Subject: Re: [patch 1/2][NETNS][RFD] store the network namespace pointer in the dst_entry structure
Posted by Benjamin Thery on Tue, 11 Dec 2007 17:24:07 GMT
View Forum Message <> Reply to Message

Eric W. Biederman wrote:
> Daniel Lezcano <dlezcano@fr.ibm.com> writes:
>
>>> Could you please place the struct net *net pointer up by the
>>> network device pointer.
>>>> };
>>> I know we need a net pointer in struct rt_table, because it
>>> is a hash table that we can't dynamically allocate so we need
>>> to place a network namespace pointer as part of the hash key.
>>>
>>> For the ipv6 fib tables I don't recall needing a net pointer as we didn't have
>>> a hash table and could instead have separate
>>> roots for different namespaces.
>> Yes don't need for the hash table but we used it to pass the network namespace
>> parameter to the underlying function which need the net parameter.
>>
>> We are facing two problems when removing the fl_net field from flowi:
>>
>> * The first one is the fl_net is used as a key. This problem can be handled
>> simply in moving the netns to the rtable.
>
> Yes.
>
>> * The second one is the usage made by the fl_net to pass through the different
>> function calls the network namespace pointer without changing all functions
>> signature. This problem can be solved if we put the netns pointer in the
>> dst_entry structure, so when we are in ipv4, we use container_of on rtable and
>> when we are in ipv6, we use the container_of on rt6_info. So everywhere with the
>> flowi, we can retrieve the netns.
>
> That doesn't work as rt6_info does not currently hold a struct flowi.
>
>>> I find this slightly odd as I didn't wind up needing to add
>>> a struct net pointer in struct dst in my proof of concept tree
>>> and struct dst doesn't have a struct flowi so that would not
>>> have prevented it.
>> The idea is to put the net in the dst_entry because it is accessible from rtable
>> or rt6_info and these ones contain a flowi field.
>
> And since that isn't true, the idea seems to fall flat on it's face.

My fault.
While talking with Daniel last week I suggested to put the net in

dst_entry instead of rtable, because dst_entry was common to rtable
and rt6_info. I thought we could factorize some code this way. I
wrongly assumed IPv6 was pretty similar to IPv4 in the way it handles
flowi. I should have checked that more carefully. Crap ;)

Benjamin

>
> I expect most of the instances of struct flowi that we would be
> looking things up with would be on the stack so the earlier concerns
> raised would likely still need to be addressed.
>
> Eric
>

--
B e n j a m i n T h e r y - BULL/DT/Open Software R&D

http://www.bull.com

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