Tailor-made Security

Building A K8s Specific Hypervisor

Andreea Florescu <fandree@amazon.com>

Samuel Ortiz <sameo@linux.intel.com>

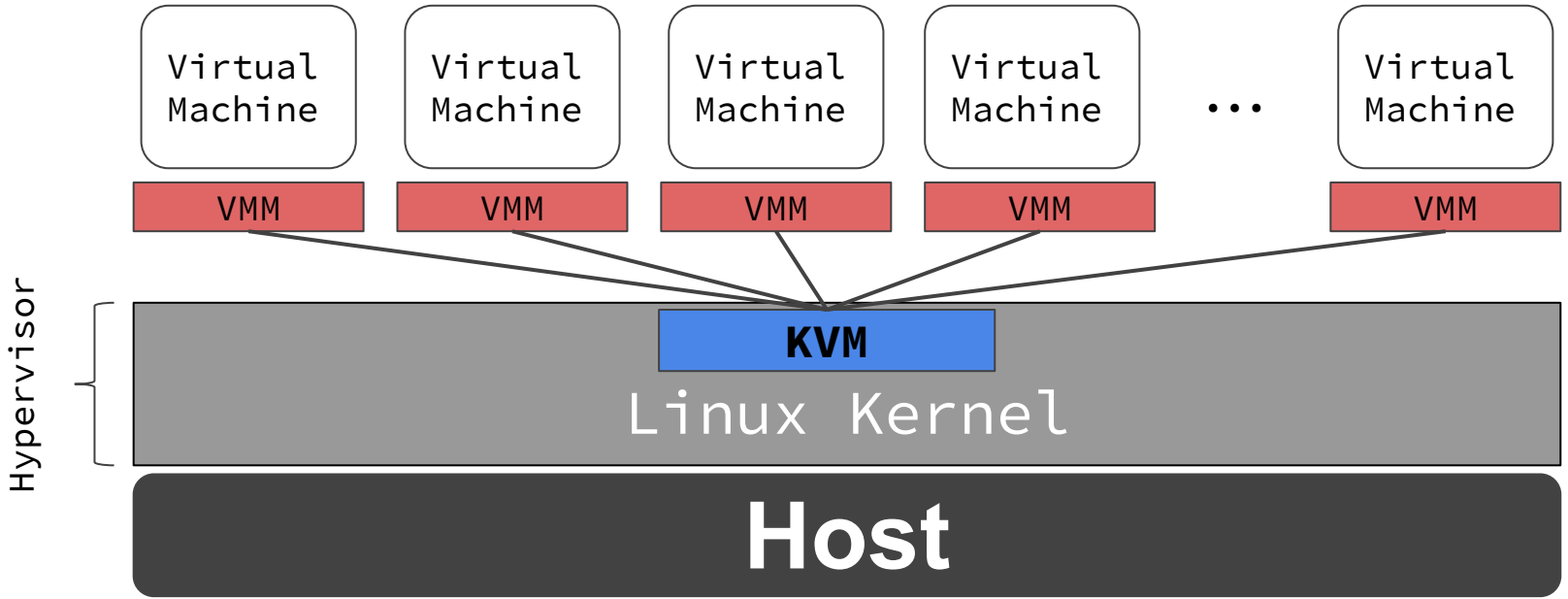
Tailor-made Security

Building A K8s Specific ~~Hypervisor~~ VMM

Andreea Florescu <fandree@amazon.com>

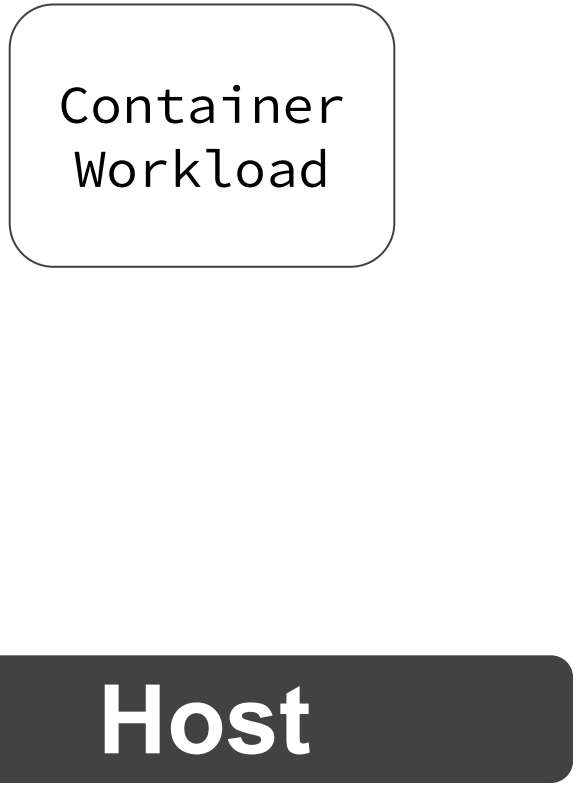
Samuel Ortiz <sameo@linux.intel.com>

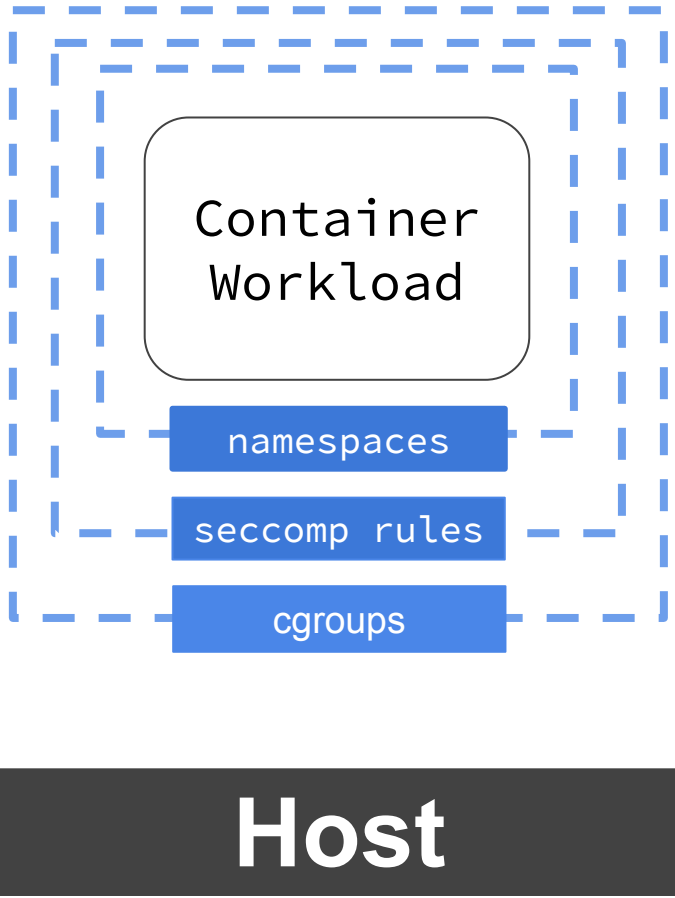
Linux Virtualization Stack

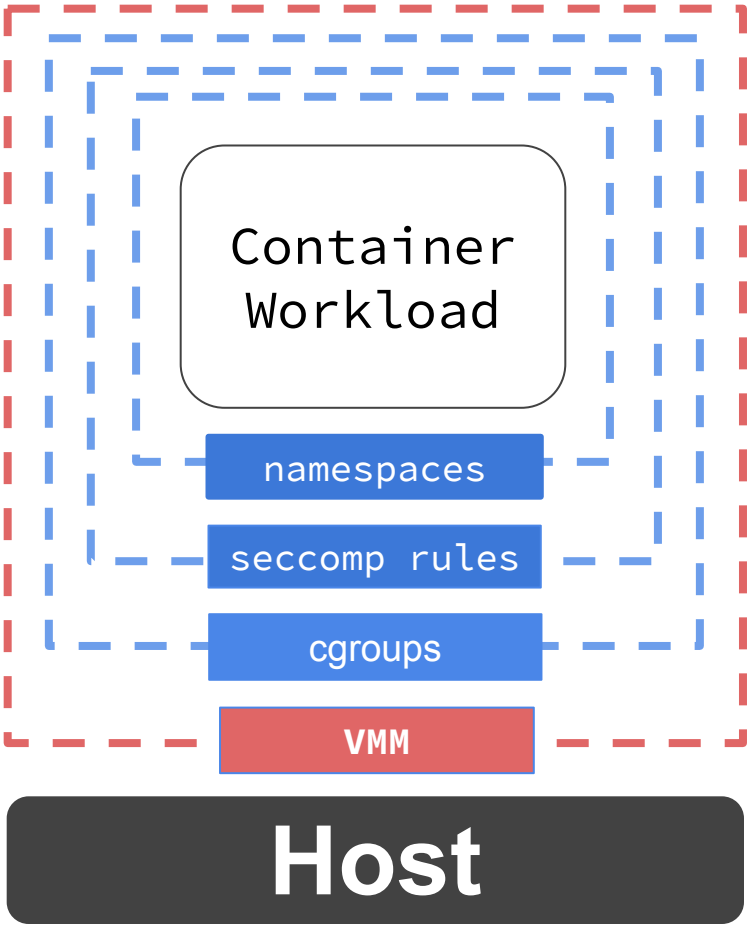


Container? VMM?


Belay Jacket Insulated Layer
Shell Layer
Fleece Layer
Performance Base Layer


Container
Workload
Host


Container
Workload
namespaces
seccomp rules
cgroups
Host


Container
Workload
namespaces
seccomp rules
cgroups
VMM
Host

Defense In Depth

Defense In Depth

Containers As Secure As VMs

Seamless Integration

Kubernetes RuntimeClass

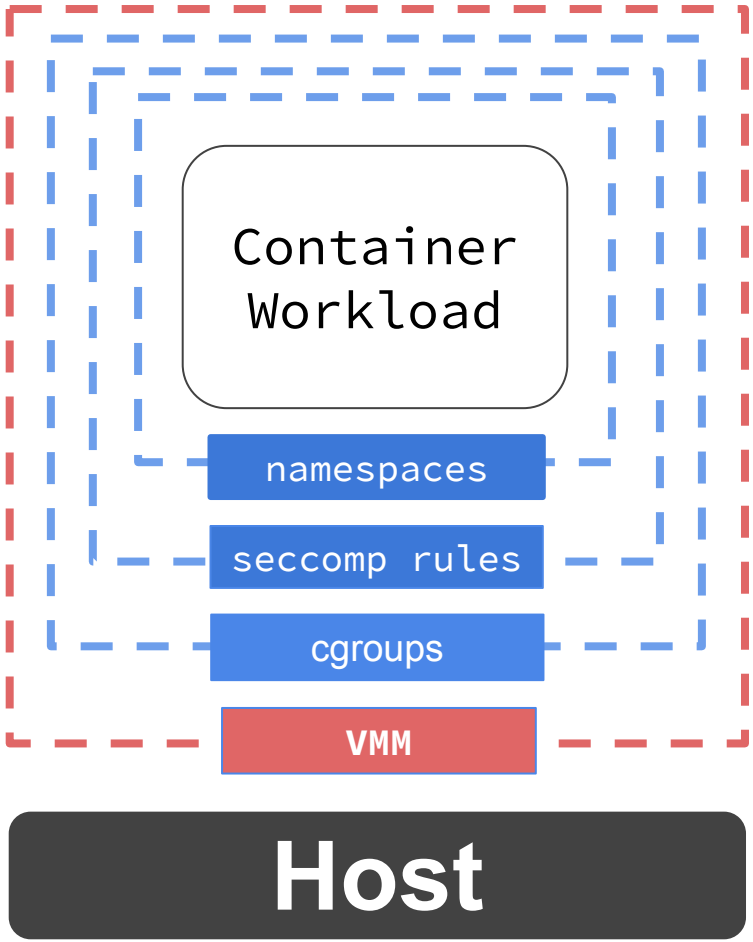
Container Runtime Interface (CRI)

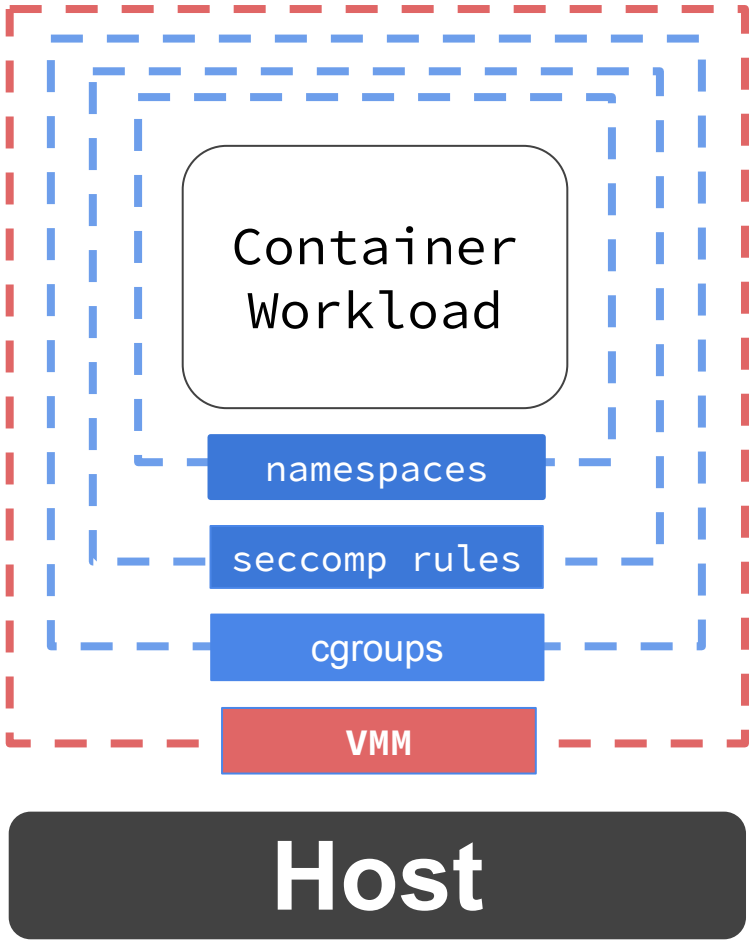
OCI Runtime

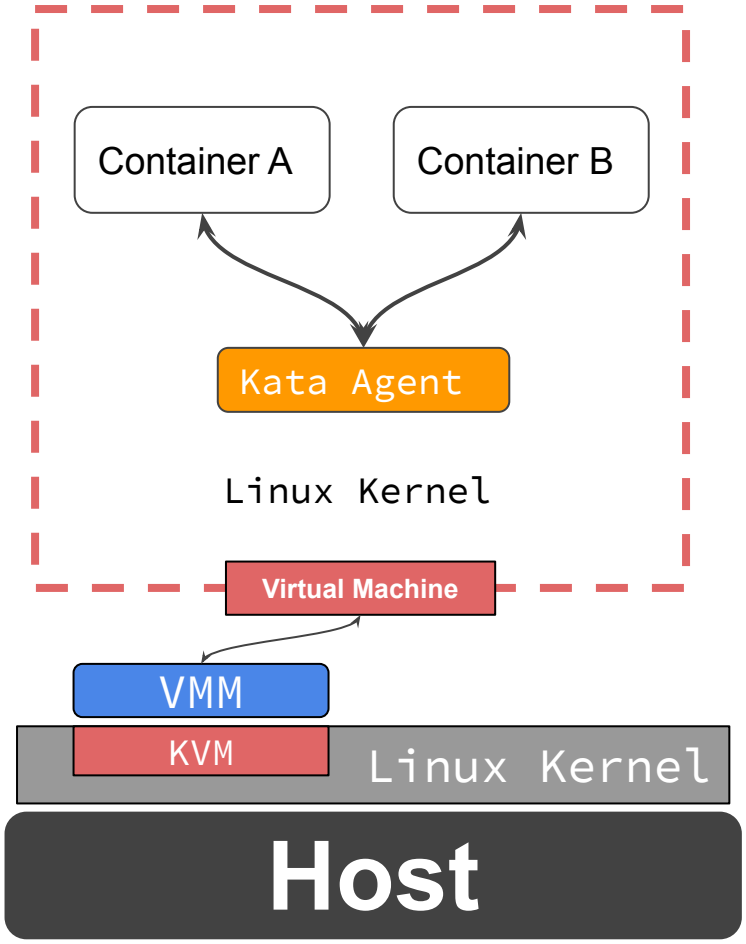

katacontainers

```
apiVersion: v1
kind: Pod
metadata:
  name: kata-pod
spec:
  containers:
    - name: busybox
      image: busybox:1.25
      command:
        - /bin/sh
  runtimeClassName: kata-qemu
```

```
apiVersion: v1
kind: Pod
metadata:
  name: kata-pod
spec:
  containers:
    - name: busybox
      image: busybox:1.25
      command:
        - /bin/sh
  runtimeClassName: kata-fc
```


Container
Workload
namespaces
seccomp rules
cgroups
VMM
Host


Container
Workload
namespaces
seccomp rules
cgroups
VMM
Host


Container A
Container B
Kata Agent
Linux Kernel
Virtual Machine
VMM
KVM
Linux Kernel
Host

Kata Containers VMMs

QEMU

NEMU

Firecracker

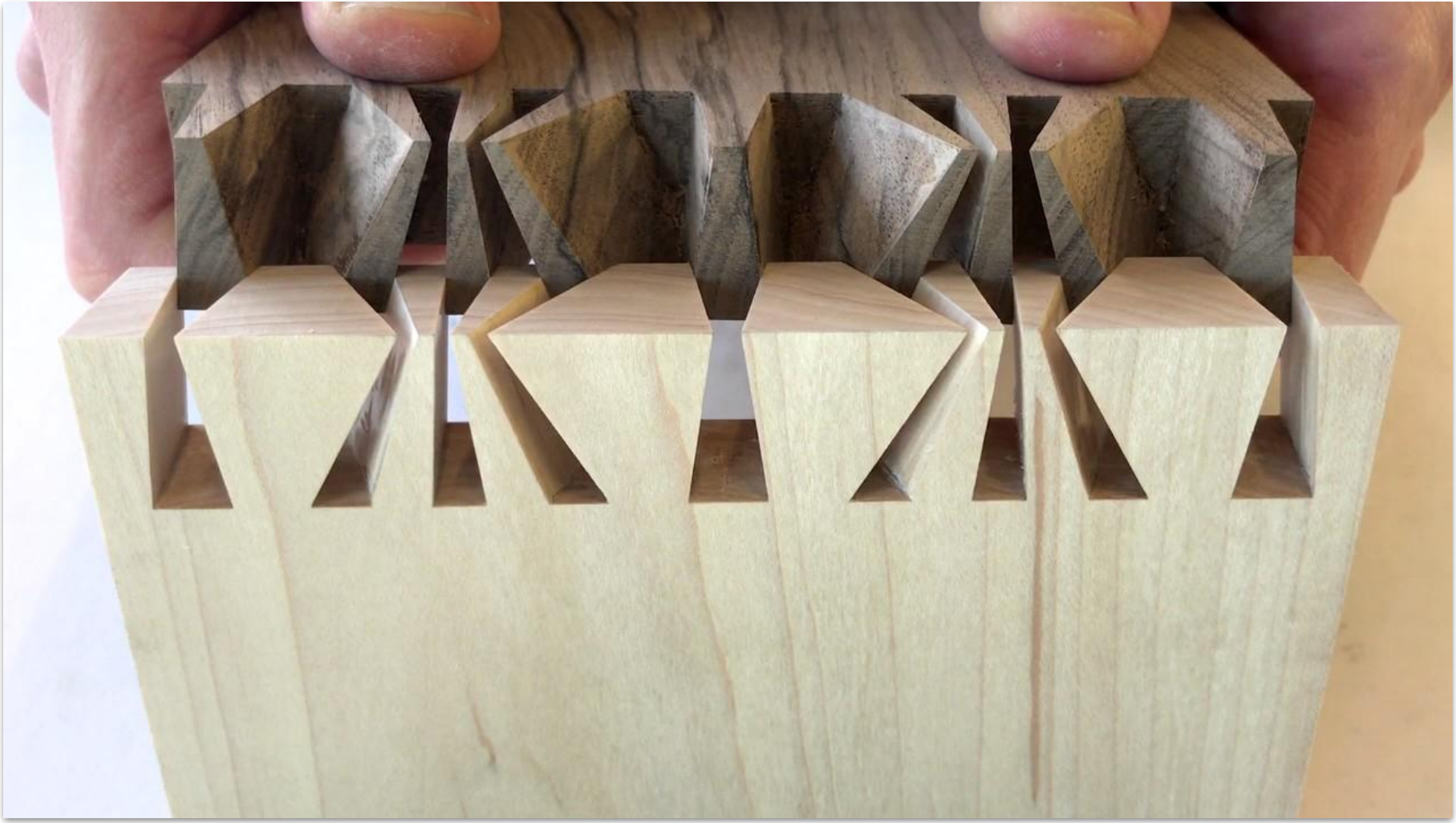
QEMU

- High Performance
- Stable
- Extremely Versatile
- Large Code Base and Footprint

Firecracker

- Laser Focused for Serverless
- Small Code Base
- Limited Scope

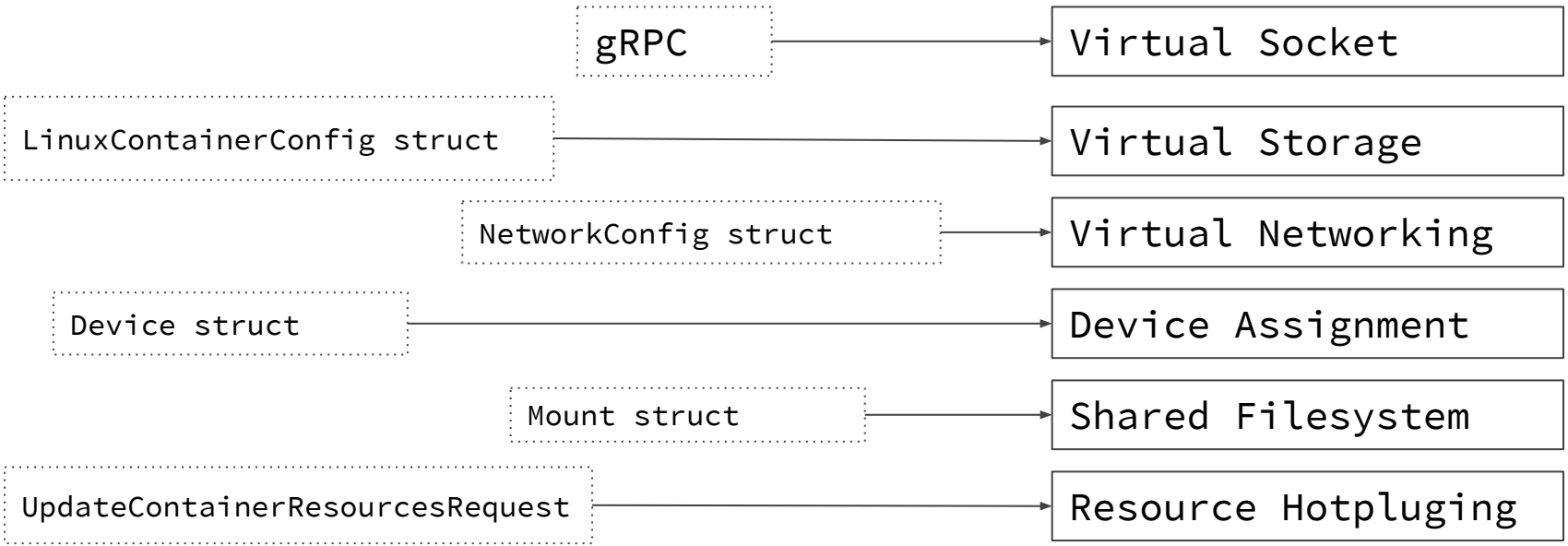
"HONESTLY? I PREFERRED WHEN WE DIDN'T TALK ABOUT THE ELEPHANT"



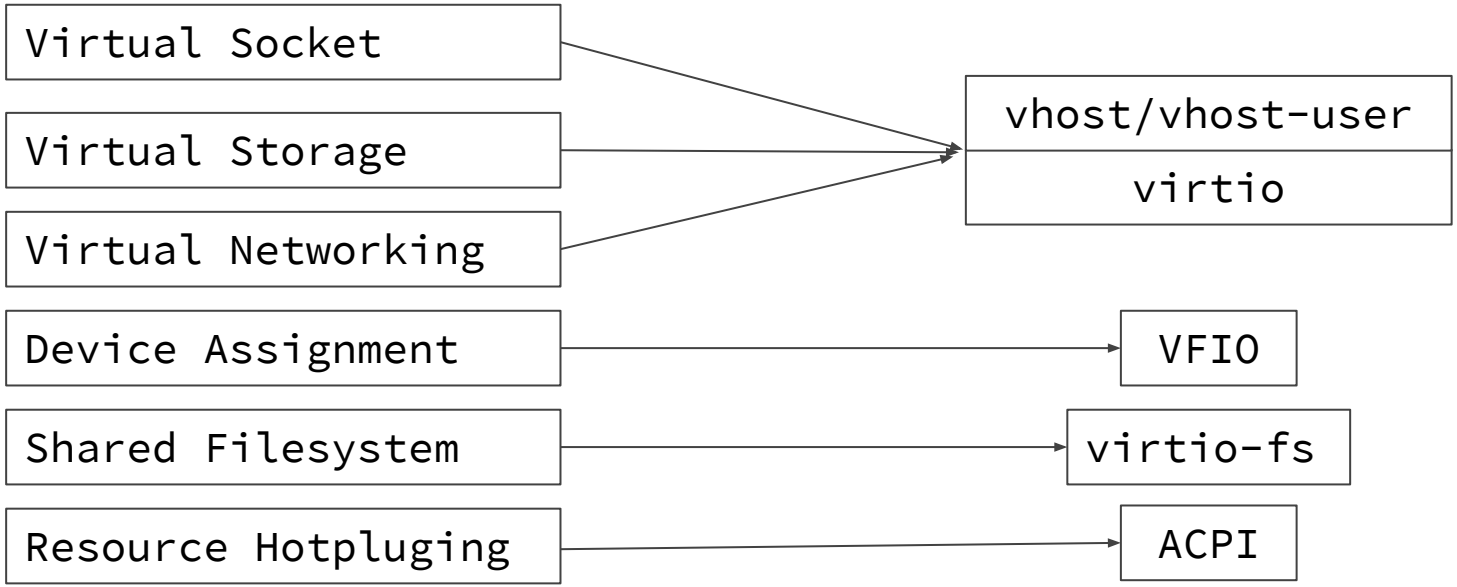
What defines a container?

- Runtime Specifications
- Open Container Initiative Runtime (OCI)
- Kubernetes Container Runtime Interface (CRI)

From CRI to Container VMM



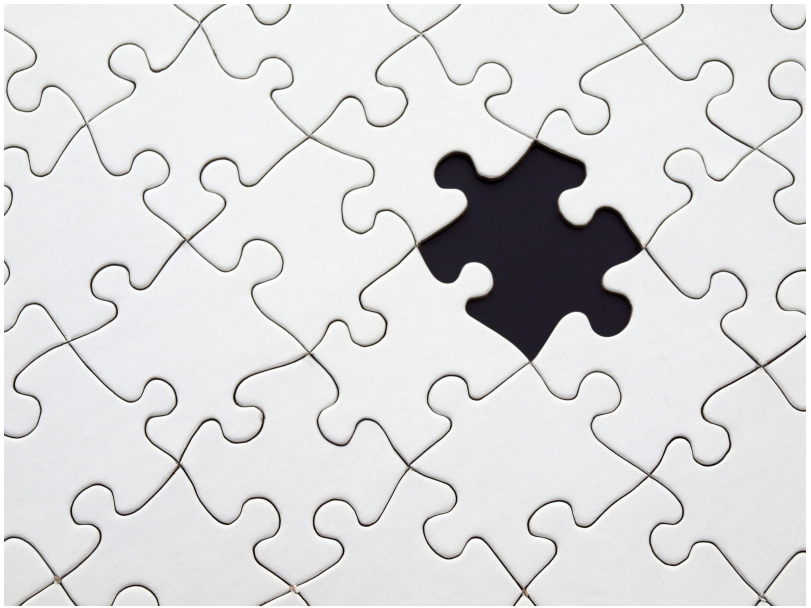
From Requirements to Technologies



Where can we get these components from?

rust-vmm

What is rust-vmm?



- Building blocks for VMMs written in Rust
- Virtualization components (crates)
- Open Source

Why rust-vmm?

- Faster development for new custom VMMs
- Security & Testability
- Clean interface
- Reduce code duplication (CrosVM & Firecracker)

rust-vmm development

Who is contributing?

Alibaba Cloud

AWS

Cloudbase Solutions

CrowdStrike

Google

Intel

RedHat

Individual Contributors

Adding crates to rust-vmm

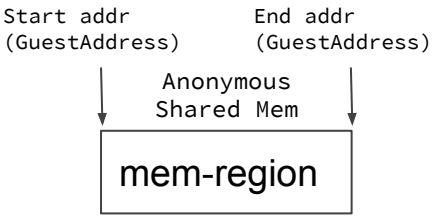
- CrosVM/Firecracker
  - Wrappers over the KVM API
  - Minimal kernel loader
  - ...
- Developing from scratch
  - vhost-user
  - ACPI
  - ...

vm-memory - Firecracker

- Guest Address

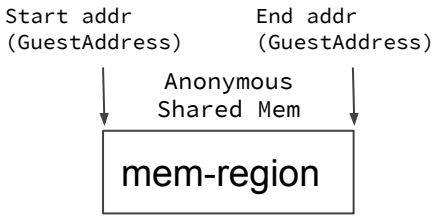
vm-memory - Firecracker

- Guest Address
- Memory Region



vm-memory - Firecracker

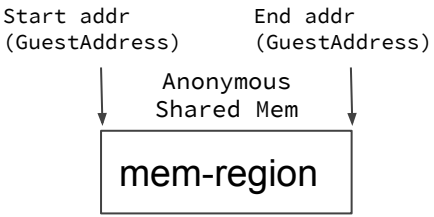
- Guest Address
- Memory Region
- Guest Memory



Guest Memory

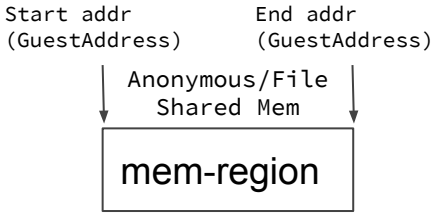
vm-memory - Firecracker

- Guest Address
- Memory Region
- Guest Memory



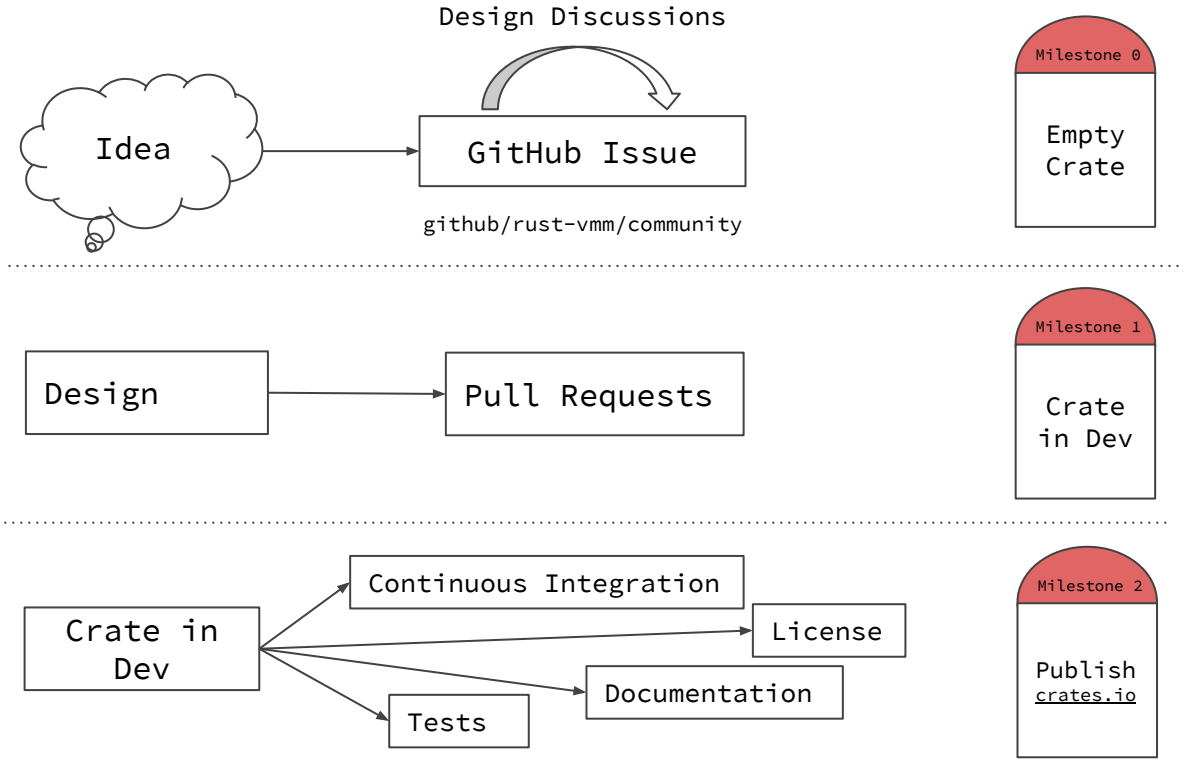
vm-memory - rust-vmm

- **Trait** Guest Address
- **Trait** Memory Region
- **Trait** Guest Memory

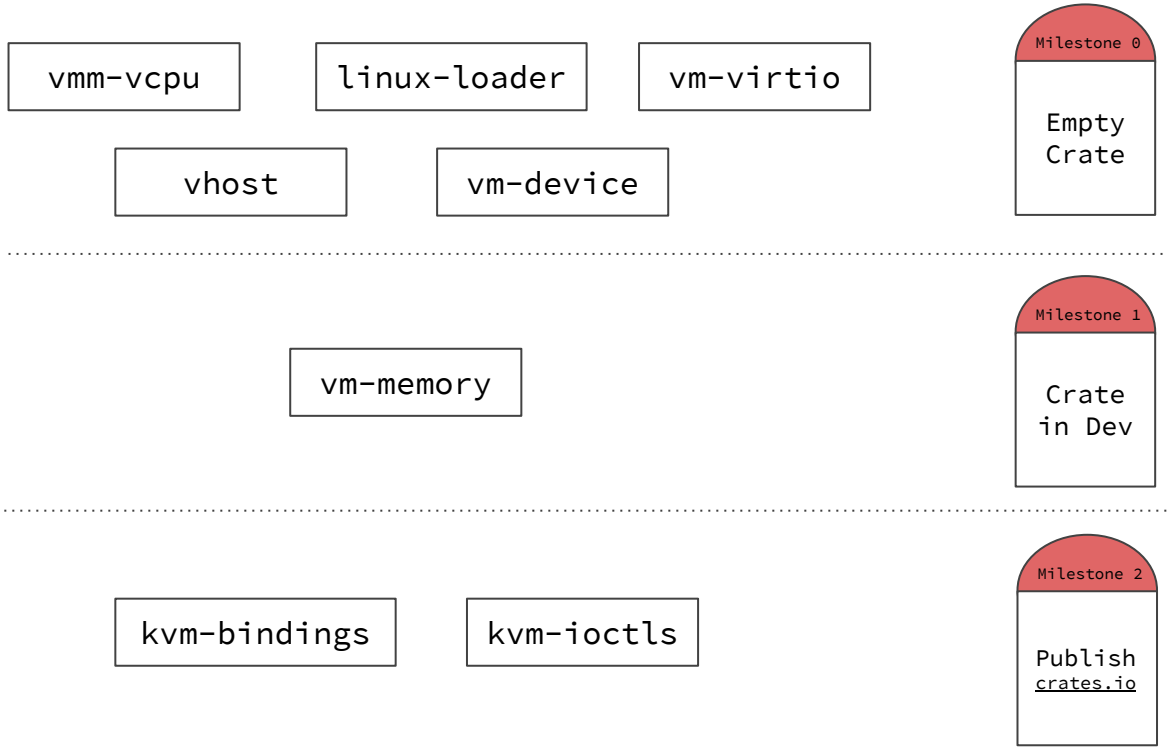


Guest Memory

From idea to published crate



Current Status



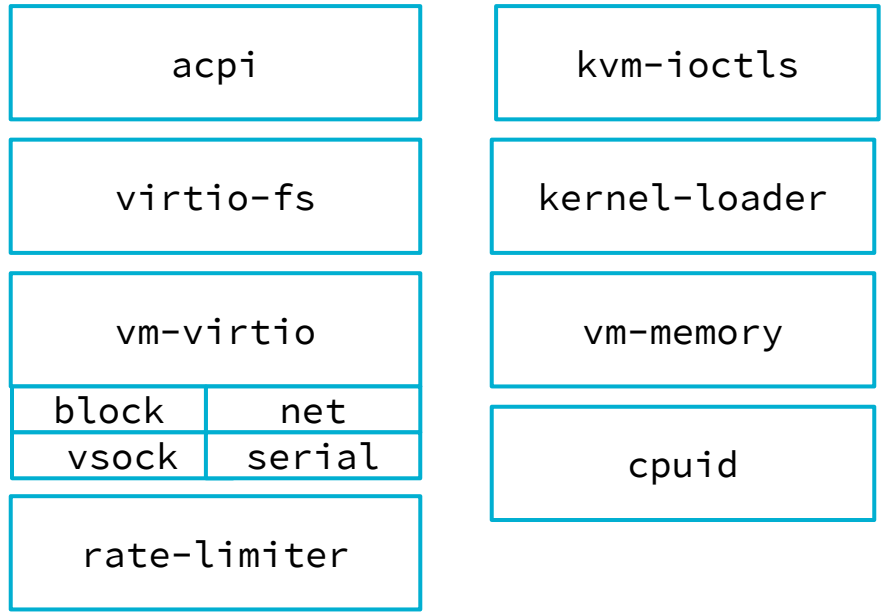
rust-vmm in practice

containers-vmm

containers-vmm

VMM API

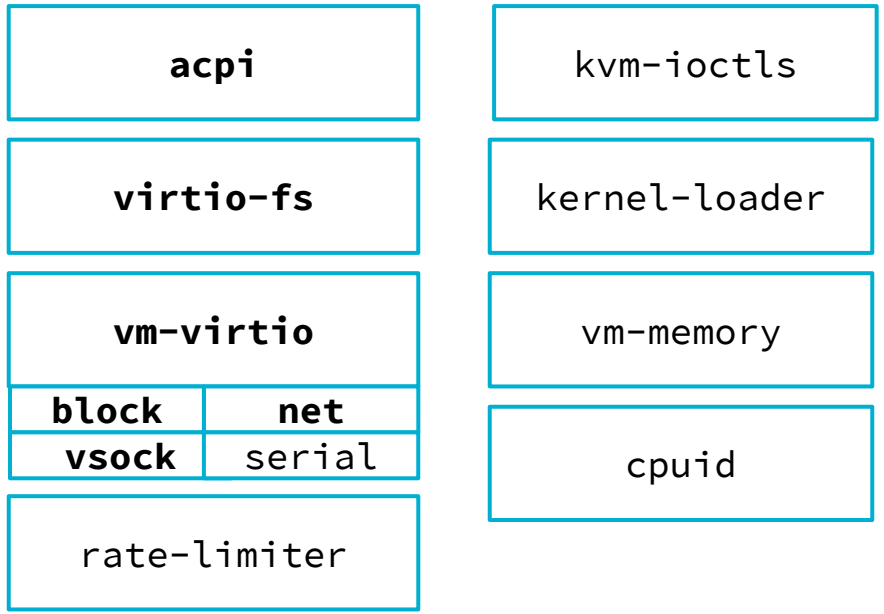
rust-vmm components



containers-vmm

VMM API

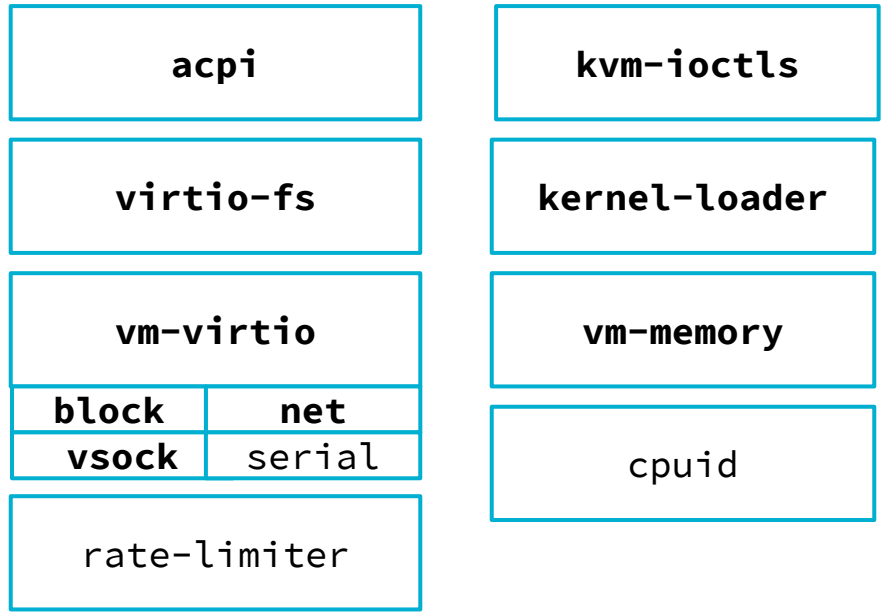
rust-vmm components



containers-vmm

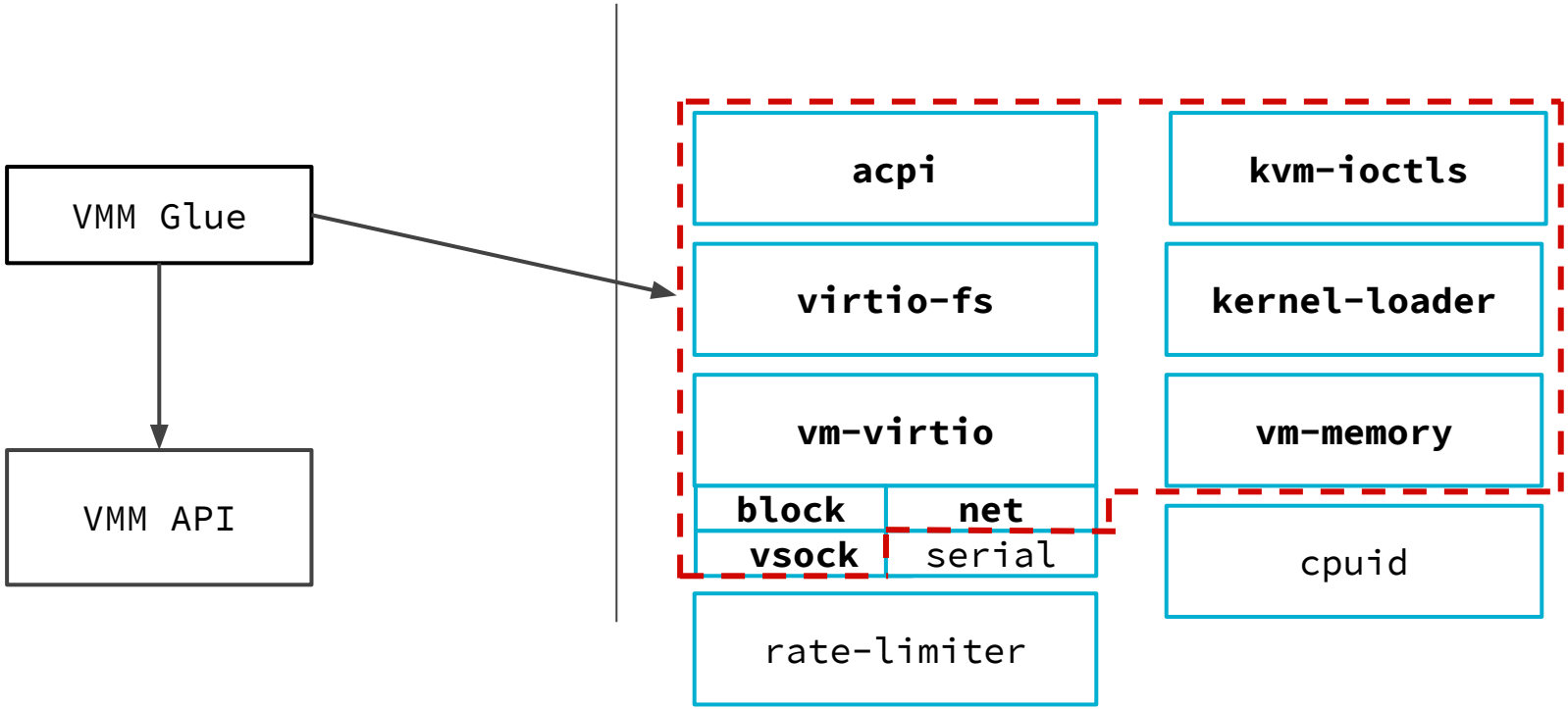
VMM API

rust-vmm components



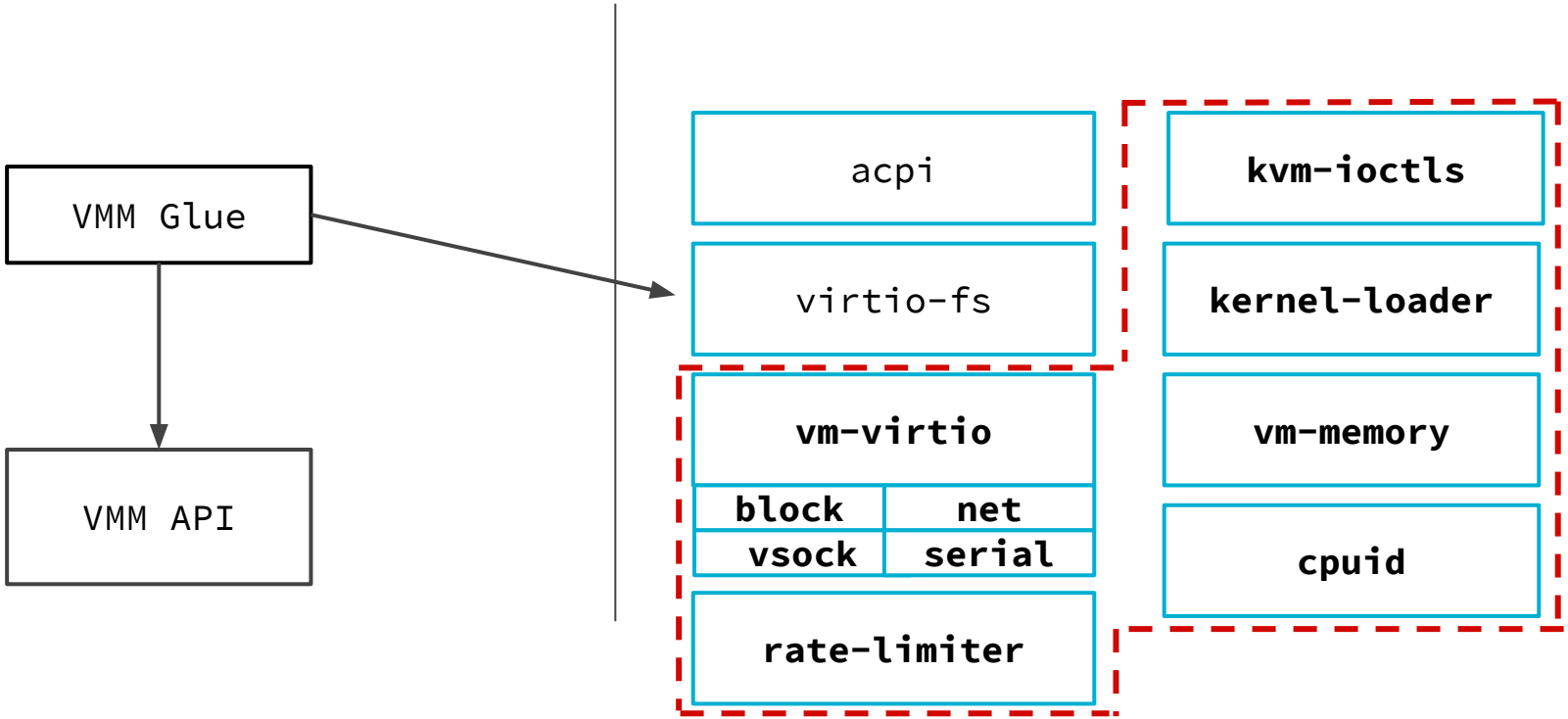
containers-vmm

rust-vmm components



serverless-vmm

rust-vmm components



What's next?

- Hypervisor-agnostic crates
- Using rust-vmm crates in existing VMMs
- Create new specialized VMMs with rust-vmm

What's next?

- Hypervisor-agnostic crates
- Using rust-vmm crates in existing VMMs
- Create new specialized VMMs with rust-vmm

Come decide with us!

Be part of rust-vmm!

- Subscribe to the rust-vmm email list
- Want feedback on your work? Submit a review request!

Conclusion

- VMs: A standard container isolation layer
- Need for a container specific VMM
- rust-vmm to provide the building blocks

https://github.com/rust-vmm

https://rust-vmm.slack.com