Running Docker, Podman, and K8s inside Podman rootless containers

Podman Community Meeting

May 2021

About us …

Cesar Talledo

Rodny Molina

- Software Engineers (ex VMware, LinkedIn)
- Developers of the Sysbox runtime
- Founders of Nestybox

Problem we are solving

Systemd, Docker, Podman, K8s, etc.

Enhancing containers to run most workloads that run in VMs, seamlessly and with strong isolation.

No special images, tricky entrypoints, complex commands, etc.

Linux User Namespace
(root in container != root in host)

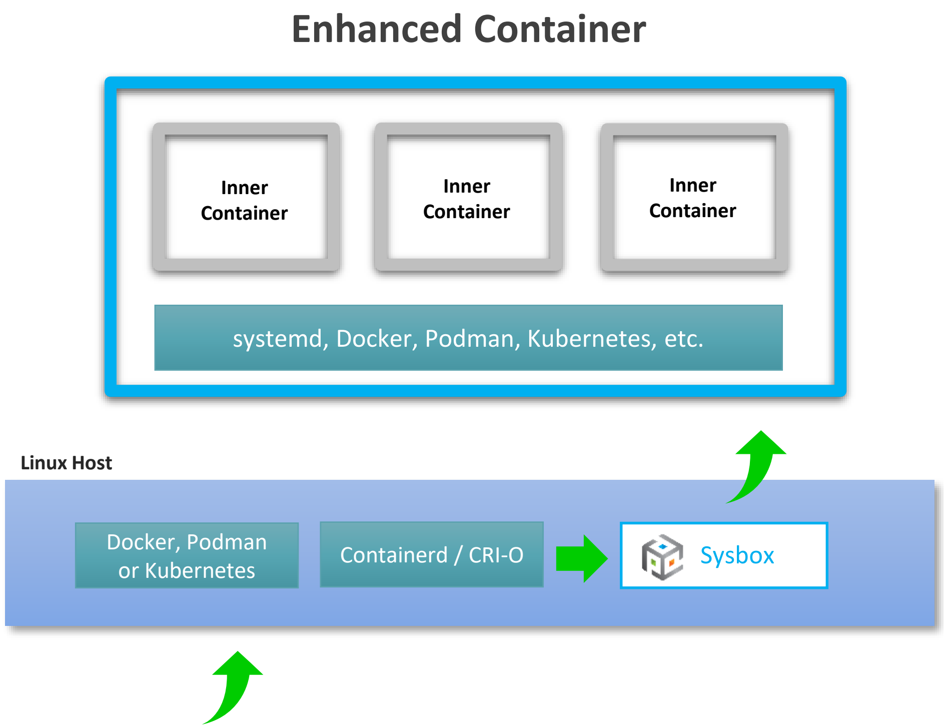
Our Goal

```
$ podman run --userns=auto:size=65536 -it any-image
```

We want this to create a container capable of running systemd, Docker, Podman, Kubernetes, etc., securely and seamlessly (like a VM).

(Make it easy, yet powerful)

Our Solution: Sysbox “runc”



Strong isolation (Linux User Namespace)

Runs same workloads as VMs, seamlessly

No special images or complex setups needed

Sysbox is an enhanced “runc”

Uses cutting-edge OS virtualization techniques (no VMs)

Open-source (Apache 2.0)

```
$ podman run --runtime=sysbox-runc --userns=auto:size=65536 –it any-image
```

Features

- User-namespace on all containers
- File-system ID shifting (shiftfs now, ID-mapped mounts soon)
- Procfs and sysfs virtualization
- Syscall interception
- Initial mount locking
- Easy preloading of inner container images
- Sharing inner container images across Sysbox containers

Limitations

- Linux only
  - Needs 5.5+ kernel
  - For Ubuntu, 5.0+ works.
- 90% OCI compatible
  - Implicitly sets up container environment to enable it to run system software
- Some workloads don't run inside the containers (yet)
  - ipvs, kernel module loading, etc.
- Sysbox is a daemon and must run as root

Demo

Summary

- Currently, running system software in containers requires
  - Insecure (privileged) containers
  - Complex container images and commands
- We need to change this.
  - Enables powerful use cases for containers (beyond micro-service deployments)
- Sysbox is a next-gen runc designed for this.
- Enterprises are using it to replace VMs in many scenarios.

Learn more …

Sysbox GitHub Repo

Nestybox Website

Nestybox Blog Site

Video #1: Running Docker inside a Container (easily & securely)

Video #2: Running Kubernetes inside a Container (easily & securely)

Contact Us!

Cesar Talledo (ctalledo@nestybox.com)

Rodny Molina (rmolina@nestybox.com)

slack

We love to hear feedback & comments!

Thank You!

Nestybox

San Jose, CA