Contribution ID: **197** Type: **not specified**

Restoring process trees with child-sub-reapers, nested pid-namespaces and inherit-only resources.

Wednesday, 14 September 2022 10:30 (30 minutes)

Re-parenting may put processes having same inherit-only resource into completely different and far away locations in the process tree, so that they don't have ancestor/descendant relations between each other anymore.

In mainstream CRIU currently we don't have nested pid-namespaces support and re-parenting to child-sub-reaper support. We just handle the most common case where task was re-parented to container init. To handle this case CRIU simply checks all children of the container init for non-session-leaders which can't inherit session from init. We "fix" the original tree by moving such children to session leader sub-tree connecting them by helper task. After that we restore tasks based on the "fixed" tree and kill helpers so that we get the right tree which we check to be the same as the dumped one.

In this talk I want to first cover how we handle in Virtuozzo more complex cases with child-sub-reapers [1], nested pid-namespaces [2], and cases where re-parenting breaks longer branches in process tree [2].

And second I want to shed some light on the problem which we can't handle in CRIU easily because of the lack of information from kernel, this problem was known from the early days of CRIU development and it is still present and vital to support arbitrary process trees.

Also I want to present one possible solution for the problem - "CABA" [3] and hope to see some feedback on it.

Links:
https://src.openvz.org/projects/OVZ/repos/criu/commits/70eee0613acf [1]
https://src.openvz.org/projects/OVZ/repos/criu/commits/aa77967c2f6c [2]
https://lore.kernel.org/lkml/20220615160819.242520-1-ptikhomirov@virtuozzo.com/ [3]

I agree to abide by the anti-harassment policy

Yes

Primary author: TIKHOMIROV, Pavel (Virtuozzo)

Presenter: TIKHOMIROV, Pavel (Virtuozzo)

Session Classification: Containers and Checkpoint/Restore MC

Track Classification: LPC Microconference: Containers and Checkpoint/Restore MC