Subject: Re: [ckrm-tech] [PATCH 3/9] Containers (V9): Add tasks file interface
Posted by Paul Menage on Thu, 10 May 2007 21:21:39 GMT
View Forum Message <> Reply to Message

On 5/8/07, Balbir Singh <balbir@linux.vnet.ibm.com> wrote:
>
> I now have a use case for maintaining a per-container task list.
> I am trying to build a per-container stats similar to taskstats.
> I intend to support container accounting of
>
> 1. Tasks running
> 2. Tasks stopped
> 3. Tasks un-interruptible
> 4. Tasks blocked on IO
> 5. Tasks sleeping
>
> This would provide statistics similar to the patch that Pavel had sent out.
>
> I faced the following problems while trying to implement this feature
>
> 1. There is no easy way to get a list of all tasks belonging to a container
>   (we need to walk all threads)

Well, walking the taks list is pretty easy - but yes, it could become
inefficient when there are many small containers in use.

I've got some ideas for a way of tracking this specifically for
containers with subsystems that want this, while avoiding the overhead
for subsystems that don't really need it. I'll try to add them to the
next patchset.

> 2. There is no concept of a container identifier. When a user issues a command
>   to extract statistics, the only unique container identifier is the container
>   path, which means that we need to do a path lookup to determine the dentry
>   for the container (which gets quite ugly with all the string manipulation)

We could just cache the container path permanently in the container,
and invalidate it if any of its parents gets renamed. (I imagine this
happens almost never.)

>
>   Adding a container id, will make it easier to find a container and return
>   statistics belonging to the container.

Not unreasonable, but there are a few questions that would have to be answered:

- how is the container id picked? Like a pid, or user-defined? Or some
kind of string?

- how would it be exposed to userspace? A generic control file
provided by the container filesystem in all container directories?

- can you give a more concrete example of how this would actually be
useful? For your container stats, it seems that just reading a control
file in the container's directory would give you the stats that you
want, and userspace already knows the container's name/id since it
opened the control file.

Paul