

ISOVALENT

# Cilium – Cloud Native Networking & Security

Cloud Native Days Stockholm



Thomas Graf

CTO & Co-Founde, Isovalent



# cilium

Created by ISOVALENT

 eBPF-based:































































- Networking
- Security
- Observability
- Service Mesh & Ingress

Foundation



Technology



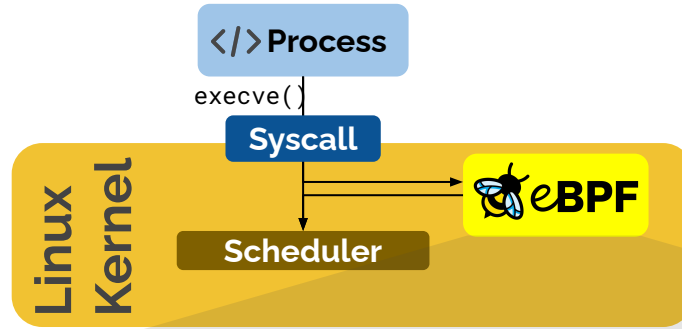
 <p>What Makes a Good Multi-Tenant Kubernetes Solution</p> <p>VIDEO 1 VIDEO 2</p>	 <p>Building High-Performance Cloud Native Pod Networks</p> <p>READ BLOG</p>	 <p>AWS picks Cilium for Networking &amp; Security on EKS Anywhere</p> <p>READ BLOG</p>	 <p>Bell uses Cilium and eBPF for telco networking</p> <p>VIDEO 1 VIDEO 2</p>	 <p>AccuKnox uses Cilium for network visibility and network policy enforcement</p>	 <p>ACOSS uses Cilium as their main CNI plugin for self-hosted Kubernetes</p>	 <p>ArangoDB Datalite uses Cilium to separate database deployments in a multi-tenant cloud environment</p>	 <p>Ayedo builds and operates cloud native platforms using Cilium</p>
 <p>Building a Secure and Maintainable Pass</p> <p>WATCH VIDEO</p>	 <p>Cloud Native Networking with eBPF</p> <p>WATCH VIDEO</p>	 <p>Datadog is using Cilium in AWS (self-hosted k8s)</p> <p>WATCH VIDEO</p>	 <p>Managed Kubernetes: 1.5 Years of Cilium Usage at DigitalOcean</p> <p>WATCH VIDEO</p>	 <p>ByteDance uses Cilium as their CNI for self-hosted Kubernetes clusters</p>	 <p>Canonical's Kubernetes distribution microK8s uses Cilium as CNI plugin</p>	 <p>Civo is offering Cilium as the CNI option for Civo users to choose it for their Civo Kubernetes clusters</p>	 <p>Cognite uses Cilium as the CNI plugin for industrial DataOps</p>
 <p>Kubernetes uses Cilium as their CNI and for load balancing</p> <p>READ BLOG</p>	 <p>Kubernetes Network Policies in Action with Cilium</p> <p>VIDEO</p>	 <p>Google chooses Cilium for Google Kubernetes Engine (GKE) networking</p> <p>READ BLOG</p>	 <p>IKEA uses Cilium for their self-hosted bare-metal private cloud</p> <p>WATCH VIDEO</p>	 <p>Elastic Path uses Cilium in their production CloudOps Kubernetes clusters</p>	 <p>FS uses Cilium VXLAN tunnel integration with BIG-IP</p>	 <p>finleap connect uses Cilium on a bare-metal/private cloud</p>	 <p>Form3 is using Cilium in their production clusters (self-hosted, bare-metal, private cloud)</p>
 <p>Scaling a Multi-Tenant Kubernetes Clusters in a Telco</p> <p>WATCH VIDEO</p>	 <p>Meltwater is using Cilium in AWS in self-hosted multi-tenant k8s clusters as the CNI plugin</p> <p>WATCH VIDEO</p>	 <p>Mobilabs uses Cilium as the CNI for their internal cloud</p> <p>READ BLOG</p>	 <p>Nexiot using Cilium as the CNI for its IoT SaaS</p> <p>READ USER STORY</p>	 <p>Infomaniak uses Cilium in self-hosted clusters on bare-metal and Openstack</p>	 <p>innoq uses Cilium to run their customer's infrastructure</p>	 <p>Cilium is the platform that powers Isovalent's enterprise networking, observability, and security solutions</p>	 <p>JUMO uses Cilium as the CNI plugin for all of their AWS-hosted EKS clusters</p>
 <p>PostFinance is using Cilium as their CNI for admission-critical, on-premise k8s clusters</p> <p>READ CASE STUDY</p>	 <p>eBPF &amp; Cilium at Sky</p> <p>WATCH VIDEO</p>	 <p>Skybet uses Cilium as their CNI</p> <p>READ BLOG</p>	 <p>Trip.com uses Cilium both on-premise and in AWS</p> <p>BLOG 1 BLOG 2</p>	 <p>Kryptos uses Cilium as the CNI for their on-prem Kubernetes clusters</p>	 <p>Kube-QVN uses Cilium to enhance the CNI service performance, security and monitoring</p>	 <p>Kubermatic uses Cilium as the CNI for its Kubernetes installer and platform</p>	 <p>KubeKey is an open-source lightweight tool for deploying Kubernetes clusters and addons</p>
 <p>Northflank uses Cilium as its CNI plugin across GCP, Azure, AWS and bare-metal</p>	 <p>Overstock uses Cilium as their CNI for self-hosted bare metal clusters</p>	 <p>Palantir is using Cilium as their main CNI plugin in AWS (self-hosted k8s)</p>	 <p>Plaid uses Cilium as the CNI for its serverless database platform</p>	 <p>Liquid Reply is a consulting firm that uses Cilium in client projects</p>	 <p>Melenion uses Cilium as the CNI for its on-premise production clusters</p>	 <p>Mux uses Cilium on self-hosted clusters in GCP and AWS to run its video streaming/analytics platforms</p>	 <p>MyFitnessPal trusts Cilium with high volume user traffic on AWS and GKE</p>
 <p>PlanetScale uses Cilium as their CNI plugin in self-hosted Kubernetes on AWS</p>	 <p>Radio France uses Cilium as their self-hosted clusters on AWS</p>	 <p>Ropports Robotics uses Cilium as their main CNI plugin for self-hosted clusters</p>	 <p>SAP uses Cilium for projects across AWS, Azure, GCP, and OpenStack</p>	 <p>Sproutfi uses Cilium as the CNI on its GKE based clusters</p>	 <p>Superorbital uses Cilium in their customer engagements</p>	 <p>Taylor Brands uses Cilium in their EKS clusters</p>	 <p>The New York Times uses Cilium on EKS to build multi-region multi-tenant shared clusters</p>
 <p>Scaleway uses Cilium as the default CNI for Kubernetes Kapsule</p>	 <p>Schuberg Philis uses Cilium as the CNI for mission-critical Kubernetes clusters they run for their customers</p>	 <p>Simple uses Cilium as default CNI for DNS</p>	 <p>SmileDirectClub uses Cilium in self-hosted clusters vSphere and EC2 for manufacturing</p>	 <p>TIS uses Cilium for its Open Sovereign Cloud product</p>	 <p>Yahoo! is using Cilium for L4 North-South Load Balancing for Kubernetes Services</p>		



# eBPF

Makes the Linux kernel programmable in a secure and efficient way.

*“What JavaScript is to the browser, eBPF is to the Linux Kernel”*



```
int syscall__ret_execve(struct pt_regs *ctx)
{
    struct comm_event event = {
        .pid = bpf_get_current_pid_tgid() >> 32,
        .type = TYPE_RETURN,
    };

    bpf_get_current_comm(&event.comm, sizeof(event.comm));
    comm_events.perf_submit(ctx, &event, sizeof(event));

    return 0;
}
```



cilium



# Cilium CNI

Scalable, Secure,  
High Performance  
CNI Plugin





## Cilium CNI

Scalable, Secure,  
High Performance  
CNI Plugin



## Cilium Service Mesh

Sidecar-free Mesh &  
Ingress





## Cilium CNI

Scalable, Secure,  
High Performance  
CNI Plugin



## Cilium Service Mesh

Sidecar-free Mesh &  
Ingress



## Hubble

Network  
Observability







## Cilium CNI

Scalable, Secure,  
High Performance  
CNI Plugin



## Cilium Service Mesh

Sidecar-free Mesh &  
Ingress



## Hubble

Network  
Observability



## Tetragon

Security Observability &  
Runtime Enforcement





cilium

The Origins....

# Then



Cilium Design Summit,  
Diavolezza, 2016

# Now



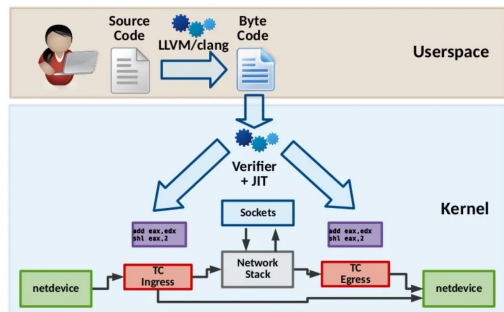
Isovalent Team,  
Punt Muragl, 2022

# First Conference Talk

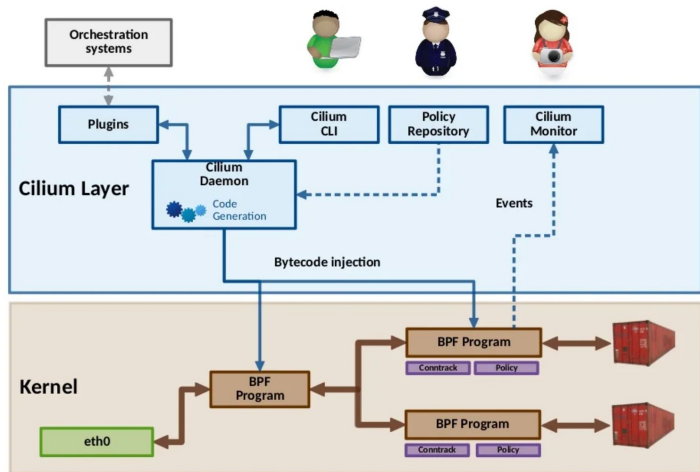


- IPv6 Only
- No concept of networks
- Policy decoupled from addressing
- Giant flat L3

## BPF - Berkley Packet Filter

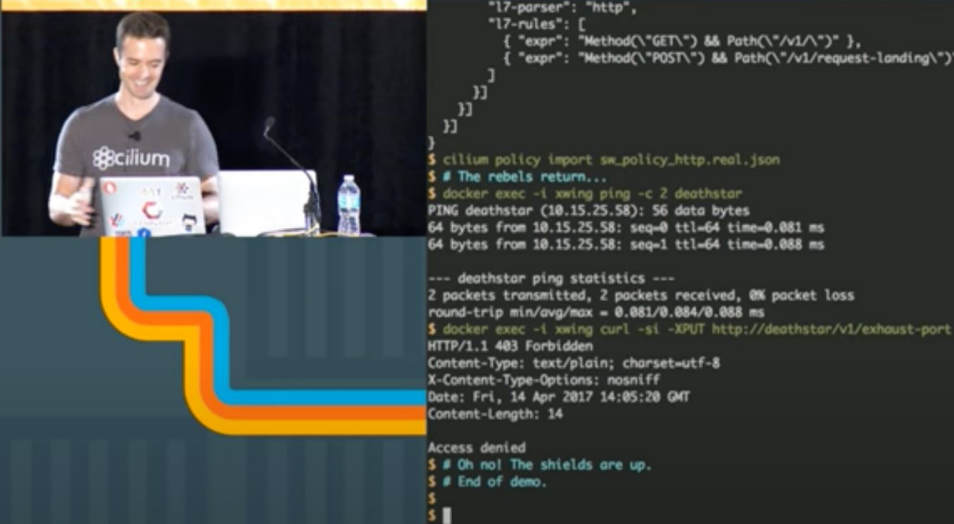


## Cilium Architecture



# DockerCon 2017

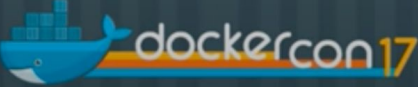
## When everything changed



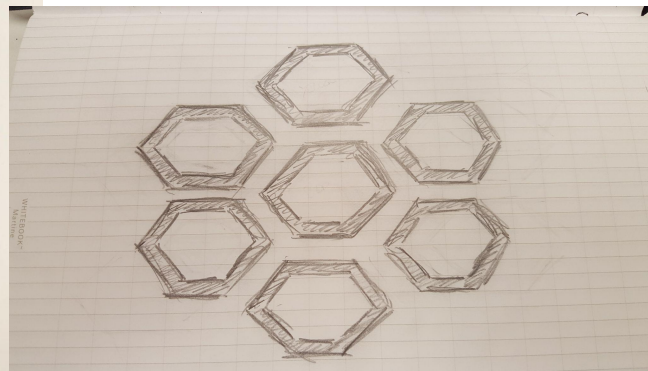
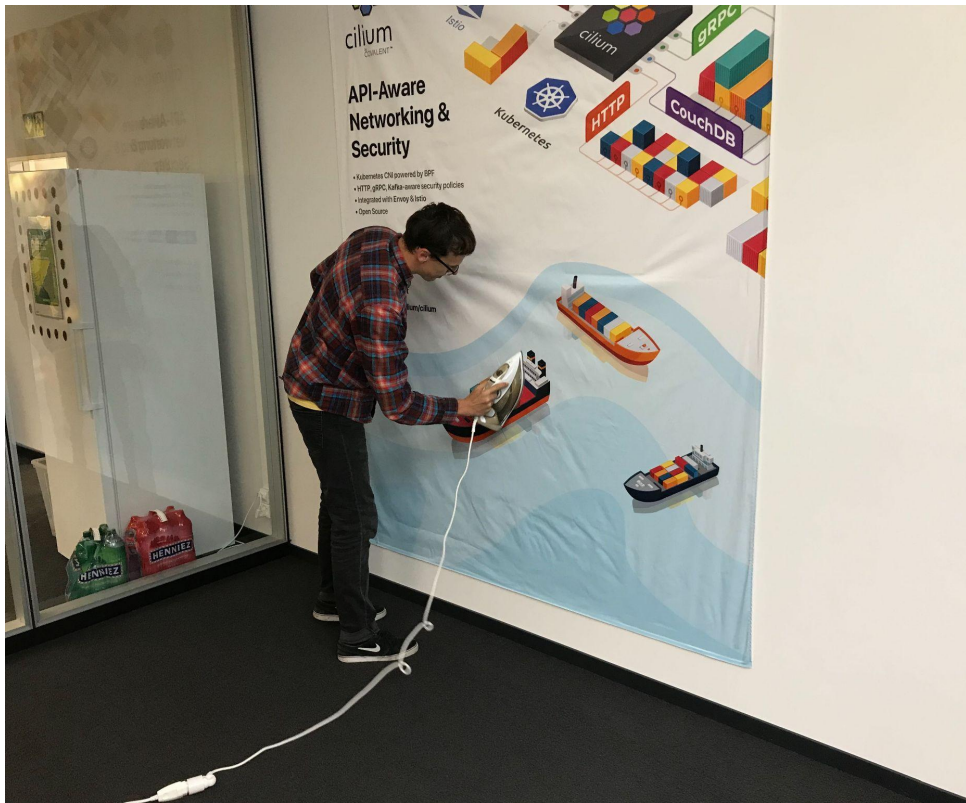
```
"l7-parser": "http",
"l7-rules": [
  { "expr": "Method(\"GET\") && Path(\"/v1/\")" },
  { "expr": "Method(\"POST\") && Path(\"/v1/request-landing\")" }
]
}
]]
]]
}
}
}
$ cilium policy import sw_policy_http.real.json
$ # The rebels return...
$ docker exec -i xwing ping -c 2 deathstar
PING deathstar (10.15.25.58): 56 data bytes
64 bytes from 10.15.25.58: seq=0 ttl=64 time=0.081 ms
64 bytes from 10.15.25.58: seq=1 ttl=64 time=0.088 ms

--- deathstar ping statistics ---
2 packets transmitted, 2 packets received, 0% packet loss
round-trip min/avg/max = 0.081/0.084/0.088 ms
$ docker exec -i xwing curl -sI -XPUT http://deathstar/v1/exhaust-port
HTTP/1.1 403 Forbidden
Content-Type: text/plain; charset=utf-8
X-Content-Type-Options: nosniff
Date: Fri, 14 Apr 2017 14:05:20 GMT
Content-Length: 14

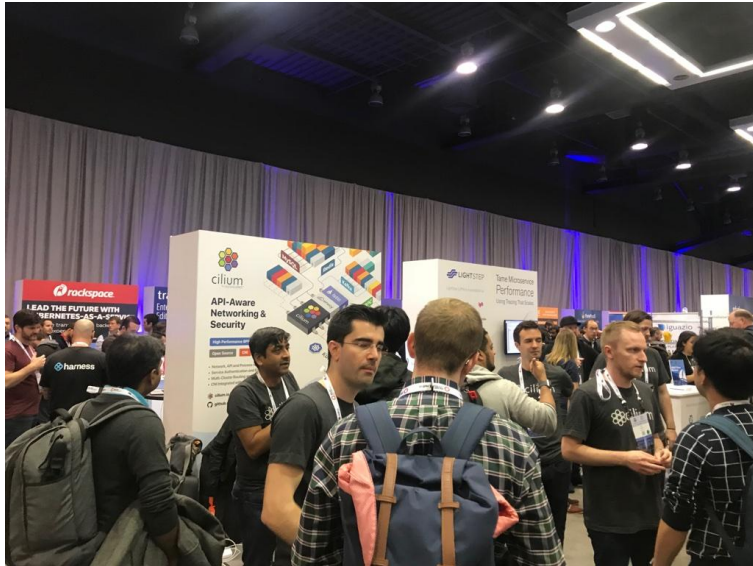
Access denied
$ # Oh no! The shields are up.
$ # End of demo.
$
$ |
```



# 2017 - First Office



# 2018 Starting to Grow



# 2019 Team Building

The Snow  
Chains  
Incident,  
Julier Pass





# 2019 Swiss Culture



























Fondue,  
AirBnB,  
Palo Alto



# 2021- Things get crazy...

Cilium joins the CNCF

Cloud Native Network

 cilium CNCI Incubating	 CNI CNCI Incubating	 ANTREA	 aviatrix	CNI-Genie	CUMULUS		
		 DANM	 FabEdge	 i.io	 flannel		
 Guardicore	 ISOVALENT	 kg	 Kube-OVN	 KUBE ROUTER	 LIGATO	 MULTUS	 Network Service Mesh
 nuagenetworks	 OvS Open vSwitch	 CALICO	 SUBMARINER	 tungstenfabric	 vmware NSX	 weave net	

AWS picks Cilium



Google picks Cilium



# 2022 - Crazy^2

## Isovalent raises \$40M Series B led by Thomvest Ventures

by Ashish Nain • September 8, 2022



Isovalent announced it has closed a \$40M Series B funding round led by Thomvest Ventures. M12 (Microsoft's Venture Fund) and Grafana Labs joined Google and Cisco as existing strategic investors in the company, highlighting the central position that Isovalent occupies in the eBPF and broader cloud native ecosystem. Additional investors include Andreessen Horowitz, Mango Capital, and Mirae Asset Capital.



**THOMVEST**



Grafana



**Microsoft**

**a16z**  
ANDREESSEN HOROWITZ

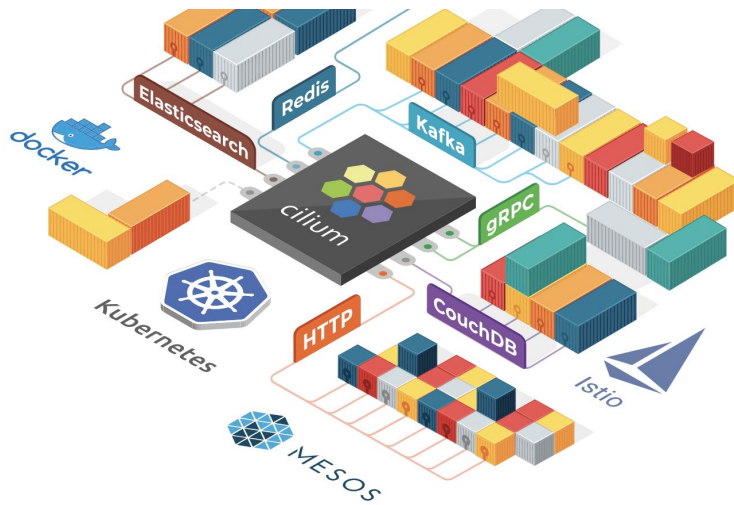
**Google**

**CISCO**



cilium

CNI



# What is Cilium CNI?

## Efficient and Scalable Kubernetes CNI

- IPv4, IPv6, NAT46, SRv6, ...
- Overlays, BGP, Cloud Provider SDNs

## High-performance Load-Balancing

- Kubernetes Services
- North-South load-balancer
- Kubernetes Ingress

## Network Policies & Encryption

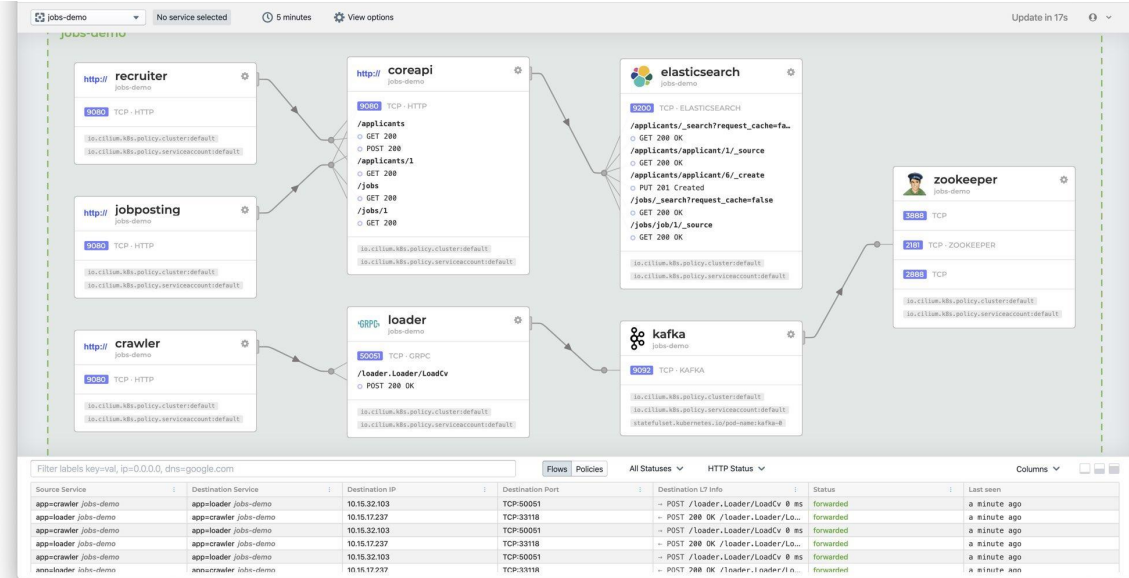
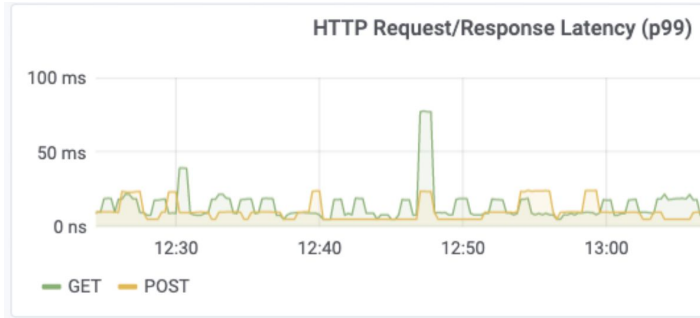
- Kubernetes Network Policy
- Cilium Network Policy (FQDN, L7, ...)
- Transparent Encryption

## Multi-Cluster & External Workloads

- Global Services, Service Discovery, Network Policy
- Integration of Metal & VMs
- Egress Gateway



# Hubble Observability



## Metrics, Logs, & Service Ma

- L3/L4
- L7 (HTTP, DNS, Kafka, ...)
- Network Policy
- ...

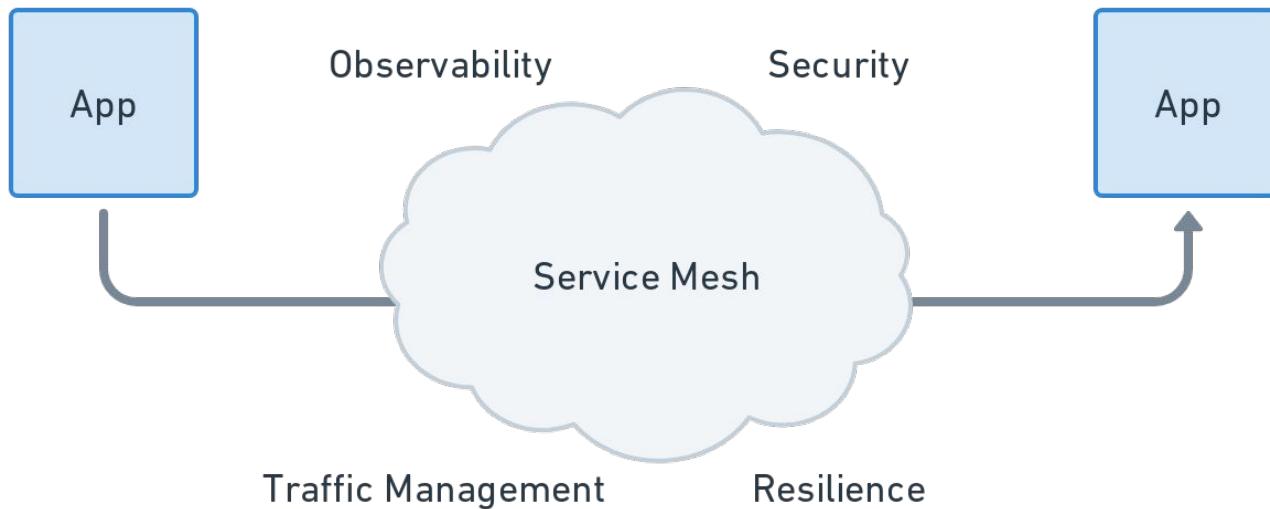




cilium

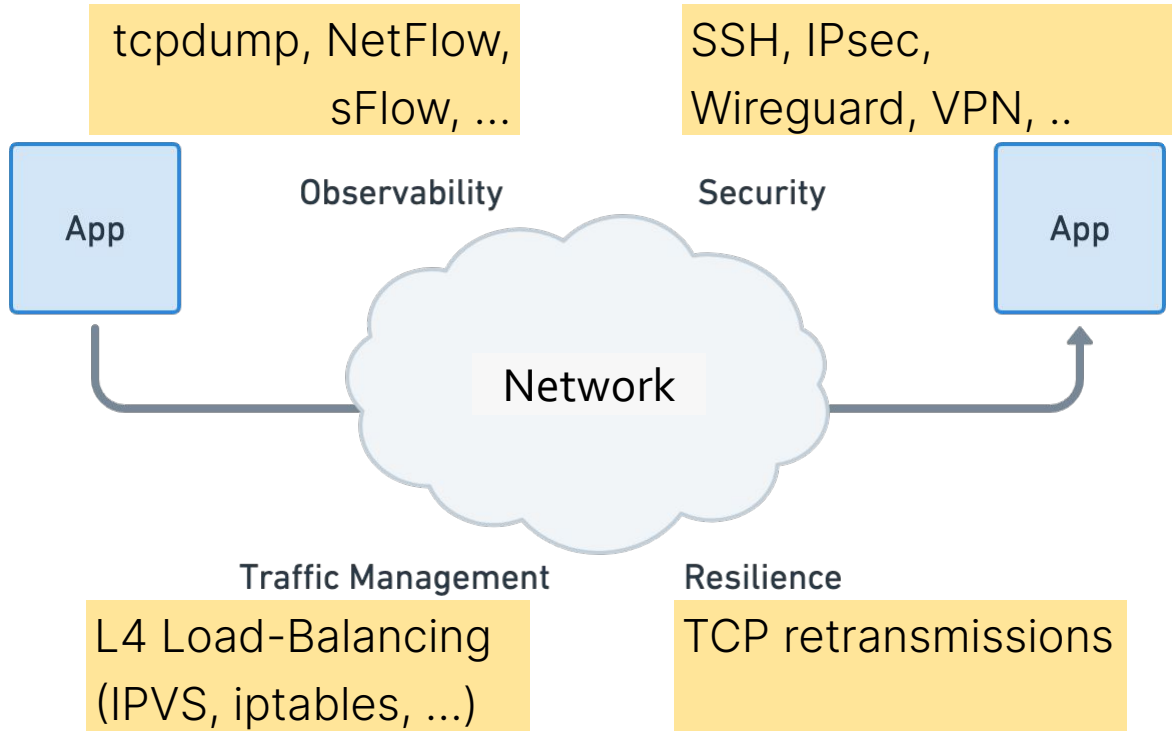
Service Mesh

# Observable, secure, resilient cloud native connectivity

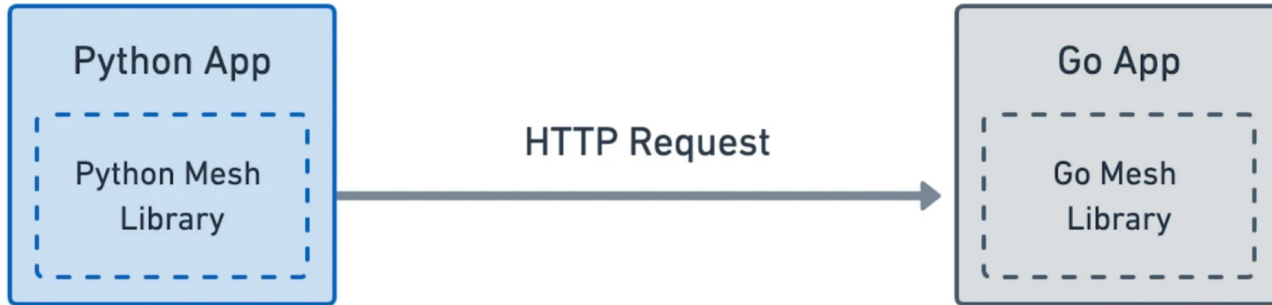




# Traditional networking is falling short

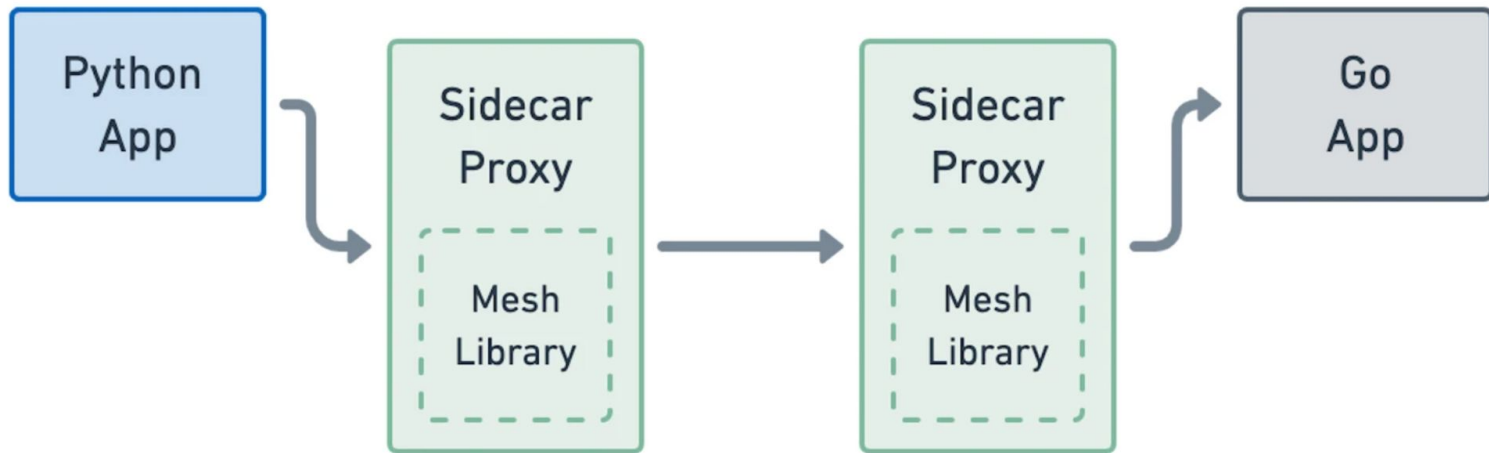


# Service Mesh Origins



Each application requires a service mesh library written in the language framework of the application.

# Service Mesh with Sidecars



Service mesh is embedded in a proxy running outside of the application.



**Kelsey Hightower** 

@kelseyhightower



service mess /'sərvəs mes/  
noun

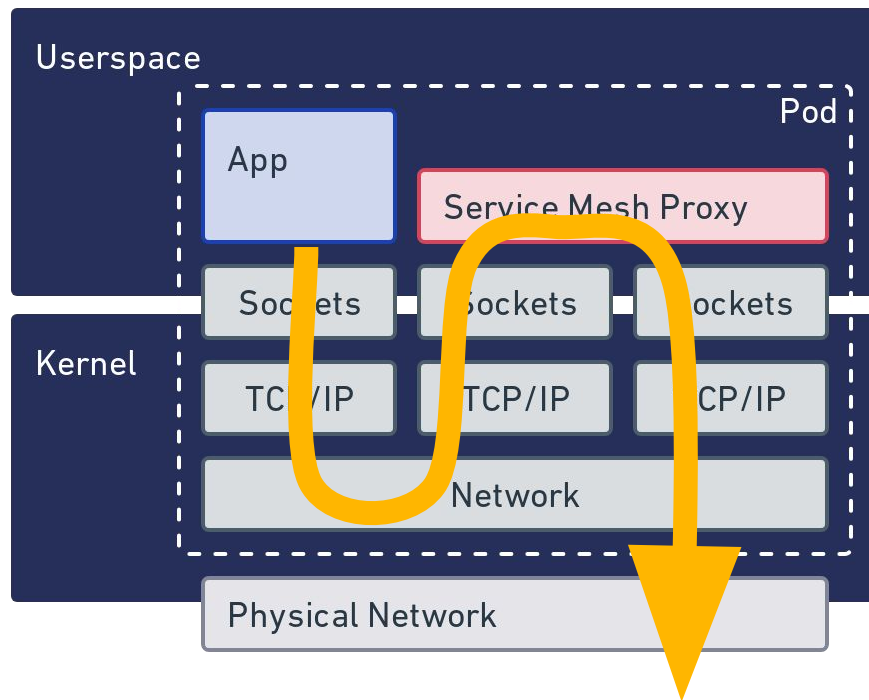
1. the result of spending more compute resources than your actual business logic dynamically generating and distributing Envoy proxy configs and TLS certificates.

11:43 PM · Jul 13, 2019 · Twitter Web App

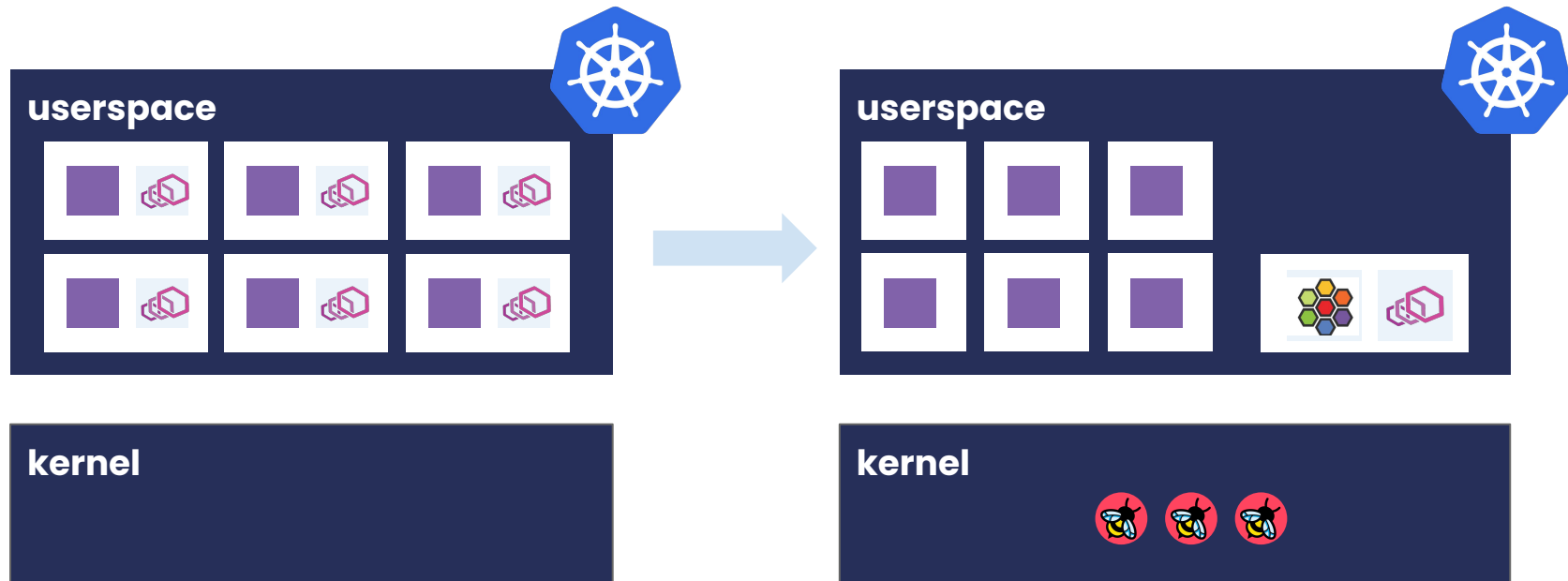
---

**397** Retweets   **30** Quote Tweets   **1,542** Likes

# The network cost of sidecar proxies



# Removing sidecars from Service Mesh





# Cilium Service Mesh

## Option 1: Sidecar-free



## Option 2: Istio Integration



## Control plane of your choice



Istio



Ingress/  
Services



Gateway  
API



SPIFFE

## Observability Integrations



fluentd



JSON

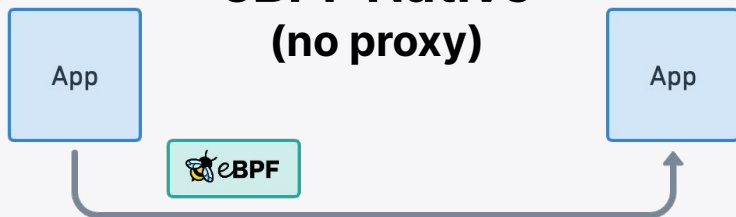


Grafana



OpenTelemetry

## eBPF Native (no proxy)



**Whenever possible**

### Traffic Management

- L3/L4 forwarding & Load-balancing
- Canary, Topology Aware Routing
- Multi-cluster

### Security

- Network Policy
- mTLS

### Observability

- Tracing, OpenTelemetry, & Metrics
- HTTP, TLS, DNS, TCP, UDP, ...

## Proxy



**When eBPF cannot do it**

### Traffic Management

- L7 Load-balancing & Ingress

### Resilience

- Retries, L7 Rate Limiting

### Security

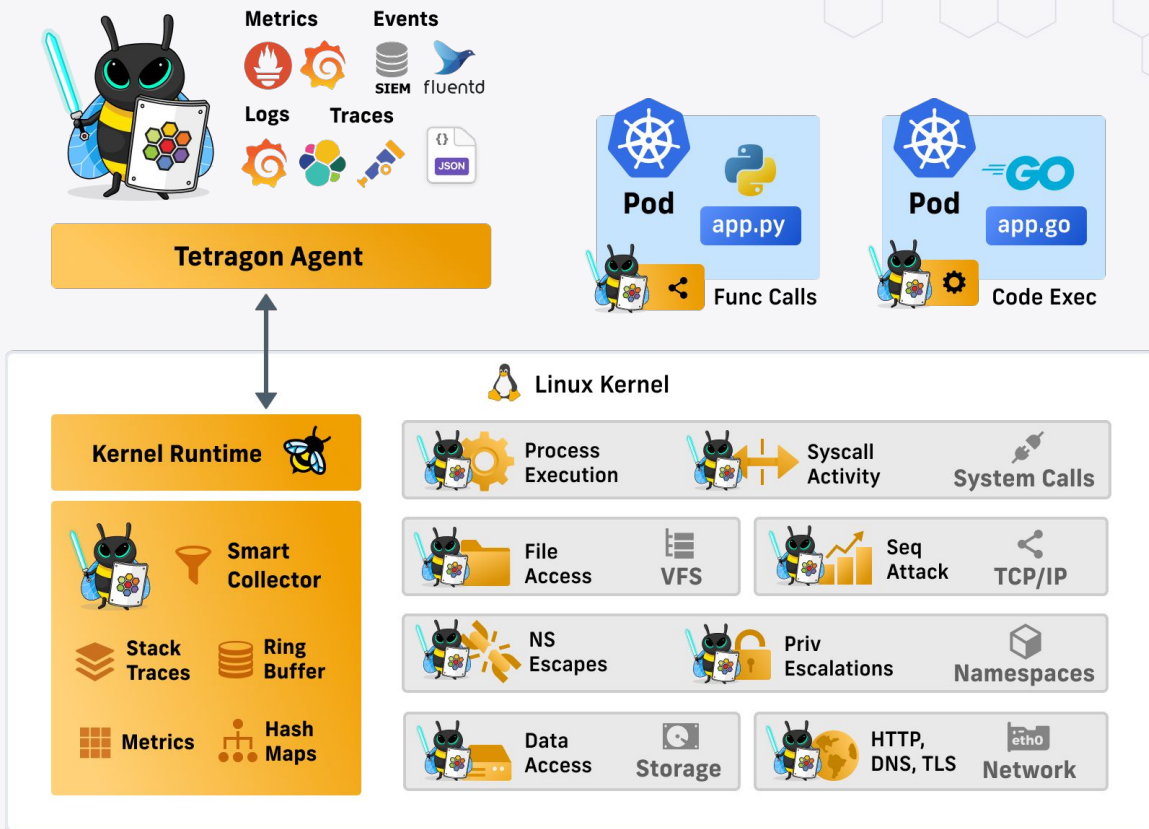
- TLS Termination & Origination
- L7 Network Policy\*



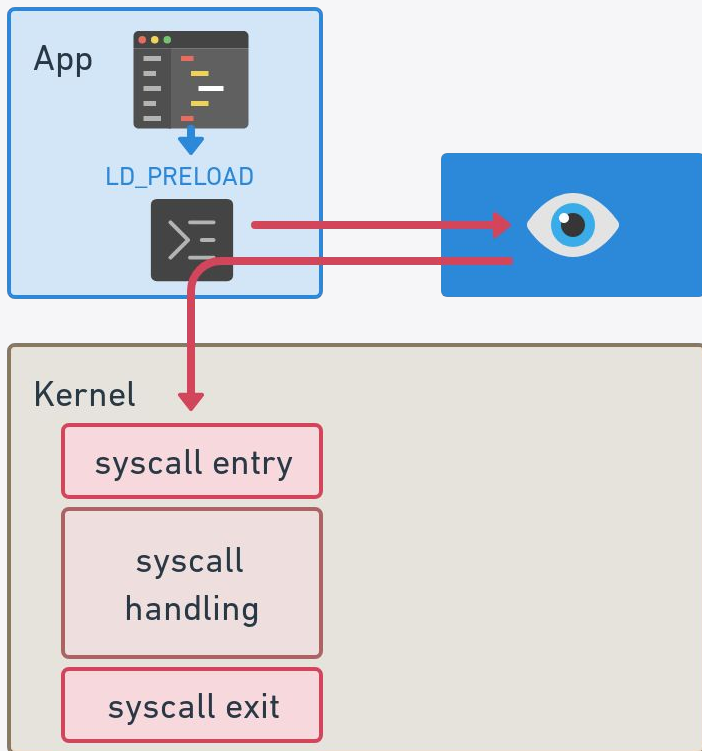


# Tetragon

Security Observability &  
Runtime Enforcement

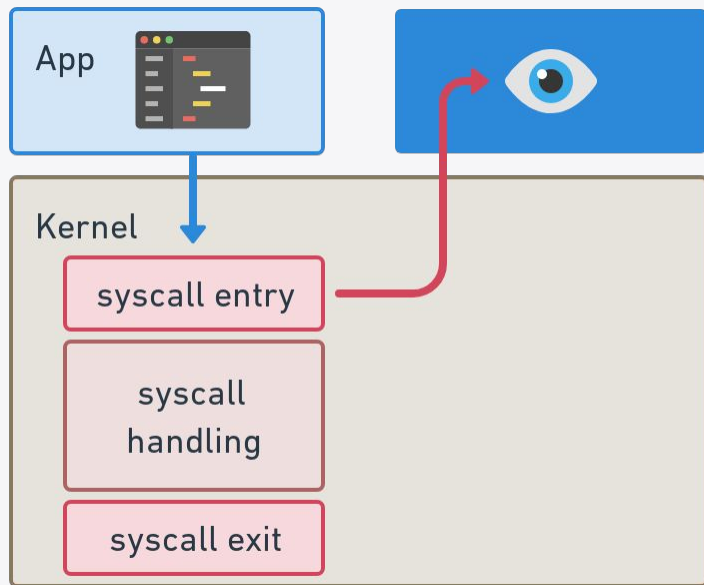


# LD\_PRELOAD



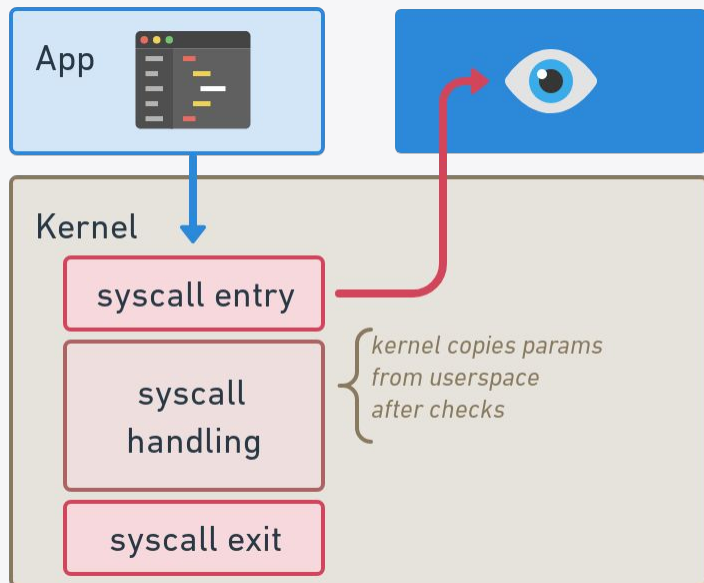
- Standard C library, dynamically linked
- System call API
- Replace the “standard” library

# Syscall checks within the kernel



ptrace,  
seccomp,  
eBPF kprobes on syscall entry

# TOCTTOU with syscalls

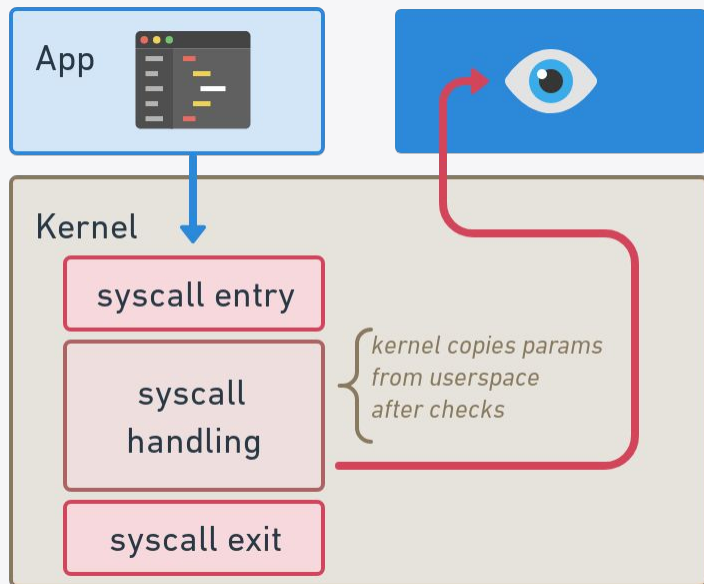


ptrace,  
seccomp,  
eBPF kprobes on syscall entry

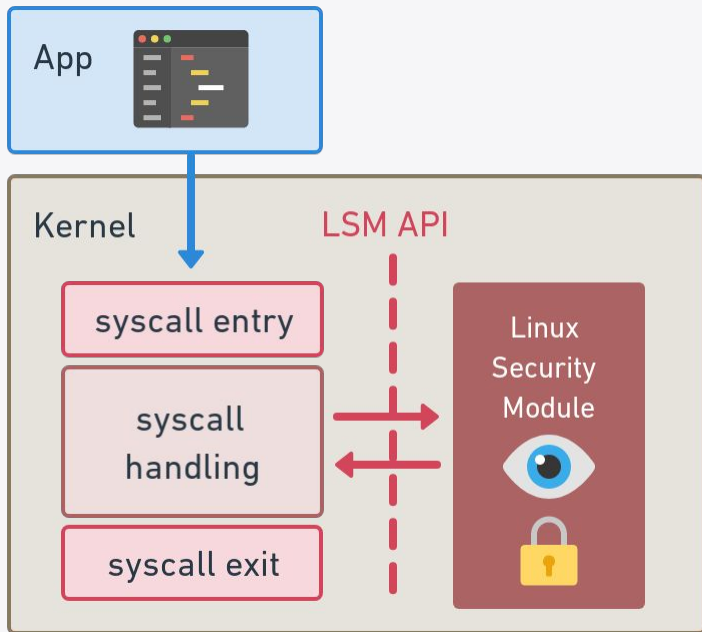
For more details

- Leo Di Donato & KP Singh at CN eBPF Day 2021
- Rex Guo & Junyuan Zeng at DEFCON 29 on Phantom attacks

# Need to make the check at the right place

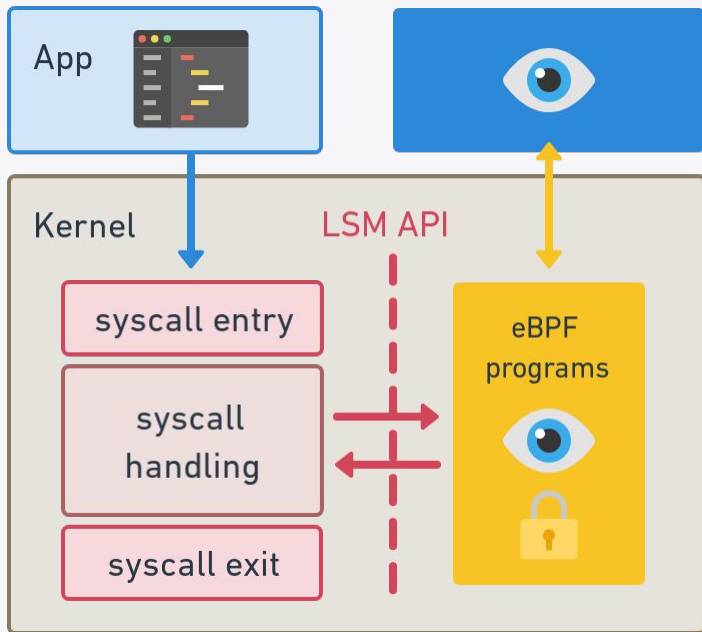


# Linux Security Modules



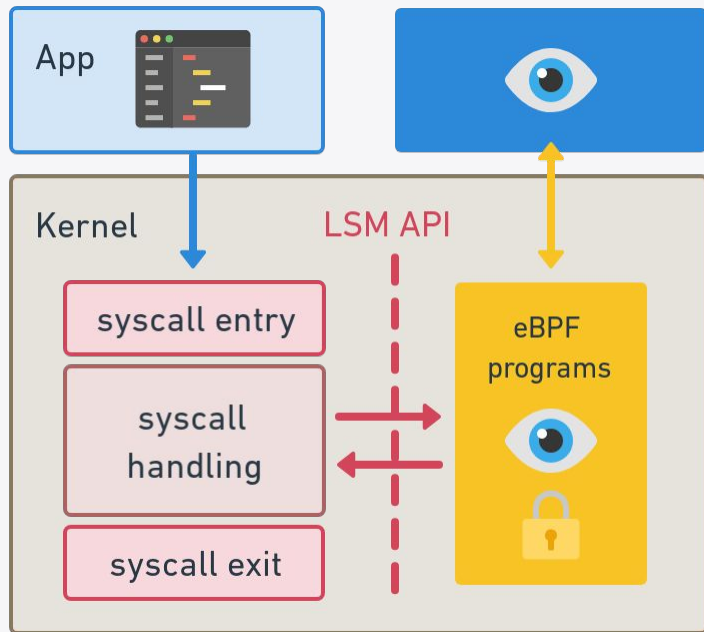
- Stable interface
- Safe places to make checks

# BPF LSM



- Stable interface
- Safe places to make checks
- eBPF makes it dynamic
- Protect pre-existing processes

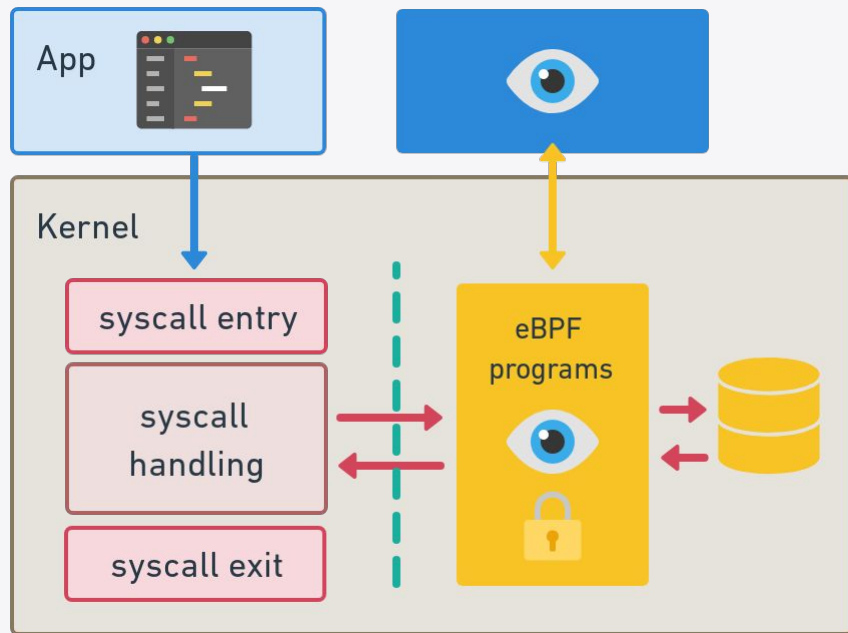
# BPF LSM



- Stable interface
- Safe places to make checks
- eBPF makes it dynamic
- Protect pre-existing processes
- Needs **kernel 5.7+**



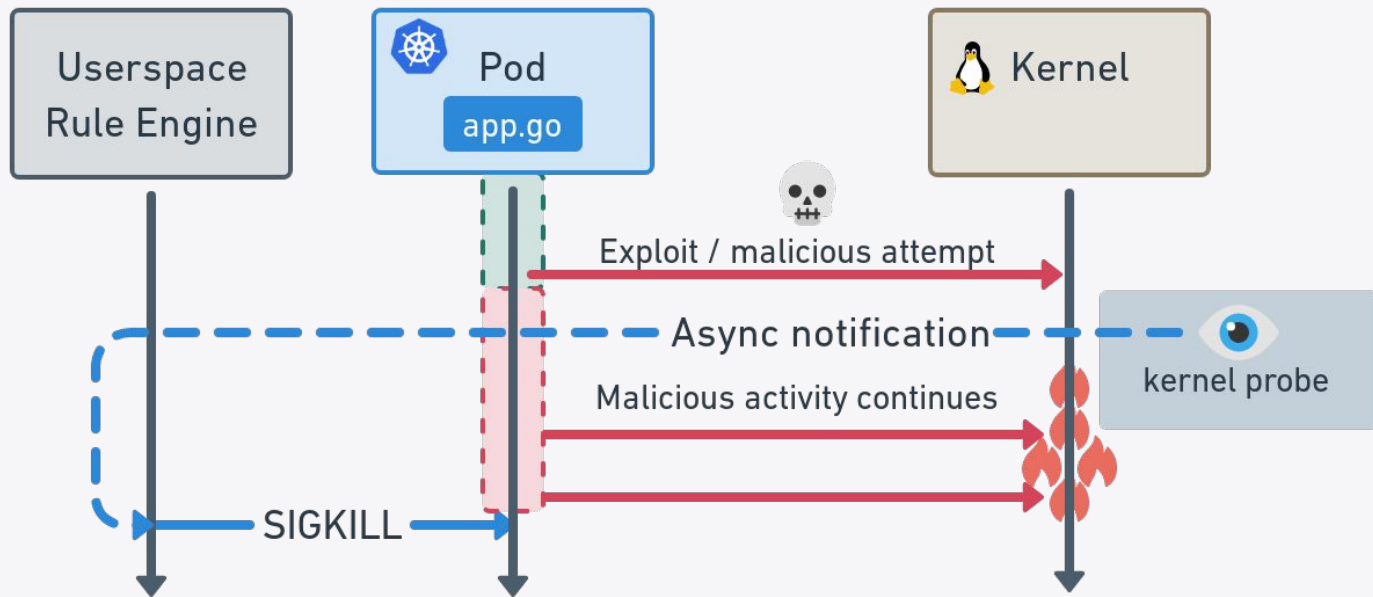
# Cilium Tetragon



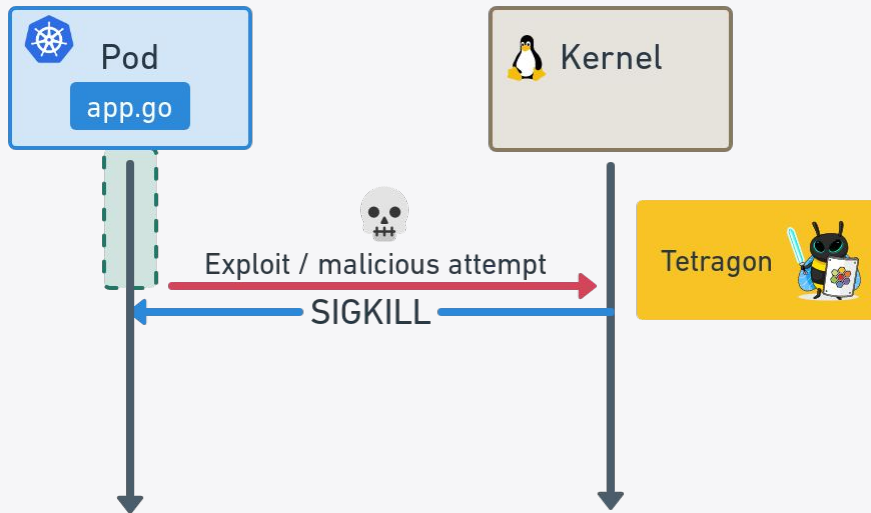
- eBPF makes it dynamic
- Protect pre-existing processes
- Uses **kernel knowledge** to hook into sufficiently stable functions
- Multiple **co-ordinated** eBPF programs
- In-kernel event **filtering**



# Reactive actions from user space



# Preventative actions from kernel





# Observability

- **Deep Visibility**

- System, network, protocols, filesystem, applications, ...

- **Transparent**

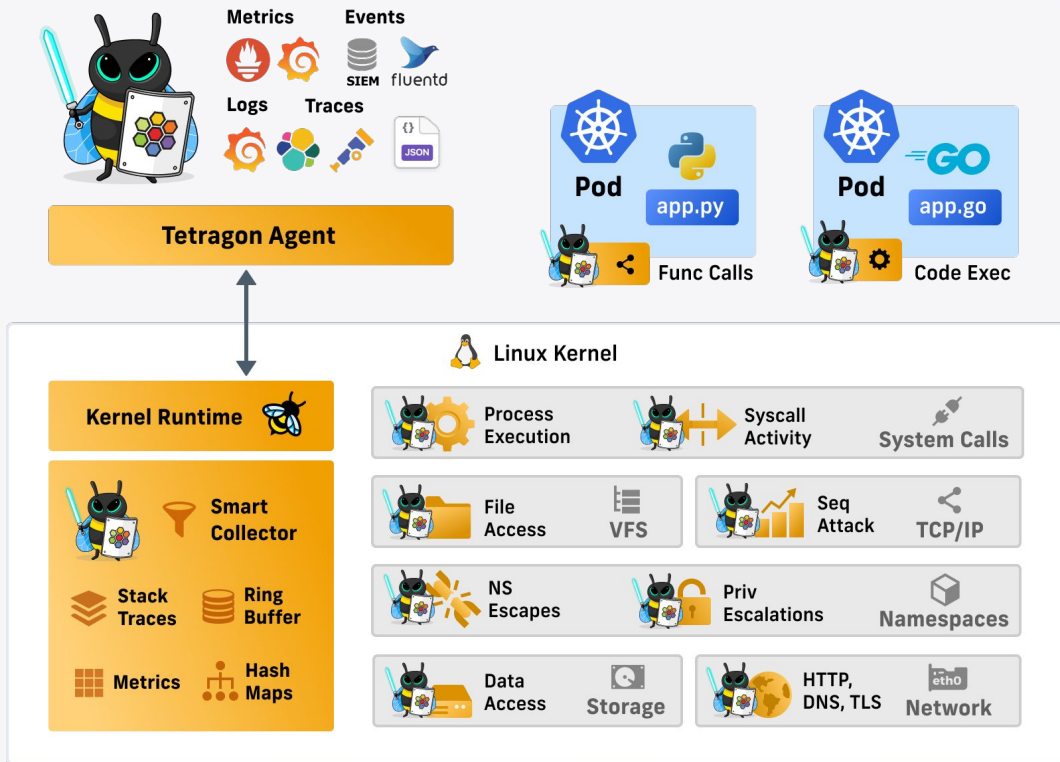
- App agonistic
- No changes to applications

- **Low-Overhead**

- Minimal overhead
- Extensive filtering & aggregation

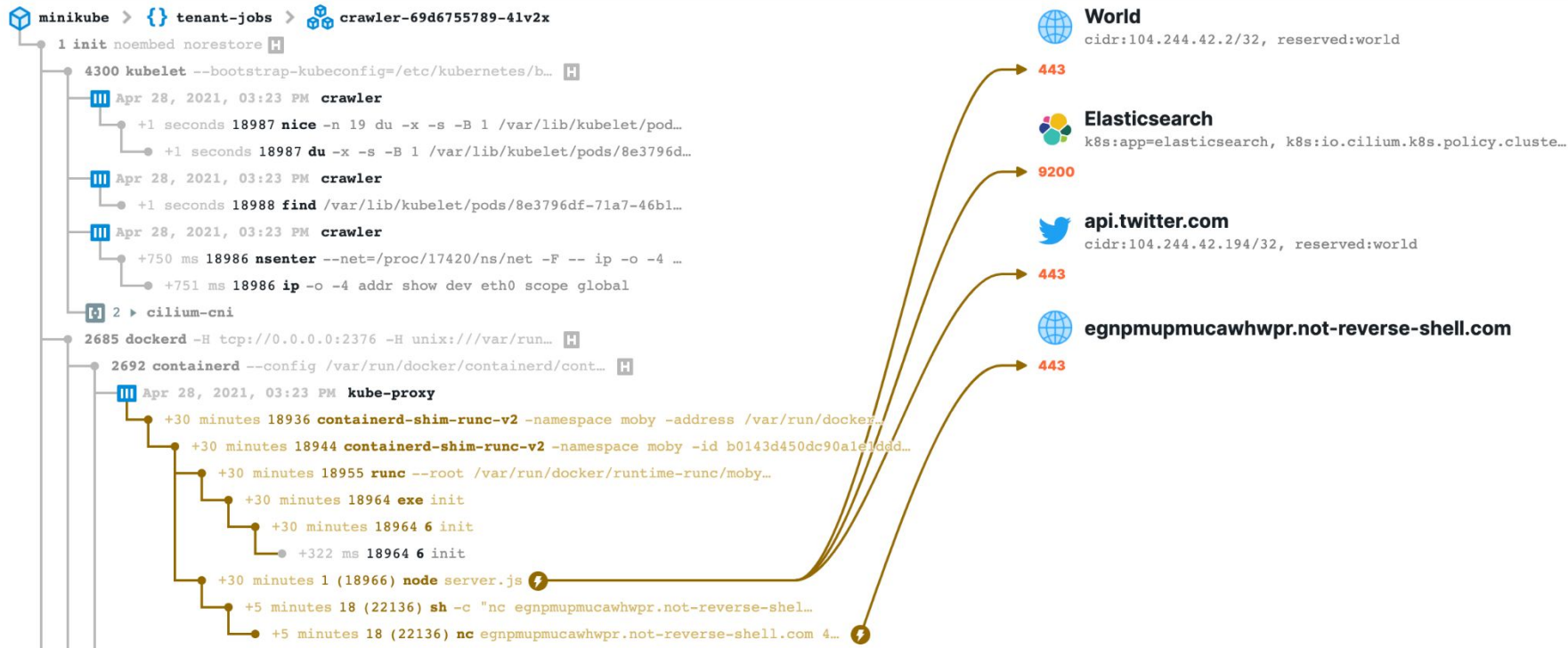
- **Integrations**

- Prometheus, Grafana, SIEM, fluentd, OpenTelemetry, elasticsearch

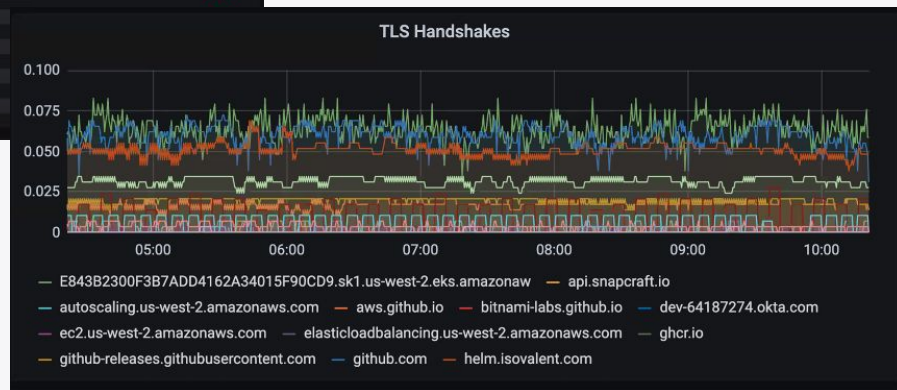
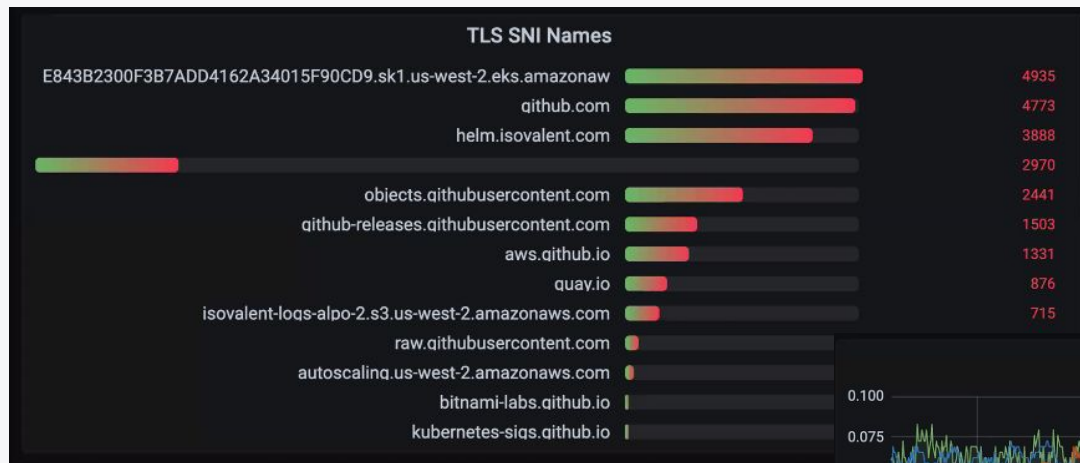


# Combined Network & Runtime Visibility

## Process tree



# TLS/SSL Visibility



# Observing DNS, HTTP, TCP, ...

```
process default/test-pod /usr/local/bin/curl cilium.io
dns default/test-pod /usr/local/bin/curl [cilium.io.default.svc.cluster.local.] => []
dns default/test-pod /usr/local/bin/curl [cilium.io.default.svc.cluster.local.] => []
dns default/test-pod /usr/local/bin/curl [cilium.io.default.svc.cluster.local.] => []
dns default/test-pod /usr/local/bin/curl [cilium.io.default.svc.cluster.local.] => []
dns default/test-pod /usr/local/bin/curl [cilium.io.svc.cluster.local.] => []
dns default/test-pod /usr/local/bin/curl [cilium.io.svc.cluster.local.] => []
dns default/test-pod /usr/local/bin/curl [cilium.io.svc.cluster.local.] => []
dns default/test-pod /usr/local/bin/curl [cilium.io.svc.cluster.local.] => []
dns default/test-pod /usr/local/bin/curl [cilium.io.cluster.local.] => []
dns default/test-pod /usr/local/bin/curl [cilium.io.cluster.local.] => []
dns default/test-pod /usr/local/bin/curl [cilium.io.cluster.local.] => []
dns default/test-pod /usr/local/bin/curl [cilium.io.cluster.local.] => []
dns default/test-pod /usr/local/bin/curl [cilium.io.c.cilium-dev.internal.] => []
dns default/test-pod /usr/local/bin/curl [cilium.io.c.cilium-dev.internal.] => []
dns default/test-pod /usr/local/bin/curl [cilium.io.c.cilium-dev.internal.] => []
dns default/test-pod /usr/local/bin/curl [cilium.io.c.cilium-dev.internal.] => []
dns default/test-pod /usr/local/bin/curl [cilium.io.google.internal.] => []
dns default/test-pod /usr/local/bin/curl [cilium.io.google.internal.] => []
dns default/test-pod /usr/local/bin/curl [cilium.io.google.internal.] => []
dns default/test-pod /usr/local/bin/curl [cilium.io.] => []
dns default/test-pod /usr/local/bin/curl [cilium.io.] => [104.198.14.52]
dns default/test-pod /usr/local/bin/curl [cilium.io.] => []
dns default/test-pod /usr/local/bin/curl [cilium.io.] => []
connect default/test-pod /usr/local/bin/curl TCP 10.80.0.12:43278 => 104.198.14.52:80 [cilium.io.]
http default/test-pod /usr/local/bin/curl cilium.io GET / 301 Moved Permanently 154.733717ms
exit default/test-pod /usr/local/bin/curl cilium.io 0
close default/test-pod /usr/local/bin/curl TCP 10.80.0.12:43278 => 104.198.14.52:80 [cilium.io.] tx 73 B rx 1.2 kB
socket default/test-pod /usr/local/bin/curl TCP 10.80.0.12:43278 => 104.198.14.52:80 [cilium.io.] tx 73 B rx 1.2 kB
```

# Monitoring & Preventing Capabilities Abuse

```
process default/test-pod /usr/bin/nsenter -t 1 -m -u -n -i -p bash ● CAP_SYS_ADMIN
process default/test-pod /usr/bin/dircolors --coreutils-prog-shebang=dircolors /usr/bin/dircolors -b /etc/DIR_COLORS ● CAP_SYS_ADMIN
exit default/test-pod /usr/bin/dircolors --coreutils-prog-shebang=dircolors /usr/bin/dircolors -b /etc/DIR_COLORS 0 ● CAP_SYS_ADMIN
setns default/test-pod /usr/bin/nsenter ipc ● CAP_SYS_ADMIN
setns default/test-pod /usr/bin/nsenter uts ● CAP_SYS_ADMIN
setns default/test-pod /usr/bin/nsenter net ● CAP_SYS_ADMIN
setns default/test-pod /usr/bin/nsenter pid ● CAP_SYS_ADMIN
setns default/test-pod /usr/bin/nsenter mnt ● CAP_SYS_ADMIN
process default/test-pod /bin/bash ● CAP_SYS_ADMIN
open default/test-pod /bin/bash /etc/passwd ● CAP_SYS_ADMIN
close default/test-pod /bin/bash /etc/passwd ● CAP_SYS_ADMIN
process default/test-pod /usr/bin/vi /etc/passwd ● CAP_SYS_ADMIN
open default/test-pod /usr/bin/vi /etc/passwd ● CAP_SYS_ADMIN
close default/test-pod /usr/bin/vi /etc/passwd ● CAP_SYS_ADMIN
open default/test-pod /usr/bin/vi /etc/passwd ● CAP_SYS_ADMIN
open default/test-pod /usr/bin/vi /etc/passwd ● CAP_SYS_ADMIN
close default/test-pod /usr/bin/vi /etc/passwd ● CAP_SYS_ADMIN
close default/test-pod /usr/bin/vi /etc/passwd ● CAP_SYS_ADMIN
open default/test-pod /usr/bin/vi /etc/passwd ● CAP_SYS_ADMIN
close default/test-pod /usr/bin/vi /etc/passwd ● CAP_SYS_ADMIN
open default/test-pod /usr/bin/vi /etc/passwd ● CAP_SYS_ADMIN
write default/test-pod /usr/bin/vi /etc/passwd 8 bytes ● CAP_SYS_ADMIN
exit default/test-pod /usr/bin/vi /etc/passwd SIGKILL ● CAP_SYS_ADMIN
```



ISOVALENT

**Thank you!**

