Subject: Re: [RFD] L2 Network namespace infrastructure
Posted by dev on Thu, 28 Jun 2007 14:53:01 GMT
View Forum Message <> Reply to Message

Eric W. Biederman wrote:
> Patrick McHardy <kaber@trash.net> writes:
>
>
>>Eric W. Biederman wrote:
>>
>>>-- The basic design
>>>
>>>There will be a network namespace structure that holds the global
>>>variables for a network namespace, making those global variables
>>>per network namespace.
>>>
>>>One of those per network namespace global variables will be the
>>>loopback device. Which means the network namespace a packet resides
>>>in can be found simply by examining the network device or the socket
>>>the packet is traversing.
>>>
>>>Either a pointer to this global structure will be passed into
>>>the functions that need to reference per network namespace variables
>>>or a structure that is already passed in (such as the network device)
>>>will be modified to contain a pointer to the network namespace
>>>structure.
>>
>>
>>I believe OpenVZ stores the current namespace somewhere global,
>>which avoids passing the namespace around. Couldn't you do this
>>as well?
>
>
> It sucks. Especially in the corner cases.   Think macvlan
> with the real network device in one namespace and the ``vlan''
> device in another device.

sorry, what's the problem here? I don't see a single problem here
related to the global current context.

> The implementation of a global is also pretty a little questionable.
> Last I looked it didn't work on the transmit path at all and
> interesting on the receive path.
>
> Further and fundamentally all a global achieves is removing the need
> for the noise patches where you pass the pointer into the various
> functions. For long term maintenance it doesn't help anything.

this is not 100% true.

Having global context also helps:
1. to account for time spent in network processing or some other activity
   (like IRQ handling) to appropriate namespace.

2. it is really usefull and helpfull for functions like printk() with virtualized
   syslog buffer. Otherwise you would need to find a context in every place where
   namespace private message should be printed.

   Actually it is the same natural as the 'current' which is global,
   but could be passed from entry.S instead to every function which needs it.

   OVZ experience just tells me that this helps to avoid all this noise
   and not harder to work with then with 'current'. But it's up to you.

Thanks,
Kirill

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