
Subject: Re: [PATCH 4/8] user ns: hook permission
Posted by [serue](#) on Wed, 24 Jan 2007 19:06:33 GMT
[View Forum Message](#) <> [Reply to Message](#)

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

> "Serge E. Hallyn" <serue@us.ibm.com> writes:
>
> > From: Serge E. Hallyn <serue@us.ibm.com>
> > Subject: [PATCH 4/8] user ns: hook permission
> >
> > Hook permission to check vfsmnt->user_ns against current.
>
> This looks wrong on several levels.
> - This should ultimately be inside generic_permission instead of
> permission as there are some distributed filesystems that know how to cope with
> multiple mount namespaces simultaneous.
>
> - As implemented the test is not what I would expect. I would
> expect comparisons of uid X == uid Y and gid X == gid Y to
> be replaced by comparing the tuples of uid namespace and uid.
> Which would allow access to world readable/writeable files,
> and it would allow users with CAP_DAC_OVERRIDE to be able to access
> everything.

Whoa - why on earth would we want that?

> All we are really saying as I understand a user namespace is that
> instead of uid's uniquely identifying a user the pair the pair uidns,
> uid is uniquely identifies a user.

Ok that would be one way to interpret it, but it is insufficient for
preventing root in one vserver from messing with users in another
vserver.

> Because you didn't pick what I would consider the obvious choice
> you now need an extra mount flag to disable the uid namespace all
> together, so you can transition through the intermediate uid namespace
> state. That really feels wrong.

Some bit of required bootstrapping seems both acceptable and expected
to me.

> All mounts should have an associated uid namespace and the only
check

> way you should be able to ignore that is to access filesystems
> that can cope with multiple uid namespaces simultaneously.

But it's my fs on my box, why shouldn't i be able to say all uid namespaces can acces this subtree read-only, just bc you feel the fs is inadequate? :)

Note that the tiniest of trees, with just a statically compiled bash, mount, pivot_mount, and initrc, should suffice, mounted readonly for all uid namespaces to use to bootstrap.

-serge

Containers mailing list
Containers@lists.osdl.org
<https://lists.osdl.org/mailman/listinfo/containers>
