ISOVALENT

Surviving Day 2

How to Troubleshoot Kubernetes Networking

Thomas Graf, Isovalent

KubeCon | CloudNativeCon

Europe 2023

Surviving Day 2
How to Troubleshoot Kubernetes Networking

Thomas Graf
Isovalent

Context:

- I'm a Cilium Maintainer
- Lessons learned helping users with Kubernetes Networking

cilium

Created by ISOVALENT

eBPF-based:

- Networking
- Security
- Observability
- Service Mesh & Ingress

Foundation

CLOUD NATIVE COMPUTING FOUNDATION

Technology

eBPF envoy

Cilium CNI

Scalable, Secure, High Perfomance Networking

Cilium Service Mesh

Sidecar-free Service Mesh, Ingress, & Gateway API

Hubble

Network Observability & Monitoring

Tetragon

Security Observability & Runtime Enforcement

Adobe
What Makes a Good Multi-tenant Kubernetes Solution
VIDEO 1 · VIDEO 2

Alibaba Cloud
Building High-Performance Cloud Native Pod Networks
READ BLOG

aws
AWS picks Cilium for Networking & Security on EKS Anywhere
READ BLOG

Bell
Bell uses Cilium and eBPF for telco networking
VIDEO 1 · VIDEO 2

AccuKnox
AccuKnox uses Cilium for network visibility and network policy enforcement

Acoss
Acoss uses Cilium as their main CNI plugin for self hosted Kubernetes

ArangoDB
ArangoDB Oasis uses Cilium to separate database deployments in a multi-tenant cloud environment

ayedo
Ayedo builds and operates cloud native platforms using Cilium

Capital One
Building a Secure and Maintainable PaaS
WATCH VIDEO

CENGN
Cloud Native Networking with eBPF
WATCH VIDEO

DATADOG
Datadog is using Cilium in AWS (self-hosted k8s)
WATCH VIDEO

DigitalOcean
Managed Kubernetes: 1.5 Years of Cilium Usage at DigitalOcean
WATCH VIDEO

ByteDance
ByteDance uses Cilium as their CNI for self-hosted Kubernetes clusters

Canonical
Canonical's Kubernetes distribution microk8s uses Cilium as CNI plugin

Civo
Civo is offering Cilium as the CNI option for Civo users to choose it for their Civo Kubernetes clusters

Cognite
Cognite uses Cilium as the CNI plugin for industrial DataOps

证通股份 E-CAPITAL TRANSFER
ect888 uses Cilium as their CNI and for load balancing
READ BLOG

GitLab
Kubernetes Network Policies in Action with Cilium
VIDEO

Google
Google chooses Cilium for Google Kubernetes Engine (GKE) networking
READ BLOG

IKEA
IKEA uses Cilium for their self-hosted bare-metal private cloud
WATCH VIDEO

elasticpath
Elastic Path uses Cilium in their production CloudOps Kubernetes clusters

F5
F5 uses Cilium VXLAN tunnel integration with BIG-IP

finleap connect
finleap connect uses Cilium on a bare metal private cloud

FORM3
Form3 is using Cilium in their production clusters (self-hosted, bare-metal, private cloud)

MÁSMÓVIL
Scaling a Multi-Tenant Kubernetes Clusters in a Telco
WATCH VIDEO

Meltwater
Meltwater is using Cilium in AWS on self-hosted multi-tenant k8s clusters as the CNI plugin
WATCH VIDEO

MOBILAB
Mobilabs uses Cilium as the CNI for their internal cloud
READ BLOG

nexxiot
Nexxiot using Cilium as the CNI plugin on EKS for its IoT SaaS
READ USER STORY

infomaniak
Infomaniak uses Cilium in self-hosted clusters on bare-metal and Openstack

INNOQ
innoQ uses Cilium to run their customer's infrastructure

ISOVALENT
Cilium is the platform that powers **Isovalent**'s enterprise networking, observability, and security solutions

JUMO
JUMO uses Cilium as the CNI plugin for all of their AWS-hosted EKS clusters

PostFinance
PostFinance is using Cilium as their CNI for all mission critical, on premise k8s clusters
READ CASE STUDY

sky
eBPF & Cilium at Sky
WATCH VIDEO

Sky BET
Skybet uses Cilium as their CNI
READ BLOG

Trip.com
Trip.com uses Cilium both on premise and in AWS
BLOG 1 · BLOG 2

KRYPTOS LOGIC
Kryptos uses Cilium as the CNI for their on-prem Kubernetes clusters

Kube-OVN
Kube-OVN uses Cilium to enhance the CNI service performance, security and monitoring

KUBERMATIC
Kubermatic uses Cilium as the CNI for its Kubernetes installer and platform

KUBESPHERE
KubeKey is an open-source lightweight tool for deploying Kubernetes clusters and addons

Northflank
Northflank uses Cilium as its CNI plugin across GCP, Azure, AWS and bare metal

overstock.com
Overstock uses Cilium as their CNI for self hosted bar metal clusters

Palantir
Palantir is using Cilium as their main CNI plugin in AWS (self hosted k8s)

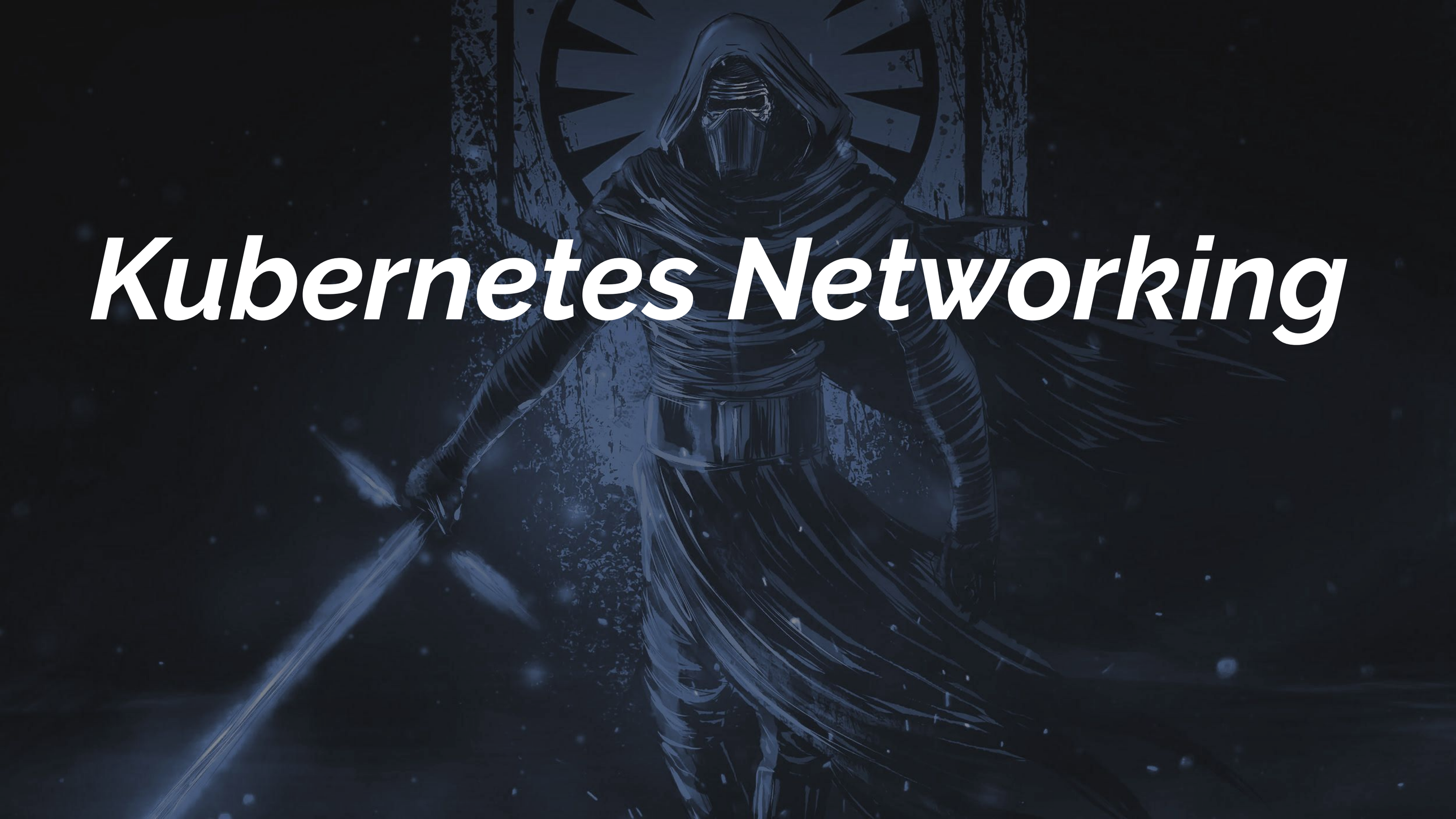
PLAID
Plaid uses Cilium as the CNI for its serverless database platform

REPLY LIQUID
Liquid Reply is a consulting firm that uses Cilium in client projects

Melenion Inc
Melenion uses Cilium as the CNI for its on-premise production clusters

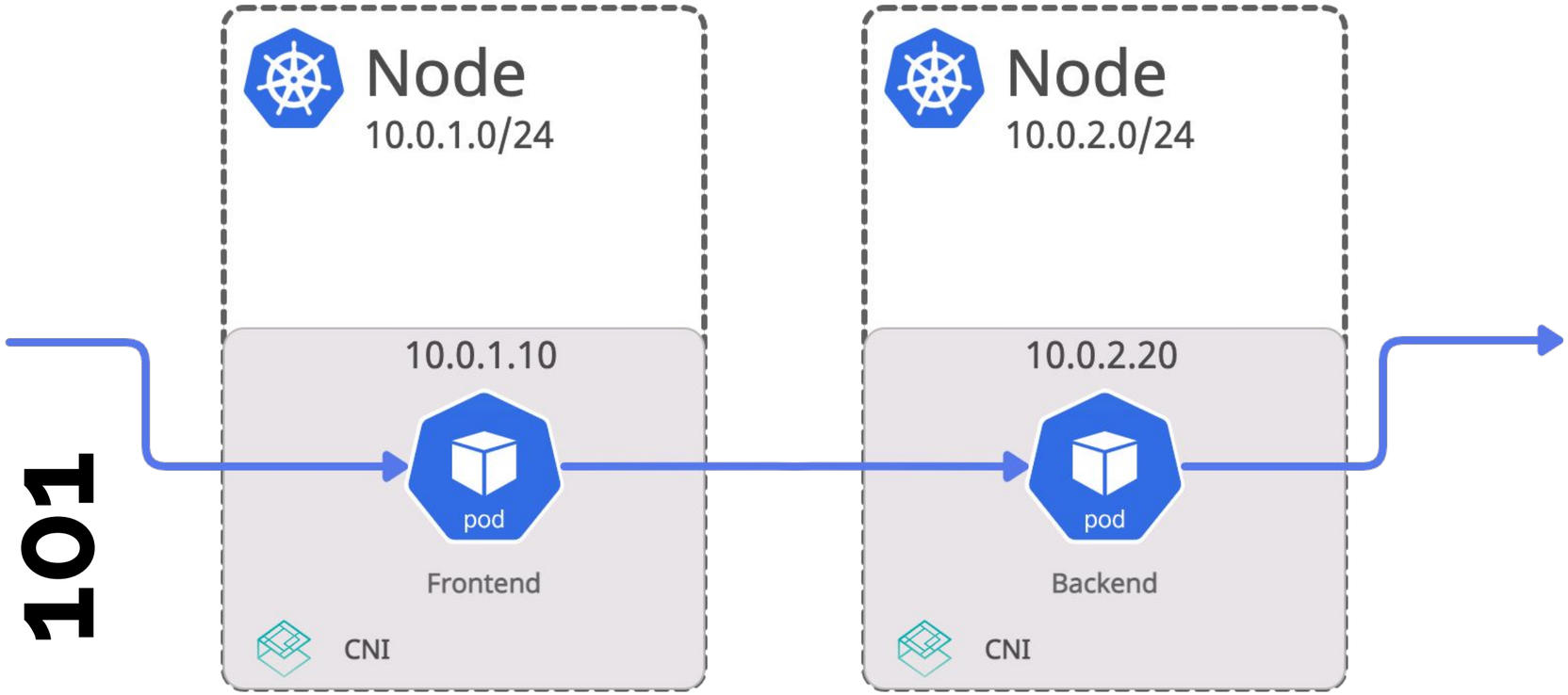
MUX
Mux uses Cilium on self-hosted clusters in GCP and AWS to run its video streaming/analytics platforms

myfitnesspal
MyFitnessPal trusts Cilium with high volume user traffic on AWS and GKE



Kubernetes Networking

Kubernetes Networking 101



- All Pods have IPs
- All Pods can talk
- PodCIDR[s] per node

- Services for load-balancing
- DNS for service-discovery
- Network Policy for segmentation

From KubeCon EU

KubeCon CloudNativeCon Europe 2021

Virtual

iptables

kube-proxy

App team scheduling 5K services

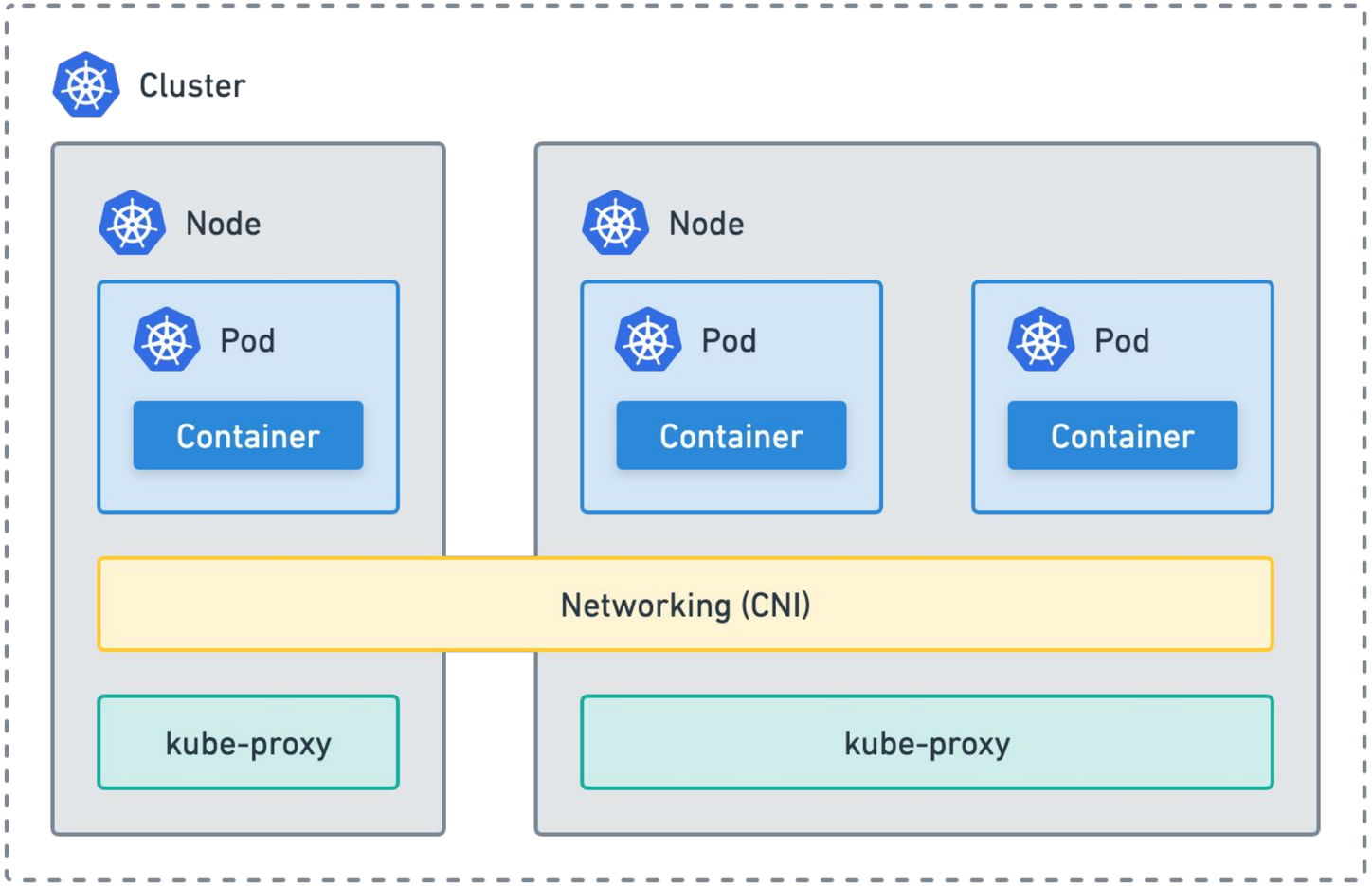
Liveness probes without network awareness

Platform team ignoring crashing CoreDNS pods

Forward Together »

CNI Chaining + kube-proxy + Ingress + CoreDNS + Service Mesh + Cloud Networking

Kubernetes Networking



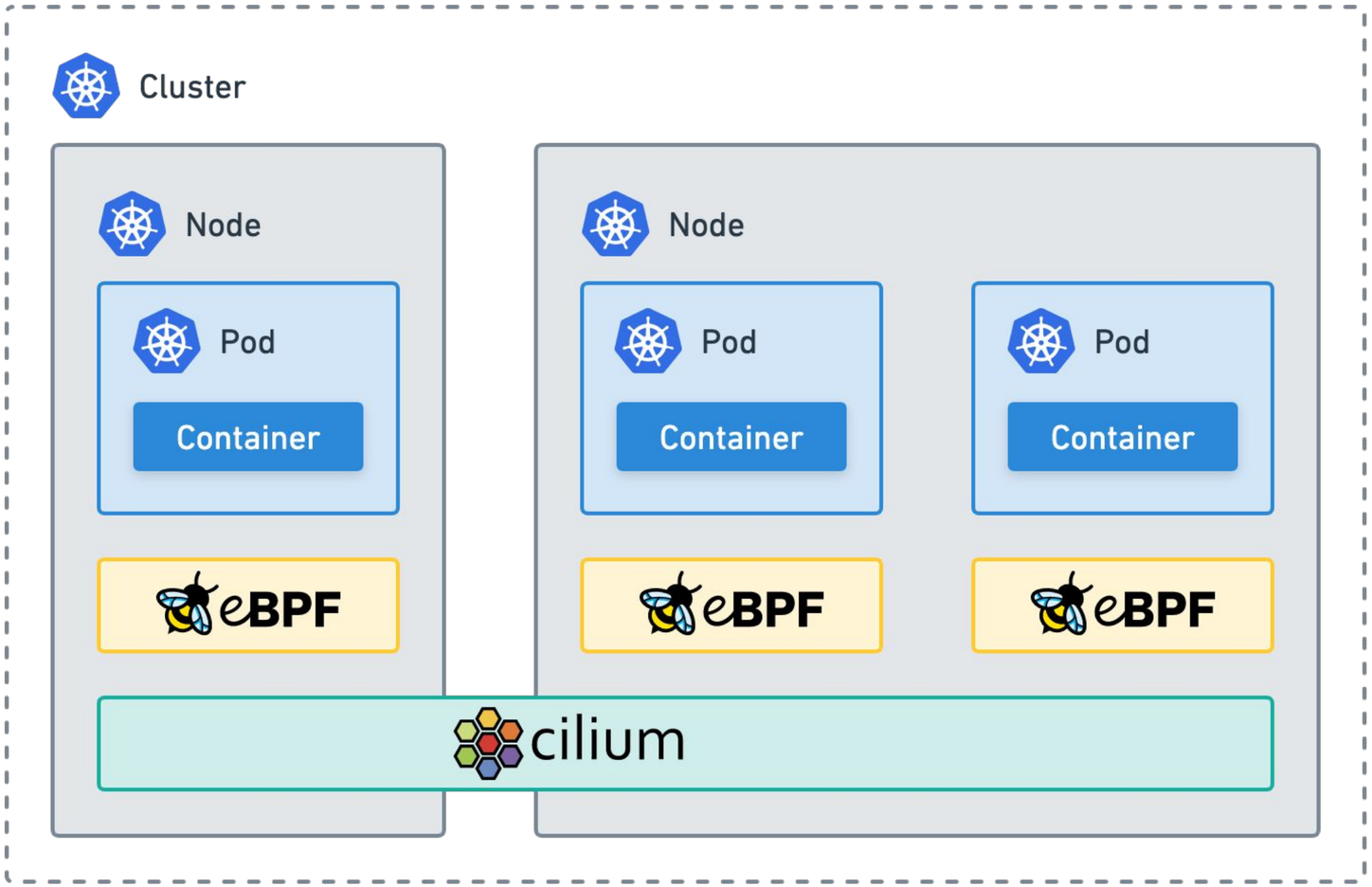
Networking plugin

- Network devices
- IP Address Management
- Intra-node connectivity
- Inter-node connectivity

Kube Proxy

- Services
- iptables or ipvs
- Service discovery

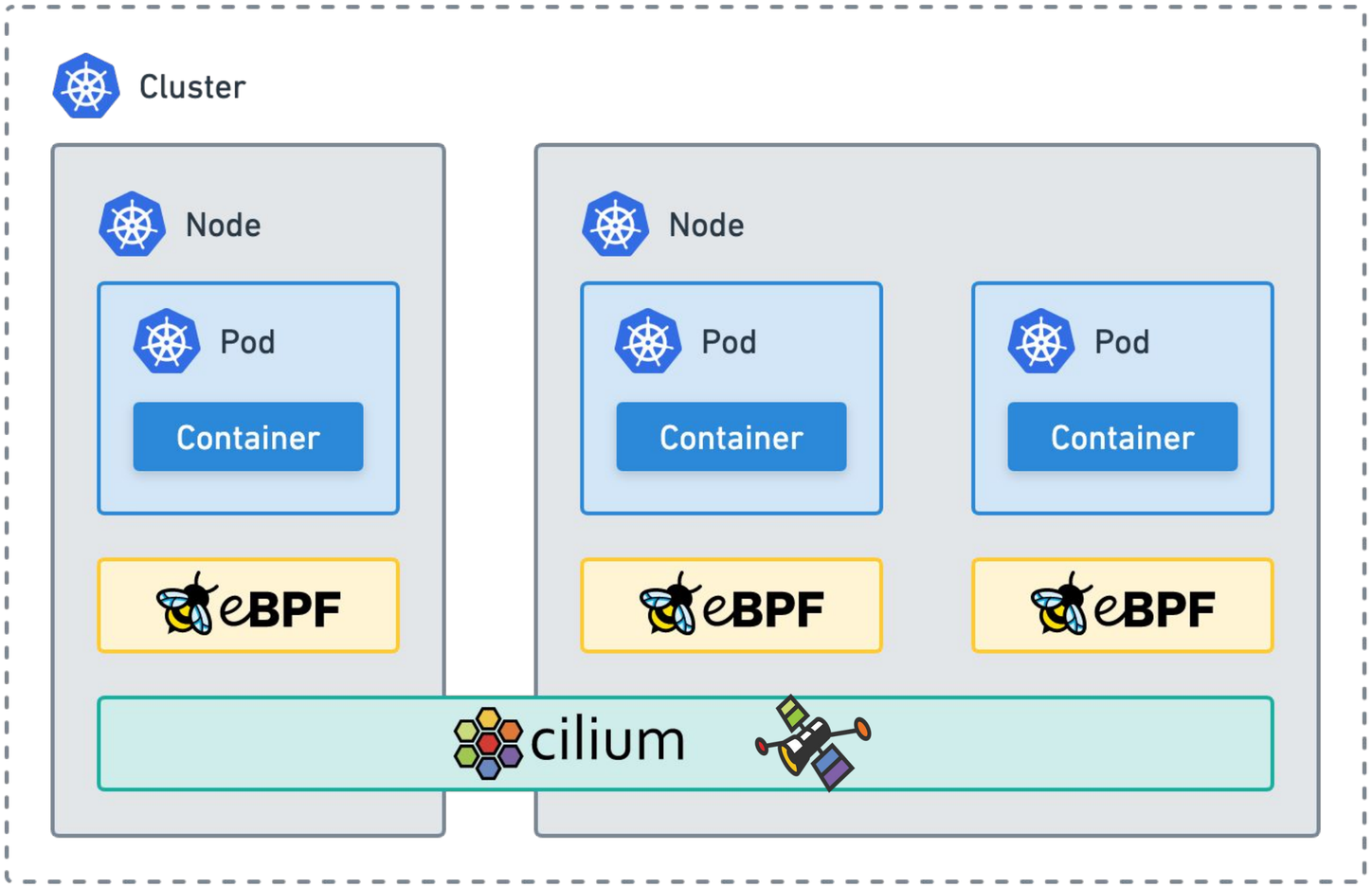
Kubernetes Networking with Cilium



The Setup we will use for these demos:

- Cilium as the CNI
- Hubble for Troubleshooting

Kubernetes Networking Observability with Hubble



Hubble

- Runs on top of Cilium
- “tcpdump for Kubernetes”
- Logs & Metrics
- Native Prometheus & Grafana Integration

Hubble

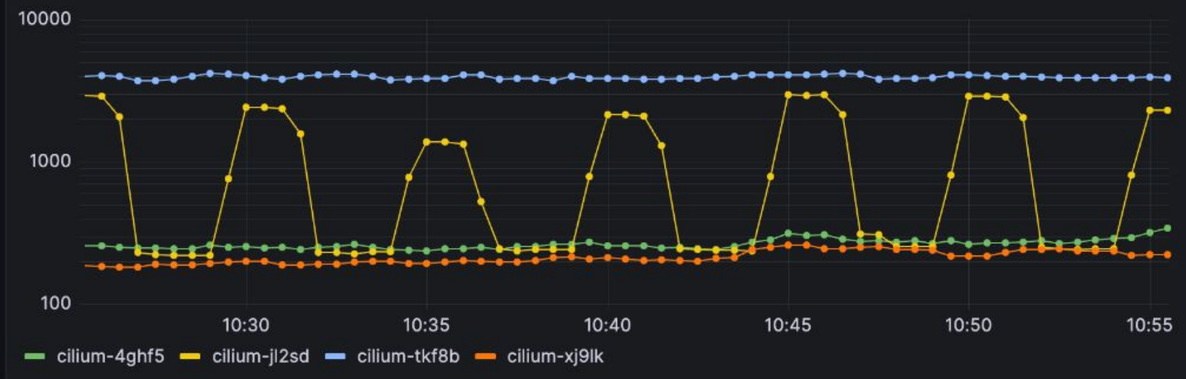
Hubble UI / Cilium / Components / L3 Policy

Last 30 minutes | 30s

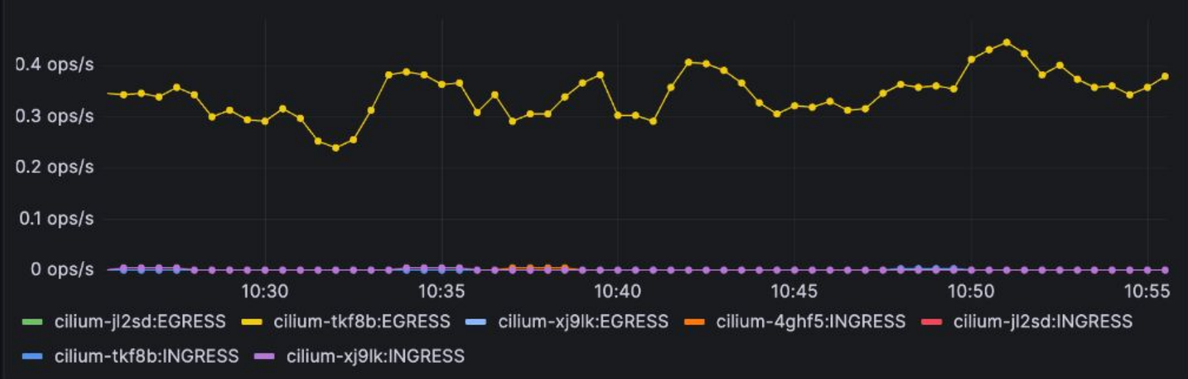
Data Source: default | cluster: All | namespace: All | pod: All | top k: 10

Cilium Overviews | Cilium Components

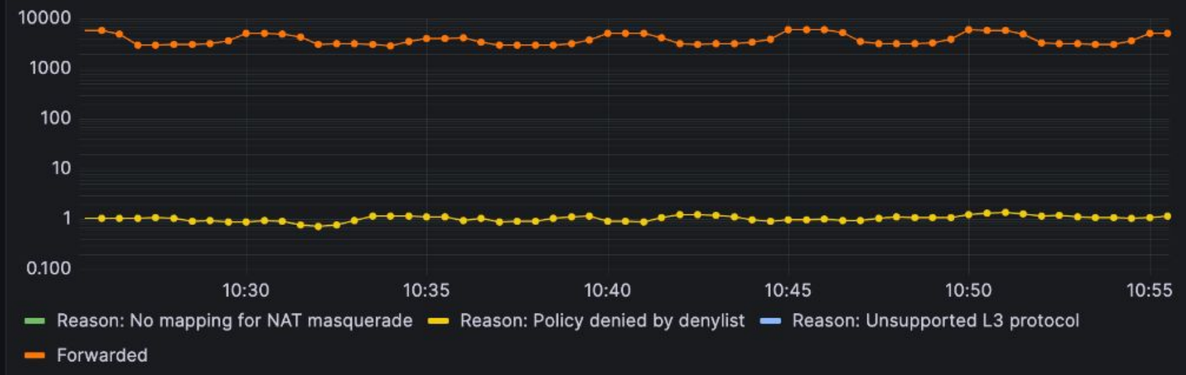
Top Pod Ingress Packet Forward Rates



Top Drop Rates

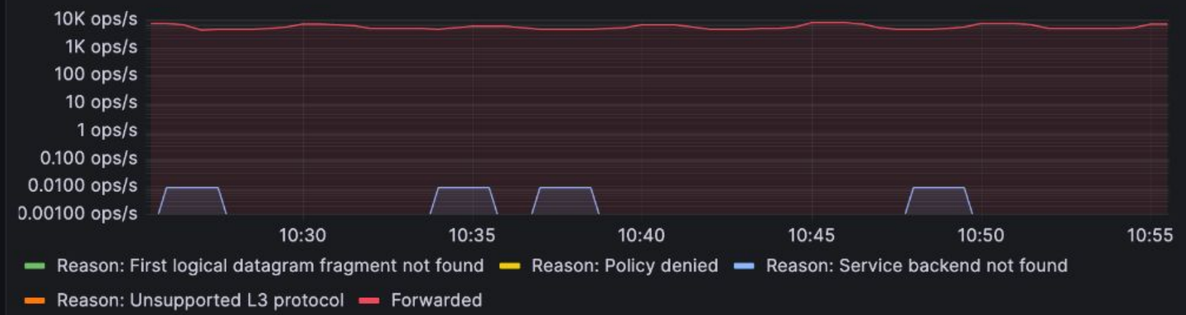


Egress Packet Drop Reasons

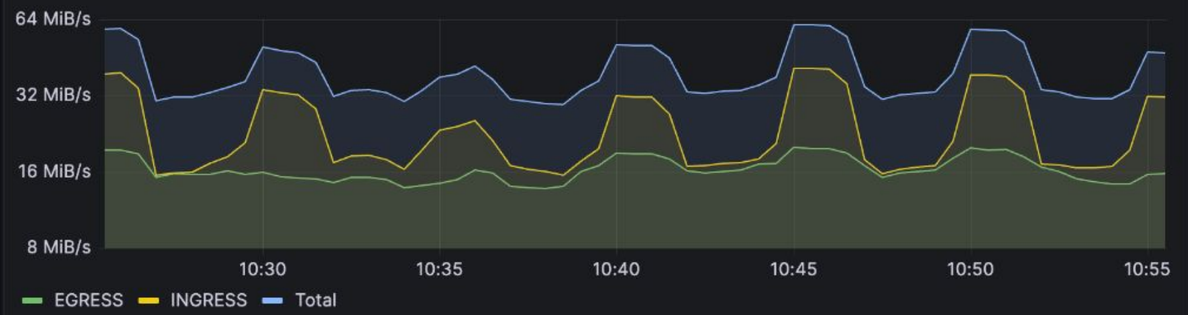


Ingress Packets by Reason

Cilium L3 Ingress Packet Counts



Total Forwarded Traffic



Live View

Namespace

Show clusterwide data

anna-otel-demo

Flows verdict

Any verdict

Forwarded

Dropped

Aggregate flows

Visual filters

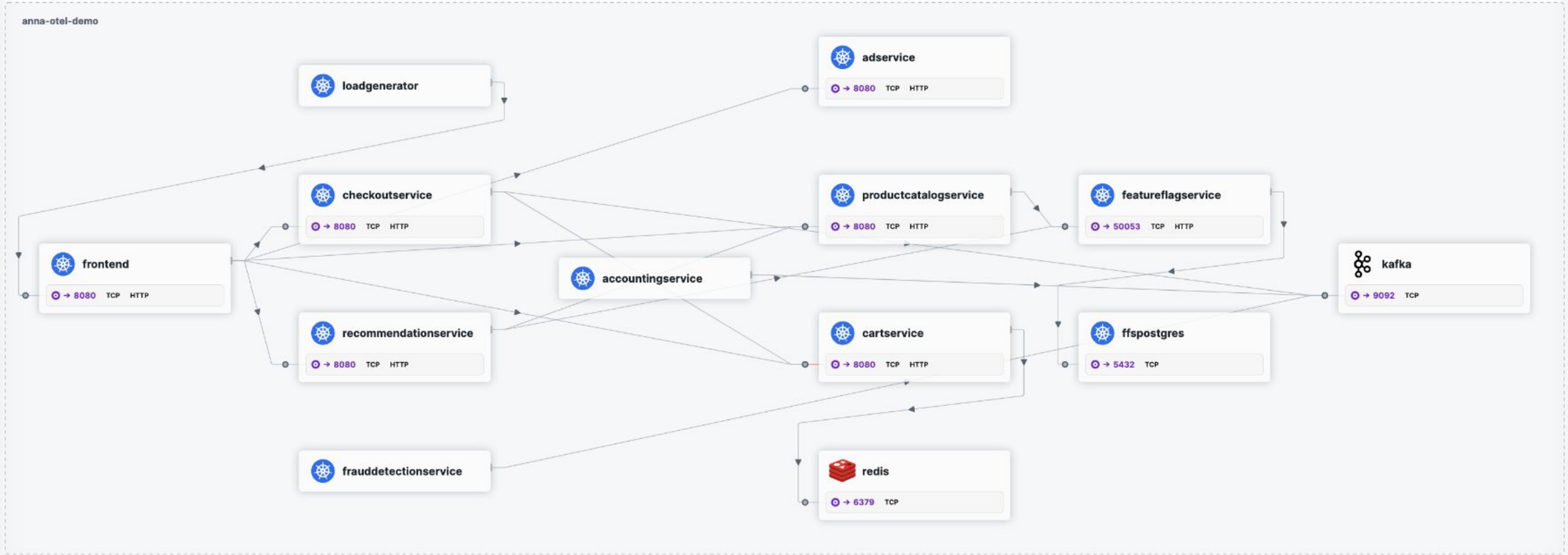
Host service

Kube-DNS:53 pod

Remote node

Prometheus app

Filter by: label key=val, ip=1.1.1.1, dns=google.com, identity=42, pod=frontend



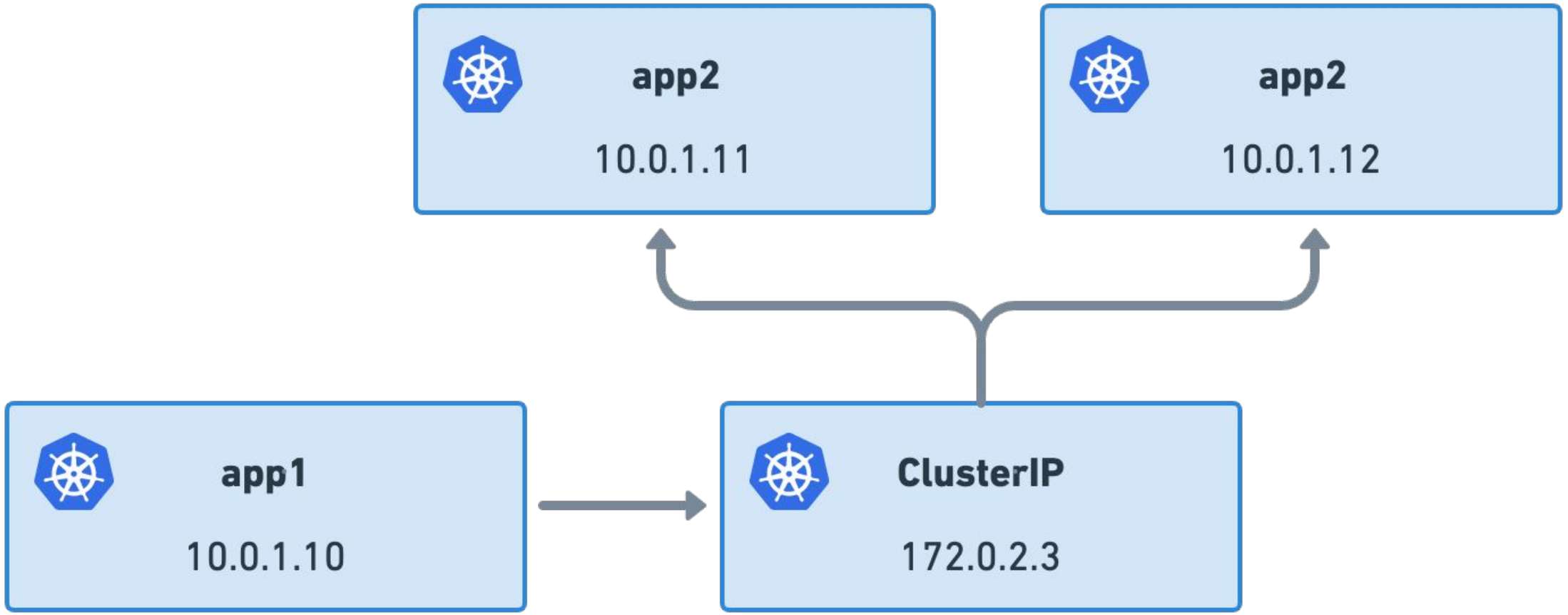
2023-04-15T08:52:22.000Z

Flows | Metrics | Services

Columns

Source Identity	Destination Identity	Destination Port	L7 info	Traffic Direction	Verdict	TCP Flags	Timestamp
loadgenerator *anna-otel-demo*	frontend *anna-otel-demo*	8080	→ GET /api/products/66VCHSJNUP 0ms	ingress	forwarded		2023/04/15 10:52:32 (+02)
checkoutservice *anna-otel-demo*	kafka *anna-otel-demo*	9092	—	egress	forwarded	ACK	2023/04/15 10:52:31 (+02)
loadgenerator *anna-otel-demo*	frontend *anna-otel-demo*	8080	→ GET / 0ms	ingress	forwarded		2023/04/15 10:52:28 (+02)
loadgenerator *anna-otel-demo*	frontend *anna-otel-demo*	8080	→ GET /api/products/L9ECAV7KIM 0ms	ingress	forwarded		2023/04/15 10:52:26 (+02)
cartservice *anna-otel-demo*	redis *anna-otel-demo*	6379	—	egress	forwarded	ACK PSH	2023/04/15 10:52:26 (+02)
productcatalogservice *anna-otel-demo*	featureflagservice *anna-otel-demo*	50053	→ POST /oteldemo.FeatureFlagService/...	ingress	forwarded		2023/04/15 10:52:25 (+02)
productcatalogservice *anna-otel-demo*	featureflagservice *anna-otel-demo*	50053	—	egress	forwarded	SYN	2023/04/15 10:52:25 (+02)
loadgenerator *anna-otel-demo*	frontend *anna-otel-demo*	8080	→ GET /api/products/OLJCESPC7Z 0ms	ingress	forwarded		2023/04/15 10:52:25 (+02)
loadgenerator *anna-otel-demo*	frontend *anna-otel-demo*	8080	→ GET /api/products/2ZYFJ3GM2N 0ms	ingress	forwarded		2023/04/15 10:52:24 (+02)
frontend *anna-otel-demo*	cartservice *anna-otel-demo*	8080	→ POST /oteldemo.CartService/GetCart...	ingress	forwarded		2023/04/15 10:52:22 (+02)
loadgenerator *anna-otel-demo*	frontend *anna-otel-demo*	8080	→ GET /api/cart 0ms	ingress	forwarded		2023/04/15 10:52:22 (+02)
frontend *anna-otel-demo*	checkoutservice *anna-otel-demo*	8080	→ POST /oteldemo.CheckoutService/Pla...	ingress	forwarded		2023/04/15 10:52:22 (+02)
loadgenerator *anna-otel-demo*	frontend *anna-otel-demo*	8080	→ POST /api/checkout 0ms	ingress	forwarded		2023/04/15 10:52:22 (+02)
frontend *anna-otel-demo*	cartservice *anna-otel-demo*	8080	→ POST /oteldemo.CartService/AddIte...	ingress	forwarded		2023/04/15 10:52:22 (+02)

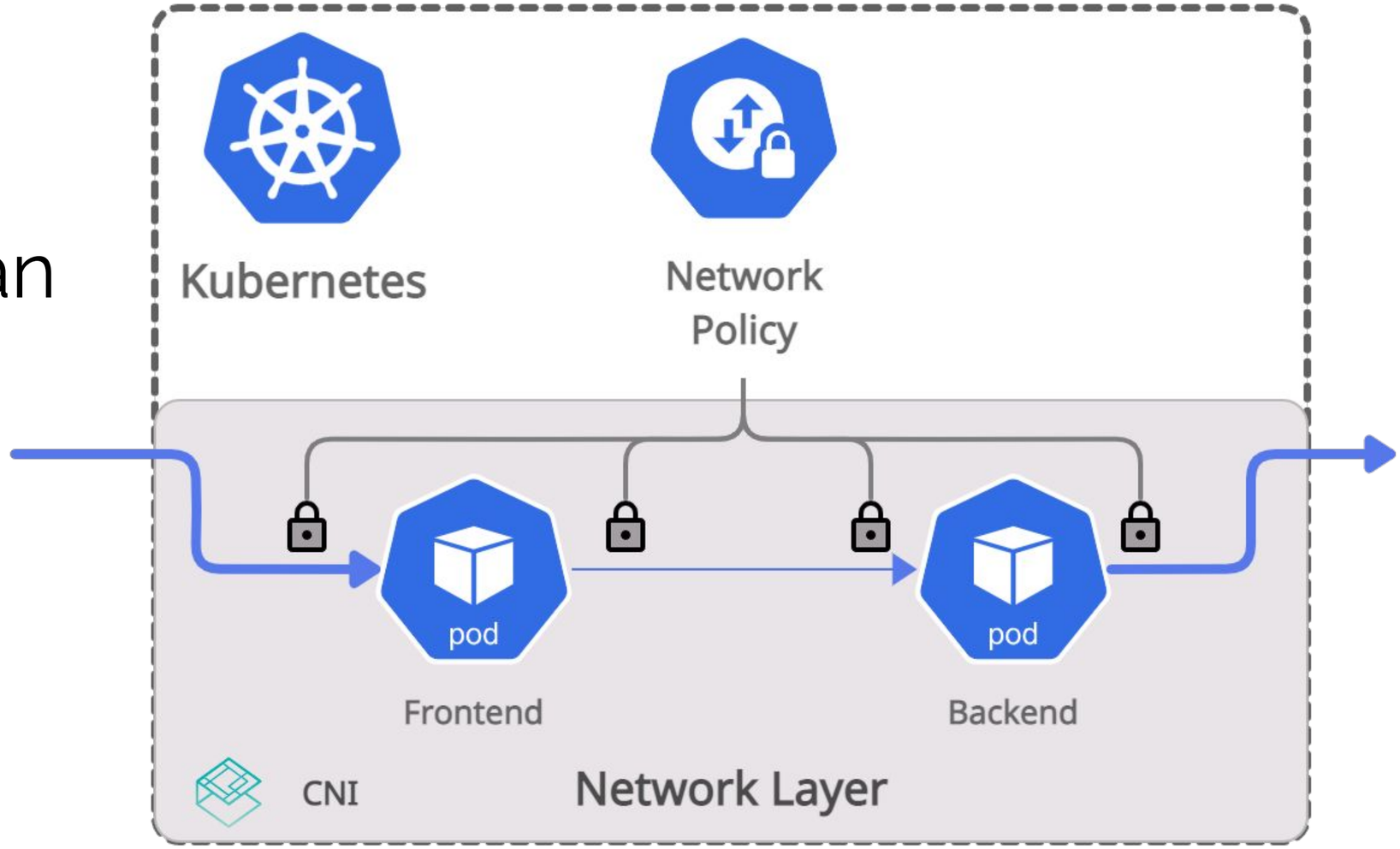
Kubernetes Services



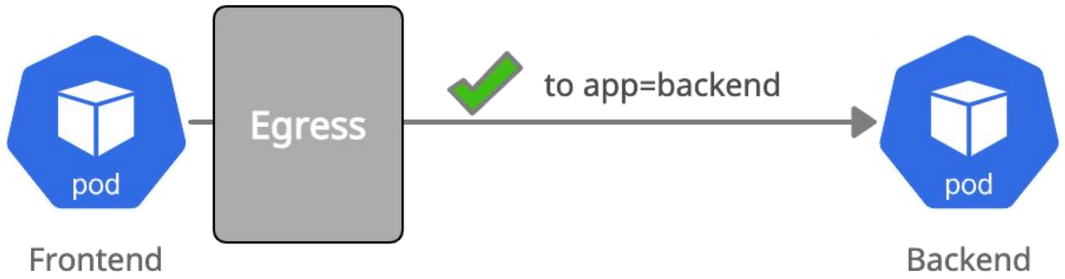
Network Policy

Network Policy

Declares who can talk to whom



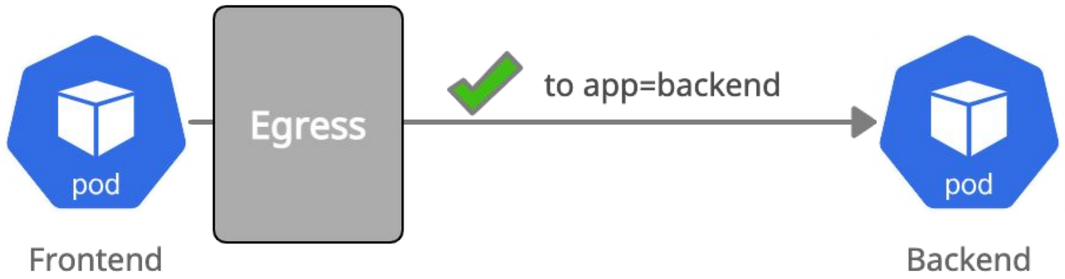
Let’s start simple



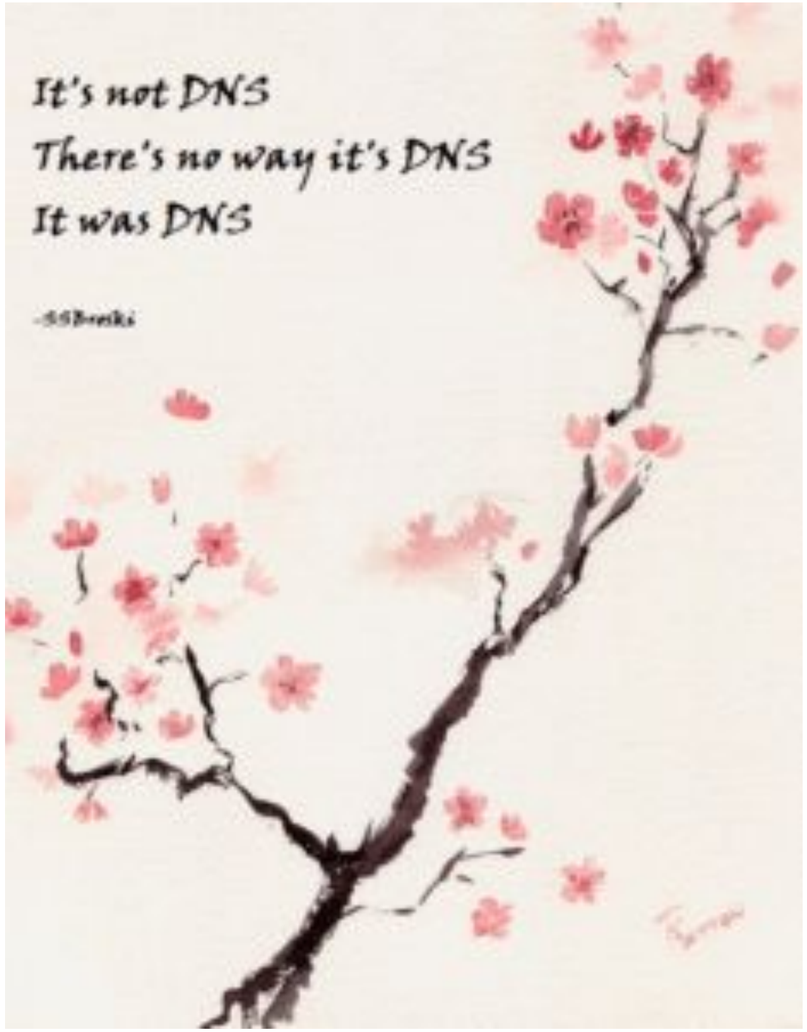

```
kind: NetworkPolicy
apiVersion: networking.k8s.io/v1
metadata:
  name: frontend-egress-allow-to-backend
spec:
  podSelector:
    matchLabels:
      app: frontend
  egress:
    - to:
      - podSelector:
          matchLabels:
            app: backend
```

Demo:
Let’s build this

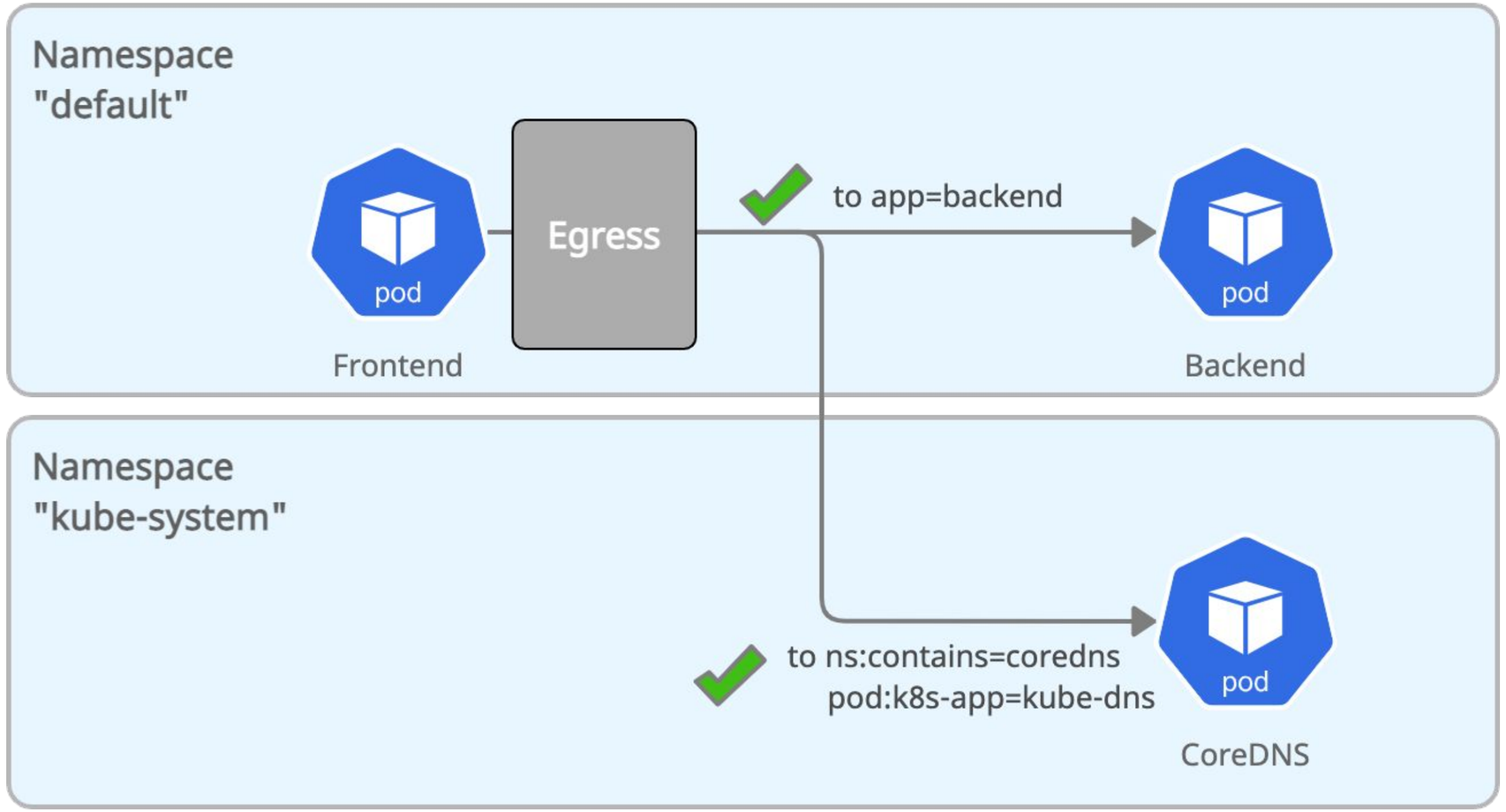
Let’s start simple




```
kind: NetworkPolicy
apiVersion: networking.k8s.io/v1
metadata:
  name: frontend-egress-allow-to-backend
spec:
  podSelector:
    matchLabels:
      app: frontend
  egress:
    - to:
      - podSelector:
          matchLabels:
            app: backend
```



Let’s start simple (fixed)



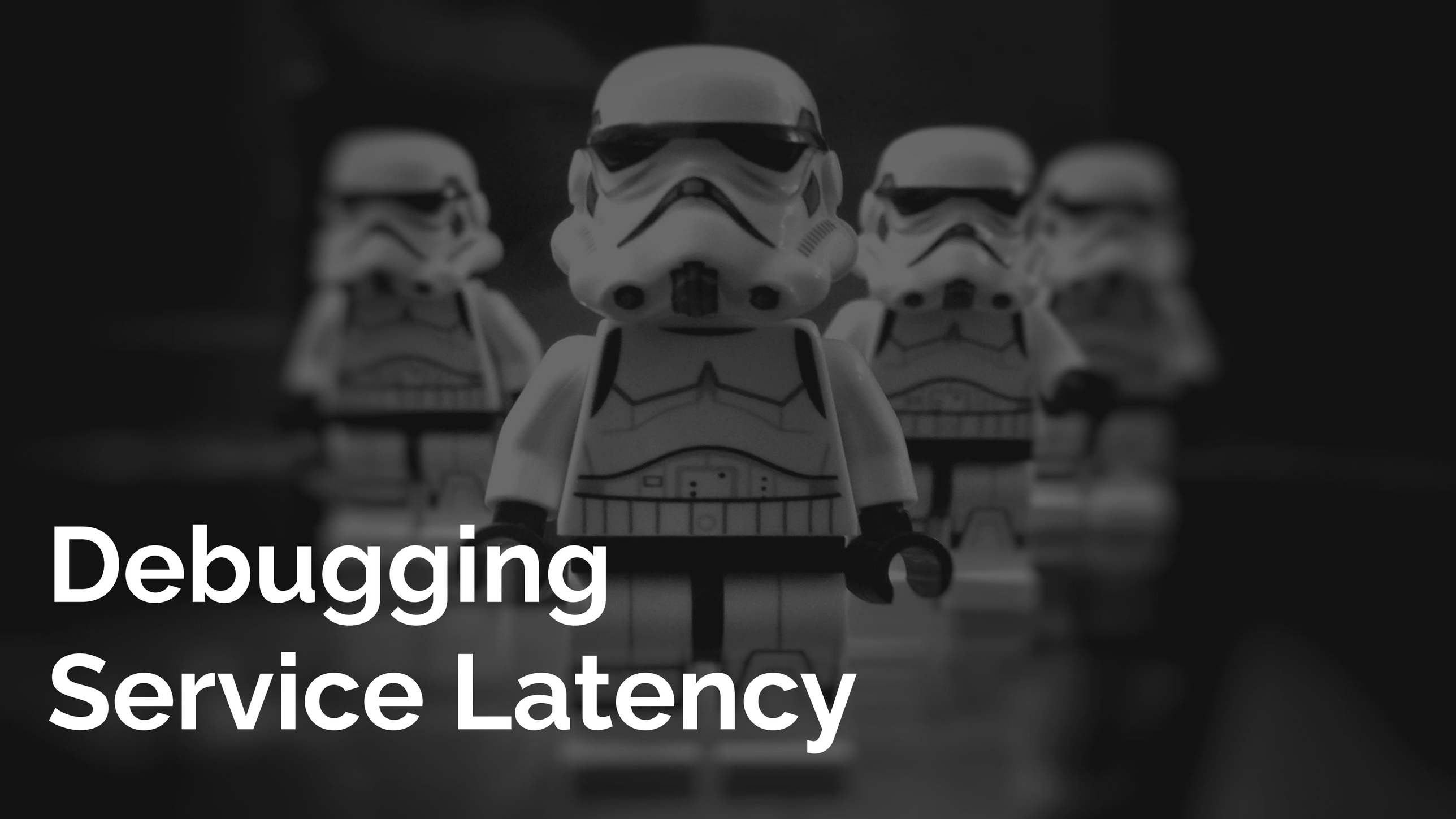
DNS

Kubernetes DNS

- Used for service discovery
- Usually implemented with CoreDNS
- Looks Simple...

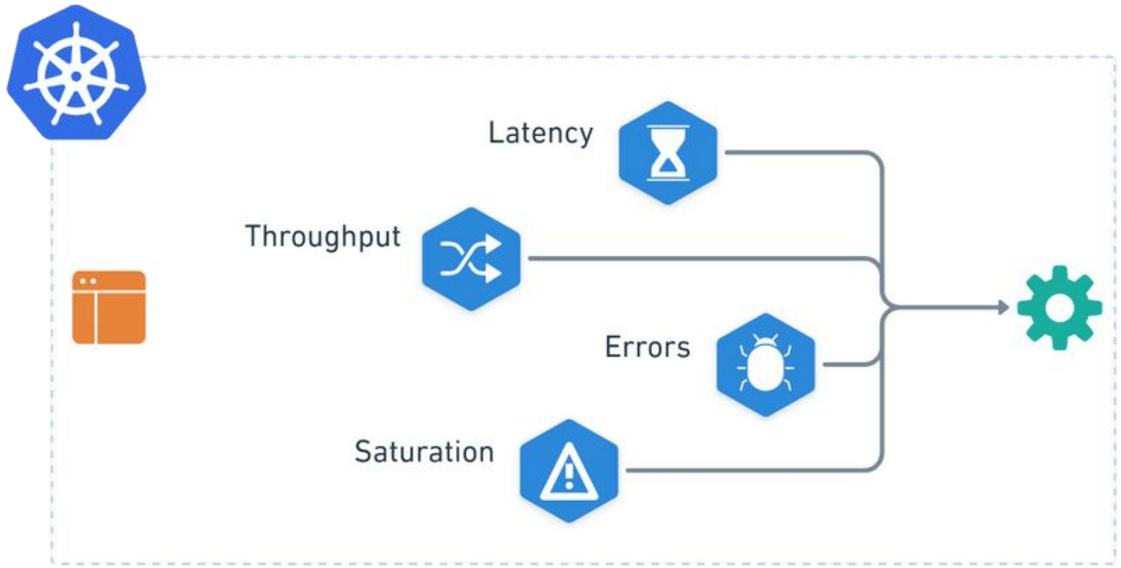
How to Troubleshoot Kubernetes DNS?

Demo:
Let’s have a look



Debugging Service Latency

Golden Signal Dashboard



- The Four Signals
  - **Latency**
  - **Traffic**
  - **Errors**
  - **Saturation**

- More information: https://sre.google/sre-book/monitoring-distributed-systems/

Hubble / Hubble HTTP Golden Signals by W...

Last 6 hours | 30s

Prometheus: Prometheus | Cluster: None | Destination Namespace: otel-demo | Destination Workload: otel-demo-featureflagservice | Reporter: server | Source Namespace: All | Source Workload: All

General

Incoming Request Volume

0.800 req/s

Incoming Request Success Rate (non-5xx responses)

114%

Request Duration

500 ms
0 s
03:00 04:00 05:00 06:00 07:00 08:00

P50 Last *: 15.7 ms P95 Last *: 41.5 ms P99 Last *: 118 ms

Requests by Source

Incoming Requests by Source and Response Code

1 req/s
0.8 req/s
0.6 req/s
0.4 req/s
0.2 req/s
0 req/s
03:00 03:30 04:00 04:30 05:00 05:30 06:00 06:30 07:00 07:30 08:00 08:30

	Max	Mean	Total	Last *
otel-demo/otel-demo-adservice: 200	0.0730 req/s	0.0218 req/s	15.7 req/s	0.0360 req/s
otel-demo/otel-demo-frontendproxy: 200	0 req/s	0 req/s	0 req/s	0 req/s
otel-demo/otel-demo-frontendproxy: 302	0 req/s	0 req/s	0 req/s	0 req/s
otel-demo/otel-demo-frontendproxy: 404	0 req/s	0 req/s	0 req/s	0 req/s

Incoming Request Success Rate (non-5xx responses) By Source

800%
600%
400%
200%
0%
03:00 03:30 04:00 04:30 05:00 05:30 06:00 06:30 07:00 07:30 08:00 08:30

	Mean	Min	Max	Last *
otel-demo/otel-demo-adservice	NaN	9.09%	∞%	36.4%
otel-demo/otel-demo-frontendproxy	NaN			NaN
otel-demo/otel-demo-productcatalogservice	∞%	14.7%	∞%	182%
otel-demo/otel-demo-recommendationservice	∞%	24.7%	∞%	103%

HTTP Request Duration by Source

1 s
800 ms
600 ms
400 ms
200 ms
0 s
03:00 03:30 04:00 04:30 05:00 05:30 06:00 06:30 07:00 07:30 08:00 08:30

CPU Usage by Source

0.03
0.02
0.01
0
03:00 03:30 04:00 04:30 05:00 05:30 06:00 06:30 07:00 07:30 08:00 08:30

ISOVALENT

Session Feedback

Thank you!

Contact:
@tgraf__

Learn about Cilium:
cilium.io