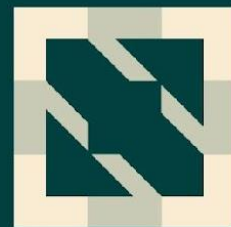




**KubeCon**



**CloudNativeCon**



**OPEN SOURCE SUMMIT**

**China 2023**



KubeCon



CloudNativeCon



OPEN SOURCE SUMMIT

China 2023

# 在裸金属或虚拟机上使用 Kubernetes？或者使用 Virtink 来兼具两者的优点

叶丰 (@fengye87), SmartX & 杨爱林 (@ailun258), Intel

# 裸金属部署 Kubernetes

## 优势

- 避免虚拟化的性能开销
- 省去虚拟化的运维费用

## 不足

- 部署、节点增删等相对不便
- 共用集群不安全
  - 容器和主机共享内核
- 共用集群不方便
  - 缺少多租户支持

# 虚拟机部署 Kubernetes

## 优势

- 隔离性更好
- 硬件使用率更高
- 集群部署、节点增删更方便
  - Cluster API

## 不足

- 商业化产品定价高昂
- 主流产品功能过于丰富
  - 过时硬件的模拟, 低效且不安全
  - 平台运维复杂



# 兼顾两者优势？

## 更高效的虚拟化技术

- 提高效率
  - 半虚拟化设备 (virtio)
  - 硬件级虚拟化支持 (SR-IOV)
- 减轻负担
  - 去掉键盘、鼠标、监视器等不必要的外设支持 ⇒ 更低的内存开销
  - 去掉过时硬件的模拟支持 ⇒ 更小的攻击面

## 更轻量的虚拟机管理

- 降低运维难度
  - 简化功能集合
  - On Kubernetes, for Kubernetes
- 控制管理开销
  - 最小化 per VM 管理所需内存

# Virtink 是什么

Virtink 是管理 Cloud Hypervisor 虚拟机的  
Kubernetes 扩展

# 关于 Cloud Hypervisor

- **VMM,专注于运行现代“云工作负载”**

- 2019年开源:

<https://github.com/cloud-hypervisor/cloud-hypervisor>

- **使用Rust VMM组件, 使用Rust语言编写**

- **目标:**

- 最少的设备模拟—改为半虚拟化设备
- 可选功能集, 确保易用性
- 设计时特别考虑安全性
- 可用于虚拟机示例以及与Kata Containers集成

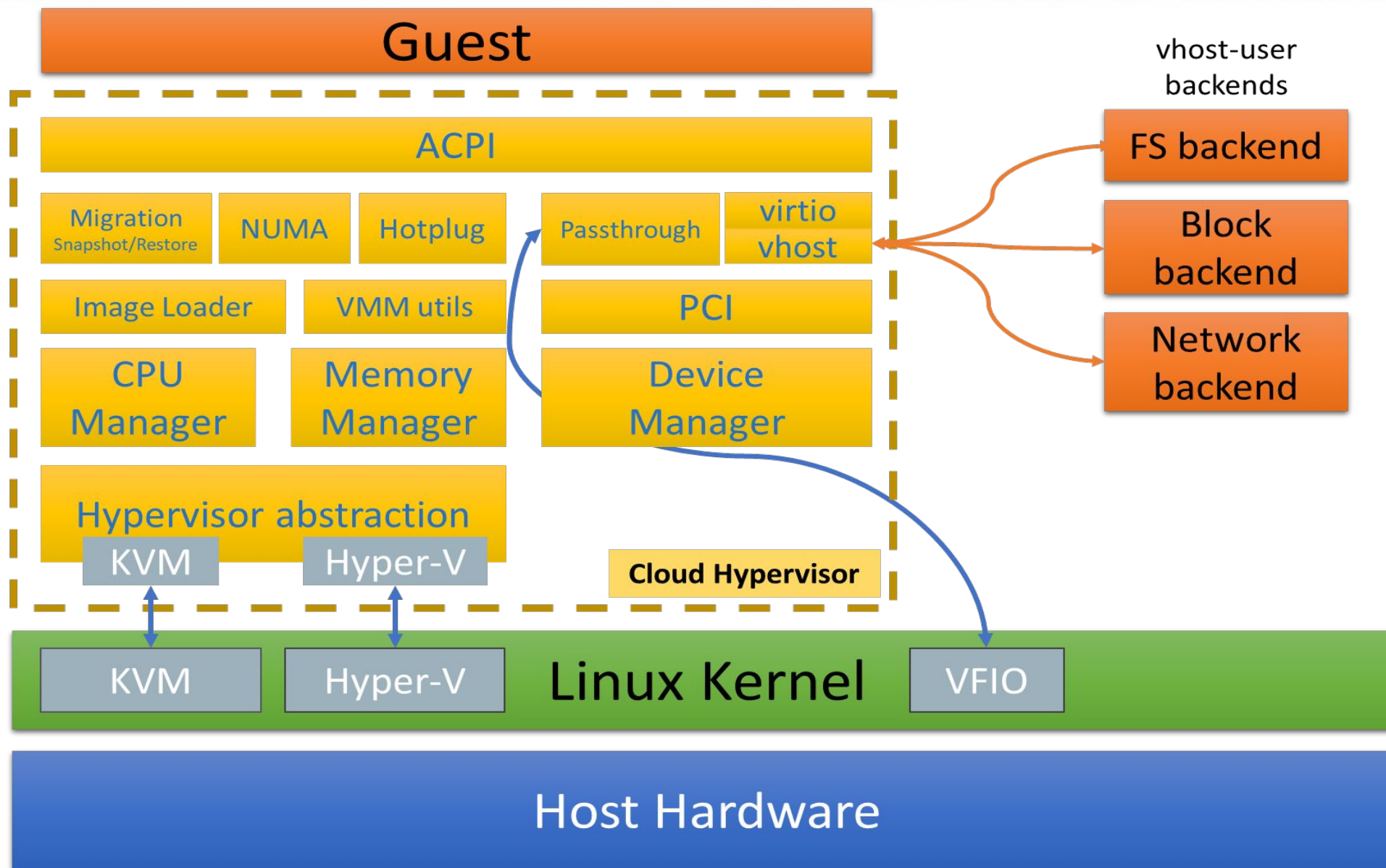
- **代码贡献来自多家公司:**

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

- **治理结构:**

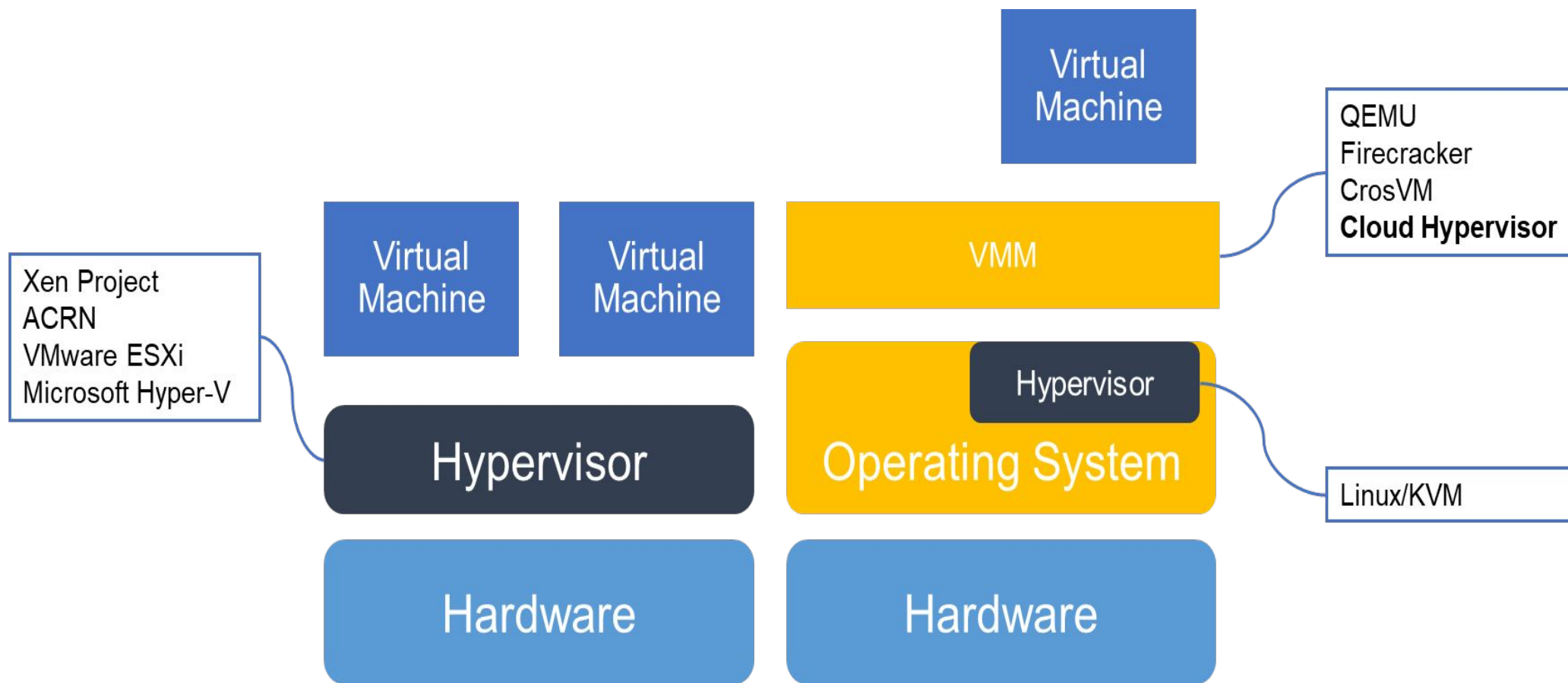
- 捐献给托管于Linux基金会
- 项目发起组织: Intel, ARM, Microsoft, Alibaba, ByteDance, +Tencent and Ampere
- 主要贡献者技术委员会
- 技术专家咨询委员会

# Cloud Hypervisor 架构图



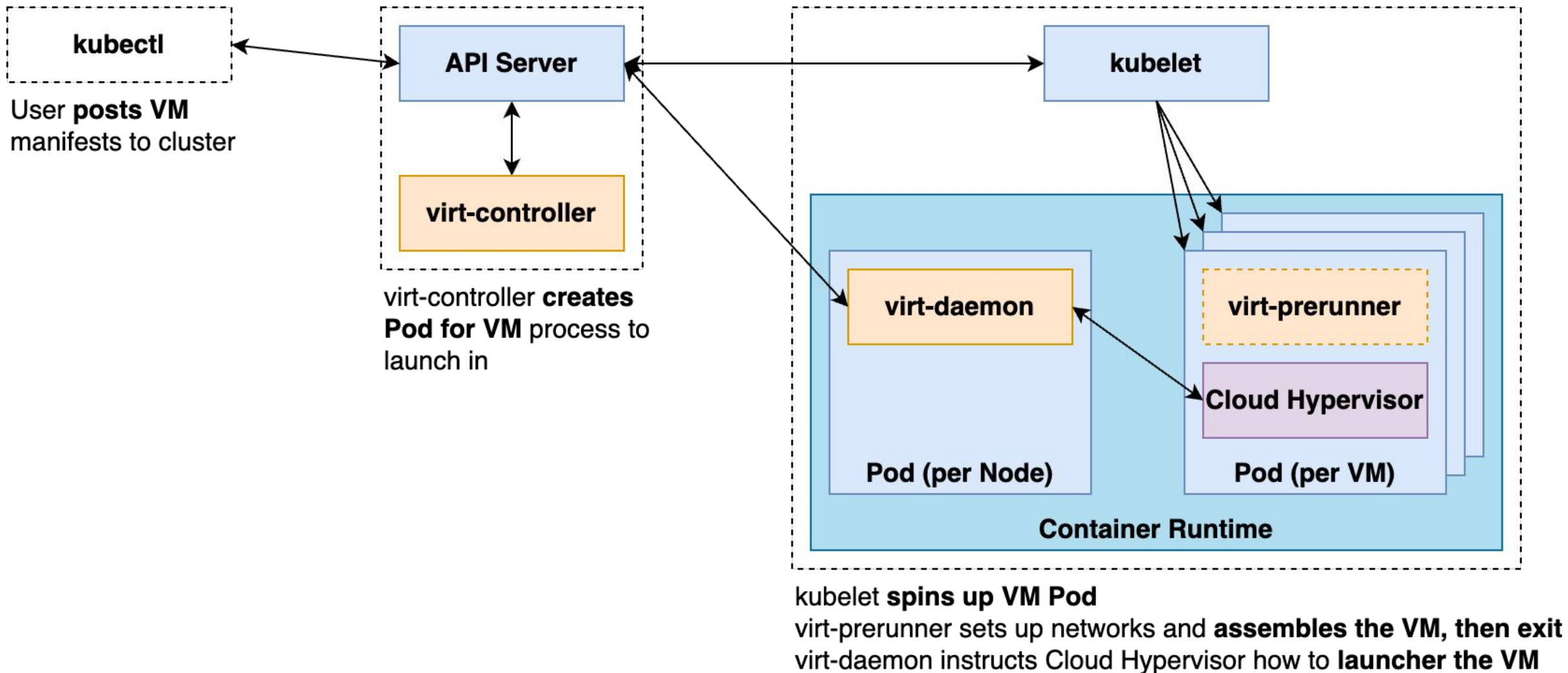


# Cloud Hypervisor 和 Others



-

# Virtink 架构图



# Virtink 对比其他

- Virtink 使用 Cloud Hypervisor 代替 QEMU (和 libvirt)
  - 减少  $\geq 30\text{MB}$  per VM 内存占用
- Virtink 的 VM Pod 中没有常驻管理进程
  - 减少  $\geq 80\text{MB}$  per VM 内存占用
- Virtink 仅提供精简功能集合
  - 不提供基于共享存储的虚拟机高可用功能
  - 不支持磁盘或网卡的热插拔

# Virtink 演示

<https://asciinema.org/a/509484>



# knest: 一键式 Nested Kubernetes 部署工具

<https://asciinema.org/a/509497>

- 裸金属或者虚拟机部署 Kubernetes 各有优劣
- 兼顾裸金属的性能和虚拟机的隔离性, 需要更高效的虚拟化技术和更轻量的虚拟机管理

## 更高效的虚拟化技术: Cloud Hypervisor

- Rust 确保内存安全
- virtio 提高 IO 效率
- 减少外设支持获得更小的攻击面

## 更轻量的虚拟机管理: Virtink

- 去掉 launcher 和 libvirt, 减少 **≥100MB** per VM 内存占用
- 精简功能集合
- On Kubernetes, for Kubernetes
- 提供 knest 一键式 guest Kubernetes 部署工具

感谢聆听🙏 & 欢迎提问🙋！