Subject: Re: [PATCH 0/7] containers (V7): Generic Process Containers
Posted by Paul Menage on Tue, 20 Feb 2007 17:55:11 GMT
View Forum Message <> Reply to Message

On 2/20/07, Eric W. Biederman <ebiederm@xmission.com> wrote:
> "Paul Menage" <menage@google.com> writes:
>
> > On 2/12/07, Sam Vilain <sam@vilain.net> wrote:
> >>
> >> I know I'm a bit out of touch, but AIUI the NSProxy *is* the container.
> >> We decided a long time ago that a container was basically just a set of
> >> namespaces, which includes all of the subsystems you mention.
> >
> > You may have done that, but the CKRM/ResGroups independently decided a
> > long time ago that the fundamental unit was the resource class, and
> > the OpenVZ folks decided that the fundamental unit was the
> > BeanCounter, and the CPUSet folks decided that the fundamental unit
> > was the CPUSet, etc ... :-)
>
> Using the container name is bad and it led to this stupid argument.
>
> The fundamental unit of what we have merged into the kernel is the
> namespace. The aggregate of all namespaces and everything is the
> container.

What are you defining here as "everything"? If you mean "all things that could be applied to a segregated group of processes such as a virtual server", then "container" seems like a good name for my patches, since it allows you to aggregate namespaces, resource control, other virtualization, etc.

Sam said "the NSProxy *is* the container". You appear to be planning to have some namespaces, possibly not aggregated within the nsproxy (pid namespace?) but are you planning to have some higher-level "container" object that aggregates the nsproxy and the other namespaces? If so, what is it? Does it track process membership, etc? What's the userspace API? In what fundamental ways would it be different from my generic containers patches with Serge Hallyn's namespace subsystem. (If these questions are answered already by designs or code, I'd be happy to be pointed at them).

I guess what it comes down to, is why is an aggregation of namespaces suitable for the name "container", when an aggregation of namespaces and other resource controllers isn't?

What do you think might be a better name for the generic process groups that I'm pushing? As I said, I'm happy to do a simple search/replace on my code to give a different name if that turned out

to be the gating factor to getting it merged. But I'd be inclined to
leave that decision up to Andrew/Linus.

> > But there's a lot of common ground between these different approaches,
> > and potential for synergy, so the point of this patch set is to
> > provide a unification point for all of them, and a stepping stone for
> > other new resource controllers and process control modules.
>
> For the case of namespaces I don't see how your code makes things
> better. I do not see a real problem that you are solving.

I'm trying to solve the problem that lots of different folks
(including us) are trying to do things that aggregate multiple process
into some kind of constrained group, and are all trying to use
different and incompatible ways of grouping/tracking those processes.

I agree that namespaces fit slightly less well into this model than
some other subsystems like resource management. But by integrating
with it you'd get automatic access to all the various different
resource controller work that's being done.

Paul

_____