Subject: Re: [patch 0/8] unprivileged mount syscall Posted by Miklos Szeredi on Mon, 09 Apr 2007 16:24:10 GMT

View Forum Message <> Reply to Message

```
>>> One thing that is missing from this series is the ability to restrict
>>> user mounts to private namespaces. The reason is that private
>>> namespaces have still not gained the momentum and support needed for
>>> painless user experience. So such a feature would not yet get enough
>>> attention and testing. However adding such an optional restriction
>>> can be done with minimal changes in the future, once private
>>> namespaces have matured.
>>> I suspect the people who developed and maintain nsproxy would disagree;)
> >
> > Well, they better show me some working and simple-to-use userspace
> > code, because I've not seen anything like that related to mount
> > namespaces.
> If you mean to test/exploit them, see
> http://lxc.sourceforge.net/patches/2.6.20/2.6.20-lxc8/broken-out/tests/
> Compile the ns exec.c program and do
> ns_exec -m /bin/sh
> to get a shell in a new mounts namespace.
```

Cool, thanks. This is a very nice utility for testing, but for the end user rather useless:

- user starts up a private namespace in a shell, mounts something
- then opens app from menu, tries to access mount, but the mount is not there
- user unhappy

BTW, looking at -mm unshare() on namespace is not privileged any more. Why is that? Or rather, what's the reason, that clone() is privileged and unshare() is not?

- > pam_namespace.so is one example of a non-working, but probably-not-too-> > hard-to-fix one.
- > Non-working? I sure hope the one used for LSPP certification is
- > working... As is the ugly version I wrote 18 mounts ago and use on my
- > laptop.

The one in pam-0.99.6.3-29.1 in opensuse-10.2 is totally broken. Are you interested in the details? I can reproduce it, but forgot to note down the details of the brokenness.

- > > I'm just saying this is not yet something that Joe Blow would just
- > > enable by ticking a box in their desktop setup wizard, and it would
- > > all work flawlessly thereafter. There's still a _long_ way towards
- > > that, and mostly in userspace.

>

- > I'm not sure there's a that long a way to go, but clearly we need to be
- > showing users what they can do, or they'll never work their way towards
- > there.

There _is_ a long way to go. Random things that spring to my mind:

- using /etc/mtab is broken with private namespaces, using /proc/mounts is missing various functionality, that /etc/mtab has, for example the "user" option, which this patchset adds
- need to set up mount propagation from global namespace to private ones, mount(8) does not yet have options to configure propagation
- user namespace setup: what if user has multiple sessions?
 - 1) namespaces are shared? That's tricky because the session needs to be a child of a namespace server, not of login. I'm not sure PAM can handle this
 - 2) or mounts are copied on login? That's not possible currently, as there's no way to send a mount between namespaces. Also it's tricky to make sure that new mounts are also shared
- > For instance, as you say, a user admin gui with a checkmark and text
- > boxes saying 'enter new namespace on login', 'create private /tmp',
- > and 'create private dmcrypted /home' would be trivial right now.

Trivial modulo the above slightly non-trivial exemptions;)

Miklos

Containere mailing list

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