



Contribution ID: 82

Type: **not specified**

Update on Task Migration at Google Using CRIU

Tuesday, September 10, 2019 4:00 PM (30 minutes)

Over the last year we have worked on expanding the task migration using CRIU in Google. The talk will discuss how in some cases the kernel interfaces are lacking for the purpose of migration:

- Lack of support for reading rseq configuration which means that it requires userspace support to migrate users of rseq properly.
- Lack of support for reading what cgroup events the users have registered for.
- Many kernel C/R interfaces are protected by CAP_SYS_ADMIN which we deemed unsafe to have for the migrator agent - CAP_RESTORE could be the solution.

We will discuss new challenges which we have encountered while developing the migration technology further:

- The lack of clean error classification in CRIU forced us to parse the migration logs.
- Lack of support for some less often used kernel features in CRIU (e.g. O_PATH, PR_SET_CHILD_SUBREAPER).
- Migrating containers while also changing the IP of the container is hard but in many cases could be done with little effort on the library or user side.
- We have finalized streaming migration support on our side and in the process we have realized that the hitless migration is infeasible for our latency sensitive users.

I agree to abide by the anti-harassment policy

Yes

Primary author: YURTSEVER, Kamil (Google)

Presenter: YURTSEVER, Kamil (Google)

Session Classification: Containers and Checkpoint/Restore MC