Subject: Re: [PATCH 3/7] containers (V7): Add generic multi-subsystem API to containers
Posted by Srivatsa Vaddagiri on Thu, 08 Mar 2007 17:46:04 GMT
View Forum Message <> Reply to Message

On Mon, Feb 12, 2007 at 12:15:24AM -0800, menage@google.com wrote:
> +static struct container_group *find_container_group(
> + struct container_group *oldcg, struct container *cont)
> +{
> + struct container_group *res;
> + struct container_subsys *ss;
> + int h = cont->hierarchy;
> + int i;
> +
> + BUG_ON(oldcg->container[h] == cont);
> + /* First see if we already have a container group that matches
> + * the desired set */
> + spin_lock(&container_group_lock);

I think this should be spin_lock_irq().

I have been bitten by lockdep warning in the rcfs patchs. But I think they should apply here as well.

Essentially, container_task_count() is called from cpuset.c after doing write_lock_irq(tasklist_lock). This creates the dependency:

tasklist_lock (irqs disabled) -> container_group_lock

However in above function, find_container_group (and several other functions) we take container_group_lock w/o blocking interrupts. An interrupt handler now trying to take tasklist_lock will cause a deadlock.

--
Regards,
vatsa