Subject: Re: [RFC][PATCH 0/7] Clone PTS namespace Posted by serue on Sat, 26 Apr 2008 13:02:43 GMT

View Forum Message <> Reply to Message

```
Quoting Eric W. Biederman (ebiederm@xmission.com):
> "Serge E. Hallyn" <serue@us.ibm.com> writes:
>
>
>> Heh, well I tried several approaches - adding tag ops to kset, to ktype,
>> etc. Finally ended up just calling sysfs enable tagging on
> > /sys/kernel/uids when that is created. It's now working perfectly.
> Sounds good.
>
>>> I suspect since you are working on this and I seem to be stuck
>>> in molasses at the moment it makes sense to figure out what it
> >> will take to handle the uid namespace before pushing these
>>> patches again.
> >
>> I had ported your patches to 2.6.25, but Benjamin in the meantime ported
> > them to 2.6.25-mm1. Since that's closer to the -net tree it's a more
> > useful port, so I'll let him post his patchset. Then I'll send the
>> userns patch on top of that. While I'm not actually able to send
> > network traffic over a veth dev (I probably am still not setting it up
> > right), I am able to pass veth devices into network namespaces, and the
> > user namespaces are properly handled.
> >
>> I believe Benjamin did notice a problem with some symlinks not existing,
> > and I think we want one more patch on top of yours removing the
> > hold_net() from sysfs_mount, which I don't think was what you really
>> wanted to do. By simply removing that, if all tasks in a netns go away,
>> the netns actually goes away and a lookup under a bind-mounted copy of
> > its /sys/class/net is empty.
> I will have to look, I need to refresh myself on where all of this code is.
> I think hold net was what I wanted. A record that there is a user
> but not something that will keep the network namespace from going away.
>
> Essentially hold_net should be a debugging check rather then a
> real limitation.
Ah, I see, I assumed it actually pinned it. Sorry, never mind then:)
-serge
> > Anyway the patches should be hitting the list next week.
> Cool. We can figure out what we need to do to merge them from
```

> there.
> >> Taking a quick look and having a clue what we will need to
> >> do for a theoretical device namespace is also a possibility.
> > I'm not sure I'm familiar enough with the kobject/class/sysfs/device
> > relationships yet to comment on that. It doesn't look like it should
> really be a problem, though simply adding tags to every directory
> under /sys/class (/sys/class/tty, /sys/class/usb_device, etc) doesn't
> seem like necessarily the nicest way to go...
> True. And the goal is something maintainable. There are still a lot
> of implications of a device namespace left unexamined so we shall see.
> Eric
Containers mailing list
Containers@lists.linux-foundation.org
https://lists.linux-foundation.org/mailman/listinfo/containers