Subject: cryo and mm->arg_start
Posted by serue on Fri, 11 Jul 2008 13:13:45 GMT
View Forum Message <> Reply to Message

What cryo does right now to restart some task (say openmp stream) is:

1. fork, ptrace_tracem(), then execute the original application (stream)
2. (some other stuff)
3. through ptrace, cause the restarted process to read the checkpointed data back into writeable maps. This includes the stack

The restarted task's filename is correctly reported through /proc/$$/cmdline. Once we rewrite the stack, it is corrupted.

The reason is that the cmdline contents are taken from mm->arg_start, which varies with each execution.

On the one hand it's kind of a "small thing." But IIUC it's like did_exec in that there is no way to fix it for userspace.

One thing we could do here is to start extending the cryo approach with Eric's checkpoint-as-a-coredump (caac?). We generate the tiniest of coredumps which, at first, contains nothing but mm->arg_start and maybe a process id. It would be simplest if it also contained a filename for the real executable, but I don't know that we could get away with that. If we *could* get away with that, then we could have a trivial fs/binfmt_cr.c "execute" such a caac file, which would mean it would exec the original executable, then change process settings in accordance with the ccac file contents.

Any other ideas? Comments?

-serge

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