Subject: Re: [patch 7/9] unprivileged mounts: allow unprivileged fuse mounts
Posted by Miklos Szeredi on Wed, 09 Jan 2008 13:16:21 GMT
View Forum Message <> Reply to Message

> ...this will break with FUSE enabled, right? (Minor security hole by
> allowing users to stop c-a-delete, where none existed before?)

Yup (or I don't know, I'm sure there was or is some problem with
ptrace, that could be used to create unkillable processes).

Fuse could actually be fixed to exit reliably for 'killall5 -9' (it
used to), but that has other problems, and it doesn't seem very
important to me. But this can be discussed.

What cannot be fixed is if one process is inside an fs operation
(e.g. unlink), holding a VFS lock (i_mutex) and another process goes
to uninterruptible sleep on that lock. There's no way (other than
rewriting the VFS) in which that second process could be killed unless
you kill the first one or the fuse server.

> I'm currently suspending by 'echo "mem" > /sys/power/state'. How
> should I do that _safely_ with FUSE enabled?

You can't. But that's only solvable with

- rewrite of VFS (see above)
- rewrite of freezer

> If I want to get rid of nasty user in multiuser system, I do
> su nastyuser 'kill -9 -1' . How do I do the equivalent with FUSE
> enabled? (Without affecting other users?)

You can still do that. If a process cannot be killed with 'kill -9',
due to being deadlocked with itself through fuse (not an easy feat to
accomplish), then it's not going to do any more harm, and you _can_
get rid of it by forced umounting the filesystem, or if it has been
detached, through the fusectl filesystem.

> Load average was never really meaningful number, but with FUSE
> enabled, users can set it to 666 without actually eating any CPU.
>
> SIGSTOP used to work, allowing you to prevent user processes from
> working while you examine them. Now SIGSTOP can be delayed for
> arbitrary time.

Making filesystem operations restartable is not easy. I would say
near impossible, but I haven't given a lot of enery into investigating.

> Heck, imagine malicious user process misbehaves. Before FUSE, you
> could at least attach it with gdb to look what it is doing. Now you
> can't.

Sure, but you can check in other ways (/proc/$PID/wchan), sysrq-t.

> I really believe FUSE vs. signals needs fixing. Either that, or
> updating all the manpages
>
> man 1 kill:
> - KILL 9 exit this signal may not be blocked
> + KILL 9 exit this signal may not be blocked, except by FUSE user mount

Heh, there are all very interesting, but most of these issues are not even on my todo list (which has grown into quite a big pile over the years), which means, that they don't seem to matter to people in practice.

You seem to be implying that fuse is worthless if these issues are not fixed, but that is very far from the truth, I think.

Miklos

_____

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