Subject: Re: [BUG][cryo] Create file on restart ?
Posted by Oren Laadan on Thu, 17 Jul 2008 23:35:19 GMT
View Forum Message <> Reply to Message

Serge E. Hallyn wrote:
> Quoting Matt Helsley (matthltc@us.ibm.com):
>> On Wed, 2008-07-16 at 14:26 -0700, sukadev@us.ibm.com wrote:
>>> Serge E. Hallyn [serue@us.ibm.com] wrote:
>>> | Quoting sukadev@us.ibm.com (sukadev@us.ibm.com):
>>> | > Serge E. Hallyn [serue@us.ibm.com] wrote:
>>> | > | Quoting sukadev@us.ibm.com (sukadev@us.ibm.com):
>>> | > | >
>>> | > | > cryo does not (cannot ?) recreate files if the application created
>>> | > |
>>> | > | I think that's for the best.
>>> | > |
>>> | > | Don't you?
>>> | >
>>> | > I can understand that configuration or data files should exist, but
>>> | > not sure about temporary or log files that an application created
>>> | > upon start-up and expects to be present. Should the admin find
>>> | > out about them and create them by hand before restart ?
>>> |
>>> | I think the admin should have set the destination environment such that
>>> | the task is restarted in the same network fs in the same directory, with
>>> | no files having been deleted.
>> [Assuming Serge meant: s/network fs/network, fs,/]
>
> Well no I meant a network filesystem - at least if you're migrating apps
> around a cluster.
>
>>> or new files created ? For instance if the application was checkpointed
>>> before it created a temporary file with O_EXCL flag, that temporary
>>> file must not exist when restarting ?
>> I think that's not a problem given my assumptions above. The filesystem
>> that the application restarts in would be the same because the admin
>> should have set up the restart environment as Serge suggested. The admin
>> can't rely on restart in an alternate environment. However, given
>> knowledge of the application and environment, using an alternate
>> environment may be a risk the admin is willing to take.
>
> Yup. But Suka is right that in the case of the checkpointed app
> continuing to run for a bit before being killed and restarted, it could
> get out of whack with respect to the file system.
>
>>> | Am I wrong?
>>>
>>> So we take a snapshot of the FS and checkpoint the application. Do they

>>> need to be atomic ?
>> If all the applications in a container are frozen then I think we can
>> get fs snapshots consistent with checkpointed applications.
>> Otherwise, yes, I think we'd be gambling that the checkpointed
>> application isn't interacting with another, running, application via an
>> intermittently-shared file.
>
> What fun :)
>
> I wonder whether the experience of users of c/r on sgi and cray could
> teach us anything here.

if you are checkpointing to migrate the application - you need not worry
about the file system, as it may not change while you migrate.

if you are checkpointing to be able to be able to recover from an error
later, you need to snapshot the file system, but you may get away with
it in some cases.

if you are checkpointing to be able to travel back in time (return to
older than last checkpoint), you certainly need to snapshot the file system.

in any event, I think this is something we may want to discuss in the
mini-summit.

Oren.

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

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