
Subject: Re: [RFC][PATCH 0/7] Clone PTS namespace

Posted by [serue](#) on Fri, 25 Apr 2008 19:21:02 GMT

[View Forum Message](#) <> [Reply to Message](#)

Quoting Eric W. Biederman (ebiederm@xmission.com):

> "Serge E. Hallyn" <serue@us.ibm.com> writes:

>

> > Quoting Serge E. Hallyn (serue@us.ibm.com):

> >> Quoting Eric W. Biederman (ebiederm@xmission.com):

> >> > "Serge E. Hallyn" <serue@us.ibm.com> writes:

> >> >>

> >> >> I'm hoping to be able to get back at this in the week or so as things

> >> >> settle down from my move. My last patches should be in my proof of

> >> >> concept network namespace tree, if they don't show up elsewhere.

> >> >

> >> > Is that the tree I'd get from

> >> >

> >> > git-fetch

> >> > git://git.kernel.org/pub/scm/linux/kernel/git/ebiederm/linux-2.6-netns.git

> >> > master:ebieder.master

> >> >

> >> > Yes.

> >> >

> >> > ? So I'd add a user_ns to the struct sysfs_tag_info?

> >> >

> >> > If so I'll give it a whirl.

> >> >

> >> > Sounds good. My apologies I keep being almost on the verge

> >> > of getting someplace.

> >>

> >> Ok I've got the sysfs relevant patches ported to 2.6.25, and am looking

> >> at how to extend it to handle /sys/kernel/uids. You have tagging tied

> >> intimately to struct class. So the question is should I generalize the

> >> taggint to deal with kobjects instead, or create a struct class user

> >> and make /sys/kernel/uids a symlink to /sys/class/user/uids?

> >

> > Heh, never mind, I was thinking class was a kobject class, not a device

> > class :) So I'll just have to generalize tagging.

>

> Yes. You just need a way to get the tags there.

>

> At the level of sysfs it is fairly general.

> Getting through the kobject layer is a different story.

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.

> 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.

Anyway the patches should be hitting the list next week.

> 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...

thanks,
-serge

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