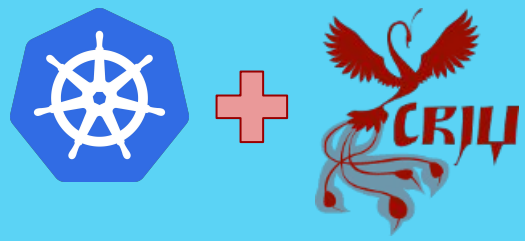
KubeCon | CloudNativeCon

Europe 2024

Enabling Coordinated Checkpointing for Distributed HPC Applications



Radostin Stoyanov - PhD Student @ Scientific Computing Group
Adrian Reber - Senior Principal Software Engineer
Supervisor: Prof. Wes Armour
Paris, 2024, March 22

Agenda

Agenda

- Background
- Integrations
- Use Cases
- Migration Demo
- Coordinated Checkpointing

Background

Checkpoint/Restore in Userspace CRIU

Integrations

Multiple Integrations Exist

OpenVZ

Borg

LXC / LXD / Incus

Docker

Podman

CRI-O

Forensic Container Checkpointing

- https://github.com/kubernetes/enhancements/pull/1990 (alpha 1.25)
- https://github.com/kubernetes/enhancements/pull/3264 (alpha 1.25)
- https://github.com/kubernetes/kubernetes/pull/104907 (alpha 1.25)
- https://kubernetes.io/blog/2022/12/05/forensic-container-checkpointing-alpha/
- https://kubernetes.io/blog/2023/03/10/forensic-container-analysis/
- https://www.opensourcerers.org/2023/09/11/forensic-container-checkpointing-in-openshift
- https://github.com/kubernetes/enhancements/pull/4288 (beta 1.30)
- https://github.com/kubernetes/kubernetes/pull/123215 (beta 1.30)

Use Cases

Migration Demo

Checkpointing Distributed HPC Applications

Checkpointing Distributed HPC Applications

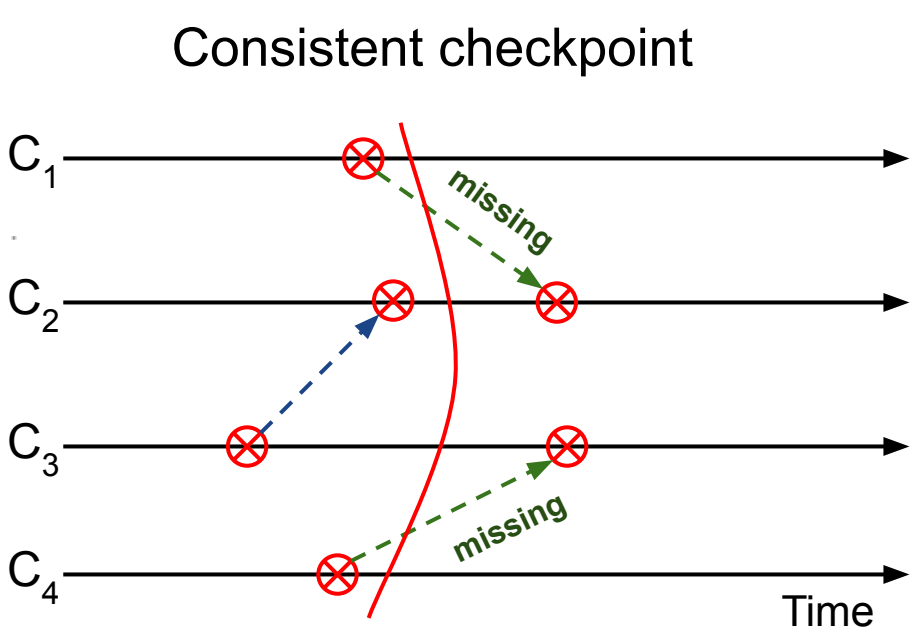
Problem definition and research questions

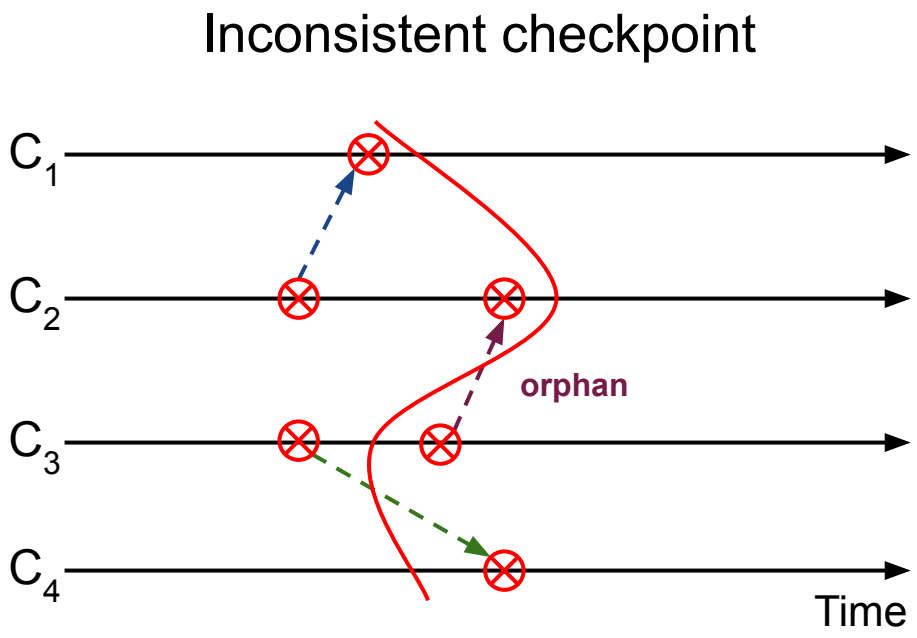
- CRIU was designed for checkpoint/restore of individual process tree, rather than distributed applications
- How to extend CRIU with support for distributed applications?
- How to enable checkpoint coordination within Kubernetes?

Related work

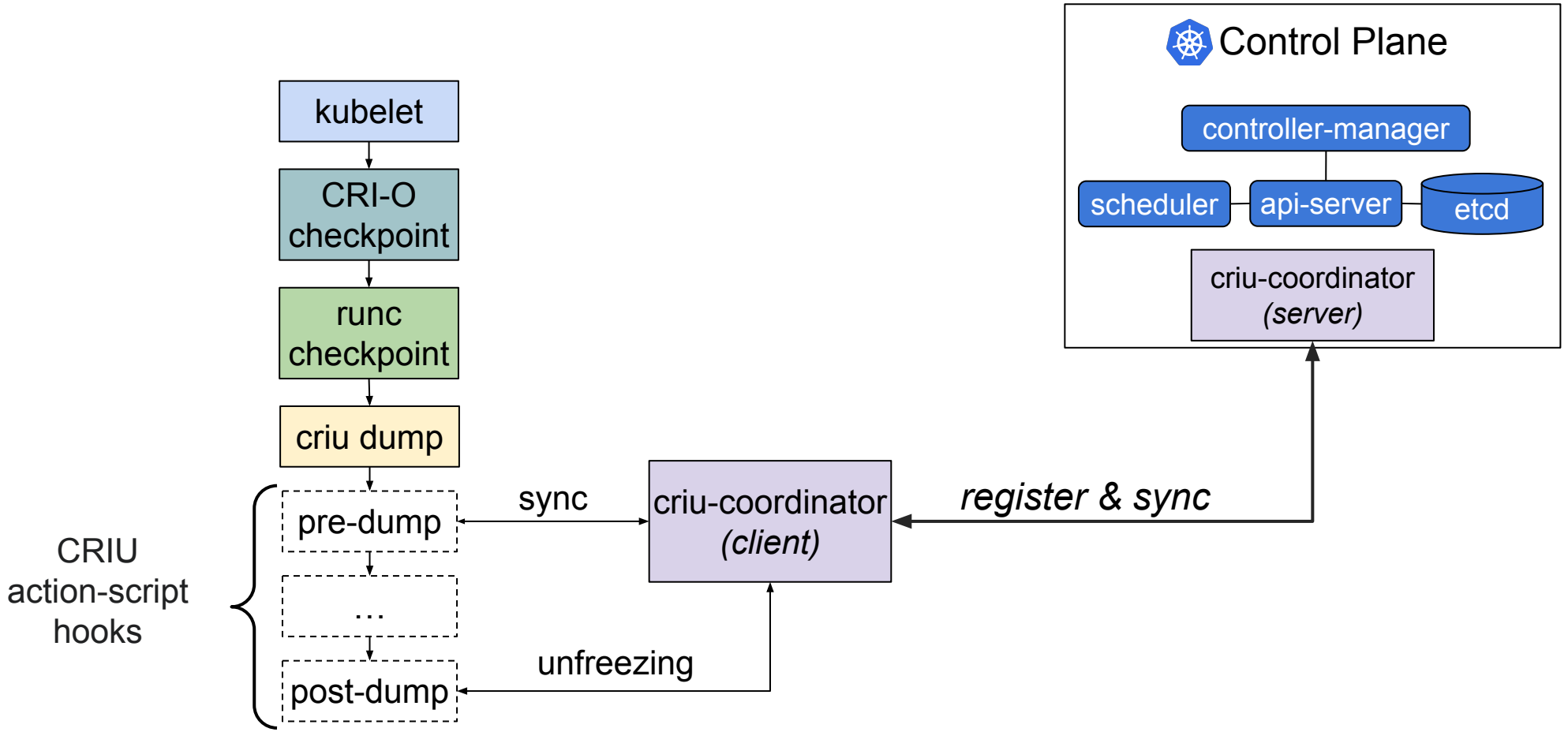
- DMTCP (Distributed MultiThreaded CheckPointing)
- Apache Flink

Coordinated Checkpointing

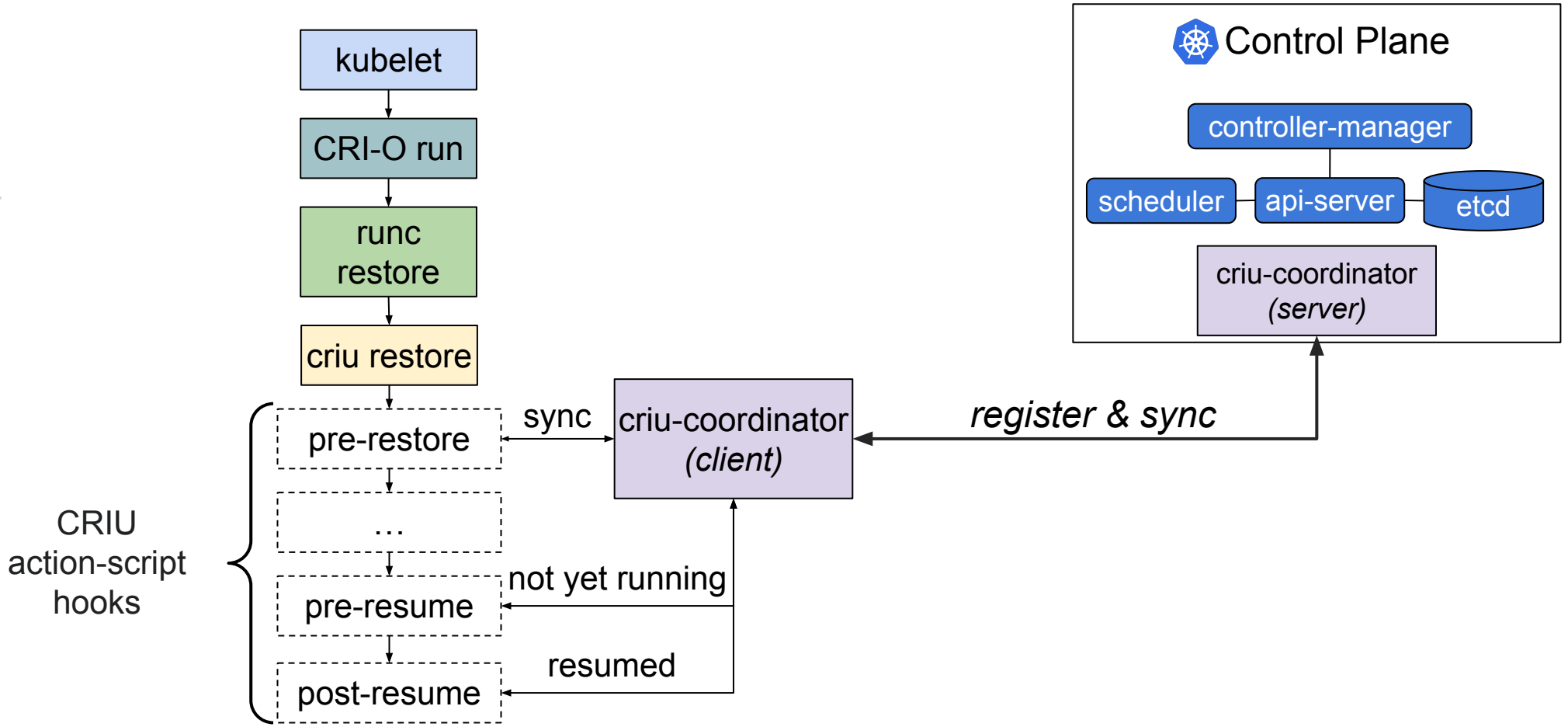




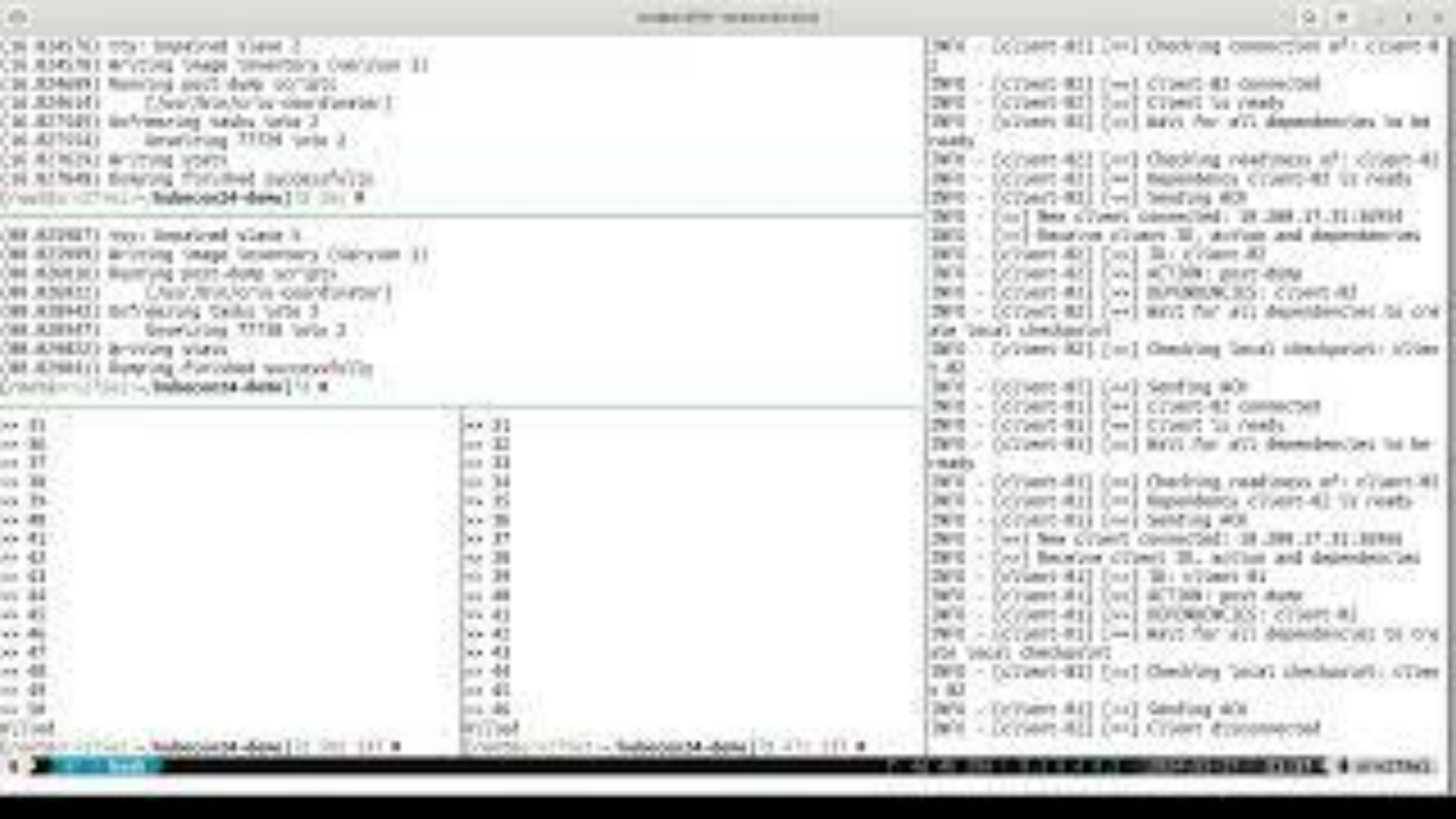
CRIU Coordination Protocol - Checkpointing



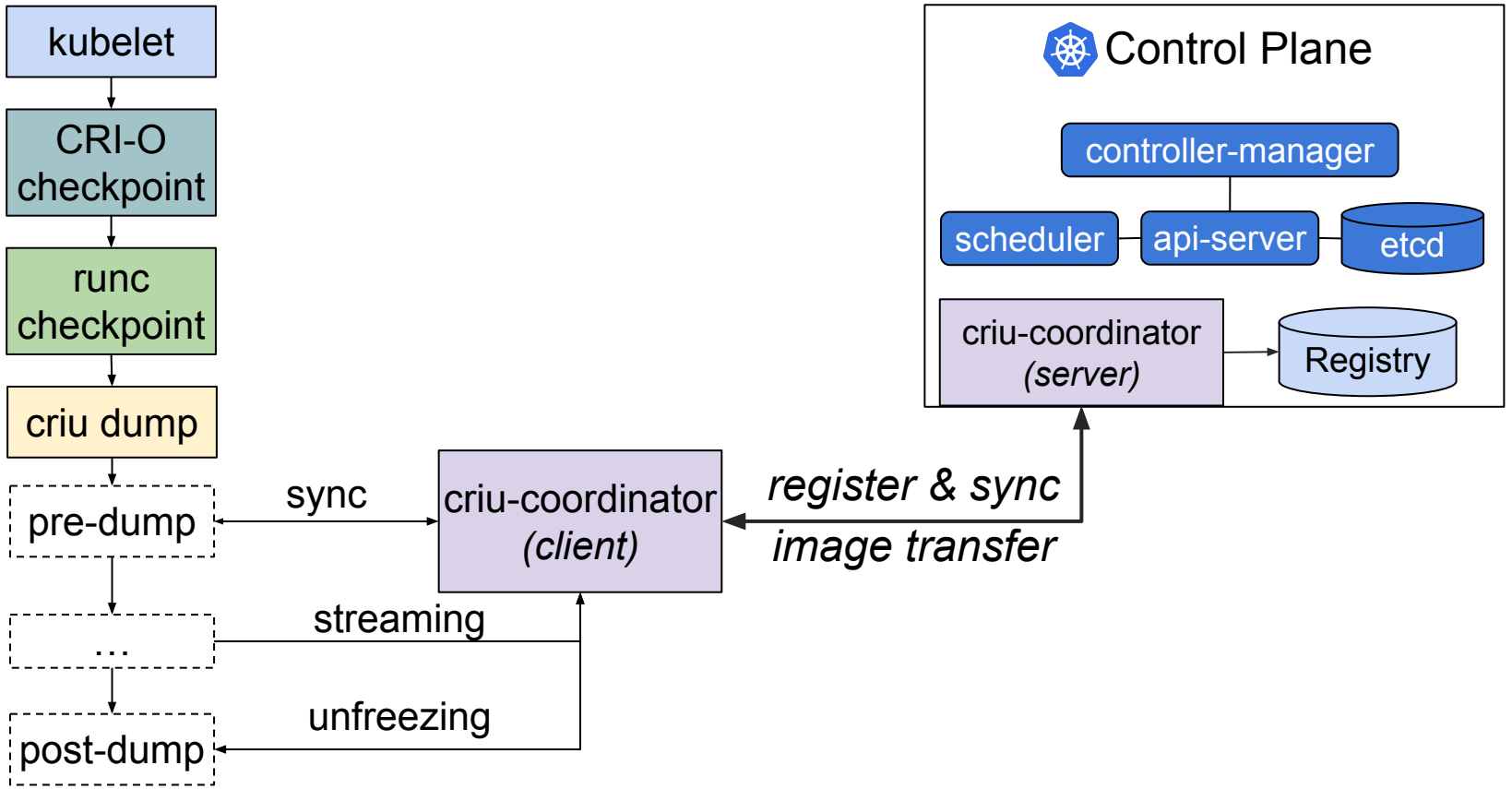
CRIU Coordination Protocol - Restore



Coordinated Checkpointing Demo



CRIU Coordination Protocol - Checkpoint Streaming



Nicolas Viennot, *Fast checkpointing with criu-image-streamer*. Linux Plumbers Conference. 2020 (https://github.com/checkpoint-restore/criu-image-streamer)

Future work

- Adding support for **criu-coordinator** in CRI-O & containerd
- Automating checkpoint dependency detection in Kubernetes
- Integrating checkpoint/restore with Kubernetes objects (e.g., Deployments, StatefulSets, etc.)

KubeCon | CloudNativeCon

Europe 2024