Subject: Re: [PATCH] namespaces: fix race at task exit
Posted by ebiederm on Thu, 25 Jan 2007 16:29:13 GMT
View Forum Message <> Reply to Message

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

> In do_exit(), the exit_task_namespaces() was placed after
> exit_notify() because exit_notify ends up using the pid
> namespace both to access the reaper, and for detaching the
> pid. However, this placement allows an nfs server to reap
> the task before exit_task_namespaces() completes.
>
> This patch moves the exit_task_namespaces() into release_task,
> below release_thread() which puts the pids(), and just above
> the call_rcu(delayed_put_task_struct). I believe this should
> solve both problems.

For the pid namespace this seems to be correct placement.
For the mount namespace this would seem to exacerbate the problem
because it now gets called after the task has been reaped!

I'd love to be convinced otherwise but I do not believe we
can safely exit both the mount and the pid namespace at the
same location in the code.

The NFS unmount currently wants a killable thread as it
uses interruptible sleeps. How does starting that process
after the process in which it lives aid this?

But thanks for remembering this. This is a real problem we
do need to solve.

Eric

_____
Containers mailing list
Containers@lists.osdl.org
https://lists.osdl.org/mailman/listinfo/containers