
Subject: Re: [PATCH] ptrace: it is fun to strace /sbin/init
Posted by [Herbert Poetzl](#) on Tue, 25 Mar 2008 19:07:43 GMT
[View Forum Message](#) <> [Reply to Message](#)

On Tue, Mar 25, 2008 at 01:06:11PM -0500, Serge E. Hallyn wrote:

> Quoting Stephen Smalley (sds@tycho.nsa.gov):

> >

> > On Tue, 2008-03-25 at 08:40 -0500, Serge E. Hallyn wrote:

> > > Quoting Stephen Smalley (sds@tycho.nsa.gov):

> > > >

> > > > On Tue, 2008-03-25 at 02:04 +0300, Oleg Nesterov wrote:

> > > > > On 03/24, Pavel Machek wrote:

> > > > > >

> > > > > > /sbin/init is important, but there are other important (and sometimes
> > > > > > much more important) services. Why it is so special so that we can't
> > > > > > debug/strace it?

> > > > > >

> > > > > > Maybe. Let's kill /sbin/init protection in 2.6.26. But making it
> > > > > > optional is wrong.

> > > > > >

> > > > > > You are right, the boot parameter is silly. How about sysctl?

> > > > > >

> > > > > > Stephen, do you see any security problems if we make /sbin/init
> > > > > > ptraceable by default?

> > > > > >

> > > > Not an issue for SELinux (we apply an orthogonal check based on security
> > > > context, so we can already block ptrace of init independent of whether
> > > > root/CAP_SYS_PTRACE can do it). I'm not sure though as to whether
> > > > people using capabilities have ever relied on this special protection of
> > > > init (e.g. custom init spawns children with lesser capabilities and
> > > > relies on the fact that they cannot ptrace init to effectively re-gain
> > > > those capabilities, even if they possess CAP_SYS_PTRACE).

> > >

> > > Still thinking it through, but it seems like special casing init isn't
> > > useful. There are likely to be other tasks with all capabilities
> > > set which the malicious task could just as well ptrace to do his
> > > mischief, right?

> >

> > Depends on the bounding set. Didn't it used to be the case that only
> > init had CAP_SETPCAP (until the meaning of it was changed by the
> > filesystem capability support)?

>

> Not quite. CAP_SETPCAP was taken out of everyone's bounding set. But
> kernel/sysctl.c allowed only init to add capabilities to the bounding
> set. (Whereas CAP_SYS_MODULE was sufficient to remove them).

>

> > Might want to double check with e.g. the vservers folks that they
> > weren't relying in any way on special handling of init.

>
> Herbert, Pavel, do you have objections to allowing ptrace of init?
> (I believe Eric has already Acked the idea iirc?)

inside a guest, by default no (i.e. there simply is
no capability for that), on the host the behaviour is
unmodified .. note that there are guests without init
where the blend through init is protected in a special
way

HTH,
Herbert

> thanks,
> -serge

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