Subject: Re: Checkpoint/restart (was Re: [PATCH 0/4] - v2 - Object creation with a specified id)
Posted by Dave Hansen on Tue, 22 Apr 2008 22:56:07 GMT

View Forum Message <> Reply to Message

On Wed, 2008-04-23 at 01:01 +0400, Alexey Dobriyan wrote:
> On Tue, Apr 22, 2008 at 11:56:20AM -0700, Dave Hansen wrote:
> > On Tue, 2008-04-22 at 23:36 +0400, Alexey Dobriyan wrote:> >
> > > a) should mainline kernel have checkpoint/restart feature at all
> > > b) if yes, should it be done from kernel- or userspace?
> > >
> > > Until agreement will be "yes/from userspace" such patches don't make
> > > sense in mainline.
> >
> > What do you mean by "from kernel" or "from userspace"?
>
> By "from userspace" I mean proposed interfaces and similar: usespace by
> special system calls puts some state of future object and then does
> normal system call which creates aforementioned object.
>
> This can be used for PIDs, OK.
>
> This can be used for SystemV shmem ids. But SystemV shmem also has
> (let's choose) uid/gid, atimes and actual content. How would restoration
> look like?
>
> How to restore struct task_struct::did_exec ? Do execve(2)?
>
> A was ptracing B at checkpoint moment...
>
> Netdevices: stats, name, all sorts of flags, hw addresses, MTU
>
> iptables rules.

Don't we already have interfaces to dump these out and restore them?

My argument is this: If we have interfaces that exist (like setting up
iptables rules) we shouldn't make a second interface *just* for
checkpoint/restart. We have a hard enough time getting *one* interface
right for things, I can't imagine getting two right, and *keeping* them
right.

If the current interface is insufficient, we should first expand it in
such a way that it can be used for checkpoint. That certainly won't
work in all cases. fork(), for instance, doesn't take any arguments and
is going to be awfully hard to expand. :)

I'd love to hear some of your insights about how things like the current

iptables interfaces are insufficient for checkpoint/restart.

> These next ids are suitable, well, only for ids which is very, very small
> part of kernel state needed to restore group of processes.

I couldn't agree more. This id setting mechanism would only be useful
for a small subset of the things we need during a restart.

-- Dave

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