Subject: Re: design of user namespaces
Posted by serue on Mon, 07 Jul 2008 15:24:06 GMT
View Forum Message <> Reply to Message

Quoting Eric W. Biederman (ebiederm@xmission.com):
> "Serge E. Hallyn" <serue@us.ibm.com> writes:
>
> > Quoting Eric W. Biederman (ebiederm@xmission.com):
> >>
> >> The very important points are that it is a remount of an existing mount
> >> so that we don't have to worry about corrupted filesystem attacks, and
> >> that authentication is performed at mount time.
> >
> > Conceptually that (making corrupted fs attacks a non-issue) is
> > wonderful. Practically, I may be missing something: When you say
> > remount, it seems you must either mean a bind mount or a remount. If
> > remount, then that will want to change superblock flags. If the
> > child userns(+child mntns) does a real remount, then that will change
> > the flags for the parent ns as well, right?
> >
> > If instead we do a bind mount we don't have that problem, but then the
> > fs can't be the one doing the user namespace work.
> >
> > I'm probably missing something.
>
> Essentially I am creating a new mount operation that is a
> cousin of a remount.
>
> Unlike a real remount you can't change the super flags.
> Unlike a bind mount you get the fs involved, and you pass in a string of flags
> that the fs can interpret in a standard way.
>
> I expect the flags you pass in would be a subset of what is allowed
> in a normal remount.
>
> Which is why I was calling it nativemount. Although usernsmount
> may be better.
>
> Eric

Ah, ok.

Now you haven't started any sort of coding for this yet, right? I'm
hoping to get some time later this week to think about/play with this.

-serge

_____
Containers mailing list