
KubeCon
CloudNativeCon
OPEN SOURCE SUMMIT
China 2023


# Kubernetes on Bare-metals or VMs?
# Or Using Virtink to Have Pros of Both

*Feng YE (@fengye87), SmartX & Ailin YANG (@ailun258), Intel*

# Kubernetes on Bare-metals



**Pros**

- No performance overhead of virtualization
- No operational costs of virtualization

**Cons**

- Not easy to deploy & scale clusters
- Unsafe to share using clusters
  - Kernel is shared between containers and hosts
- Inconvenient sharing clusters
  - No built-in multi-tenant support

# Kubernetes on VMs



**Pros**

- Highly isolated
- Higher hardware utilization
- Easy to deploy & scale clusters
  - Cluster API

**Cons**

- Enterprise products can be costly
- Overrich feature set
  - Legacy hardware emulation is less efficient and unsafe
  - Complicated to operate

# Pros of both?



**Highly efficient virtualization**

- Lower overhead
  - Para-virtualization (virtio)
  - SR-IOV
- Less burden
  - Avoid keyboard, mouse or monitor
  - Remove legacy hardware support

**Lightweight VM management**

- Easier operations
  - Minimal feature set
  - On Kubernetes, for Kubernetes
- Lower management overhead
  - Minimize per VM management memory overhead

# What is Virtink



Virtink is a Cloud Hypervisor Add-on for Kubernetes

# About Cloud Hypervisor



**•VMM with focus on running modern "Cloud Workloads" only**

**•Open Source since 2019:**
https://github.com/cloud-hypervisor/cloud-hypervisor

**•Written in Rust utilising Rust-VMM components**

**•Goals:**

•Minimal device emulation – paravirtualised devices instead

•Opinionated feature set to ensure ease of use

•Designed with security in mind

•Usable for "pet" VMs as well as for integrating with Kata Containers

**•Contributions from multiple companies:**

•Intel, Alibaba, Microsoft, ARM, Bytedance, Oracle, Ericsson, Phytium, Red Hat, Ant Financial, ZTE, Smartx,Tencent

**•Governance:**

•Linux Foundation Project

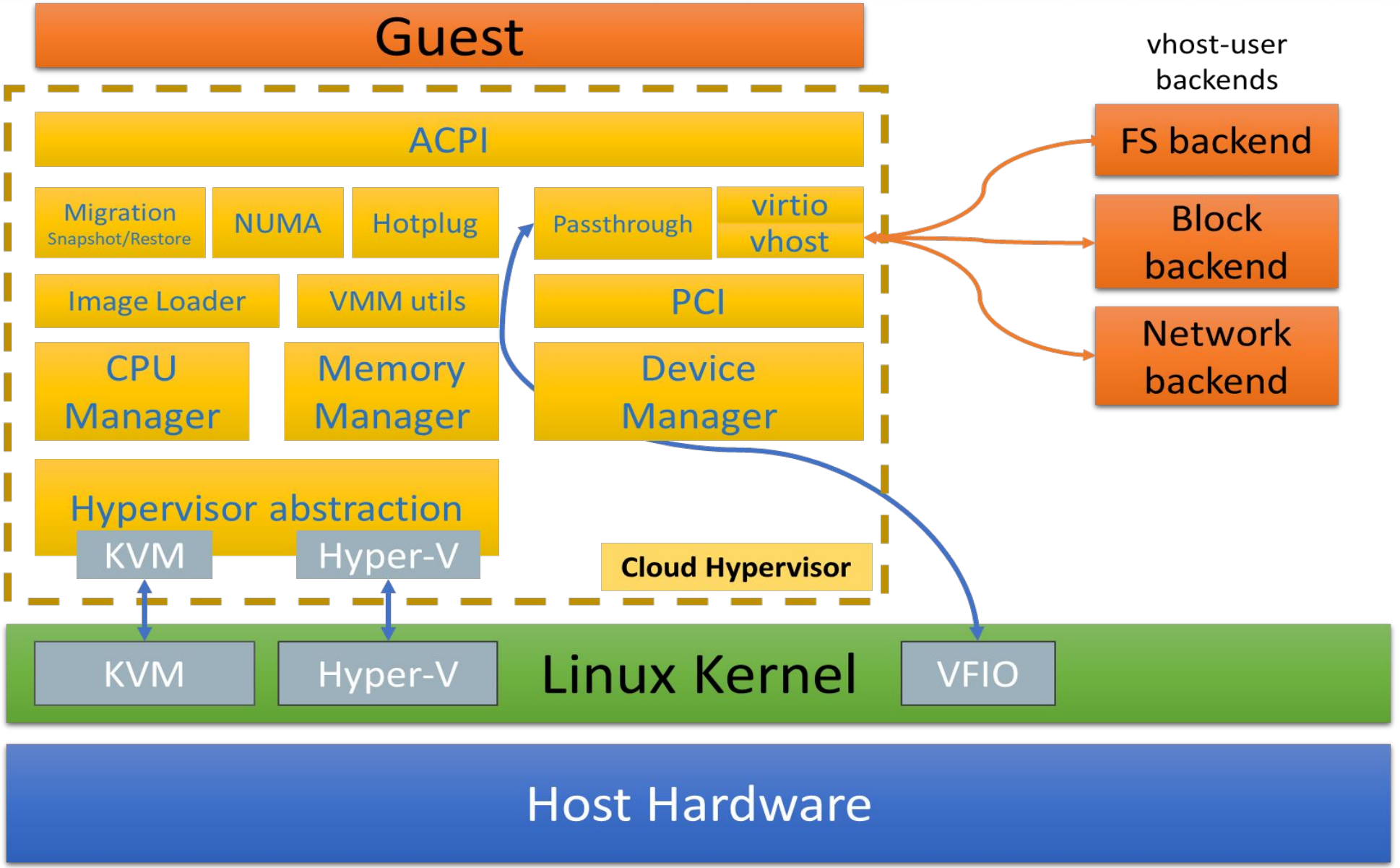
•Founding members: Intel, ARM, Microsoft, Alibaba, ByteDance, +Tencent and Ampere

•Technical committee of key contributors to settle potential disputes

• Advisory Board of technologists

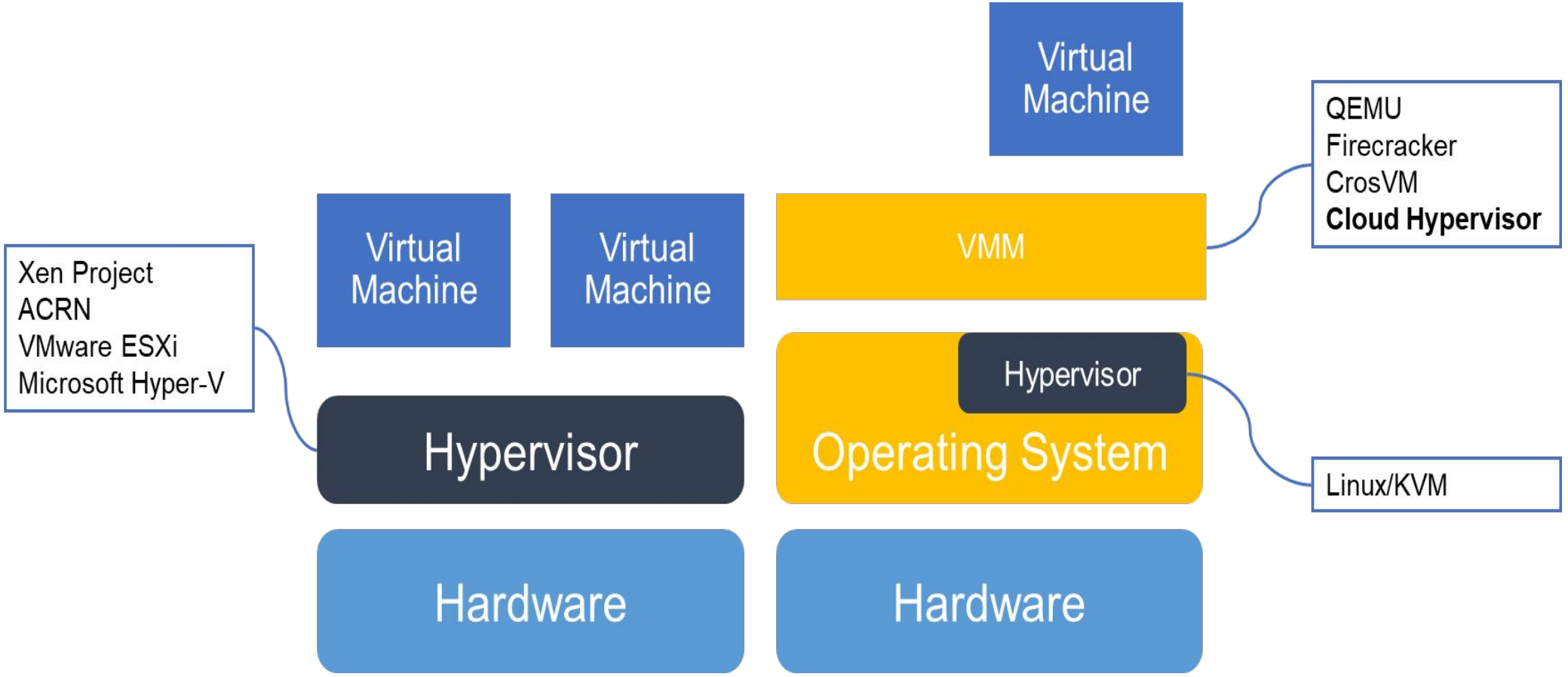
# Cloud Hypervisor Architecture





# Comparing Cloud Hypervisor to Others

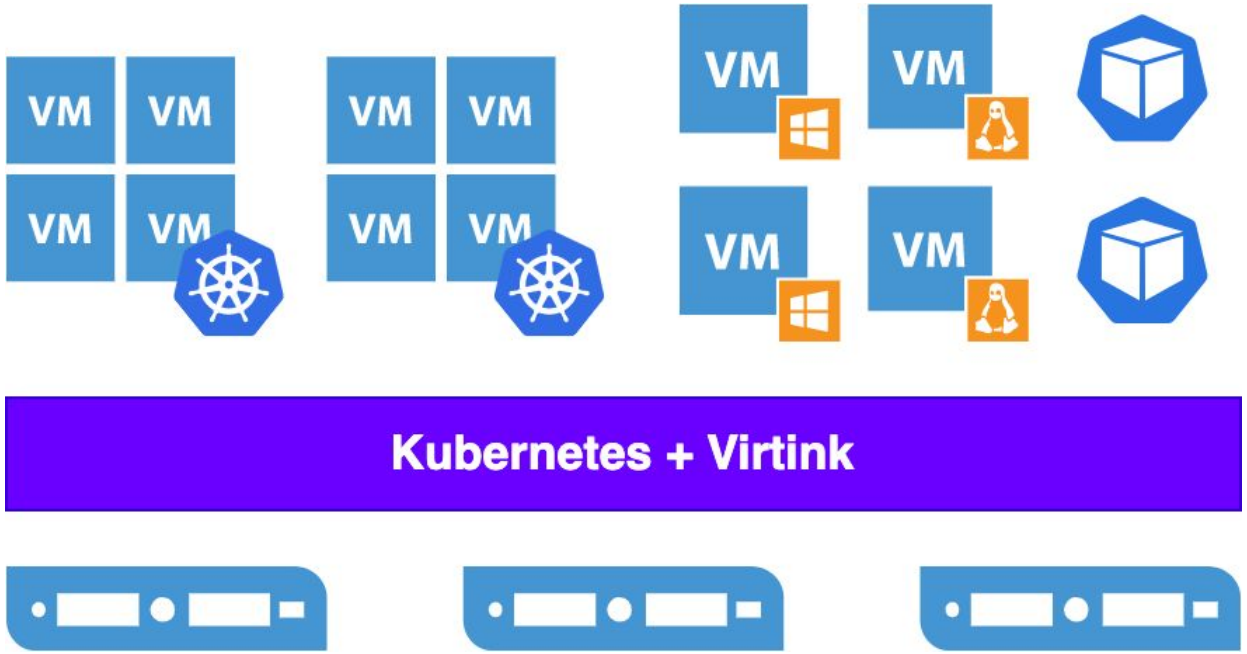




# Virtink Design Goals

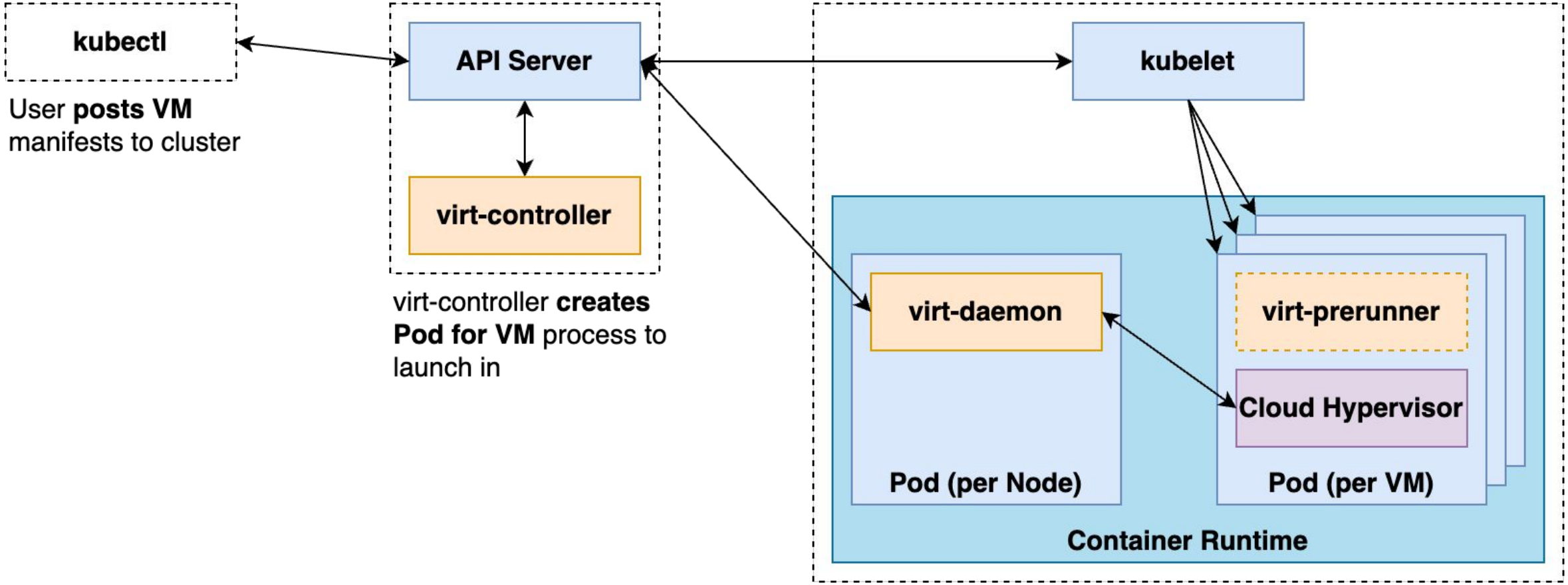


- Runs on (almost) any Kubernetes clusters
  - No CRI replacement
  - Non-invasion
- Kubernetes native
  - CSI for VM storage
  - CNI for VM network
- Support nested Kubernetes



# Virtink Architecture





# Comparing Virtink to Others



- Virtink replaces QEMU (and libvirt) with Cloud Hypervisor
  - **≥30MB** less per VM memory overhead
- No long-running management process in VM Pod
  - **≥ 80MB** less per VM management memory overhead
- Minimized feature set
  - No shared-storage based VM HA
  - No device hot-plugging

# Virtink Demo



https://asciinema.org/a/509484

# knest: Turnkey Nested Kubernetes Tool



https://asciinema.org/a/509497

# Summary



- Kubernetes on bare-metals or VMs both have pros & cons
- To achieve high performance and isolation at the same time, a more efficient and lightweight virtualization is required

**More efficient virtualization: Cloud Hypervisor**

- Rust for memory safety
- virtio for efficient IO
- Minimized attack face

**More lightweight VM management: Virtink**

- **≥100MB** per VM memory overhead
- Minimized feature set
- On Kubernetes, for Kubernetes
- A turnkey guest Kubernetes tool

# Thanks 🙏 & Questions 🙋?