

**containerd**

# Project Update and Deep Dive

Akihiro Suda, NTT

Kohei Tokunaga, NTT

Kirtana Ashok, Microsoft

Akhil Mohan, VMware by Broadcom



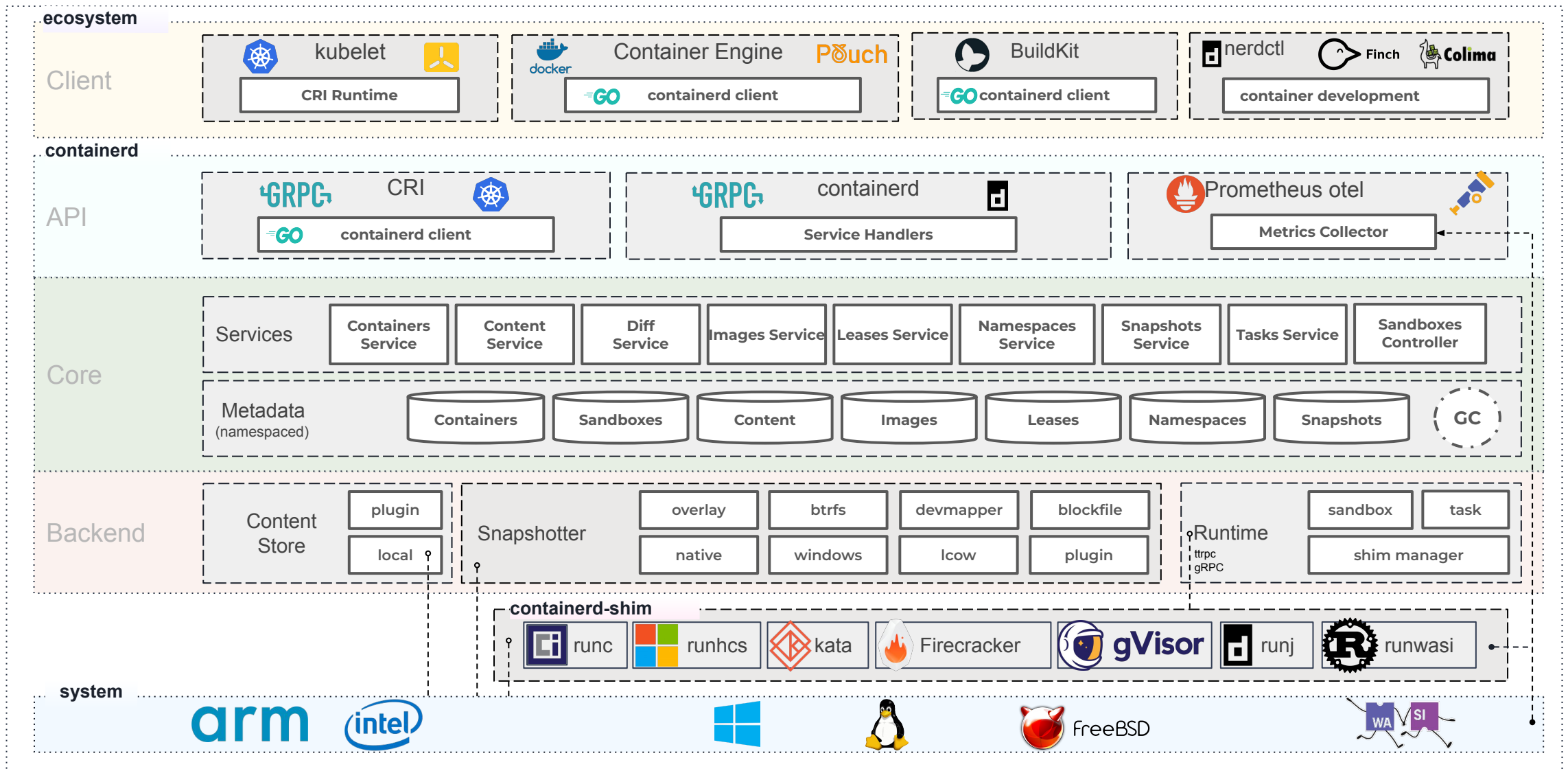
- Intro
  - v2.1 recap
  - New release cadence
  - Changes to the LTS model
  - New KEP Process
  - v2.2 expectations
  - Filesystems & Snapshotters
  - NRI
- 
- Maintainer panel discussion!

- The industry's standard container runtime, supporting both Docker and Kubernetes
- Adopted by well-known enterprise Kubernetes services and distributions
  - Amazon EKS, Azure Kubernetes Service, Google Kubernetes Engine, ...

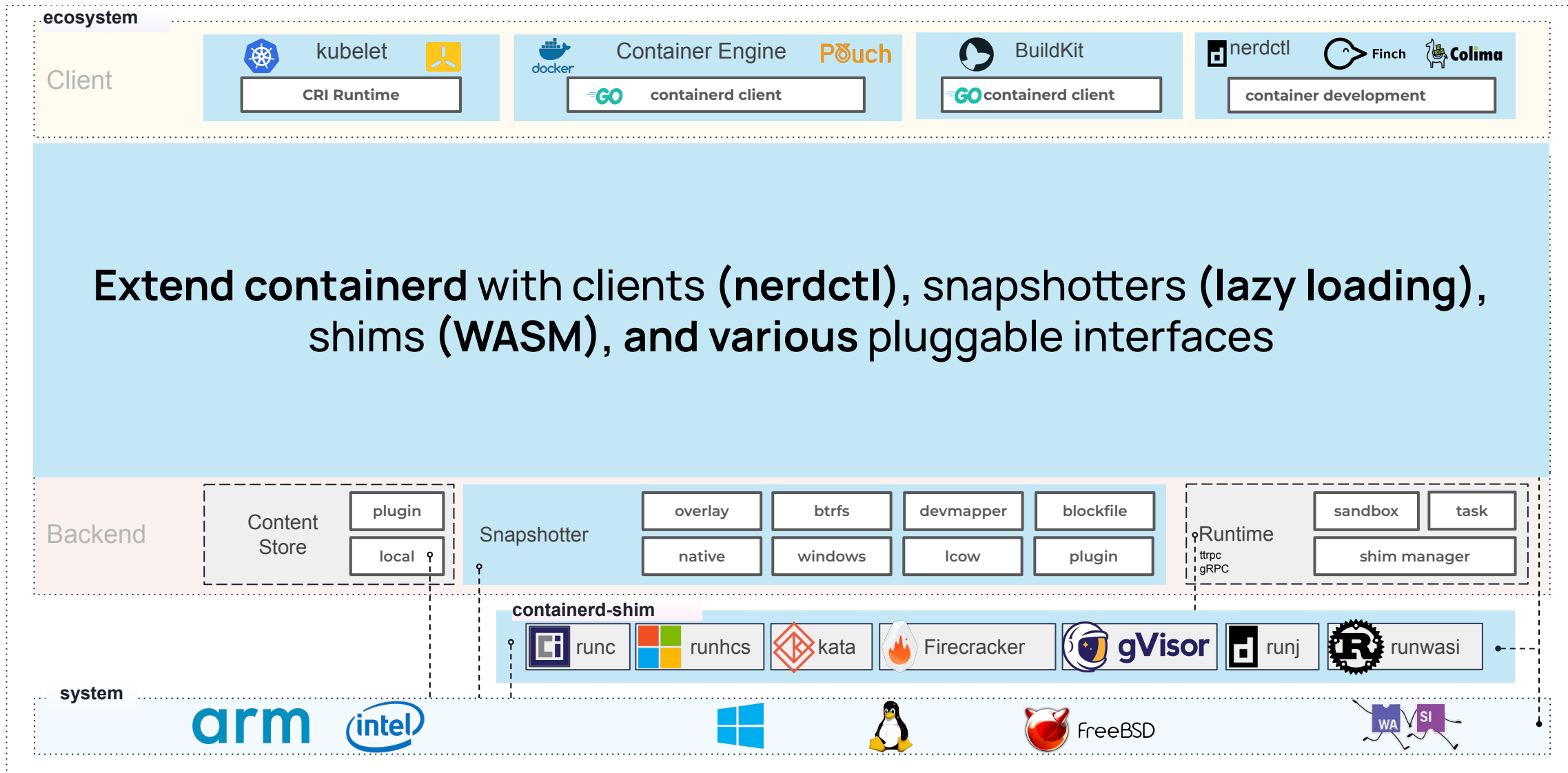
container 

<https://containerd.io/>

# Introduction to containerd



# Introduction to containerd



# v2.1 Recap

- The first release after the new release cadence was defined

<https://containerd.io/releases/#release-cadence>

<https://endoflife.date/containerd>

- List of changes:

<https://github.com/containerd/containerd/releases/tag/v2.1.0>

- EROFS (Enhanced Read-Only File System) ([#10705](#))
  - Needs linux kernel 6.12+
  - Suitable for images with many layers
- Image volumes ([KEP 4639](#)) ([#10579](#))
  - Allow mounting a container image as a Kubernetes volume
  - Split data image from code image
  - Example: Share config files, machine learning models, binary artifacts between containers in a pod.



- Writable cgroups for non-privileged containers ([#11131](#))
  - Allows containers to manage their own cgroups (CPU, memory, ...)
  - Better resource isolation and management.
- More flexible support for UserNS-remap ([#10307](#), [#10722](#))
  - Allows multiple UID/GID entries.

```
ctr run --uidmap 0:100000:1000 --uidmap 1000:201795:1 \  
--uidmap 1001:10000000:64535 --gidmap 0:100000:100 \  
--gidmap 100:1001:1 --gidmap 101:10000000:65435 \  
--user foo <image> test bash
```
  - The UID mapping range can be now non-contiguous  
e.g., uidmap=0:666:1000,1000:6666:64536

- Update CRI to use transfer service by default
- Container restore support added through CRI
- Removed support for schema1 images and go plugins.

# New Release Cadence

# New release cadence!

- New minor releases (2.x) every 6 months
- Starting with the 2.1 release in **May** and 2.2 in **November**  
(and continuing May and November after that)
- Beta builds start 8-10 weeks before a minor release (March and September)
- Feature freeze and RCs start 2-4 weeks before a minor release  
(April and October)

2025/05	2.1 release!
2025/06 - 2025/08	Development focused on 2.2
2025/09	2.2 Beta
2025/10	2.2 feature freeze and RC
2025/11	2.2 release!

# LTS Model and KEP Process

# Changed Long-Term Supported model

- No periodic cadance to ship LTS releases
- An existing release is promoted to LTS when volunteer maintainers step in
- LTS releases are supported for a minimum of 2 years
  - Can be longer at discretion of branch owners
- containerd 1.6 EOL soon
- containerd 1.7 is now LTS
- containerd 2.x LTS is yet to be determined

Version	Status	Release	EOL
1.6	LTS	2022-02-15	2025-07-23
1.7	LTS	2023-03-10	2026-03-10
2.0	Active	2024-11-05	2025-11-08
2.1	Active	2025-05-07	2026-05-05 (tentative)

## KEP: Kubernetes Enhancement Proposals

### Problems

- poor visibility of proposals
- lack of communication across Kubernetes maintainers and containerd maintainers

### Recent Improvements:

- Tracking issues in containerd/containerd repo
- “KEP shepherd” containerd maintainer to help move PRs along
- “SIG-Node liaison” containerd maintainer to improve the containerd project <-> SIG-Node communication

# v2.2 Expectations



- Quota for snapshots (overlayfs etc.) ([#5859](#))
- New mount management interface, with garbage collection ([#11303](#))
- CRI: Pause-less pod sandboxes ([#10505](#))
- CRI: Image pull per runtime class ([#11807](#), [KEP 4216](#))
- nerdctl: HEALTHCHECK support ([nerdctl#4302](#))



KubeCon



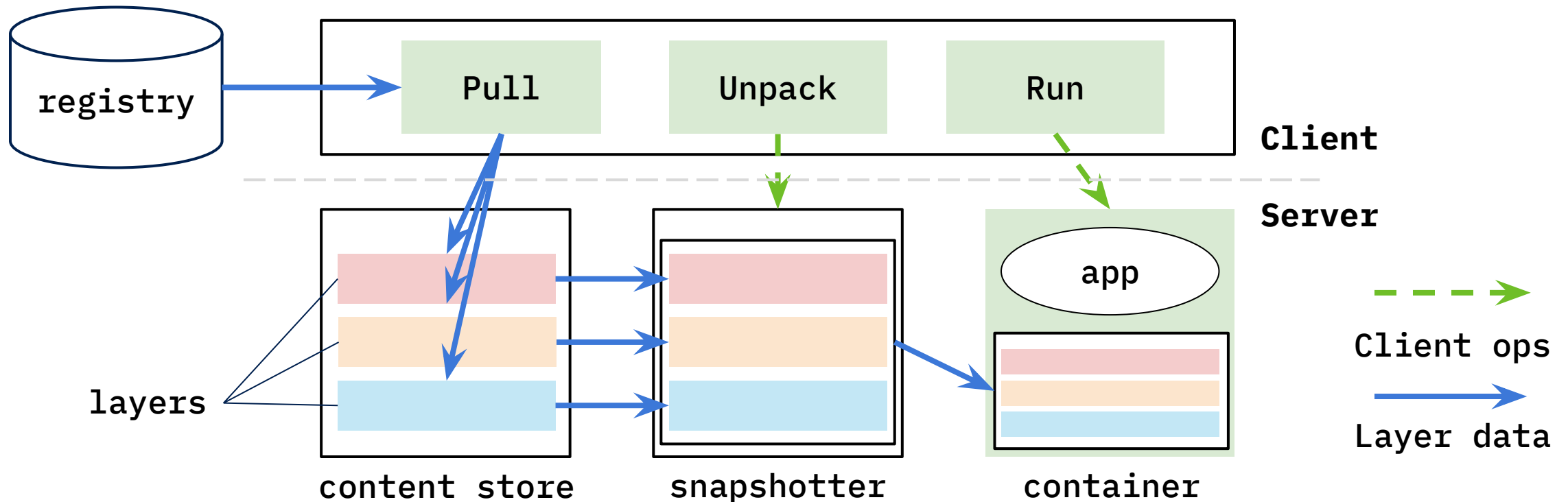
CloudNativeCon

Japan 2025

# containerd Snapshotters

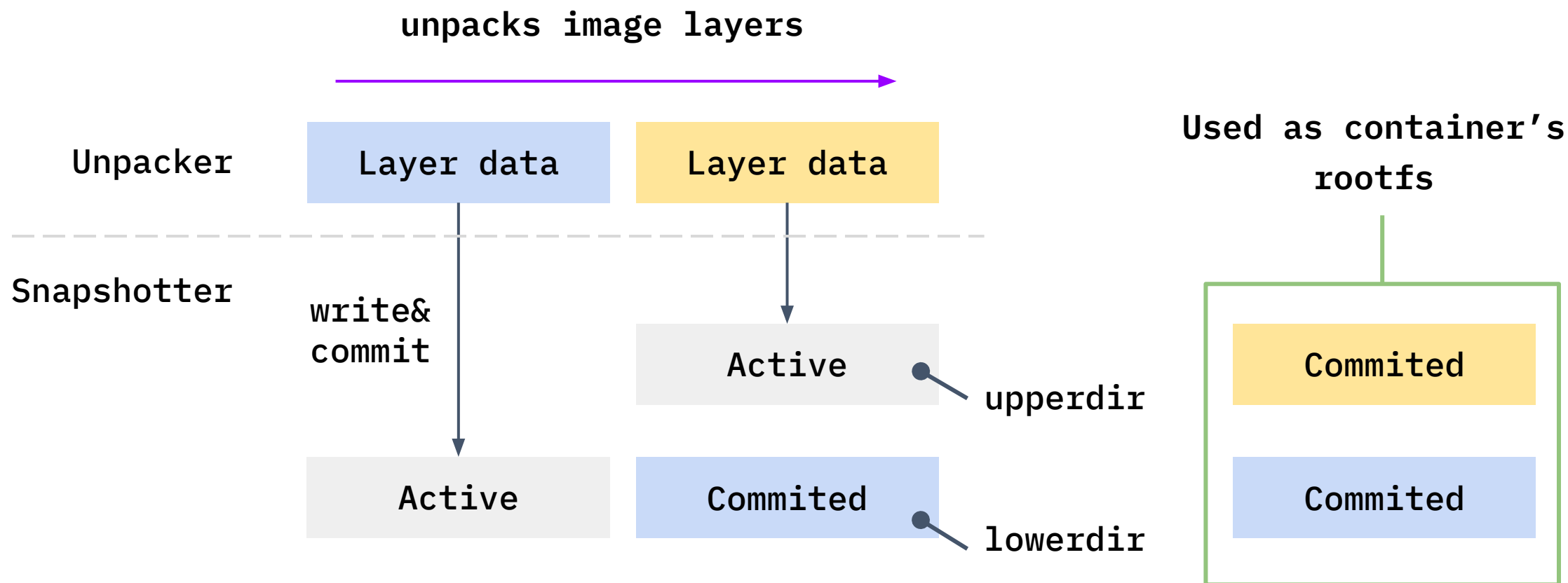
# containerd Snapshotters

- Manages unpacked image layers = snapshots
- Supports a variety of filesystems, including overlayfs, erofs, btrfs, devmapper, etc.



# Example: overlayfs snapshotter

- Uses Linux's overlayfs to stack layers
- Default snapshotter on Linux

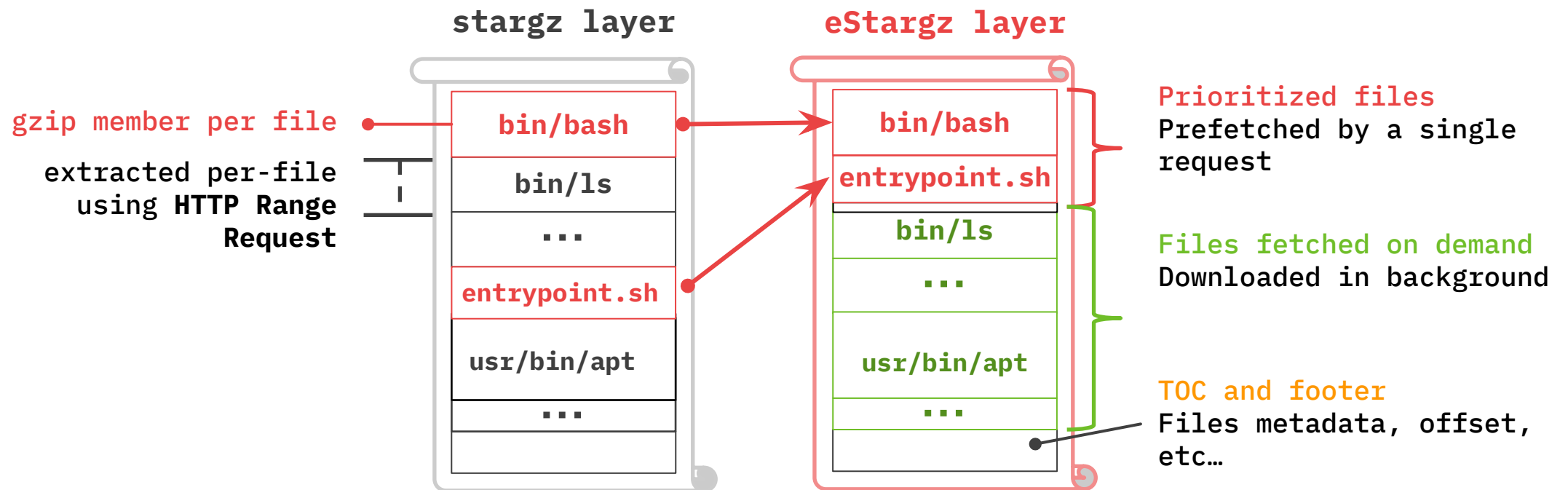


- Bring your own snapshotter, plugging it into containerd via gRPC
- Stargz Snapshotter: <https://github.com/containerd/stargz-snapshotter>
- SOCI Snapshotter: <https://github.com/awslabs/soci-snapshotter>
- Nydus Snapshotter: <https://github.com/containerd/nydus-snapshotter>
- OverlayBD Snapshotter: <https://github.com/containerd/overlaybd>
- CVMFS Snapshotter:  
<https://github.com/cvmfs/cvmfs/tree/devel/snapshotter>

<https://github.com/containerd/stargz-snapshotter>

- Snapshotter plugin with **lazy pulling**
  - Starting a container before the image is fully downloaded
- Uses OCI-compatible eStargz image format
  - FUSE-mounted from the registry to the node
- Enables containerd to use eStargz mounts instead of fully pulling layers

- OCI-compatible image format for lazy pulling
- Extended Google's [stargz by CRFS project](#) to support prefetching
- 100% compatible to the current OCI Image format
  - i.e. Usable as a normal image on eStargz-unaware runtimes



# Lazy pulling performance of eStargz

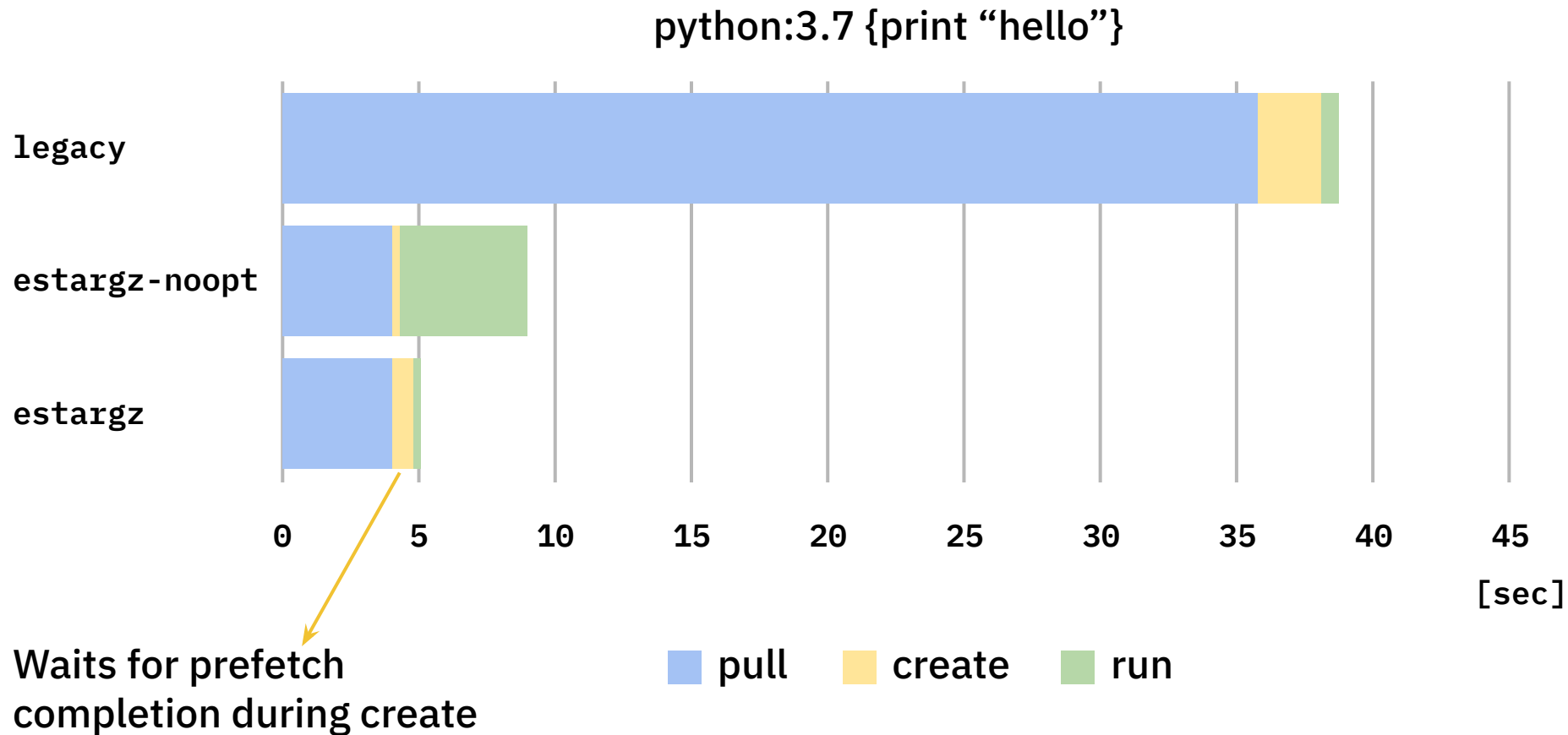


Figure from “Faster Container Image Distribution on a Variety of Tools with Lazy Pulling” - Kohei Tokunaga & Tao Peng. KubeCon+CloudNativeCon North America 2021. <https://sched.co/lV2a>





KubeCon

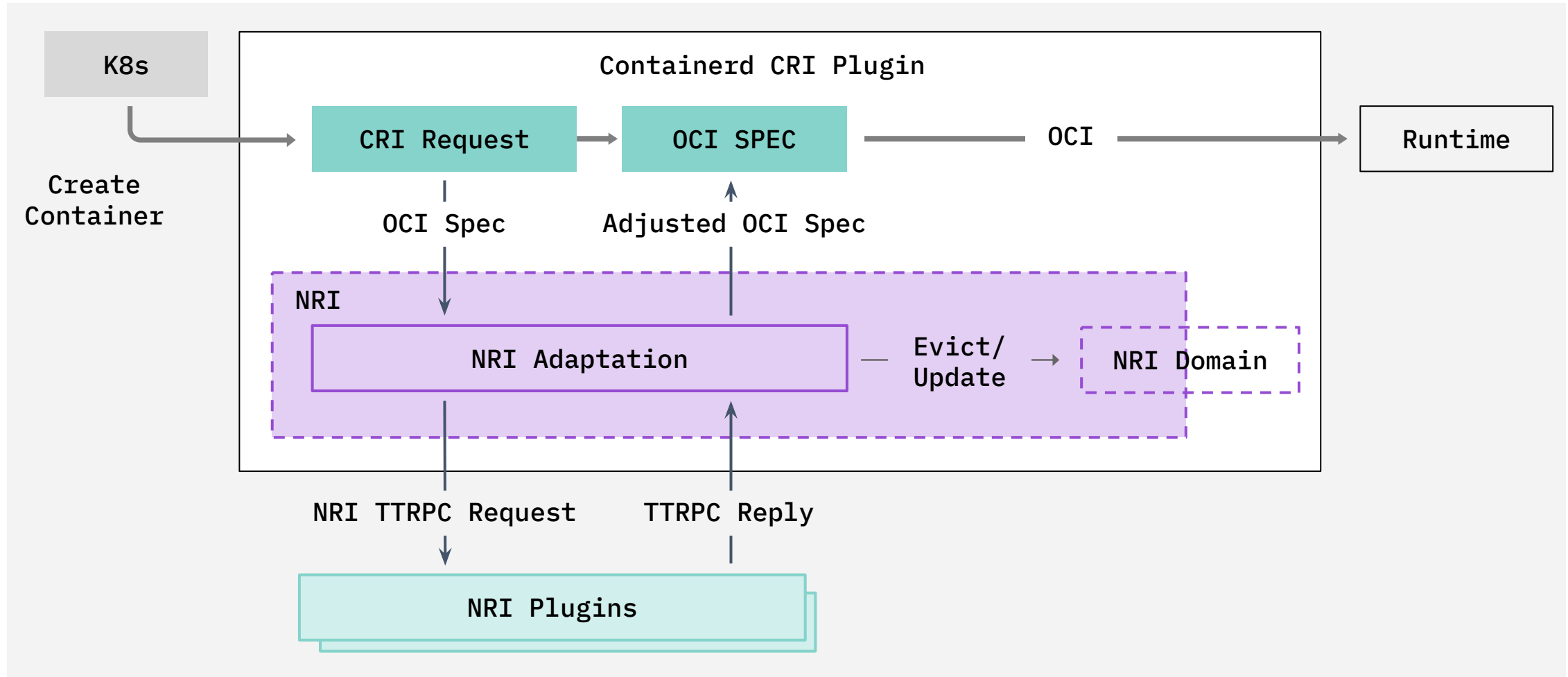


CloudNativeCon

Japan 2025

# Node Resource Interface (NRI)

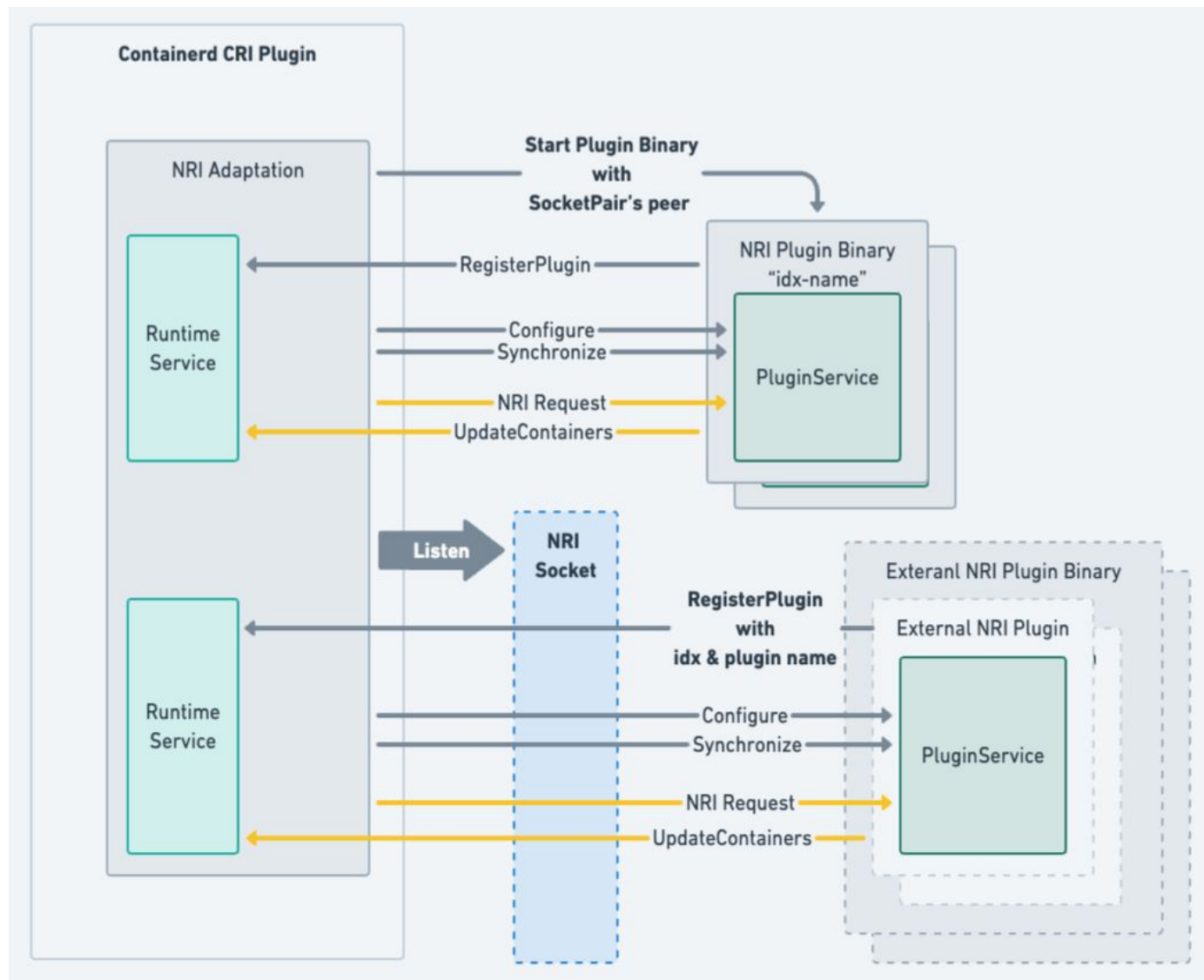
## Middleware extension between CRI and OCI



```
[plugins.'io.containerd.nri.v1.nri']  
# Disable NRI support in containerd  
disable = false  
# path of NRI socket to create for plugins to connect to  
socket_path = '/var/run/nri/nri.sock'  
# Directory to search for plugins to launch on startup  
plugin_path = '/opt/nri/plugins'  
# Directory to search for plugin specific configuration  
plugin_config_path = '/etc/nri/conf.d'  
# timeout for a plugin to register after connection  
plugin_registration_timeout = '5s'  
# timeout for a plugin to handle an event/request  
plugin_request_timeout = '2s'  
# Allow connections from externally launched NRI plugins  
disable_connections = false
```

## Plugin Registration

- Plugin Binary
- External Plugin via socket support



## Maintained at containerd/nri/plugins

- ulimit adjuster
- network device injector
- device injector

- Policy controls
- Improved Observability
- Consistent Failure Characteristics

# Panel discussion



KubeCon



CloudNativeCon

Japan 2025

# How did you get involved with containerd?





KubeCon



CloudNativeCon

Japan 2025

# What is something you are excited about in the container space?

- **GitHub:** <https://github.com/containerd>
- **Slack:** <https://slack.cncf.io/>  
(Channels: `#containerd` `#containerd-dev`)
- **X (former Twitter):** [@containerd](https://twitter.com/containerd)
- **Bluesky:** [@containerd.dev](https://bsky.app/profile/containerd.dev)
- Community meetings on [cncf.io/calendar](https://cncf.io/calendar)



**KubeCon**



**CloudNativeCon**

**Japan 2025**