
Subject: Re: containers (was Re: -mm merge plans for 2.6.23)

Posted by [Peter Zijlstra](#) on Wed, 11 Jul 2007 11:30:40 GMT

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

On Wed, 2007-07-11 at 13:24 +0200, Peter Zijlstra wrote:

> On Wed, 2007-07-11 at 04:10 -0700, Paul Jackson wrote:

> > Srivatsa wrote:

> > > So Ingo was proposing we use cpuset as that user interface to manage
> > > task-groups. This will be only for 2.6.23.

> >

> > Good explanation - thanks.

> >

> > In short, the proposal was to use the task partition defined by cpusets
> > to define CFS task-groups, until the real process containers are
> > available.

> >

> > Or, I see in the next message, Ingo responding favorably to your
> > alternative, using task uid's to partition the tasks into CFS
> > task-groups.

> >

> > Yeah, Ingo's preference for using uid's (or gid's ??) sounds right to
> > me - a sustainable API.

> >

> > Wouldn't want to be adding a cpuset API for a single 2.6.N release.

> >

> > gid's -- why not?

>

>

> Or process or process groups, or all of the above :-)

>

> One thing to think on though, we cannot have per process,uid,gid,pgrp
> scheduling for one release only. So we'd have to manage interaction with
> process containers. It might be that a simple weight multiplication
> scheme is good enough:

>

> weight = uid_weight * pgp_weight * container_weight

>

> Of course, if we'd only have a single level group scheduler (as was
> proposed IIRC) it'd have to create intersection sets (as there might be
> non trivial overlaps) based on these various weights and schedule these
> resulting sets instead of the initial groupings.

Lets illustrate with some ASCII art:

so we have this dual level weight grouping (uid, container)

uid: a a a a a b b b b c c c c

container: A A A A A A B B B B B B B

set: 1 1 1 1 1 2 2 3 3 3 4 4 4 4 4

resulting in schedule sets 1,2,3,4

so that (for instance) weight_2 = weight_b * weight_A

Containers mailing list

Containers@lists.linux-foundation.org

<https://lists.linux-foundation.org/mailman/listinfo/containers>
