katacontainers*

The speed of containers, the security of VMs

Xu Wang, Hyper <xu@hyper.sh>
Samuel Ortiz, Intel <samuel.ortiz@intel.com>

*Other names and brands may be claimed as the property of others.

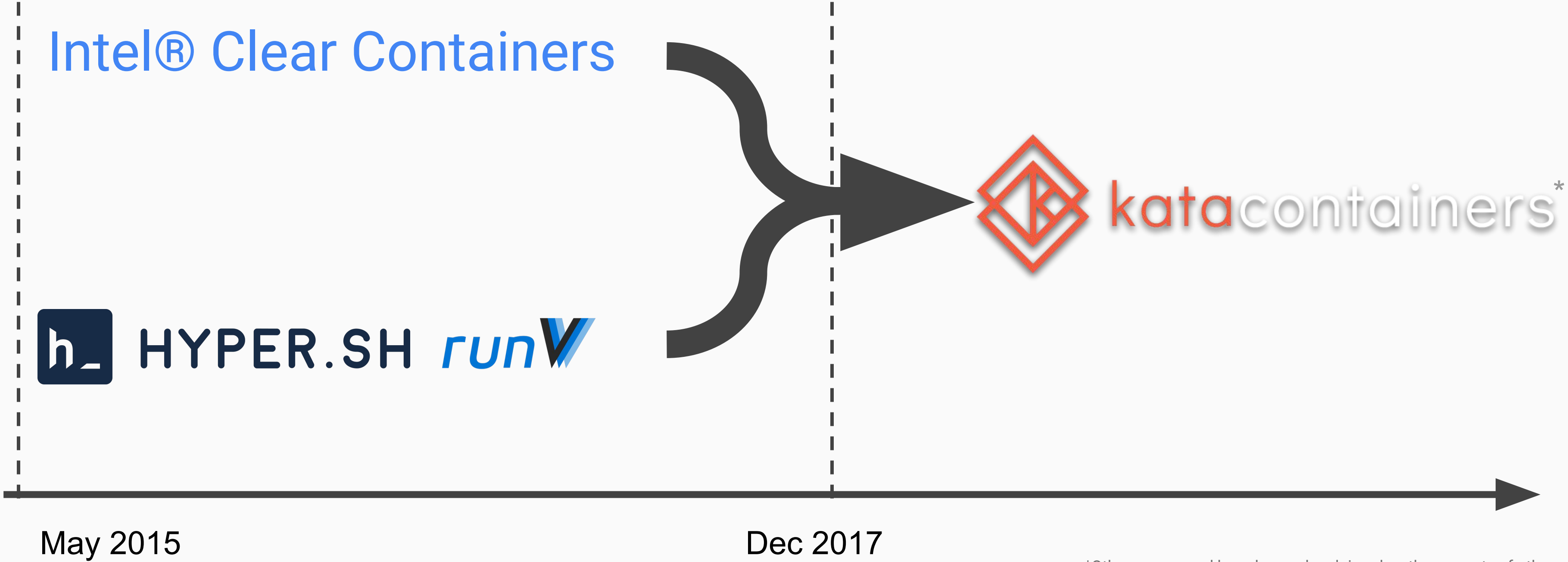
Contents

Project Overview

Technical Details

Get Involved

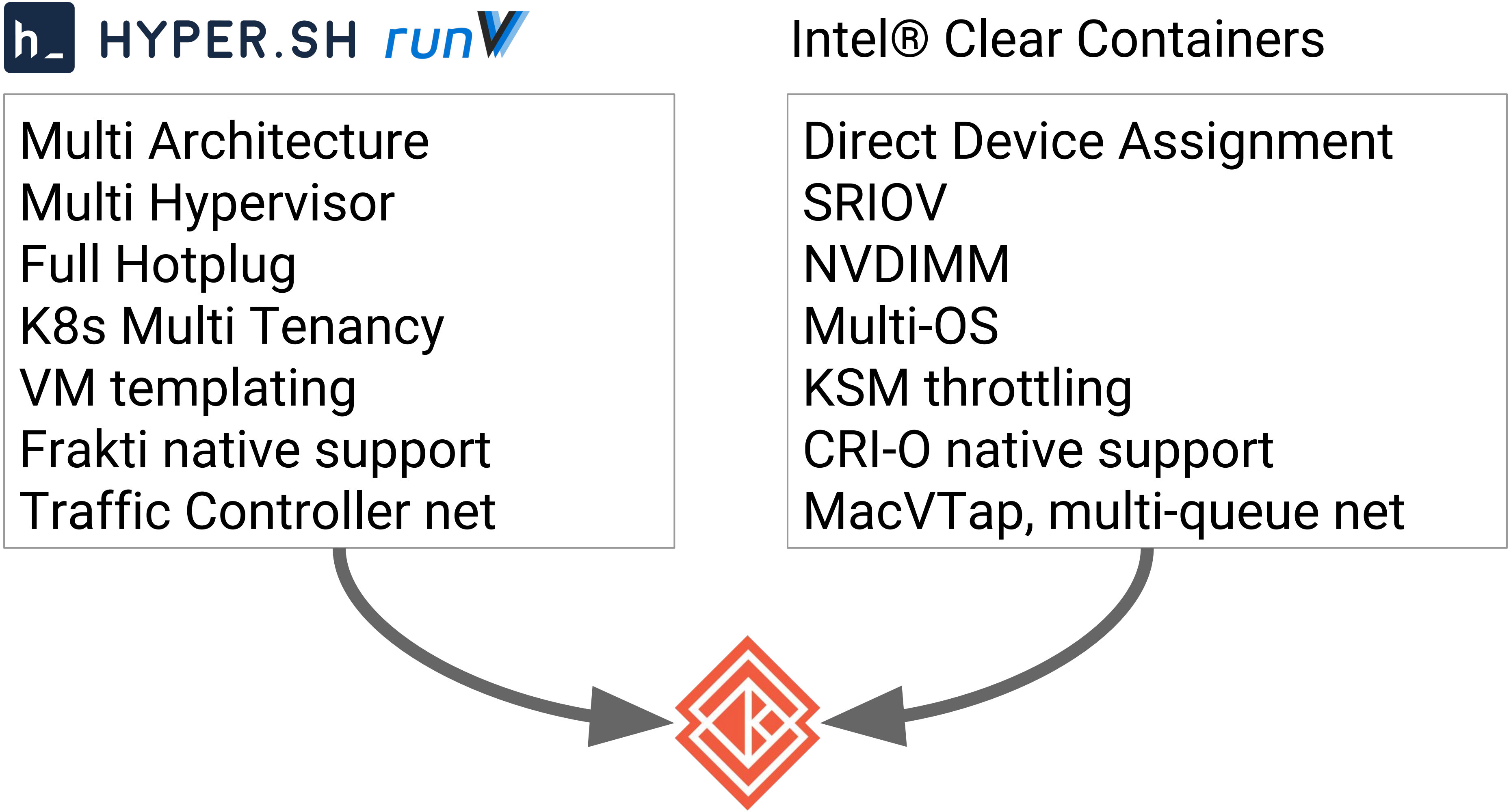
History



*Other names and brands may be claimed as the property of others.

Technical Vision

- Light and fast VM-based containers
- Merge Intel® Clear Containers and Hyper runV technologies
- Seamless integration with Kubernetes (CRI), Docker and Openstack
- Support multiple architectures (x86 today; others to come in the future)
- Support multiple hypervisors (KVM today; others to come in the future)



HYPER.SH runV

- Multi Architecture
- Multi Hypervisor
- Full Hotplug
- K8s Multi Tenancy
- VM templating
- Frakti native support
- Traffic Controller net

Intel® Clear Containers

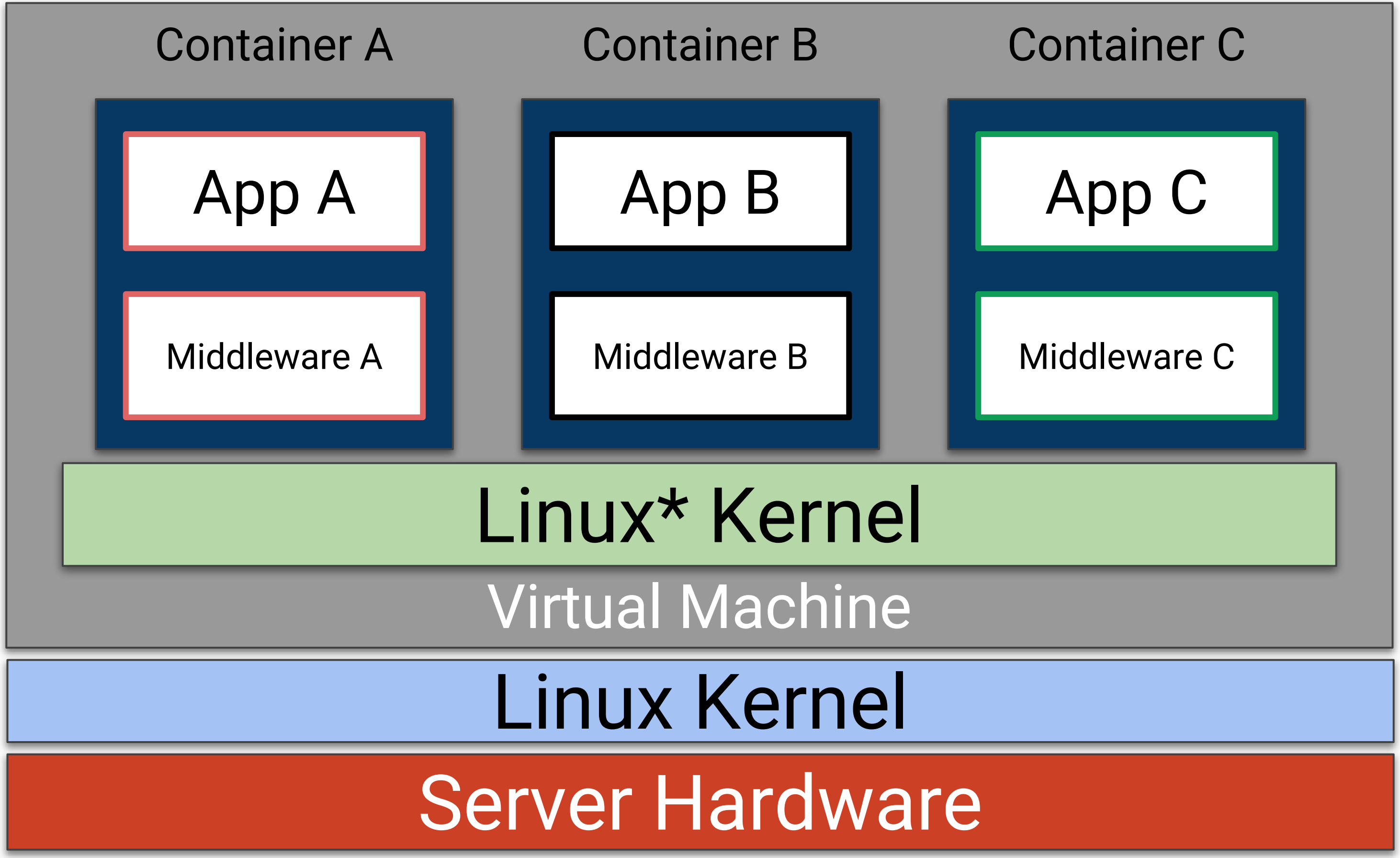
- Direct Device Assignment
- SRIOV
- NVDIMM
- Multi-OS
- KSM throttling
- CRI-O native support
- MacVTap, multi-queue net

Non-Technical Goals

- Open and vendor-neutral project
- All VM based containers, users and consumers under the same project
- Managed **at** the OpenStack Foundation*
- Independent from the OpenStack* software project

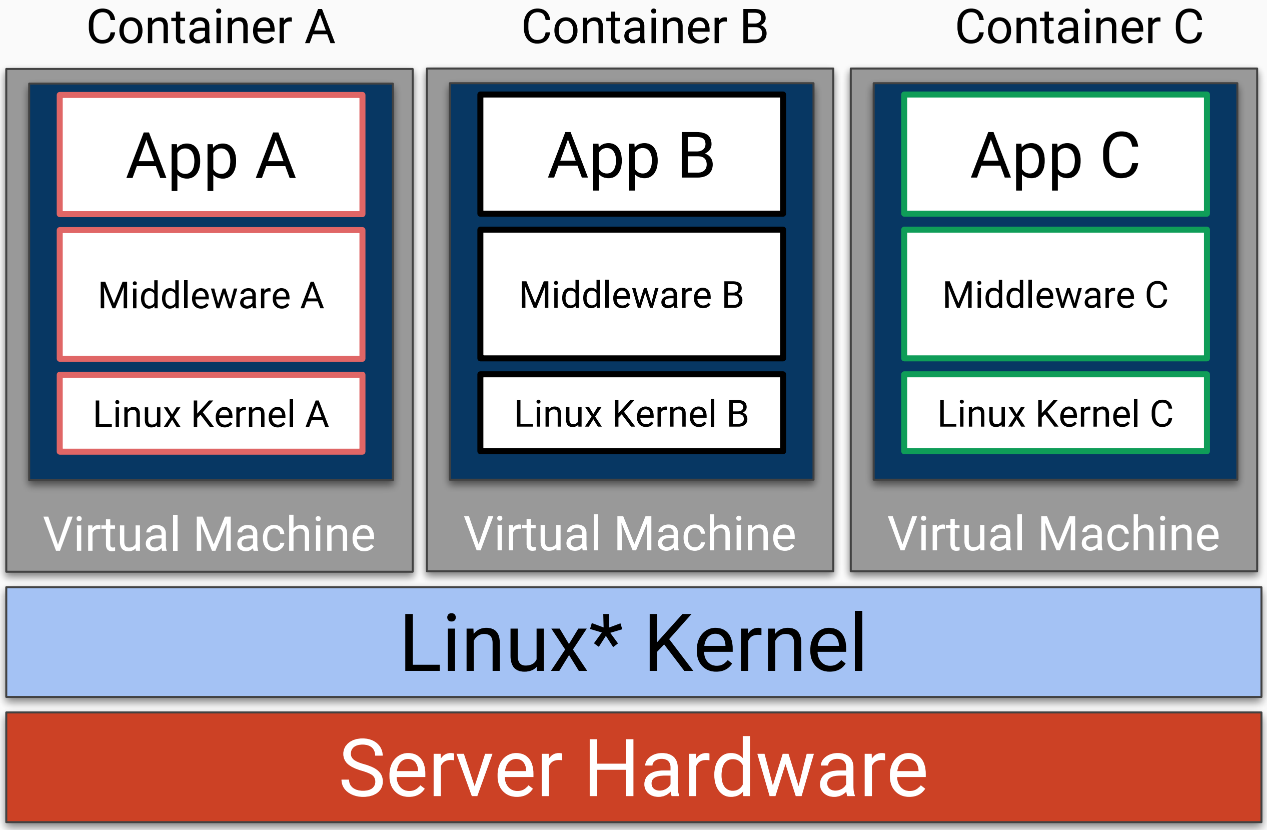
*Other names and brands may be claimed as the property of others.

Containers in Cloud



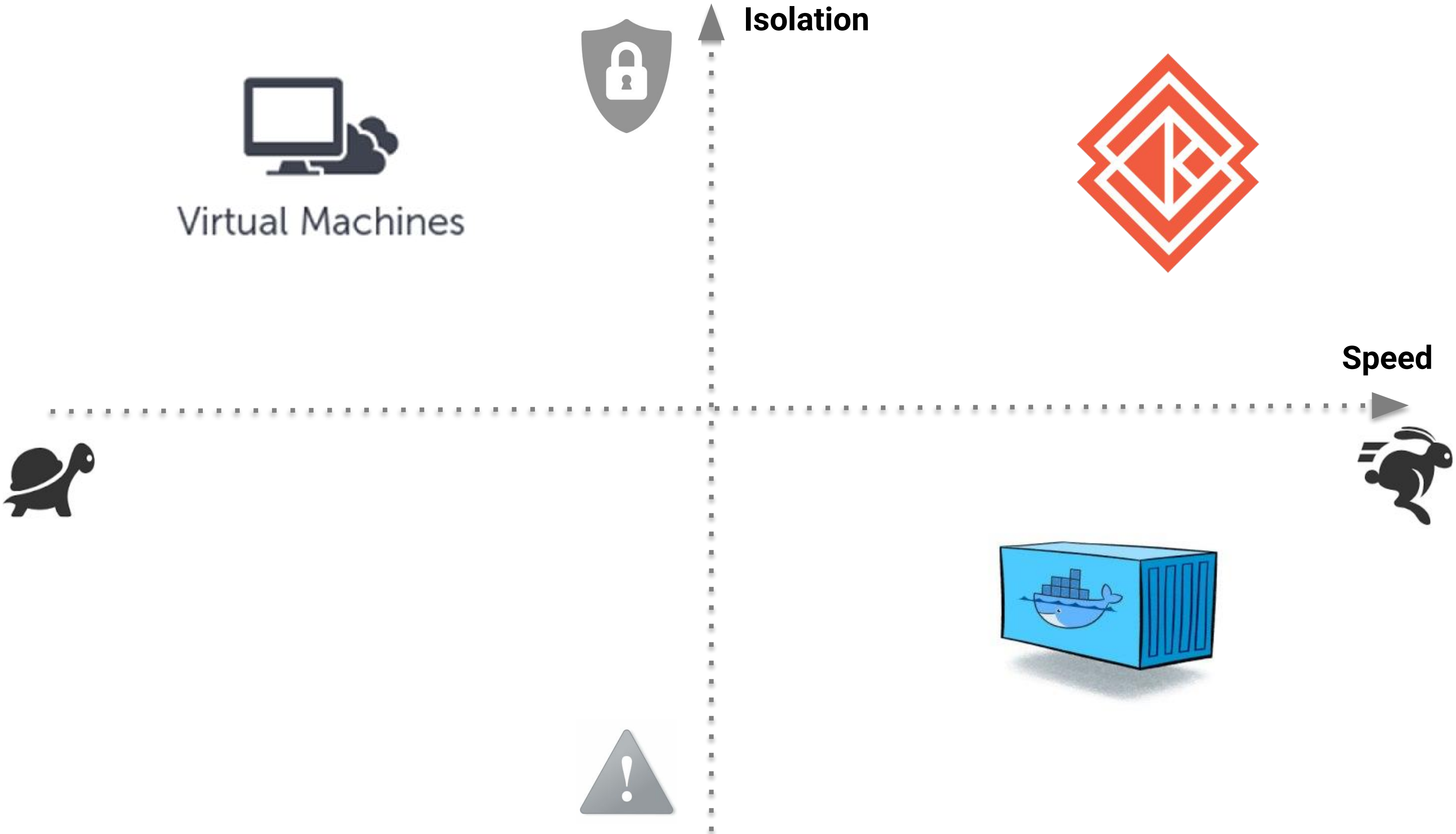
*Other names and brands may be claimed as the property of others.

Hypervisor Based Containers



- Each container/pod is hypervisor isolated
- As secure as a VM
- As fast as a container
- Seamless integration with the container ecosystem and management layers

*Other names and brands may be claimed as the property of others.

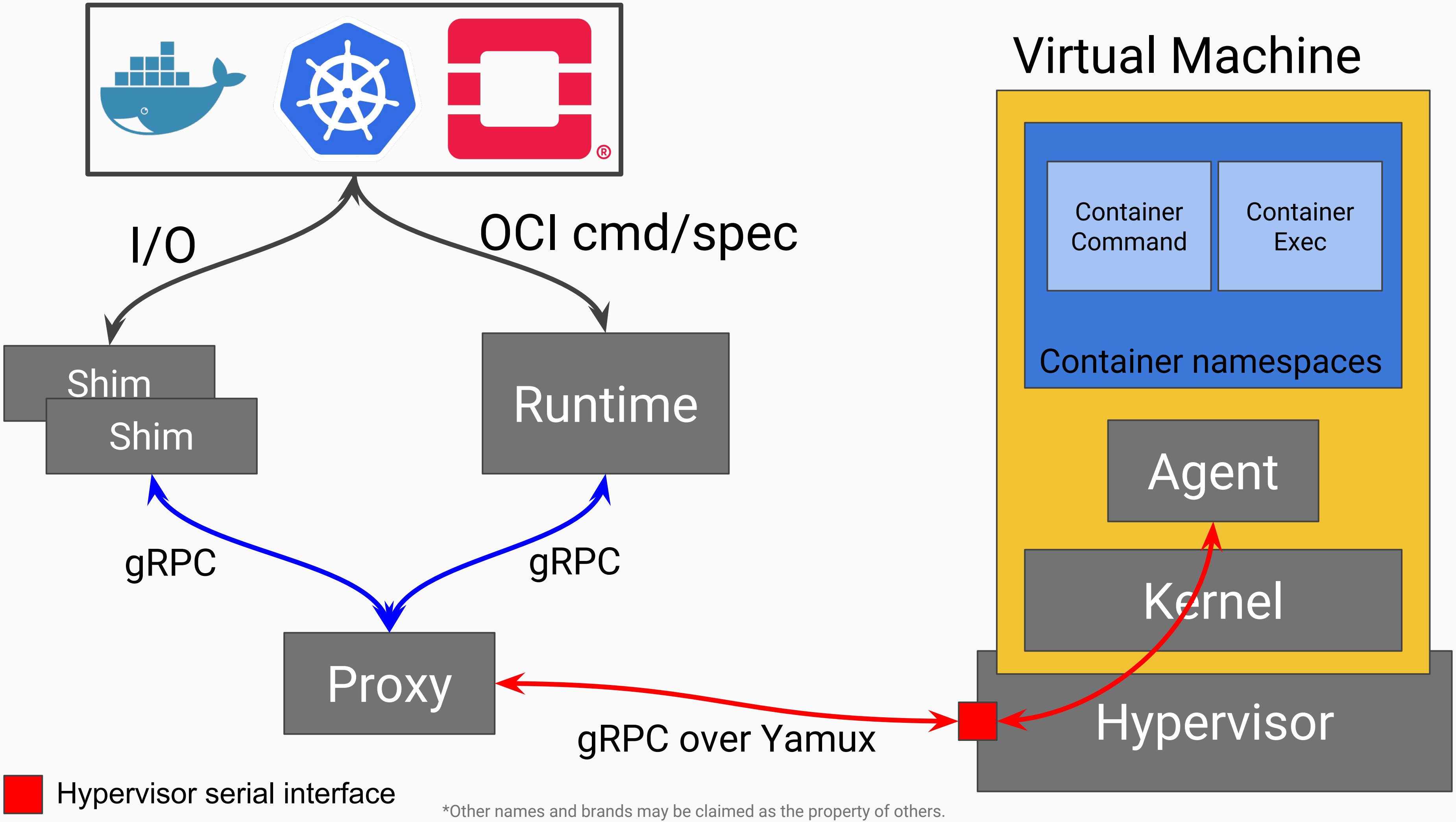


Isolation

Speed

Virtual Machines

Technical Details



Virtual Machine

Container Command

Container Exec

Container namespaces

Agent

Kernel

Hypervisor

I/O

OCI cmd/spec

Shim

Shim

Runtime

gRPC

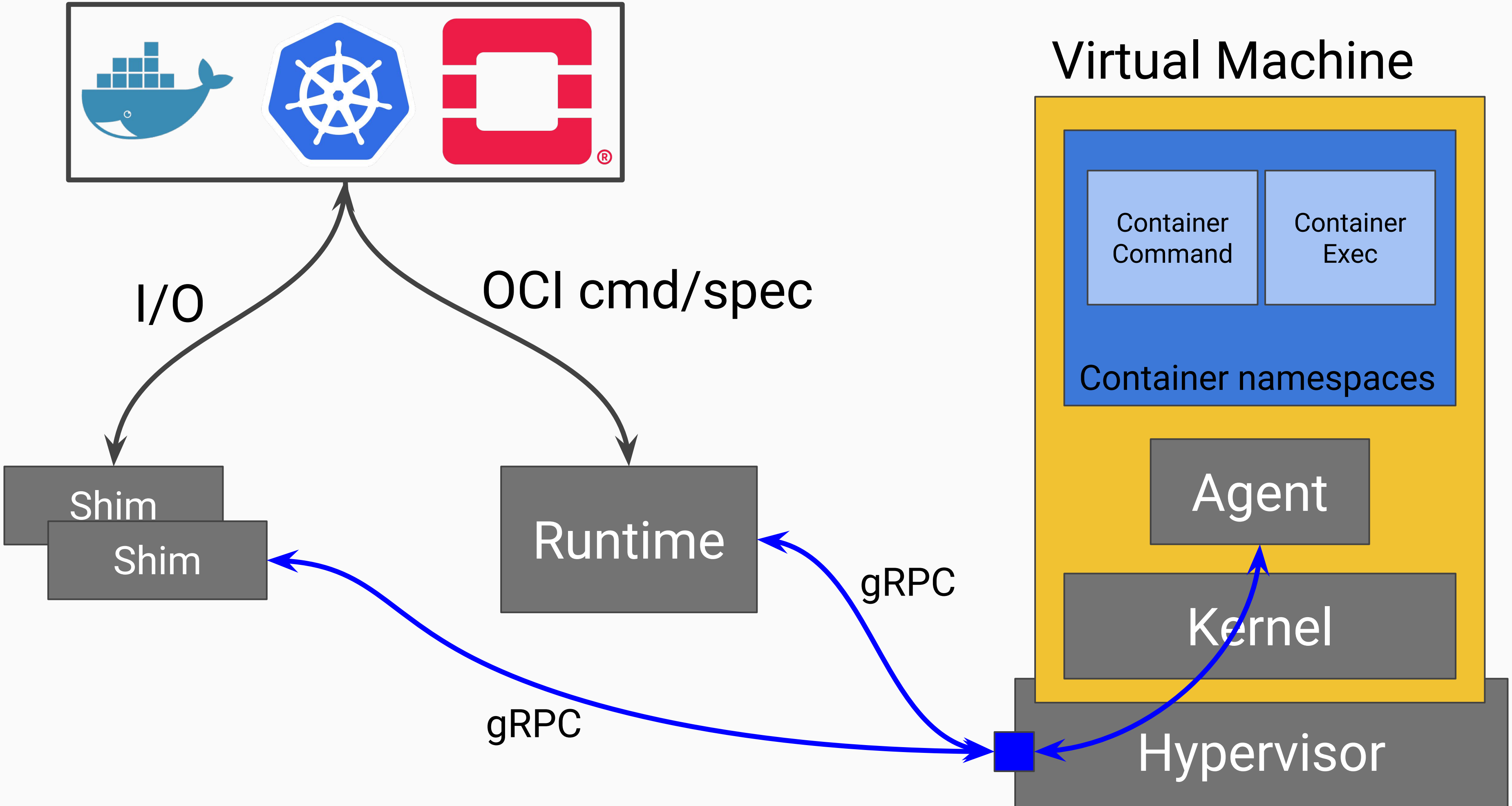
gRPC

Proxy

gRPC over Yamux

🟥 Hypervisor serial interface

*Other names and brands may be claimed as the property of others.



Virtual Machine

Container Command

Container Exec

Container namespaces

Agent

Kernel

Hypervisor

I/O

OCI cmd/spec

Shim

Shim

Runtime

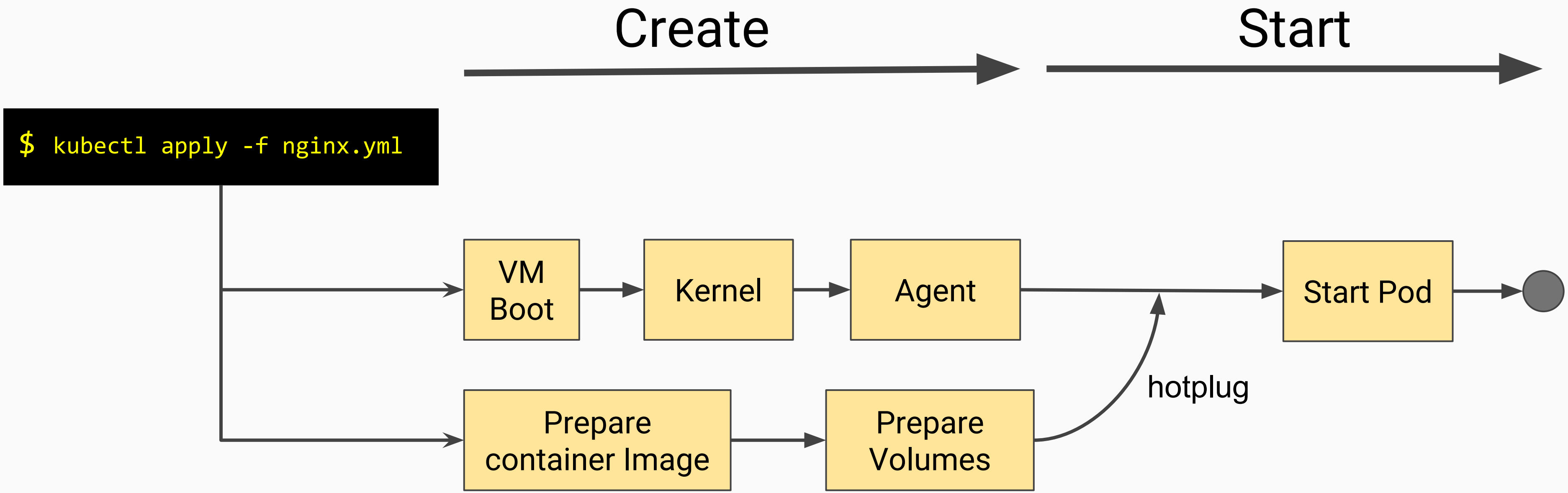
gRPC

gRPC

🟦 Hypervisor VSOCK socket

*Other names and brands may be claimed as the property of others.

Fast as a Container



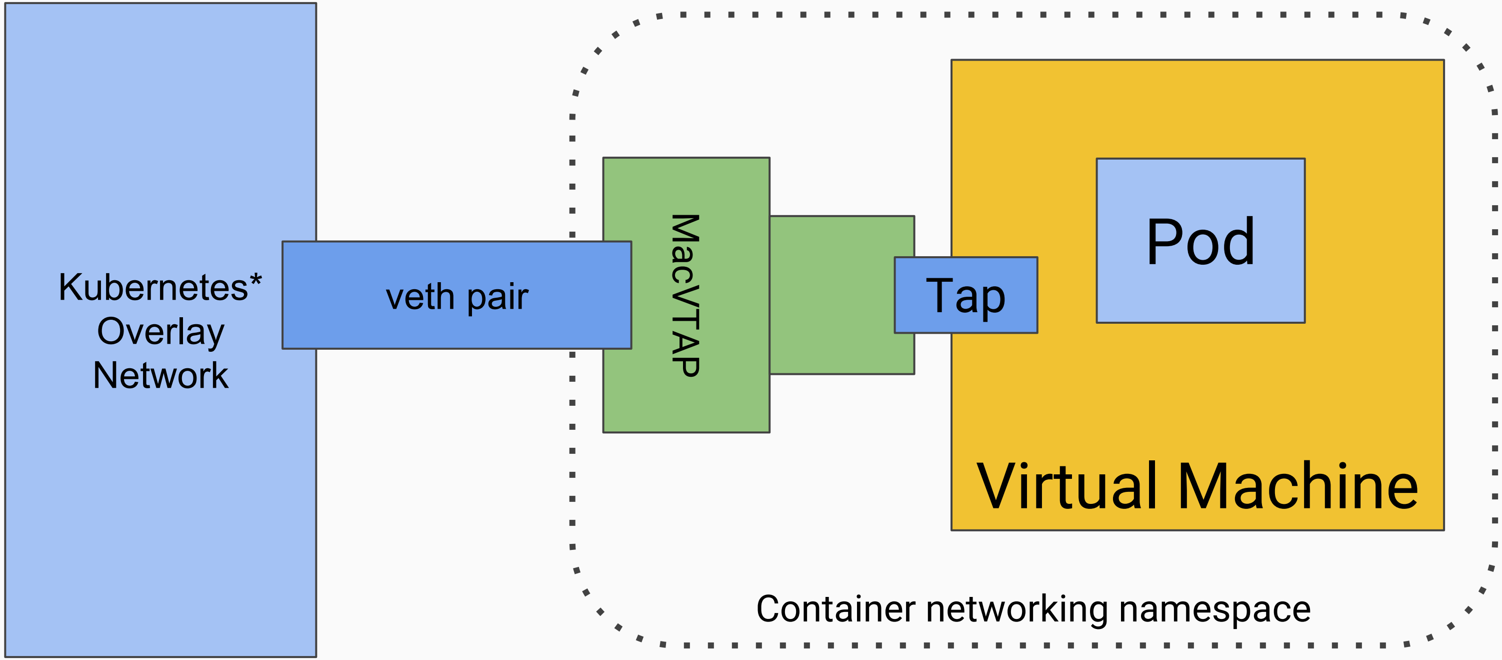
```
$ kubectl apply -f nginx.yml
```

Small as a Container

- Minimize memory footprint
  - Minimal rootfs
  - Minimal kernel
  - VM Template
  - DAX/nvdimm

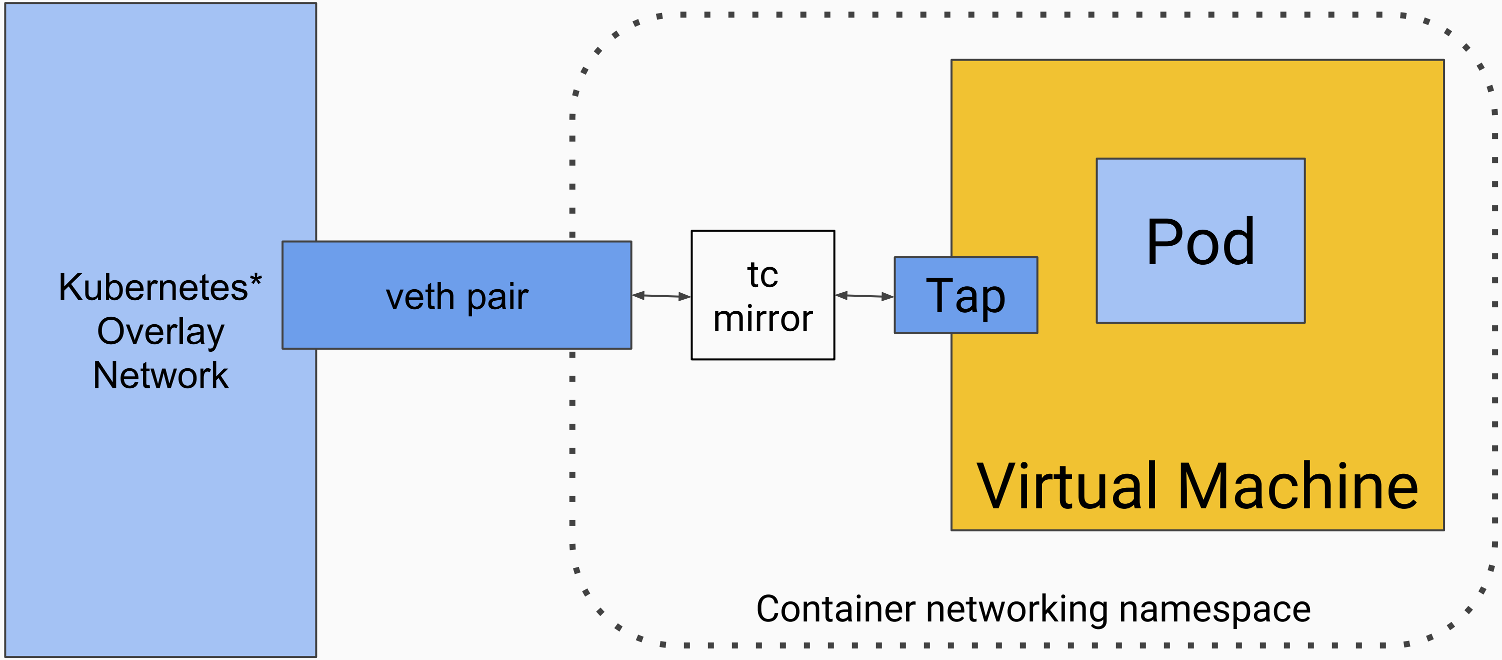
- De-duplicate memory across VMs
  - KSM (with throttling)

Networking



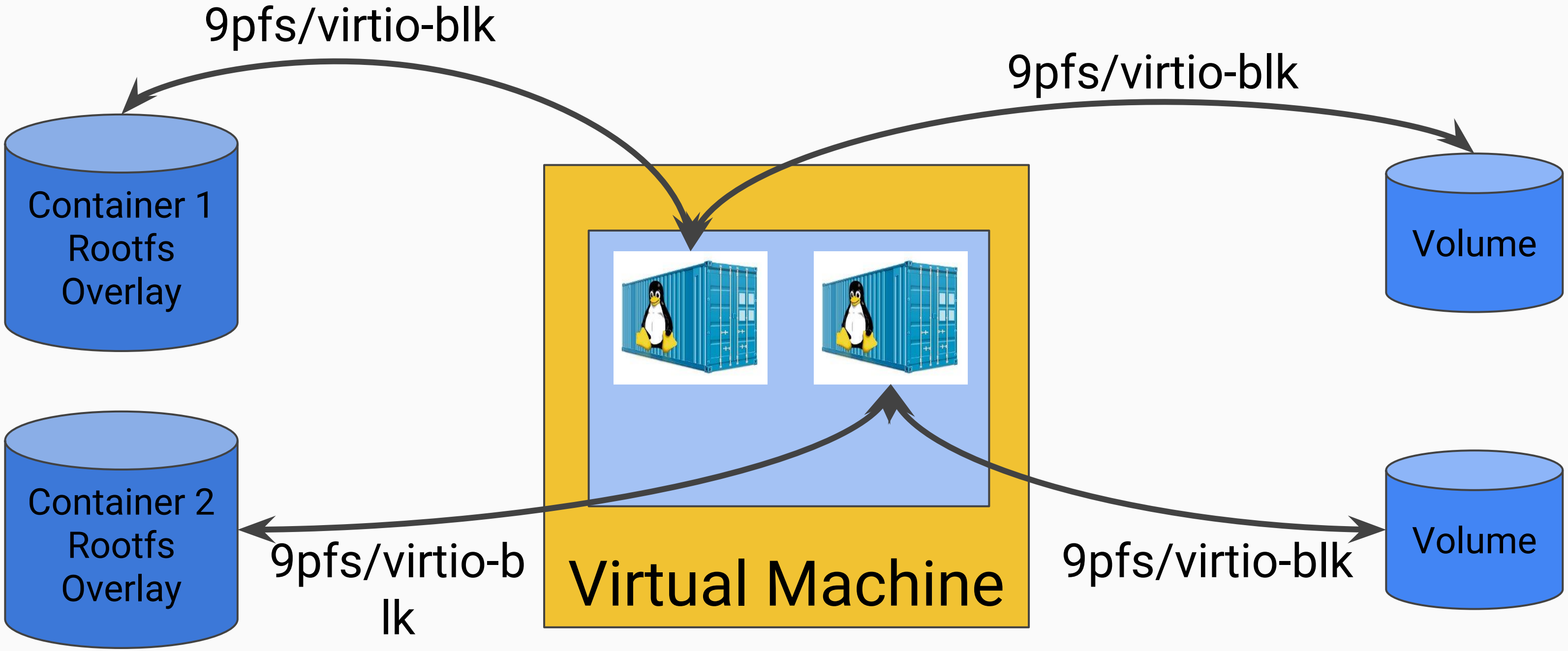
*Other names and brands may be claimed as the property of others.

Networking

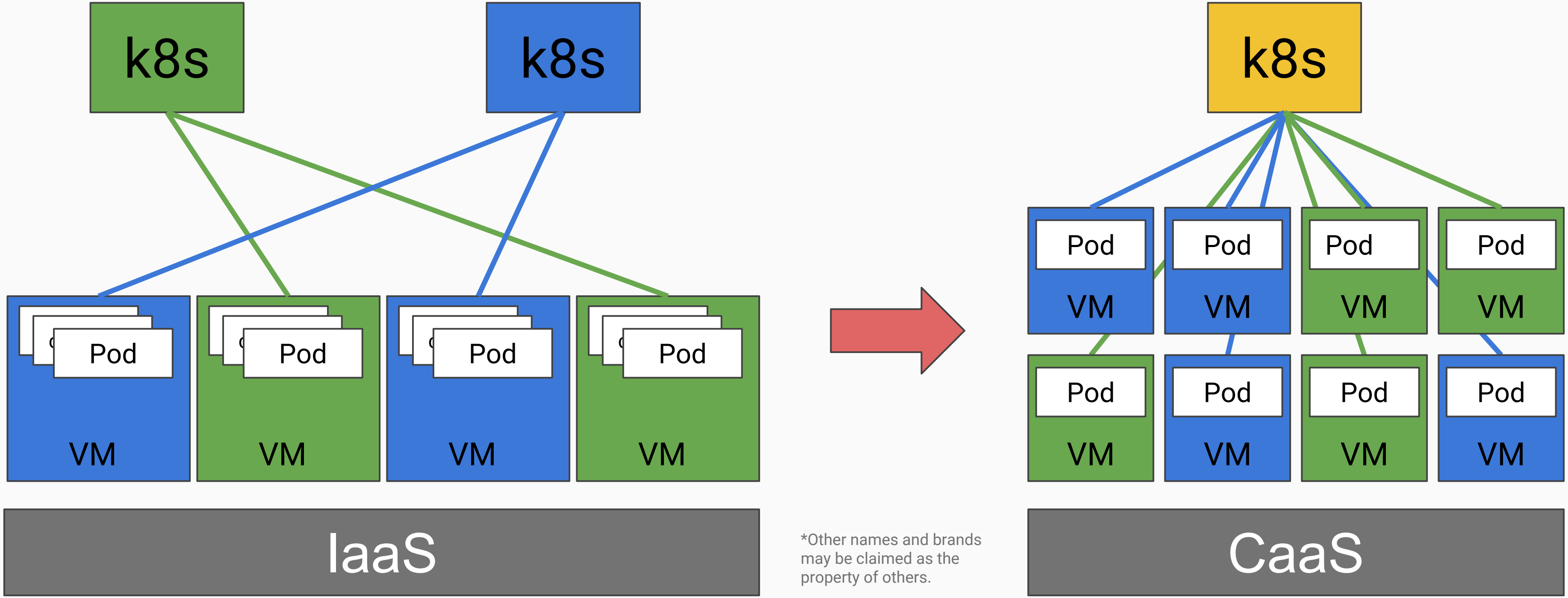


*Other names and brands may be claimed as the property of others.

Storage


9pfs/virtio-blk
9pfs/virtio-blk
Container 1
Rootfs
Overlay
Volume
Container 2
Rootfs
Overlay
9pfs/virtio-b
lk
Virtual Machine
9pfs/virtio-blk
Volume

Multi-tenant Kubernetes*



k8s k8s

Pod Pod Pod Pod

VM VM VM VM

IaaS

k8s

Pod Pod Pod Pod

VM VM VM VM

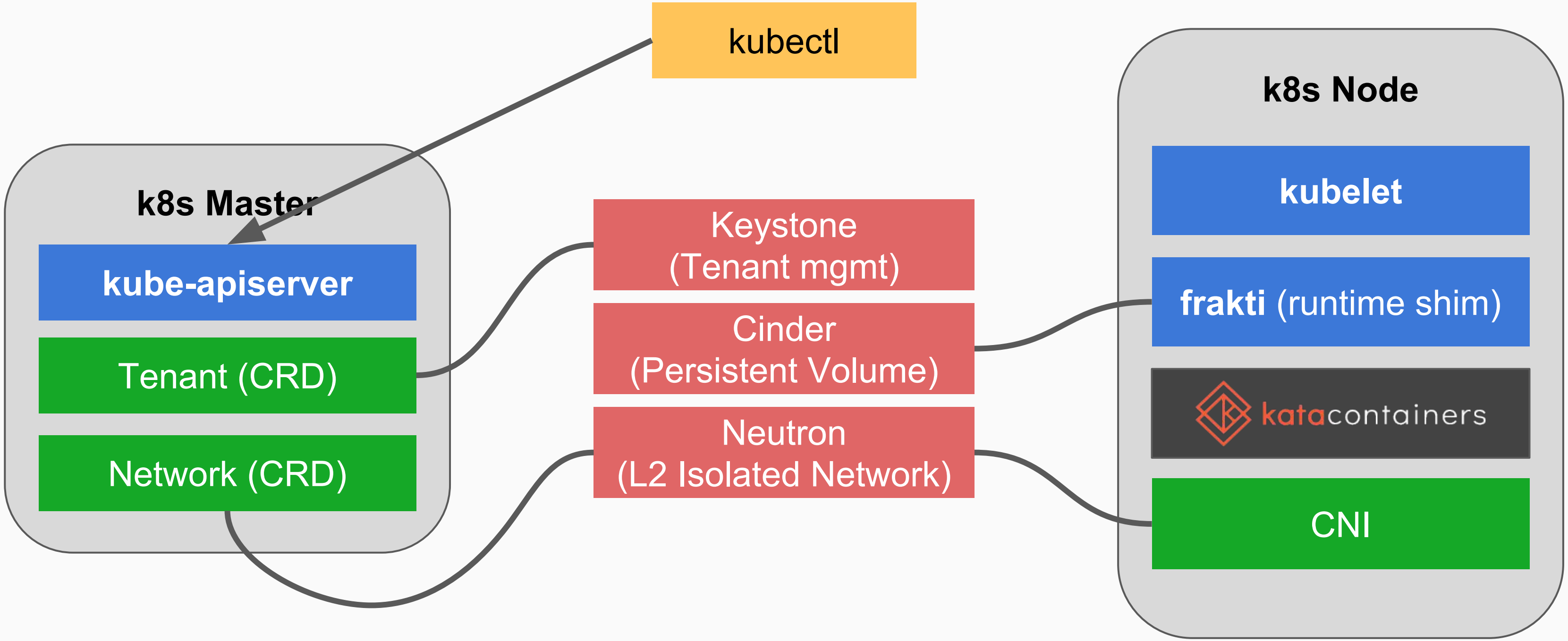
Pod Pod Pod Pod

VM VM VM VM

CaaS

*Other names and brands
may be claimed as the
property of others.

Demo: Stackube - K8S with Hard Multi-tenancy



What’s Next?

1H’2018 Horizon

- 1.0 Release (parity with RunV and CC 3.0 with upgrade path)
- CRI integration: Frakti, CRI-O, containerd-cri
- OCI runtime spec support for hypervisor based containers
- OSV support
- Documented case studies

Get Involved

Contribute

- Code and documentation hosted on https://github.com/kata-containers/
- Major releases managed through Github* Projects
- Intel (Intel® Clear Containers) & Hyper (runV) contributing initial IP
- Apache 2 license
- Slack: katacontainers.slack.com
- IRC: #kata-dev@freenode
- Mailing-list: kata-dev@lists.katacontainers.io

*Other names and brands may be claimed as the property of others.

Where To Contribute?

	Role	Language	Upstream version	Host/Guest
Shim	I/O and signal handling between the host and the VM	Go	N/A	Host
Proxy	I/O and signal multiplexing (optional, serial connection)	Go	N/A	Host
Runtime	OCI commands handling. VM, shim, and proxy startup	Go	N/A	Host
QEMU	Hypervisor	C	2.9	Host
Agent	Guest containers manager	Go	N/A	Guest
Guest Kernel	Boot to systemd/Boot initrd	C	4.13.13	Guest
Guest image	Minimal Linux root filesystem that starts the agent	N/A	*Pick your image*	Guest

Open Governance

- **Contributors**
    - At least one github contribution for the past 12 months
- **Maintainers**
    - Active contributor, nominated by fellow maintainers
    - Can merge code
- **Architecture Committee**
    - Take high level architecture and roadmap decisions
    - 5 seats, elected by contributors

Thank you!

Disclaimer