

---

Subject: Re: [RFC][PATCH] Pid namespaces vs locks interaction

Posted by [gblond](#) on Wed, 12 Dec 2007 17:42:41 GMT

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

---

On 12 December 2007 20:31:15 Serge E. Hallyn wrote:

> Quoting Vitaliy Gusev (vgusev@openvz.org):

> > Hello

> >

> > On 6 December 2007 18:51:30 Serge E. Hallyn wrote:

> > > fl\_pid is used by nfs, fuse and gfs2. For instance nfs keeps in

> > > fl\_pid some unique id to identify locking process between hosts - it

> > > is not a process pid.

> > >

> > > Ok, but so the struct user\_flock->fl\_pid is being set to the task's

> > > virtual pid, while the struct kernel\_flock->fl\_pid is being set to

> > > task->tgid for nfsd use.

> > >

> > > Why can't nfs just generate a uniqueid from the struct pid when it

> > > needs it?

> >

> > I think it is hard. lockd uses struct nlm\_host to get process unique id

> > (see \_\_nlm\_alloc\_pid() function).

>

> Looks pretty simple though... That whole set of code could even stay

> the same except for in \_\_nlm\_alloc\_pid():

>

> option 1: compare struct pid\* instead of uint32\_t pid

> option 2: use the "global pid" out of the stored struct pid,

> something like pid->numbers[0].nr.

We can't use process pid. Process pid is circulated! NFS (lockd) needs unique process id between hosts which can't repeat oneself.

>

> > > Fuse just seems to copy the pid to report it to userspace, so it would

> > > just copy pid\_vnr(kernel\_flock->pid) into user\_flock->fl\_pid.

> > >

> > > Anyway I haven't looked at all the uses of struct fl\_pid, but you

> > > can always get the pidnr back from the struct pid if needed so there

> > > should be no problem.

> > >

> > > The split definately seems worthwhile to me, so that

> > > user\_flock->fl\_pidnr can always be said to be the pid in the acting

> > > process' namespace, and flock->fl\_pid can always be a struct pid,

> > > rather than having fl\_pid sometimes be current->tgid, or sometimes

> > > pid\_vnr(flock->fl\_nspid)...

> > >

> > > -serge

> > > -  
> > > 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>  
> >  
> > --  
> > Thank,  
> > Vitaliy Gusev  
>  
> -  
> 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>

--  
Thank,  
Vitaliy Gusev

---

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

---