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

Posted by [serue](#) on Tue, 22 Apr 2008 14:25:39 GMT

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

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

> On Wed, 2008-04-09 at 20:59 -0500, Serge E. Hallyn wrote:

>
> > Feh, so of course sysfs would have the most interactions for a device
> > namespace, but now we have pty, network, and user namespace all needing
> > some sort of sysfs solution. For a quickfix for
> > CONFIG_USER_SCHED+CONFIG_USER_NS, I just moved /sys/kernel/uids/<uid>
> > to /sys/kernel/uids/<userns_address>/<uid>. But what would be a *good*
> > general solution?
> >
> > In -s /sys /proc/self/sys?
>
> LOL
>
> In /proc I prefer the /proc/self approach because we can do it and it is
> just much easier to setup and use. (Plus we have weird problems if we
> try and capture more then the pid namespace).
>
> For other filesystems the only really viable option is to capture
> namespaces at mount time, as we are doing for devpts and proc with
> respect to the pid namespace.
>
> For the network namespace where it is very much more then a single
> directory with symlinks from physical devices pointing at logical
> network interfaces.
>
> My last effort in that area was ok'd by Tejun lightly tested by a few
> others and misplaced by gregkh so I don't think we have a real problem
> with resurrecting those patches cleaning them up a bit and merging them.
>
> The biggest gotcha with sysfs is that the VFS locking for the dcache
> is in the wrong order for distributed filesystems, where we would like
> to make the change atomically on the server and then make the change in
> the local cache. Or in this case the sysfs internal data structures.
> The truly nasty case is supporting rename (as sysfs does) as the VFS
> is not at all happy if you just punt and shoot down the dentries and
> instantiate new ones.
>
> 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

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

If so I'll give it a whirl.

thanks,
-serge

> It isn't my perception that we have a design problem rather, we just
> need to move an important piece of code in a subtle and moderately
> uninteresting direction for it's primary maintainer.
>
> Further what I did for the network namespace should easily handle the
> uid/gid namespace and should be a good starting place for a general
> device namespace.
>
> Eric
>

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