



Exploring Kata-Containers

Enhancing Cloud Security
and Performance

Anastassios Nanos

Systems Researcher @ Nubificus LTD



OpenInfraDays Hungary, Budapest, Jun 3rd, 2024

- Introduction - Kata-containers Overview
- Installation walkthrough
- Go vs Rust runtime
- Example use-cases
 - Expose Hardware acceleration through vAccel
 - Serverless Sandboxes



- Spent some time in academia
- Spent some time consulting & worked for a deep-tech startup

Primary focus:

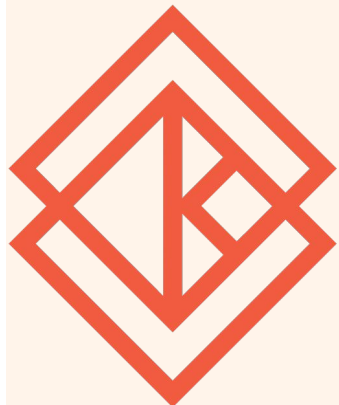
- low-level systems software
 - hypervisors
 - hardware-acceleration
 - minimizing OS overhead
- Started using kata-containers as a means to sandbox workloads using AWS Firecracker
 - Continued trying to maintain AWS Firecracker for Go runtime
 - As of April 2024, joined the AC of kata-containers to assist in the aarch64 CI, AWS Firecracker support for the Rust runtime and (hopefully) many more interesting things!



Systems Researcher



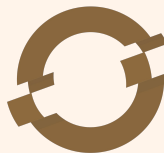
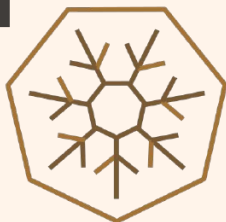
nubificus



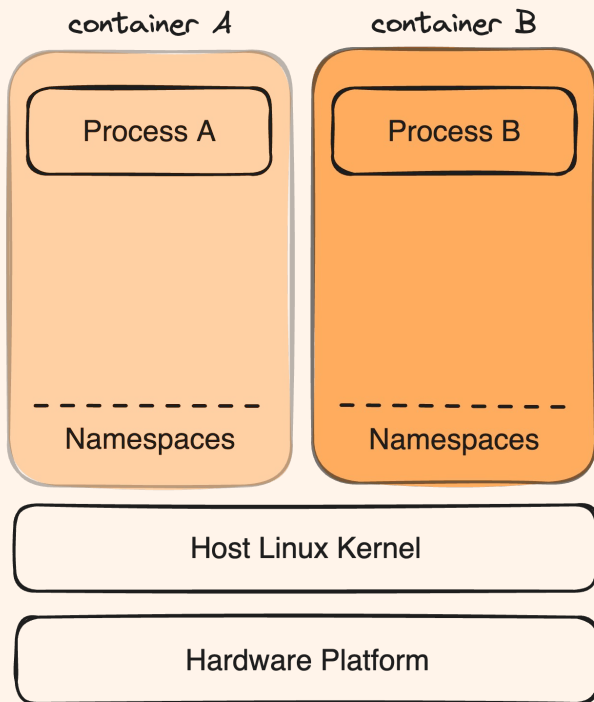
Kata Containers is an open-source project designed to provide the benefits of both containers and virtual machines (VMs) for workload isolation and security.

Key points:

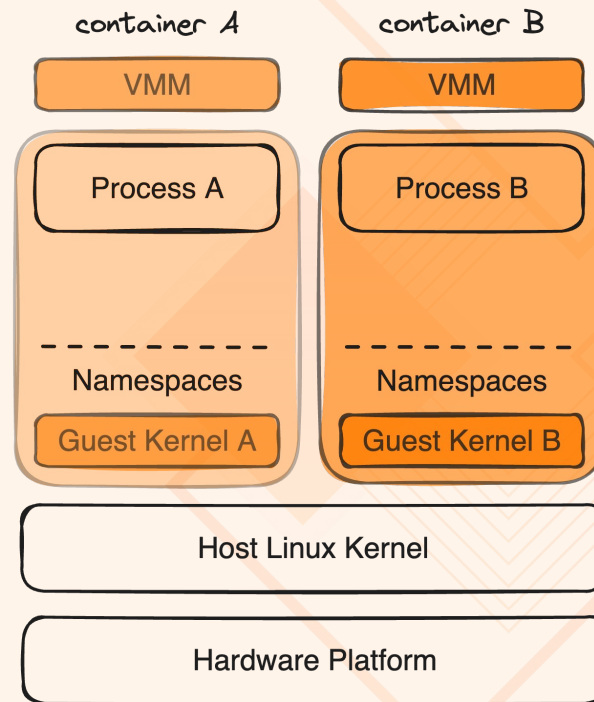
- Combines the speed of containers with the security of VMs.
- Ideal for multi-tenant environments, edge computing, and highly regulated industries.
- Compatible with various high-level runtimes & orchestrators



Traditional Containers



Kata Containers



- **Enhanced Security:** Isolation provided by VM-level separation improves security posture.
- **Performance:** Lightweight architecture minimizes overhead compared to traditional VMs.
- **Use Cases:**
 - **Secure Multi-Tenancy:** Ideal for cloud providers hosting multiple customers' workloads.
 - **Edge Computing:** Ensures security and isolation in edge environments.
 - **Compliance:** Meets stringent security and compliance requirements in industries like finance and healthcare.



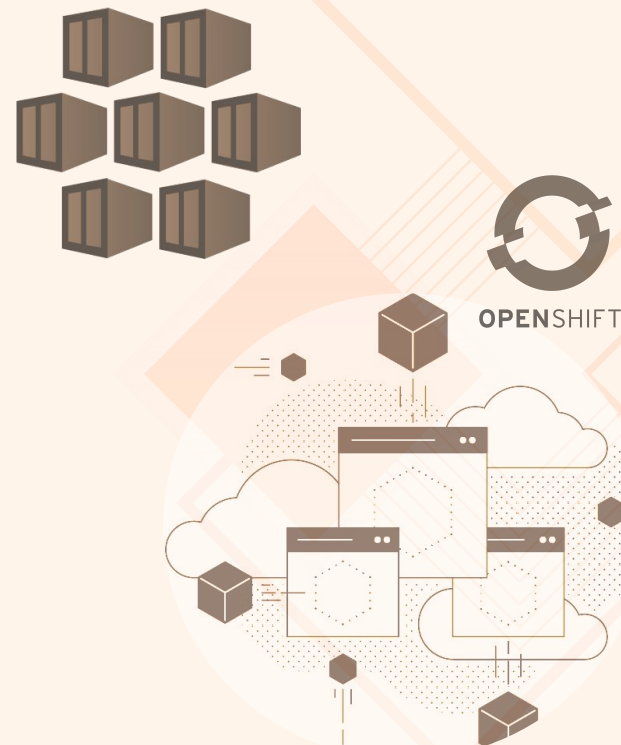
- RedHat OpenShift sandboxed containers

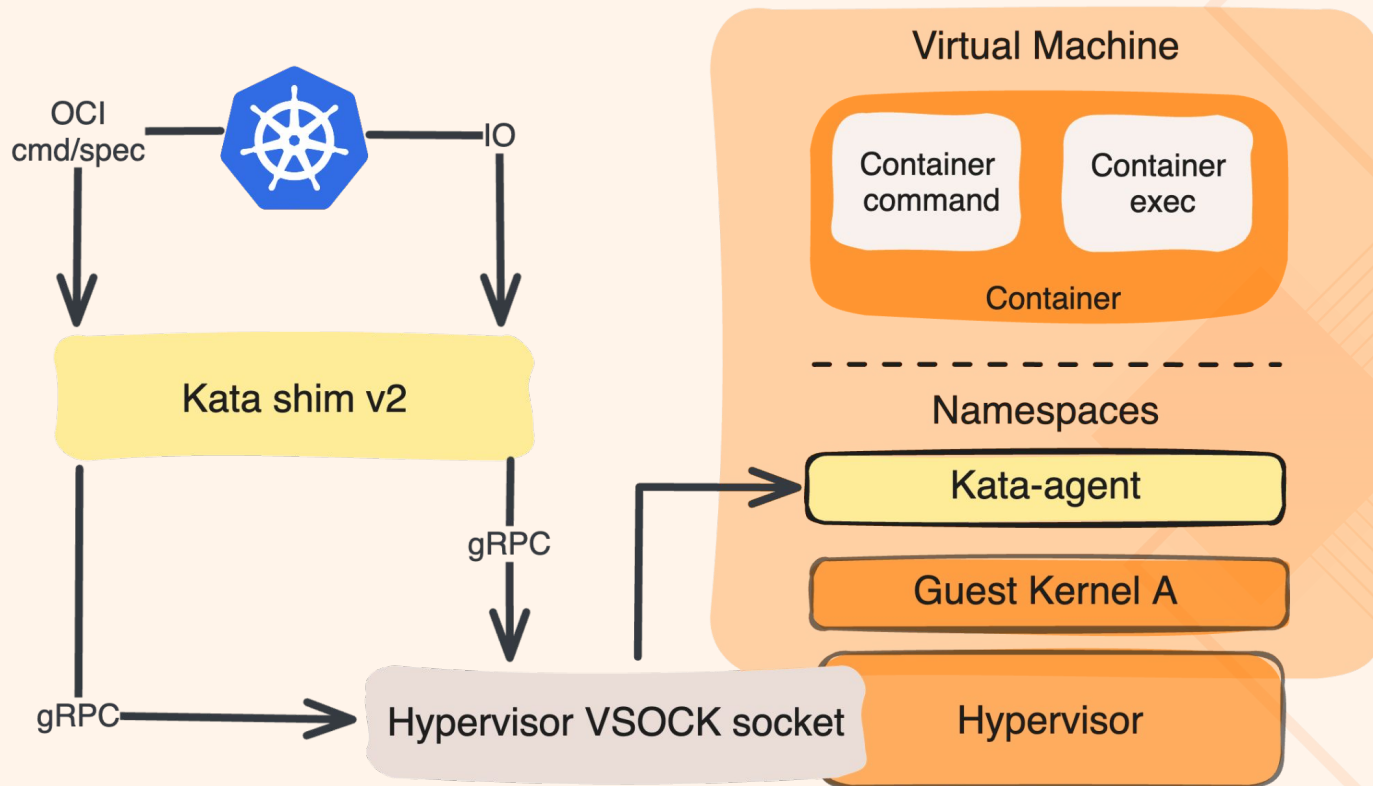
- Microsoft Azure Pod sandboxing on AKS

- Alibaba cloud



- Huawei cloud







kata-deploy



static release



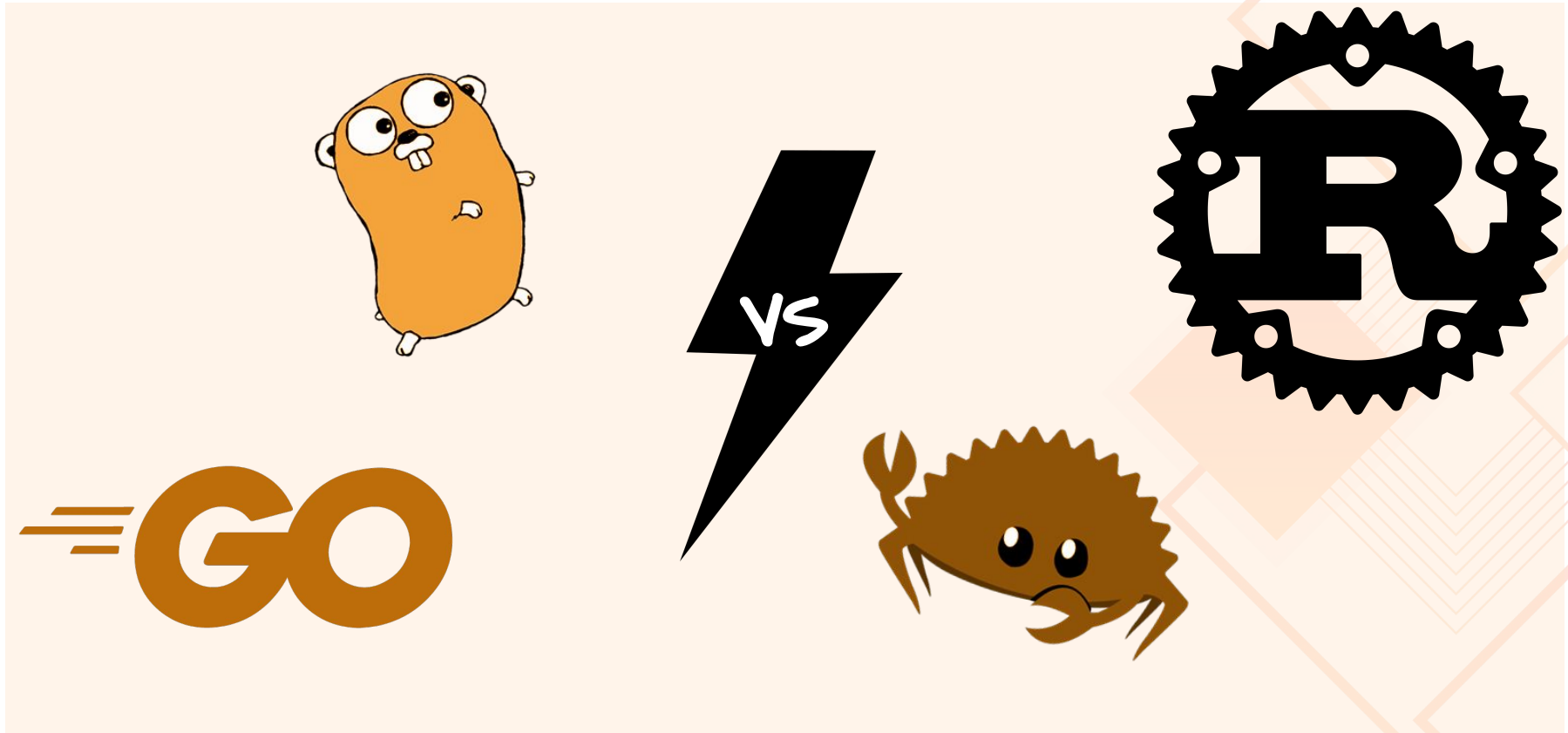
- Fresh Ubuntu 22.04 (cloud image)
 - Install k3s
 - Install a CNI (calico)
 - Install kata-deploy manifest
 - Run a simple container with kata!



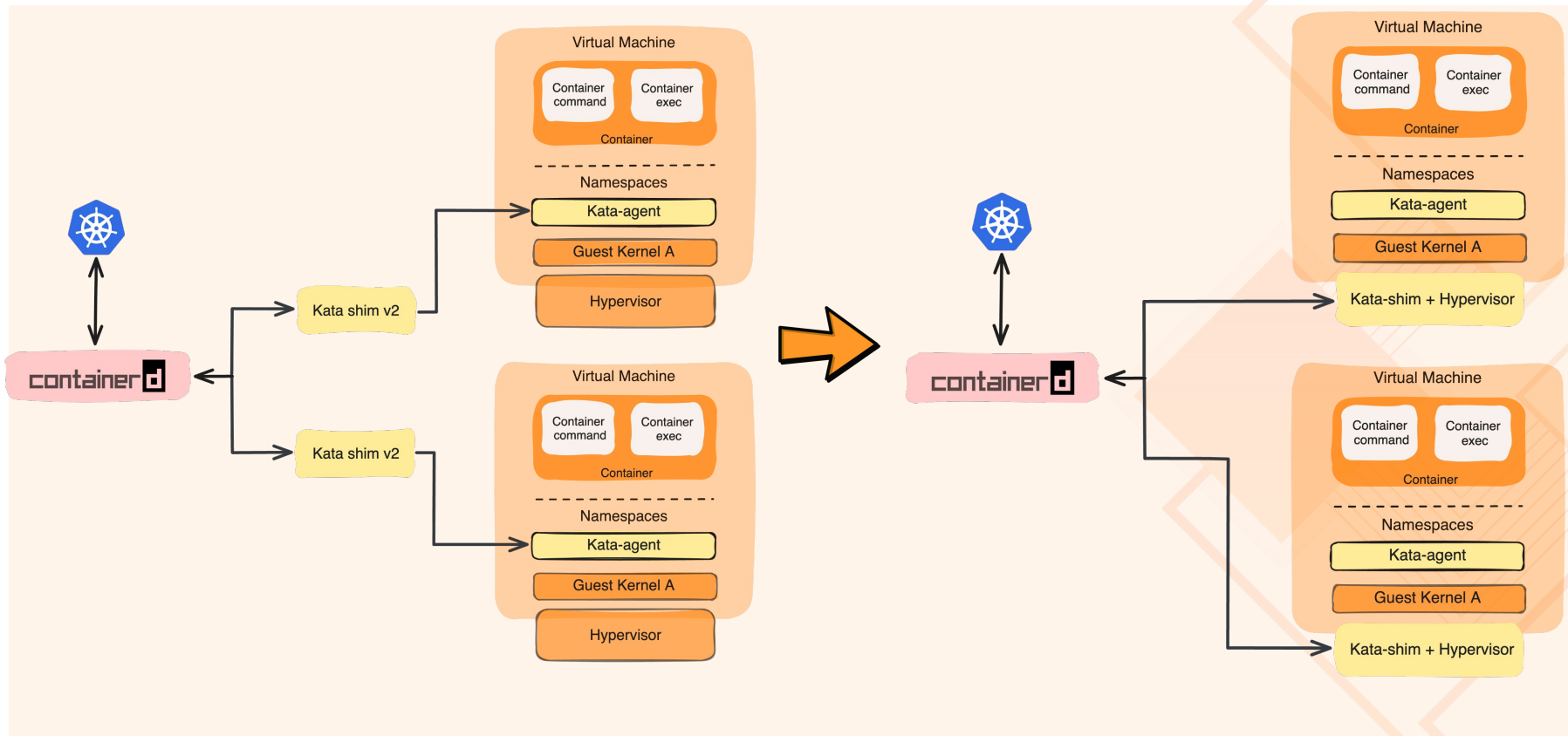
- Fresh Ubuntu 22.04 (cloud image)
 - Install containerd
 - Install CNI
 - Download & unpack release binaries
 - Run a simple container with kata!



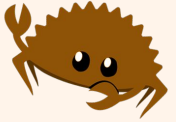
[nubificus/openinfradayshu-demos](https://github.com/nubificus/openinfradayshu-demos)



runtime vs runtime-rs



- Unify the runtime + hypervisor
- Reduce Kata Containers resource consumption and management complexity
- Integrated Rust hypervisor ensures that Kata Containers only spawn one host component for each POD.
- Aligns with the popular trends in the Linux community to rustify core software stacks.



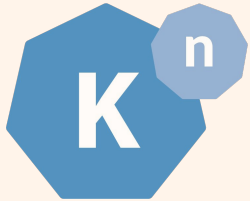
Hypervisor	runtime	runtime-rs
QEMU	★	☆
Cloud-hypervisor	★	☆
Firecracker	★	☆
Dragonball	—	★



Cloud
Hypervisor



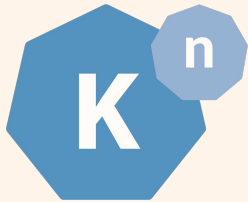
- A comparison of feature differences between Kata 3.0/runtime-rs and Kata 2.x and alignment status [#8702](#)
- Developers identified 66 distinct features of the Go runtime that should be available on the rust runtime: of those, only 15 are not **yet** available
- Most Probably, release v4.0 will come with runtime-rs as the default runtime!



Serverless Sandboxes



Expose hardware
acceleration functions



Sandbox user-submitted code using kata-containers:

- Protect the infrastructure from malicious users
- Extend Knative's threat model

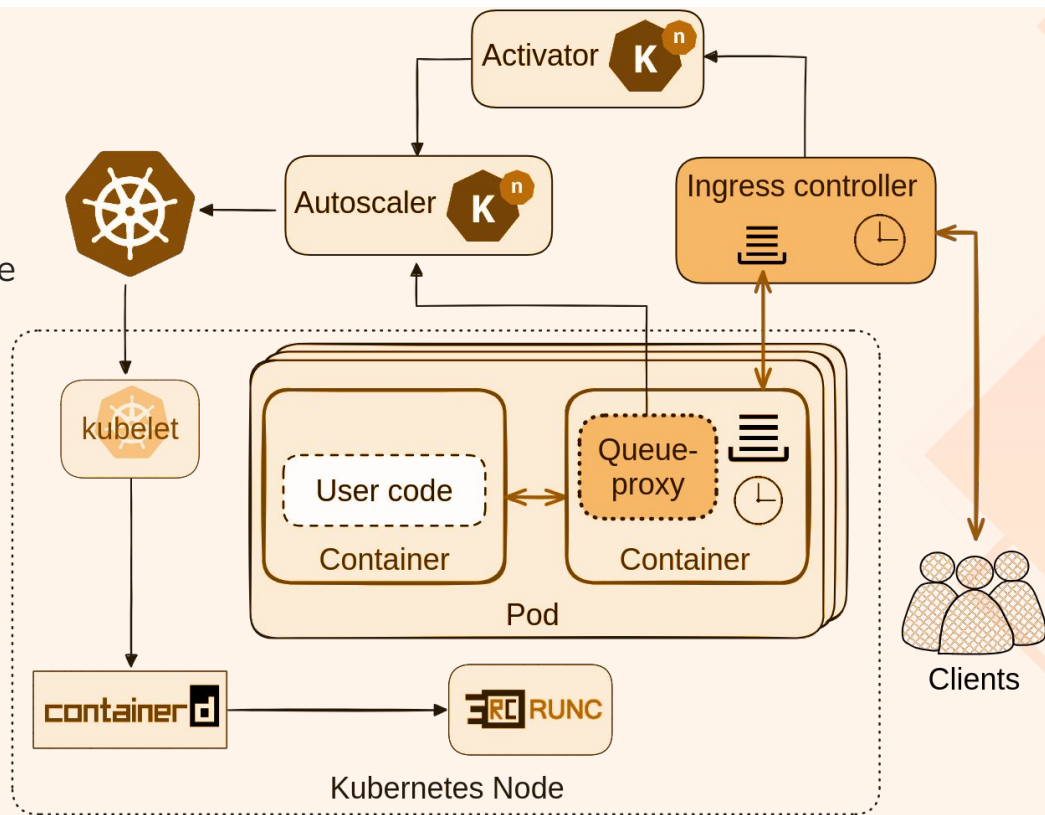


API-remoting for sandboxed workloads:

- User-code never touches the accelerator
- Accelerator sharing without PCI/mediated passthrough

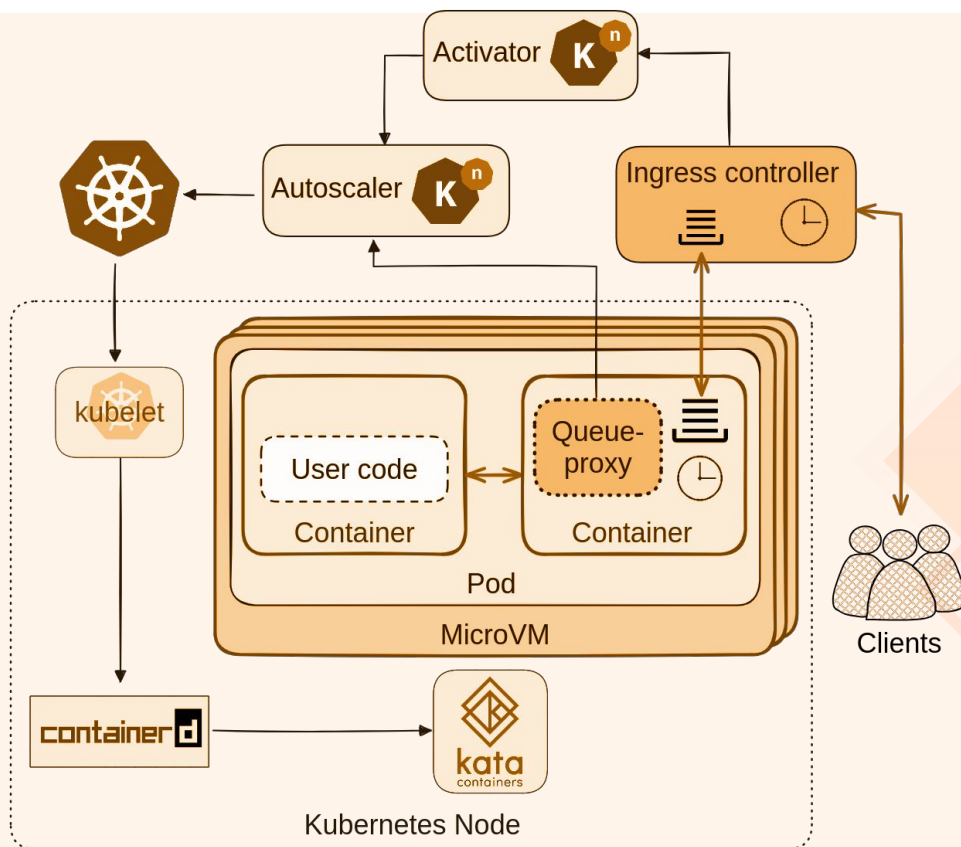
Pods consist of generic containers:

- ✗ no true isolation from the rest of the infrastructure



Pods run containers inside the microVM sandbox:

- ✓ protect the rest of the infrastructure from user-submitted code



- ✓ setup k3s cluster
- ✓ setup kata (kata-deploy)
- ✓ install knative
- ✓ setup ingress/DNS

- ✓ deploy helloworld service
 - simple HTTP header echo

Point to:

<https://hellocontainer.openinfra.nbfci.io>

<https://hellors.openinfra.nbfci.io/>

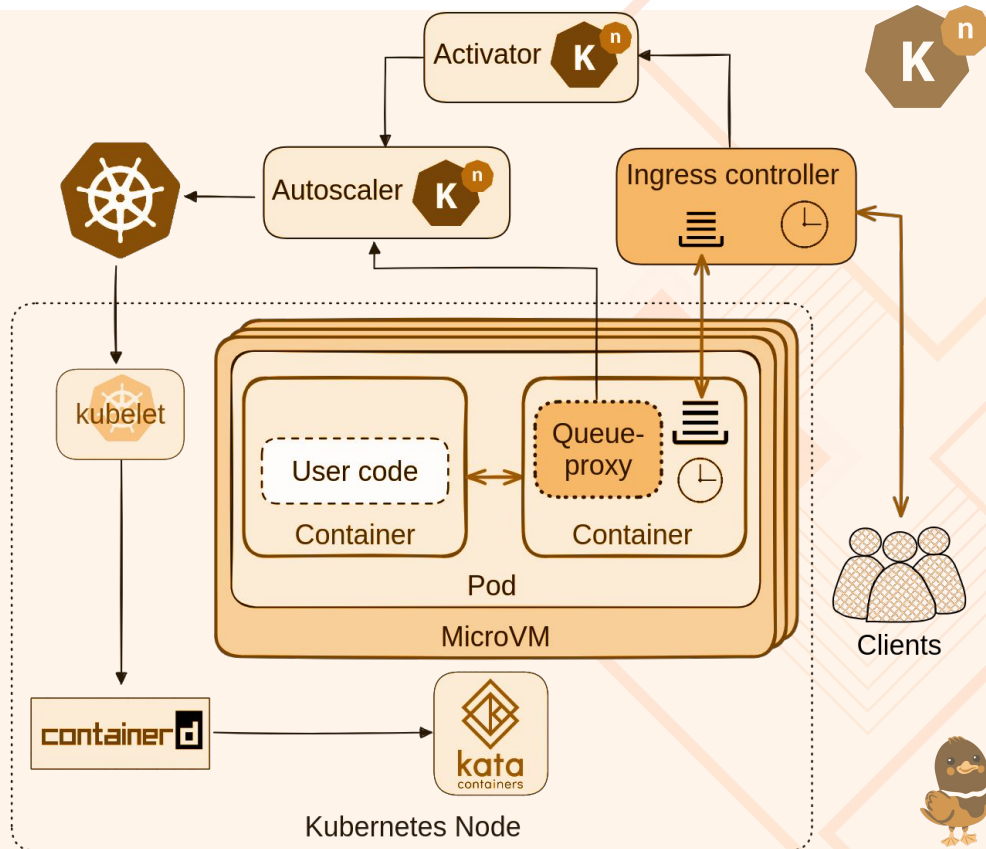
<https://helloclh.openinfra.nbfci.io/>

<https://hellogemu.openinfra.nbfci.io/>

<https://hellofc.openinfra.nbfci.io/>



[nubificus/openinfradayshu-demos](https://github.com/nubificus/openinfradayshu-demos)



API-remoting for sandboxed workloads:

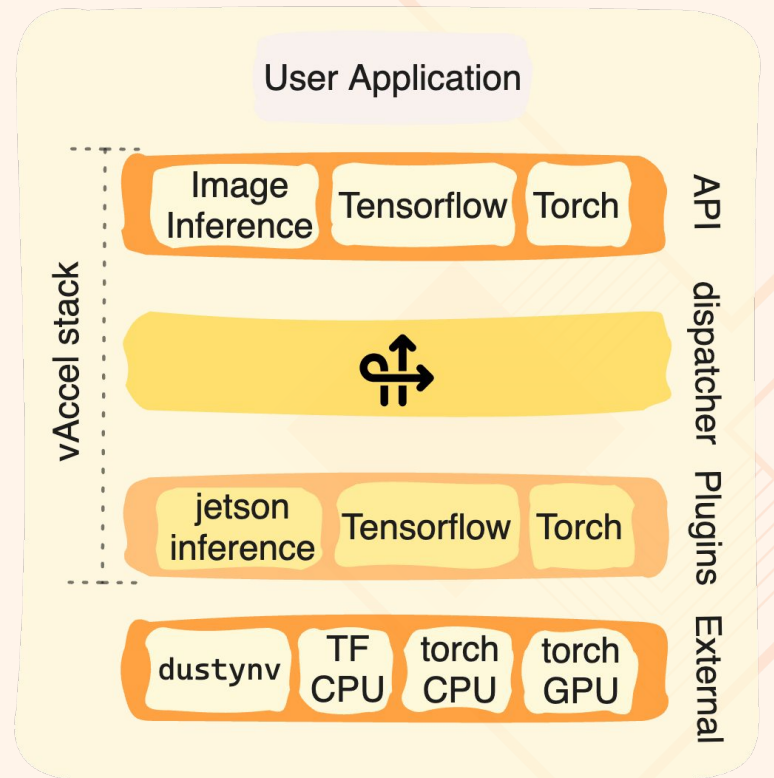
- User-code never touches the accelerator
- Accelerator sharing without PCI/mediated passthrough

WiP, under development!

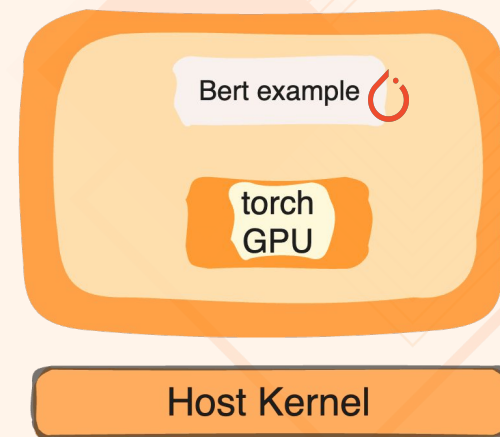
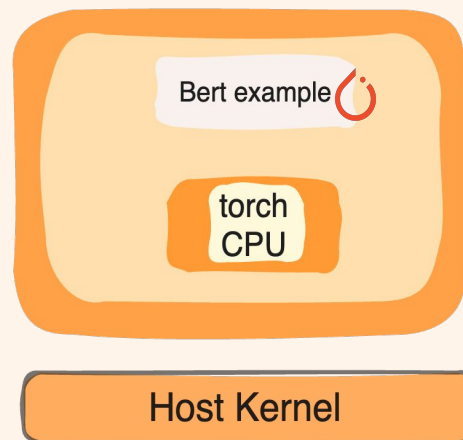
<https://docs.vaccl.org/>

 [cloudkernels/vacclrt](https://github.com/cloudkernels/vacclrt)



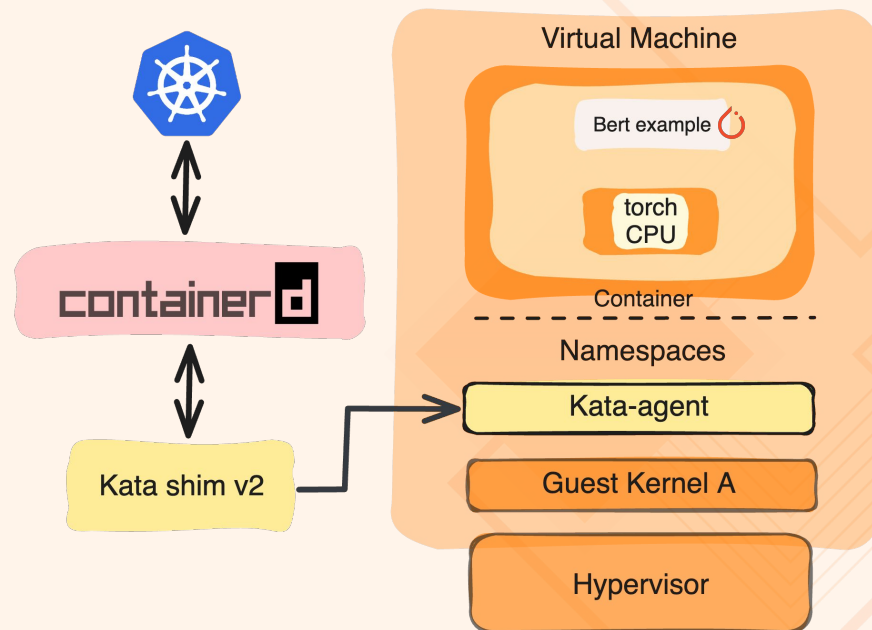


- Simple Torch example:
 - BERT model, speech classification
 - hate-speech
 - offensive-language
 - neutral
 - CPU / GPU implementation
- 1000 tweets
- Run locally (CPU/GPU)
- Run in a sandbox container (CPU, no GPU)
- Run in a sandbox container (GPU, vAccel)



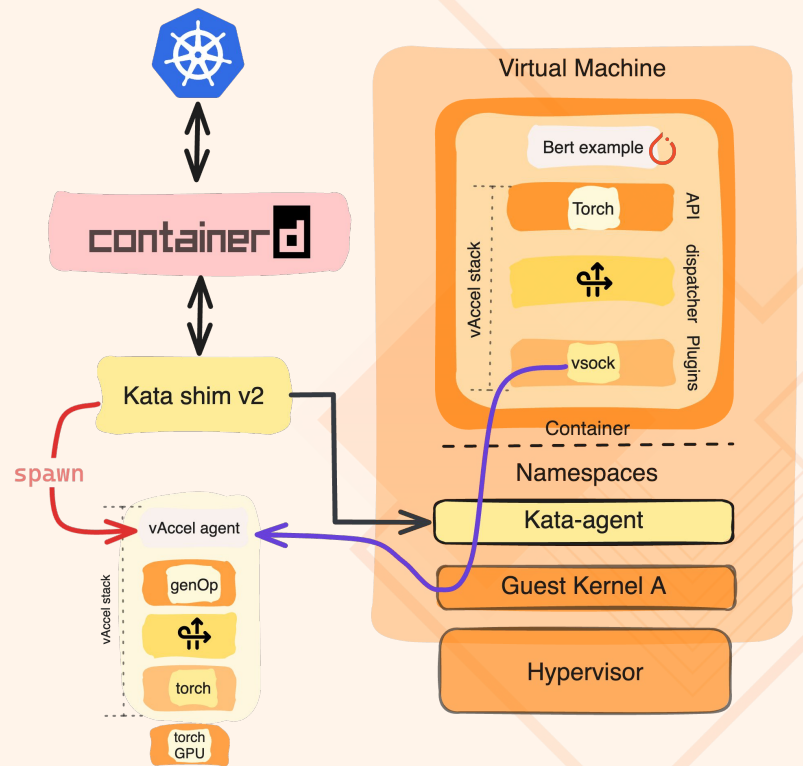
[nubificus/openinfra-dayshu-demos](https://github.com/nubificus/openinfra-dayshu-demos)

- Simple Torch example:
 - BERT model, speech classification
 - hate-speech
 - offensive-language
 - neutral
 - CPU / GPU implementation
- 1000 tweets
- Run locally (CPU/GPU)
- Run in a sandbox container (CPU, no GPU)
- Run in a sandbox container (GPU, vAccel)



- Simple Torch example:
 - BERT model, speech classification
 - hate-speech
 - offensive-language
 - neutral
 - CPU / GPU implementation
- 1000 tweets
- Run locally (CPU/GPU)
- Run in a sandbox container (CPU, no GPU)
- Run in a sandbox container (GPU, vAccel)


vaccel





Part of the work presented is supported by Horizon Europe RIA actions, MLSysOps (GA: 101092912), DESIRE6G (GA: 101096466), and EMPYREAN (GA: 101136024)



- Kata-containers Overview
- Installation: `kata-deploy` / static release
- Go vs Rust runtime
- Use-cases: sandboxing / Hardware acceleration (vAccel) / Serverless Sandboxes
- Try it out:
 - <https://katacontainers.io>
 -  [kata-containers/kata-containers](https://github.com/kata-containers/kata-containers)
- Release v4.0 is coming soon!
 - runtime-rs
 - Enhanced hypervisor support
 - Enhanced Confidential Containers



<https://confidentialcontainers.org/>

OpenInfraDays Hungary, Budapest, Jun 3rd, 2024



Thanks!

Anastassios Nanos

Systems Researcher @ Nubificus LTD

