Subject: Re: [patch 05/10] add "permit user mounts in new namespace" clone flag
Posted by Herbert Poetzl on Fri, 13 Apr 2007 04:16:12 GMT
View Forum Message <> Reply to Message

On Thu, Apr 12, 2007 at 03:32:08PM -0500, Serge E. Hallyn wrote:
> Quoting Miklos Szeredi (miklos@szeredi.hu):
> > From: Miklos Szeredi <mszeredi@suse.cz>
> >
> > If CLONE_NEWNS and CLONE_NEWNS_USERMNT are given to clone(2) or
> > unshare(2), then allow user mounts within the new namespace.

> > This is not flexible enough, because user mounts can't be enabled for
> > the initial namespace.
> >
> > The remaining clone bits also getting dangerously few...

ATM I think we do not have that many CLONE flags
available, so that this feature will have to wait
for a clone2/64 or similar ...

> > Alternatives are:
> >
> > - prctl() flag
> > - setting through the containers filesystem

> Sorry, I know I had mentioned it, but this is definately my least
> favorite approach.
>
> Curious whether are any other suggestions/opinions from the containers
> list?

question: how is mounting filesystems (loopback,
fuse, etc) secured in such way that the user
cannot 'create' device nodes with 'unfortunate'
permissions?

TIA,
Herbert

> thanks,
> -serge
>
> > Signed-off-by: Miklos Szeredi <mszeredi@suse.cz>
> > ---
> >
> > Index: linux/fs/namespace.c
> > ===
> > --- linux.orig/fs/namespace.c 2007-04-12 13:46:19.000000000 +0200

> > +++ linux/fs/namespace.c 2007-04-12 13:54:36.000000000 +0200
> > @@ -1617,6 +1617,8 @@ struct mnt_namespace *copy_mnt_ns(int fl
> > return ns;
> >
> > new_ns = dup_mnt_ns(ns, new_fs);
> > + if (new_ns && (flags & CLONE_NEWNS_USERMNT))
> > + new_ns->flags |= MNT_NS_PERMIT_USERMOUNTS;
> >
> > put_mnt_ns(ns);
> > return new_ns;
> > Index: linux/include/linux/sched.h
> > ===
> > --- linux.orig/include/linux/sched.h 2007-04-12 13:26:48.000000000 +0200
> > +++ linux/include/linux/sched.h 2007-04-12 13:54:36.000000000 +0200
> > @@ -26,6 +26,7 @@
> > #define CLONE_STOPPED 0x02000000 /* Start in stopped state */
> > #define CLONE_NEWUTS 0x04000000 /* New utsname group? */
> > #define CLONE_NEWIPC 0x08000000 /* New ipcs */
> > +#define CLONE_NEWNS_USERMNT 0x10000000 /* Allow user mounts in ns? */
> >
> > /*
> > * Scheduling policies
> > Index: linux/kernel/fork.c
> > ===
> > --- linux.orig/kernel/fork.c 2007-04-11 18:27:46.000000000 +0200
> > +++ linux/kernel/fork.c 2007-04-12 13:59:10.000000000 +0200
> > @@ -1586,7 +1586,7 @@ asmlinkage long sys_unshare(unsigned lon
> > err = -EINVAL;
> > if (unshare_flags & ~(CLONE_THREAD|CLONE_FS|CLONE_NEWNS|CLONE_SIGHAND|
> > CLONE_VM|CLONE_FILES|CLONE_SYSVSEM|
> > - CLONE_NEWUTS|CLONE_NEWIPC))
> > + CLONE_NEWUTS|CLONE_NEWIPC|CLONE_NEWNS_USERMNT))
> > goto bad_unshare_out;
> >
> > if ((err = unshare_thread(unshare_flags)))
> >
> > --
> > -
> > 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
> _____
> Containers mailing list
> Containers@lists.linux-foundation.org
> https://lists.linux-foundation.org/mailman/listinfo/containers

_____
Containers mailing list
Containers@lists.linux-foundation.org