Subject: netns refcounting
Posted by den on Fri, 16 Nov 2007 14:04:08 GMT
View Forum Message <> Reply to Message

Hello, All!

During port of Eric's patches I have noticed an interesting thing.
The patch "net: Teach the ipv4 route cache to handle multiple network
namespaces" call hold_net for each IPv4 DST cache entry.

Though it is not possible to stop a namespace without stopping all the
devices inside. Additionally, the device can't be unregistered if there
are dst entries to it. These entries are moved to a namespace loopback
and the namespace will block until these entries will gone from a loopback.

So, I do not see a necessity to have an extra atomic on this hot path,
i.e. hold_net can re moved away for this. Are there any holes?

Regards,
Den

Subject: Re: netns refcounting
Posted by Daniel Lezcano on Fri, 16 Nov 2007 15:50:13 GMT
View Forum Message <> Reply to Message

Denis V. Lunev wrote:
> Hello, All!
>
> During port of Eric's patches I have noticed an interesting thing.
> The patch "net: Teach the ipv4 route cache to handle multiple network
> namespaces" call hold_net for each IPv4 DST cache entry.
>
> Though it is not possible to stop a namespace without stopping all the
> devices inside. Additionally, the device can't be unregistered if there
> are dst entries to it. These entries are moved to a namespace loopback
> and the namespace will block until these entries will gone from a loopback.
>
> So, I do not see a necessity to have an extra atomic on this hot path,
> i.e. hold_net can re moved away for this. Are there any holes?

That seems reasonable to remove, good catch.

Subject: Re: netns refcounting
Posted by ebiederm on Fri, 16 Nov 2007 17:14:40 GMT
View Forum Message <> Reply to Message

"Denis V. Lunev" <den@sw.ru> writes:

> Hello, All!
>
> During port of Eric's patches I have noticed an interesting thing.
> The patch "net: Teach the ipv4 route cache to handle multiple network
> namespaces" call hold_net for each IPv4 DST cache entry.
>
> Though it is not possible to stop a namespace without stopping all the
> devices inside. Additionally, the device can't be unregistered if there
> are dst entries to it. These entries are moved to a namespace loopback
> and the namespace will block until these entries will gone from a loopback.
>
> So, I do not see a necessity to have an extra atomic on this hot path,
> i.e. hold_net can re moved away for this. Are there any holes?

So all hold_net is good for is to catch logic errors where people are
still using a network namespace but have freed it. It is a bit like
the device reference count in that regard.

The usage of the dst cache from ipv4 looked complicated enough that making
some small mistake easy to have so I included the reference count for good
measure.

If the movement to loopback device is sufficient to prevent us from goofing
up in this area I am happy to avoid the reference count.

Eric