

---

Subject: Re: [RFC PATCH 0/5] Resend - Use procs to change a syscall behavior  
Posted by [Pavel Machek](#) on Tue, 08 Jul 2008 21:58:21 GMT

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

---

> > > > An alternative to this solution consists in defining a new field in the  
> > > > task structure (let's call it next\_syscall\_data) that, if set, would change  
> > > > the behavior of next syscall to be called. The sys\_fork\_with\_id() previously  
> > > > cited can be replaced by  
> > > > 1) set next\_syscall\_data to a target upid nr  
> > > > 2) call fork().  
> > > >  
> > > > ...bloat task struct and  
> > > >  
> > > > A new file is created in procs: /proc/self/task/<my\_tid>/next\_syscall\_data.  
> > > > This makes it possible to avoid races between several threads belonging to  
> > > > the same process.  
> > > >  
> > > > ...introducing this kind of ugliness.  
> > > >  
> > > > Actually, there were proposals for sys\_indirect(), which is slightly  
> > > > less ugly, but IIRC we ended up with adding syscalls, too.  
> > >  
> > > Silly question...  
> > >  
> > > Oren, would you object to defining sys\_fork\_with\_id(),  
> > > sys\_msgget\_with\_id(), and sys\_semget\_with\_id()?  
> > >  
> > > Eric, Pavel (Emelyanov), Dave, do you have preferences?  
> > >  
> > > For the cases Nadia has implemented here I'd be tempted to side with  
> > > Pavel Machek, but once we get to things like open() and socket(), (a)  
> > > the # new syscalls starts to jump, and (b) the per-syscall api starts to  
> > > seem a lot more cumbersome.  
> >  
> > You should not need to modify open/socket. You can already select fd  
> > by creatively using open/dup/close...  
>  
> That's what we do right now in cryo. And if we end up patching up every  
> API with separate syscalls, then we wouldn't create open\_with\_id(). But  
> so long as the next\_id were to exist, exploiting it in open is nigh on  
> trivial and much nicer.

Ok, so ignore previous email. You know how unix works.

I believe you should just introduce syscalls you need. Yes,  
introducing new syscalls is hard/expensive, but changing existing  
syscalls is simply bad idea.

So what new syscalls do you really need? Not open\_this\_fd, nor socket\_this\_fd.

Pavel

--

(english) <http://www.livejournal.com/~pavelmachek>

(cesky, pictures) <http://atrey.karlin.mff.cuni.cz/~pavel/picture/horses/blog.html>

---

Containers mailing list

Containers@lists.linux-foundation.org

<https://lists.linux-foundation.org/mailman/listinfo/containers>

---