
Subject: Re: Re: [patch 05/10] add "permit user mounts in new namespace" clone flag

Posted by [serue](#) on Tue, 17 Apr 2007 14:28:45 GMT

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

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

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

> >>

> >> Why are directory permissions not sufficient to allow/deny non-privileged
> > mounts?

> >> I don't understand that contention yet.

> >

> > The same scenarios laid out previously in this thread. I.e.

> >

> > 1. user hallyn does mount --bind / /home/hallyn/root

> > 2. (...)

> > 3. admin does "deluser hallyn"

> >

> > and deluser starts wiping out root

> >

> > Or,

> >

> > 1. user hallyn does mount --bind / /home/hallyn/root

> > 2. backup daemon starts backing up /home/hallyn/root/home/hallyn/root/home...

> >

> > So we started down the path of forcing users to clone a new namespace

> > before doing user mounts, which is what the clone flag was about. Using

> > per-mount flags also suffices as you had pointed out, which is being

> > done here. But directory permissions are inadequate.

>

> Interesting....

>

> So far even today these things can happen, however they are sufficiently

> unlikely the tools don't account for them.

>

> Once a hostile user can cause them things are more of a problem.

>

> > (Unless you want to tackle each problem legacy tool one at a time to

> > remove problems - i.e. deluser should umount everything under

> > /home/hallyn before deleting, backup should be spawned from it's own

> > namespace cloned right after boot or just back up on one filesystem,

> > etc.)

>

> I don't see a way that backup and deluser won't need to be modified

> to work properly in a system where non-privileged mounts are allowed,

> at least they will need to account for /share.

>

> That said it is clearly a hazard if we enable this functionality by

> default.
>
> If we setup a pam module that triggers on login and perhaps when
> cron and at jobs run to setup an additional mount namespace I think
> keeping applications locked away in their own mount namespace is
> sufficient to avoid hostile users from doing unexpected things to
> the initial mount namespace. So unless I am mistake it should be
> relatively simple to prevent user space from encountering problems.
>
> That still leaves the question of how we handle systems with an old
> user space that is insufficiently robust to deal with mounts occurring
> at unexpected locations.
>
>
> I think a simple sysctl to enable/disable of non-privileged mounts
> defaulting to disabled is enough.

There is a sysctl for max_user_mounts which can be set to 0.

So a simple on/off sysctl is unnecessary, but given that admins might wonder whether 0 means infinite :), and I agree on/off is important, a second one wouldn't hurt.

> Am I correct or will it be more difficult than just a little pam
> module to ensure non-trusted users never run in the initial mount
> namespace?
>
> Eric

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