Kubernetes Security with eBPF

Mongkol Thongkraikaew
Head of Platform Engineering
Ascend Money

About Me

Mongkol Thongkraikaew

Head of Platform Engineering

Ascend Money

Open-Source Lovers, Cloud-Native Lovers, Tech Blogger

: Mongkol Thongkraikaew

Medium : @mongkol.ttm

: Mongkol Thongkraikaew

Agenda

- Kubernetes is Just Linux
- Security Practices for Kubernetes
- Overview of eBPF
- Using eBPF to Secure Linux Systems
- Kubernetes Security with eBPF

Game Changer

OpenTelemetry

opentelemetry.io

CLOUD NATIVE
COMPUTING FOUNDATION

Game Changer

OpenTelemetry

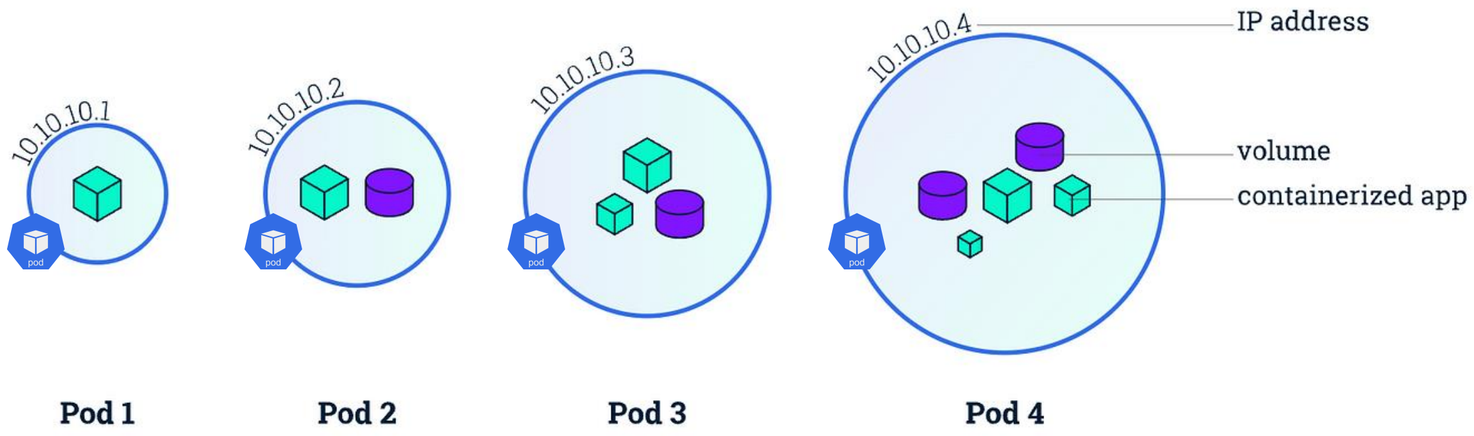
opentelemetry.io

eBPF

ebpf.io

CLOUD NATIVE
COMPUTING FOUNDATION

kubernetes



POD

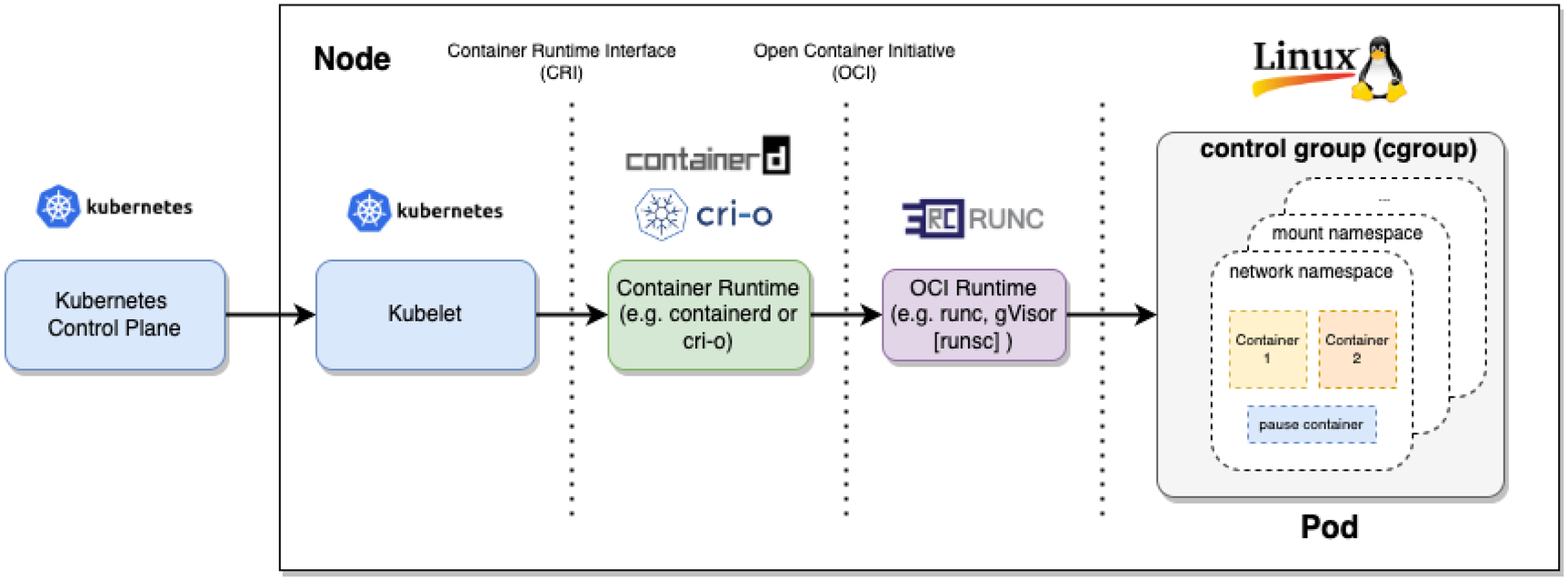
Container Engine

docker

containerd

cri-o

Behind the scenes of Container



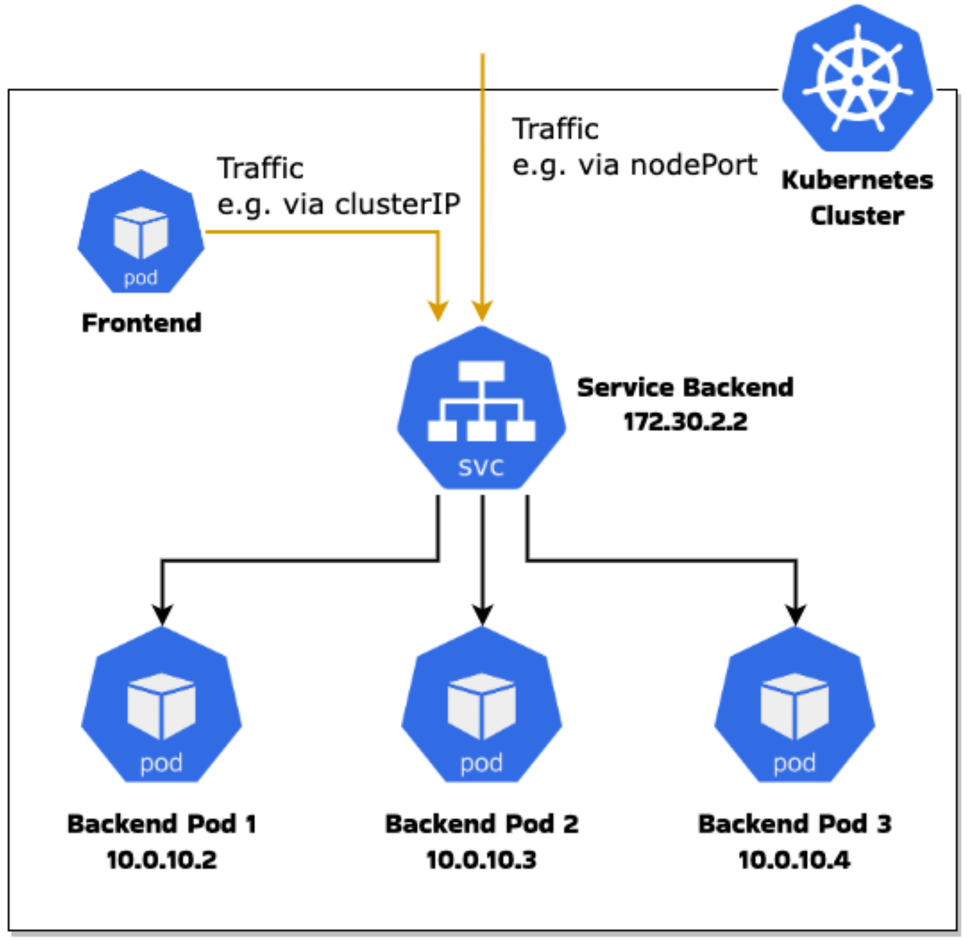
Container built from two main Linux kernels:

Linux Namespaces : The main feature that ensures the process is completely separate from other processes.

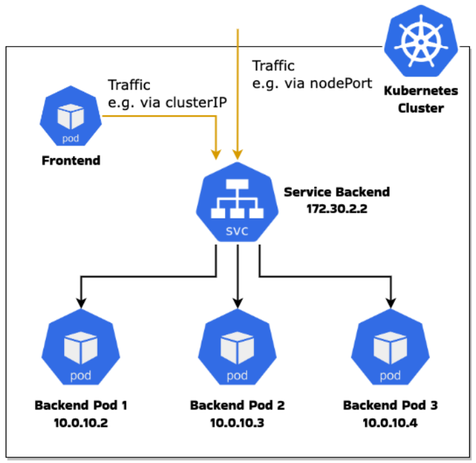
Control Group (Cgroups) : Allows you to limit the system resources (CPU, memory, disk I/O, network, etc) of a process.

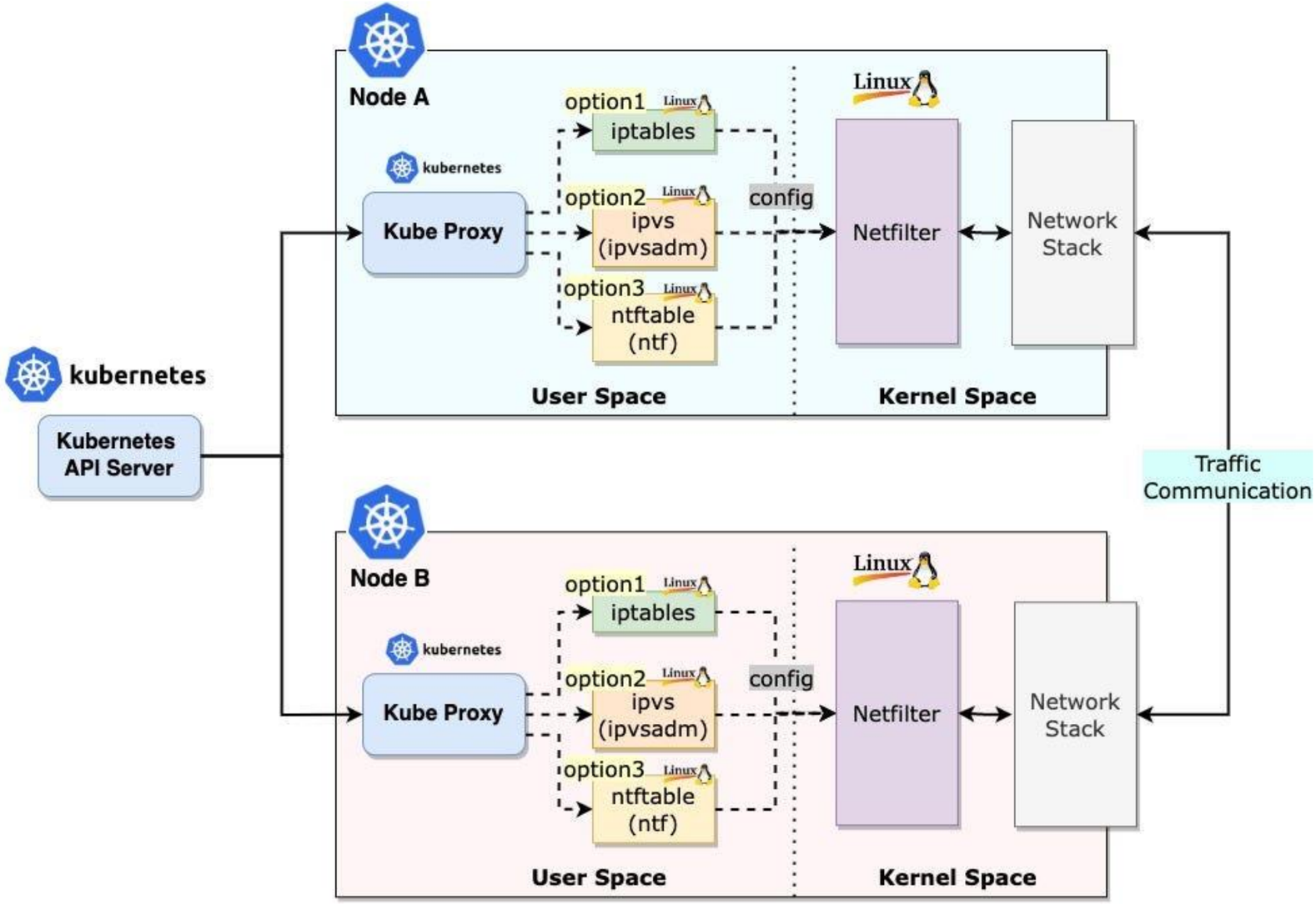
"Container are just collection of processes"

Kubernetes Networking - Service



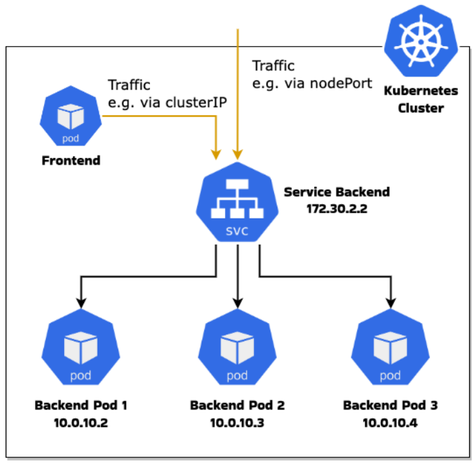
Kubernetes Networking - Service

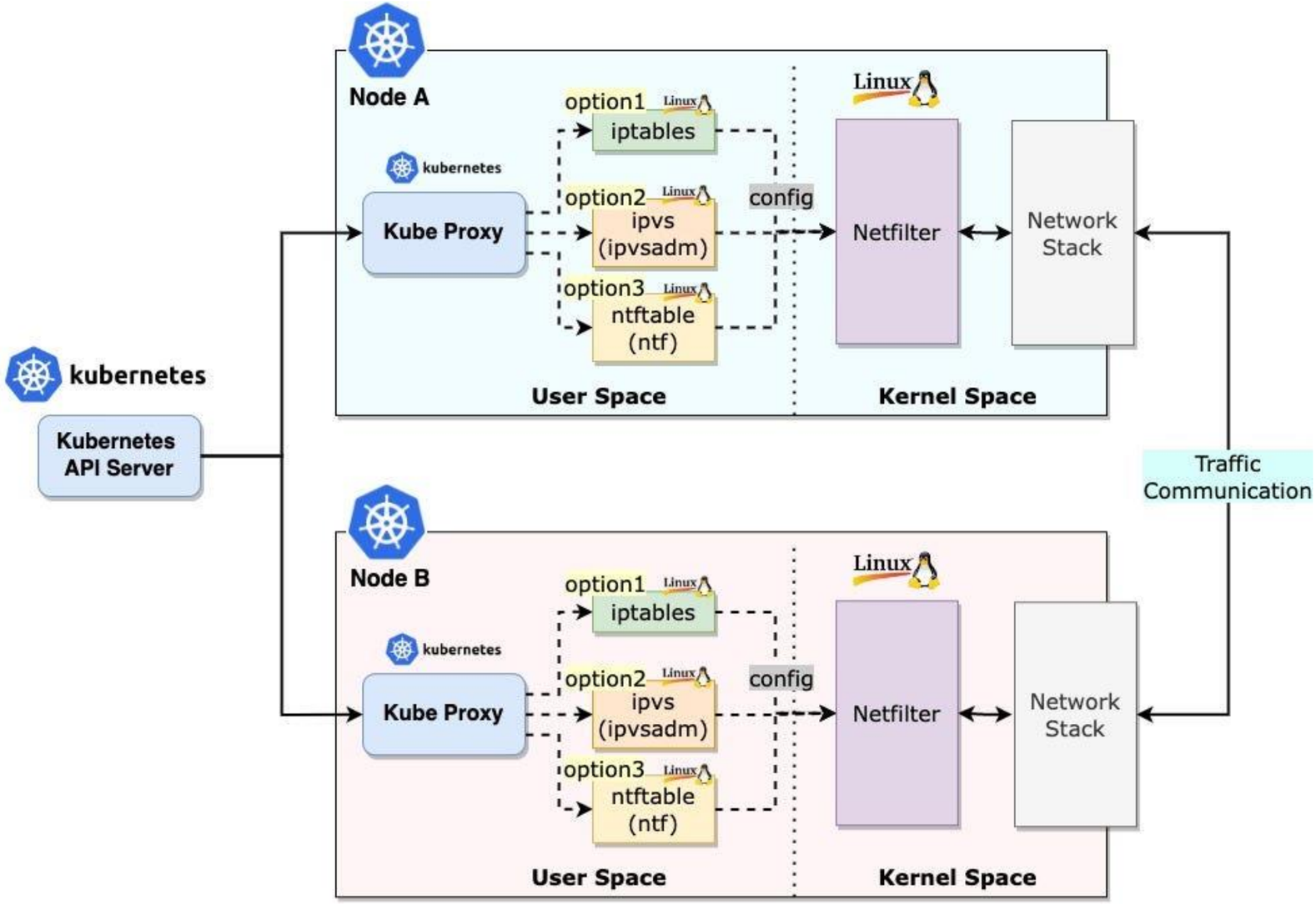




Behind the scenes

Kubernetes Networking - Service





Behind the scenes

Kubernetes capabilities

Resource Limit

cgroup

Network Policy

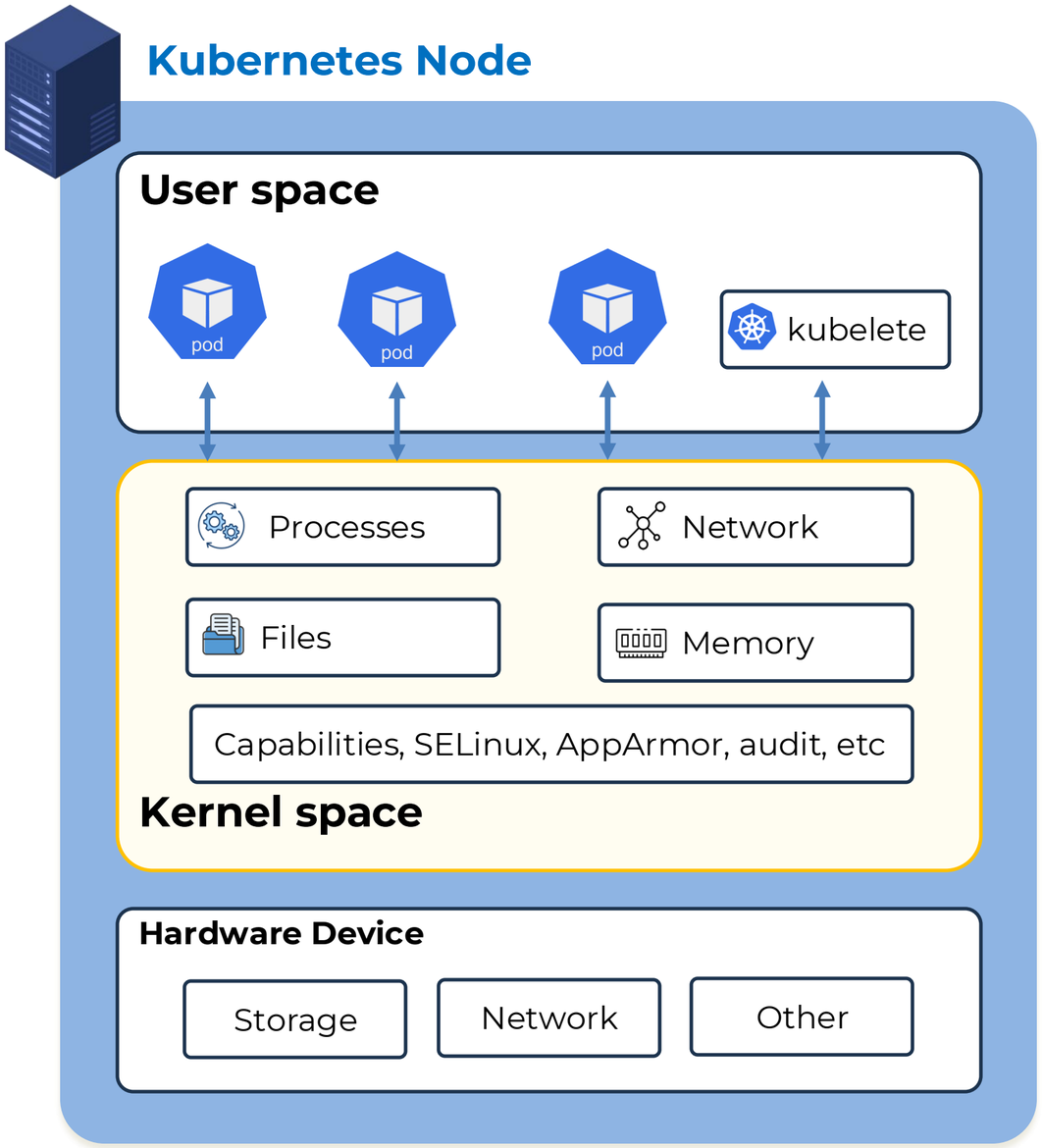
iptable

Persistent Volume

Virtual file system
(VFS)

Kubernetes is essentially a leveraged feature of ...

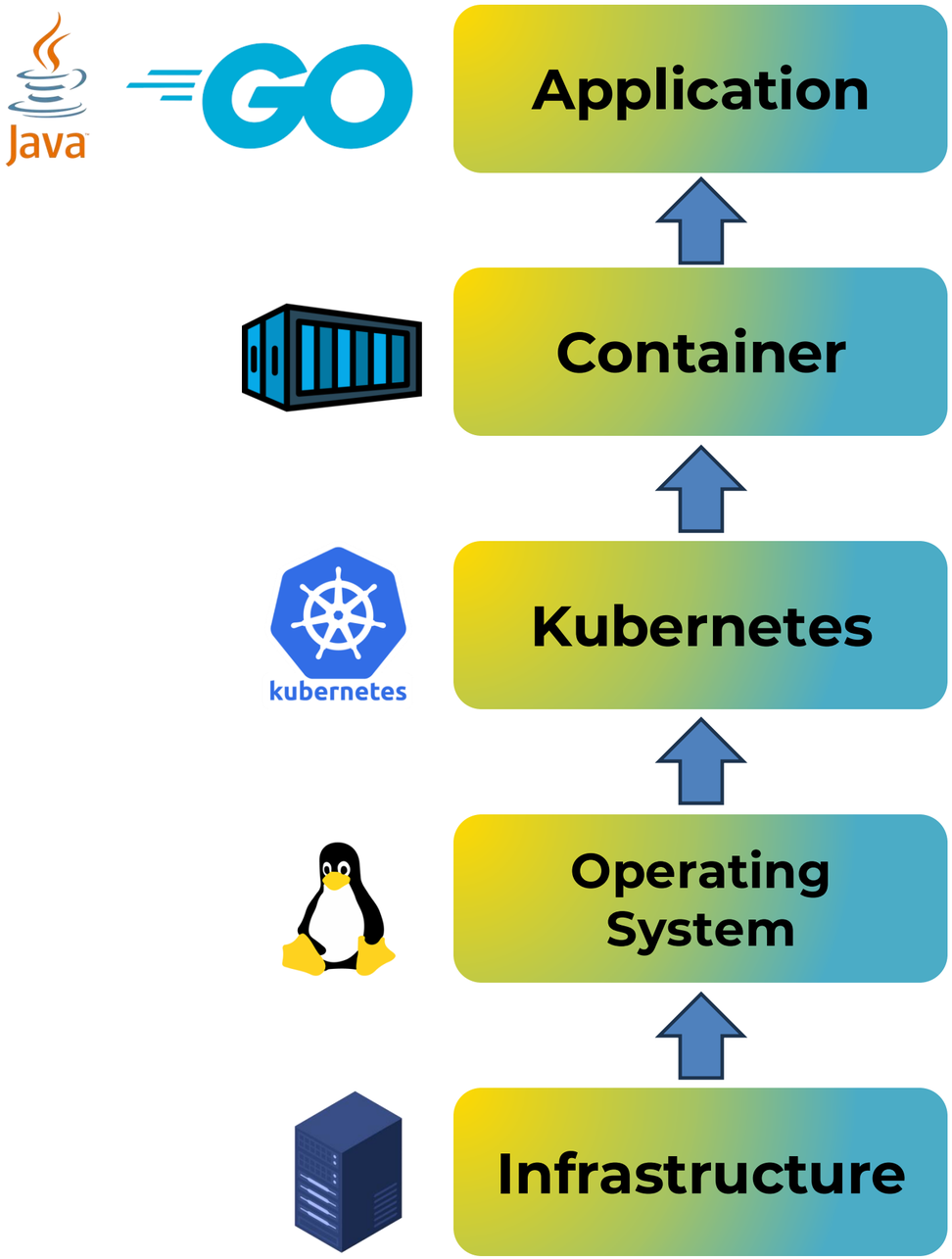
The Linux Kernel

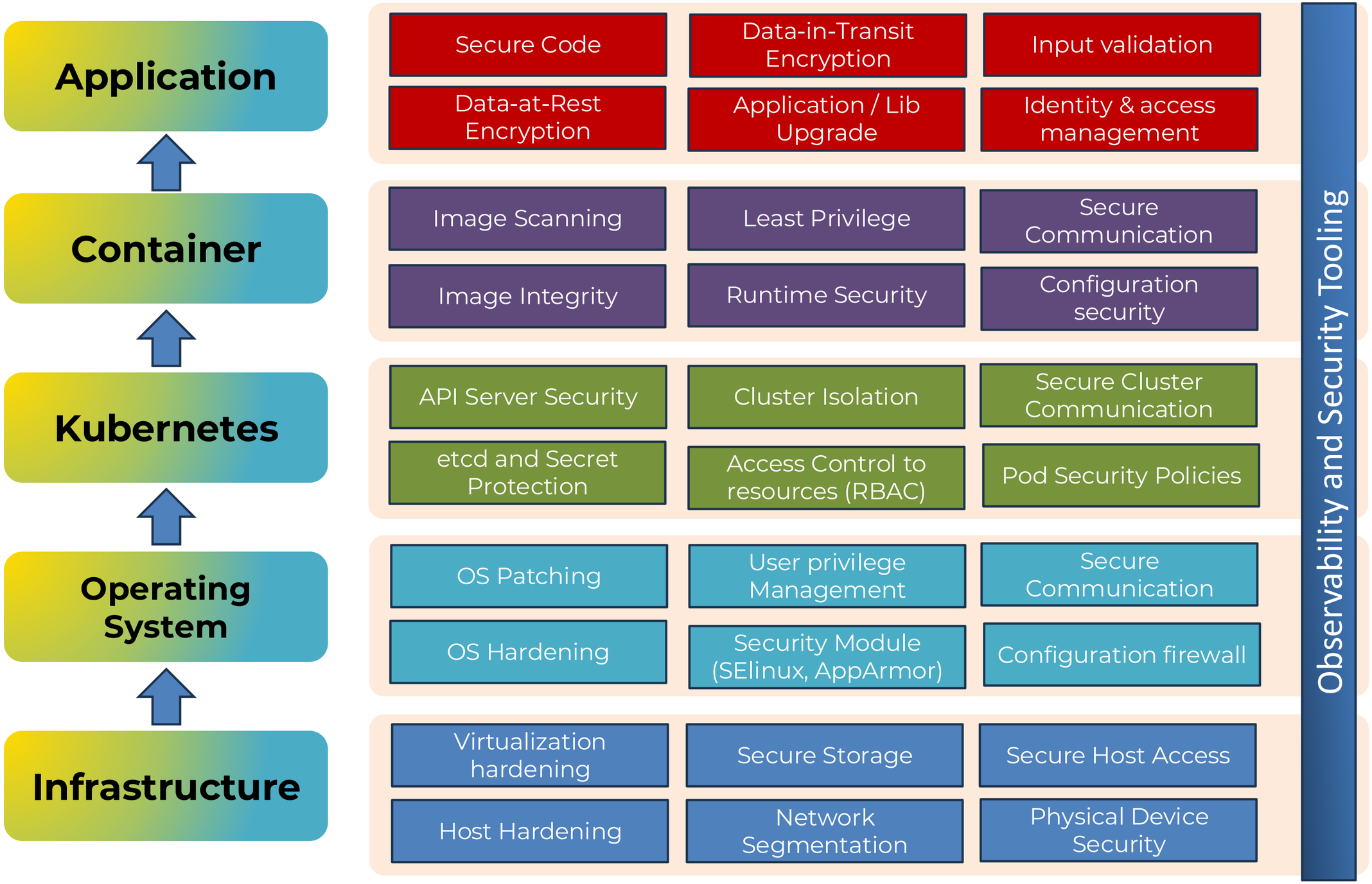


Read More

Medium

Security Practice for Kubernetes


Java
GO
Application
Container
kubernetes
Kubernetes
Operating System
Infrastructure


Application
Secure Code
Data-in-Transit Encryption
Input validation
Data-at-Rest Encryption
Application / Lib Upgrade
Identity & access management
Container
Image Scanning
Least Privilege
Secure Communication
Image Integrity
Runtime Security
Configuration security
Kubernetes
API Server Security
Cluster Isolation
Secure Cluster Communication
etcd and Secret Protection
Access Control to resources (RBAC)
Pod Security Policies
Operating System
OS Patching
User privilege Management
Secure Communication
OS Hardening
Security Module (SElinux, AppArmor)
Configuration firewall
Infrastructure
Virtualization hardening
Secure Storage
Secure Host Access
Host Hardening
Network Segmentation
Physical Device Security
Observability and Security Tooling

Recommend tools for security validation in Kubernetes

CLOUD NATIVE TELECOM INITIATIVE
Cloud Native Telecom Initiative

Tool that validates telco application's adherence to cloud native principles and best practices.

https://github.com/cnti-testcatalog/testsuite

- Category: Security Tests

[Container socket mounts] | [Privileged Containers] | [External IPs] | [SELinux Options] | [Sysctls] | [Privilege escalation] | [Symlink file system] | [Application credentials] | [Host network] | [Service account mapping] | [Ingress and Egress blocked] | [Insecure capabilities] | [Non-root containers] | [Host PID/IPC privileges] | [Linux hardening] | [CPU limits] | [Memory limits] | [Immutable File Systems] | [HostPath Mounts]

aqua kube-bench

kube-bench is a tool that checks whether Kubernetes is deployed securely by running the checks documented in the CIS Kubernetes Benchmark.

https://github.com/aquasecurity/kube-bench

Kubescape

Kubescape is an open-source Kubernetes security platform that ensures comprehensive security throughout the development and deployment lifecycle, offering hardening, posture management, and runtime security for robust protection of Kubernetes environments.

https://github.com/kubescape/kubescape

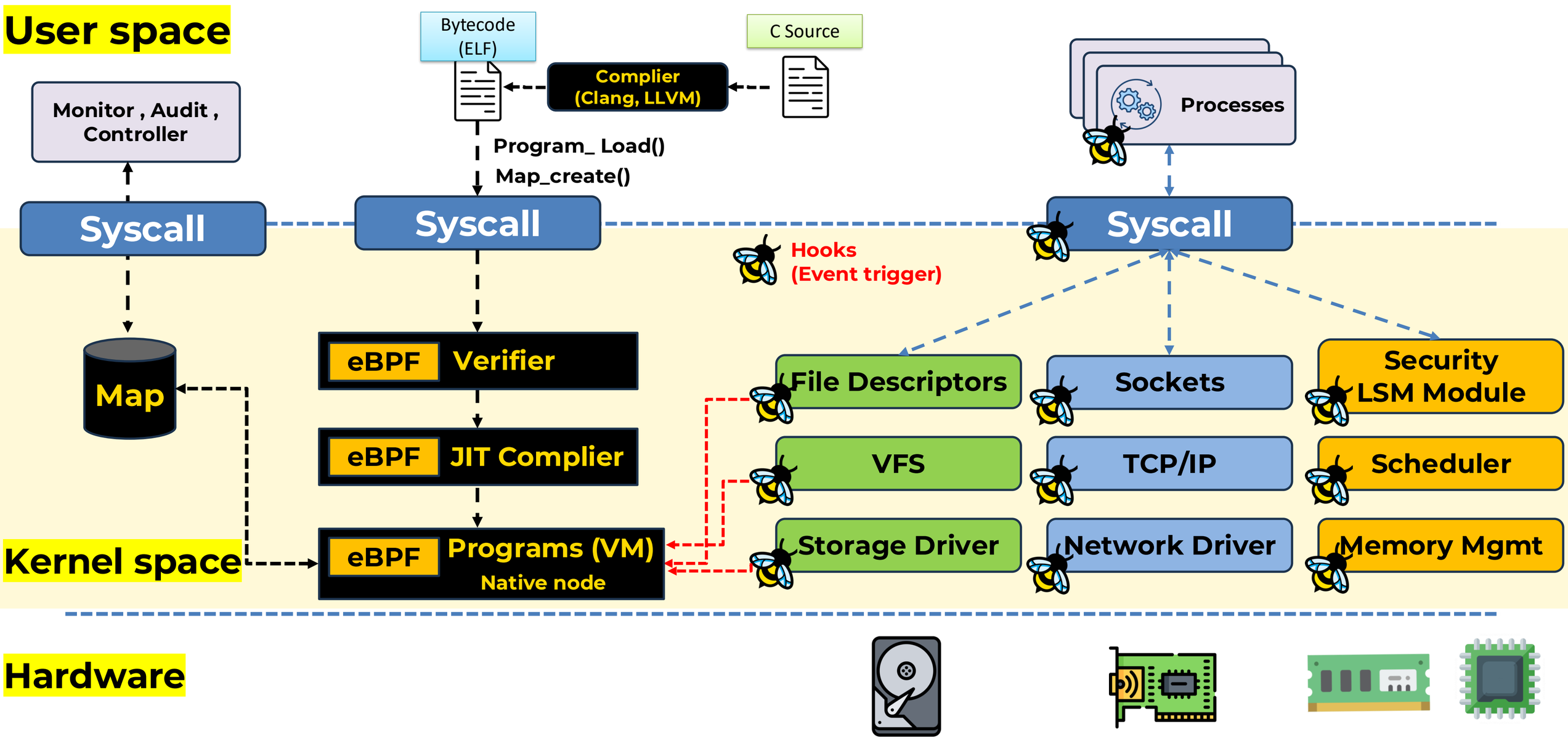
Overview of eBPF

Accurate Definition

eBPF is a programming language & runtime to extend Operating Systems

eBPF refers to extended BPF
BPF - Berkeley Packet Filtering

eBPF Diagram and Capability



Popular Attach point or Event Hook of eBPF

eBPF **Programs (VM)**
Native node

Kprobe

Attach a probe to a specific kernel function. Enabling you to inspect or modify the behavior of the kernel. Ex. **do_sys_open** hooked when open file

Uprobe

Similar to kprobes, but for user-space applications.
Ex. **free ,** function for monitor memory deallocation. identify memory leaks.

Tracepoint

Way to gather information about various kernel events, such as scheduling or network packet processing.

Perf Event

Used for **performance monitoring**. Track CPU cycles, cache misses, or other performance metrics, Allowing you to analyze the performance of your applications or the kernel itself.

XDP
(eXpress Data Path)

XDP is a high-performance packet processing to run eBPF programs in the earliest point of network stack or at network interface card level
Capability : Packet filter, Packet modify, Load balancing, Traffic Monitoring

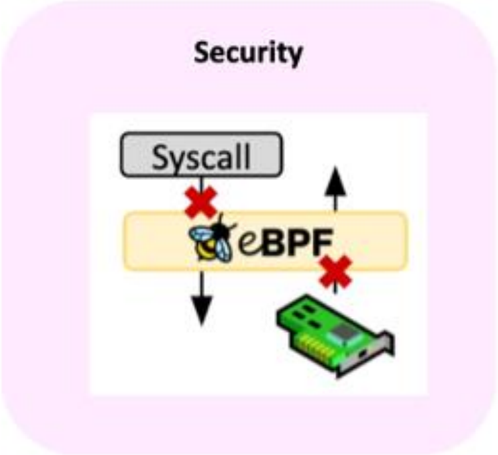
Socket Filter

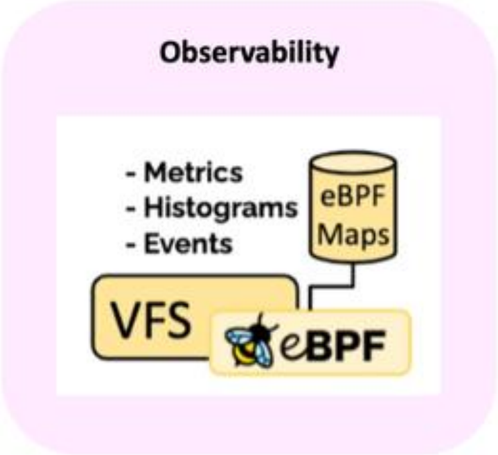
To filter incoming and outgoing packets at the socket level.
- security policies, monitoring traffic, optimizing application performance by reducing unnecessary packet processing.

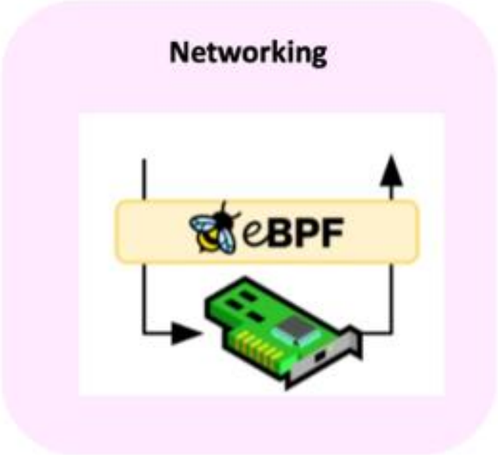
LSM
(Linux Security Module)

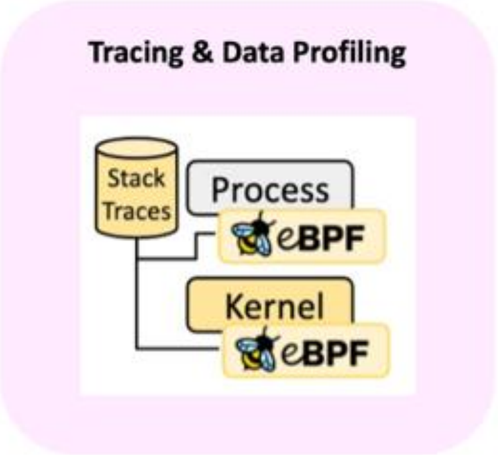
Implementation of various security policies and mechanisms involves attaching programs to "**LSM hooks**", allowing for the creation of custom security checks and controls at multiple points within the kernel.

eBPF Use Cases









Networking activities at the packet and socket levels. And enforce security policies by blocking or modifying certain system events based on predefined rules.

Trace network packets, system calls, function calls, and other events in real-time, enabling deep visibility into system behavior.

Firewalls, load balancers, and network monitoring tools. The low overhead and dynamic loading capabilities of eBPF

Profiling and tracing the runtime behavior of systems and applications

Using eBPF to Secure Linux Systems

Visibility and monitoring
Detect unusual behavior, identify potential threats, and analyze the attack surface.

Network Security
implementation of custom security traffic policies.

Container Security
Monitor and enforce security policies within containerized by attaching programs to **cgroup** hooks.

Intrusion Detection & Prevention
Monitor file system activity, restricting unauthorized access or modification of sensitive data.

Policy enforcement
Enforce access controls, regulate resource usage, and protect sensitive system components

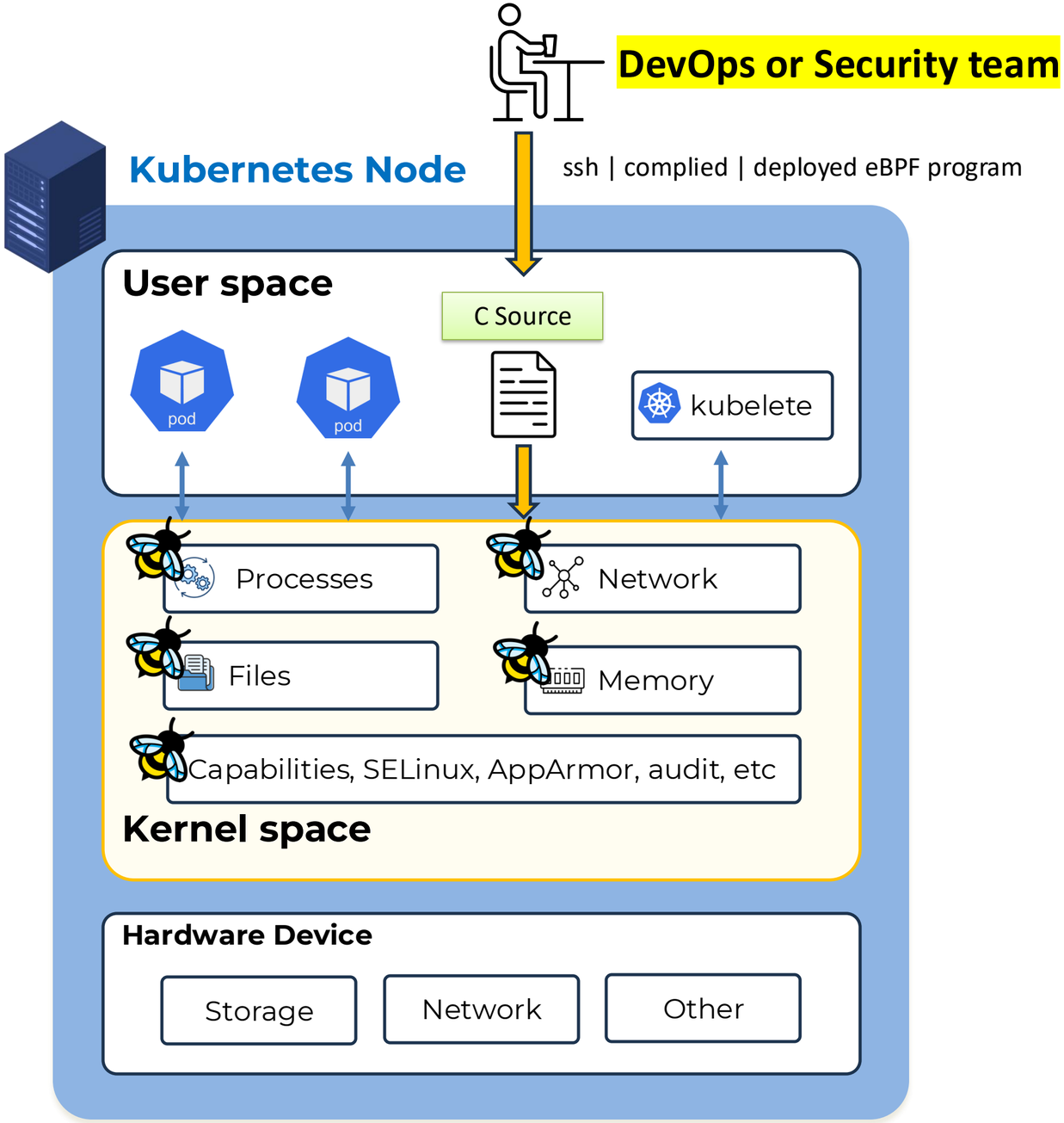
Incident response & forensics
Collect detailed information about system activity and state during an incident

Kubernetes Security with eBPF

Security capabilities related to eBPF in kubernetes

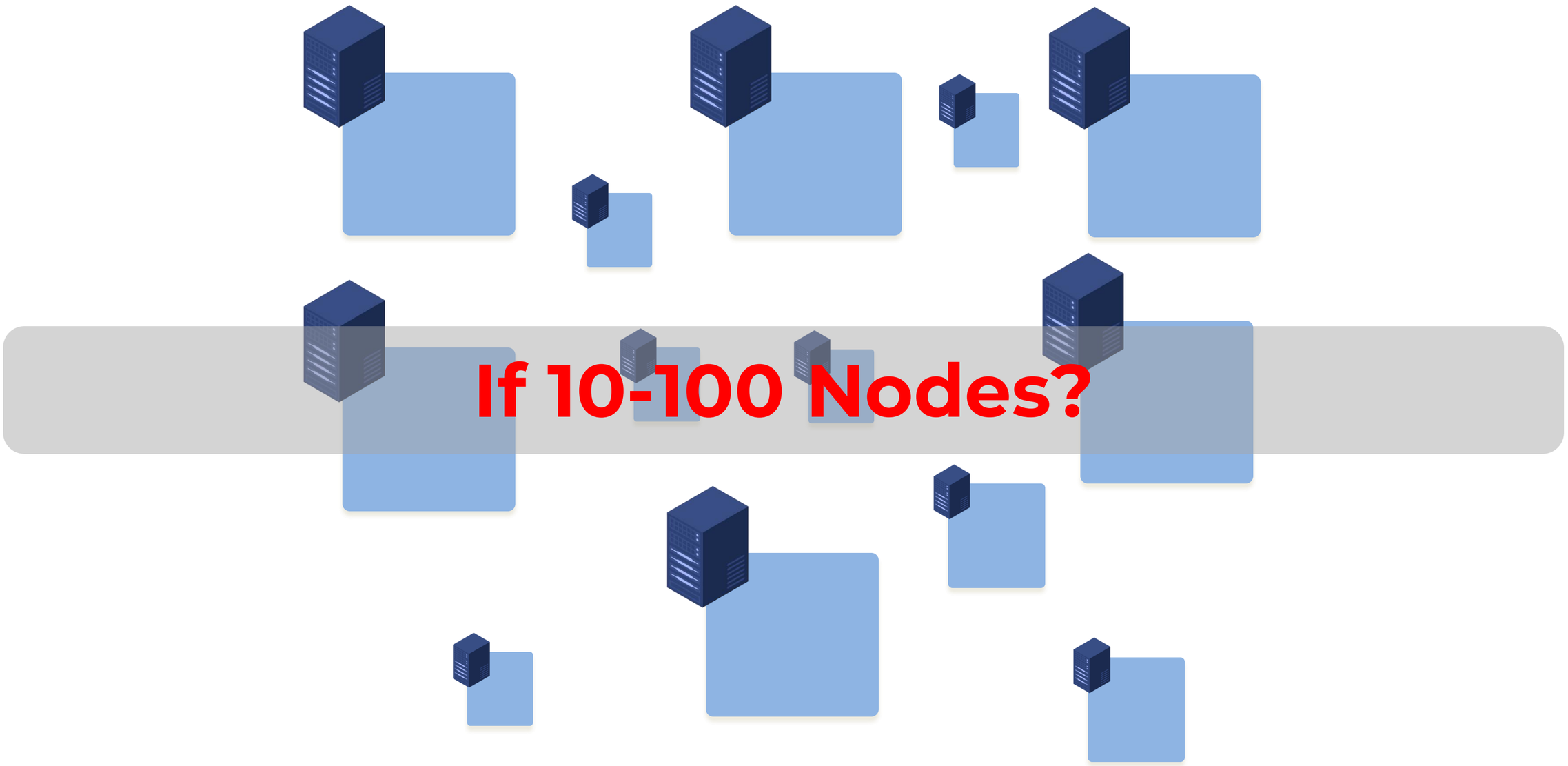
- Network Policy Enforcement
- Security Observability
- Runtime Security Monitoring
- Intrusion Detection Systems (IDP)
- Cgroup & Namespace Isolation
- Policy Enforcement
- Service Mesh

How to working with eBPF in Kubernetes

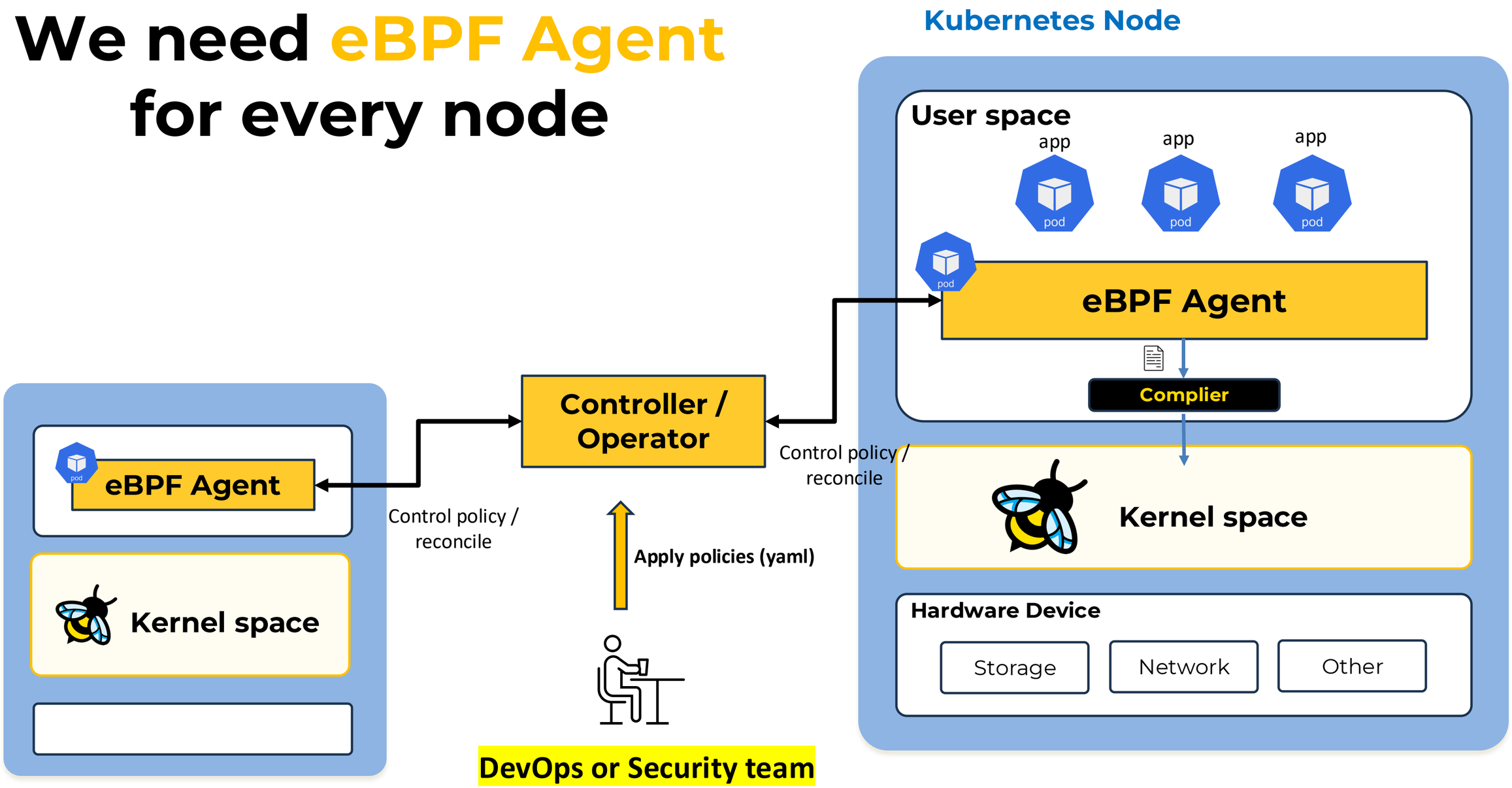


Just 1 Node

Kubernetes Node



We need eBPF Agent for every node



Most Popular Tools for Security

★ 19.6 k

cilium
Service Mesh

cilium
Network Encryption (IPSec)

★ 5.9 k

PROJECT CALICO

tetragon
★ 3.5 k

kubearmor
★ 1.3 k

falco
★ 7.2 k

aqua tracee
★ 3.5 k

Network Policy, Network Security, Service Mesh

Runtime Security, Runtime Enforcement, Real-time threat detection,

Security Observability

ebpf.io/applications

tetragon

Use Cases

Process lifecycle
Tetragon observes by default the process lifecycle via exec and exit

Filename access
Monitor filename access using kprobe hooks

Network observability
Monitor TCP connect using kprobe hooks

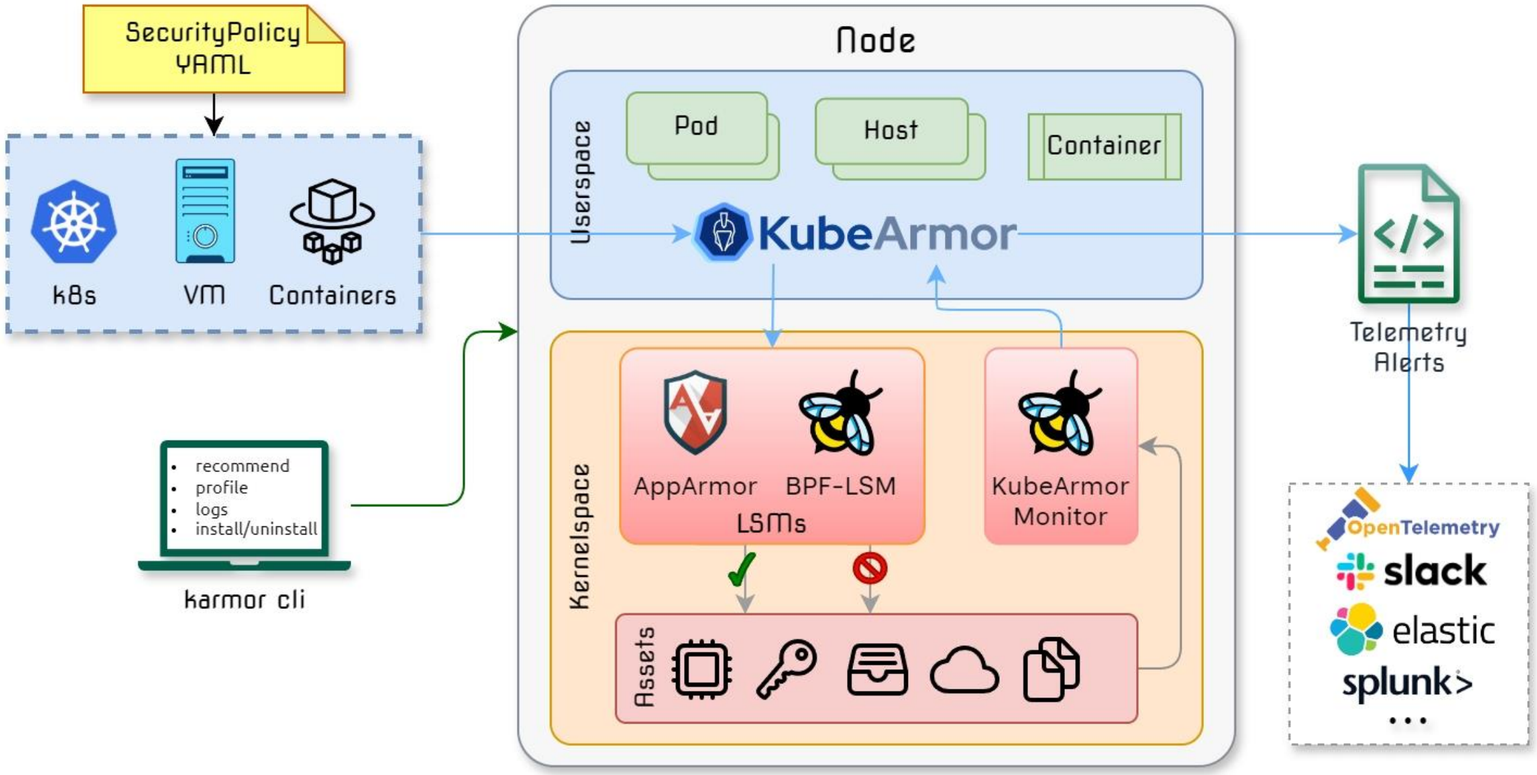
Linux process credentials
Monitor Linux process credentials

Host System Changes
Monitor Host System changes

Security Profiles
Observe and record security events

tetragon.io

kubearmor



Use Cases

Harden Infrastructure

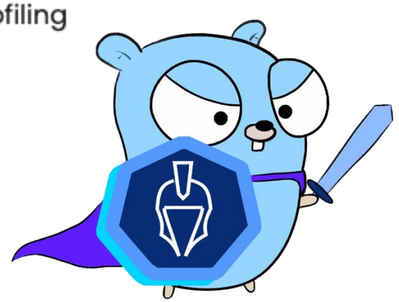
- Protect critical paths such as cert bundles
- MITRE, STIGs, CIS based rules
- Restrict access to raw DB table

Least Permissive Access

- Process Whitelisting
- Network Whitelisting
- Control access to sensitive assets

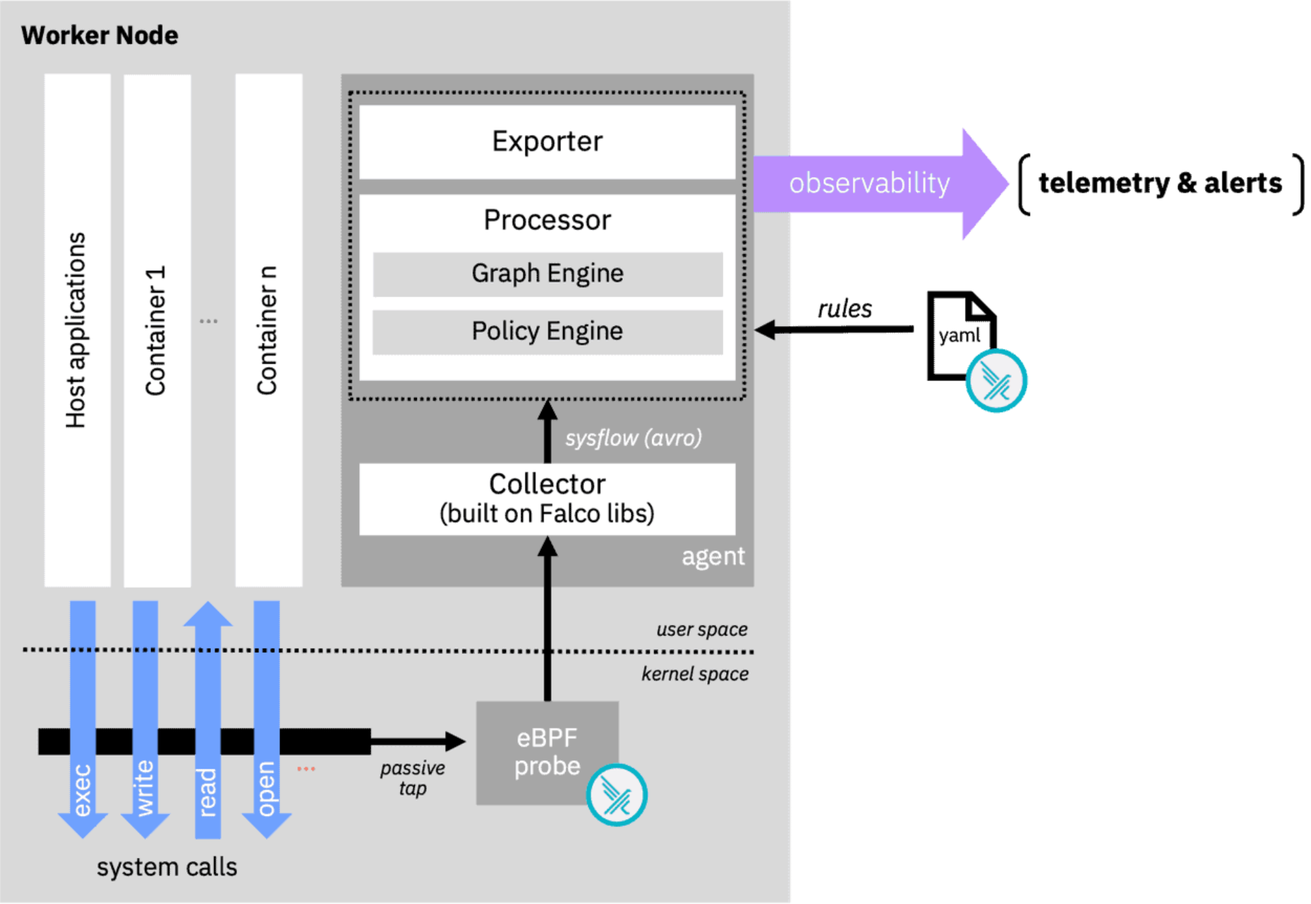
Application Behavior

- Process execs, File System accesses
- Service binds, Ingress, Egress connections
- Sensitive system call profiling



kubearmor.io

falco



Network Connection outside Local Subnet	incubating	WARNING	enabled
Mount Launched in Privileged Container	incubating	WARNING	enabled
Launch Ingress Remote File Copy Tools in Container	incubating	NOTICE	enabled
Read environment variable from /proc files	incubating	WARNING	enabled
Exfiltrating Artifacts via Kubernetes Control Plane	incubating	NOTICE	enabled
Adding ssh keys to authorized_keys	incubating	WARNING	enabled
Potential Local Privilege Escalation via Environment Variables Misuse	incubating	NOTICE	enabled
Directory traversal monitored file read	stable	WARNING	enabled
Read sensitive file trusted after startup	stable	WARNING	enabled
Read sensitive file untrusted	stable	WARNING	enabled
Run shell untrusted	stable	NOTICE	enabled
System user interactive	stable	INFO	enabled
Terminal shell in container	stable	NOTICE	enabled
Contact K8S API Server From Container	stable	NOTICE	enabled
Netcat Remote Code Execution in Container	stable	WARNING	enabled
Search Private Keys or Passwords	stable	WARNING	enabled
Clear Log Activities	stable	WARNING	enabled

Example Falco Default Rules (more than 90 rules)

falco.org

Bonus

Bonus

Interesting eBPF project

bpftrace
High-level tracing language for Linux eBPF

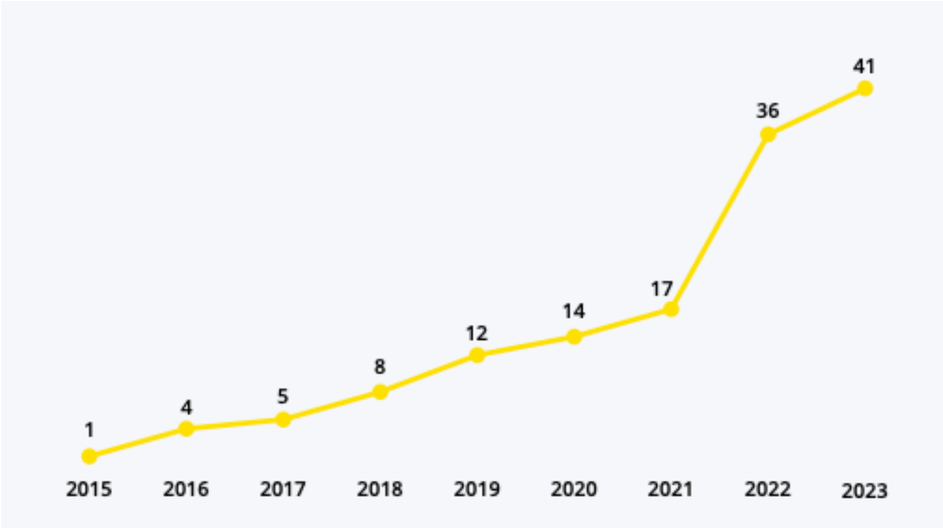
bpftop
Real-time eBPF Program Monitoring and Performance Statistics

Pixie
Scriptable observability for Kubernetes

Hubble
Network, Service & Security Observability for Kubernetes using eBPF

kubectl trace
Schedule bpftrace programs on your Kubernetes cluster

Caretta
eBPF based Kubernetes service map



Active eBPF Landscape project by year
Source: github

Bonus

Interesting eBPF project

bpftrace
High-level tracing language for Linux eBPF

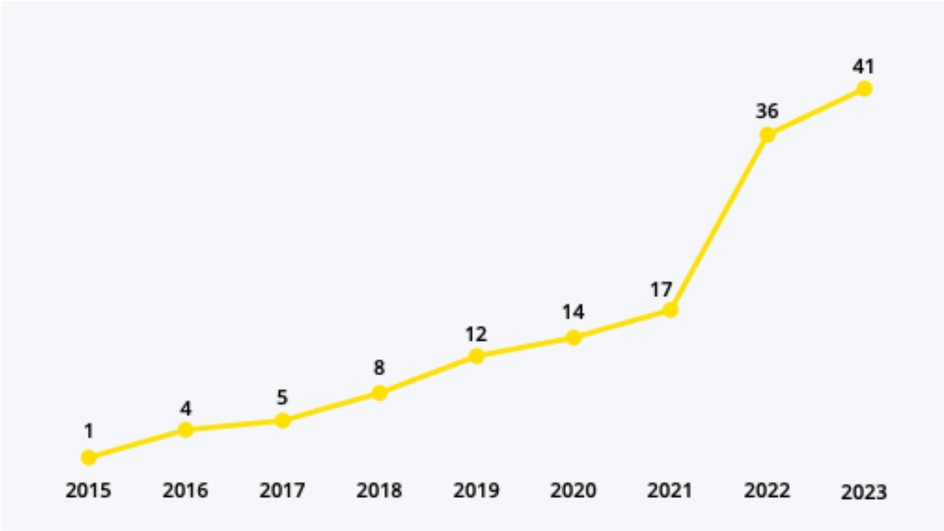
bpftop
Real-time eBPF Program Monitoring and Performance Statistics

Pixie
Scriptable observability for Kubernetes

Hubble
Network, Service & Security Observability for Kubernetes using eBPF

kubectl trace
Schedule bpftrace programs on your Kubernetes cluster

Caretta
eBPF based Kubernetes service map



Active eBPF Landscape project by year
Source: github

eBPF Case Studies

Cloudflare uses eBPF for network security, performance monitoring, and network observability

Meta uses eBPF to process and load balance every packet coming into their data centers

Red Hat uses eBPF at scale for load balancing and tracing in their private cloud

Google uses eBPF for security auditing, packet processing, and performance monitoring

Netflix uses eBPF at scale for network insights

Alibaba uses eBPF through Cilium to provide networking in their cloud

Datadog uses eBPF for networking and security in their SaaS product

Line Corporation uses eBPF at scale for load balancing and tracing in their private cloud

ebpf.io/case-studies

eBPF

“eBPF will be for everyone, but we don’t expect everyone to know about it”

Thank You

Mongkol Thongkraikaew

: Mongkol Thongkraikaew

Medium : @mongkol.ttm

: Mongkol Thongkraikaew