

Transparent Container Checkpointing and Rollback-Recovery with Kubernetes

Open Challenges & Research Directions

Radostin Stoyanov

DPhil Candidate, Scientific Computing Group

Supervisor: Prof. Wes Armour

Checkpoint/Restore in Userspace



- Open-source tool for transparent checkpoint/restore of Linux processes
- Implemented entirely in userspace
- Integrated with:
 - runc / crun / youki
 - Docker / Podman / LXD
 - CRI-O / containerd
 - Kubernetes

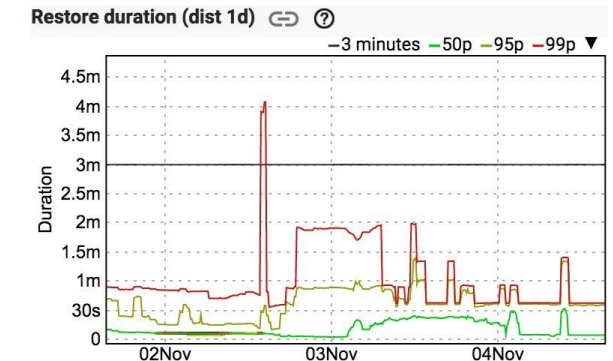
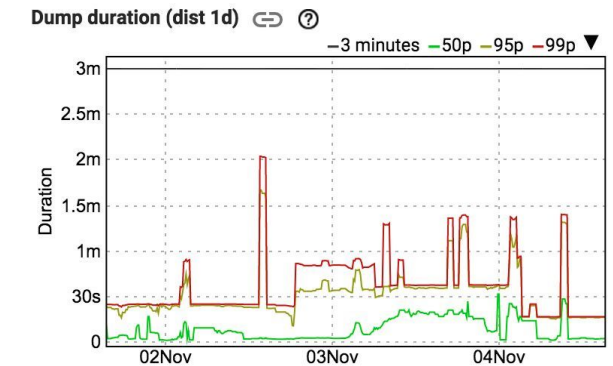
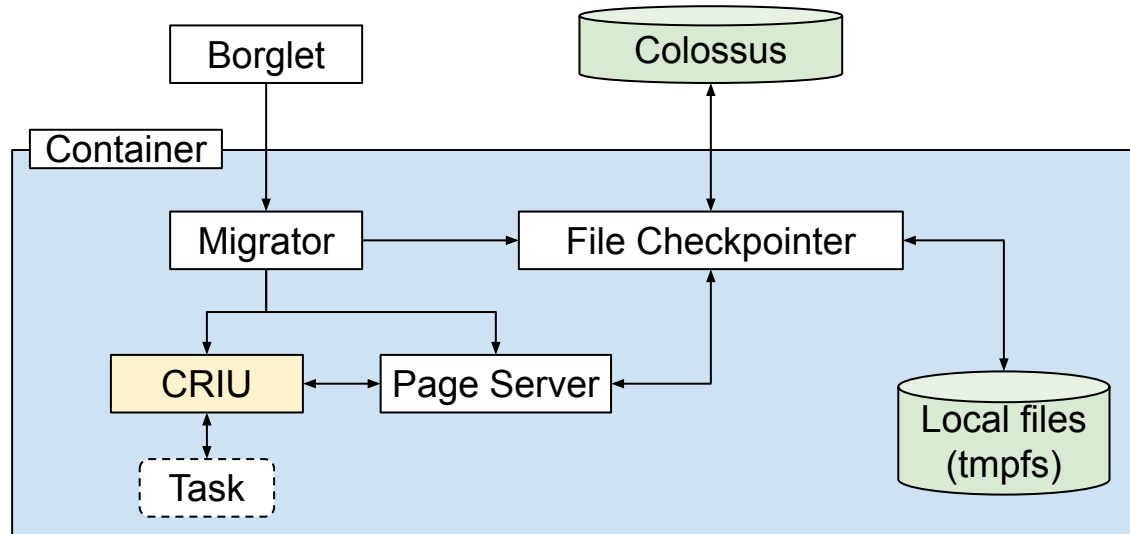
Container Checkpoint / Restore



Common Use cases and Scenarios for Container Checkpointing

Task Migration at Google Using CRIU

Task Migration at Google Using CRIU



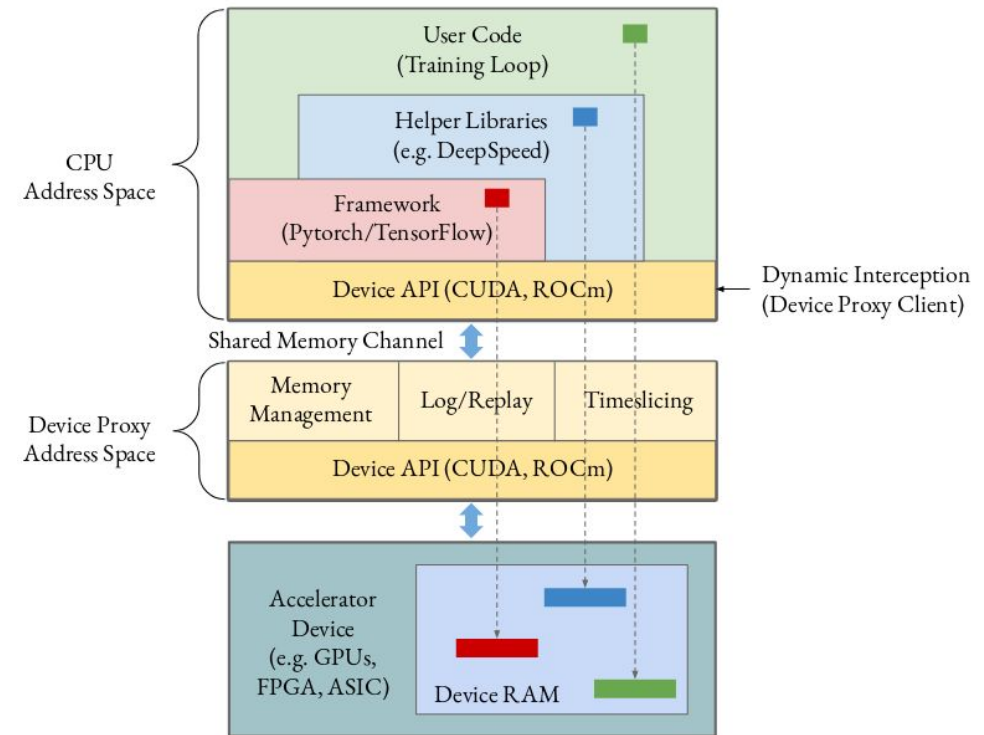
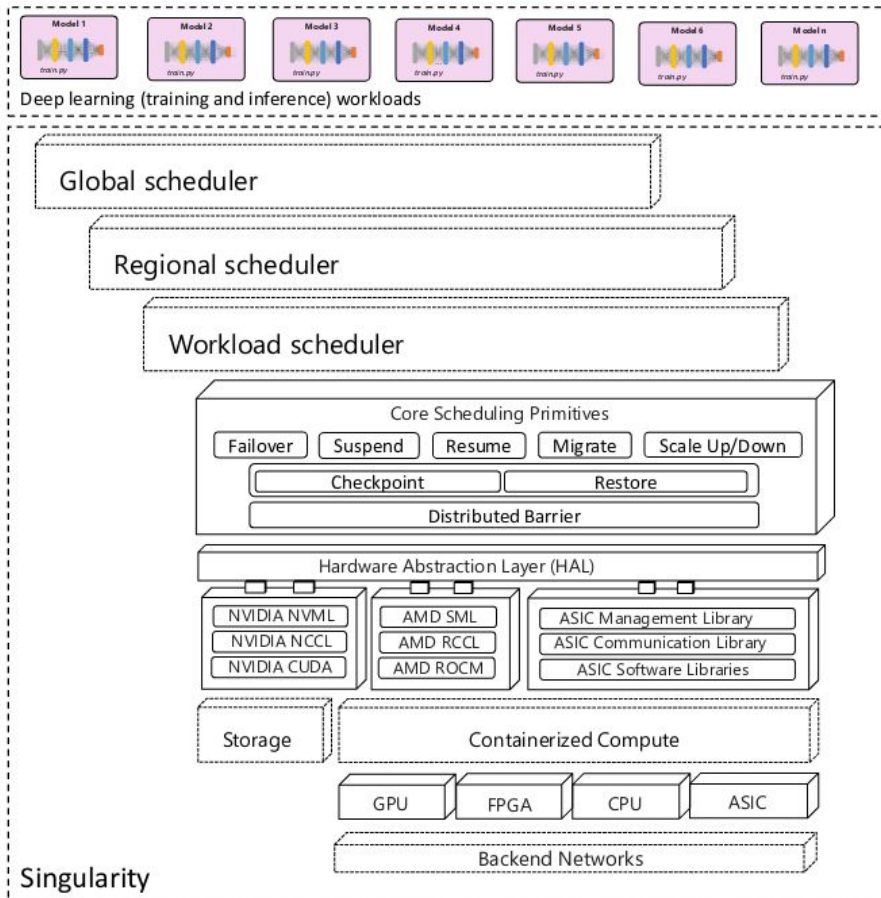
Performance:

- Streaming is significantly faster than staging on local disk or tmpfs
- Scaling issues for tasks with large numbers (~1000s) of threads and sockets
- 1-2 minutes blackout time for most tasks

Kamil Yurtsever (Google). *Update on Task Migration at Google Using CRIU*. Linux Plumbers Conference. 2019
Radoslaw Burny (Google). *Securely Migrating Untrusted Workloads with CRIU*. Linux Plumbers Conference. 2018
Andy Tucker (Google). *Task Migration at Google Using CRIU*. Linux Plumbers Conference. 2018
Victor Marmol (Google). *Task Migration at Scale Using CRIU*. Linux Plumbers Conference. 2018

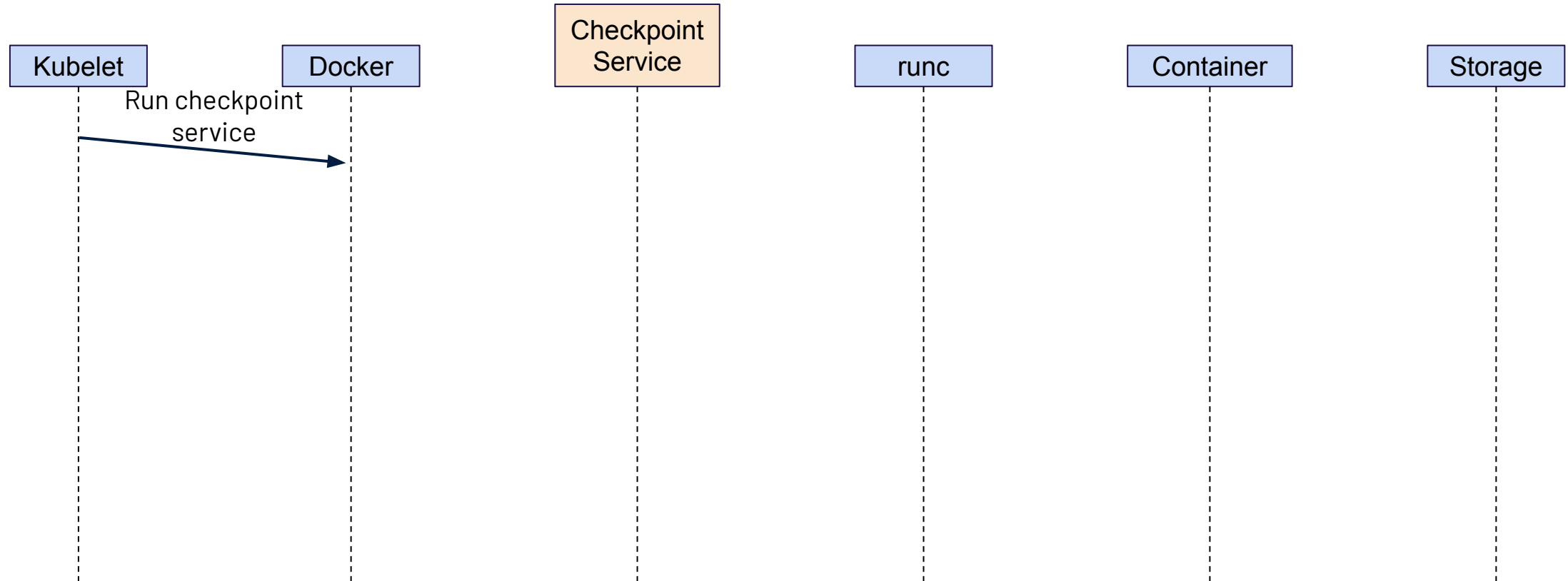
Scheduling of AI Workloads

Scheduling of AI Workloads

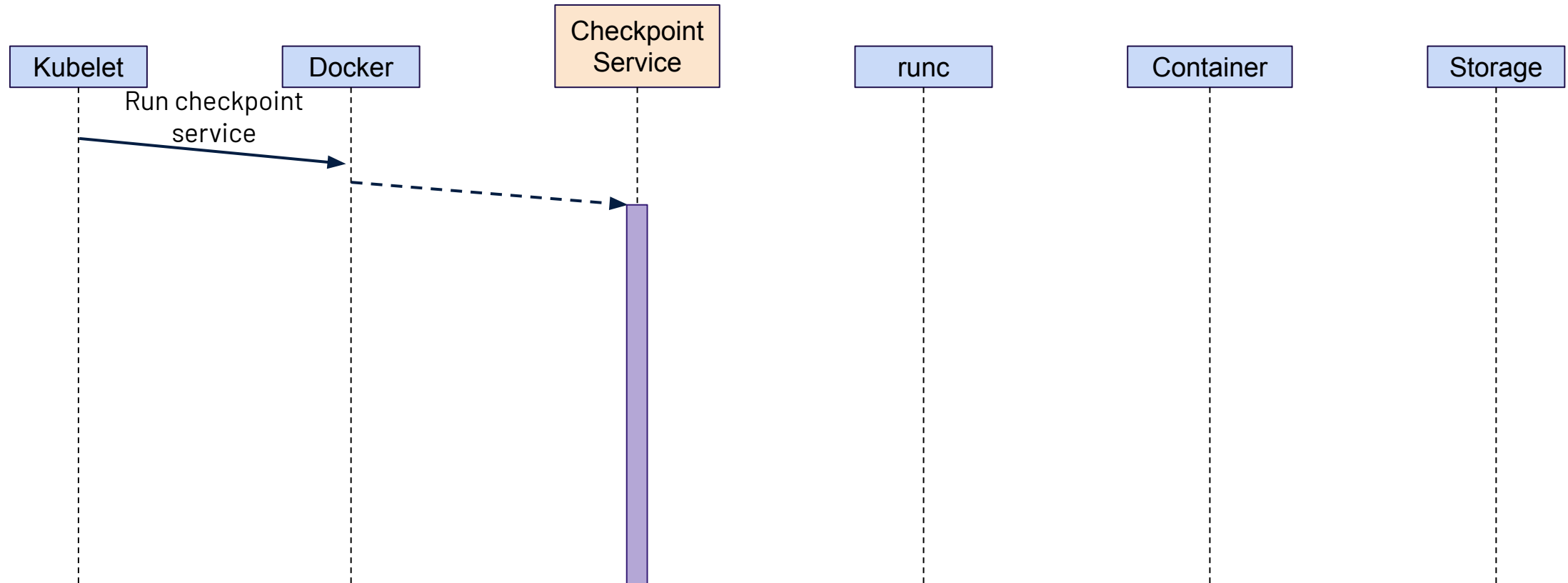


Scaling Microservices Horizontally

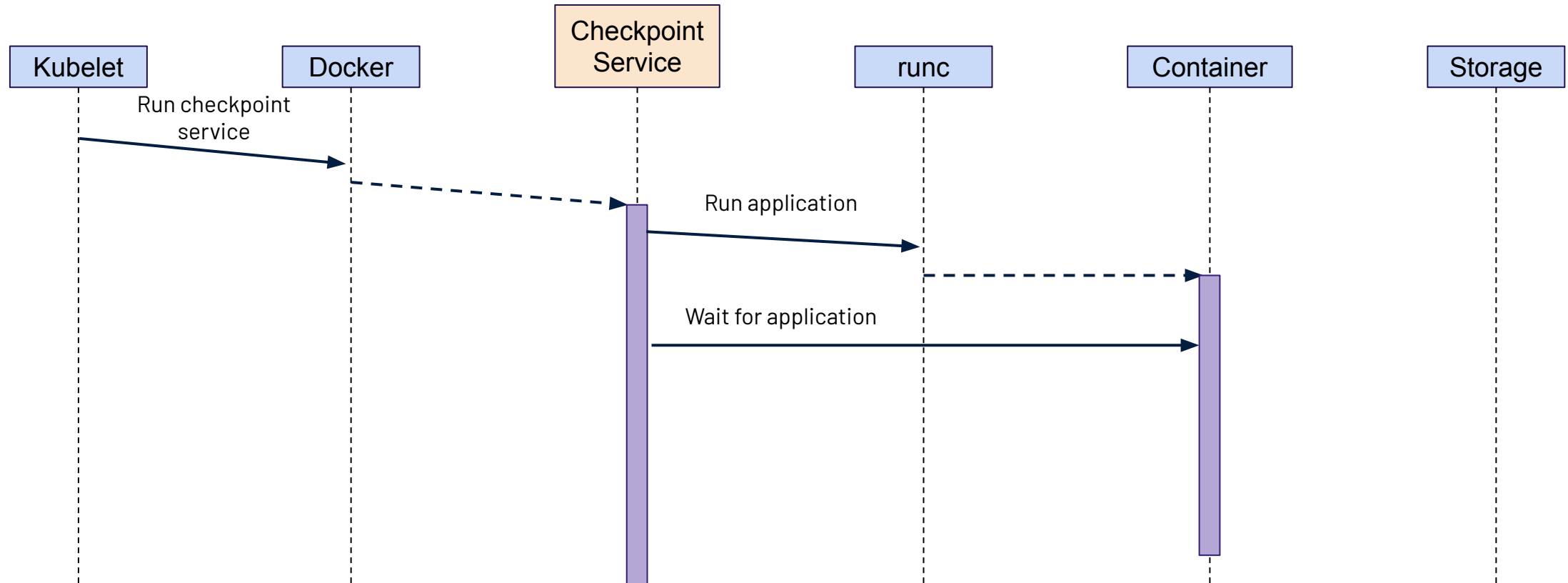
Scaling Microservices Horizontally



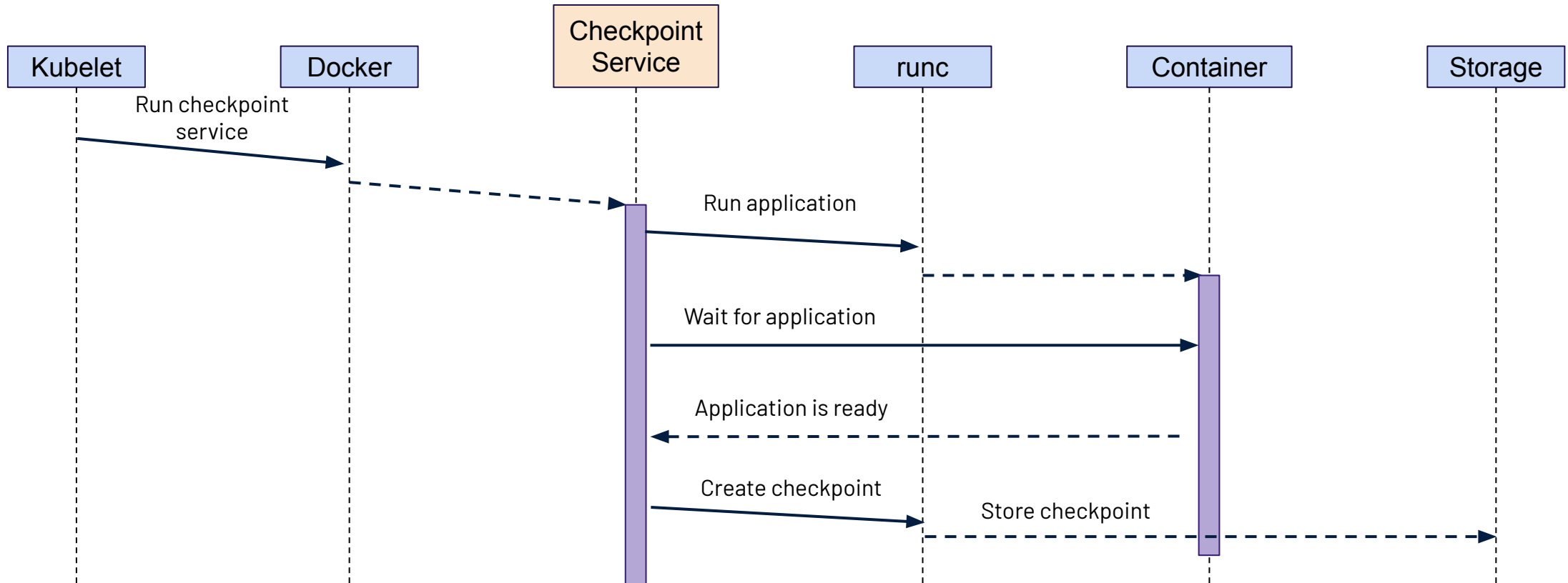
Scaling Microservices Horizontally



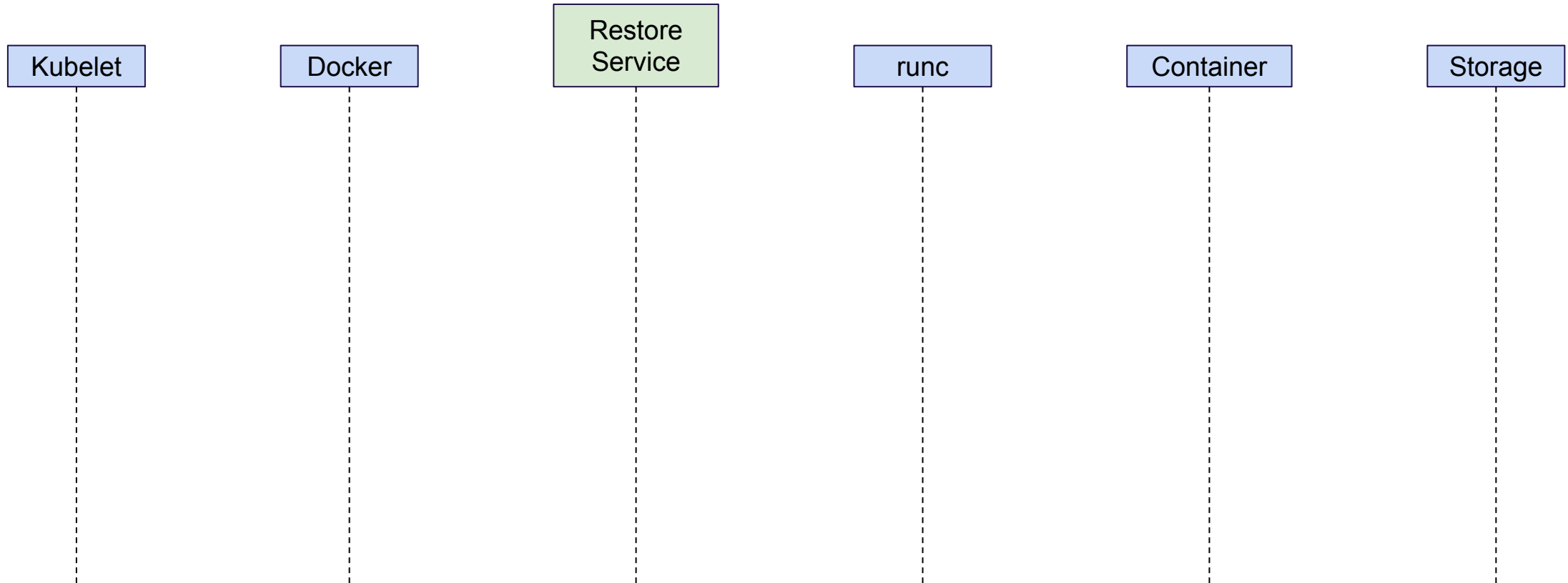
Scaling Microservices Horizontally



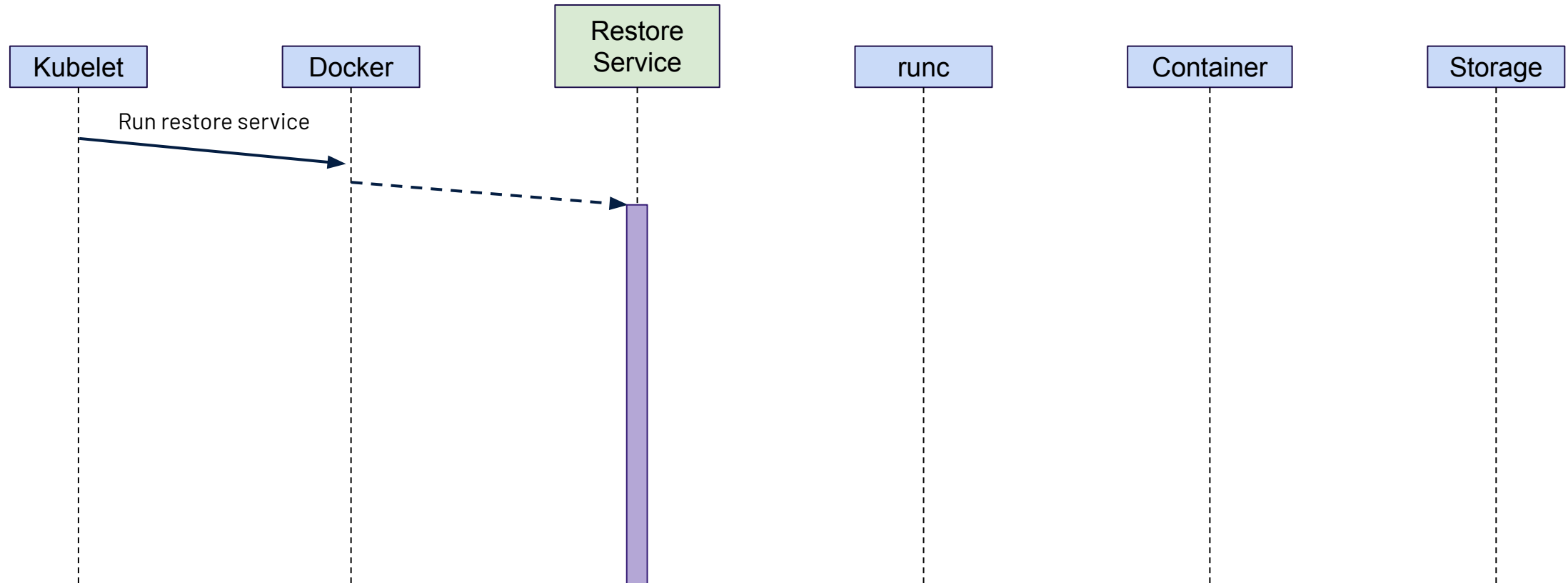
Scaling Microservices Horizontally



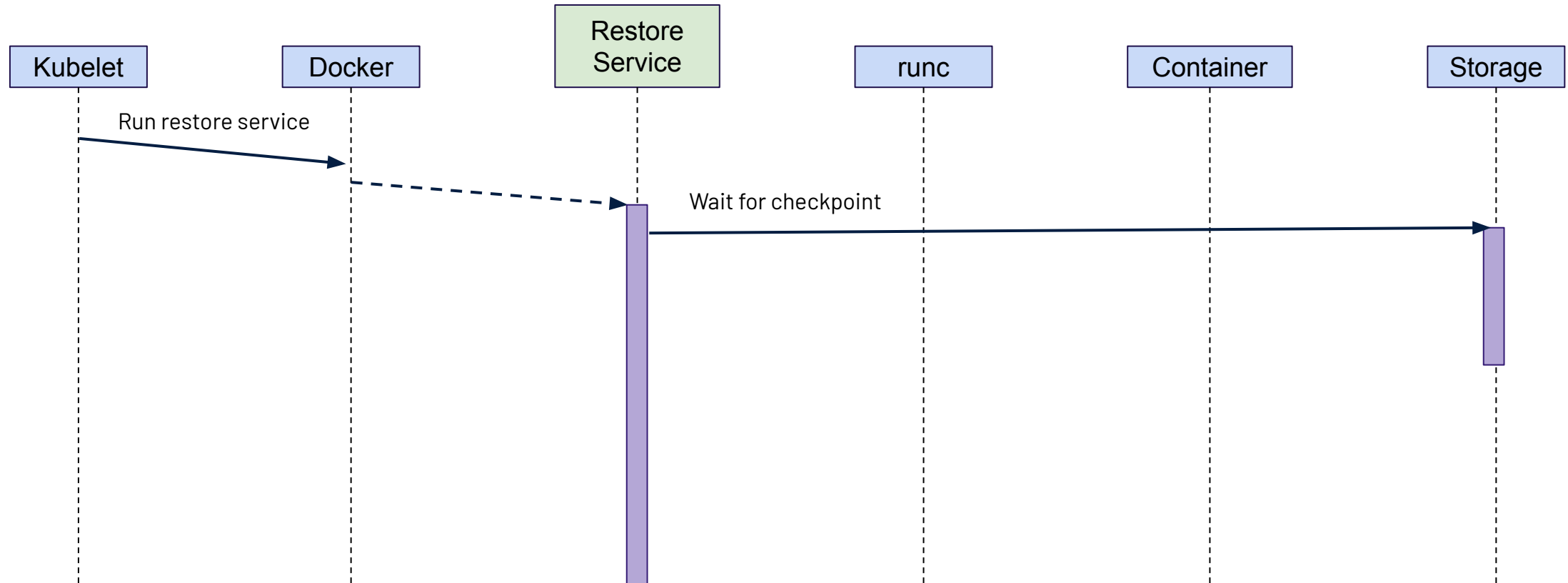
Scaling Microservices Horizontally



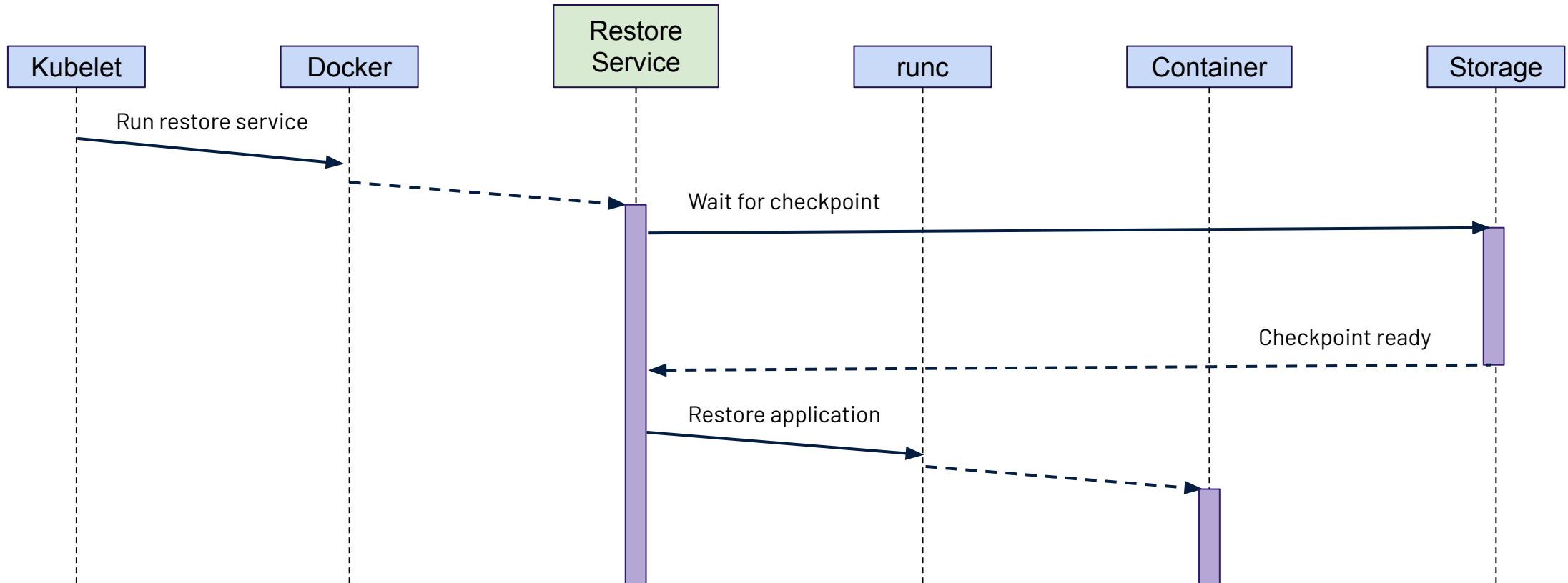
Scaling Microservices Horizontally



Scaling Microservices Horizontally



Scaling Microservices Horizontally

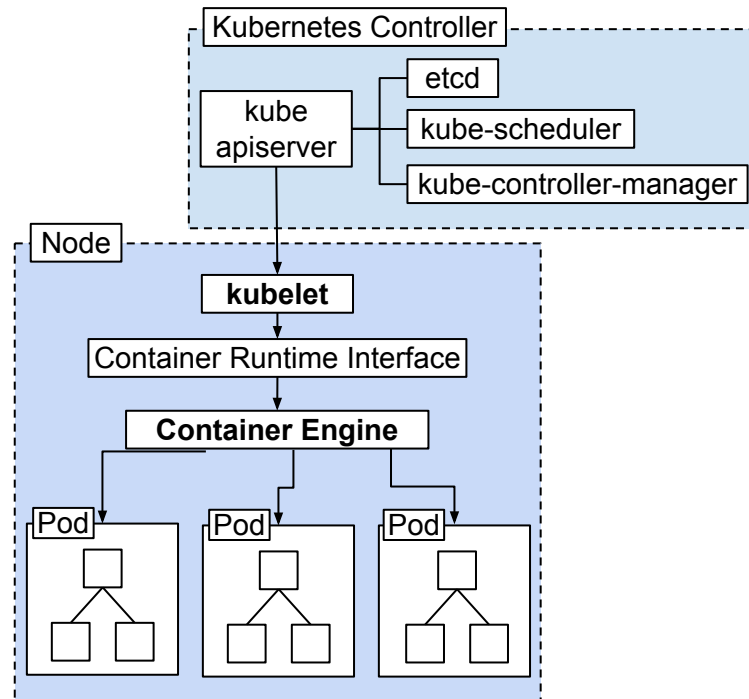


Container Checkpointing in Kubernetes

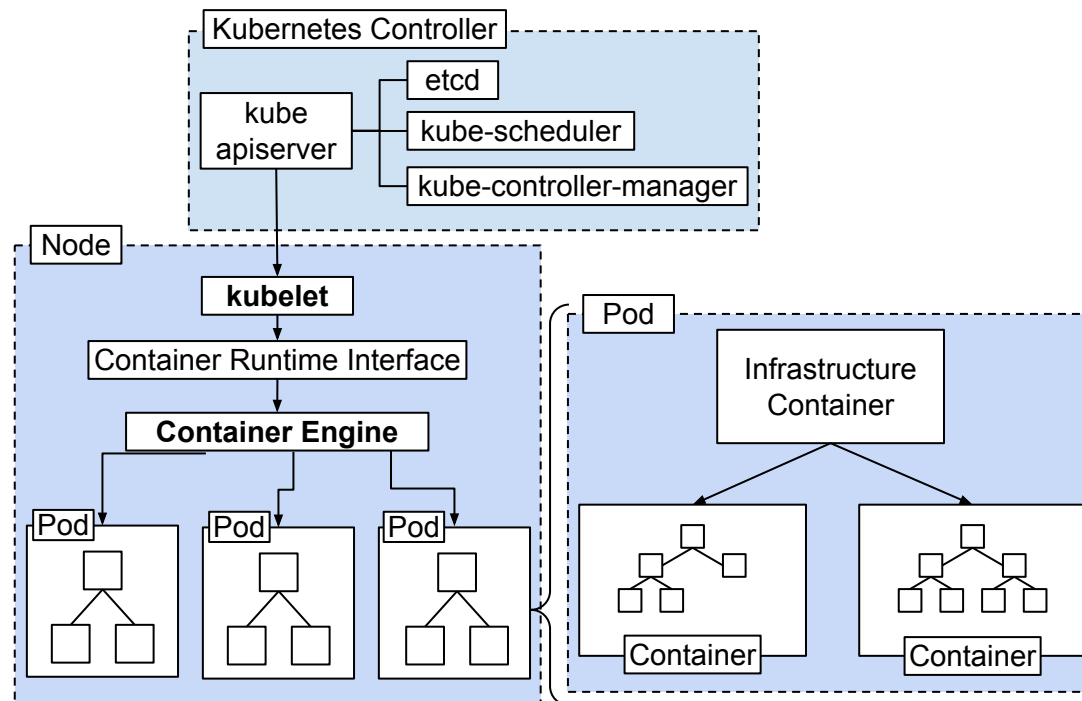
An alpha feature in Kubernetes v1.25

<https://kubernetes.io/blog/2022/12/05/forensic-container-checkpointing-alpha/>

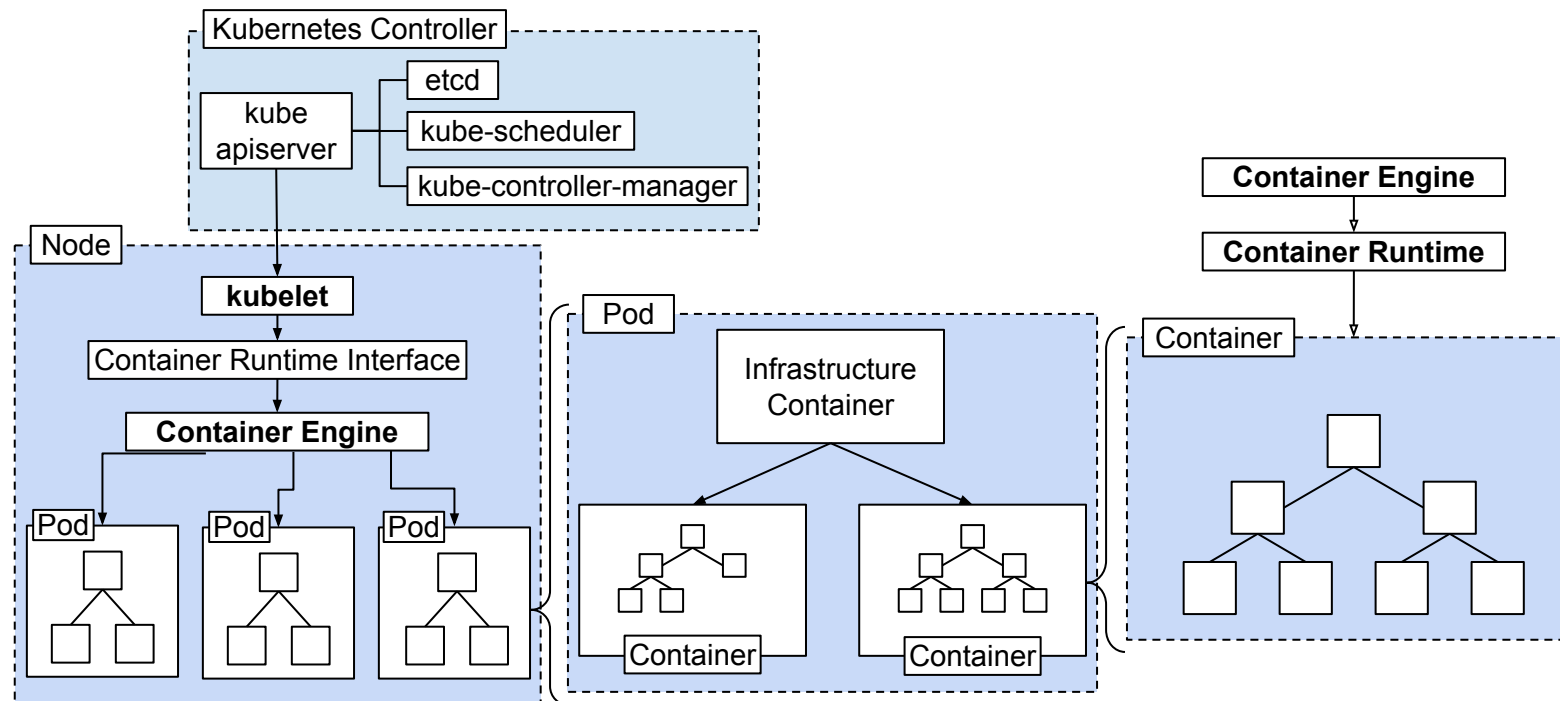
Container Checkpointing in Kubernetes



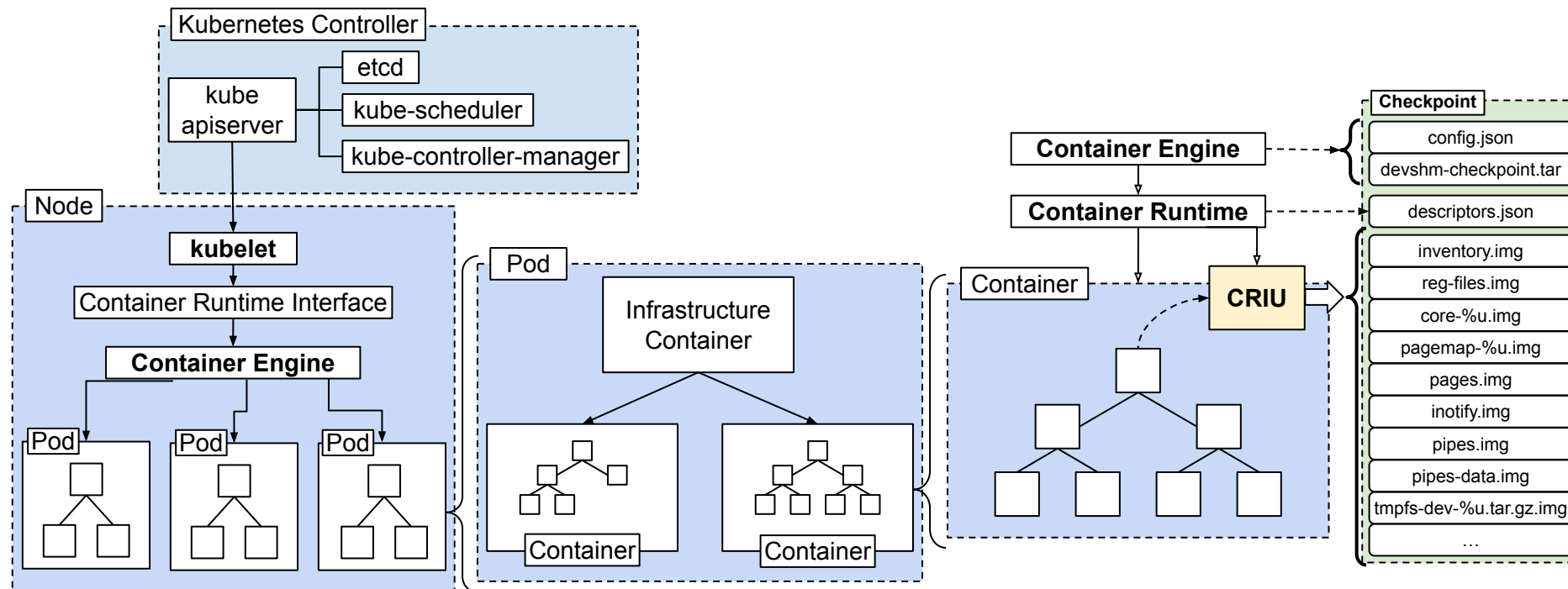
Container Checkpointing in Kubernetes



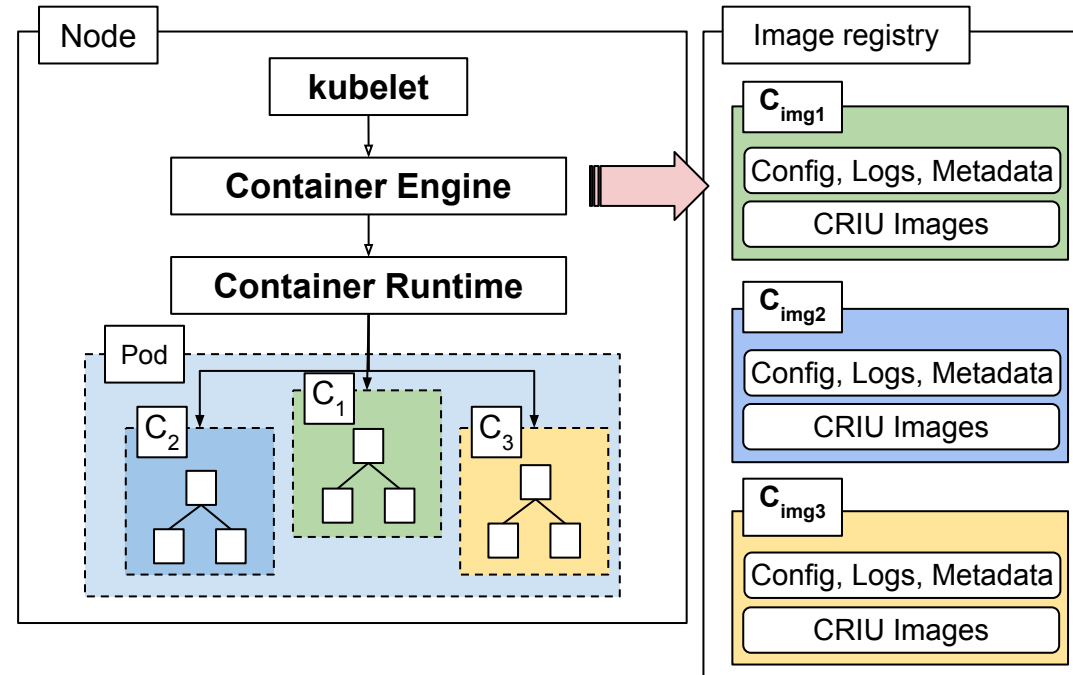
Container Checkpointing in Kubernetes



Container Checkpointing in Kubernetes



Checkpoint Images



Mar, 2022. Radostin Stoyanov. Add support for checkpoint image. <https://github.com/containers/podman/pull/13505>

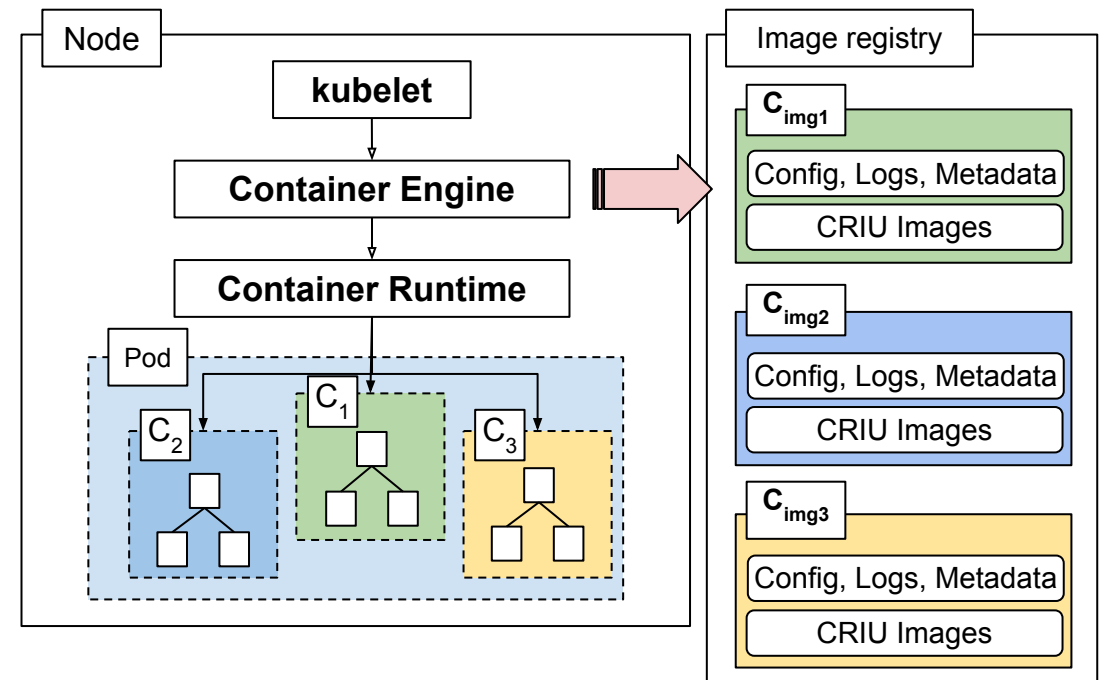
Aug, 2022. Adrian Reber. Support checkpointing to OCI image. <https://github.com/cri-o/cri-o/pull/6181>

Nov, 2022. Radostin Stoyanov. Add support for checkpoint images with 'podman run'. <https://github.com/containers/podman/pull/16569>

Dec, 2022. Adrian Reber. Forensic container checkpointing in Kubernetes. <https://kubernetes.io/blog/2022/12/05/forensic-container-checkpointing-alpha/>

Checkpoint Images

```
$ podman container checkpoint --create-image <img>
```



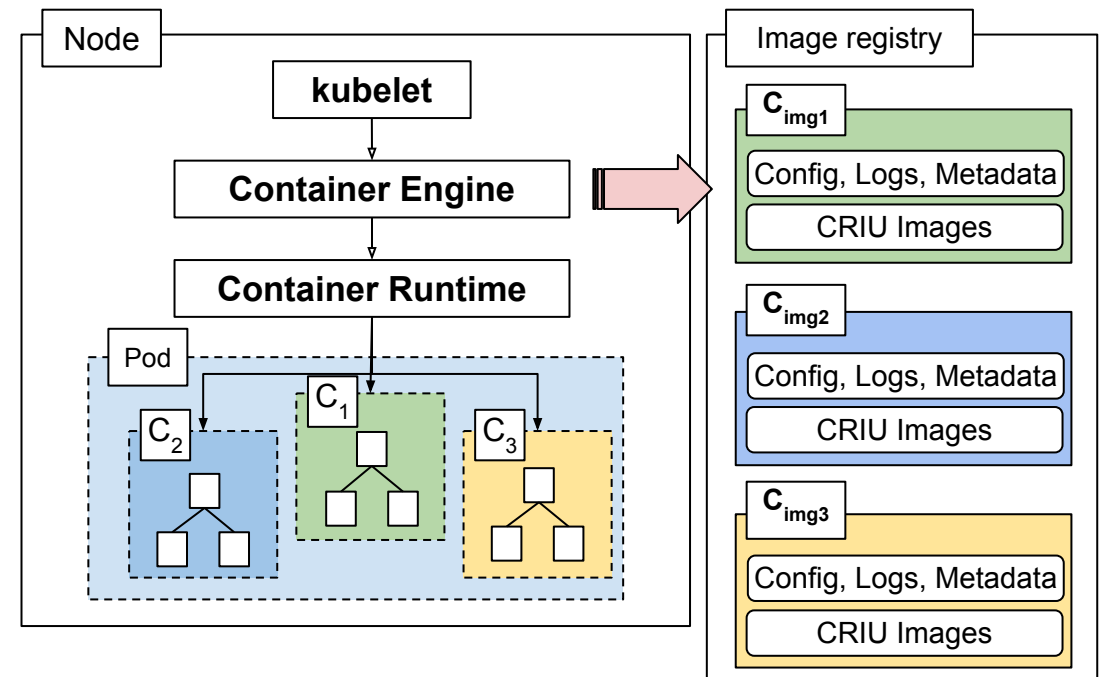
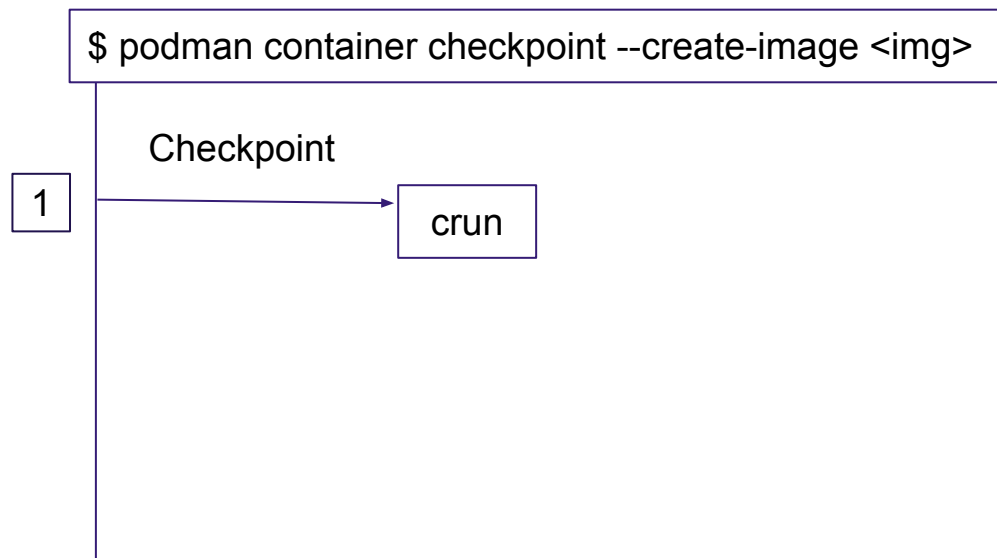
Mar, 2022. Radostin Stoyanov. Add support for checkpoint image. <https://github.com/containers/podman/pull/13505>

Aug, 2022. Adrian Reber. Support checkpointing to OCI image. <https://github.com/cri-o/cri-o/pull/6181>

Nov, 2022. Radostin Stoyanov. Add support for checkpoint images with 'podman run'. <https://github.com/containers/podman/pull/16569>

Dec, 2022. Adrian Reber. Forensic container checkpointing in Kubernetes. <https://kubernetes.io/blog/2022/12/05/forensic-container-checkpointing-alpha/>

Checkpoint Images



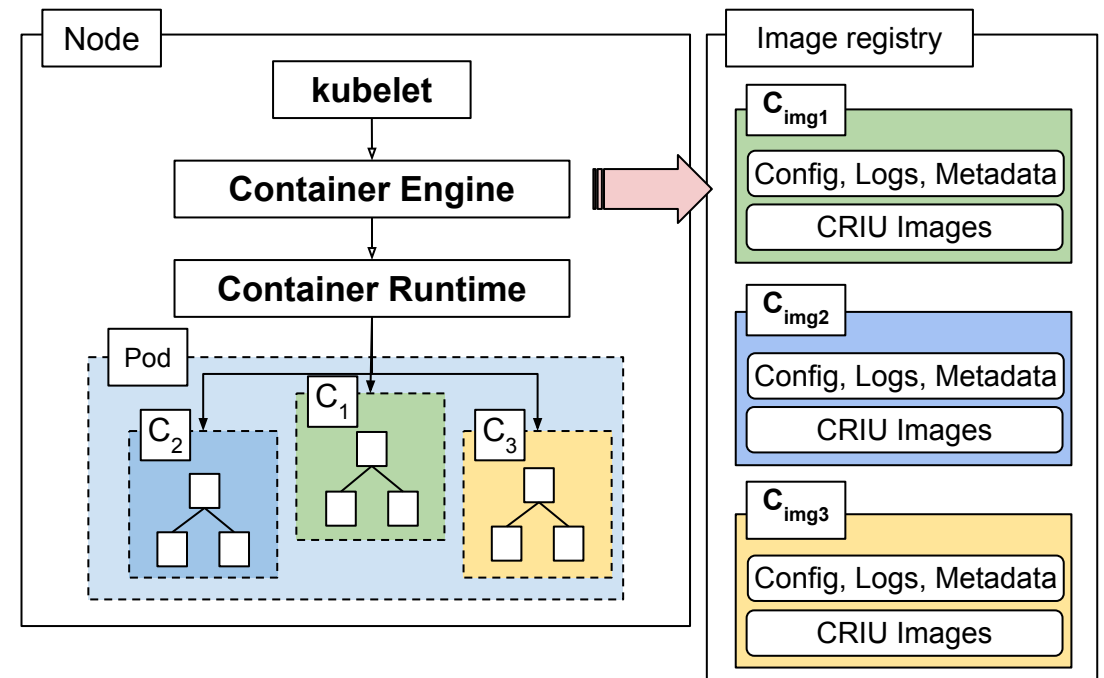
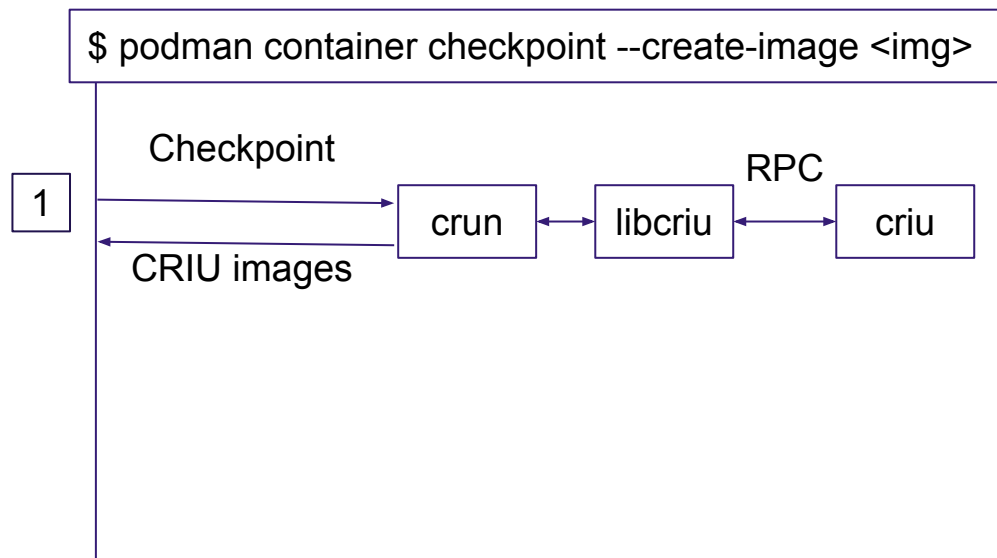
Mar, 2022. Radostin Stoyanov. Add support for checkpoint image. <https://github.com/containers/podman/pull/13505>

Aug, 2022. Adrian Reber. Support checkpointing to OCI image. <https://github.com/cri-o/cri-o/pull/6181>

Nov, 2022. Radostin Stoyanov. Add support for checkpoint images with 'podman run'. <https://github.com/containers/podman/pull/16569>

Dec, 2022. Adrian Reber. Forensic container checkpointing in Kubernetes. <https://kubernetes.io/blog/2022/12/05/forensic-container-checkpointing-alpha/>

Checkpoint Images



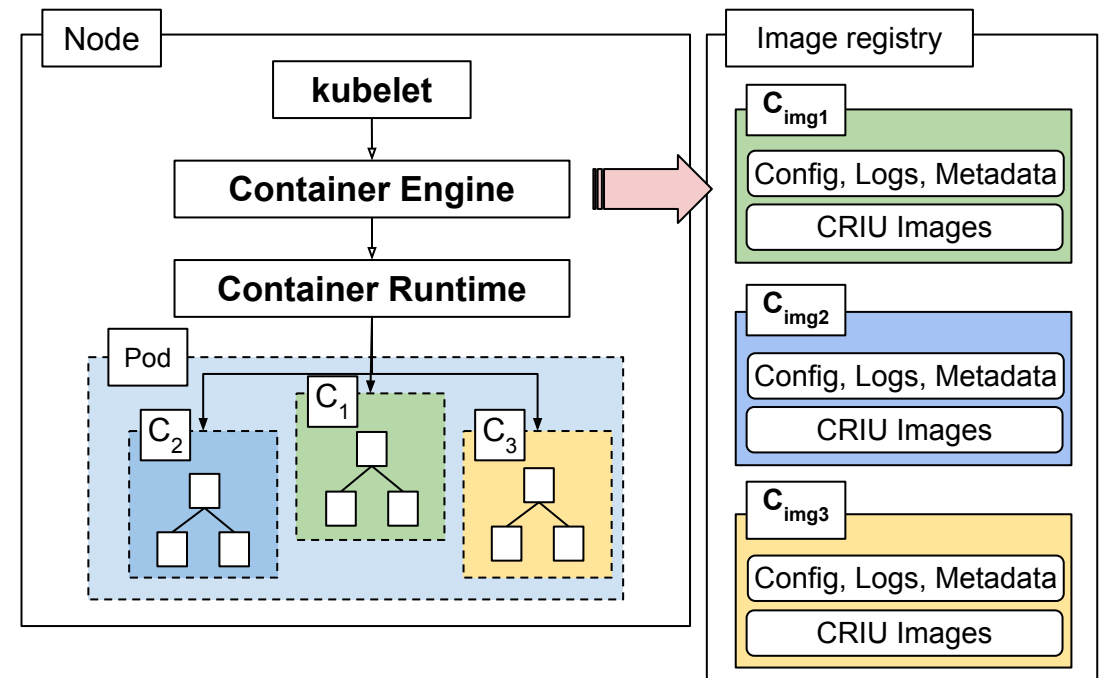
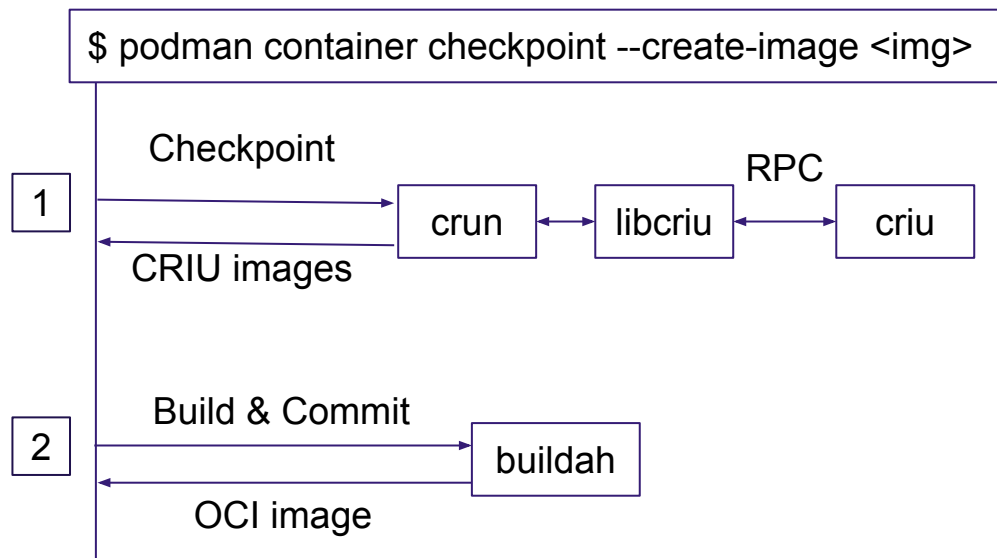
Mar, 2022. Radostin Stoyanov. Add support for checkpoint image. <https://github.com/containers/podman/pull/13505>

Aug, 2022. Adrian Reber. Support checkpointing to OCI image. <https://github.com/cri-o/cri-o/pull/6181>

Nov, 2022. Radostin Stoyanov. Add support for checkpoint images with 'podman run'. <https://github.com/containers/podman/pull/16569>

Dec, 2022. Adrian Reber. Forensic container checkpointing in Kubernetes. <https://kubernetes.io/blog/2022/12/05/forensic-container-checkpointing-alpha/>

Checkpoint Images



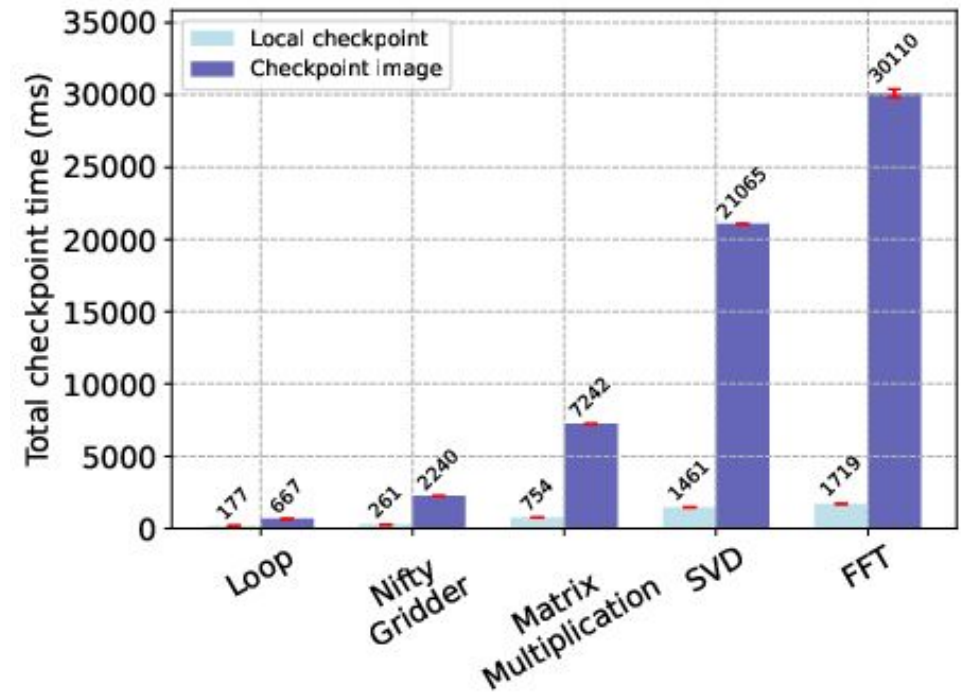
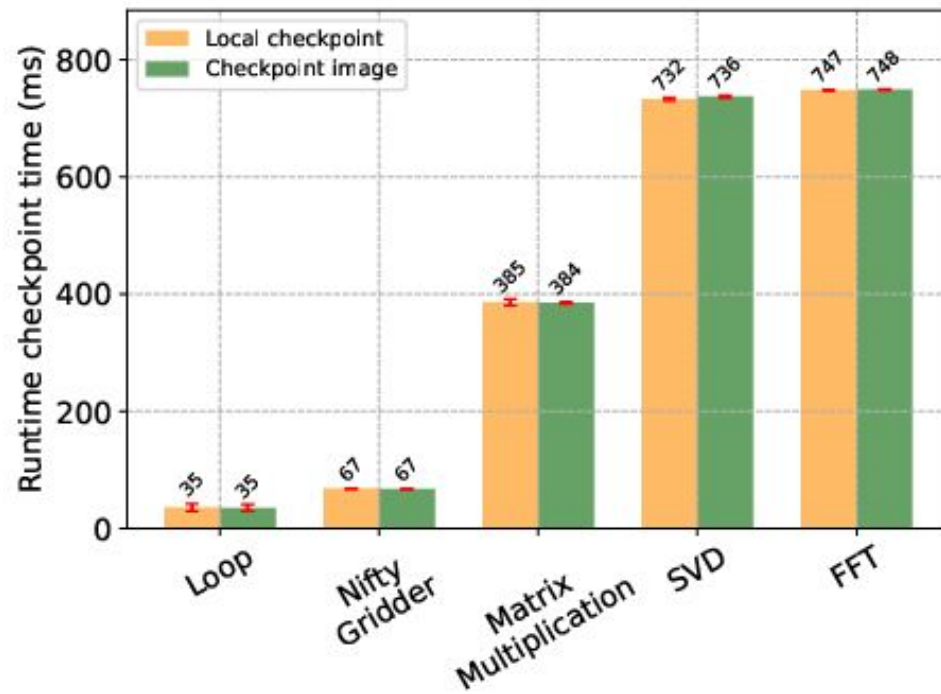
Mar, 2022. Radostin Stoyanov. Add support for checkpoint image. <https://github.com/containers/podman/pull/13505>

Aug, 2022. Adrian Reber. Support checkpointing to OCI image. <https://github.com/cri-o/cri-o/pull/6181>

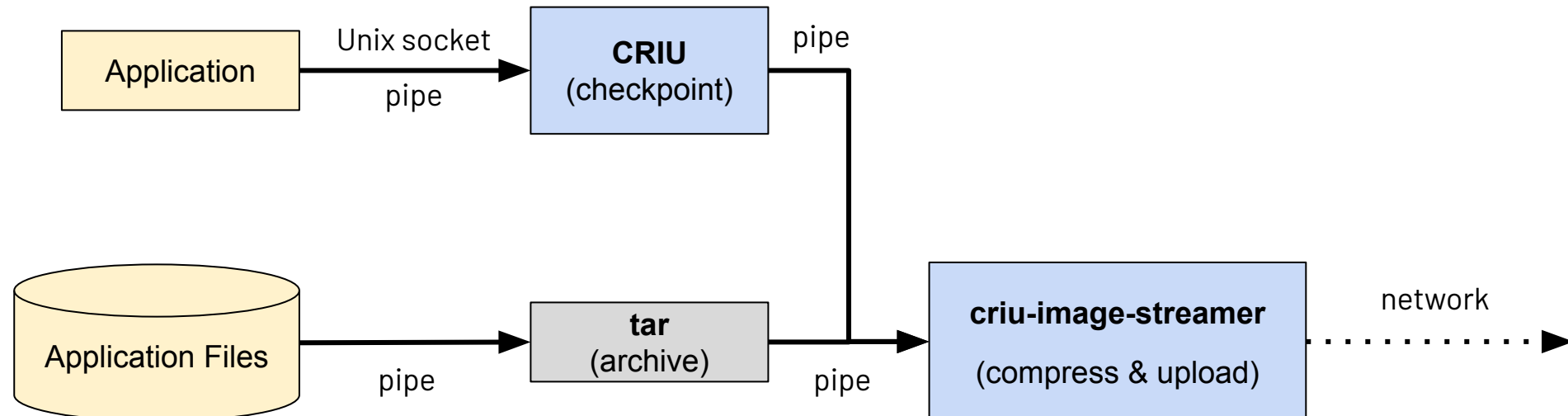
Nov, 2022. Radostin Stoyanov. Add support for checkpoint images with 'podman run'. <https://github.com/containers/podman/pull/16569>

Dec, 2022. Adrian Reber. Forensic container checkpointing in Kubernetes. <https://kubernetes.io/blog/2022/12/05/forensic-container-checkpointing-alpha/>

Performance Trade-offs



Checkpoint Streaming



<https://github.com/checkpoint-restore/criu-image-streamer>

Open Challenges & Research Directions



- Checkpoint/restore of *nested* namespaces
- Checkpoint/restore of hardware-accelerated applications
- Cluster scheduling with preemption and relocation

Summary & Questions

<https://criu.org/>

<https://github.com/checkpoint-restore/criu>

<https://kubernetes.io/blog/2022/12/05/forensic-container-checkpointing-alpha/>