[View Forum Message](#) <> [Reply to Message](#)

Kirill Korotaev <dev@sw.ru> writes:

>> Ideally we can optimize the bridge code or something equivalent to
>> it so that we can take one look at the destination mac address and
>> know which network namespace we should be in. Potentially moving this
>> work to hardware when the hardware supports multiple queues.
> yes, we can hack the bridge, so that packets coming out of eth devices
> can go directly to the container and get out of veth devices from
> inside the container.
>
>> If we can get the overhead out of the routing code that would be
>> tremendous. However I think it may be more realistic to get the
>> overhead out of the ethernet bridging code where we know we don't need
>> to modify the packet.
> Why not optimize both? :)

If the optimizations are safe and correct I don't have a problem.

When we seem to have multiple copies of a packet in circulation and we skip a what appears to be a required copy on write, I'm dubious.

Although the more I look at suggested optimization the less dubious I am as it appears all we are skipping is a ttl decrement and the cow flag exclusively applies to the data chunk and not the header chunk of the packet whatever that means.

However we still need to guard against a loop in our routing table setup between multiple guests.

Eric

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