Subject: Re: [ckrm-tech] [PATCH 0/2] resource control file system - aka containers on top of nsproxy!
Posted by serue on Mon, 12 Mar 2007 15:56:43 GMT
View Forum Message <> Reply to Message

Quoting Srivatsa Vaddagiri (vatsa@in.ibm.com):
> On Fri, Mar 09, 2007 at 02:09:35PM -0800, Paul Menage wrote:
> > > 3. This next leads me to think that 'tasks' file in each directory doesnt make
> > > sense for containers. In fact it can lend itself to error situations (by
> > > administrator/script mistake) when some tasks of a container are in one
> > > resource class while others are in a different class.
> > >
> > > Instead, from a containers pov, it may be usefull to write
> > > a 'container id' (if such a thing exists) into the tasks file
> > > which will move all the tasks of the container into
> > > the new resource class. This is the same requirement we
> > > discussed long back of moving all threads of a process into new
> > > resource class.
> >
> > I think you need to give a more concrete example and use case of what
> > you're trying to propose here. I don't really see what advantage
> > you're getting.
>
> Ok, this is what I had in mind:
>
>
> mount -t container -o ns /dev/namespace
> mount -t container -o cpu /dev/cpu
>
> Lets we have the namespaces/resource-groups created as under:
>
> /dev/namespace
> |-- prof
> | |- tasks <- (T1, T2)
> | |- container_id <- 1 (doesnt exist today perhaps)
> |
> |-- student
> | |- tasks <- (T3, T4)
> | |- container_id <- 2 (doesnt exist today perhaps)
>
> /dev/cpu
> |-- prof
> | |-- tasks
> | |-- cpu_limit (40%)
> |
> |-- student
> | |-- tasks
> | |-- cpu_limit (20%)

> |
> |
>
>
> Is it possible to create the above structure in container patches?
> /me thinks so.
>
> If so, then accidentally someone can do this:
>
> echo T1 > /dev/cpu/prof/tasks
> echo T2 > /dev/cpu/student/tasks
>
> with the result that tasks of the same container are now in different
> resource classes.

What's wrong with that?

> Thats why in case of containers I felt we shldnt allow individual tasks
> to be cat'ed to tasks file.
>
> Or rather, it may be nice to say :
>
> echo "cid 2" > /dev/cpu/prof/tasks
>
> and have all tasks belonging to container id 2 move to the new resource
> group.

Adding that feature sounds fine, but don't go stopping me from putting
T1 into /dev/cpu/prof/tasks and T2 into /dev/cpu/student/tasks just
because you have your own notion of what each task is supposed to be.
Just because they're in the same namespaces doesn't mean they should get
the same resource allocations. If you want to add that kind of policy,
well, it should be policy - user-definable.

-serge

_____
Containers mailing list
Containers@lists.osdl.org
https://lists.osdl.org/mailman/listinfo/containers

Subject: Re: [ckrm-tech] [PATCH 0/2] resource control file system - aka containers on top of nsproxy!
Posted by Srivatsa Vaddagiri on Mon, 12 Mar 2007 16:20:45 GMT
View Forum Message <> Reply to Message

On Mon, Mar 12, 2007 at 10:56:43AM -0500, Serge E. Hallyn wrote:
> What's wrong with that?

I had been asking around on "what is the fundamental unit of res mgmt
for vservers" and the answer I got (from Herbert) was "all tasks that are
in the same pid namespace". From what you are saying above, it seems to
be that there is no such "fundamental" unit. It can be a random mixture
of tasks (taken across vservers) whose resource consumption needs to be
controlled. Is that correct?

> > echo "cid 2" > /dev/cpu/prof/tasks
>
> Adding that feature sounds fine,

Ok yes ..that can be a optional feature.

--
Regards,
vatsa

_____
Containers mailing list
Containers@lists.osdl.org
https://lists.osdl.org/mailman/listinfo/containers

Subject: Re: [ckrm-tech] [PATCH 0/2] resource control file system - aka containers on top of nsproxy!
Posted by Sam Vilain on Mon, 12 Mar 2007 21:15:45 GMT
View Forum Message <> Reply to Message

Srivatsa Vaddagiri wrote:
> On Mon, Mar 12, 2007 at 10:56:43AM -0500, Serge E. Hallyn wrote:
>
>> What's wrong with that?
>>
>
> I had been asking around on "what is the fundamental unit of res mgmt
> for vservers" and the answer I got (from Herbert) was "all tasks that are
> in the same pid namespace". From what you are saying above, it seems to
> be that there is no such "fundamental" unit. It can be a random mixture
> of tasks (taken across vservers) whose resource consumption needs to be
> controlled. Is that correct?
>

Sure, for instance, all postgres processes across all servers might be
put in a different IO and buffercache use container by the system
administrator.

Sam.

_____
Containers mailing list
Containers@lists.osdl.org
https://lists.osdl.org/mailman/listinfo/containers

Subject: Re: [ckrm-tech] [PATCH 0/2] resource control file system - aka containers on top of nsproxy!
Posted by Herbert Poetzl on Mon, 12 Mar 2007 23:31:13 GMT
View Forum Message <> Reply to Message

On Mon, Mar 12, 2007 at 09:50:45PM +0530, Srivatsa Vaddagiri wrote:
> On Mon, Mar 12, 2007 at 10:56:43AM -0500, Serge E. Hallyn wrote:
> > What's wrong with that?
>
> I had been asking around on "what is the fundamental unit of res mgmt
> for vservers" and the answer I got (from Herbert) was "all tasks that are
> in the same pid namespace". From what you are saying above, it seems to
> be that there is no such "fundamental" unit. It can be a random mixture
> of tasks (taken across vservers) whose resource consumption needs to be
> controlled. Is that correct?

just means that the current Linux-VServer behaviour
is a subset of that, no problem there as long as
it really _is_ a subset :) we always like to provide
more features in the future, no problem with that :)

best,
Herbert

> > > echo "cid 2" > /dev/cpu/prof/tasks
> >
> > Adding that feature sounds fine,
>
> Ok yes ..that can be a optional feature.
>
> --
> Regards,
> vatsa
> _____
> Containers mailing list
> Containers@lists.osdl.org
> https://lists.osdl.org/mailman/listinfo/containers

_____
Containers mailing list
Containers@lists.osdl.org
https://lists.osdl.org/mailman/listinfo/containers

Subject: Re: [ckrm-tech] [PATCH 0/2] resource control file system - aka containers on top of nsproxy!
Posted by Srivatsa Vaddagiri on Tue, 13 Mar 2007 02:22:34 GMT
View Forum Message <> Reply to Message

On Tue, Mar 13, 2007 at 12:31:13AM +0100, Herbert Poetzl wrote:
> just means that the current Linux-VServer behaviour
> is a subset of that, no problem there as long as
> it really _is_ a subset :) we always like to provide
> more features in the future, no problem with that :)

Considering the example Sam quoted, doesn't it make sense to split resource classes (some of them atleast) independent of each other? That would also argue for providing multiple hierarchy feature in Paul's patches.

Given that and the mail Serge sent on why nsproxy optimization is usefull given numbers, can you reconsider your earlier proposals as below:

- pid_ns and resource parameters should be in a single struct
  (point 1c, 7b in [1])

- pointers to resource controlling objects should be inserted
  in task_struct directly (instead of nsproxy indirection)
  (points 2c in [1])

[1] http://lkml.org/lkml/2007/3/12/138

--
Regards,
vatsa

_____
Containers mailing list
Containers@lists.osdl.org
https://lists.osdl.org/mailman/listinfo/containers