KubeCon | CloudNativeCon
North America 2023

Pods and Circumstance: CRI-O’s Graduation Celebration

Peter Hunt & Sascha Grunert

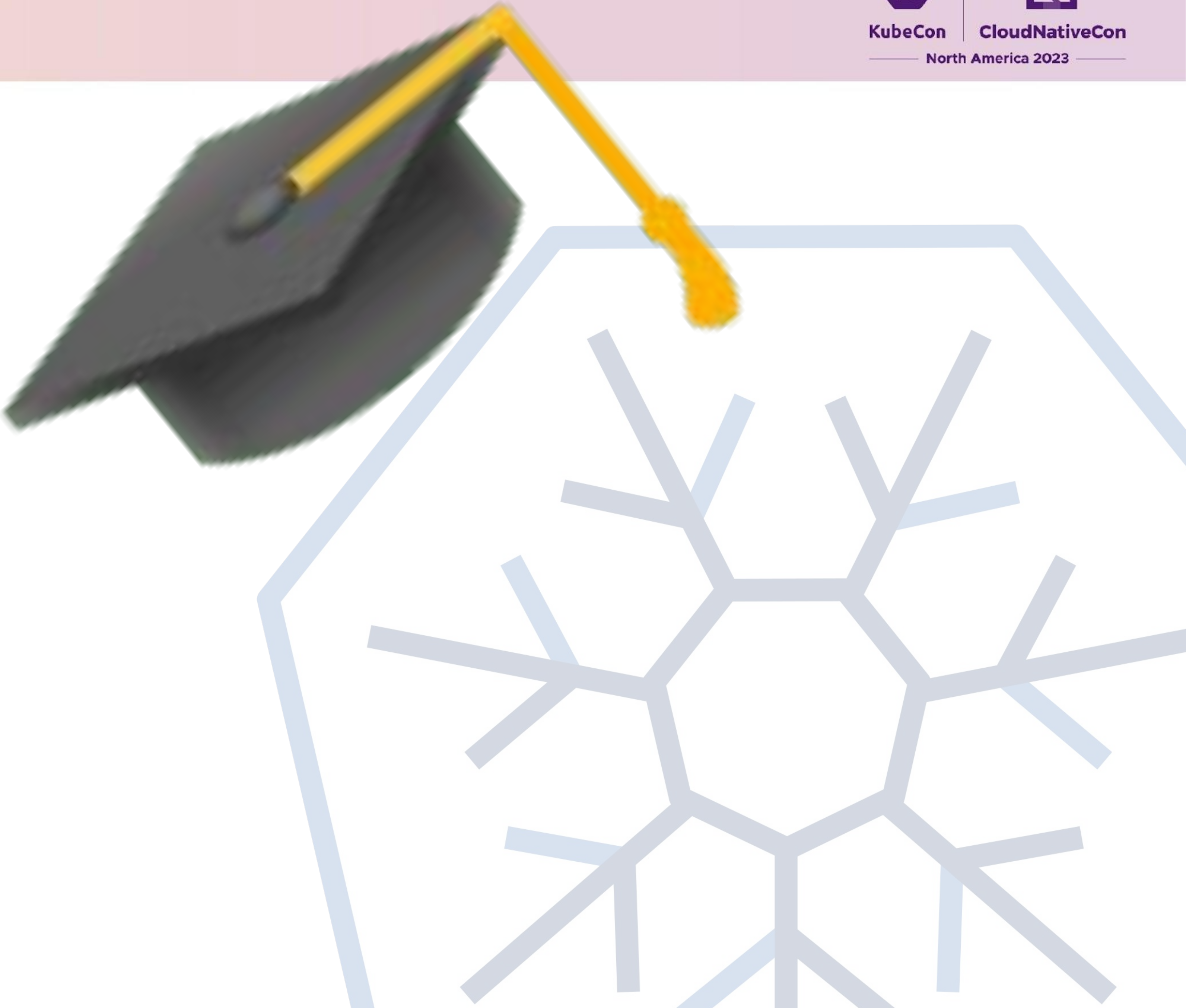
Image by pikisuperstar on Freepik

Contents

1. How to graduate a project in the CNCF
2. Getting new maintainers: CRI-O mentorship programs
3. Container Runtime Interface metrics enhancements
4. What’s new about conmon-rs?
5. Sigstore signature verification
6. Recent enhancement to CRI-O’s packaging efforts
7. What we plan for the future

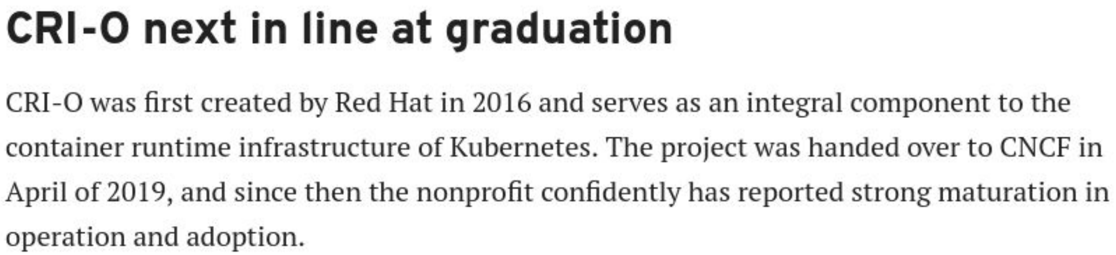
CRI-O Graduated!

CRI-O Graduated!



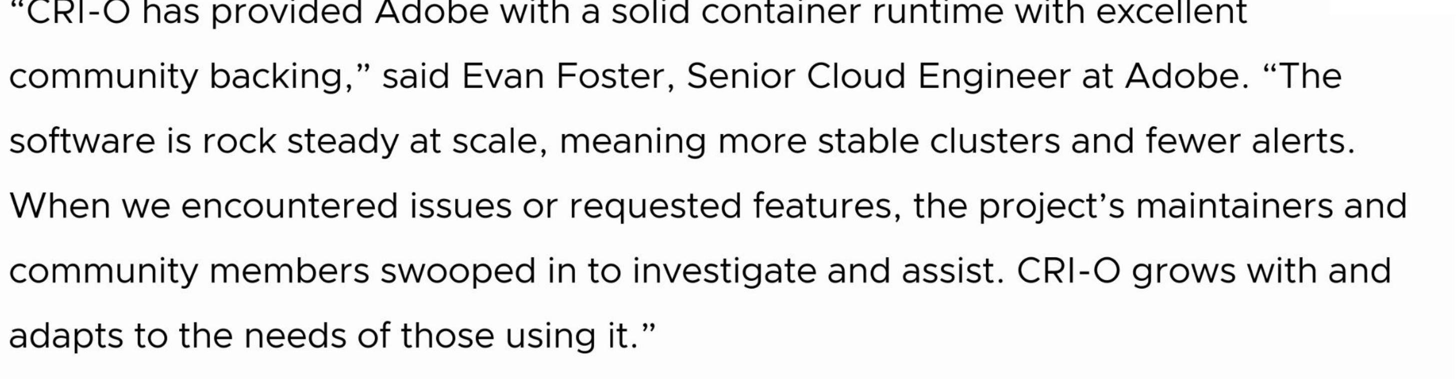
"CRI-O has remained focused on creating a simple and lightweight container runtime optimized for Kubernetes only in large-scale production environments," said Chris Aniszczyk, CTO at CNCF. "At the end of the day, it's great to have options and competition in the container runtime space. We look forward to seeing even more achievements and growth from the project team as a graduated project."

ANNOUNCEMENT
Cloud Native Computing Foundation Announces Graduation of CRI-O
CLOUD NATIVE
COMPUTING FOUNDATION



CRI-O next in line at graduation

CRI-O was first created by Red Hat in 2016 and serves as an integral component to the container runtime infrastructure of Kubernetes. The project was handed over to CNCF in April of 2019, and since then the nonprofit confidently has reported strong maturation in operation and adoption.



“CRI-O has provided Adobe with a solid container runtime with excellent community backing,” said Evan Foster, Senior Cloud Engineer at Adobe. “The software is rock steady at scale, meaning more stable clusters and fewer alerts. When we encountered issues or requested features, the project’s maintainers and community members swooped in to investigate and assist. CRI-O grows with and adapts to the needs of those using it.”

The Road to Graduation

- Have committers from at least two organizations
- Paperwork
  - CII best practices badge
  - GOVERNANCE.md/OWNERS.md/MAINTAINERS.md
  - List of adopters
- Have completed an independent and third party security audit
- Receive a supermajority vote from the TOC to move to graduation stage

"The Long and Windy Road" by Rodney Campbell is licensed under CC BY-NC-ND 2.0.

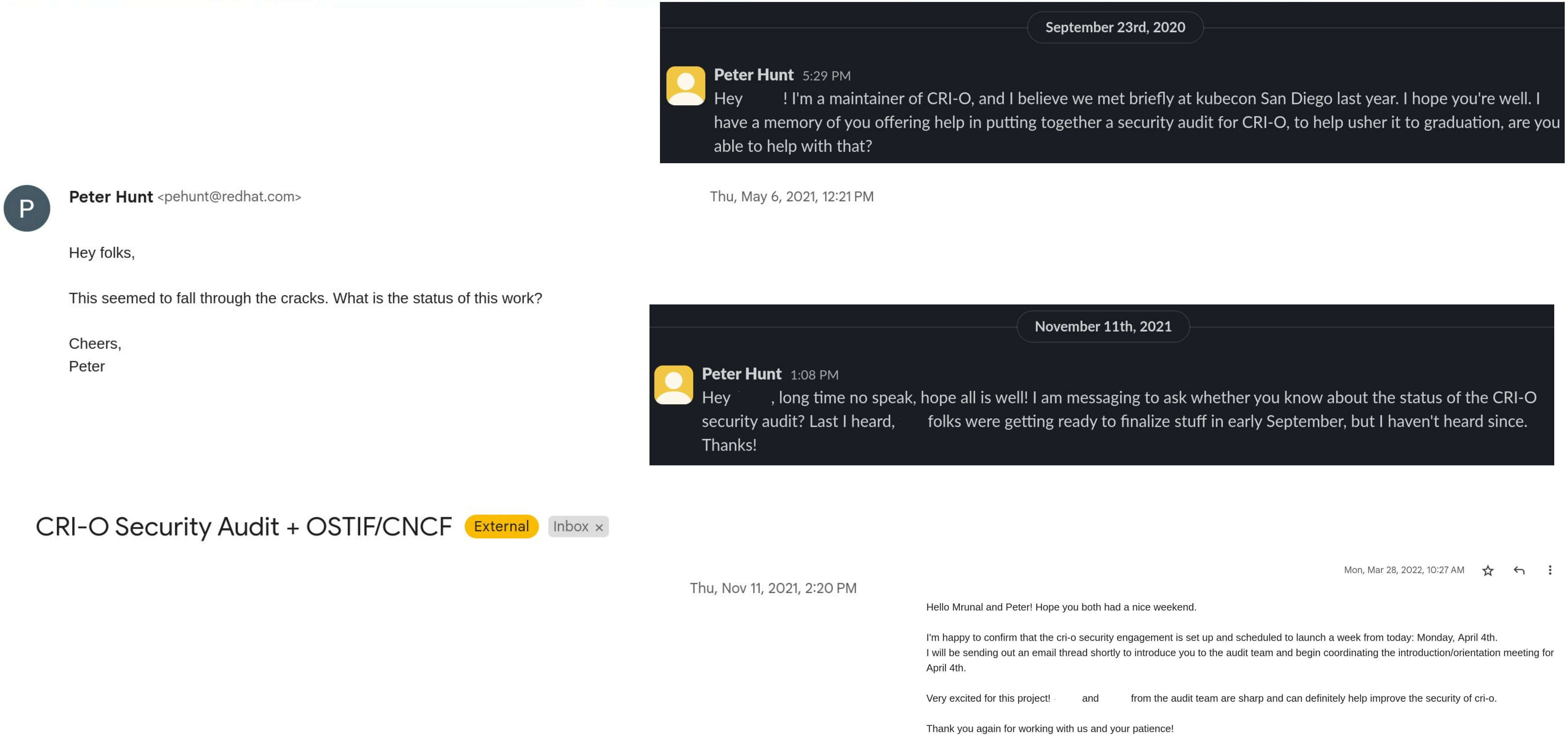
The Road to Graduation

- Have committers from at least two organizations
- Paperwork
  - CII best practices badge
  - GOVERNANCE.md/OWNERS.md/MAINTAINERS.md
  - List of adopters
- Have completed an independent and third party security audit
- Receive a supermajority vote from the TOC to move to graduation stage

Easy… right?

"The Long and Windy Road" by Rodney Campbell is licensed under CC BY-NC-ND 2.0.

The Road to Graduation



September 23rd, 2020

Peter Hunt 5:29 PM
Hey [redacted]! I'm a maintainer of CRI-O, and I believe we met briefly at kubecon San Diego last year. I hope you're well. I have a memory of you offering help in putting together a security audit for CRI-O, to help usher it to graduation, are you able to help with that?

Peter Hunt <pehunt@redhat.com> Thu, May 6, 2021, 12:21 PM

Hey folks,

This seemed to fall through the cracks. What is the status of this work?

Cheers,
Peter

November 11th, 2021

Peter Hunt 1:08 PM
Hey [redacted], long time no speak, hope all is well! I am messaging to ask whether you know about the status of the CRI-O security audit? Last I heard, [redacted] folks were getting ready to finalize stuff in early September, but I haven't heard since. Thanks!

CRI-O Security Audit + OSTIF/CNCF

Thu, Nov 11, 2021, 2:20 PM

Mon, Mar 28, 2022, 10:27 AM

Hello Mrunal and Peter! Hope you both had a nice weekend.

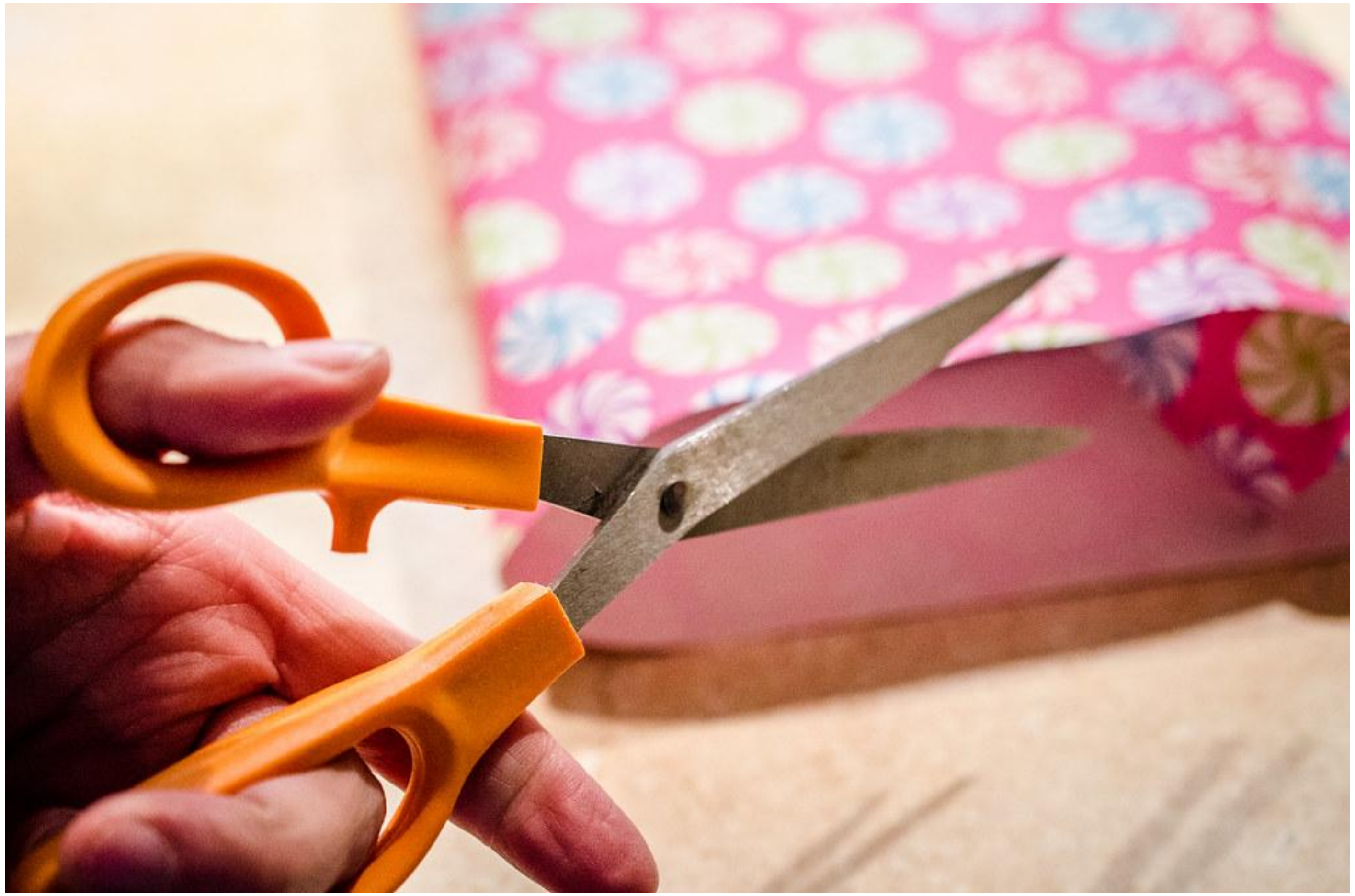
I'm happy to confirm that the cri-o security engagement is set up and scheduled to launch a week from today: Monday, April 4th.
I will be sending out an email thread shortly to introduce you to the audit team and begin coordinating the introduction/orientation meeting for April 4th.

Very excited for this project! [redacted] and [redacted] from the audit team are sharp and can definitely help improve the security of cri-o.

Thank you again for working with us and your patience!

The Road to Graduation

- Due Diligence document
    - Project introduction
    - architectural diagram
    - Project scope
    - Project alignment
    - Contributor breakdown
        - Company agnostic requirement
- TOC session times
    - Be aware of Kubecon breaks!



"Cutting wrapping paper with scissors" by m01229 is licensed under CC BY 2.0.

CRI-O Mentee Programs

- GSoC
- LFX
- and more… maybe you?

CRI Stats/Metrics Update

cAdvisor

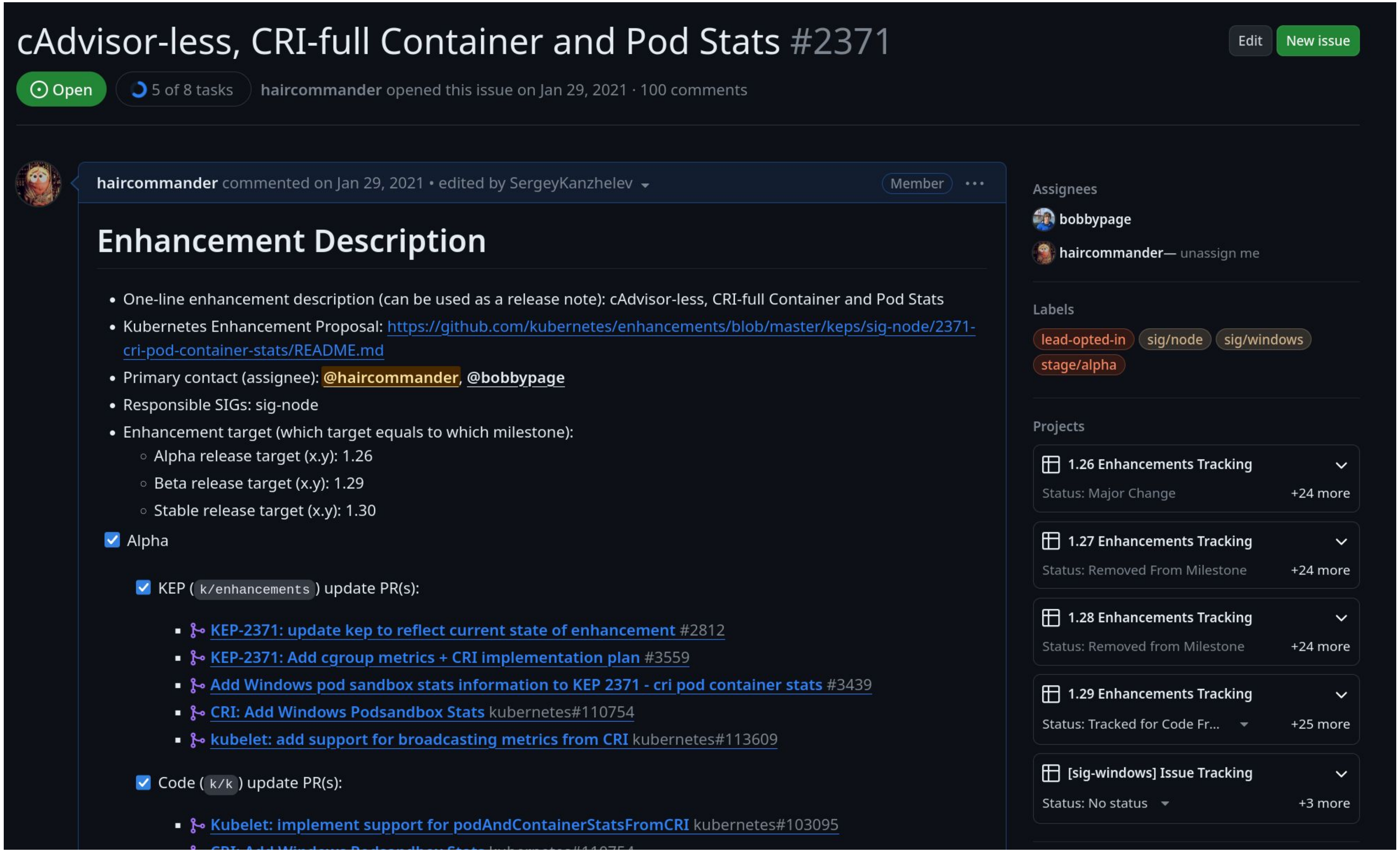
WE SHOULD TAKE METRICS COLLECTION
AND PUSH IT TO THE CRI
imgflip.com

https://imgflip.com/memegenerator/Put-It-Somewhere-Else-Patrick

cri-o

containerd

CRI Stats/Metrics Update



cAdvisor-less, CRI-full Container and Pod Stats #2371

haircommander opened this issue on Jan 29, 2021 · 100 comments

haircommander commented on Jan 29, 2021 • edited by SergeyKanzhelev

Enhancement Description

- One-line enhancement description (can be used as a release note): cAdvisor-less, CRI-full Container and Pod Stats
- Kubernetes Enhancement Proposal: https://github.com/kubernetes/enhancements/blob/master/keps/sig-node/2371-cri-pod-container-stats/README.md
- Primary contact (assignee): @haircommander, @bobbypage
- Responsible SIGs: sig-node
- Enhancement target (which target equals to which milestone):
  - Alpha release target (x.y): 1.26
  - Beta release target (x.y): 1.29
  - Stable release target (x.y): 1.30

☑ Alpha

- ☑ KEP (`k/enhancements`) update PR(s):
  - KEP-2371: update kep to reflect current state of enhancement #2812
  - KEP-2371: Add cgroup metrics + CRI implementation plan #3559
  - Add Windows pod sandbox stats information to KEP 2371 - cri pod container stats #3439
  - CRI: Add Windows Podsandbox Stats kubernetes#110754
  - kubelet: add support for broadcasting metrics from CRI kubernetes#113609
- ☑ Code (`k/k`) update PR(s):
  - Kubelet: implement support for podAndContainerStatsFromCRI kubernetes#103095
  - CRI: Add Windows Podsandbox Stats kubernetes#110754

Assignees

bobbypage

haircommander— unassign me

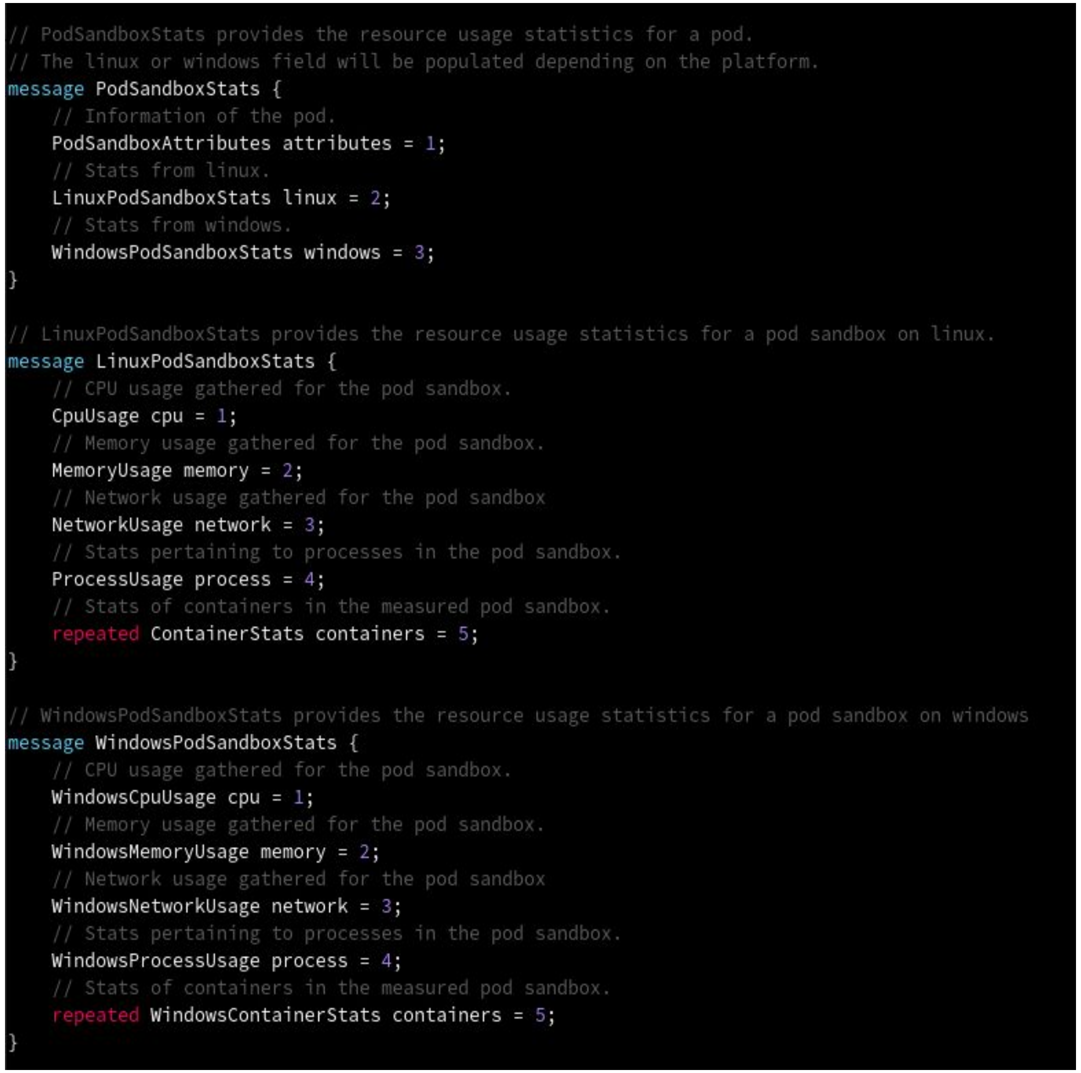
Labels

lead-opted-in sig/node sig/windows stage/alpha

Projects

1.26 Enhancements Tracking	Status: Major Change	+24 more
1.27 Enhancements Tracking	Status: Removed From Milestone	+24 more
1.28 Enhancements Tracking	Status: Removed from Milestone	+24 more
1.29 Enhancements Tracking	Status: Tracked for Code Fr...	+25 more
[sig-windows] Issue Tracking	Status: No status	+3 more

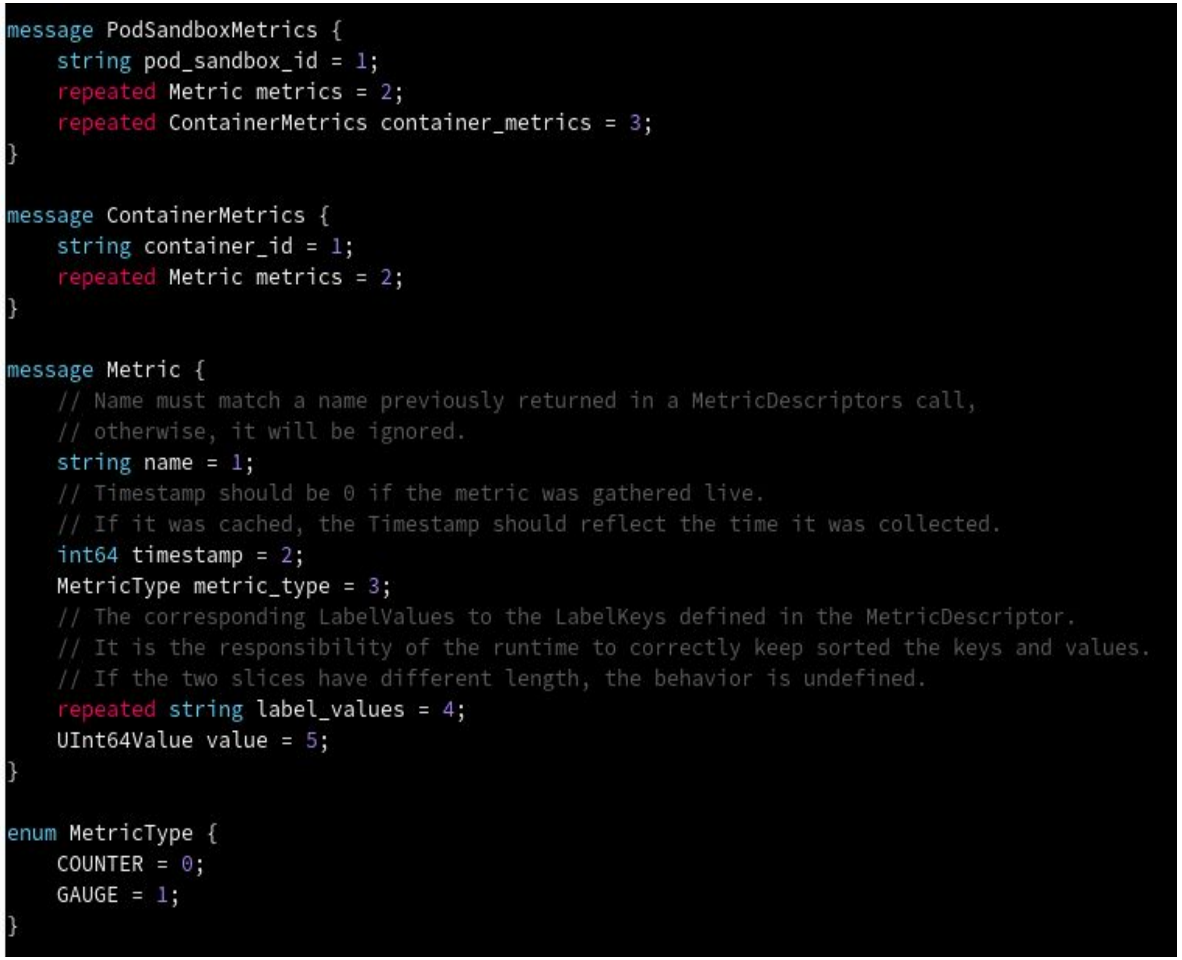
CRI Stats/Metrics Update



```
// PodSandboxStats provides the resource usage statistics for a pod.
// The linux or windows field will be populated depending on the platform.
message PodSandboxStats {
    // Information of the pod.
    PodSandboxAttributes attributes = 1;
    // Stats from linux.
    LinuxPodSandboxStats linux = 2;
    // Stats from windows.
    WindowsPodSandboxStats windows = 3;
}

// LinuxPodSandboxStats provides the resource usage statistics for a pod sandbox on linux.
message LinuxPodSandboxStats {
    // CPU usage gathered for the pod sandbox.
    CpuUsage cpu = 1;
    // Memory usage gathered for the pod sandbox.
    MemoryUsage memory = 2;
    // Network usage gathered for the pod sandbox
    NetworkUsage network = 3;
    // Stats pertaining to processes in the pod sandbox.
    ProcessUsage process = 4;
    // Stats of containers in the measured pod sandbox.
    repeated ContainerStats containers = 5;
}

// WindowsPodSandboxStats provides the resource usage statistics for a pod sandbox on windows
message WindowsPodSandboxStats {
    // CPU usage gathered for the pod sandbox.
    WindowsCpuUsage cpu = 1;
    // Memory usage gathered for the pod sandbox.
    WindowsMemoryUsage memory = 2;
    // Network usage gathered for the pod sandbox
    WindowsNetworkUsage network = 3;
    // Stats pertaining to processes in the pod sandbox.
    WindowsProcessUsage process = 4;
    // Stats of containers in the measured pod sandbox.
    repeated WindowsContainerStats containers = 5;
}
```



```
message PodSandboxMetrics {
    string pod_sandbox_id = 1;
    repeated Metric metrics = 2;
    repeated ContainerMetrics container_metrics = 3;
}

message ContainerMetrics {
    string container_id = 1;
    repeated Metric metrics = 2;
}

message Metric {
    // Name must match a name previously returned in a MetricDescriptors call,
    // otherwise, it will be ignored.
    string name = 1;
    // Timestamp should be 0 if the metric was gathered live.
    // If it was cached, the Timestamp should reflect the time it was collected.
    int64 timestamp = 2;
    MetricType metric_type = 3;
    // The corresponding LabelValues to the LabelKeys defined in the MetricDescriptor.
    // It is the responsibility of the runtime to correctly keep sorted the keys and values.
    // If the two slices have different length, the behavior is undefined.
    repeated string label_values = 4;
    UInt64Value value = 5;
}

enum MetricType {
    COUNTER = 0;
    GAUGE = 1;
}
```

Stats/Metrics Today

Kubelet exposes the cAdvisor metrics via

- /metrics/cadvisor (direct prometheus)

```
kubectl get --raw "/api/v1/nodes/kind-worker/proxy/metrics/cadvisor"
```

- /stats/summary (json)

```
kubectl get --raw "/api/v1/nodes/kind-worker/proxy/stats/summary"
```

- /metrics/resource (metrics server)

Kubelet also depends on cAdvisor for:

- Gathering node level stats
- Eviction Manager

Stats/Metrics Tomorrow

Kubelet exposes the CRI metrics via

- /metrics/cadvisor

Interpreted from Metrics object of CRI

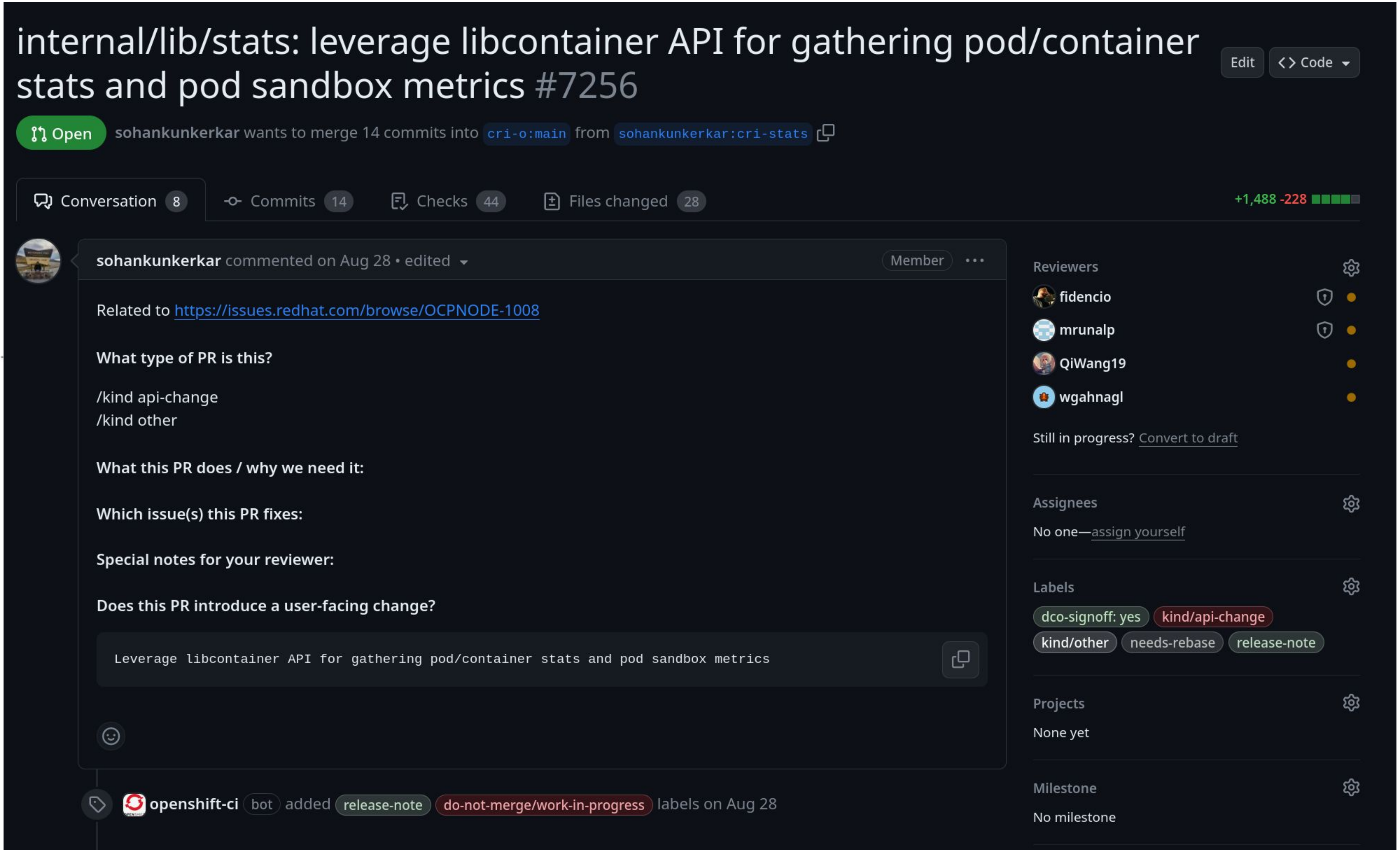
- /stats/summary , /metrics/resource

Interpreted from Stats object of CRI

Kubelet still depends on cAdvisor for:

- Gathering node level stats
- Eviction Manager

CRI Stats/Metrics Update



internal/lib/stats: leverage libcontainer API for gathering pod/container stats and pod sandbox metrics #7256

Open sohankunkerkar wants to merge 14 commits into `cri-o:main` from `sohankunkerkar:cri-stats`

sohankunkerkar commented on Aug 28 • edited

Related to https://issues.redhat.com/browse/OCPNODE-1008

What type of PR is this?

/kind api-change
/kind other

What this PR does / why we need it:

Which issue(s) this PR fixes:

Special notes for your reviewer:

Does this PR introduce a user-facing change?

```
Leverage libcontainer API for gathering pod/container stats and pod sandbox metrics
```

openshift-ci bot added release-note do-not-merge/work-in-progress labels on Aug 28

Reviewers
fidencio
mrunalp
QiWang19
wgahnagl

Still in progress? Convert to draft

Assignees
No one—assign yourself

Labels
dco-signoff: yes kind/api-change kind/other needs-rebase release-note

Projects
None yet

Milestone
No milestone

Optimizations

- Collect once, use everywhere
- Periodic collection
- Configuration for which metrics to collect
- Reduce object churn

https://www.unravelsitecore.com/wp-content/uploads/2018/02/optimize-before-you-automate.jpg

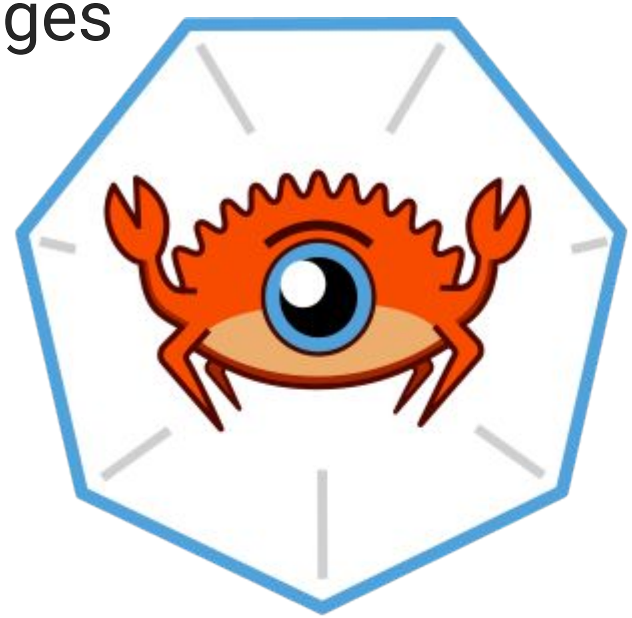
CRI Stats/Metrics Update

- Aiming to support CRI Metrics in 1.29.0
    - Could be 1.30.0 instead
- Advocating for containerd support
- KEP to Beta after that

What’s new about conmon-rs?

Released various versions since KubeCon EU 2023 (latest is v0.6.1):

- conmon-rs has a logo!
- Various bug fixes and stabilizations
- Added multi-arch static binaries for `amd64`, `arm64` and `ppc64le`
- Integration into default CRI-O static bundle and packages
- Support for JSON logger as part of the LFX mentoring program



Looking forward to integration into Podman!

Sigstore signature verification

- Kubernetes signs container image-based artifacts since v1.24, which aims for stabilization in one of the upcoming releases
- Verifying the signatures manually via the CLI is cumbersome
- CRI-O supports basic pull-based sigstore validation since v1.28

Compared to an admission controller based verification (like the sigstore policy-controller), CRI-O aims to validate signatures directly from the container runtime level.

Sigstore signature verification

CRI-O reads a policy.json, which contains the rules to evaluate:

```
{
  "default": [{ "type": "reject" }],
  "transports": {
    "docker": {
      "quay.io/crio/signed": [
        {
          "type": "sigstoreSigned",
          "signedIdentity": { "type": "matchRepository" },
          "fulcio": {
            "oidcIssuer": "https://github.com/login/oauth",
            "subjectEmail": "sgrunert@redhat.com",
            "caData": "LS0…="
          },
          "rekorPublicKeyData": "LS0…="
        }
      ]
    }
  }
}
```

Sigstore signature verification

CRI-O v1.28 supports namespaced policies by using the options:

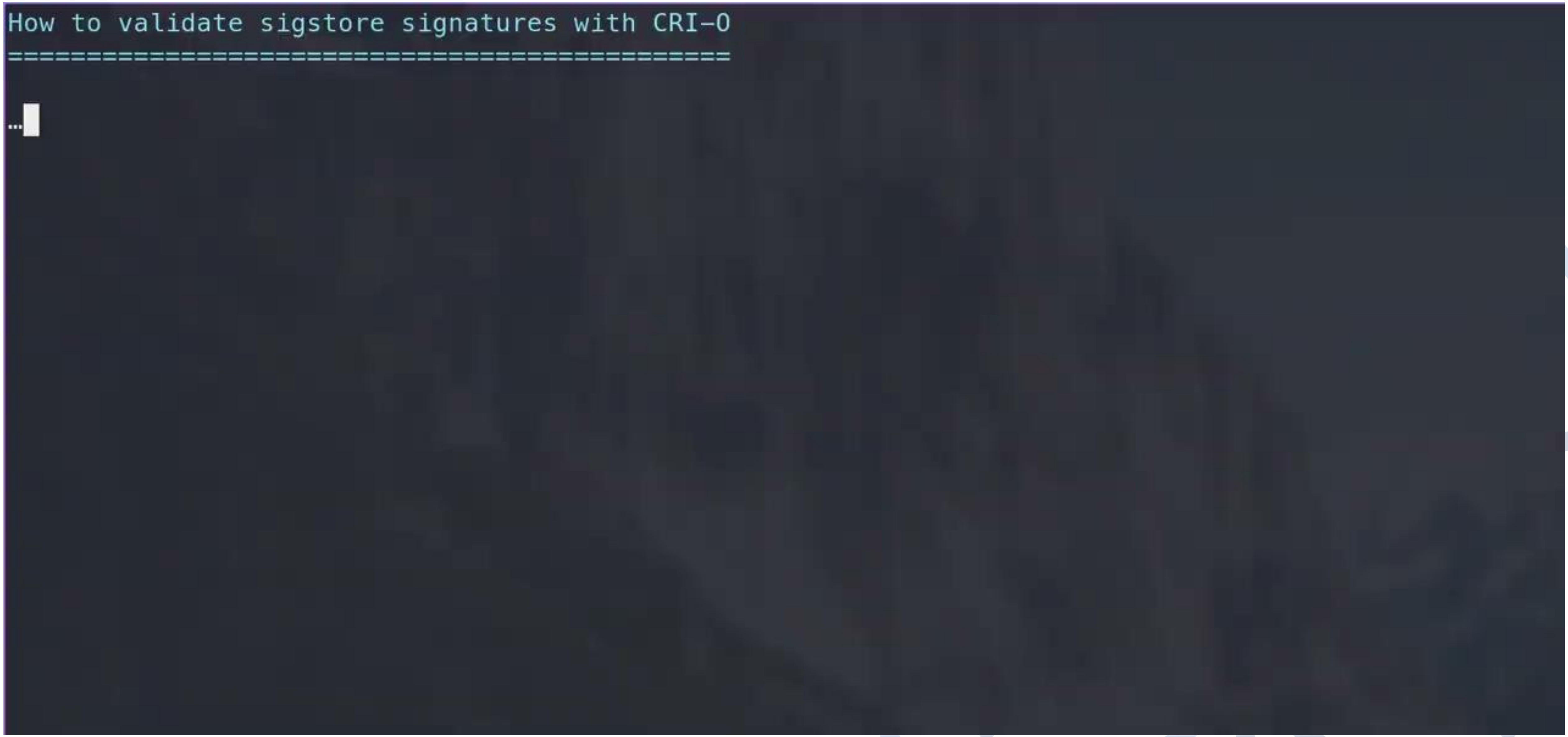
`--signature-policy-dir` / `signature_policy_dir`

- defines the root path for pod namespace-separated signature policies
- CRI-O will lookup that path and assemble a policy like `<SIGNATURE_POLICY_DIR>/<NAMESPACE>.json`, which will be used on image pull (if existing)
- Global policy will be used as fallback

Future work:

- Integration of container creation based signature validation
- Custom Resources as abstraction of policies

Sigstore signature verification



```
How to validate sigstore signatures with CRI-O
==============================================

…
```

CRI-O’s packaging efforts

- Kubernetes announced that legacy package repositories are frozen back in August
- Community owned deb/rpm repositories powered by the OpenBuildService (OBS): https://build.opensuse.org

CRI-O already used OBS for packaging the traditional way:

- Building from sources
- Maintaining dependencies
- Updating new versions manually

CRI-O's packaging efforts

We already ship static binary bundles:

- Contain everything CRI-O needs to run on a certain platform (os/arch)
- Automatically published for release branches and tagged releases
- Reproducible builds using Nix packages (nixpkgs)
- Using static binaries as input for packages is the Kubernetes way

Integrating into the existing OBS project layout

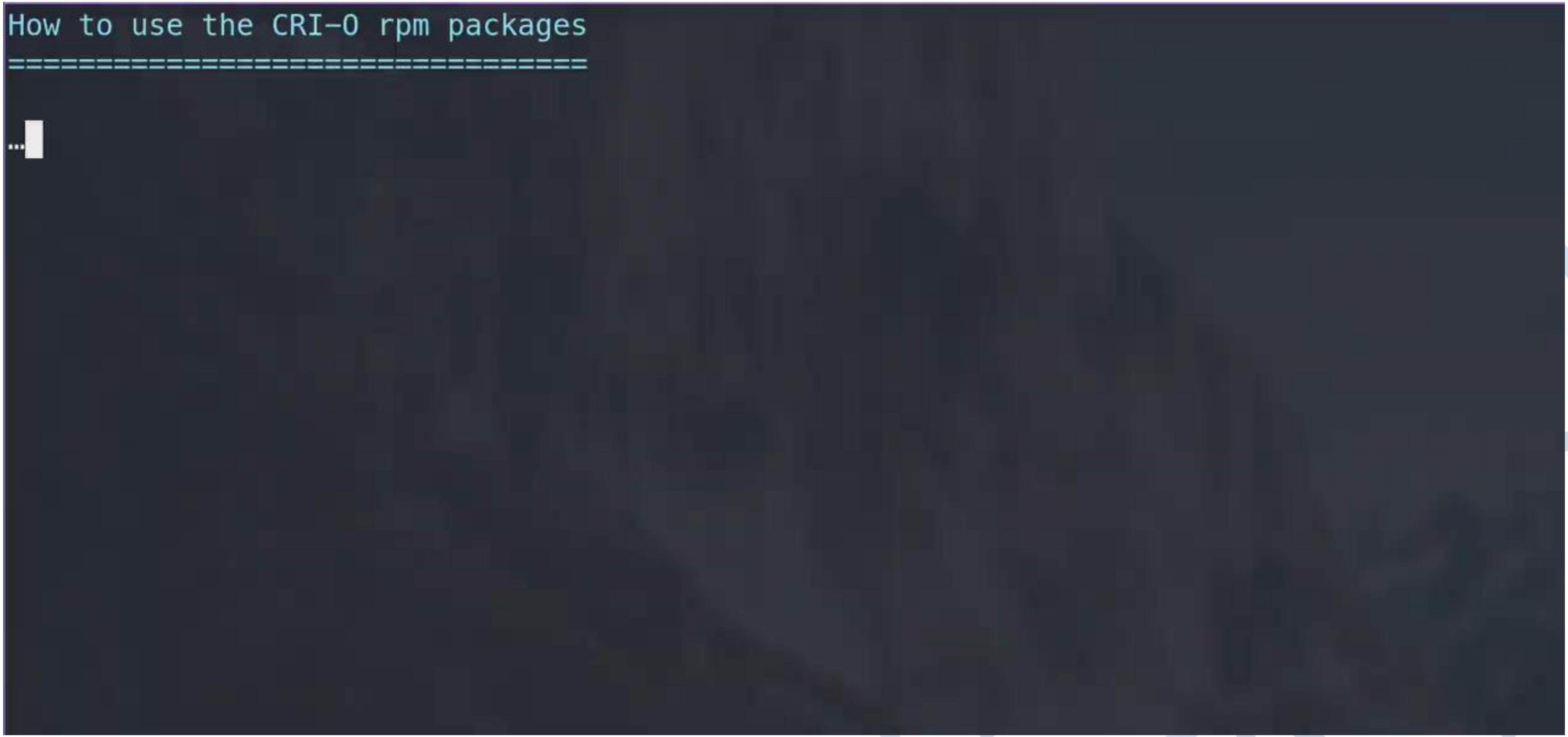
seems to be a low hanging fruit… 🍎

CRI-O's packaging efforts

All future CRI-O packages will be shipped as part of the officially supported Kubernetes infrastructure hosted on pkgs.k8s.io! 📦

- Package builds are now in a dedicated repository: https://github.com/cri-o/packaging
- Daily reconciliation for supported branches
- Package test pipeline for various distributions
- Staging and releasing packages powered by the Kubernetes Release Toolbox (krel)

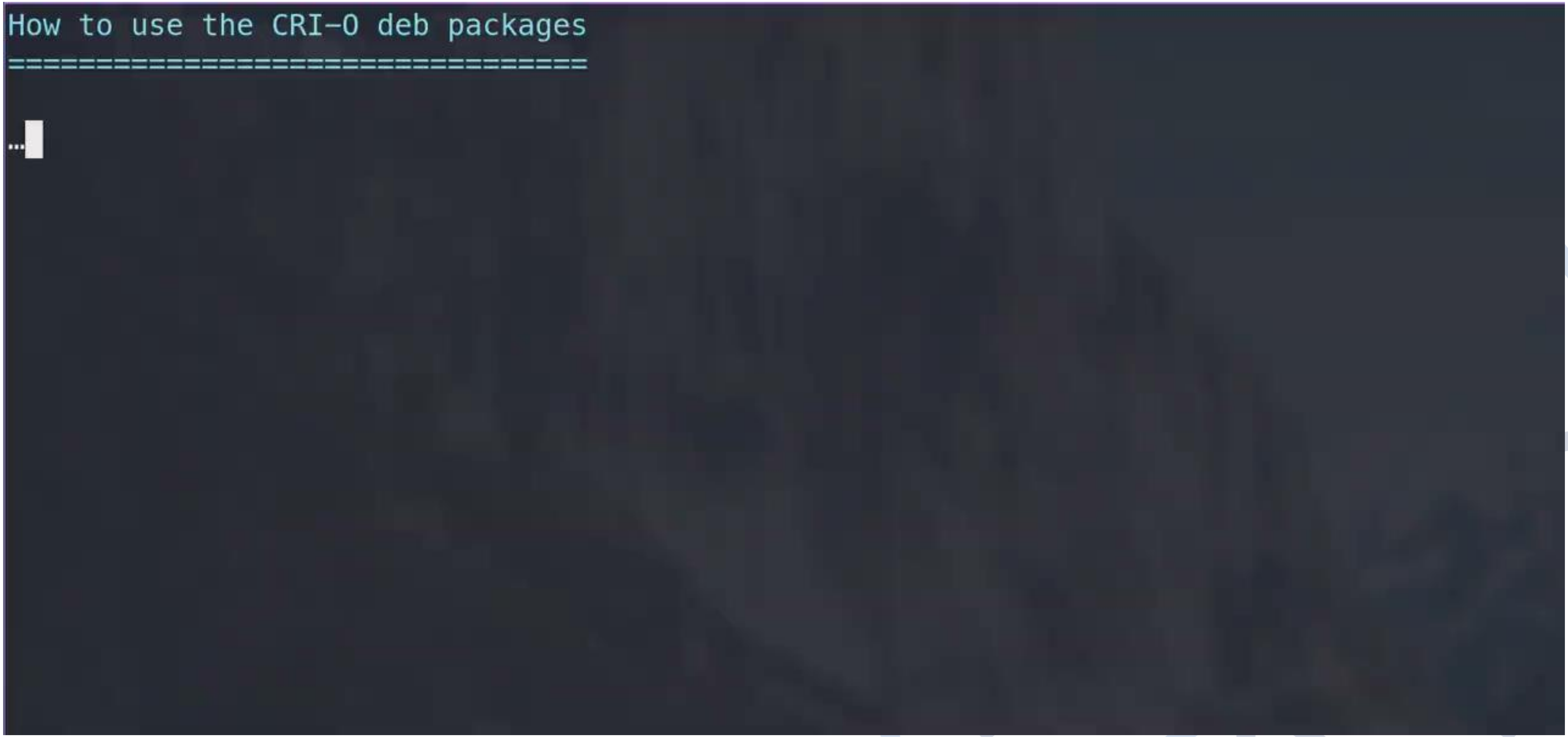
CRI-O’s packaging efforts



```
How to use the CRI-O rpm packages
=================================


...
```

CRI-O’s packaging efforts



```
How to use the CRI-O deb packages
=================================


...
```

CRI-O’s packaging efforts

Demo sources:

https://github.com/saschagrunert/kubecon-na-2023-cri-o

What we plan for the future

CRI-O's feature roadmap invites everyone to contribute:

https://github.com/orgs/cri-o/projects/1

We've plans for:

- A Rust based NRI framework
- WASM plugins loaded directly into CRI-O (instead of NRI)
- Increasing release automation
- Documentation enhancements / a cri-o.io blog
- Making CRI-O more portable for non Linux

Reach out to us

GitHub

Slack

KubeCon | CloudNativeCon

North America 2023

Thank you!