Subject: Re: [RFC][PATCH 1/5] Virtualization/containers: startup
Posted by ebiederm on Tue, 07 Feb 2006 01:57:05 GMT
View Forum Message <> Reply to Message

Cedric Le Goater <clg@fr.ibm.com> writes:

> Eric W. Biederman wrote:
>
>> As someone said to me a little bit ago, for migration or checkpointing
>> ultimately you have to capture the entire user/kernel interface if
>> things are going to work properly. Now if we add this facility to
>> the kernel and it is a general purpose facility. It is only a matter
>> of time before we need to deal with nested containers.
>>
>> Not considering the case of having nested containers now is just foolish.
>> Maybe we don't have to implement it yet but not considering it is silly.
>
> That could be restricted. Today, process groups are not nested. Why do you
> think nested containers are inevitable ?

process groups are a completely different kind of beast.
A closer analogy are hierarchical name spaces and mounts.
If we didn't need things like waitpid outside one pid namespace
to wait for a nested namespace they would be complete disjoint
and the implementation would be trivial.

>> As far as I can tell there is a very reasonable chance that when we
>> are complete there is a very reasonable chance that software suspend
>> will just be a special case of migration, done complete in user space.
>
> Being able to sofware suspend one container among many would be a very
> interesting feature to have.

That is what checkpointing. And that is simply the persistent form of
migration.

Eric