# containerd v2.0
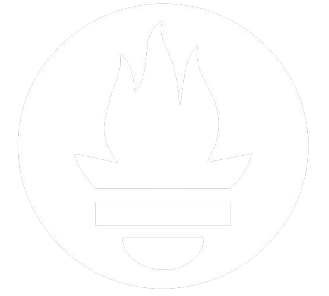
## An industry-standard container runtime
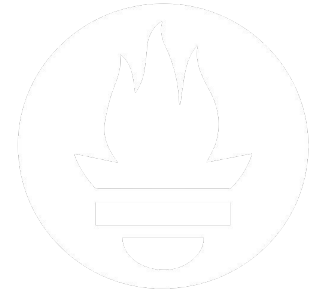
- CNCF Graduated since 2019

- Adopted by well-known enterprise Kubernetes services

  and distributions

- Designed to be pluggable to support cutting-edge features

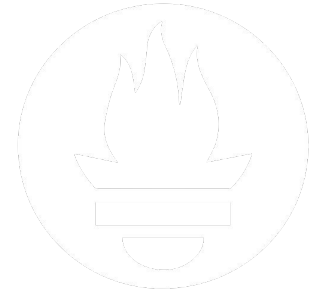  (remote snapshotters, WASM runtimes, …)

# Runtime plugins

- **runc**: the regular runtime for Linux

- **runhcs**: Windows

- **runj**: FreeBSD jail

- **runwasi**: WASM

- **kata**: VM

- **runsc (gvisor)**: ptrace sandbox, etc.

# Snapshotter plugins

Regular snapshotters: **overlayfs**, **btrfs**, **zfs**, **devmapper**, ...

"Remote" snapshotters support pulling image contents on demand to shorten the container startup time

- **stargz**: Forward compatible with OCI v1 tar.gz images
- **nydus**: Uses an alternate image format
- **overlaybd**: Uses block devices as container images

# containerd v2.0

- The focus is on cleaning up deprecated features

  (runtime v1 API, AUFS snapshotter, non-sandboxed pods…)

- Huge improvements on the support for User Namespaces

  - UserNS maps the root user in containers to a non-root user on the host to

    mitigate potential container break-out

  - Overhead for chowning is now eliminated

Snapshotter is overlayfs.

| Image | Size | Inodes | Chown without metacopy | Chown with metacopy | IDMapped Mount |
|---|---|---|---|---|---|
| tensorflow/tensorflow:2.13.0 | 1.6GiB | 37460 | 2m58.18635105s | 1.425115168s | 21.969618ms |
| library/node:20.11.1 | 1.1GiB | 33210 | 2m40.427234391s | 1.047708275s | 17.713161ms |
| library/ubuntu:22.04 | 84MiB | 3517 | 14.986225994s | 101.244062ms | 9.268065ms |

4

# Deprecation Warning

- ctr deprecations list

  - Supported in v1.6.27+, v1.7.12+

  - For each warning, you can see

    - What is being removed

    - A suggestion for migration

    - When it was last used
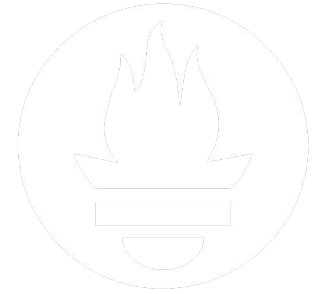
Deprecation Warnings in containerd - Getting Ready for 2.0!

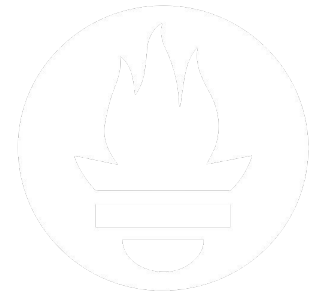# nerdctl: contaiNERD CTL

- https://github.com/containerd/nerdctl
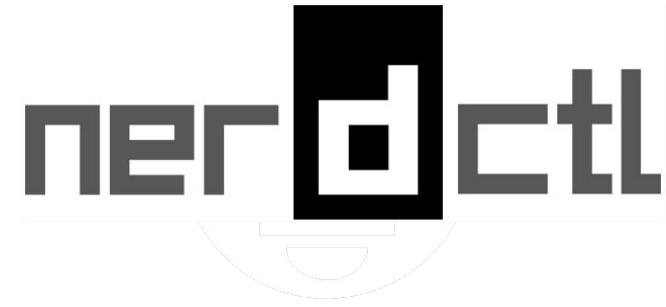
- Same UI/UX as the docker CLI (including Compose)

- Made for facilitating new experiments in the containerd platform (e.g., stargz, fast rootless with bypass4netns)

- Useful for debugging Kubernetes nodes too

```
$ nerdctl run hello-world
$ nerdctl compose up
```

# nerdctl v2.0

- Rootless networking is refactored to support "detach-netns"
  - Faster pull/push
    - Previously limited to less than 10Gbps due to user-mode TCP/IP; Now as fast as rootful
    - Containers can be accelerated too with bypass4netns (experimental)
  - Proper support for --net=host
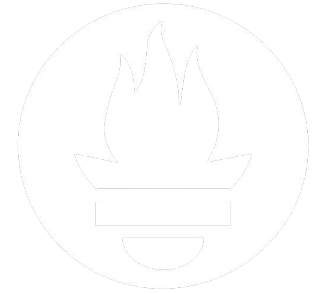  - Proper support for localhost registries

# nerdctl v2.0
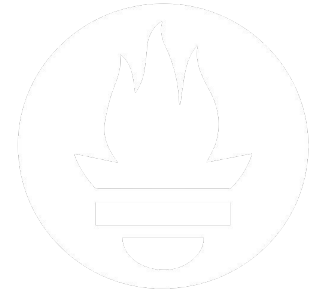
- Support running systemd in a container without --privileged
    - "Plain old VM"-like user experience
    - Often considered to be an anti-pattern
    - Useful for testing, etc.

```
$ nerdctl run --systemd=true ...
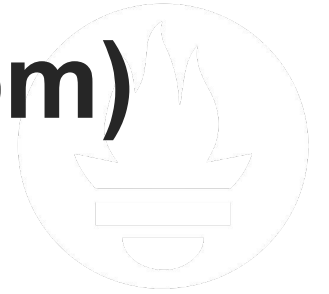```

- Massive refactoring and testing

# Join the community!

- **GitHub:** https://github.com/containerd
- **Slack**:   https://slack.cncf.io/
  (Channels: #containerd #containerd-dev)
- **X (former Twitter)**: @containerd
- **Bluesky:** @containerd.dev

# Visit us again on Friday (2:55pm-3:30pm)

## What's Going on in the Containerd Neighborhood?

Phil Estes, AWS

Samuel Karp, Google

Akihiro Suda, NTT

Michael Brown, IBM

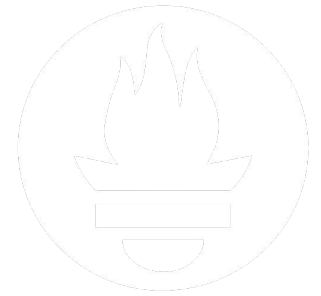Kirtana Ashok, Microsoft

Hyatt Regency | Level 4 | Regency Ballroom A

# Learn about containerd 2.0
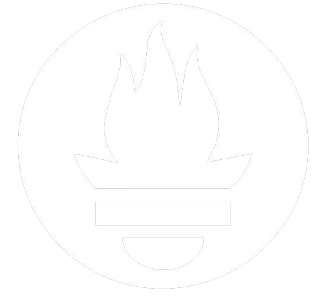
**[What containerd 2.0 means for you](#)**

**Samuel Karp, Google**

**Friday, 11:55am - 12:30pm**

**Salt Palace | Level 2 | 254 B**

# Advanced topics

# New features, newly stable features, defaults in v2.0

- Transfer service <span style="color:red">NOW STABLE</span>
- Sandbox service (and sandboxed CRI) <span style="color:red">NOW STABLE</span>
- Faster image extraction with igzip <span style="color:red">NEW</span>
- Improved OTEL configuration <span style="color:red">NEW</span>
- NRI enabled by default <span style="color:red">NEW</span>
- Image verifier plugins <span style="color:red">NEW</span>
- Plugin introspection <span style="color:red">NEW</span>
- CDI enabled by default <span style="color:red">NEW</span>
- CRI support for user namespaces <span style="color:red">NEW</span>

# Highlight: Node Resource Interface (NRI)

- Akin to a mutating webhook, but for container configuration
  - Middleware between CRI and OCI
- Use cases
  - Injection (devices, network devices, OCI hooks)
  - Resource modification/management (ulimits, topology/NUMA, advanced QoS, SGX memory)
  - Policy enforcement
- Plugins can run in containers or as system services
- Enabled by default
- Community plugins
  - https://github.com/containerd/nri/tree/main/plugins
  - https://github.com/containers/nri-plugins

# Highlight: Image verifier plugins
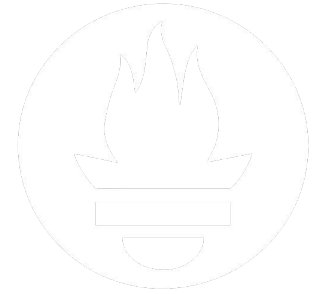
- Exec-based plugins containerd invokes during image pull

- Policy enforcement use-cases

  - Container image signature verification

  - Trust for particular signers

  - Allow only specific registries/repositories

- Integrated with the Transfer service (not supported for legacy pulls)