Red Hat

FUTURE OF CONTAINERS

VINCENT BATTS

Office of the CTO - Emerging Technologies

CONTAINER ECOSYSTEM

📦 CONTAINER ECOSYSTEM

parts of containers are "boring" now!

📦 CONTAINER ECOSYSTEM

parts of containers are "boring" now!

standardization! OCI specifications

📦 CONTAINER ECOSYSTEM

parts of containers are "boring" now!

standardization! OCI specifications

conversation is at (and has been) at the orchestration layer (kubernetes/openshift)

📦 CONTAINER ECOSYSTEM

parts of containers are "boring" now!

standardization! OCI specifications

conversation is at (and has been) at the orchestration layer (kubernetes/openshift)

less and less dependence on `docker`

📦 CONTAINER ECOSYSTEM

parts of containers are "boring" now!

standardization! OCI specifications

conversation is at (and has been) at the orchestration layer (kubernetes/openshift)

less and less dependence on `docker`

foundational community has bounded and rebounded, and is solid

MORE & MORE ISOLATION

MORE & MORE ISOLATION

(source)

MORE & MORE ISOLATION

containers inside containers

challenges of non-root everywhere (vfs and namespaces)

smarter and smarter CAPS, syscall awareness

MORE & MORE ISOLATION

		`--privileged`		non-privileged	
		root	non-root	root	non-root
On Host		✔	✔	✔	✔
In container	root	✔ (with vfs or fuse-overlayfs)		✔ (with vfs or fuse-overlayfs)	✘ (setuid)
	non-root (userns ?)			✘ (clone and setuid)	✘

The goal is for consistent experience
in all arrangements

CONTROL GROUPS

cgroups v2 are coming!

"hybrid" cgroups was no good

BPF will be a new technology to settle down

CONTROL GROUPS

v1 controllers are each their own hierarchy

(pids, memory, freezer, devices, net_cls, net_prio, cpu, cpuacct, cupset, hugetlb, blkio, perf_event)

v2 controllers are a unified hierarchy

(cpu, memory, io, pids, cpuset, rdma, perf_event)

Some controllers are only via BPF now

(devices)

https://medium.com/nttlabs/cgroup-v2-596d035be4d7

BPF

BPF

Super Cool Tech that we'll all hear more and more about

BPF

Super Cool Tech that we'll all hear more and more about

Originally focused on network filtering logic,

BPF

Super Cool Tech that we'll all hear more and more about

Originally focused on network filtering logic,

Now is kernel instrumented bytecode run in-kernel VM

Currently requires a compiler and is largely not relocatable 😰

Will likely become the defacto tracing tool for containers at large

🚚 DISTRIBUTION 🚚

🚚 DISTRIBUTION 🚚

OCI distribution spec is underway (formerly Docker Registry API v2)

🚚 DISTRIBUTION 🚚

OCI distribution spec is underway (formerly Docker Registry API v2)

Consolidate all of our registry stories (all eyes on Quay 👀)

🚚 DISTRIBUTION 🚚

OCI distribution spec is underway (formerly Docker Registry API v2)

Consolidate all of our registry stories (all eyes on Quay 👀)

signing has more options now

Source code of the container image

PROJECT QUAY

- Open sourced as of November 2019
- projectquay.io
- Integration tests will benefit the ecosystem for registry conformance

SOURCE CODE IMAGE

- part of "container first" builds
- fundamental shift in Red Hat's focus on RPMs
- drastic audit improvement of "what's built into this container?"
- Significant like ftp.redhat.com

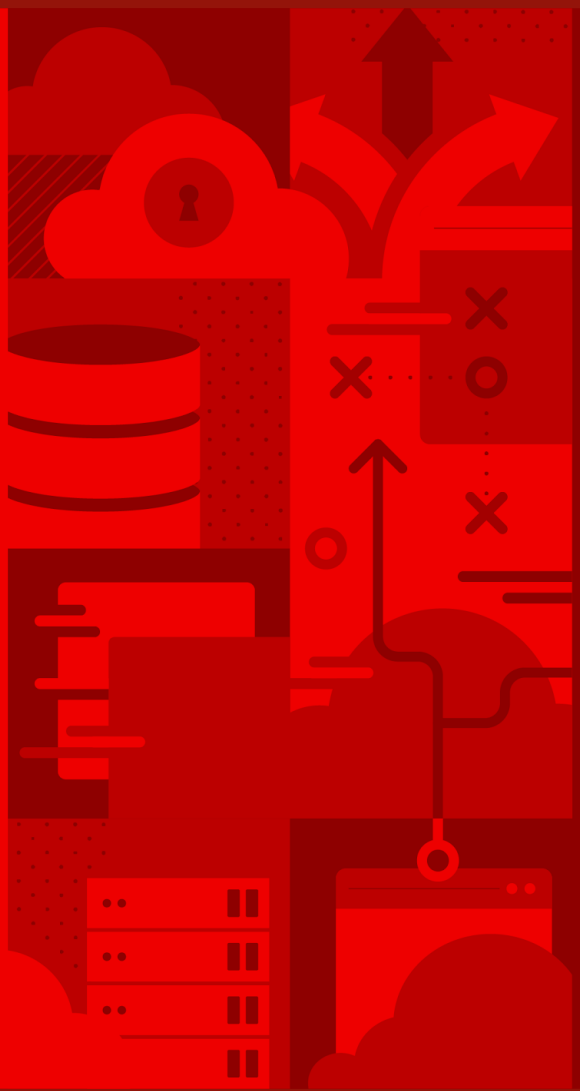
🔦LIGHTS; 📽️CAMERA; 🕺ACTION!

`podman generate` and `podman play` (for system and k8s)

Try out Fedora 31 (which defaults to cgroups-v2-only)

I want to hear your software audit requirements!

get clarity on the nuance of non-root container requests



STAY ENGAGED

Developers. redhat.com

Your access point for no-cost developer tools and product subscriptions, how-tos, and demos

Red Hat User Groups

Meetups for networking and tech deep dives
www.meetup.com/Dallas-Red-Hat-Users-Group/

DevNation

Virtual and live events
Catch replays at
https://developers.redhat.com/devnation/

Next.redhat.com

Stay in touch with the Office of the CTO

Red Hat

Red Hat

THANK YOU!