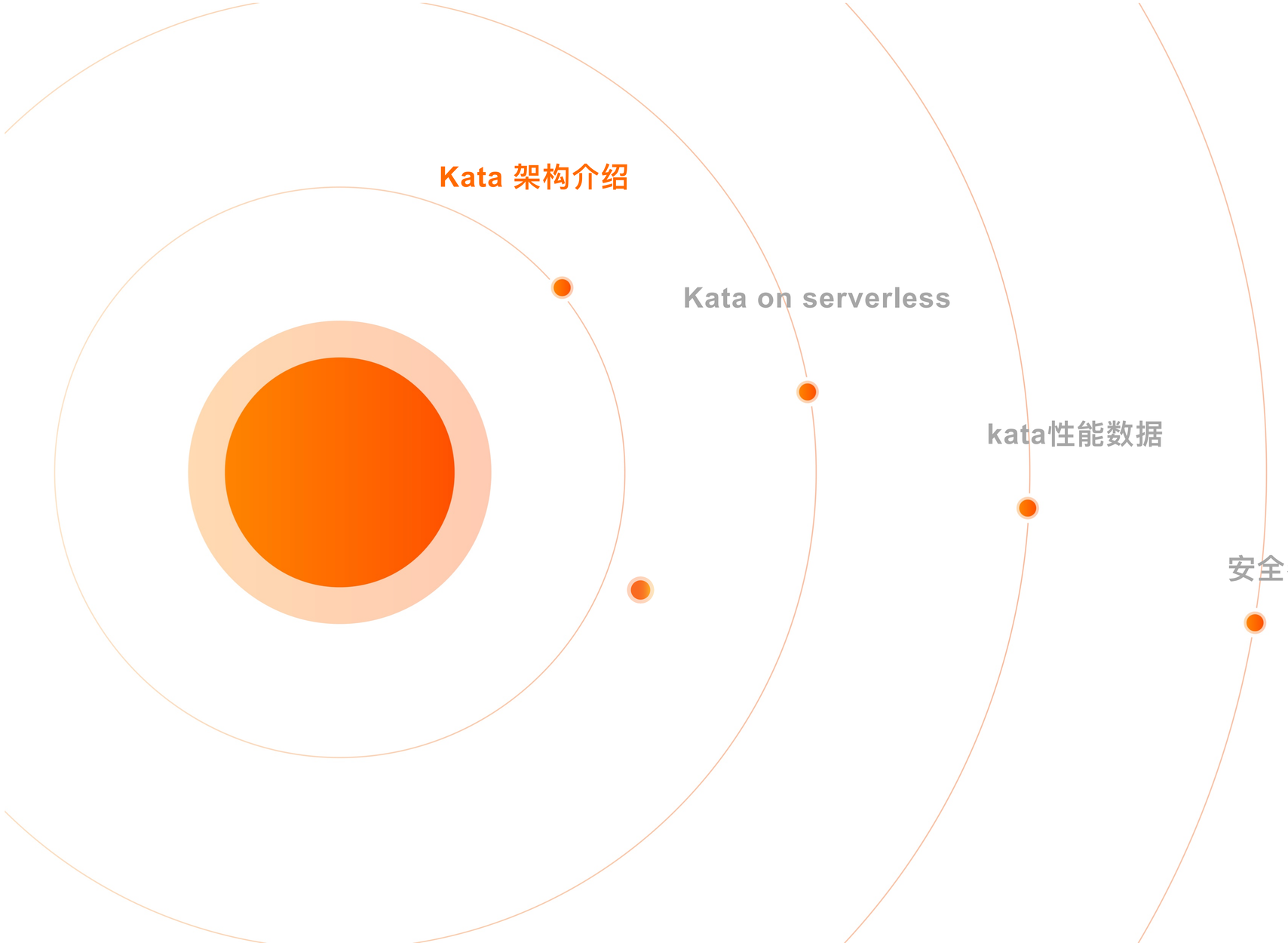
阿里云 | 奥运会全球指定云服务商

Kata Containers on Serverless

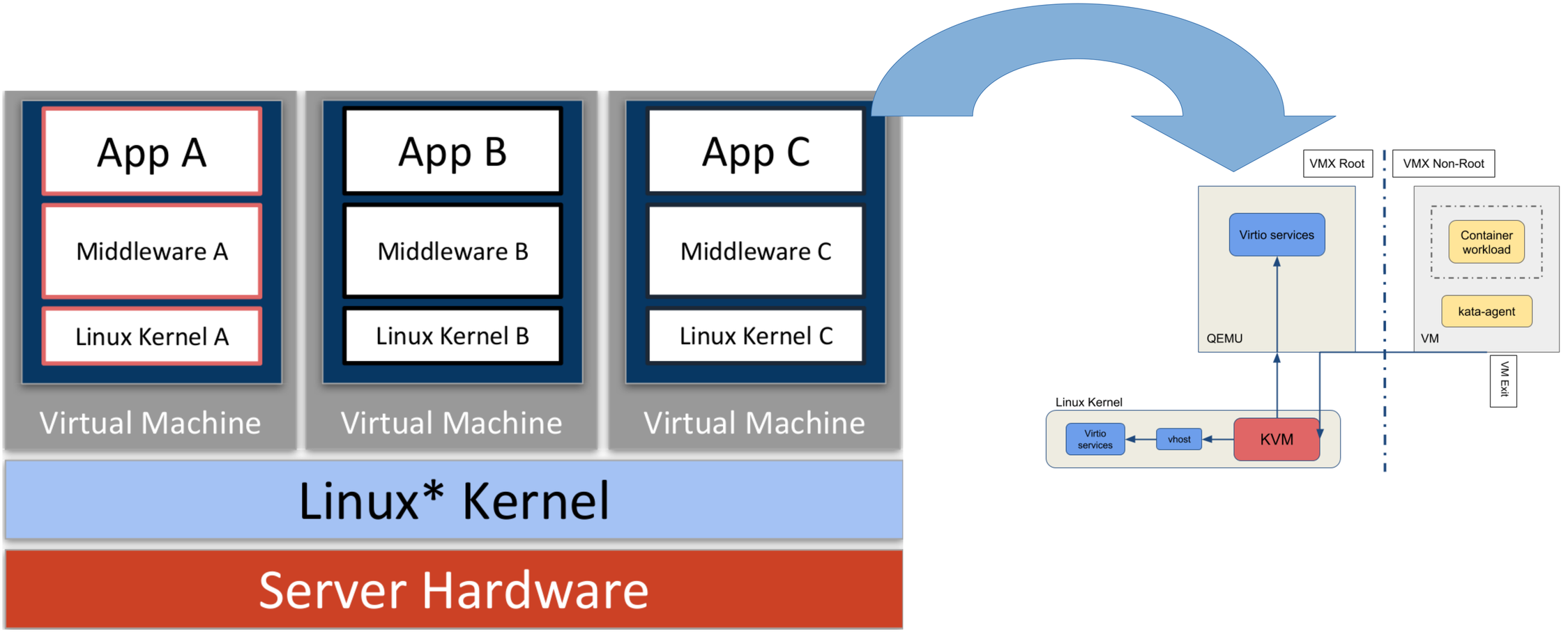
华敏

目录

阿里云 | 奥运会全球指定云服务商



Kata Container 架构



A example show how to run a Kata Container

Kata模块介绍

Kata Containers主要有6个模块

Kata agent： 该模块运行在虚拟机内部，负责在虚拟机内部传建容器和进程。

Kata runtime: 负责容器操作，创建/停止/更新等

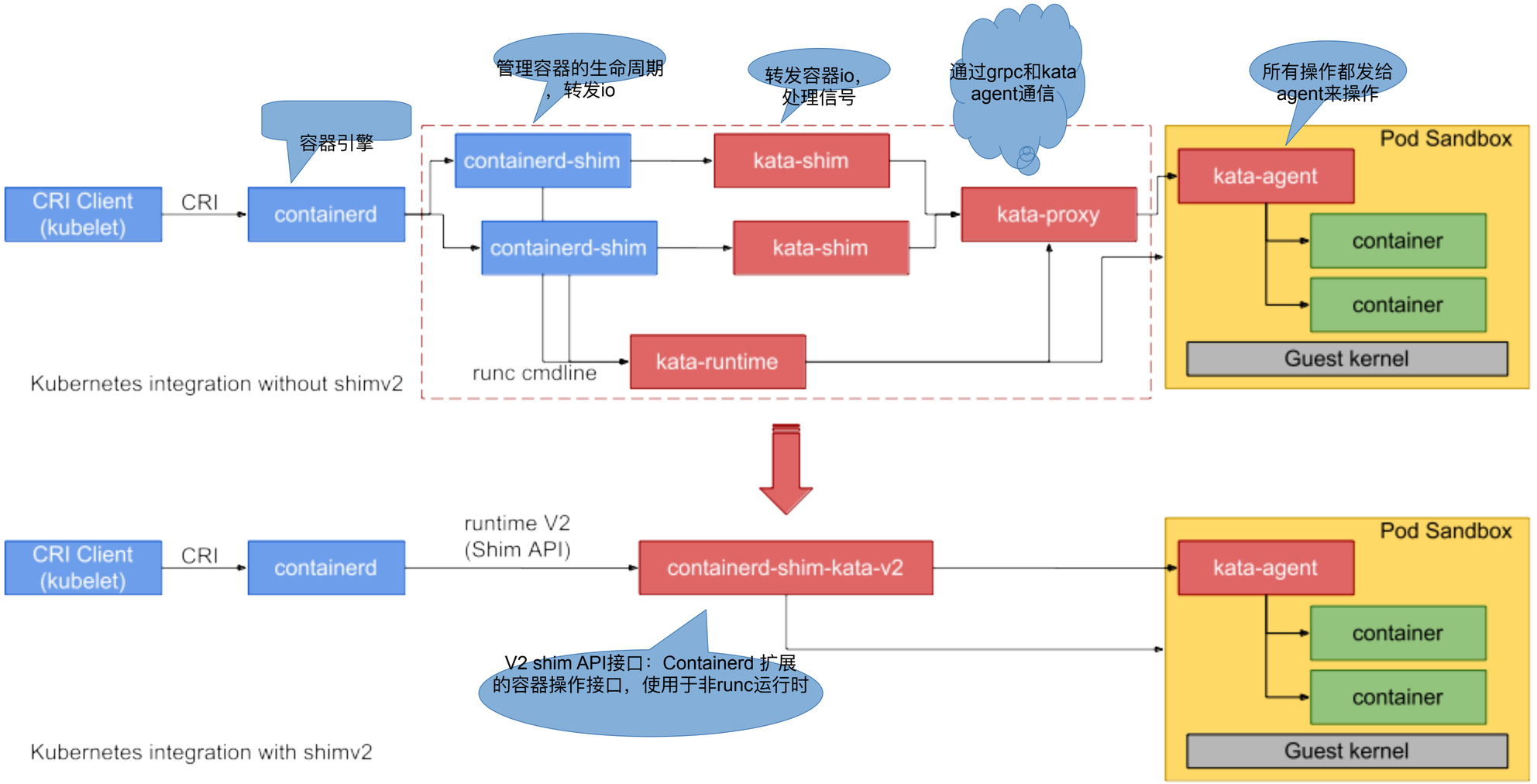
Kata proxy: 连接vm内的kata agent

Kata shim: 转发容器进程的I/O和信号

Guest Kernel：精简过的linux kernel

QEMU：用户态的虚拟机工作

Shim V2 API



kata Container On Serverless

阿里云 | 奥运会全球指定云服务商

Serverless 是什么：

1. 用户无需关心 server 端运维
2. 平台提供极致的自动扩缩容服务
3. 平台提供按需付费的特性

Kata container为Serverless 带来的优势

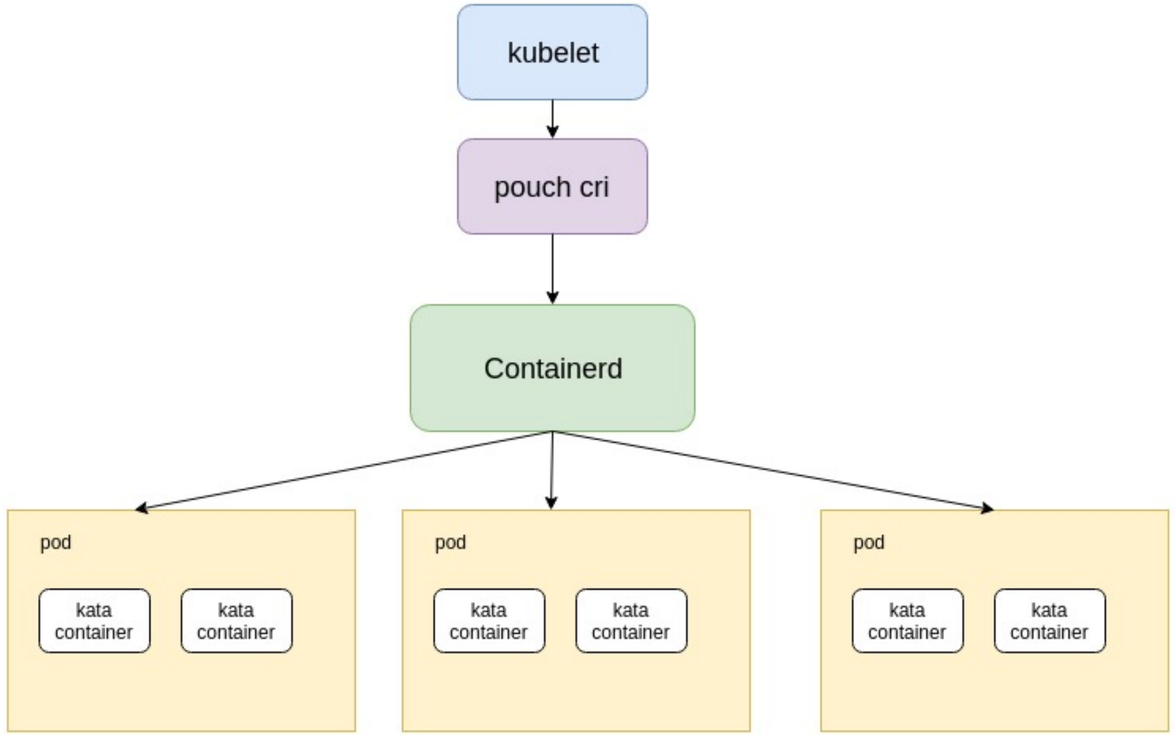
1. 更快的启动时间
2. 更安全的隔离

Kata 在serverless场景下的使用

阿里云 | 奥运会全球指定云服务商

kata带来的优势

1. 虚拟机提供更好的隔离性
2. 裁剪后guest kernel使启动时间更快，可以持平runc
3. 和容器技术结合，生产细粒度的资源



Kata boot time

阿里云 | 奥运会全球指定云服务商

kata容器 vm侧的启动时间

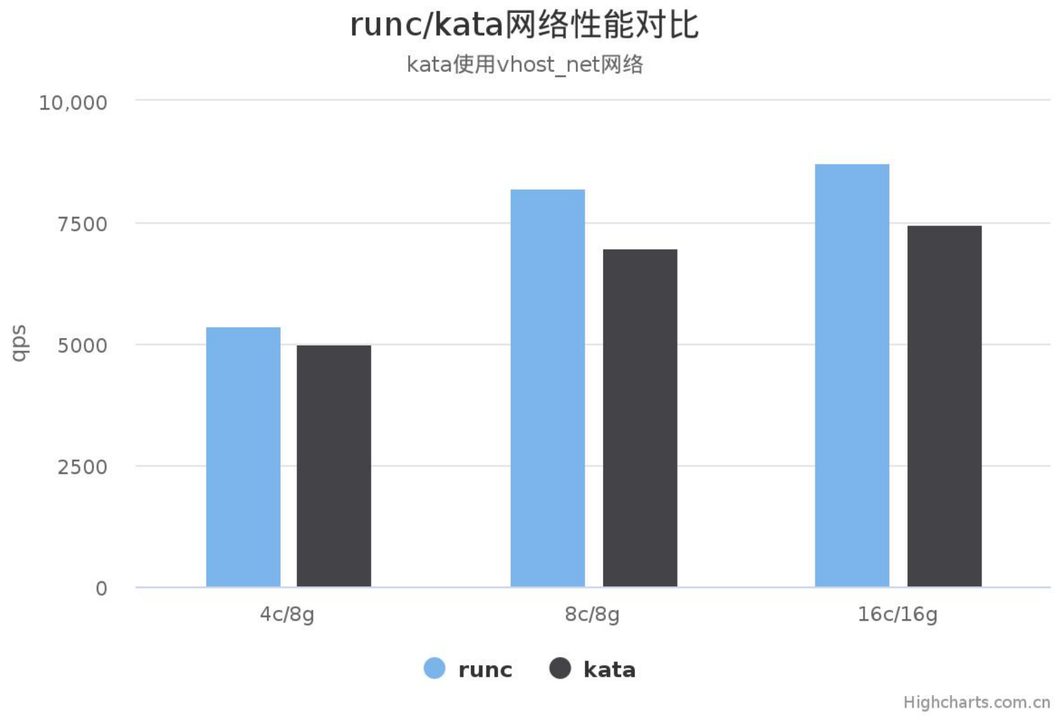
内核	第一次 total time（guest kernel/ usersapce）	第二次	第三次	第四次	平均
4.9 kernel	578.94(109.8/189.9)	567.86(113.9/188.6)	569.61(110.6/192.2)	568.35(109.6/191.0)	571.04(110.98/190.43)
clear container	525.22(67.2/190.3)	537.73(71.6/189.0)	537.35(63.9/212.5)	521.10(62.5/183.9)	530.35(66.3/193.92)

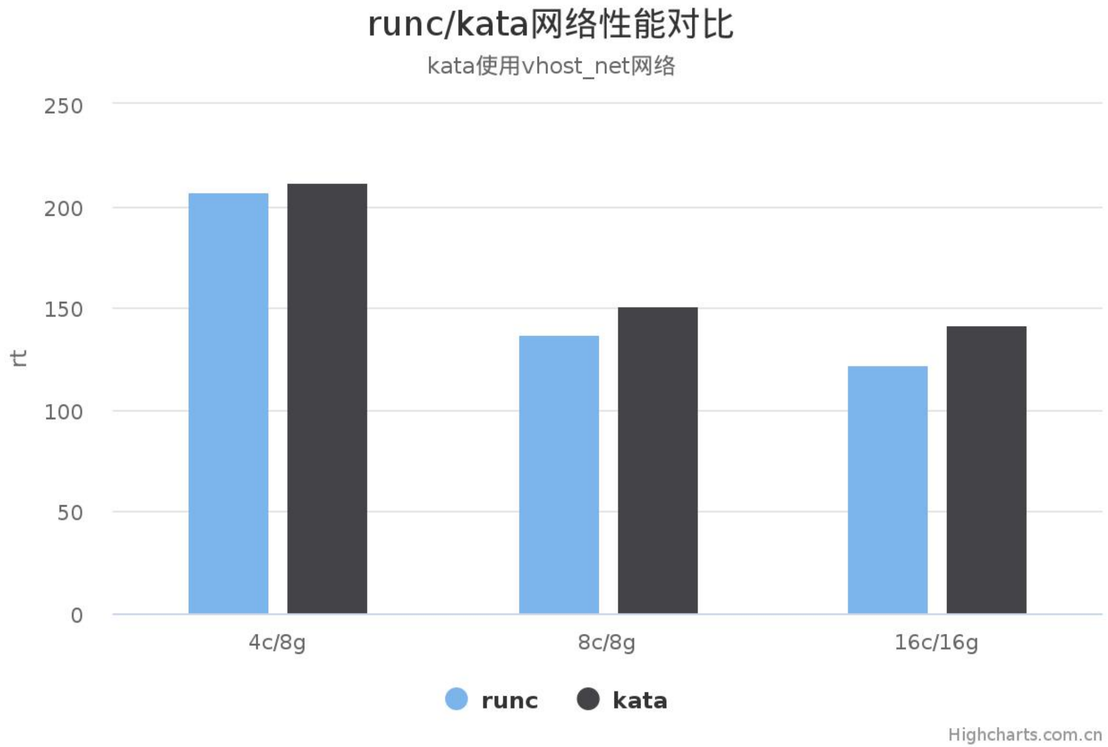
这里的启动时间由3部分构成: qemu boot time + guest kernel time + userspace time

网络性能测试

阿里云 | 奥运会全球指定云服务商

优化后，vhost_net网络开销和runc相比还是会相差5-15%，但是qemu使用直通网卡的话性能上和runc可以持平





IO性能

dd if=$dest_mnt/testfile of=/dev/null bs=4k count=16k iflag=direct

dd if=/dev/zero of=$dest_mnt/testfile bs=4k count=16k oflag=direct

storage	read	write
nfs	22.5	14.7
9pfs	147	120
vhost-9p	216	179

bs=128k, iodepth=128, numjobs=1, rw={read, write}

bs=4k, iodepth=32, numjobs=16, rw={randread, randwrite}

bs=4k, iodepth=1, numjobs=1, rw={read, write, randread, randwrite}

storage	randread	randwrite
nfs	40201	38048
9pfs	58797	97687
vhost-9p	126511	132077

IO性能

阿里云 | 奥运会全球指定云服务商

使用块设备作为rootfs来替代9pfs 优化io

Ngnix 测试性能

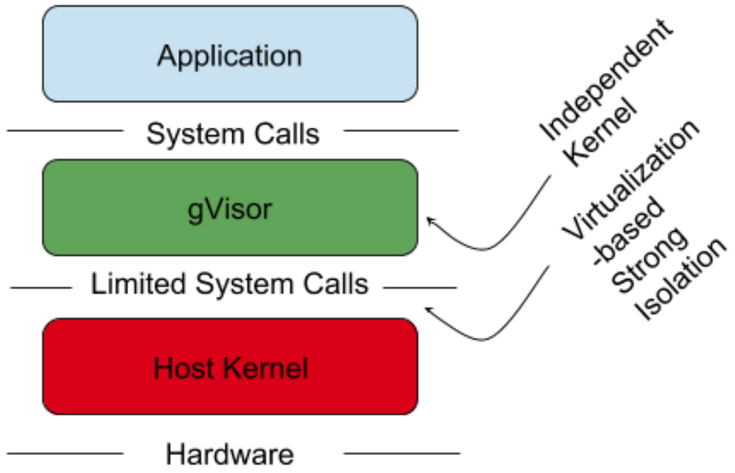
测试参数： 压测参数：ab -kn 100000 -c 100

	多核单nginx进程
runc	46509.73
kata+9p	1236.88
kata+block	47702.50（开启了host侧缓存）

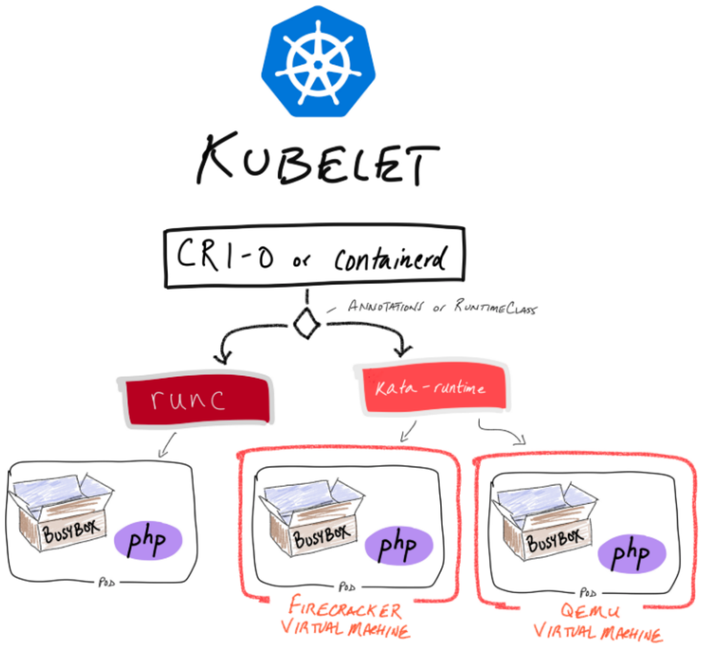
安全容器的未来

在我看来有2个方向

Kernel 优化 → 用户态内核 gvisor



Hypervisor 优化 → 更轻量的vmm，firecracker



阿里云 | 奥运会全球指定云服务商

Q & A