Kubernetes VM Solutions for Multi-Tenant Applications

Guangxu Li, Senior Software Engineer, ZTE
li.guangxu@zte.com.cn
Container and VM Ecosystem

- Kubernetes
- Docker Swarm
- Marathon
- Nomad
- OpenStack
- Others
Why We Run VM on Kubernetes?

- Traditional Applications
- No linux based Applications
- Functions provided by host kernel are not satisfied
- OpenStack is too complex
- Unified infrastructure
- Better isolation
VM related Projects

- Virtlet
- KubeVirt
- RancherVM
- Kata Container

Focus: deploy REAL vm (traditional vm app)

Focus: container security
Virtlet is a Kubernetes runtime server which allows you to run VM workloads, based on QCOW2 images.

https://github.com/Mirantis/virtlet
Virtlet compares with other CRI
Virtlet Architecture
Virtlet Deploying Objects

- DaemonSet
- ConfigMap
- ClusterRole/Role
- Service Account

virtlet solution
Virtlet Pros

- define VM as Pod
- supports using multiple interfaces
- SR-IOV

**NFV Environments**
Virtlet Cons

- limited storage options
- more configurations
- VM actions limited by Pod
Building a virtualization API for Kubernetes

https://github.com/kubevirt
KubeVirt Application Layout

KubeVirt Components
• virt-controller
• virt-handler
• libvirtd

KubeVirt Managed Pods
• VMI Foo
• VMI Bar
KubeVirt Pros & Cons

Pros
• Kubernetes cluster addon
• freedom - not limited by Pod definition

Cons
• VMs need to be managed separately from kubelet
• a new controller
• much bigger codebase
Package and run KVM images as Kubernetes pods, run at scale.

https://github.com/rancher/vm
RancherVM Architecture

Bare Metal Server 1
- ui-frontend
- ui-backend
- ip-controller
- kubelet
- apiserver

Bare Metal Server 2
- ui-frontend
- ui-backend
- ip-controller
- kubelet
- apiserver
RancherVM Networking
Container Security

kata containers + gVisor = NFV?
Kata Container

The speed of containers, the security of VMs

https://github.com/kata-containers
Kata Container Architecture

Diagram showing the architecture with I/O and OCI cmd/spec flowing through a Shim, Runtime, Proxy, virtual machine, and Hypervisor.
How to use kata container?
k8s + docker + kata not easy

kubernetes(dockershim) does not support to choose OCI runtime
k8s + docker + kata not easy

kata container network hotplug (support now)

Dockershim / Docker

- create pause container
- get container netns
- create net resources in netns

Containerd

- create netns
- create net resources in netns
- create pause container and app container

Cri-o
k8s + docker + kata create pod
k8s + docker + runc create pod
How ZTE Uses kata container in NFV

ZTE OpenPalette

`kubernetes based PAAS`

+ kata container 1.3
How ZTE Use kata container in NFV

ZTE Knitter
CNI based networking solution
gVisor is a user-space kernel that implements a substantial portion of the Linux system surface

https://github.com/google/gvisor
Why does gVisor exist?

- A single, shared kernel also mean that container escape is possible
- gVisor implements Linux by way of Linux
- Another approach to enhance container isolation
gVisor is special

Machine-level virtualization

Rule-based execution

gVisor
Technology landscape
DEMO