Microservices, Containerization, and Orchestration:

Implications on Cloud-Native Security and Performance

Hui Lu, The University of Texas at Arlington (UTA)

Nathan Daughety, Air Force Research Laboratory (AFRL)

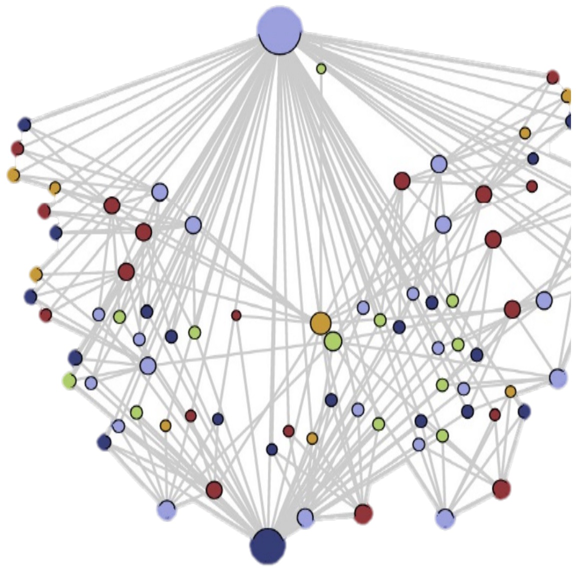
Approved for Public Release; Distribution Unlimited: PA
CLEARANCE PENDING – DO NOT DISTRIBUTE

The “Cloud-Native” Era

- **Cloud native** – via **decomposition** and **automation** – offer means to develop and manage cloud applications more effectively.
- A monolithic application is **decomposed** into graphs of ***single-purpose, loosely-coupled*** microservices

Monolithic Applications

»



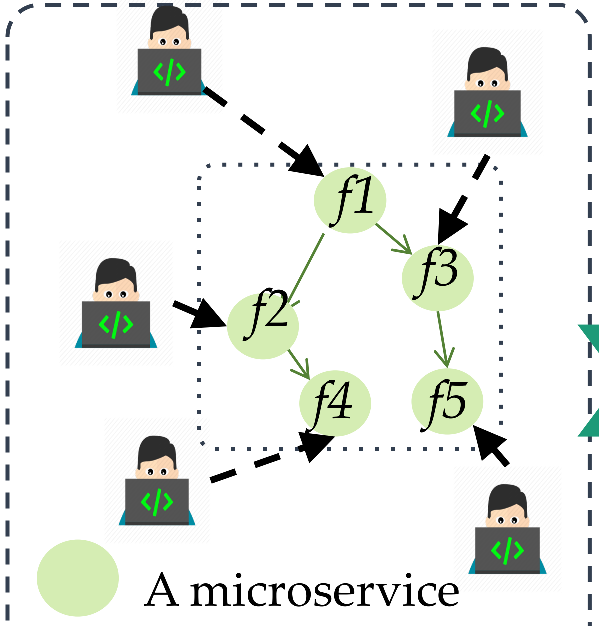
Microservices

Reduced **development** complexity

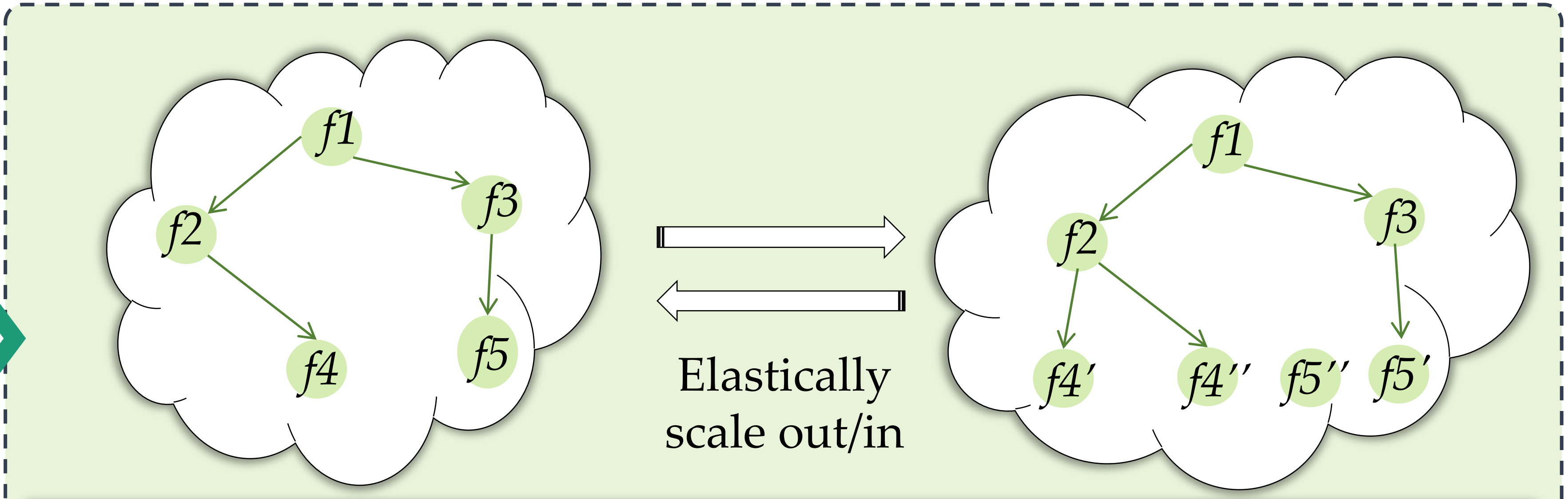
Increased **code** velocity

The “Cloud-Native” Era

- **Cloud native** – via **decomposition** and **automation** – offer means to develop and manage cloud applications more effectively
- The deployment and management of microservices-based cloud-native applications can be **automatically** handled by cloud-native platforms



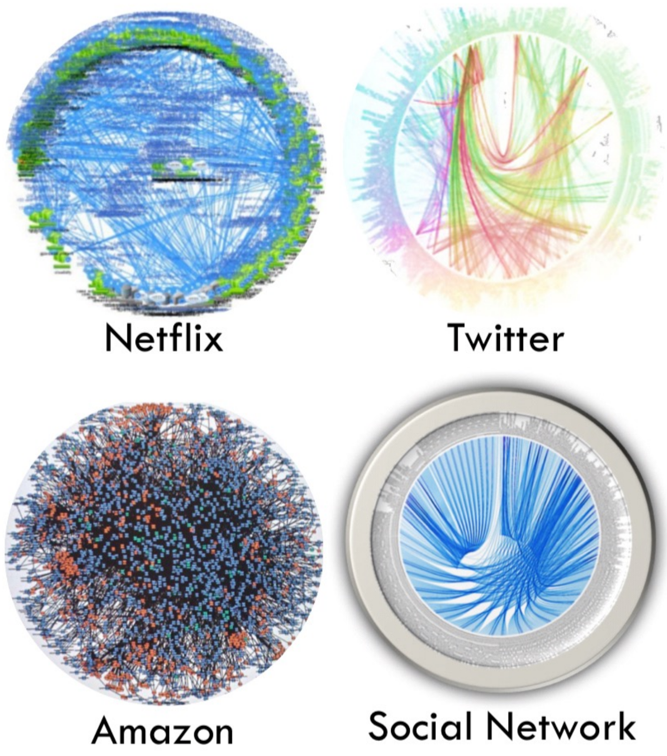
A microservices based cloud-native **application**



The cloud-native application is deployed, managed and automatically-scaled by **cloud-native platforms**

The “Cloud-Native” Era

- The use of cloud-native technologies is becoming pervasive

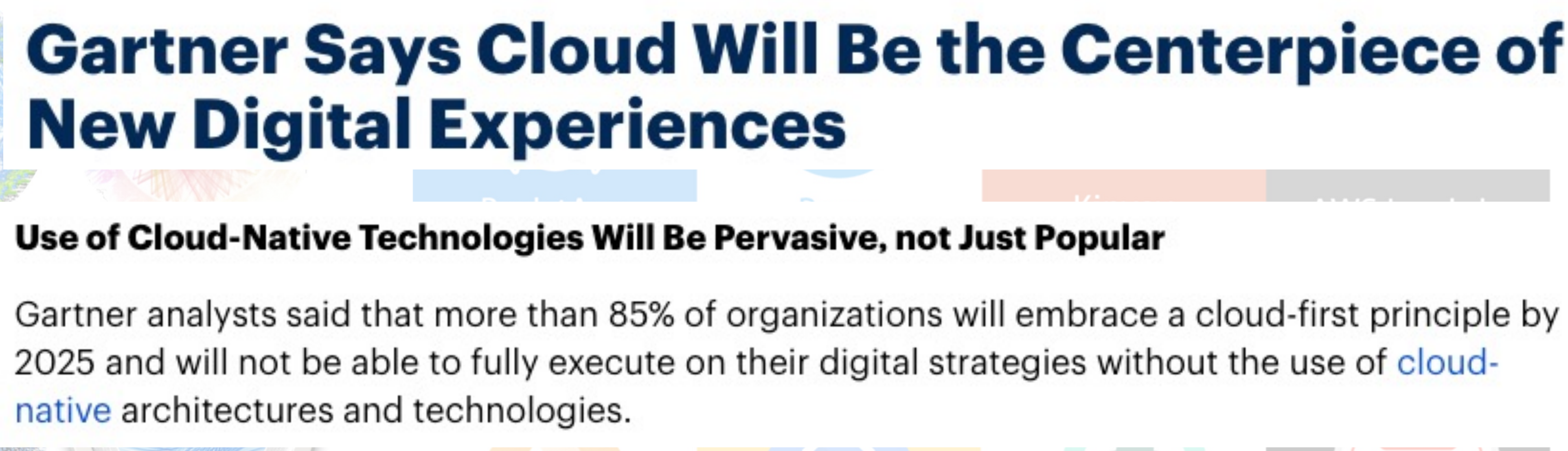


Microservices

Cloud-native platforms

The "Cloud-Native" Era

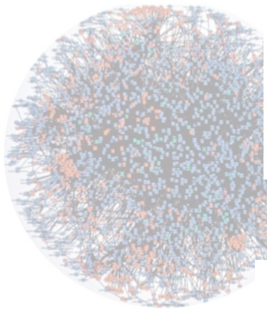
- The use of cloud-native technologies is becoming pervasive

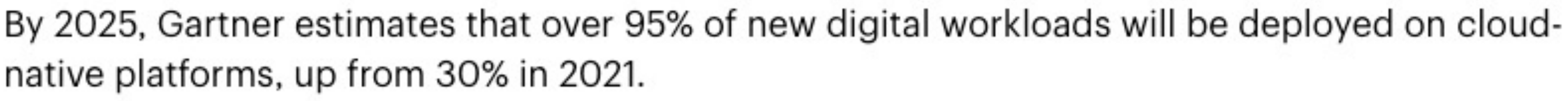


Gartner Says Cloud Will Be the Centerpiece of New Digital Experiences

Use of Cloud-Native Technologies Will Be Pervasive, not Just Popular

Gartner analysts said that more than 85% of organizations will embrace a cloud-first principle by 2025 and will not be able to fully execute on their digital strategies without the use of cloud-native architectures and technologies.



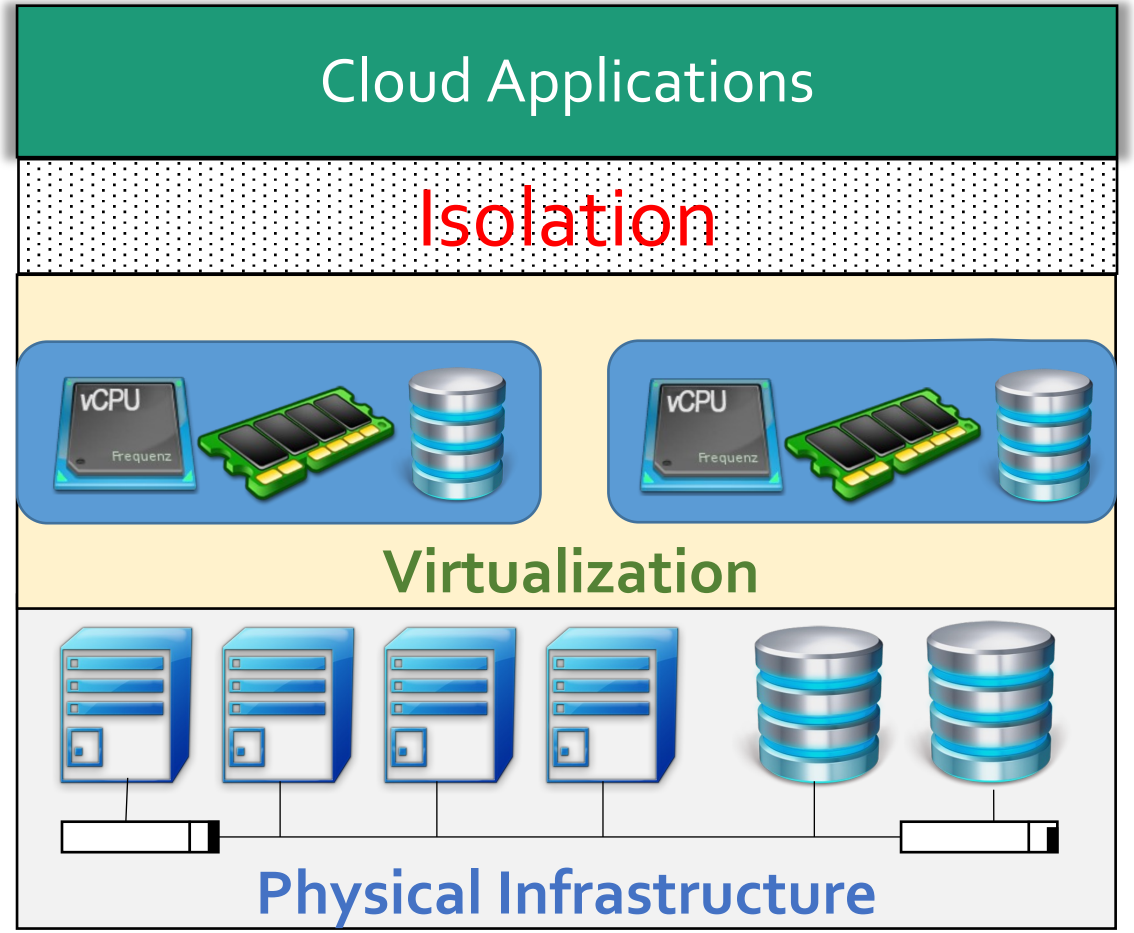


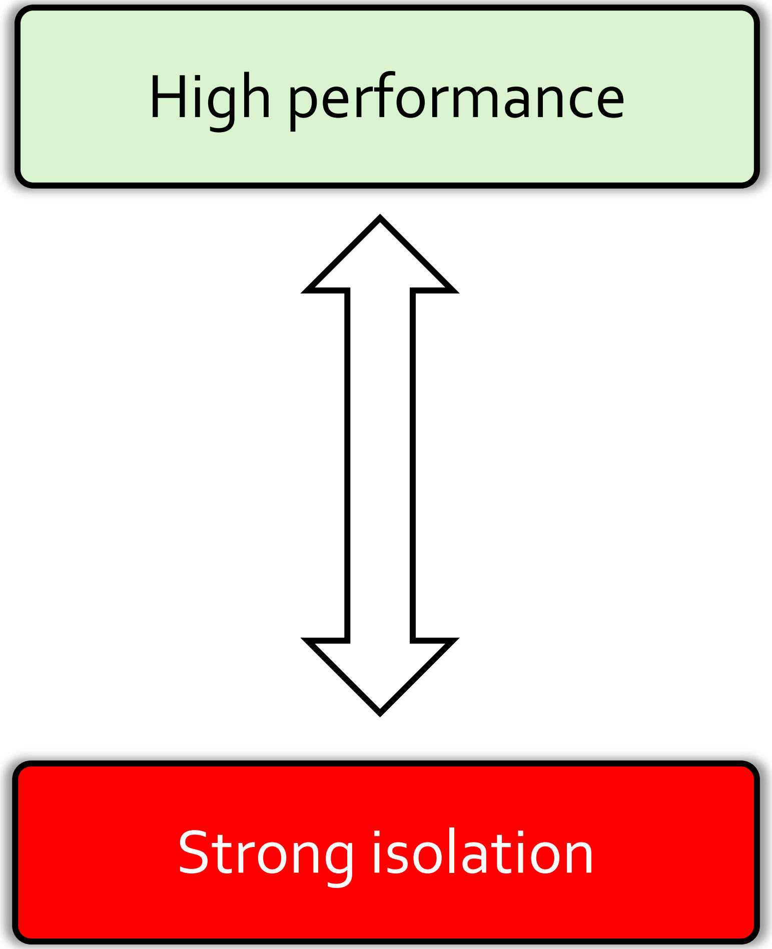
By 2025, Gartner estimates that over 95% of new digital workloads will be deployed on cloud-native platforms, up from 30% in 2021.

Microservices

Cloud-native platforms

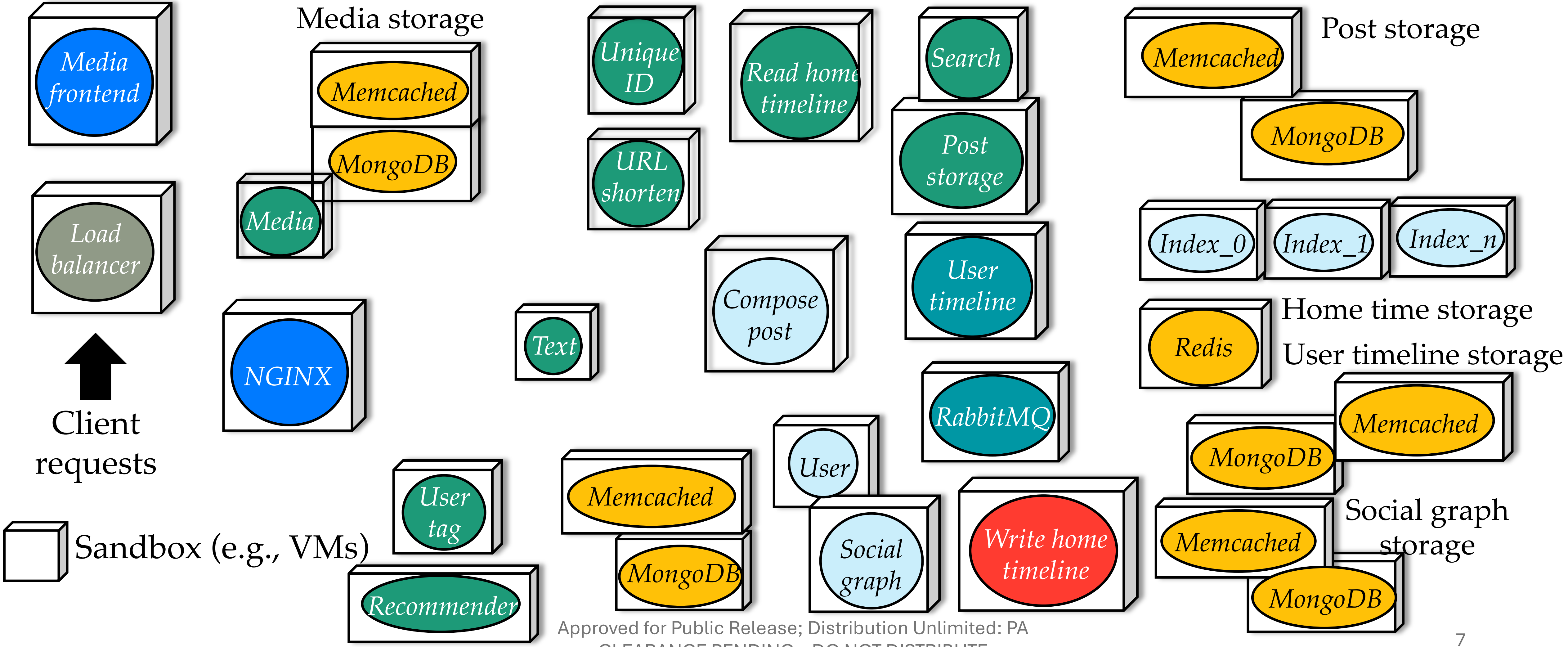
Performance and Security in Cloud (Native)





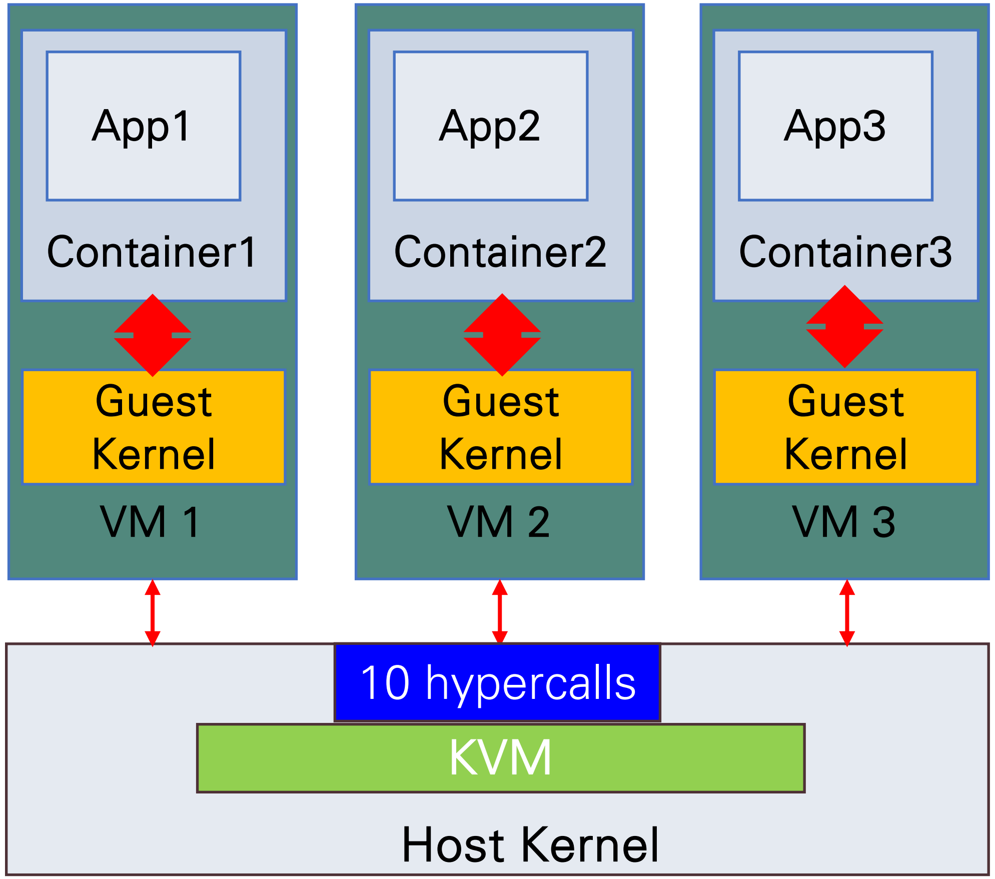
Microservices are **Containerized**

- Through sandboxes such as VMs and containers

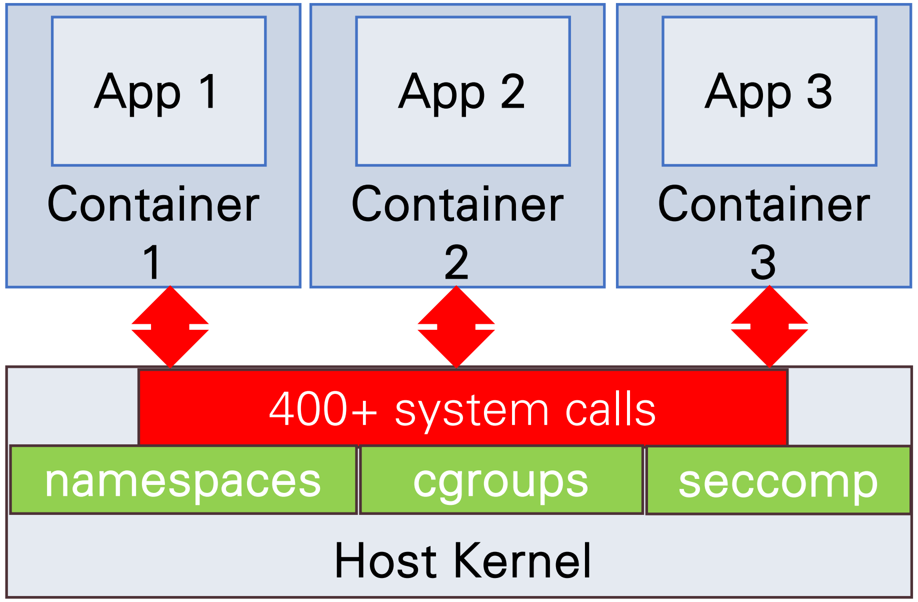


Approved for Public Release; Distribution Unlimited: PA CLEARANCE PENDING – DO NOT DISTRIBUTE

Isolation and Performance



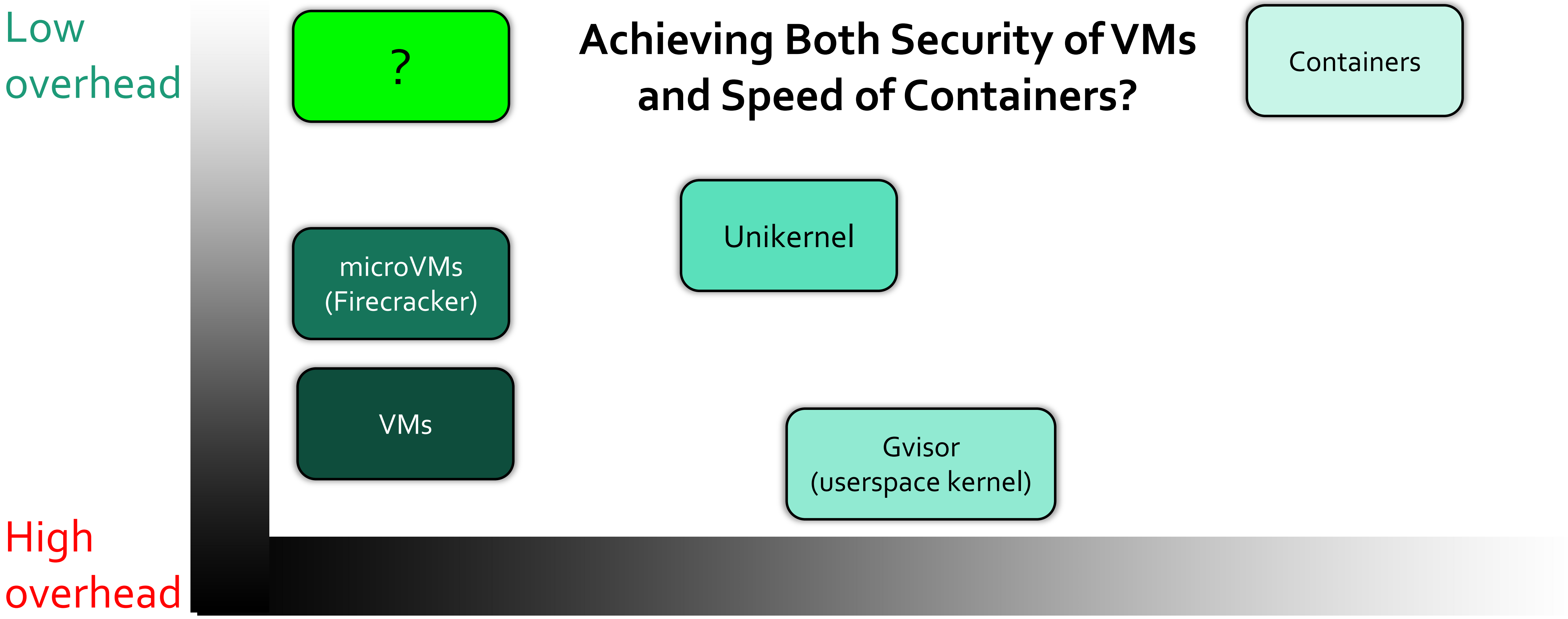
VMs



Containers

Approved for Public Release; Distribution Unlimited: PA
CLEARANCE PENDING – DO NOT DISTRIBUTE

Isolation and Performance – More Efforts



Low overhead

?

Achieving Both Security of VMs and Speed of Containers?

Containers

Unikernel

microVMs (Firecracker)

VMs

Gvisor (userspace kernel)

High overhead

Strong isolation

Weak isolation

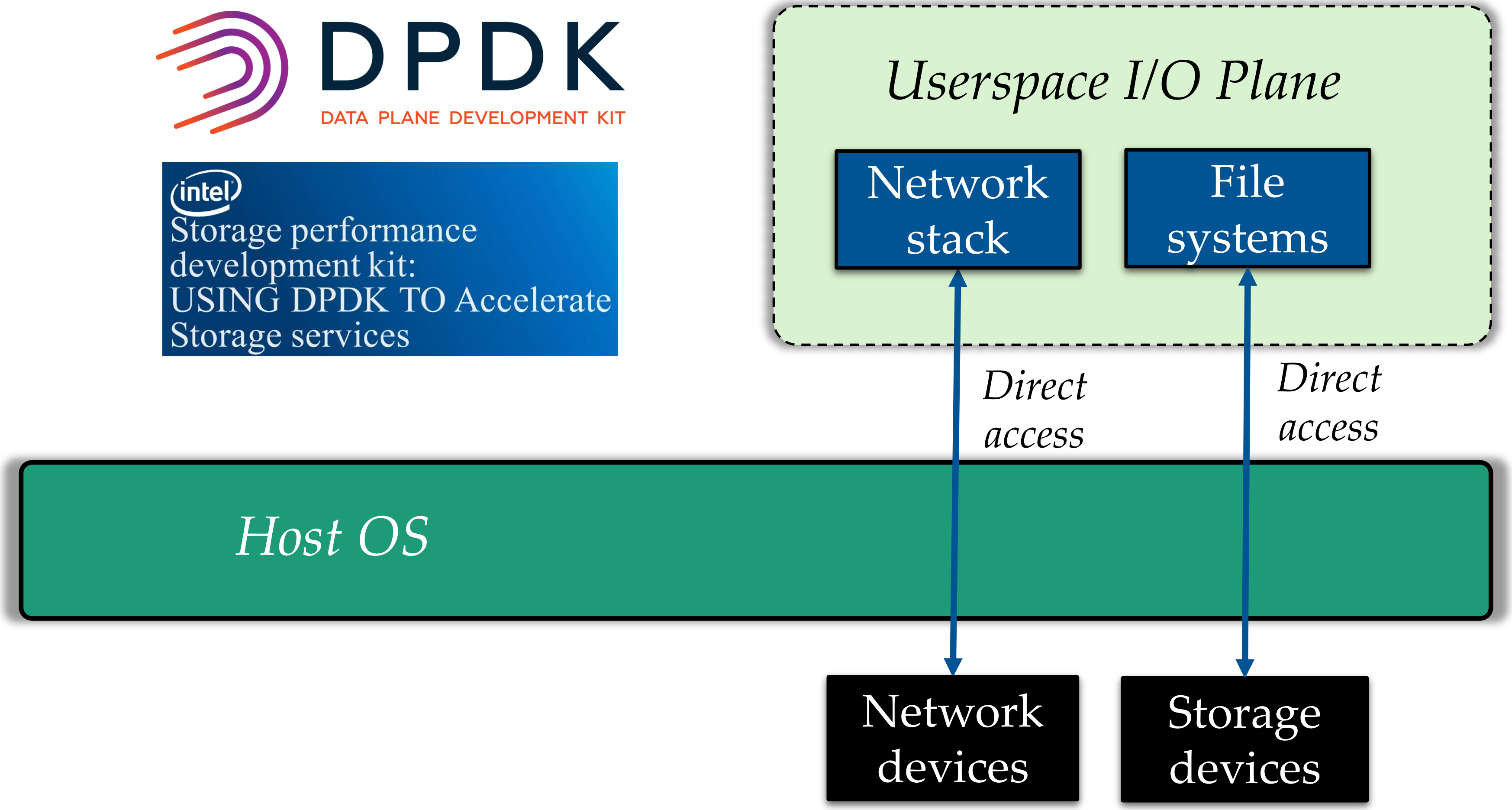
Approved for Public Release; Distribution Unlimited: PA CLEARANCE PENDING – DO NOT DISTRIBUTE

Userspace High-Performance I/O Data Plane

- Exploring Userspace Solutions

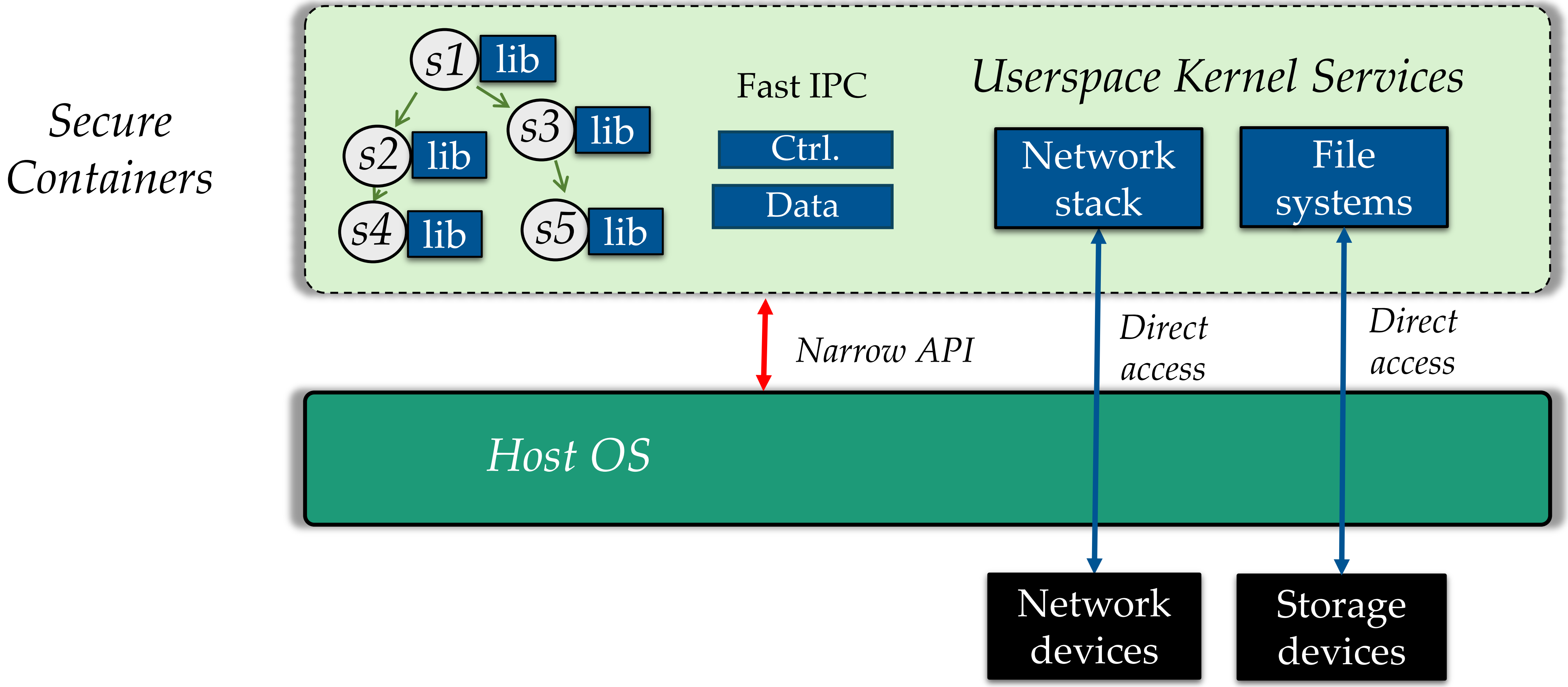
DPDK
DATA PLANE DEVELOPMENT KIT

intel
Storage performance development kit:
USING DPDK TO Accelerate Storage services



Strong yet Lightweight Isolation

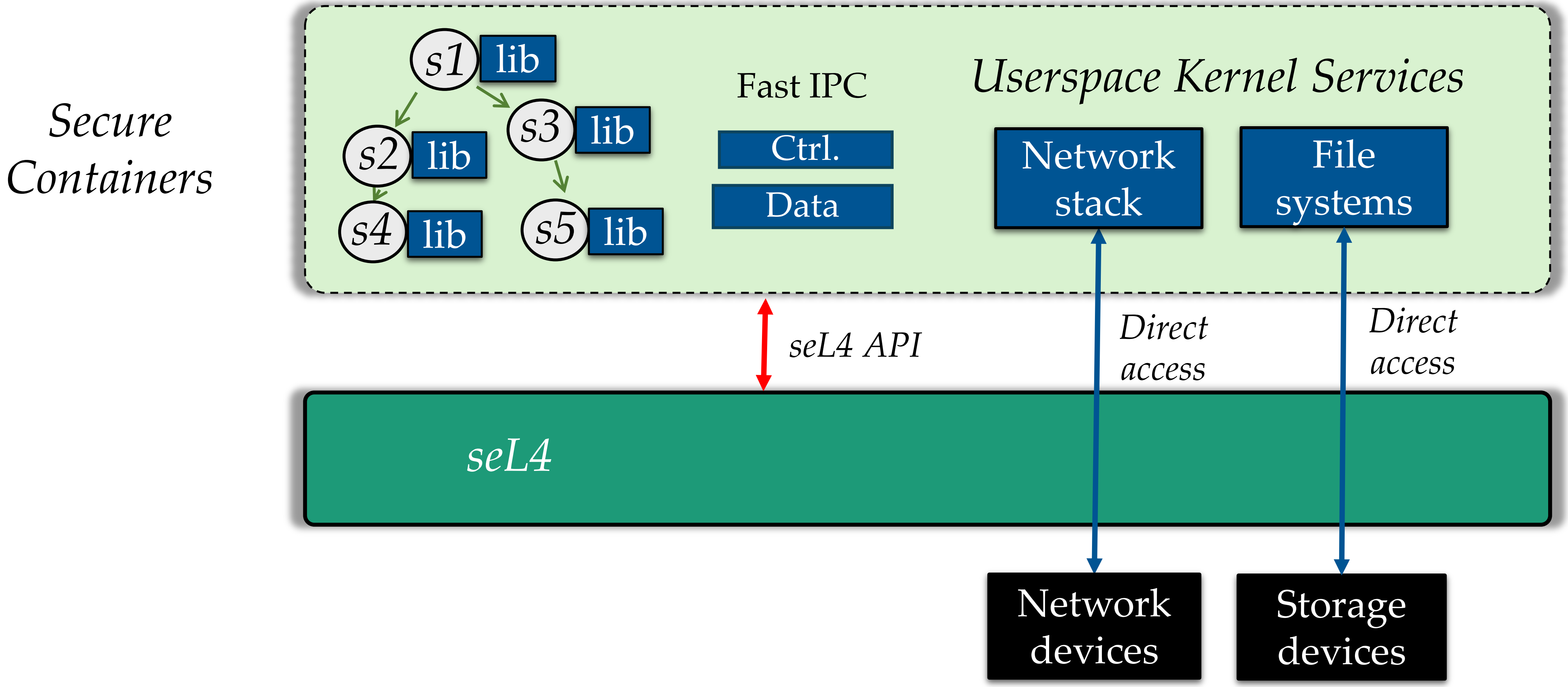
- Exploring Userspace Solutions



Approved for Public Release; Distribution Unlimited: PA CLEARANCE PENDING – DO NOT DISTRIBUTE

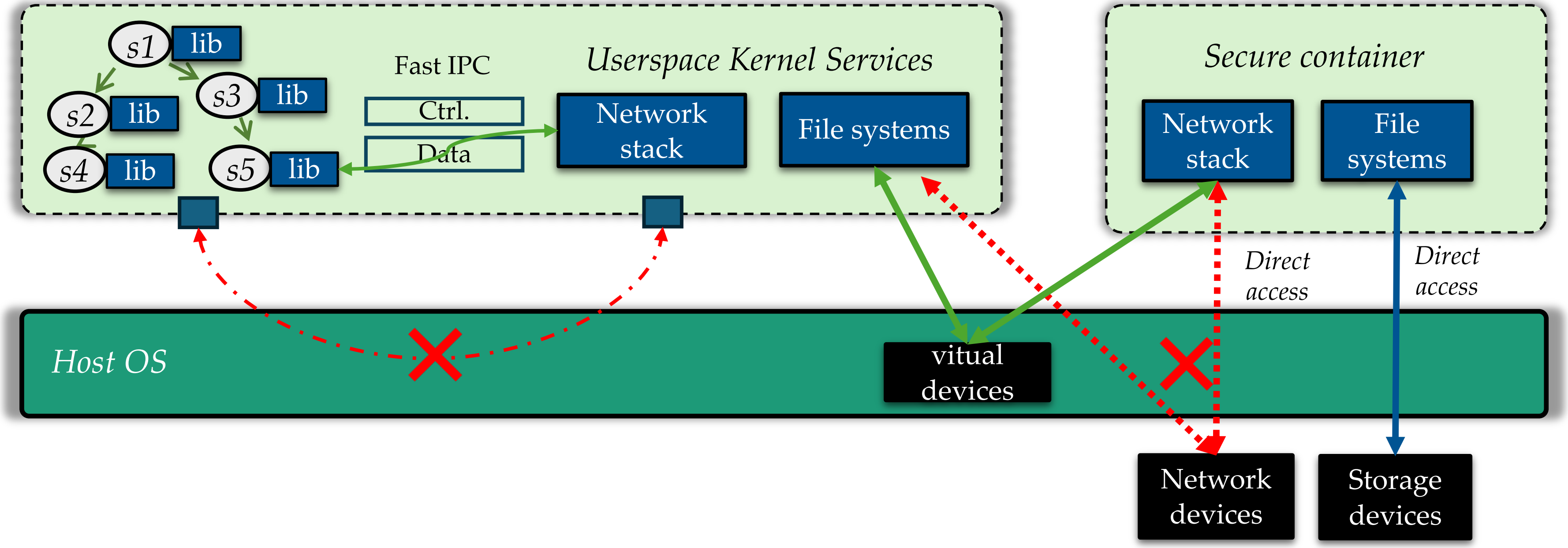
Strong yet Lightweight Isolation

- Exploring Userspace Solutions



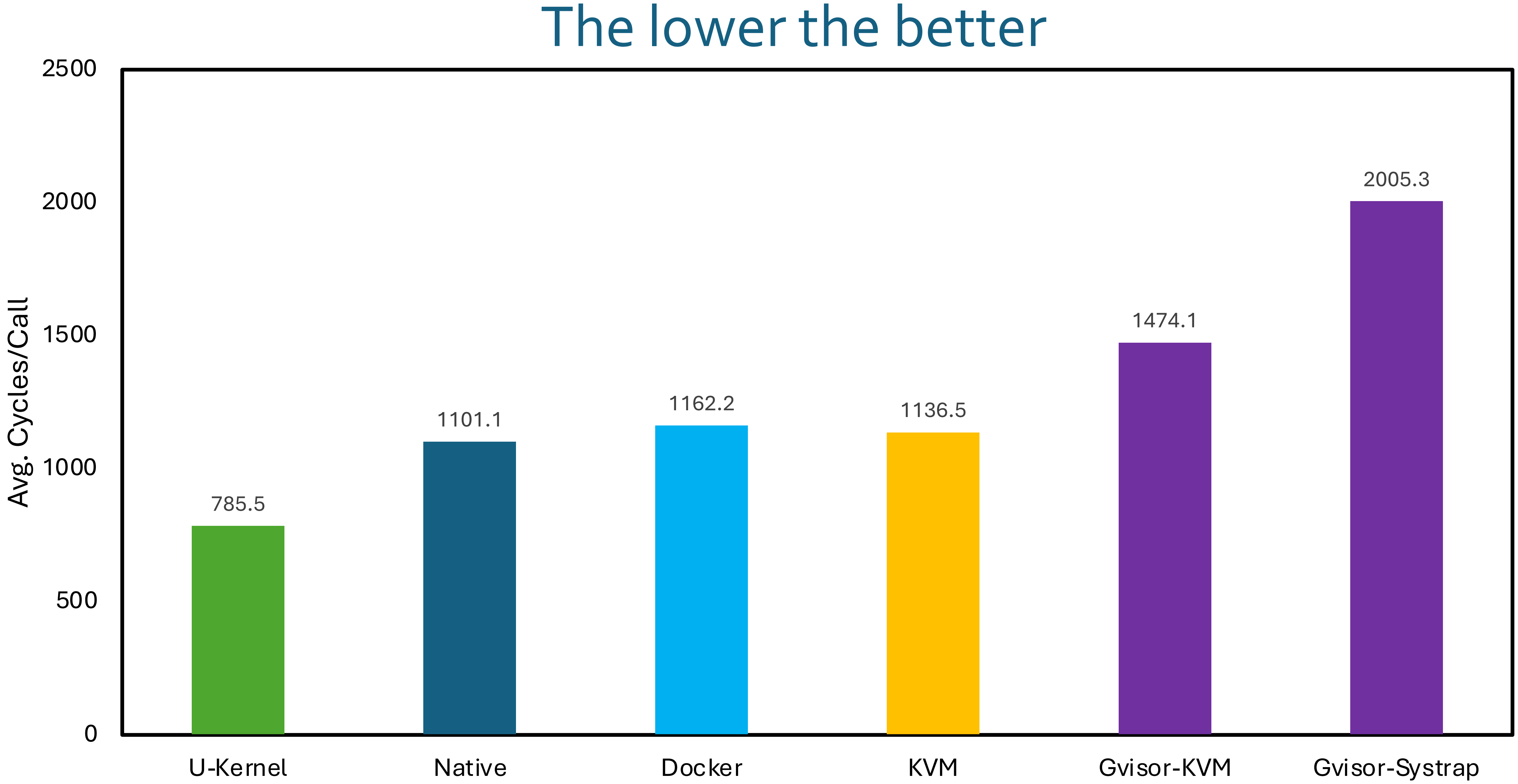
Key Challenges

- Support legacy applications
- Efficient userspace IPCs
- Resource sharing



Approved for Public Release; Distribution Unlimited: PA
CLEARANCE PENDING – DO NOT DISTRIBUTE

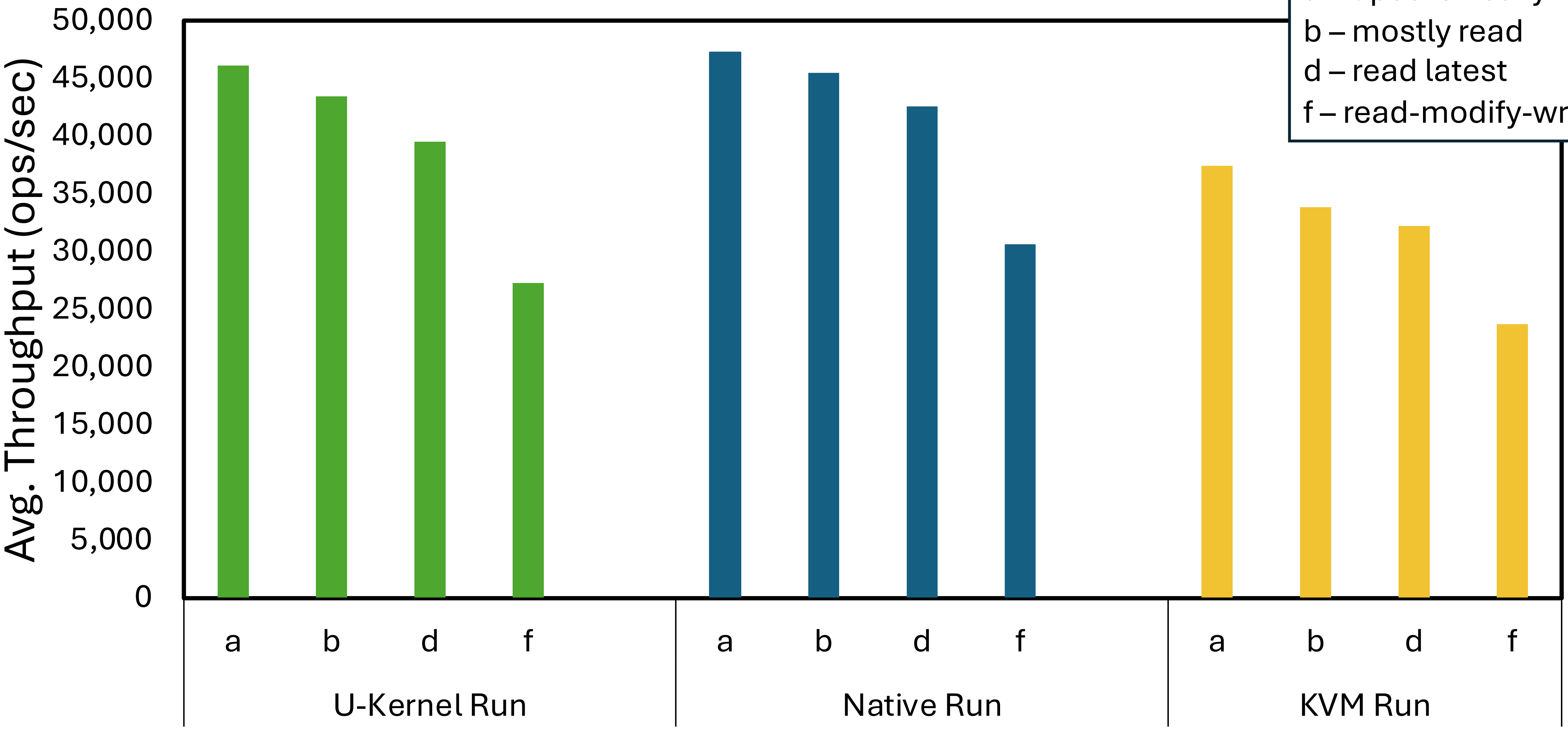
Preliminary Results – Syscall Latency



Approved for Public Release; Distribution Unlimited: PA
CLEARANCE PENDING – DO NOT DISTRIBUTE

Preliminary Results – YCSB over Redis

The higher the better

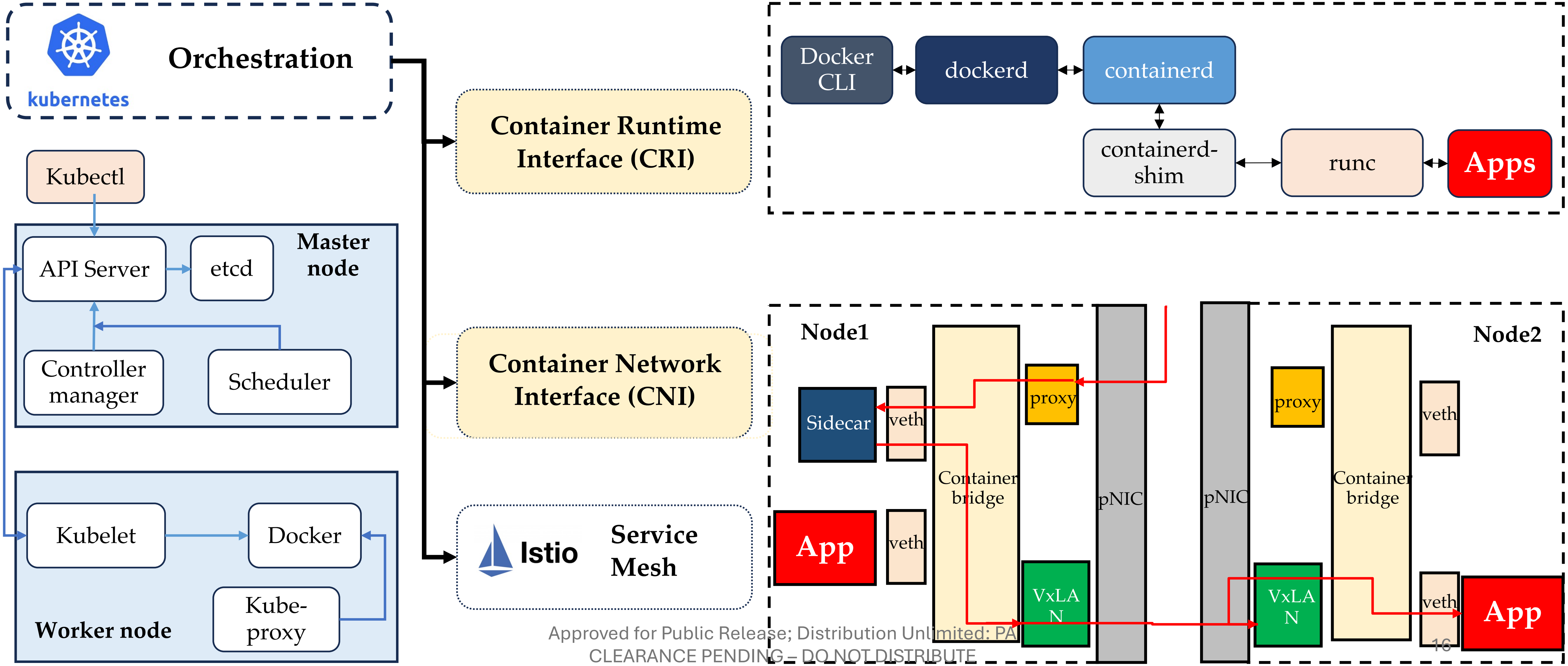


Workload Description:
a – update heavy
b – mostly read
d – read latest
f – read-modify-write

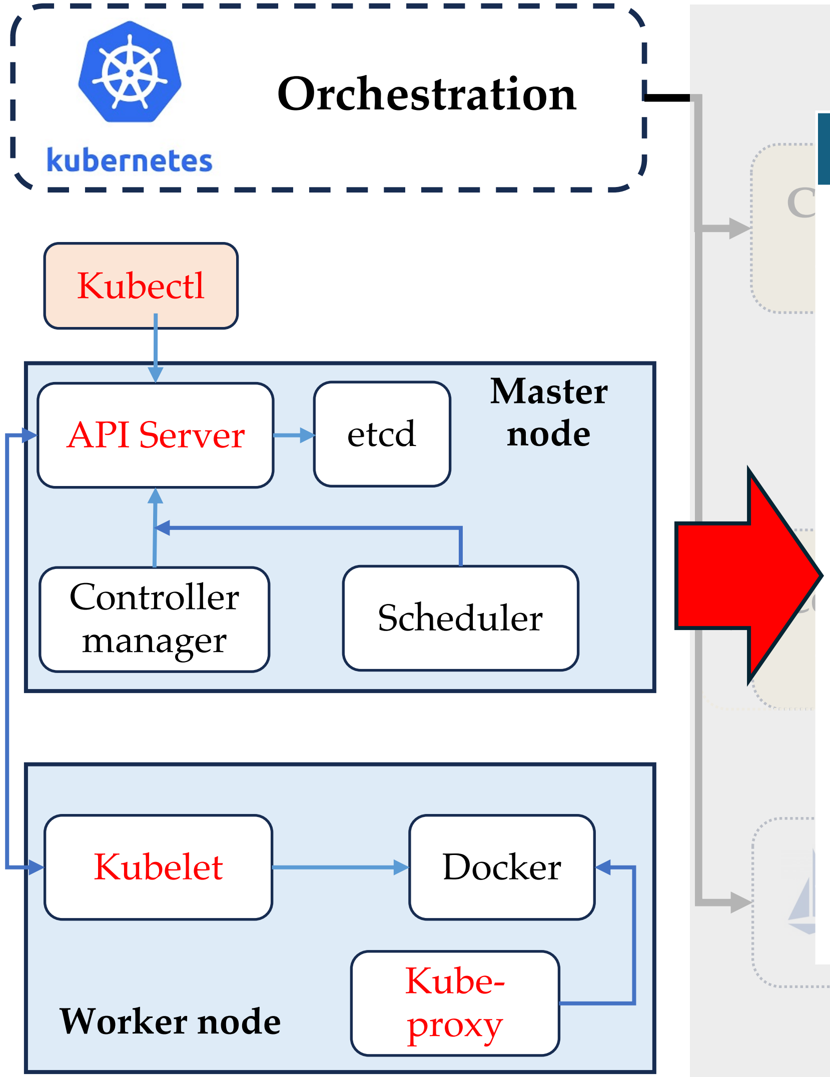
Approved for Public Release; Distribution Unlimited: PA
CLEARANCE PENDING – DO NOT DISTRIBUTE

Containerized Microservices are **Orchestrated**

- Through orchestration platforms, such as Kubernetes

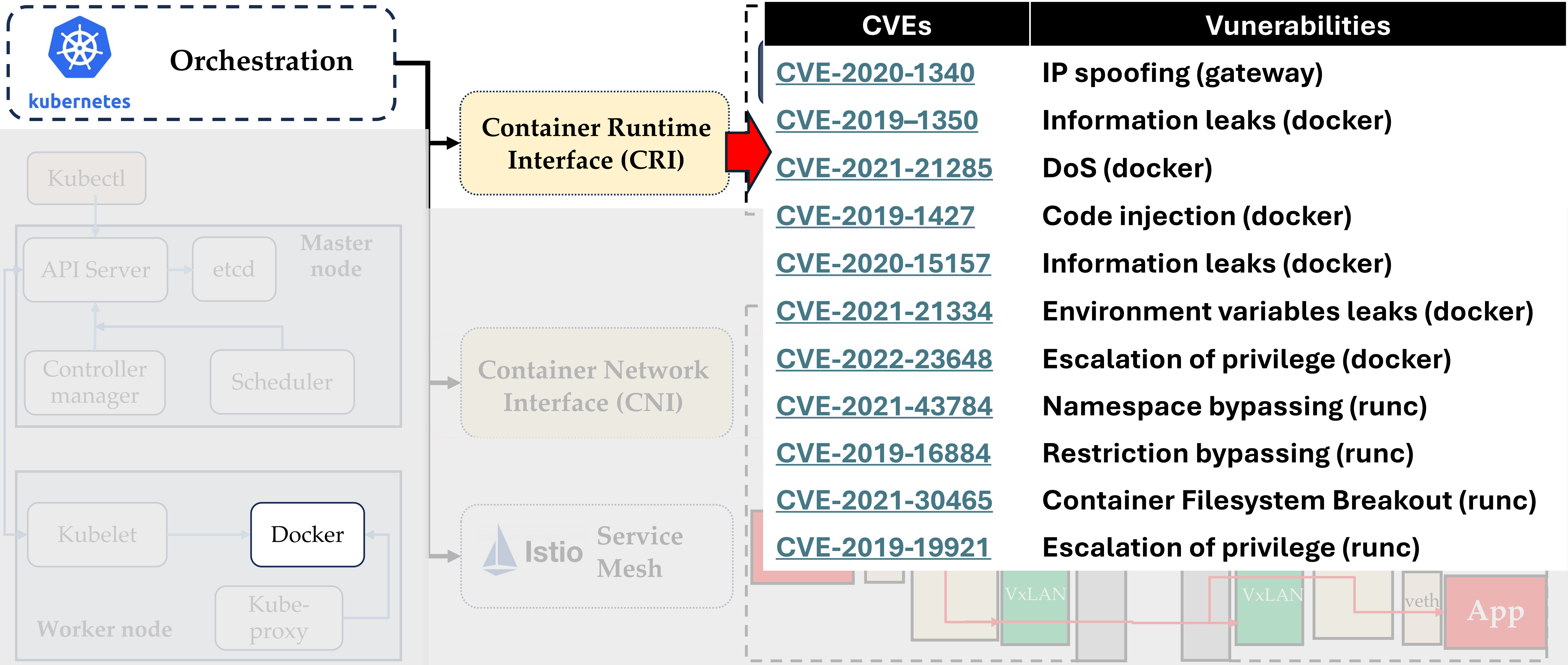


Vulnerability Analysis: Control Plane



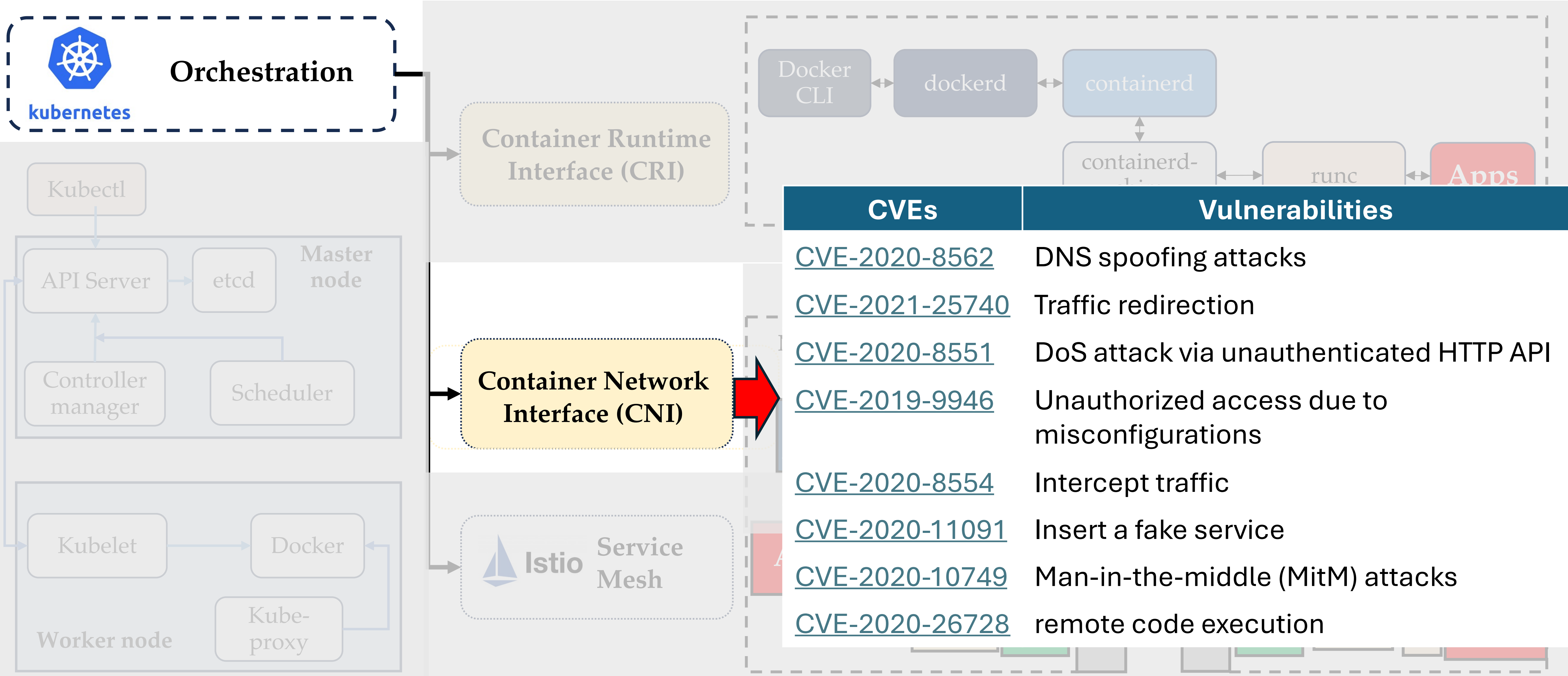
CVEs	Vunerabilities
CVE-2020-8555	Server Side Request Forgery (SSRF) (kube-controller-manager)
CVE-2021-25735	Security check bypass (kube-apiserver)
CVE-2019-11254	DoS (kube-apiserver)
CVE-2019-11247	Namespace isolation breakthrough (kube-apiserver)
CVE-2021-25743	Compromised services (Kubectl)
CVE-2019-11251	Unauthorized file operations (Kubectl)
CVE-2020-8558	unauthenticated requests (Kubelet and Kube-proxy)
CVE-2020-8557	DoS - Fill up disk space (Kubelet)
CVE-2020-8565	Leak of sensitive data (logger)
CVE-2020-8563	Leak of credentials (logger)

Vulnerability Analysis: Sandboxing Service



CVEs	Vunerabilities
CVE-2020-1340	IP spoofing (gateway)
CVE-2019–1350	Information leaks (docker)
CVE-2021-21285	DoS (docker)
CVE-2019-1427	Code injection (docker)
CVE-2020-15157	Information leaks (docker)
CVE-2021-21334	Environment variables leaks (docker)
CVE-2022-23648	Escalation of privilege (docker)
CVE-2021-43784	Namespace bypassing (runc)
CVE-2019-16884	Restriction bypassing (runc)
CVE-2021-30465	Container Filesystem Breakout (runc)
CVE-2019-19921	Escalation of privilege (runc)

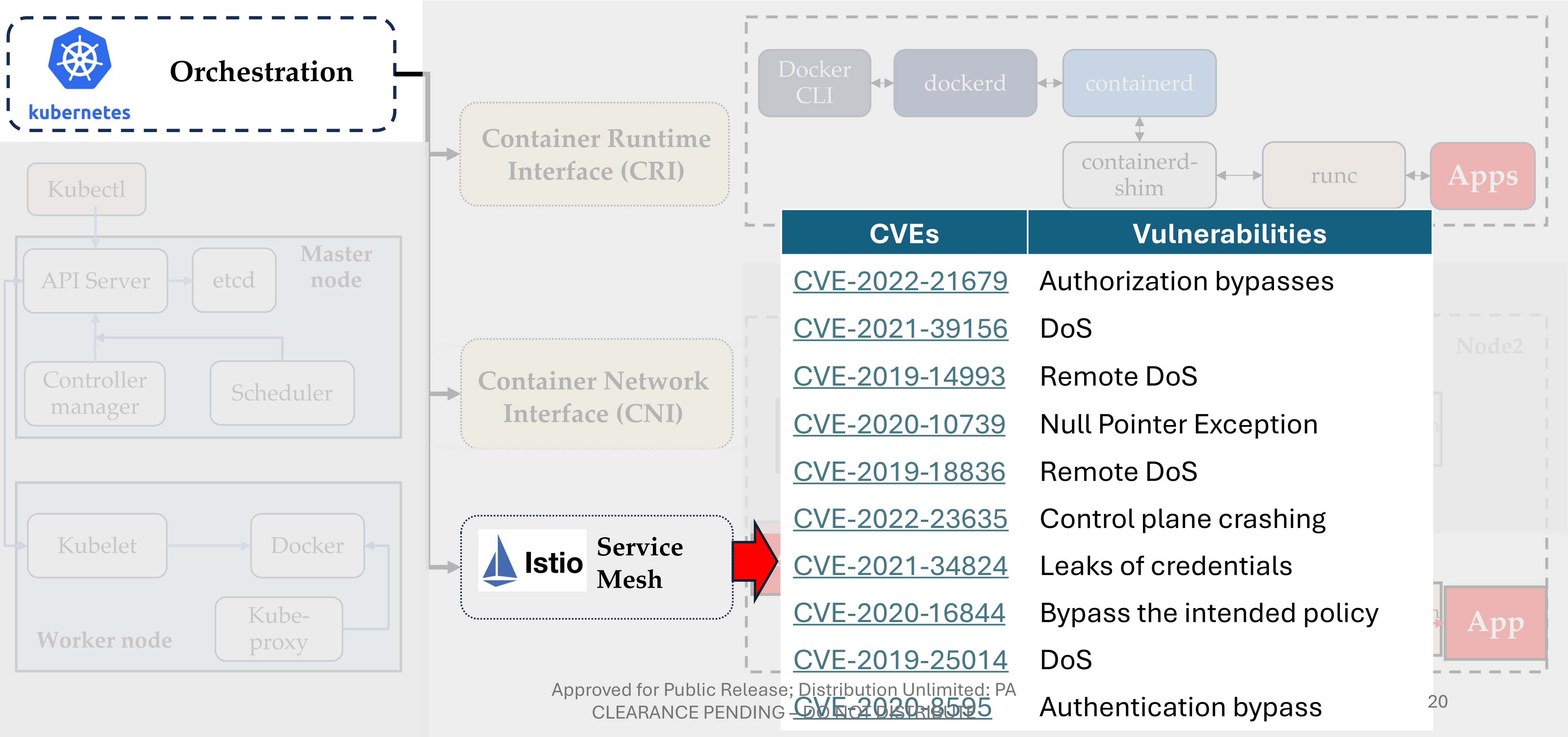
Vulnerability Analysis: Data Plane



CVEs	Vulnerabilities
CVE-2020-8562	DNS spoofing attacks
CVE-2021-25740	Traffic redirection
CVE-2020-8551	DoS attack via unauthenticated HTTP API
CVE-2019-9946	Unauthorized access due to misconfigurations
CVE-2020-8554	Intercept traffic
CVE-2020-11091	Insert a fake service
CVE-2020-10749	Man-in-the-middle (MitM) attacks
CVE-2020-26728	remote code execution

Approved for Public Release; Distribution Unlimited: PA CLEARANCE PENDING – DO NOT DISTRIBUTE

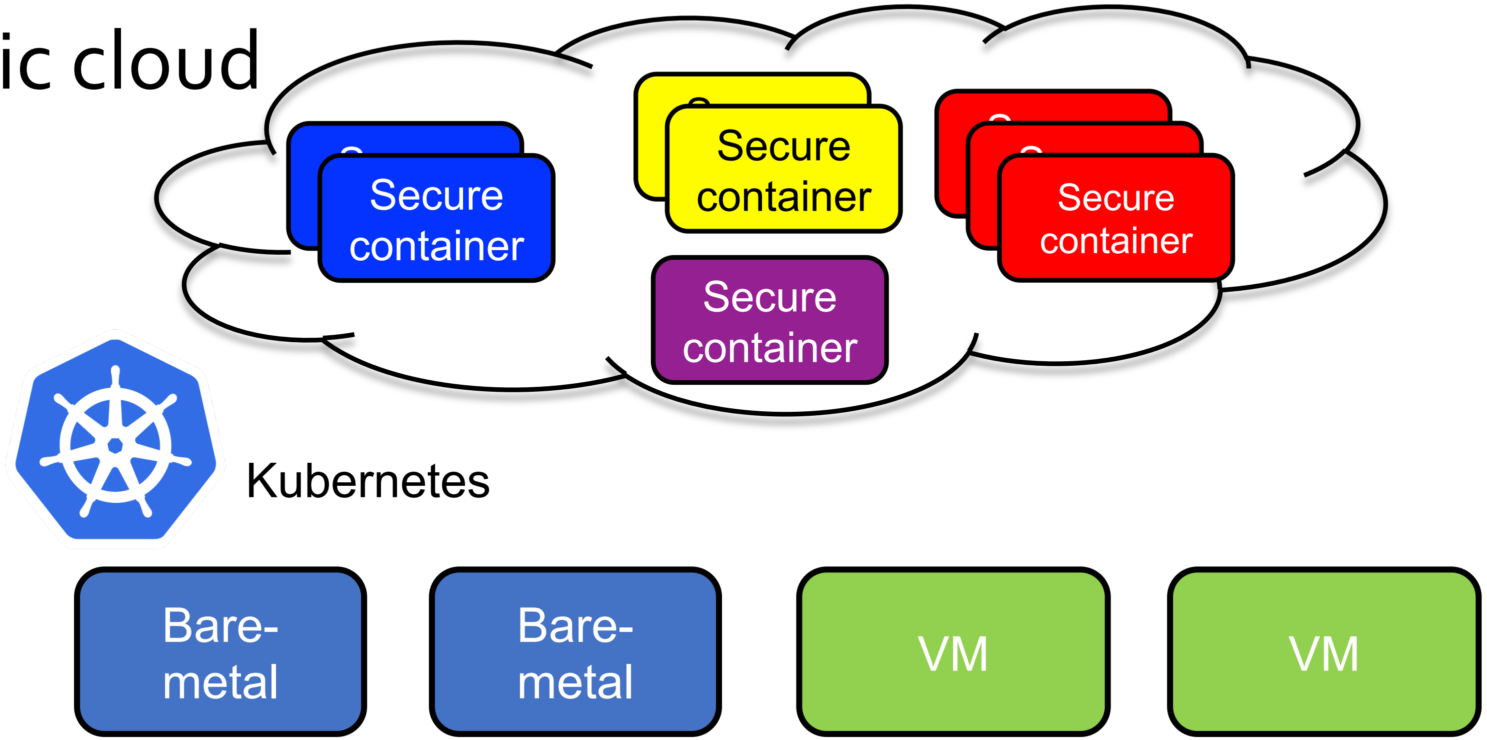
Vulnerability Analysis: Service Mesh



CVEs	Vulnerabilities
CVE-2022-21679	Authorization bypasses
CVE-2021-39156	DoS
CVE-2019-14993	Remote DoS
CVE-2020-10739	Null Pointer Exception
CVE-2019-18836	Remote DoS
CVE-2022-23635	Control plane crashing
CVE-2021-34824	Leaks of credentials
CVE-2020-16844	Bypass the intended policy
CVE-2019-25014	DoS
CVE-2020-8595	Authentication bypass

Another Level of Isolation

- Nested Virtualization
  - Adopted by public cloud service providers



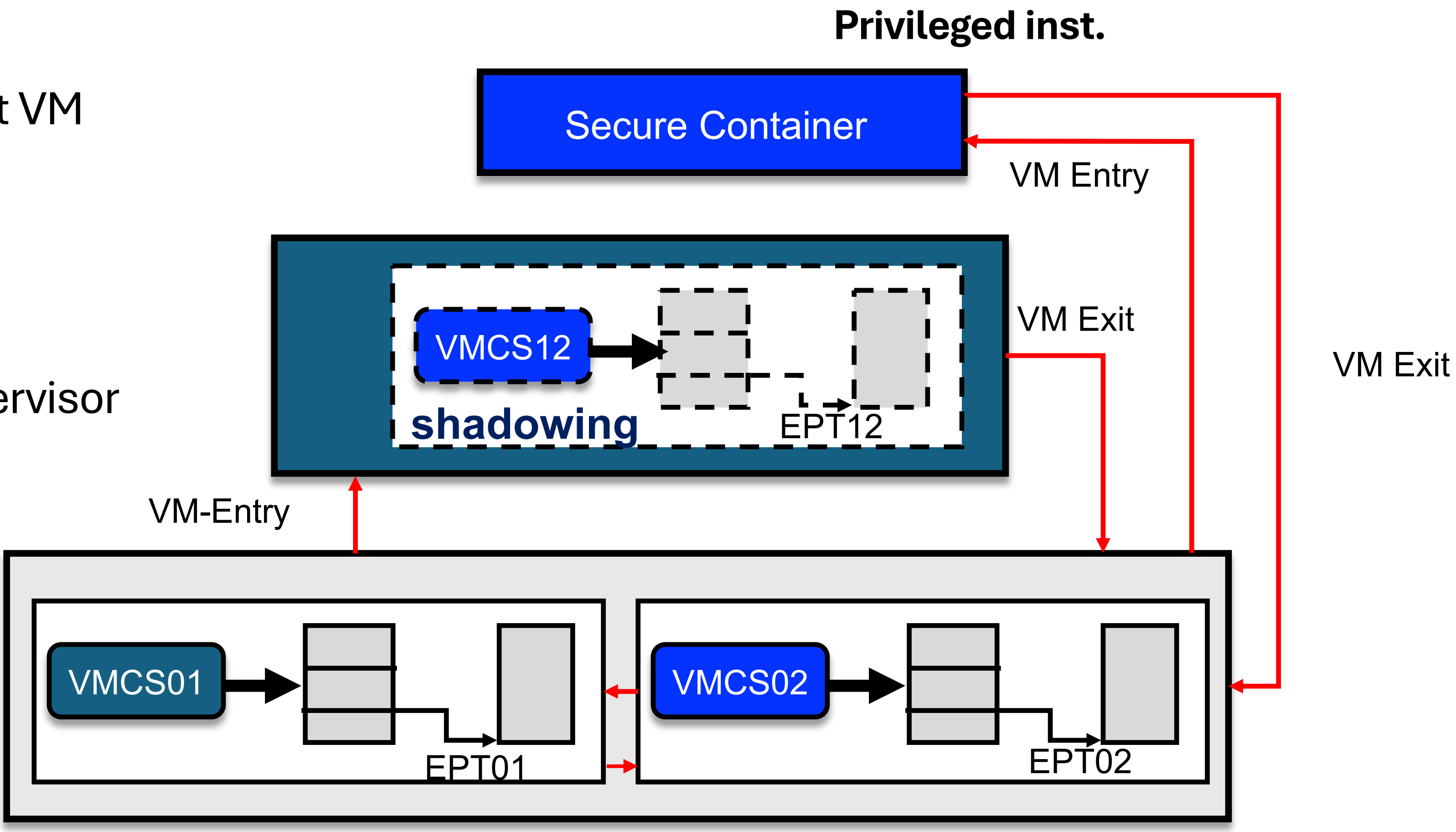
Hosting secure containers orchestration platforms in VMs leased from clouds:
Need nested virtualization!

Another Level of Performance Overhead

L2 (non-root mode): lightweight VM running in an L1 VM

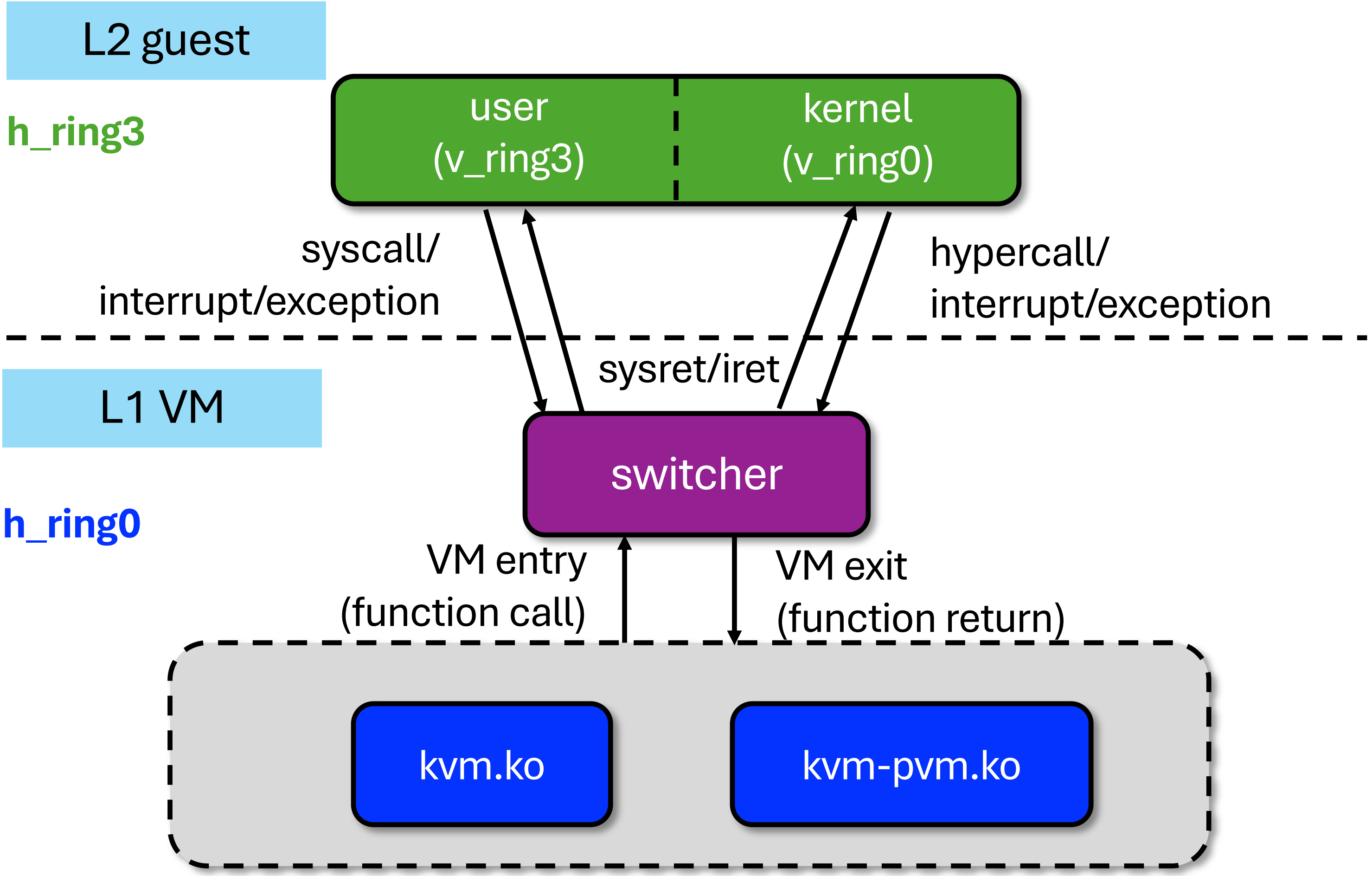
L1 (non-root mode): guest hypervisor running in a cloud VM

L0 (root mode): host hypervisor on a bare-metal machine



High isolation/security overhead

Another Level of Lightweight Containerization!



PVM hypervisor on a cloud VM

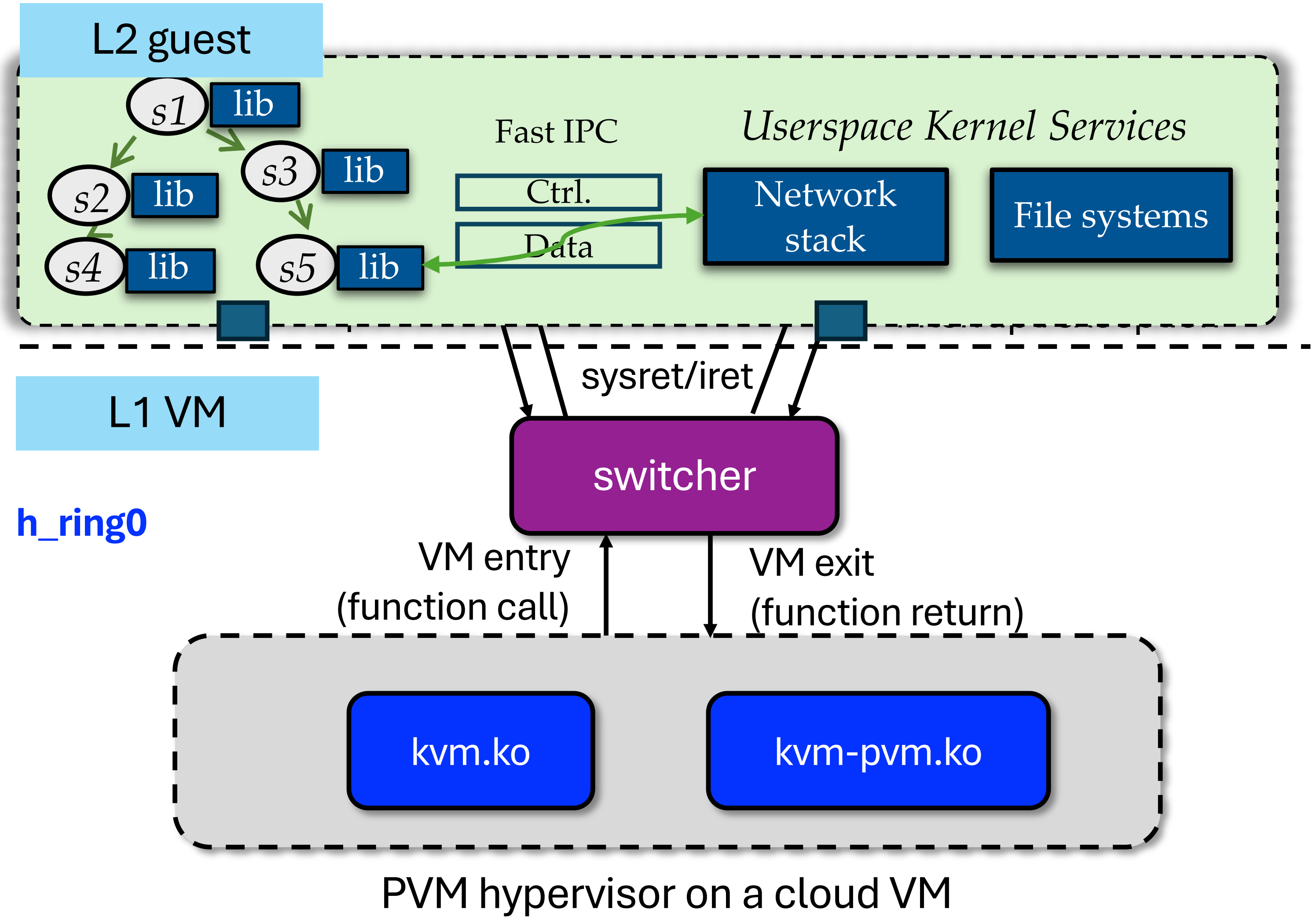
L2 guest runs entirely at **ring3** featuring a **para-virtualized** kernel.

Switcher enables efficient world switches.

PVM hypervisor handles CPU, **memory**, and I/O virtualization.

Hang Huang, Jiangshan Lai, Jia Rao, Hui Lu, Wenlong Hou, Hang Su, Quan Xu, Jiang Zhong, Jiahao Zeng, Xu Wang, Zhengyu He, Weidong Han, Jiang Liu, Tao Ma, and Song Wu. 2023. PVM: Efficient Shadow Paging for Deploying Secure Containers in Cloud-native Environment. In Proceedings of the 29th Symposium on Operating Systems Principles (SOSP '23).

Another Level of Lightweight Containerization!



PVM hypervisor on a cloud VM

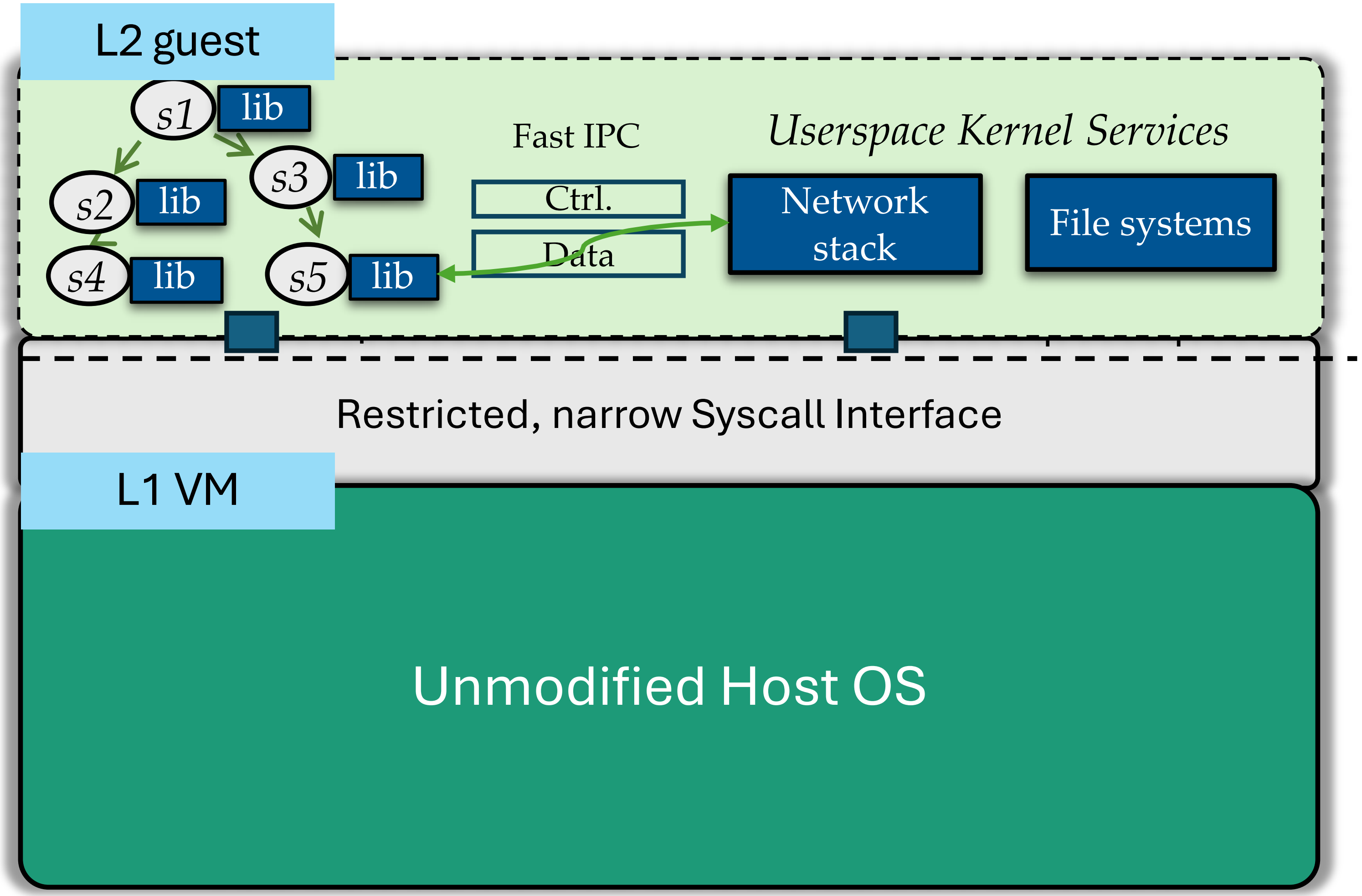
L2 guest runs u-kernel.

Switcher enables efficient world switches.

PVM hypervisor handles CPU, **memory**, and I/O virtualization.

Approved for Public Release; Distribution Unlimited: PA
CLEARANCE PENDING – DO NOT DISTRIBUTE

Another Level of Lightweight Containerization!



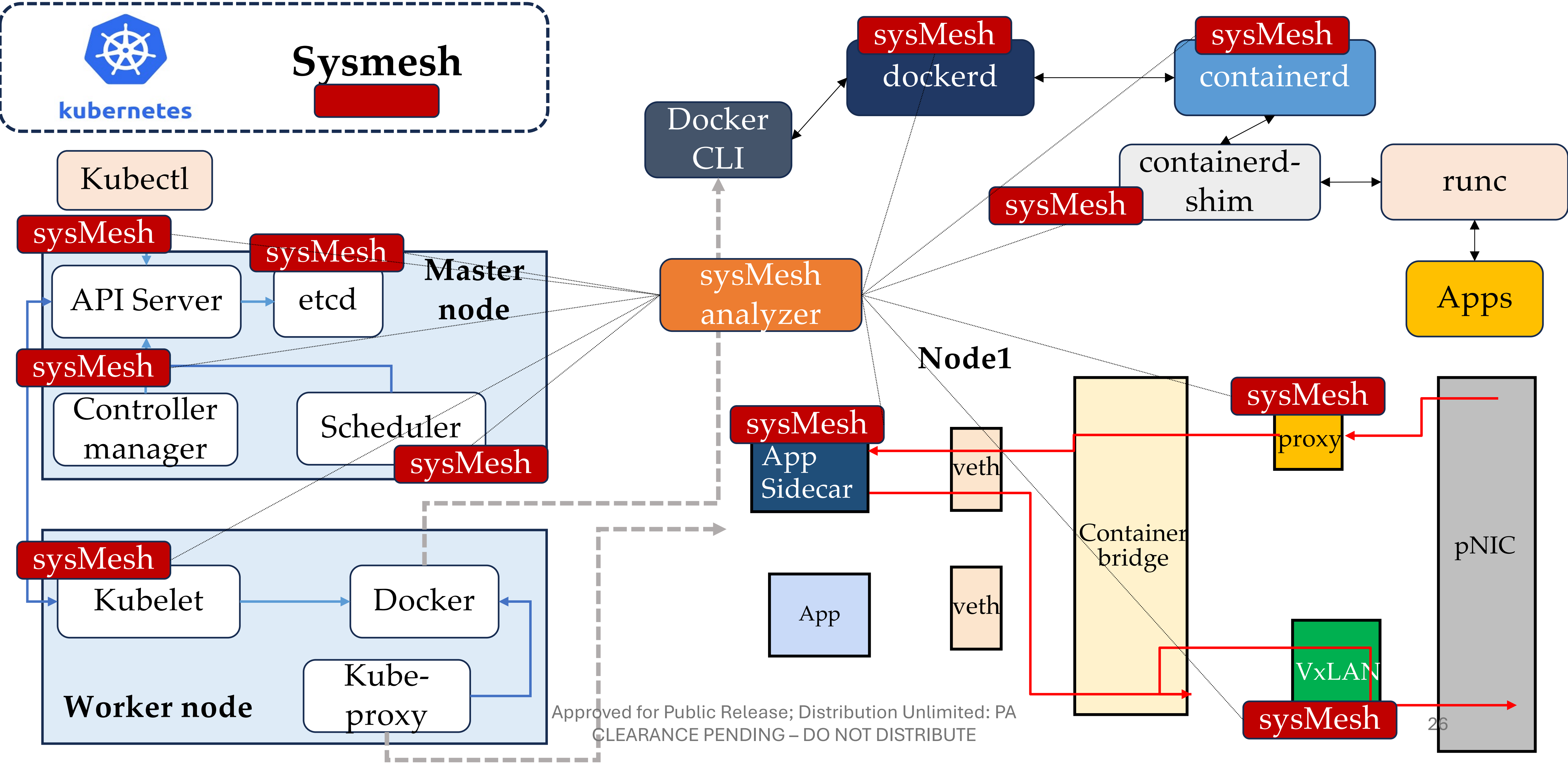
U-kernel on a cloud VM

L2 guest runs u-kernel

L1 runs unmodified guest kernel

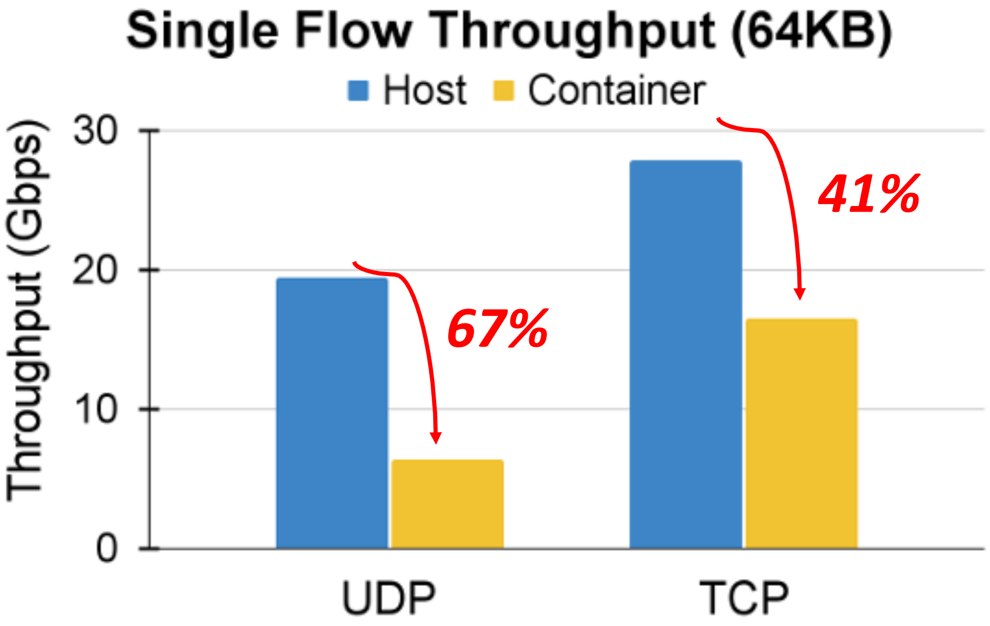
Approved for Public Release; Distribution Unlimited: PA
CLEARANCE PENDING – DO NOT DISTRIBUTE

Platform-Level Security Services

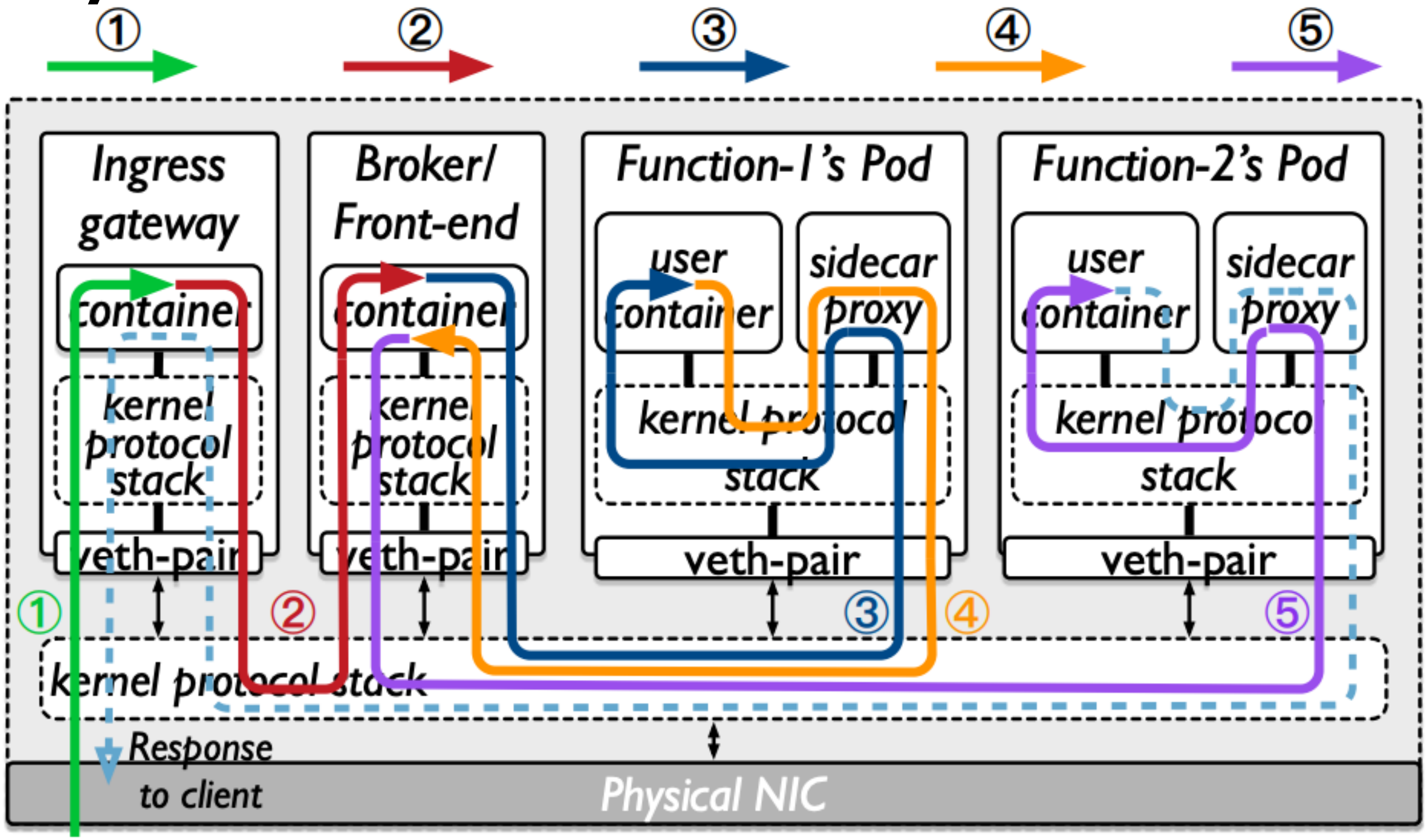


Key Challenge: Prolonged Data Plane

- Excessively, prolonged data plane in orchestration platforms, **due to increasingly added security services**



Performance comparisons between container overlay networks and the native.



Jiaxin Lei, Manish Munikar, Kun Suo, Hui Lu, and Jia Rao. 2021. Parallelizing packet processing in container overlay networks. In Proceedings of the Sixteenth European Conference on Computer Systems (EuroSys '21).

Qi S, Monis L, Zeng Z, Wang IC, Ramakrishnan KK. SPRIGHT: extracting the server from serverless computing! high-performance eBPF-based event-driven, shared-memory processing. ACM SIGCOMM 2022

Conclusions

- While cloud-native technologies offer advanced and effective means to develop and manage today's ubiquitous cloud applications, new security – and performance – challenges also arise

- We need to rethink virtualization techniques for building highly **secure** and **efficient** cloud-native systems. Examples include:
  - Strong-yet-lightweight *isolation* architecture
  - *Fast data planes* tailored for interactive microservices and their containerization and orchestration platforms
  - System-level security support for enhancing security measures and features

Approved for Public Release; Distribution Unlimited: PA
CLEARANCE PENDING – DO NOT DISTRIBUTE

Thank you!

Hui Lu, The University of Texas at Arlington (UTA)

Nathan Daughety, Air Force Research Laboratory (AFRL)