
Subject: Re: [RFC][PATCH] Add child reaper to struct pspace

Posted by [serue](#) on Fri, 08 Sep 2006 14:36:08 GMT

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

Quoting Eric W. Biederman (ebiederm@xmission.com):

> "Serge E. Hallyn" <serue@us.ibm.com> writes:

>

> > Quoting Eric W. Biederman (ebiederm@xmission.com):

> >> Yes. We all so need something like that list to support kill -1.

> >> Although walking the list of all processes may be sufficient for a first

> >> pass.

> >>

> >> The real trick is handing nested pid namespaces, properly.

> >

> > Not if, as you've suggested in the past, pid_ns 5 has valid pids in its

> > own pid_ns for every process in pid_namespaces nested under it.

> >

> > It should be simple to implement, should not impact the non-container

> > cases, and should only start to impact performance as the nesting gets

> > deep, which AFAIK we all believe won't happen (max nesting of 2 AFAICS,

> > one checkpointable application container under one vserver-thingie)

> >

> > And it makes kill -1 trivial, as in pid_ns 5 we just kill all processes

> > in pid_ns 5, without worrying about finding the ones in it's decendent

> > pid namespaces.

>

> If you do it correctly I agree. But you have to be very careful where

> you put the list.

>

> My point being not that we can't get this correct with simple code, but

> that it is easy to get it wrong.

True.

I should think the list has to go into the struct pid, not the task_struct. There is one struct pid per (pid_ns, pid), so we can keep just one simple list to walk the pid_ns processes. If we were to put it in the task_struct, well there's really no clean way to go about it then :)

thanks,

-serge

Containers mailing list

Containers@lists.osdl.org

<https://lists.osdl.org/mailman/listinfo/containers>
