Subject: Re: Re: [patch 05/10] add "permit user mounts in new namespace" clone flag
Posted by serue on Mon, 16 Apr 2007 19:56:52 GMT
View Forum Message <> Reply to Message

Quoting Eric W. Biederman (ebiederm@xmission.com):
> Miklos Szeredi <miklos@szeredi.hu> writes:
>
> >> > That depends. Current patches check the "unprivileged submounts
> >> > allowed under this mount" flag only on the requested mount and not on
> >> > the propagated mounts. Do you see a problem with this?
> >>
> >> I think privileges of this sort should propagate. If I read what you
> >> just said correctly if I have a private mount namespace I won't be able
> >> to mount anything unless when it was setup the unprivileged submount
> >> command was explicitly set.
> >
> > By design yes. Why is that a problem?
>
> It certainly doesn't match my intuition.
>
> Why are directory permissions not sufficient to allow/deny non-priveleged 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.

(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.)

-serge

> I should probably go back and look and see how plan9 handles mount/unmount
> permissions. Plan9 gets away with a lot more because it doesn't have
> a suid bit and mount namespaces were always present, so they don't have
> backwards compatibility problems.
>
> My best guess at the moment is that plan9 treated mount/unmount as
> completely unprivileged and used the mount namespaces to limit the
> scope of what would be affected by a mount/unmount operation. I think
> that may be reasonable in linux as well but it will require the
> presence of a mount namespace to limit the affects of what a user can
> do.
>
> So short of a more thorough audit I believe the final semantics should
> be:
> - mount/unmount for non-priveleged processes should only be limited
>   by the mount namespace and directory permissions.
> - CLONE_NEWNS should not be a privileged operation.
>
> What prevents us from allowing these things?
>
> - Unprivileged CLONE_NEWNS and unprivileged mounts needs resource
>   accounting so we don't have a denial of service attack.
>
> - Unprivileged mounts must be limited to directories that we have
>   permission to modify in a way that we could get the same effect
>   as the mount or unmount operation in terms of what files are visible
>   otherwise we can mess up SUID executables.
>
> - Anything else?
>
> There are user space issues such as a reasonable pam module and how
> to do backups. However those are user space issues.
>
> What am I missing that requires us to add MNT_USER and MNT_USERMNT?
>
> Eric
> -
> To unsubscribe from this list: send the line "unsubscribe linux-fsdevel" in
> the body of a message to majordomo@vger.kernel.org
> More majordomo info at http://vger.kernel.org/majordomo-info.html