Subject: Re: Re: [patch 05/10] add "permit user mounts in new namespace" clone flag
Posted by Ram Pai on Tue, 17 Apr 2007 19:28:31 GMT
View Forum Message <> Reply to Message

On Tue, 2007-04-17 at 19:44 +0200, Miklos Szeredi wrote:
> > I'm a bit lost about what is currently done and who advocates for what.
> >
> > It seems to me the MNT_ALLOWUSERMNT (or whatever :) flag should be
> > propagated. In the /share rbind+chroot example, I assume the admin
> > would start by doing
> >
> > mount --bind /share /share
> > mount --make-slave /share
> > mount --bind -o allow_user_mounts /share (or whatever)
> > mount --make-shared /share
> >
> > then on login, pam does
> >
> > chroot /share/$USER
> >
> > or some sort of
> >
> > mount --bind /share /home/$USER/root
> > chroot /home/$USER/root
> >
> > or whatever. In any case, the user cannot make user mounts except under
> > /share, and any cloned namespaces will still allow user mounts.
>
> I don't quite understand your method. This is how I think of it:
>
> mount --make-rshared /
> mkdir -p /mnt/ns/$USER
> mount --rbind / /mnt/ns/$USER
> mount --make-rslave /mnt/ns/$USER
> mount --set-flags --recursive -oallowusermnt /mnt/ns/$USER
> chroot /mnt/ns/$USER
> su - $USER
>
> I did actually try something equivalent (without the fancy mount
> commands though), and it worked fine. The only "problem" is the
> proliferation of mounts in /proc/mounts. There was a recently posted
> patch in AppArmor, that at least hides unreachable mounts from
> /proc/mounts, so the user wouldn't see all those. But it could still
> be pretty confusing to the sysadmin.

unbindable mounts were designed to overcome the proliferation problem.

Your steps should be something like this:

```
mount --make-rshared /
mkdir -p /mnt/ns
mount --bind /mnt/ns /mnt/ns
mount --make-unbindable /mnt/ns
mkdir -p /mnt/ns/$USER
mount --rbind / /mnt/ns/$USER
mount --make-rslave /mnt/ns/$USER
mount --set-flags --recursive -oallowusermnt /mnt/ns/$USER
chroot /mnt/ns/$USER
su - $USER
```

try this and your proliferation problem will disappear. :-)

>
> So in that sense doing it the complicated way, by first cloning the
> namespace, and then copying and sharing mounts individually which need
> to be shared could relieve this somewhat.

the unbindable mount will just provide you permanent relief.

>
> Another point: user mounts under /proc and /sys shouldn't be allowed.
> There are files there (at least in /proc) that are seemingly writable
> by the user, but they are still not writable in the sense, that
> "normal" files are.
>
> Anyway, there are lots of userspace policy issues, but those don't
> impact the kernel part.
>
> As for the original question of propagating the "allowusermnt" flag, I
> think it doesn't matter, as long as it's consistent and documented.
>
> Propagating some mount flags and not propagating others is
> inconsistent and confusing, so I wouldn't want that. Currently
> remount doesn't propagate mount flags, that may be a bug,

For consistency reason, one can propagate all the flags. But
propagating only those flags that interfere with shared-subtree
semantics should suffice.

wait...Dave's read-only bind mounts infact need the ability to
selectively make some mounts readonly. In such cases propagating
the read-only flag will just step on Dave's feature. Wont' it?

RP

>
> Miklos