Subject: Re: [PATCH] Hookup group-scheduler with task container infrastructure
Posted by Srivatsa Vaddagiri on Fri, 14 Sep 2007 15:51:13 GMT
View Forum Message <> Reply to Message

On Fri, Sep 14, 2007 at 11:41:58AM +0200, Ingo Molnar wrote:
> * Srivatsa Vaddagiri <vatsa@linux.vnet.ibm.com> wrote:
> > Add interface to control cpu bandwidth allocation to task-groups.
>
> btw., just in case it was not obvious, i'll repeat my older assessment
> of your patch: the general picture approach looks good to me and the
> code is upstream-worthy.

Thanks for the feedback!

> (suggestion: if you want more people to test it then you might want to
> do some add-on "put all users into separate groups" patch and .config
> option - which could be tried without people having to know anything
> about container setup.)

I do want more people to test it and I agree that hooking onto
user-id based groups is the best way to get that done. How do we
implement that? I have two choices:

1. Do a kernel patch, as you suggest above, which defines task-groups
   based on user-id and hook that group definition with group scheduler.
   We need to provide some means for the admin to tune relative
   nice-value of each user (perhaps thr' sysctl?).

   This user-id based grouping will have to be mutually exclusive with
   task-container based grouping. Hence we need to ensure that only one
   form of grouping is selected and not both at compile time.

2. Enable only one form of grouping, which is task-container based. Provide a
   user-space daemon (attached) which can automatically put tasks of different
   users in different task-containers. The daemon will need to be started at
   early boot-time. It can also be extended to support a configuration file
   (ex: inittab) where cpu allocation for different users are specified.
   The fact that daemon is managing to provide fair allocation to users
   should be transparent.

   I hope that task-containers (aka cgroups) will go into 2.6.24, in which case
   the second option seems to be more attractive to me.

I will neverthless try to work out Option 1, just to see how it looks.

--
Regards,
vatsa

Containers mailing list
Containers@lists.linux-foundation.org
https://lists.linux-foundation.org/mailman/listinfo/containers