Subject: Re: [PATCH] namespaces: fix race at task exit
Posted by serue on Thu, 25 Jan 2007 17:36:55 GMT
View Forum Message <> Reply to Message

Quoting Oleg Nesterov (oleg@tv-sign.ru):
> On 01/25, Serge E. Hallyn wrote:
> >
> > 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.
> >
> > Signed-off-by: Serge E. Hallyn <serue@us.ibm.com>
> >
> > ---
> >
> > kernel/exit.c | 2 +-
> > 1 files changed, 1 insertions(+), 1 deletions(-)
> >
> > 765277a4170d7bbd1c4613de661ec6ac64d5580a
> > diff --git a/kernel/exit.c b/kernel/exit.c
> > index 3540172..ab9ae30 100644
> > --- a/kernel/exit.c
> > +++ b/kernel/exit.c
> > @@ -174,6 +174,7 @@ repeat:
> >   write_unlock_irq(&tasklist_lock);
> >   proc_flush_task(p);
> >   release_thread(p);
> > + exit_task_namespaces(p);
> >   call_rcu(&p->rcu, delayed_put_task_struct);
>
> Probably I missed some other patches in this area, but I can't understand
> this fix.
>
> With this change we are doing __put_mnt_ns() when we surely don't have ->sighand,
> no? Could you please explain?

Explanation: it's wrong :)

we'll just need to break exit_task_namespaces() up.

thanks,

-serge

_____

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