Subject: Re: [patch 0/8] unprivileged mount syscall
Posted by Ram Pai on Wed, 11 Apr 2007 18:28:36 GMT
View Forum Message <> Reply to Message

On Wed, 2007-04-11 at 12:44 +0200, Miklos Szeredi wrote:

```
> > 1. clone the master namespace.
> >
> > 2. in the new namespace
> >
> >  move the tree under /share/$me to /
> >        for each ($user, $what, $how) {
> >            move /share/$user/$what to /$what
> >      if ($how == slave) {
> >                 make the mount tree under /$what as slave
> >            }
> >        }
> >
> > 3. in the new namespace make the tree under
> >       /share as private and unmount /share
>
> Thanks.  I get the basic idea now: the namespace itself need not be
> shared between the sessions, it is enough if "share" propagation is
> set up between the different namespaces of a user.
>
> I don't yet see either in your or Viro's description how the trees
> under /share/$USER are initialized.  I guess they are recursively
> bound from /, and are made slaves.
```

yes. I suppose, when a userid is created one of the steps would be

```
mount --rbind / /share/$USER
mount --make-rslave /share/$USER
mount --make-rshared /share/$USER
```

RP

> Miklos

Subject: Re: [patch 0/8] unprivileged mount syscall
Posted by Miklos Szeredi on Fri, 13 Apr 2007 11:58:59 GMT
View Forum Message <> Reply to Message

```
> On Wed, 2007-04-11 at 12:44 +0200, Miklos Szeredi wrote:
> > > 1. clone the master namespace.
> > > 
> > > 2. in the new namespace
> > > 
> > >  move the tree under /share/$me to /
> > >       for each ($user, $what, $how) {
> > >           move /share/$user/$what to /$what
> > >     if ($how == slave) {
> > >                make the mount tree under /$what as slave
> > >           }
> > >       }
> > > 
> > > 3. in the new namespace make the tree under
> > >      /share as private and unmount /share
> > 
> > Thanks.  I get the basic idea now: the namespace itself need not be
> > shared between the sessions, it is enough if "share" propagation is
> > set up between the different namespaces of a user.
> > 
> > I don't yet see either in your or Viro's description how the trees
> > under /share/$USER are initialized.  I guess they are recursively
> > bound from /, and are made slaves.
> 
> yes. I suppose, when a userid is created one of the steps would be
> 
> mount --rbind / /share/$USER
> mount --make-rslave /share/$USER
> mount --make-rshared /share/$USER
```

Thinking a bit more about this, I'm quite sure most users wouldn't even want private namespaces. It would be enough to

```
  chroot /share/$USER
```

and be done with it.

Private namespaces are only good for keeping a bunch of mounts referenced by a group of processes. But my guess is, that the natural behavior for users is to see a persistent set of mounts.

If for example they mount something on a remote machine, then log out from the ssh session and later log back in, they would want to see their previous mount still there.

Miklos

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

Subject: Re: [patch 0/8] unprivileged mount syscall
Posted by Karel Zak on Fri, 13 Apr 2007 20:07:20 GMT

View Forum Message <> Reply to Message

```
On Fri, Apr 13, 2007 at 01:58:59PM +0200, Miklos Szeredi wrote:
> > On Wed, 2007-04-11 at 12:44 +0200, Miklos Szeredi wrote:
> > > > 1. clone the master namespace.
> > > >
> > > > 2. in the new namespace
> > > >
> > > >  move the tree under /share/$me to /
> > > >       for each ($user, $what, $how) {
> > > >           move /share/$user/$what to /$what
> > > >     if ($how == slave) {
> > > >                make the mount tree under /$what as slave
> > > >           }
> > > >       }
> > > >
> > > > 3. in the new namespace make the tree under
> > > >       /share as private and unmount /share
> > > 
> > > Thanks.  I get the basic idea now: the namespace itself need not be
> > > shared between the sessions, it is enough if "share" propagation is
> > > set up between the different namespaces of a user.
> > > 
> > > I don't yet see either in your or Viro's description how the trees
> > > under /share/$USER are initialized.  I guess they are recursively
> > > bound from /, and are made slaves.
> > 
> > yes. I suppose, when a userid is created one of the steps would be
> > 
> > mount --rbind / /share/$USER
> > mount --make-rslave /share/$USER
> > mount --make-rshared /share/$USER
> 
> Thinking a bit more about this, I'm quite sure most users wouldn't
> even want private namespaces.  It would be enough to
> 
>   chroot /share/$USER
> 
> and be done with it.
```

I don't think so. How to you want to implement non-shared /tmp directories? The chroot is overkill in this case. See:

http://www.coker.com.au/selinux/talks/sage-2006/PolyInstantiatedDirectories.html
http://danwalsh.livejournal.com/

> Private namespaces are only good for keeping a bunch of mounts
> referenced by a group of processes. But my guess is, that the natural
> behavior for users is to see a persistent set of mounts.
>
> If for example they mount something on a remote machine, then log out
> from the ssh session and later log back in, they would want to see
> their previous mount still there.

They can mount to /mnt where the directory is shared ("mount --make-shared /mnt") and visible and all namespaces.

I think /share/$USER is an extreme example. You can found more situations when private namespaces are nice solution.

Karel

--
Karel Zak <kzak@redhat.com>

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

---

Subject: Re: [patch 0/8] unprivileged mount syscall
Posted by Miklos Szeredi on Sun, 15 Apr 2007 20:21:05 GMT

View Forum Message <> Reply to Message

> > Thinking a bit more about this, I'm quite sure most users wouldn't
> > even want private namespaces. It would be enough to
> >
> > chroot /share/$USER
> >
> > and be done with it.
>
> I don't think so. How to you want to implement non-shared /tmp
> directories?

```
 mount --bind /.tmp/$USER /share/$USER/tmp
```

or whatever else this polyunsaturated thingy does within the cloned namespace.

> The chroot is overkill in this case.

What do you mean it's an overkill? clone(CLONE_NS) duplicates all the
mounts, just as mount --rbind does.

```
> > Private namespaces are only good for keeping a bunch of mounts
> > referenced by a group of processes.  But my guess is, that the natural
> > behavior for users is to see a persistent set of mounts.
> >
> > If for example they mount something on a remote machine, then log out
> > from the ssh session and later log back in, they would want to see
> > their previous mount still there.
>
>  They can mount to /mnt where the directory is shared ("mount
>  --make-shared /mnt") and visible and all namespaces.
>
>  I think /share/$USER is an extreme example. You can found more
>  situations when private namespaces are nice solution.
```

Private to a single login session? I'd like to hear examples.

Thanks,
Miklos

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

Subject: Re: [patch 0/8] unprivileged mount syscall
Posted by Ram Pai on Mon, 16 Apr 2007 07:59:06 GMT
View Forum Message <> Reply to Message

On Fri, 2007-04-13 at 13:58 +0200, Miklos Szeredi wrote:

```
> > On Wed, 2007-04-11 at 12:44 +0200, Miklos Szeredi wrote:
> > > > 1. clone the master namespace.
> > > >
> > > > 2. in the new namespace
> > > >
> > > >  move the tree under /share/$me to /
> > > >        for each ($user, $what, $how) {
> > > >            move /share/$user/$what to /$what
> > > >     if ($how == slave) {
> > > >                 make the mount tree under /$what as slave
> > > >            }
> > > >        }
```

> > > >
> > > > 3. in the new namespace make the tree under
> > > >   /share as private and unmount /share
> > >
> > > Thanks. I get the basic idea now: the namespace itself need not be
> > > shared between the sessions, it is enough if "share" propagation is
> > > set up between the different namespaces of a user.
> > >
> > > I don't yet see either in your or Viro's description how the trees
> > > under /share/$USER are initialized. I guess they are recursively
> > > bound from /, and are made slaves.
> >
> > yes. I suppose, when a userid is created one of the steps would be
> >
> > mount --rbind / /share/$USER
> > mount --make-rslave /share/$USER
> > mount --make-rshared /share/$USER
>
> Thinking a bit more about this, I'm quite sure most users wouldn't
> even want private namespaces. It would be enough to
>
>   chroot /share/$USER
>
> and be done with it.
>
> Private namespaces are only good for keeping a bunch of mounts
> referenced by a group of processes. But my guess is, that the natural
> behavior for users is to see a persistent set of mounts.
>
> If for example they mount something on a remote machine, then log out
> from the ssh session and later log back in, they would want to see
> their previous mount still there.

They will continue see their previous mount tree.
Even if all the namespaces belonging to the different sessions of the
user get dismantled when all the sessions exit, the a mirror of those
mount trees continue to exist under /share/$USER in the original
namespace. So I don't think we have a issue.

NOTE: when I say 'original namespace' I mean the admin namespace; the
first namespace that gets created when the machine boots.

RP

>
> Miklos