
Subject: Re: [PATCH 3/7] containers (V7): Add generic multi-subsystem API to containers

Posted by [Paul Menage](#) on Thu, 15 Feb 2007 01:17:09 GMT

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

On 2/13/07, Srivatsa Vaddagiri <vatsa@in.ibm.com> wrote:

>
> Well, we already bump up reference count in fork() w/o grabbing those
> mutexes don't we? Also if rmdir() sees container->count to be zero, then
> it means no task is attached to the container. How will then a function
> like bc_file_charge() bump up the reference count to such a container
> (presuming it wanted to do so w/o manage/callback mutexes -and- that the
> container pointer in bc_file_charge is derived from some task in
> that container). I think it is safe to bump up container->count in
> bc_file_charge w/o grabbing manage/callback mutexes.

Right, I was never suggesting that we take either of those mutexes for this operation. The spin lock in css_get() was an attempt to avoid that. But I think you're right that it was too heavyweight, and can be avoided with atomic operations. See my other email to Pavel.

>
> Are you talking about (un)bind of subsystem to/from hierararchies that
> have non-zero containers in them? That sounds very icky. Anyway that
> doesnt seem to be supported in current patches.

The bind/unbind from active hierarchies is supported in the user-space API, and it's implemented for hierarchies that have no child containers. Hence it's important, at least conceptually, for the reference count to be held by the subsystem state rather than the container.

Implementing a full bind/unbind for arbitrary subsystems and hierarchies will indeed be a lot of work, which is why I'm not trying to do it at this point.

Paul
