Hello,

Eric W. Biederman wrote:
> How close are we to the point where we can get mount sysfs multiple
> times and get multiple dentry trees with different super blocks?

Yeah, that sounds much better. We only have to pay attention to getting
sysfs_dirent tree correct. The rest can be done by just looking up the
correct sysfs_dirent in sysfs_lookup(). We would still need to pin all
shadows to keep sysfs_get_dentry() working.

> That really does sound like the right way to go. Especially as it
> simplifies the monitoring of containers. If you want to watch what
> the view looks like in some container your bind mount his sysfs and
> look at that.
>
> If we can do that the dcache side at least will be beautiful. And
> with a little care we may be able to reduce the work to a special case
> in lookup, some extra handling to mark directories as belonging only
> to a certain mount of sysfs.
>
> If we can find something that is stupid and simple I'm all for that.

Amen.

> To reach the no-kobj utopia we may also need a special device_migrate
> that is a super set of device_rename (because sometimes we need to
> rename devices when we move them between namespaces).

One thing I'm curious about is which semantic is appropriate behavior
when a node is migrated from one namespace to another - renaming or
deactivation followed by activation in new name space. I guess it
doesn't really matter.

> So are we close to having a sysfs that we can have multiple super
> blocks for?

Sorry but I dunno. It sounds much more appealing than other approaches tho.

--
tejun
_____
Containers mailing list

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