

Kata Containers: When Virtualization Meets Cloud Native

Xu Wang @ Ant Group

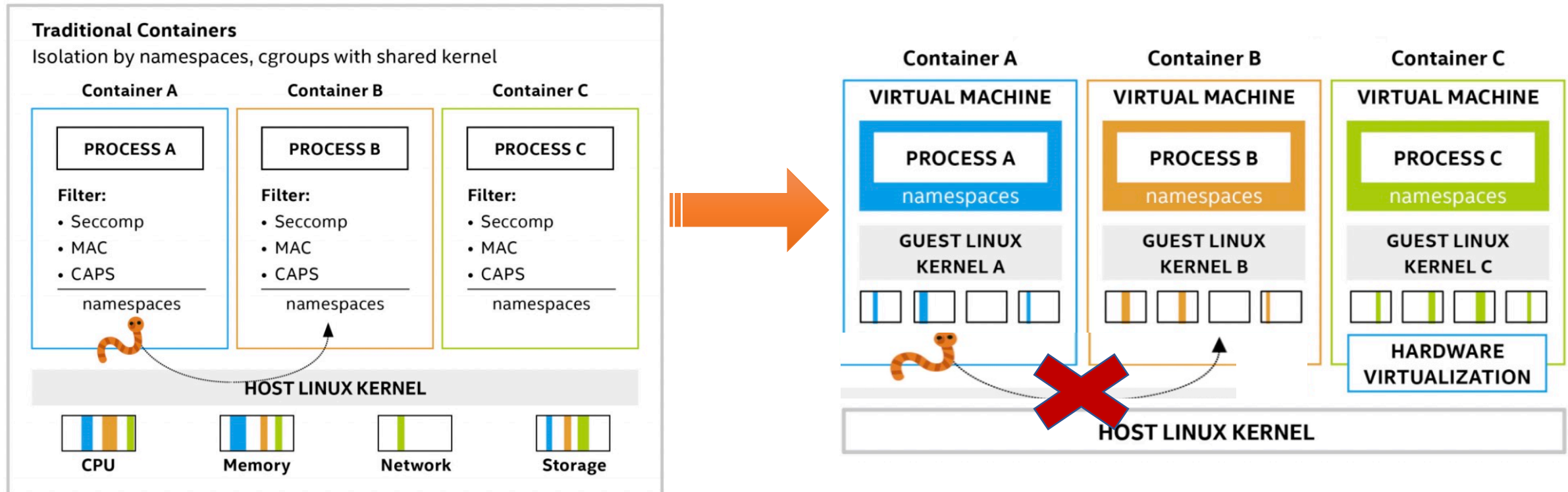
Kata Containers: When Virtualization Meets Cloud Native

Xu Wang @ Ant Group

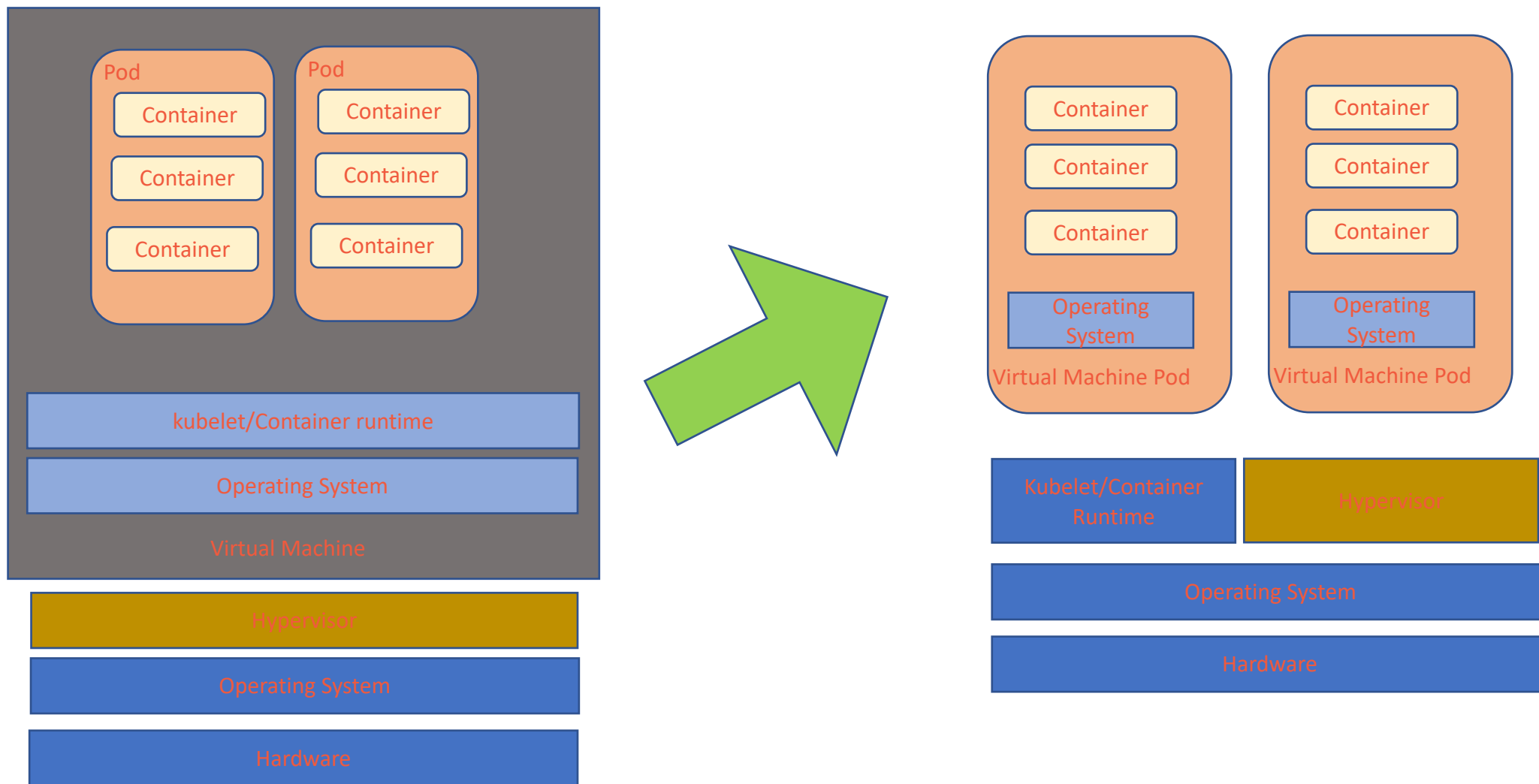
Index

- Kata Containers Introduction
- Kata Containers 2.0
- Kata @ Ant
- Future Works

What Is Kata Containers

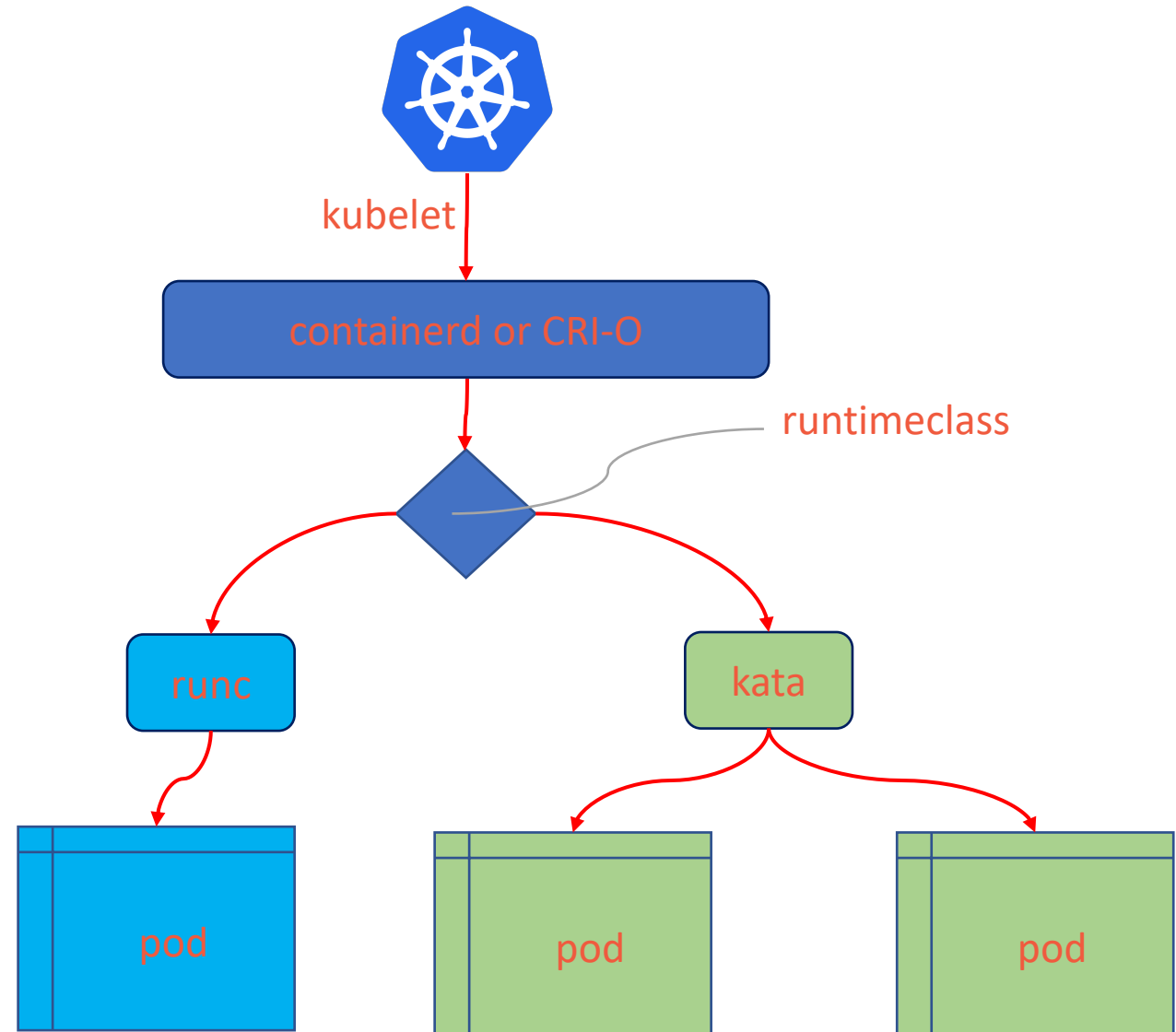


Virtual Machine as a Pod

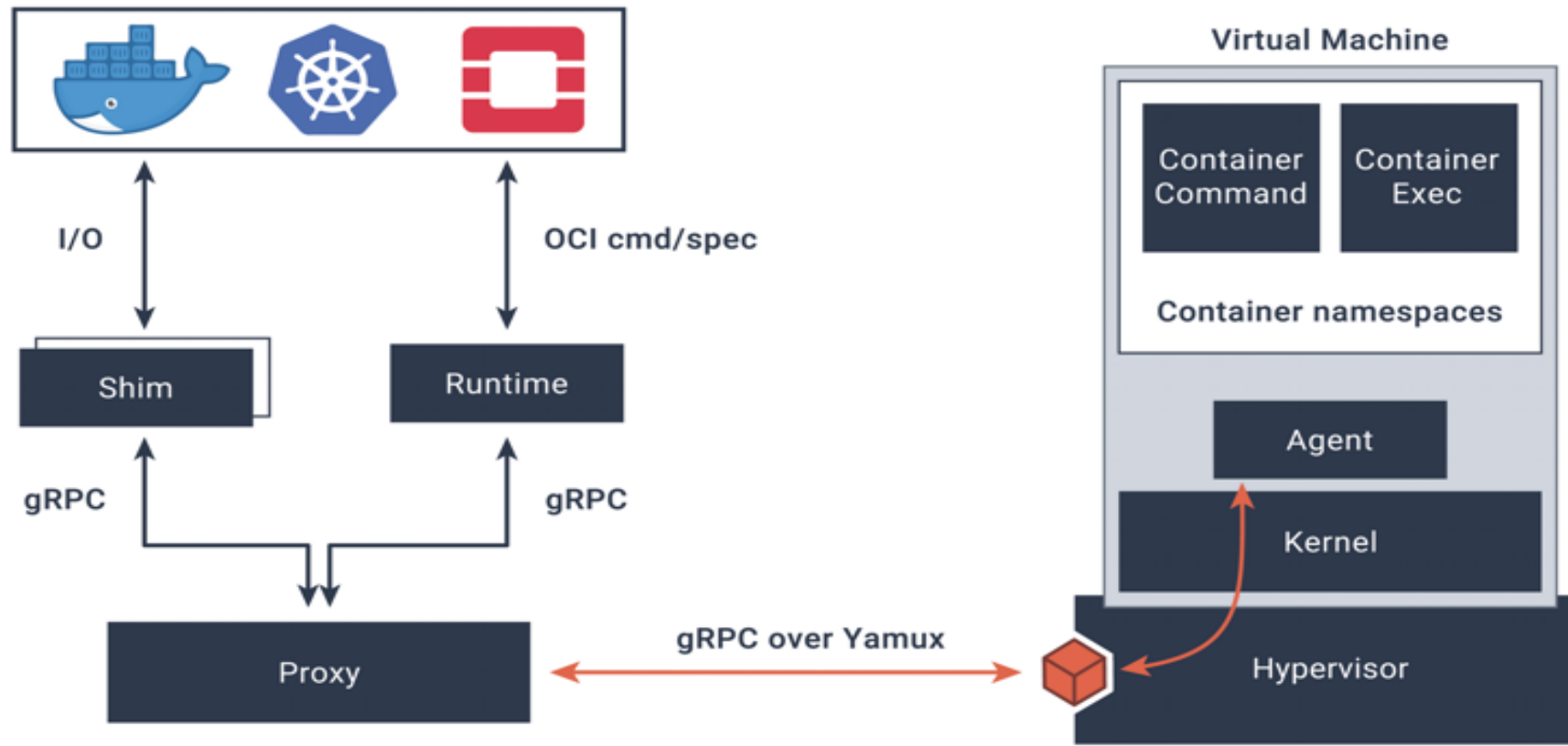


Kata & Kubernetes

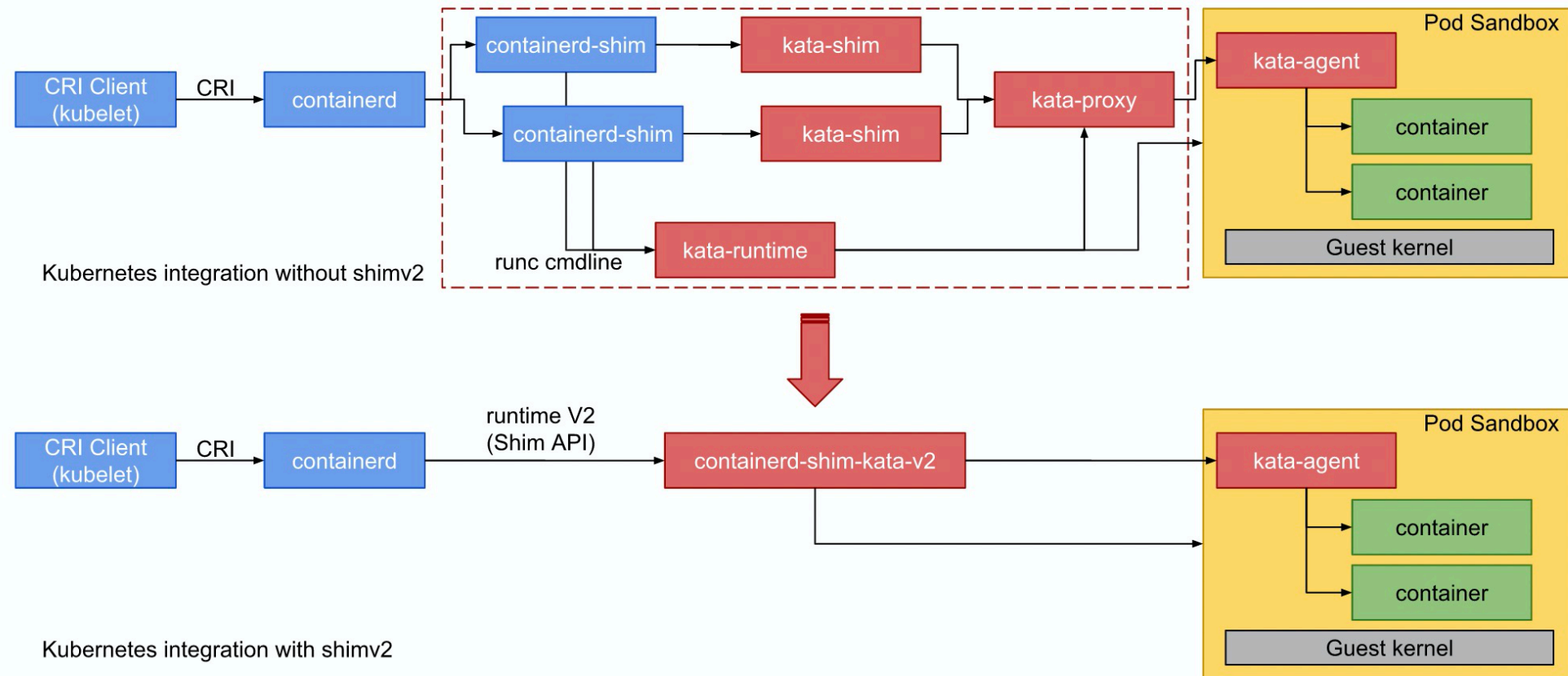
- RuntimeClass Resource
- Run runc and kata containers in the Same cluster/node
- Beta feature since v1.14
- runtimeClassName in pod spec



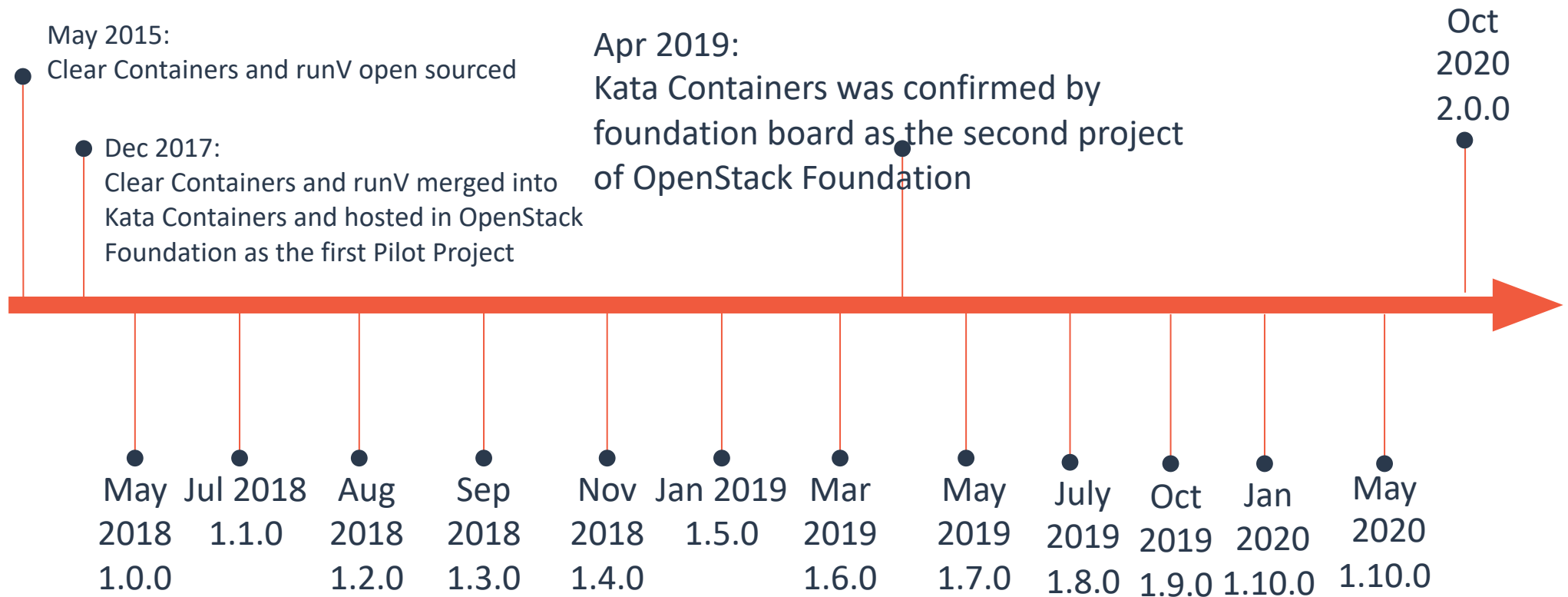
Kata Containers 1.x Architecture



Architecture Evolvment



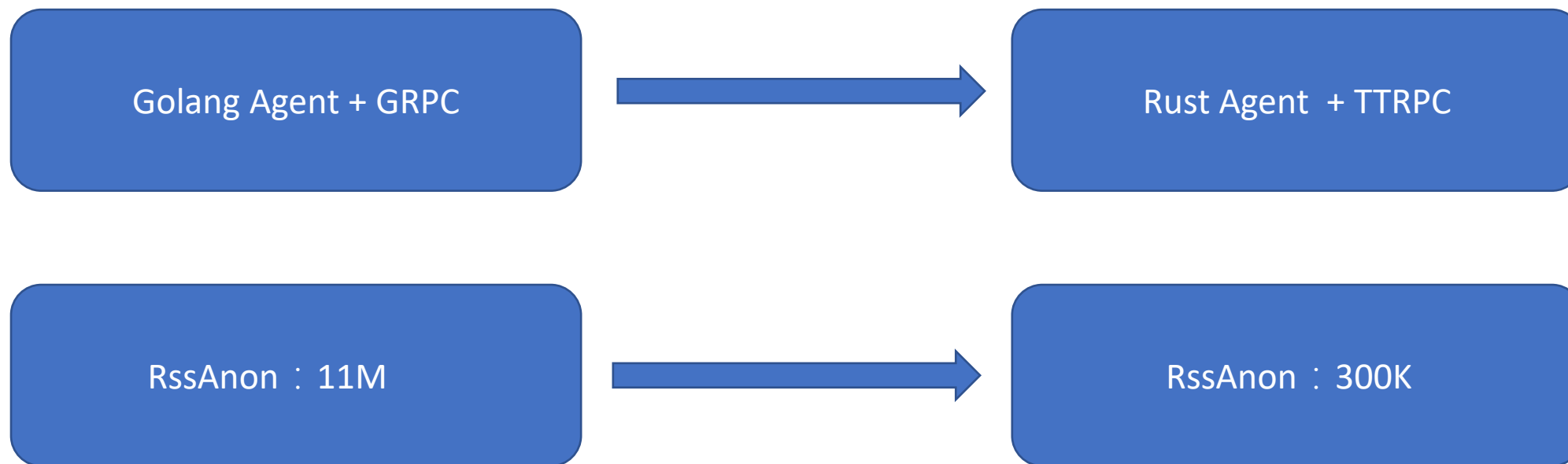
A bit history



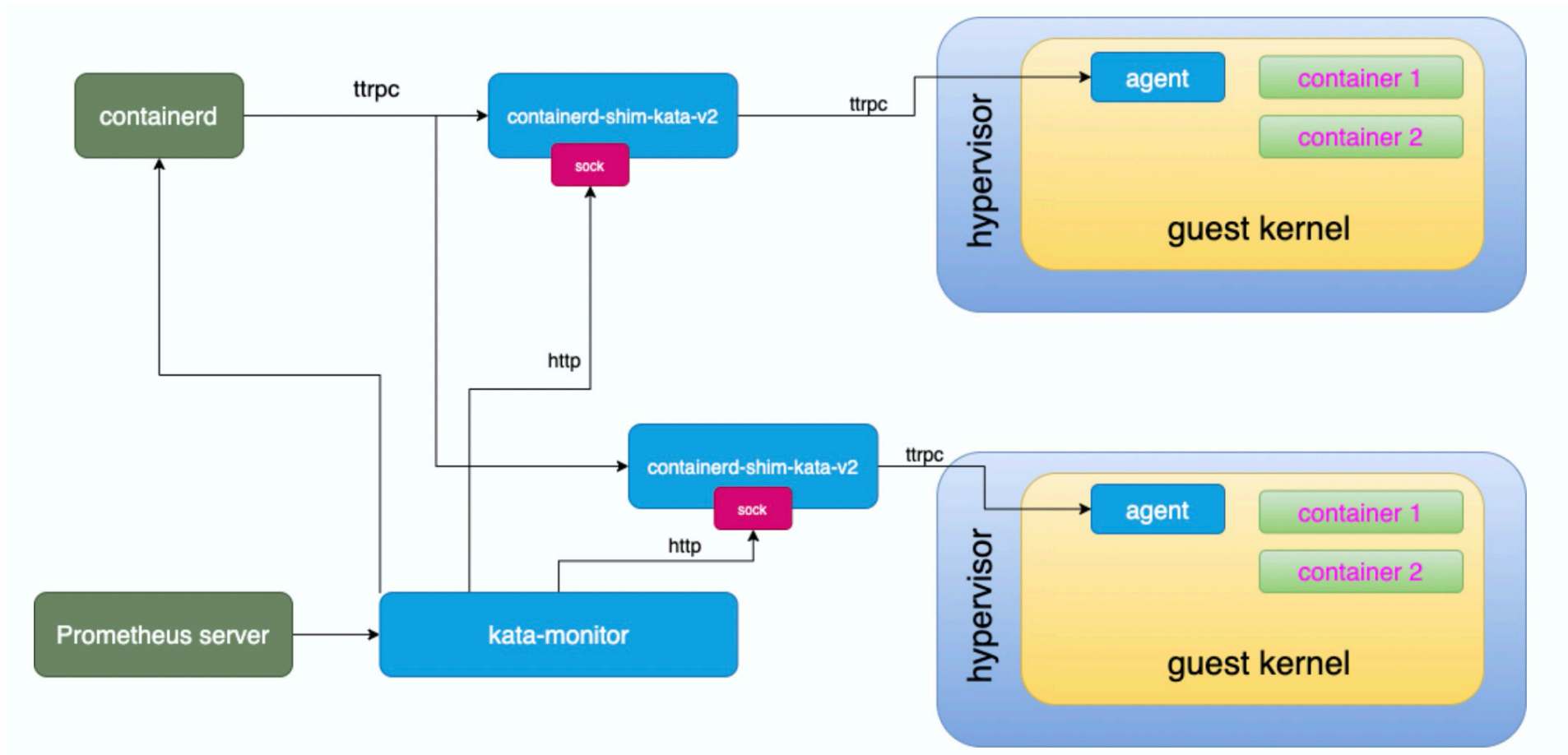
What's New In 2.0

- Kata-agent has been rewritten in rust to significantly reduce memory overhead overall attack surface.
- Agent protocol has been simplified to use ttRPC from grpc.
- Component called Kata-monitor has been introduced to improve observability and manageability
- New component called agent-ctl has been introduced to help validate agent API
- We have moved to a mono-repo model, all code and document repositories consolidated into one single repo.
- Virtio-fs is now the default shared file system type which mean better POSIX compliance compared to virtio-9p
- Latest Cloud Hypervisor support on par with QEMU
- Move to support solely shimv2 api to simplify code and reduce attack surface further. This also means kata-shim and kata-proxy used in 1.x are no longer required.
- Guest kernel updated to v5.4.71
- QEMU updated to v5.0.0

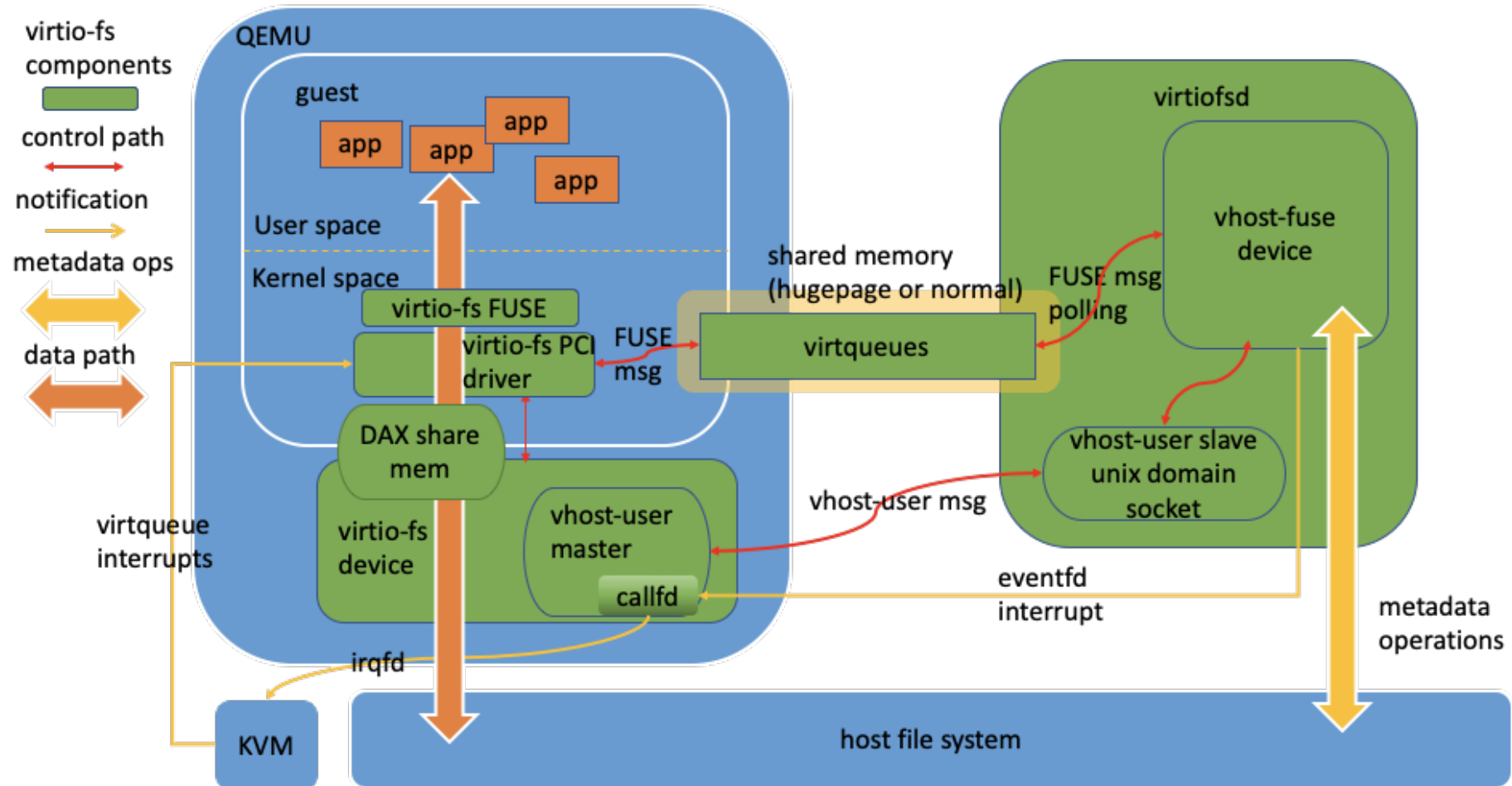
rust-agent + ttRPC



Kata Metrics



virtio-fs



Kata Containers & Ant

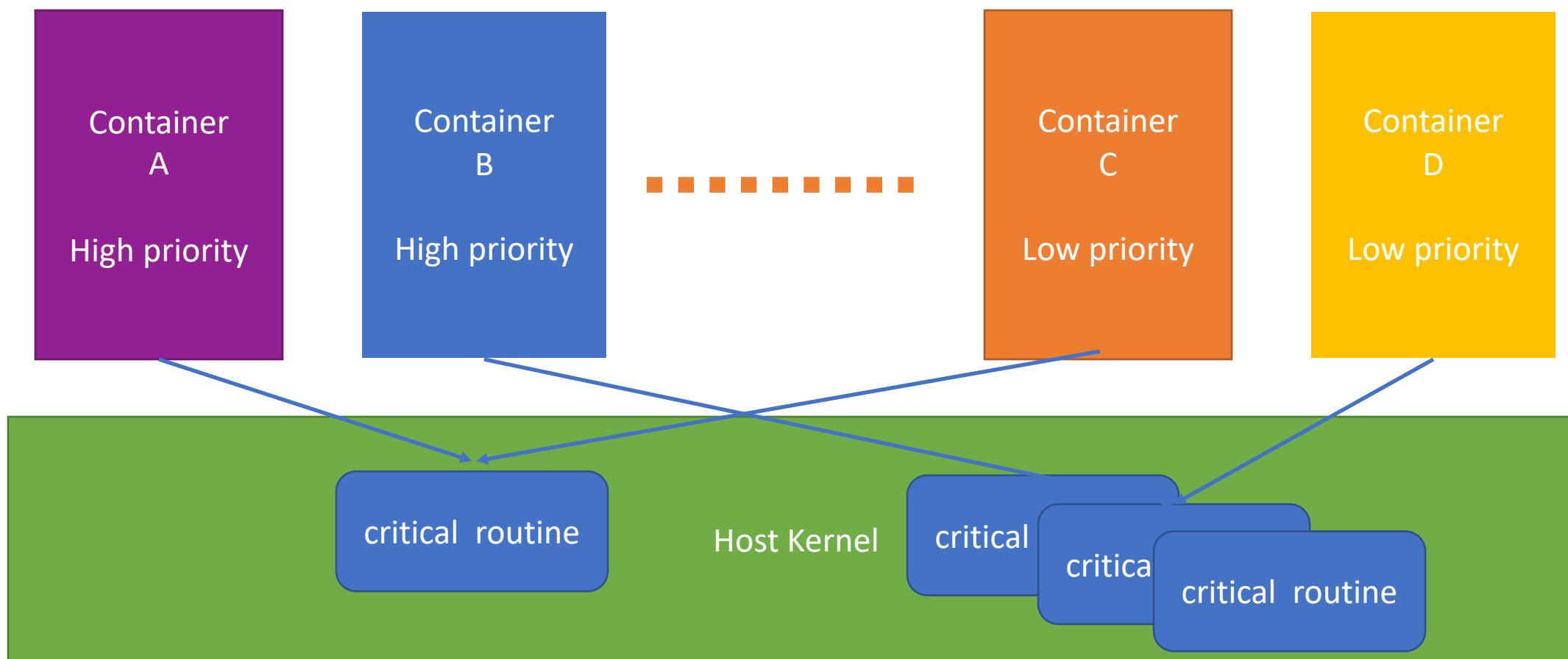


Is only suitable for
untrusted workloads?

A large, bold red 'X' is superimposed over the text, indicating a negative answer or a warning.

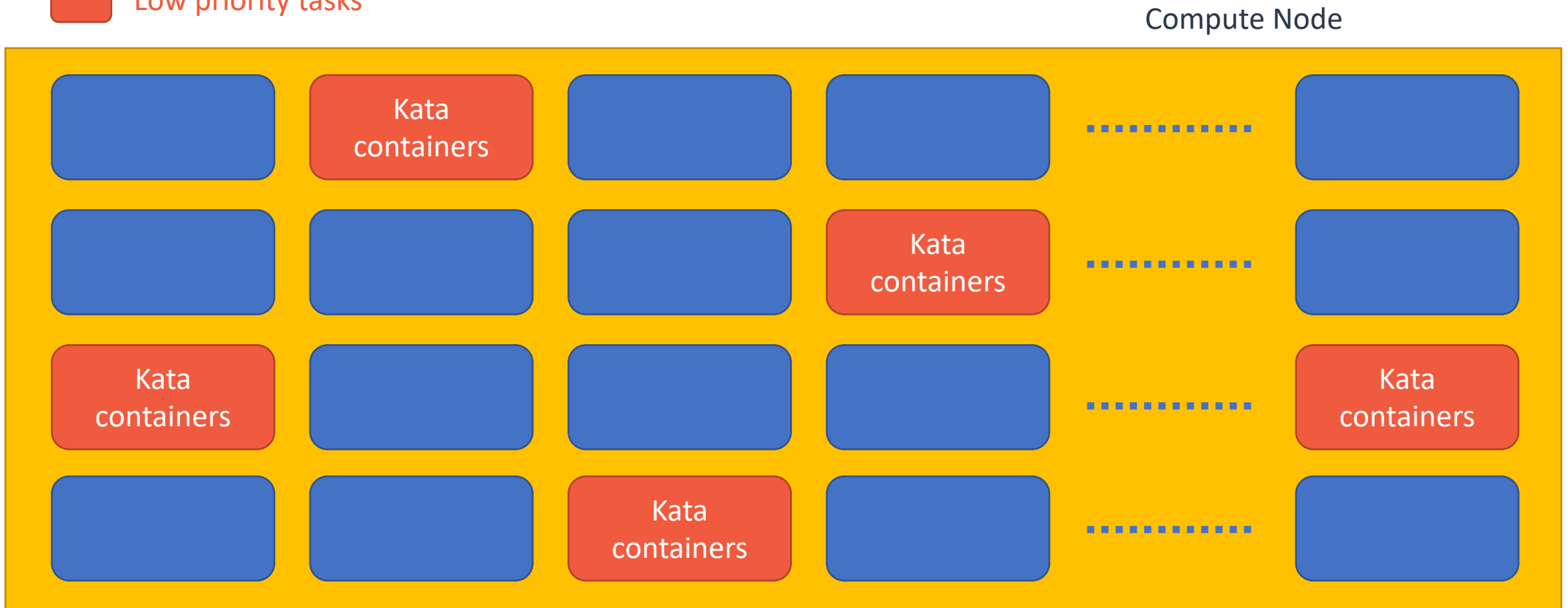
Conventional container workload isolation

- Container A's response time would be delayed

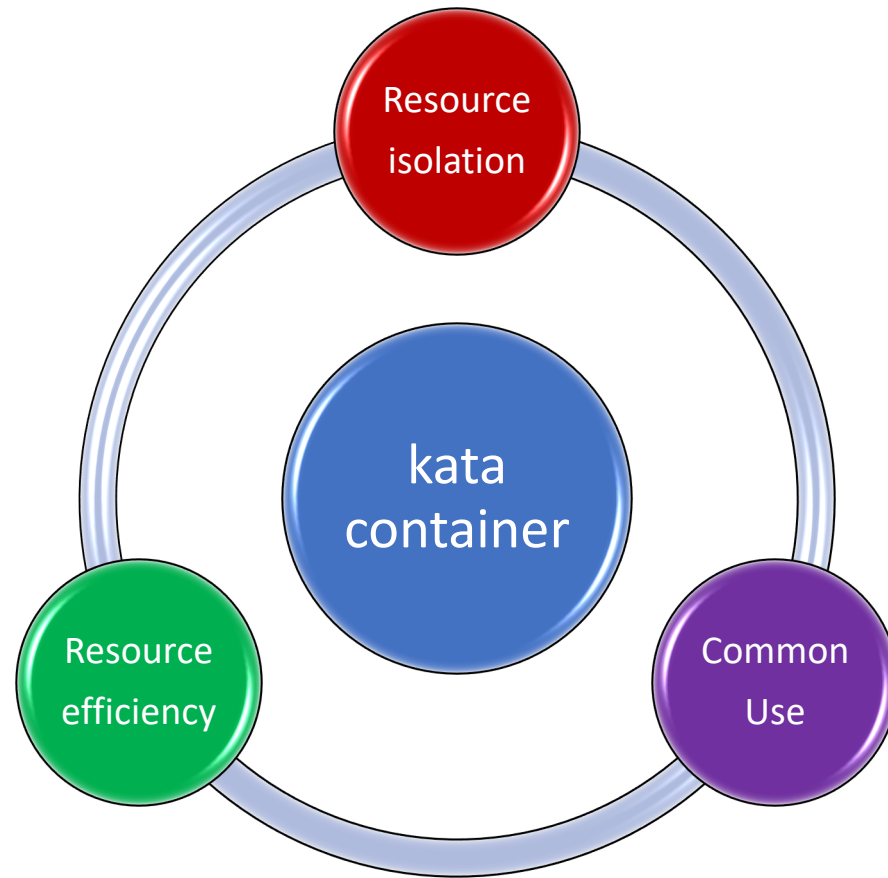


Colocation of high/low priority workloads

- High priority tasks
- Low priority tasks



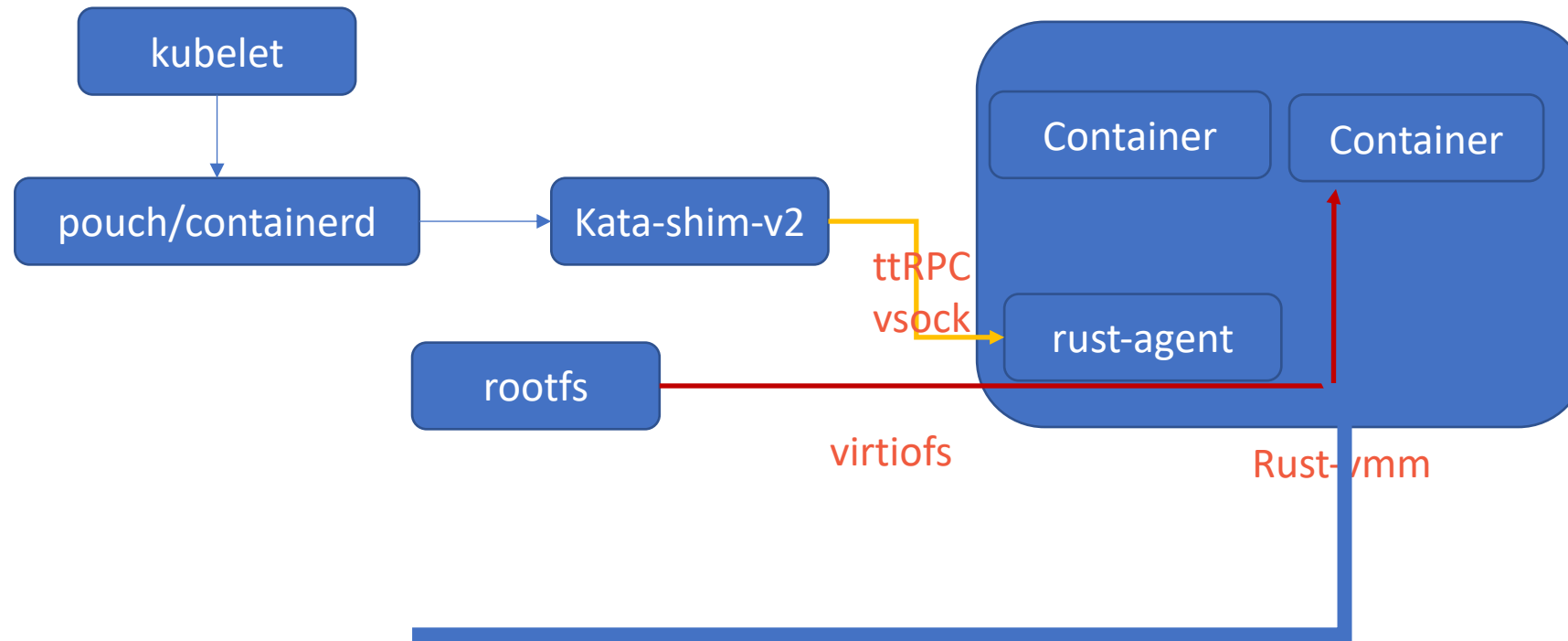
Art of Ballance



Kata Resource Improvements

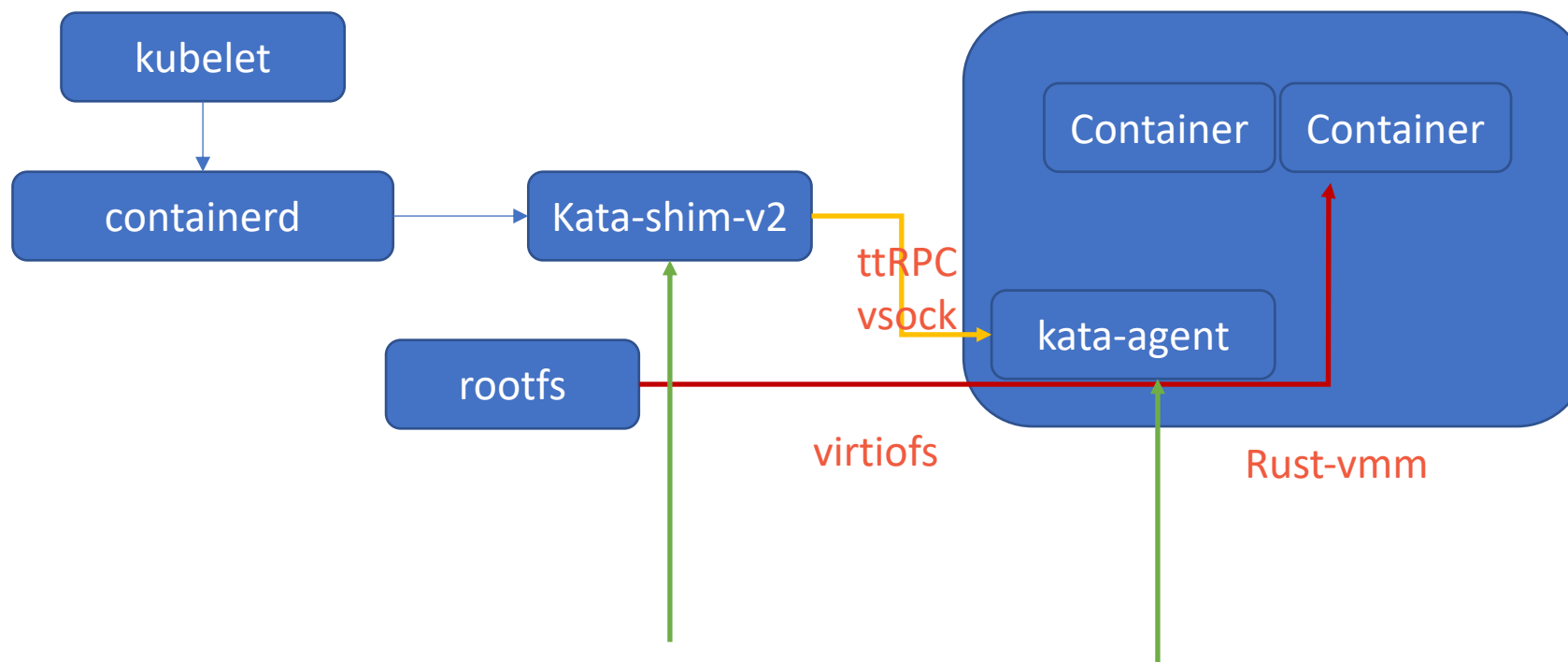
- Rust-VMM Hypervisor
- Rewrite kata-agent from Go to rust
- Replace GRPC with ttRPC
- Replace virtio-serial with vsock
- Replace 9pfs with virtiofs

Extra Ops Channel



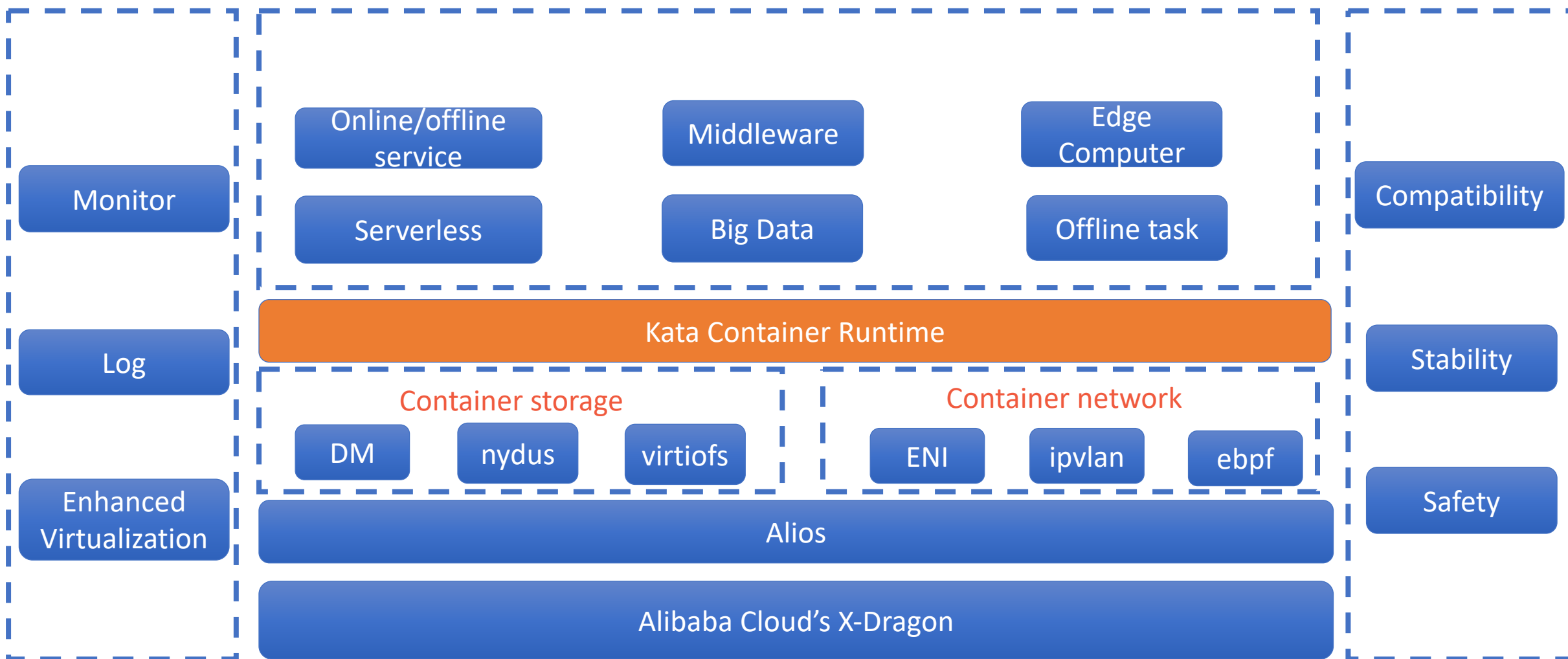
Add an Ops Channel

Hot Upgrade Kata Components



Kata-shim-v2 and kata-agent can be hot upgraded without stopping container tasks

Trust Native Infrastructure Conerstone



What's Next

- More isolation enhancements
- Container image operation improvements
- IO stream improvements
- Networking (port forwarding) improvements
- Debugability improvements

Links

- website: <http://katacontainers.io/>
- github: <https://github.com/kata-containers>
- IRC freenode: **#kata-dev**
- Slack (bridged to IRC): <http://bit.ly/katacontainersslack>
- Twitter: **@katacontainers**
- Mailing list: **lists.katacontainers.io**