
Subject: Re: [RFC][PATCH] Do not set /proc inode->pid for non-pid-related inodes
Posted by [serue](#) on Thu, 22 Mar 2007 14:33:50 GMT

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

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

...

> Back to the main subject I still don't understand the idea of running
> a kernel daemon as pid == 1. What would that buy us?

I think the idea is that for lightweight application containers, where there is no explicit /sbin/init process, the kthread would act as reaper for the pid_ns so that the first userspace process could freely exit while other processes continued.

I still prefer that we forego that kthread, and just work toward allowing pid1 to exit. Really I think the crufty /proc/<pid> handling is the only reason we were going to punt on that for now. So for our first stab I think we should have pid=1 exiting cause all other processes in the same pid_ns to be killed. Then when we get /proc fixed up, we can change the semantics so that pid=1 exiting just switches the pid_namespace's reaper to either the parent of the killed pid=1, or to the global init.

-serge

Containers mailing list

Containers@lists.linux-foundation.org

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