Subject: Re: [PATCH] cgroup: fix a race condition in manipulating tsk->cg_list
Posted by Paul Menage on Thu, 17 Apr 2008 04:28:33 GMT
View Forum Message <> Reply to Message

On Wed, Apr 16, 2008 at 9:18 PM, Paul Menage <menage@google.com> wrote:
>
> My only thought about the downside of this is that an exiting task
> that gets stuck somewhere between setting PF_EXITING and calling
> cgroup_exit() won't show up in its cgroup's tasks file, since we'll
> enable cgroup links but skip it. I guess that's not a big deal.
>

How about this as an alternative approach? We can take advantage of the indirection in tsk->cgroups to create an additional distinguished css_set that indicates the task has passed the point of checking tsk->cg_list:

- create a new css_set, called exit_css_set; it has the same cgroup pointer set as init_css_set.

- in cgroup_exit(), set current->cgroups to &exit_css_set rather than &init_css_set

- in cgroup_enable_task_cg_list(), ignore any task where p->cgroups == &exit_css_set

That way we're synchronizing directly with the task_lock()-protected section in cgroup_exit(), rather than with the setting of PF_EXITING at the beginning of do_exit().

Paul