
Subject: Re: Checkpoint/restart (was Re: [PATCH 0/4] - v2 - Object creation with a specified id)

Posted by [ebiederm](#) on Thu, 10 Jul 2008 18:55:34 GMT

[View Forum Message](#) <> [Reply to Message](#)

"Serge E. Hallyn" <serue@us.ibm.com> writes:

>> So, the checkpoint-as-a-corefile idea sounds good to me, but it
>> definitely leaves a lot of questions about exactly how we'll need to do
>> the restore.
>
> Talking with Dave over irc, I kind of liked the idea of creating a new
> fs/binfmt_cr.c that executes a checkpoint-as-a-coredump file.
>
> One thing I do not like about the checkpoint-as-coredump is that it begs
> us to dump all memory out into the file. Our plan/hope was to save
> ourselves from writing out most memory by:
>
> 1. associating a separate swapfile with each container
> 2. doing a swapfile snapshot at each checkpoint
> 3. dumping the pte entries (/proc/self/)
>
> If we do checkpoint-as-a-coredump, then we need userspace to coordinate
> a kernel-generated coredump with a user-generated (?) swapfile snapshot.
> But I guess we figure that out later.

Well it is a matter of which VMAs you dump. For things that are file backed you need to dump them.

I don't know that even a binfmt for per process level checkpoints is sufficient but I do know having something of that granularity looks much easier. Otherwise it takes a bazillian little syscalls to do things no one else is interested in doing.

Eric

Containers mailing list

Containers@lists.linux-foundation.org

<https://lists.linux-foundation.org/mailman/listinfo/containers>
