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

Posted by [Miklos Szeredi](#) on Wed, 18 Apr 2007 14:03:23 GMT

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

> > > I've tried to make this unprivileged mount thing as simple as
> > > possible, and no simpler. If we can make it even simpler, all the
> > > better.
> > >
> > > We are certainly much more complex then the code in plan9 (just
> > > read through it) so I think we have room for improvement.
> > >
> > > Just for reference what I saw in plan 9 was:
> > > - No super user checks in it's mount, unmount, or namespace creation paths.
> > > - A flag to deny new mounts but not new bind mounts (for administrative purposes
> > > the comment said).
> > >
> > > Our differences from plan9.
> > > - suid capable binaries. (SUID please go away).
> > > - A history of programs assuming only root could call mount/unmount.
> >
> > I hate suid as well. _The_ motivation behind this patchset was to get
> > rid of "fusermount", a suid mount helper for fuse.
> >
> > But I don't think suid is going away, and definitely not overnight.
> > Also I don't think we want to require auditing userspace before
> > enabling user mounts.
> >
> > If I understand correctly, your proposal is to get rid of MNT_USER and
> > MNT_ALLOWUSERMNT and allow/deny unprivileged mounts and umounts based
> > on a boolean sysctl flag and on a check if the target namespace is the
> > initial namespace or not. And maybe add some extra checks which
> > prevent ugliness from happening with suid programs. Is this correct?
> >
> > If so, how are we going to make sure this won't break existing
> > userspace without doing a full audit of all suid programs in every
> > distro that wants this feature?
> >
> > Also how are we going to prevent the user from creating millions of
> > mounts, and using up all the kernel memory for vfsmounts?
>
> Don't forget that almost all mount flags are per-superblock. How are you
> planning on dealing with the case that one user mounts a filesystem
> read-only, while another is trying to mount the same one read-write?

Yeah, I forgot, the per-mount read-only patches are not yet in.

That doesn't really change my argument though. _If_ the flag is per

mount, then it makes sense to be able to change it on a master and not on a slave. If mount flags are propagated, this is not possible.

Miklos

Containers mailing list

Containers@lists.linux-foundation.org

<https://lists.linux-foundation.org/mailman/listinfo/containers>
