ISOVALENT

Cilium Gateway API

The New L7 Kubernetes Ingress Standard

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

Solutions Architect, Isovalent | CNCF Ambassador

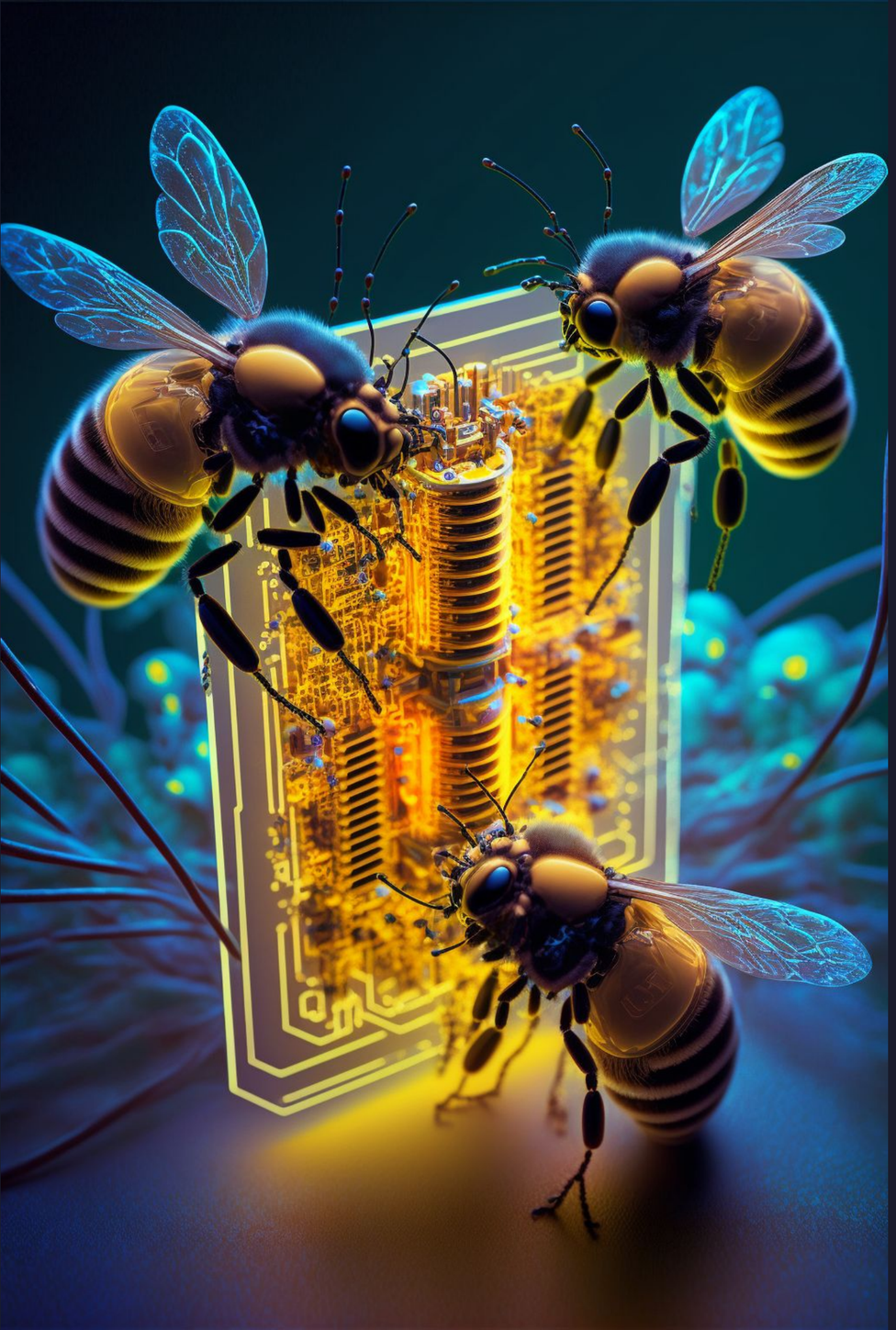
ISOVALENT

Who am I

Raphaël Pinson

Solutions Architect @ Isovalent

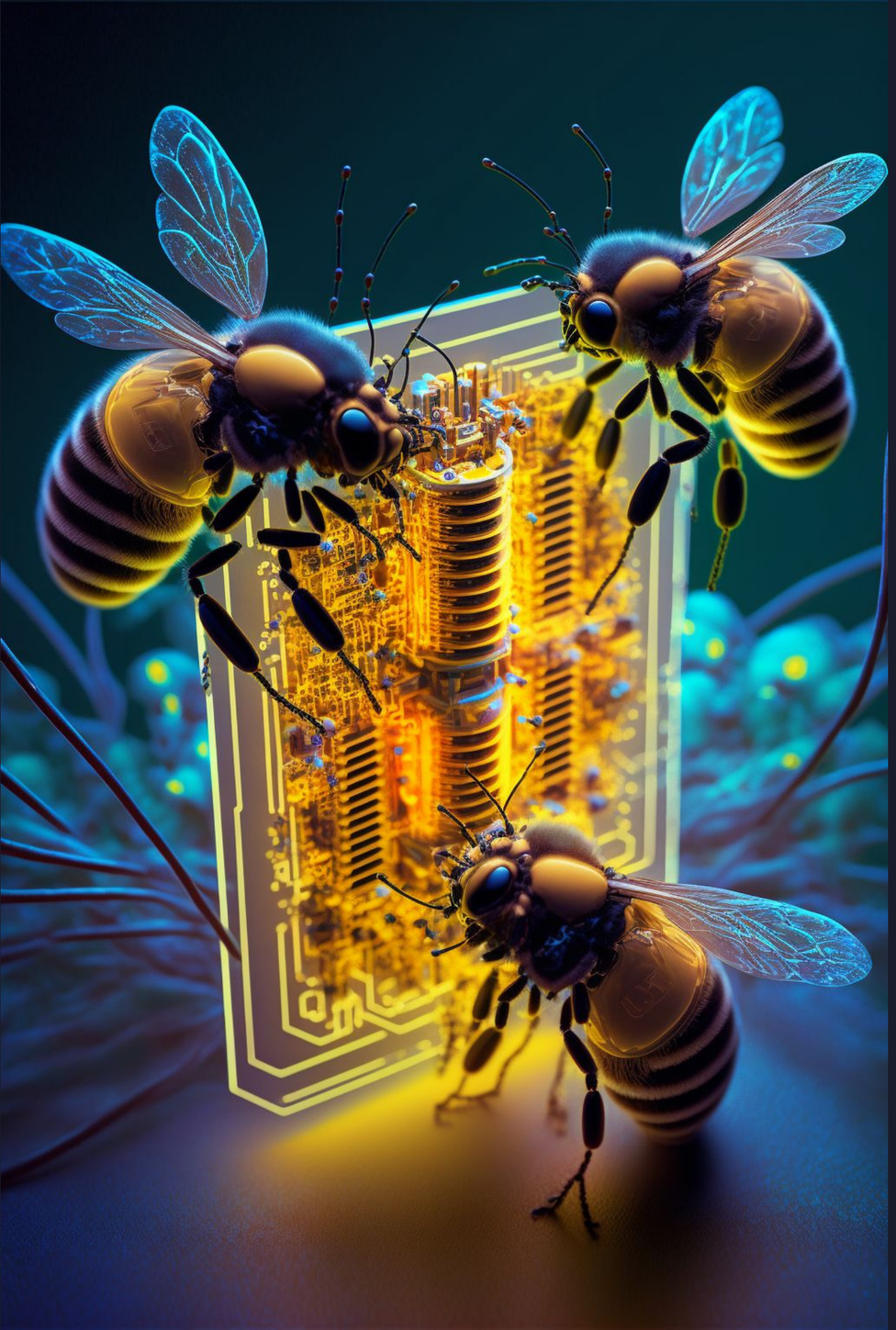
CNCF Ambassador



Cilium Gateway API

The New L7 Kubernetes Ingress Standard

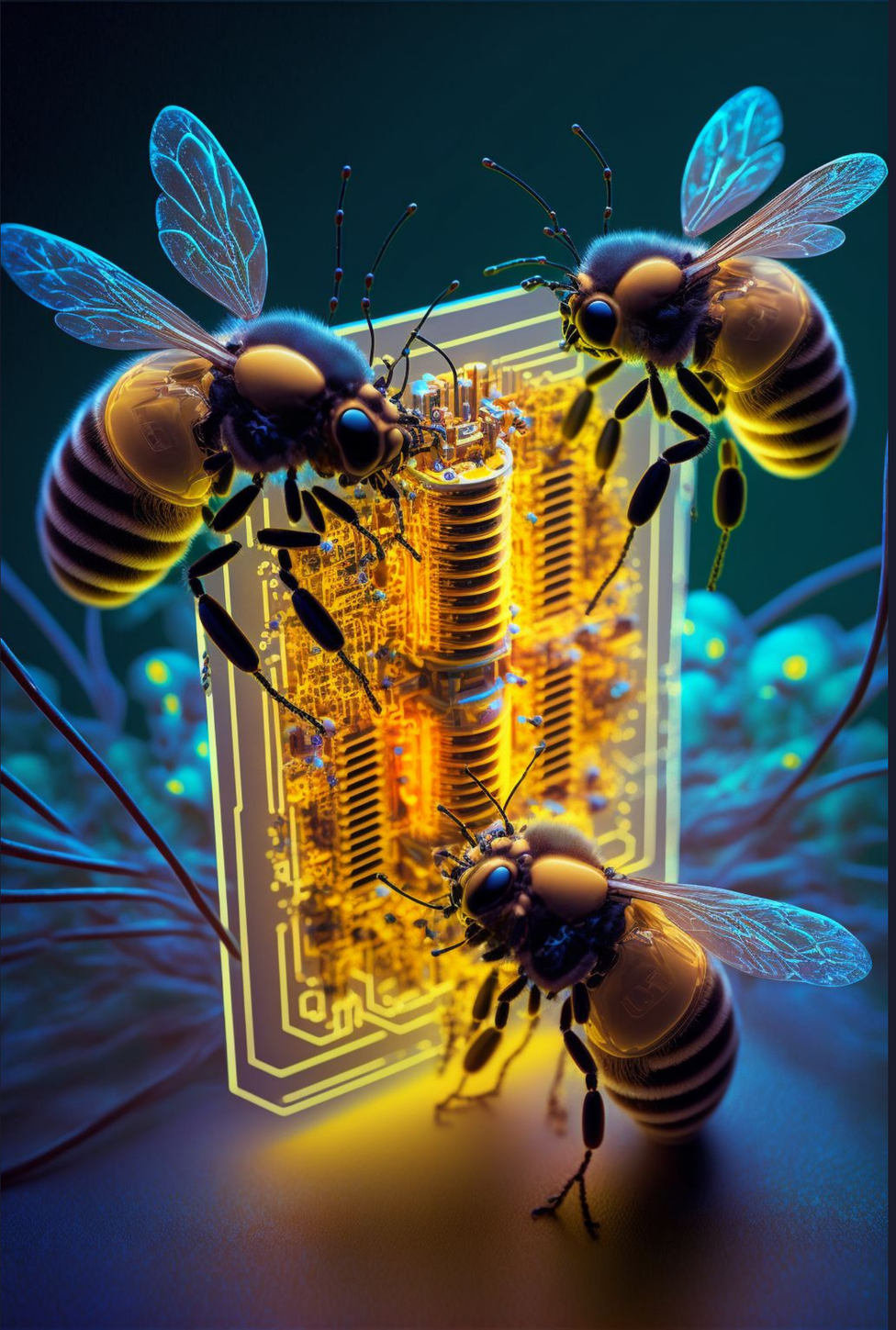
- **Cilium & eBPF**



Cilium Gateway API

The New L7 Kubernetes Ingress Standard

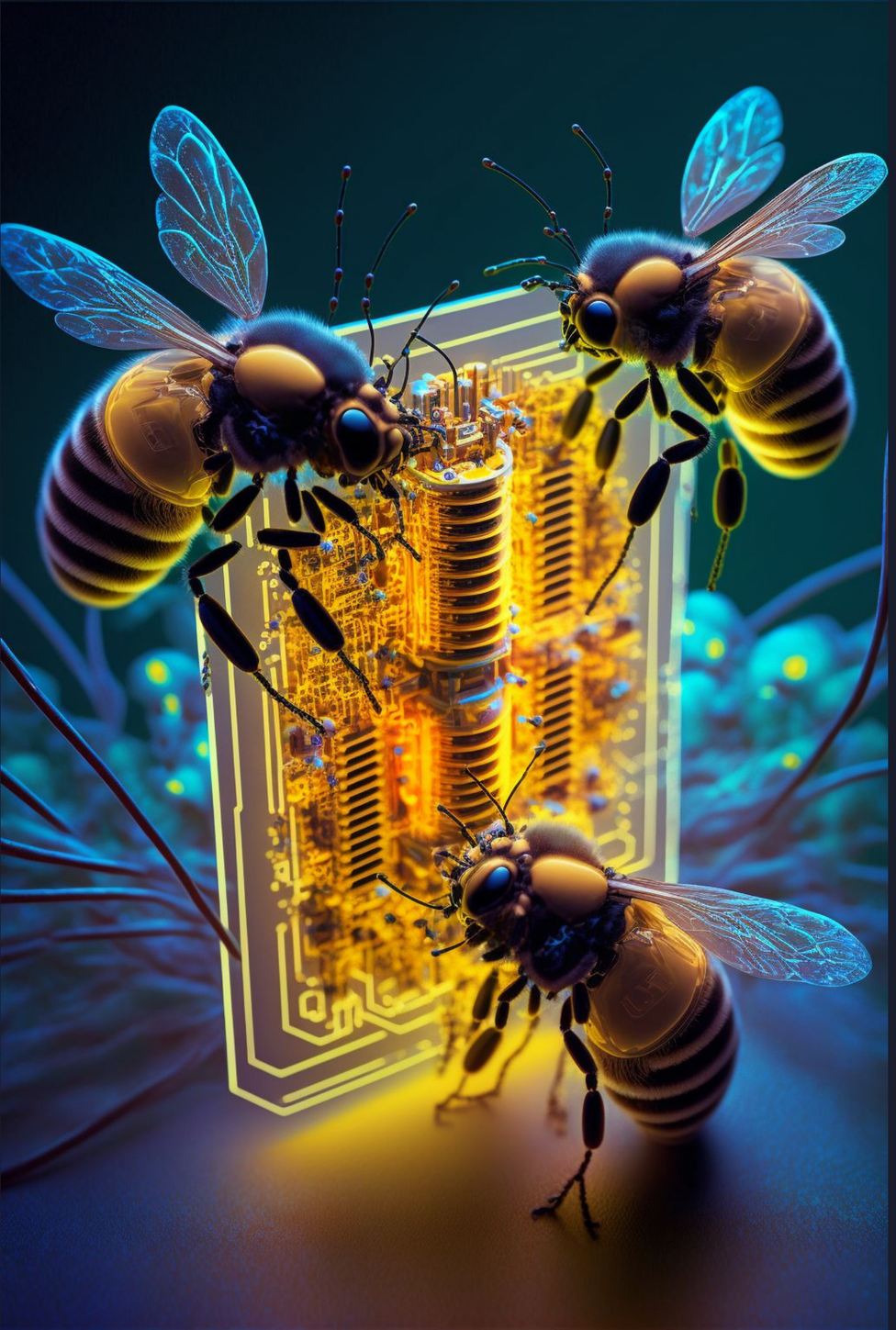
- Cilium & eBPF
- Kubernetes Services



Cilium Gateway API

The New L7 Kubernetes Ingress Standard

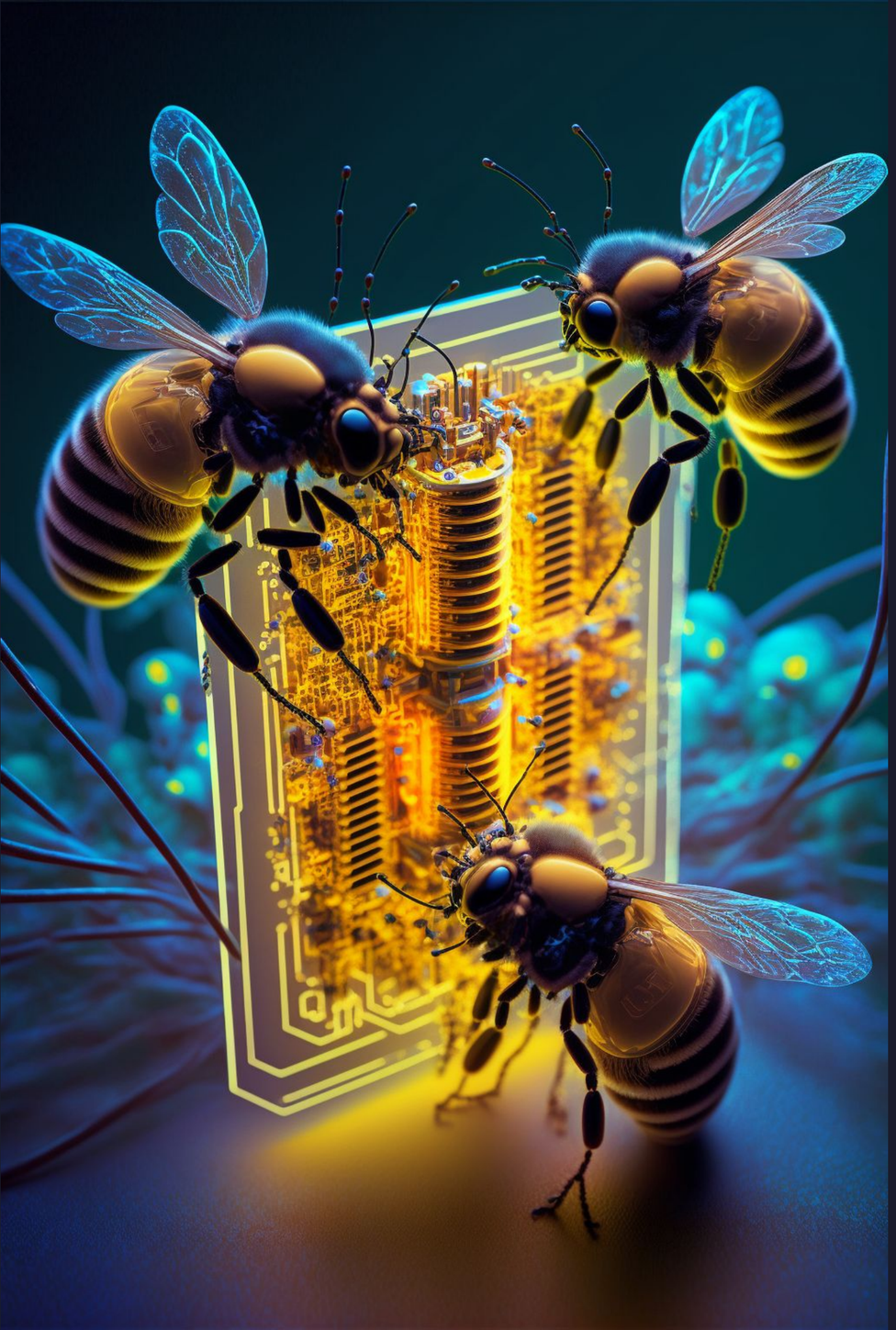
- Cilium & eBPF
- Kubernetes Services
- Ingress



Cilium Gateway API

The New L7 Kubernetes Ingress Standard

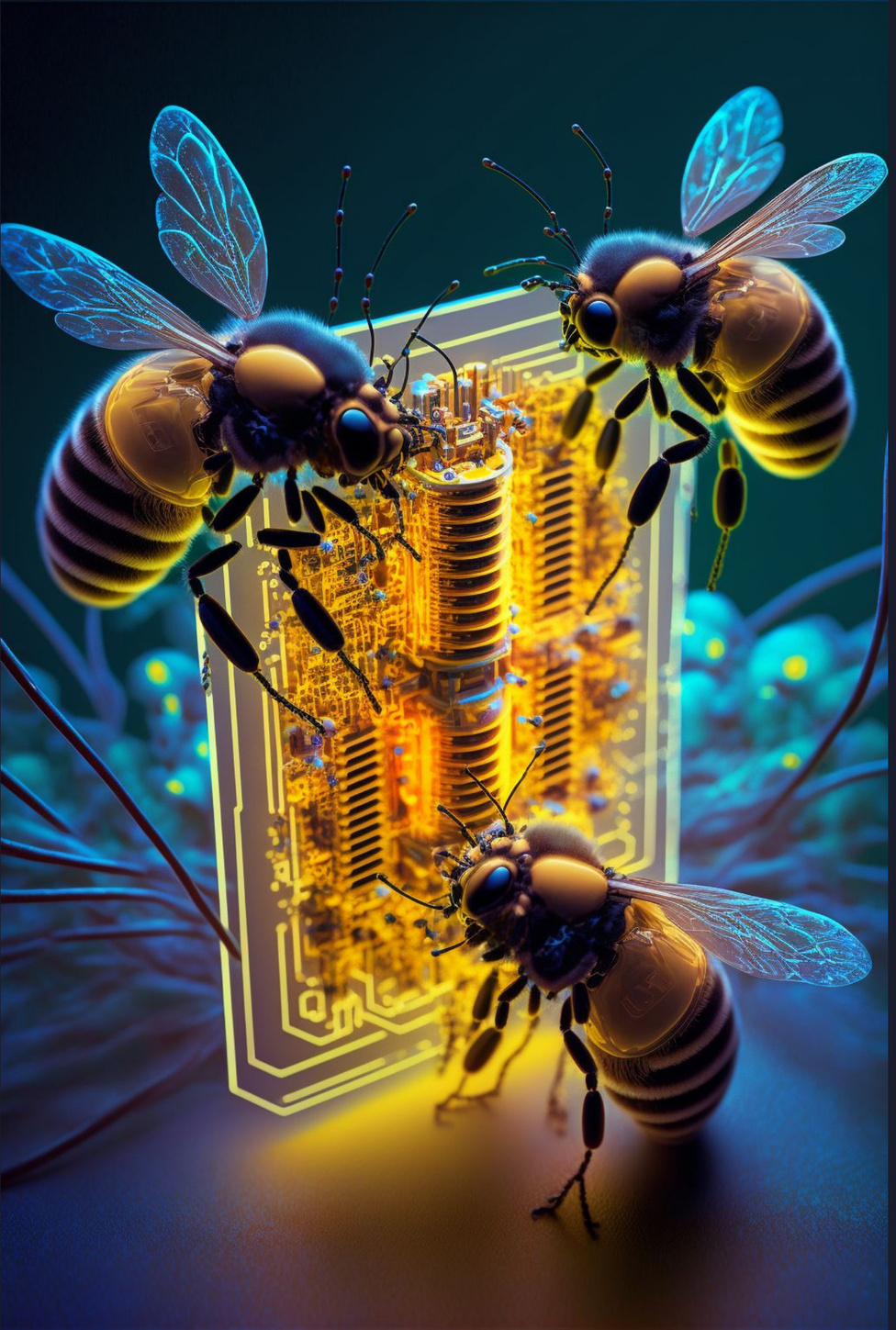
- Cilium & eBPF
- Kubernetes Services
- Ingress
- Cilium Envoy Config



Cilium Gateway API

The New L7 Kubernetes Ingress Standard

- Cilium & eBPF
- Kubernetes Services
- Ingress
- Cilium Envoy Config
- Gateway API



Cilium Gateway API

The New L7 Kubernetes Ingress Standard

- Cilium & eBPF
- Kubernetes Services
- Ingress
- Cilium Envoy Config
- Gateway API
- Cilium Service Mesh

Cilium Gateway API

The New L7 Kubernetes Ingress Standard

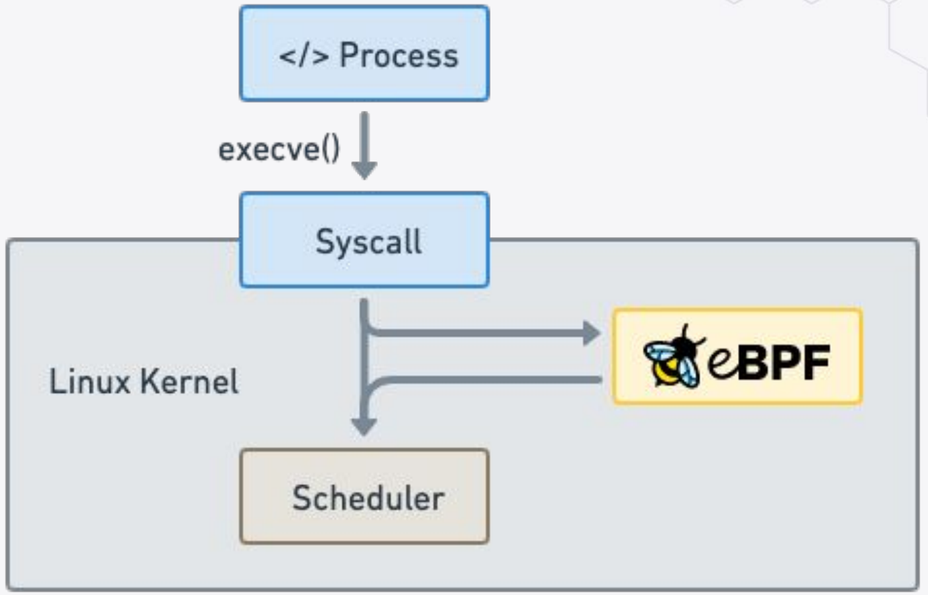
- **Cilium & eBPF**
- Kubernetes Services
- Ingress
- Cilium Envoy Config
- Gateway API
- Cilium Service Mesh


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”

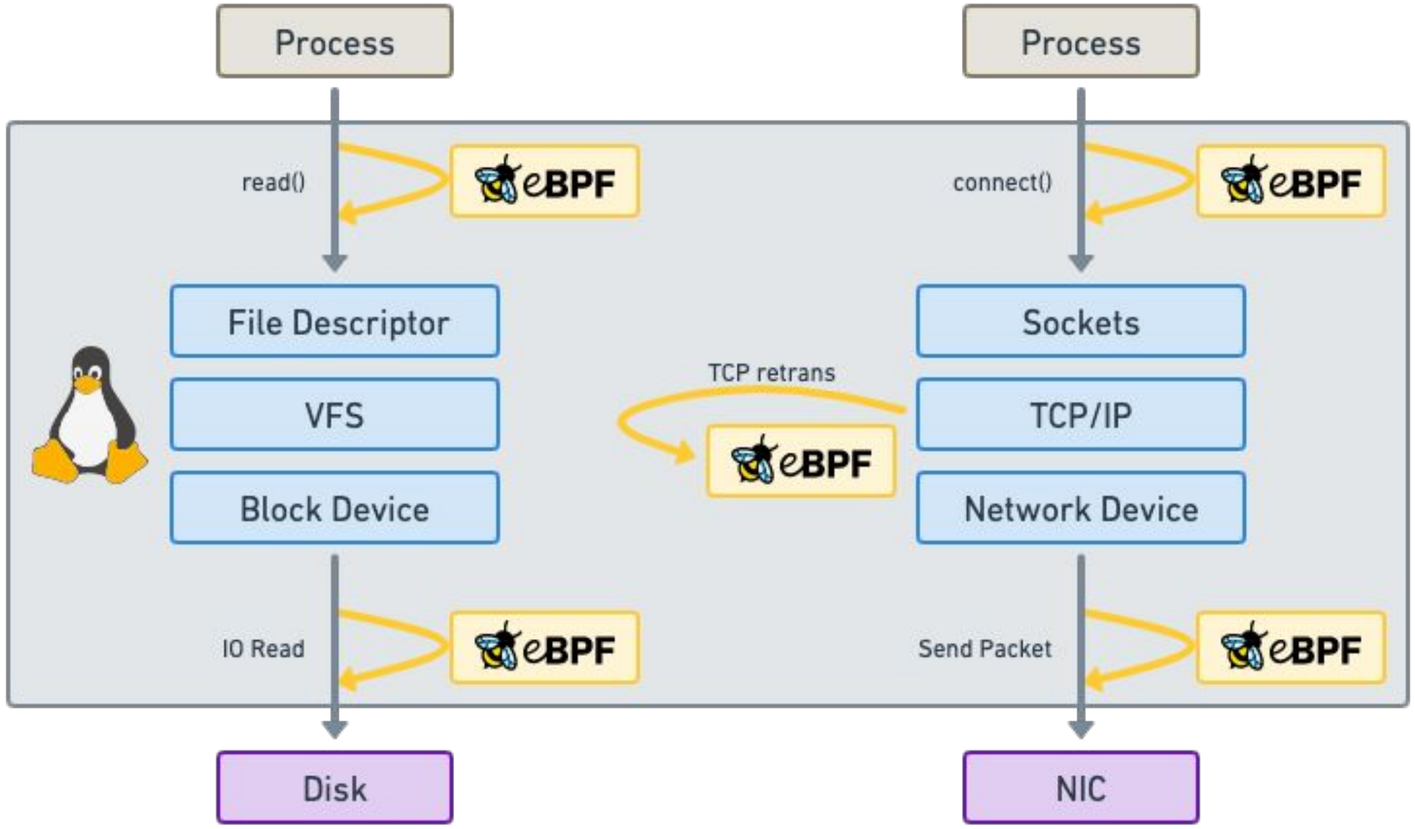



```c
int syscall__ret_execve(struct pt_regs *ctx)
{
        struct comm_event event = {
                .pid = bpf_get_current_pid_tgid() >> 32,
                .type = TYPE_RETURN,
        };

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

        return 0;
}
```

Run eBPF programs on events



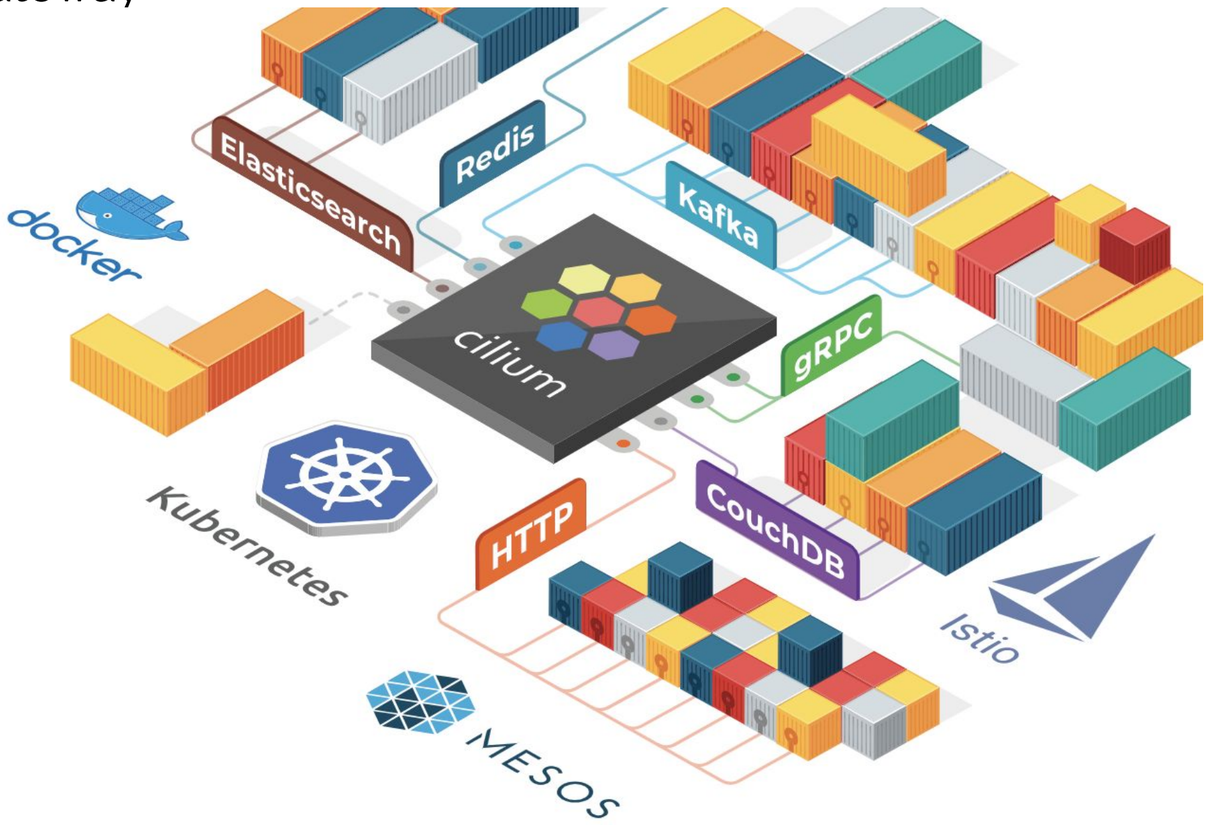
Attachment points

- 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



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

Networking: Multi-Cluster | 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

Building a Global Multi Cluster Gaming Infrastructure with Cilium

---

What Makes a Good Multi-tenant Kubernetes Solution

Building a Secure and Maintainable PaaS

Building High-Performance Cloud-Native Pod Networks

Scaling a Multi-Tenant k8s Cluster in a Telco

First step towards cloud native networking

Cloud Native Networking with eBPF

Managed Kubernetes: 1.5 Years of Cilium Usage at DigitalOcean

---

Google chooses Cilium for Google Kubernetes Engine (GKE) networking

---

Why eBPF is changing the Telco networking space?

Kubernetes Network Policies in Action with Cilium

AWS picks Cilium for Networking & Security on EKS Anywhere

Scaleway uses Cilium as the default CNI for Kubernetes Kapsule

---

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

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

---

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

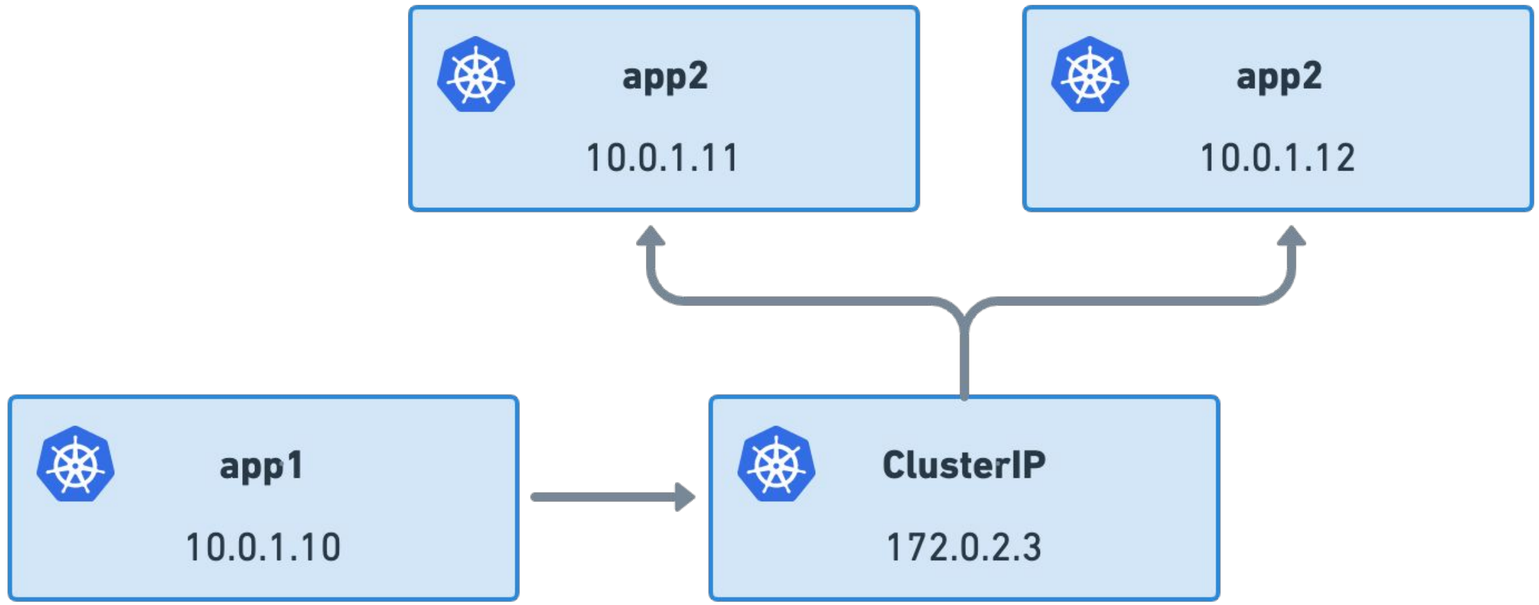
ISOVALENT

Cilium Gateway API

The New L7 Kubernetes Ingress Standard

- Cilium & eBPF
- **Kubernetes Services**
- Ingress
- Cilium Envoy Config
- Gateway API
- Cilium Service Mesh

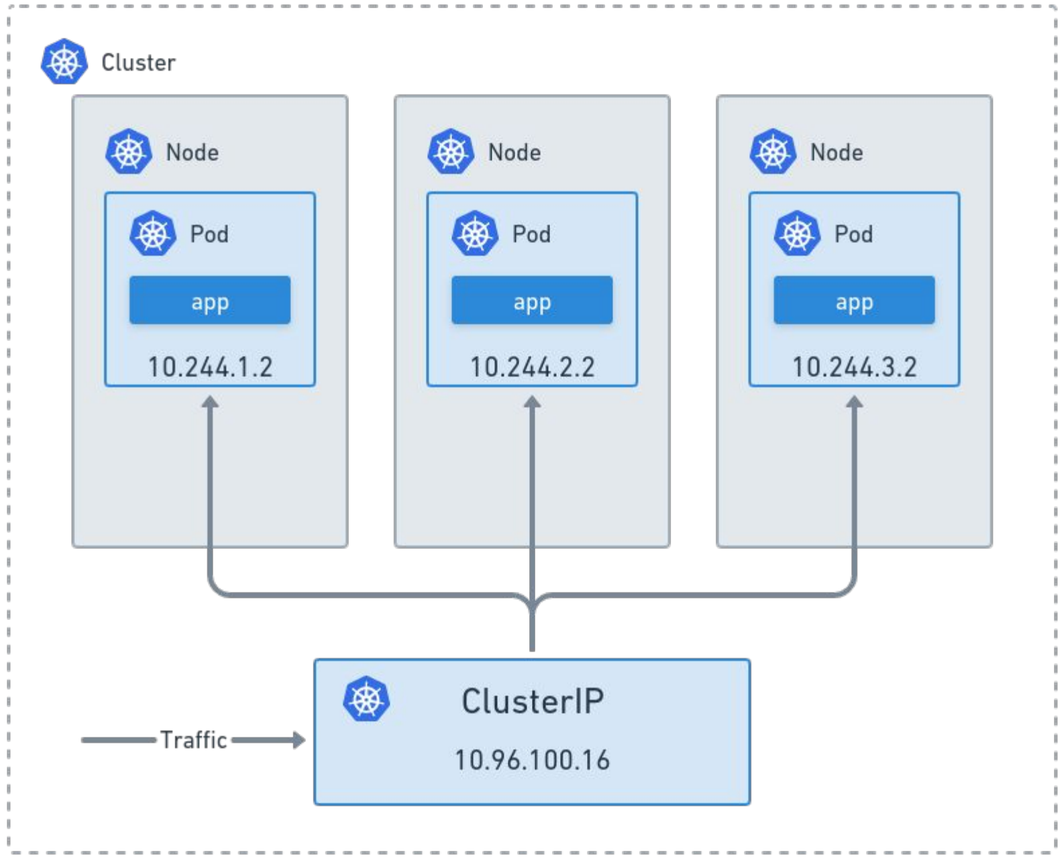
Kubernetes Services



East-west connectivity

- Durable abstraction
- Connect applications
- Ephemeral addresses
- High churn
- Iptables or ipvs

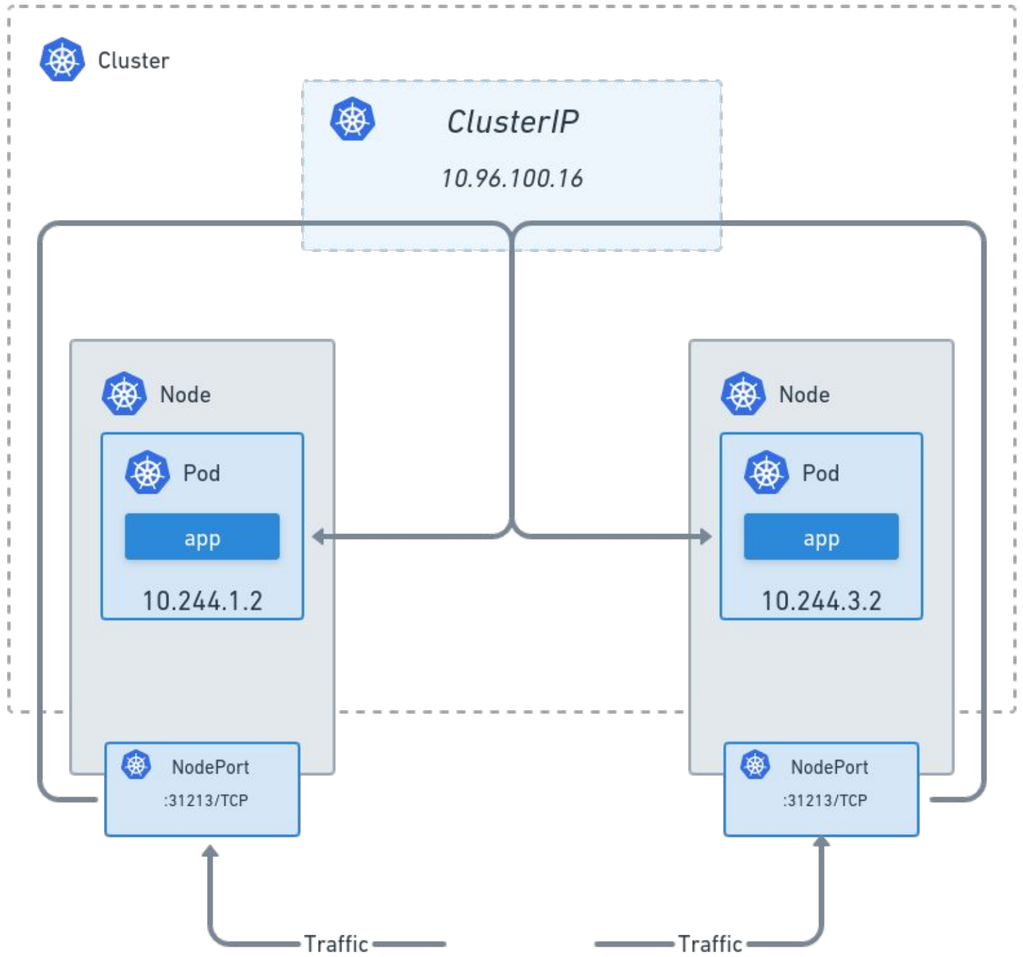
Service: ClusterIP



East-West connectivity

- Exposes the Service on a cluster-internal IP
- Only reachable from within the cluster

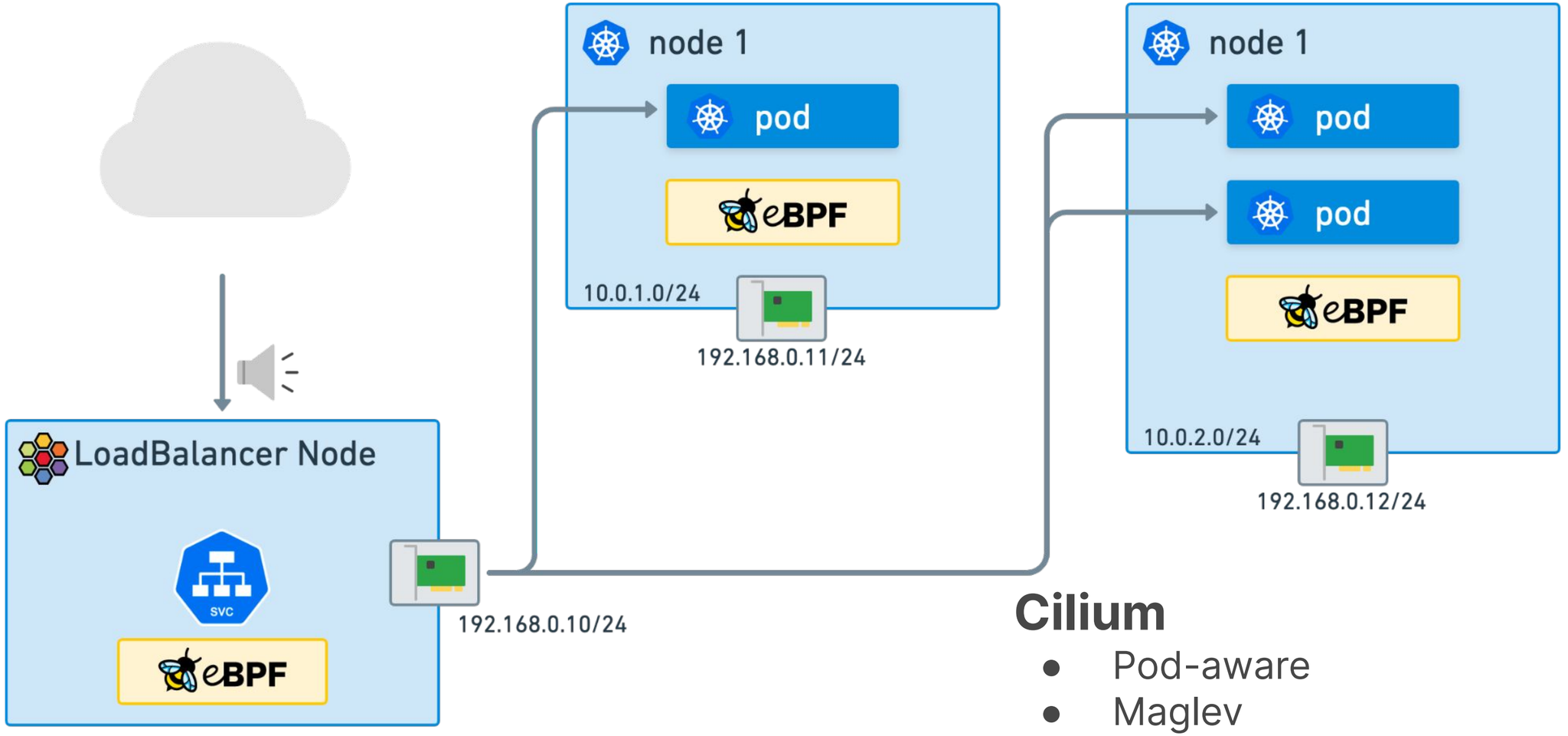
Service: NodePort



North-South connectivity

- Exposes the Service on each Node's IP at a static port
- Kubernetes automatically sets up a cluster IP too

Load Balancing



Cilium

- Pod-aware
- Maglev
- Standalone or distributed

Cilium Gateway API

The New L7 Kubernetes Ingress Standard

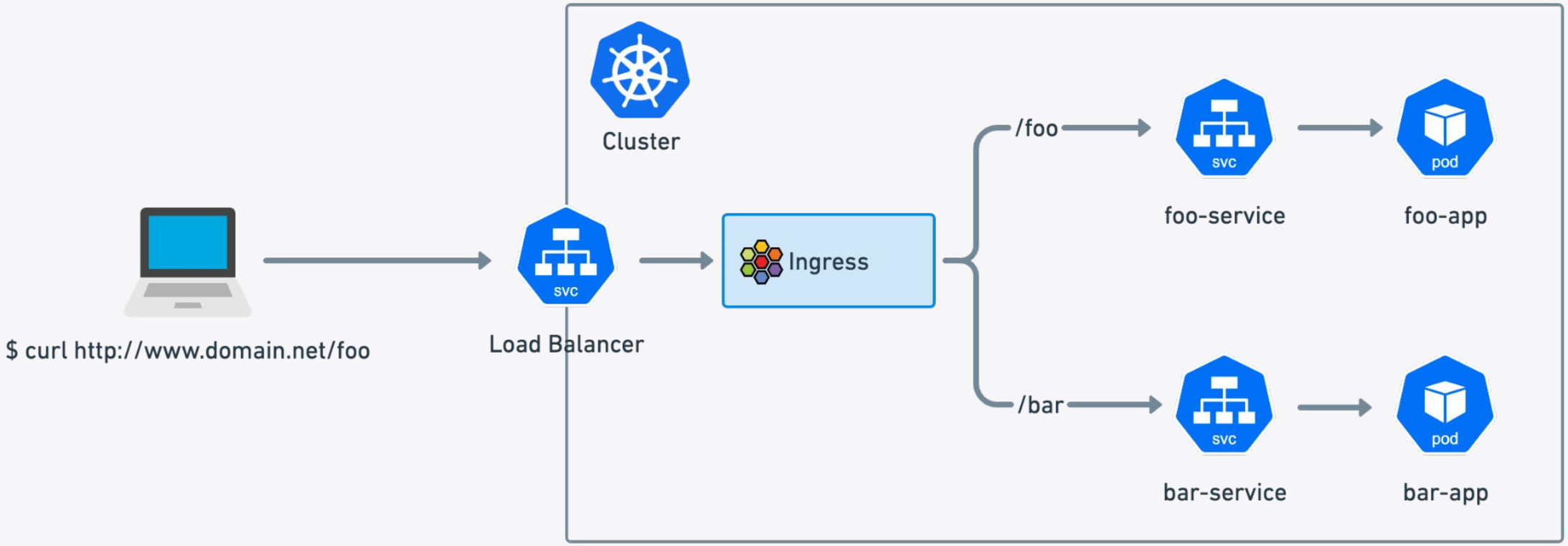
- Cilium & eBPF
- Kubernetes Services
- **Ingress**
- Cilium Envoy Config
- Gateway API
- Cilium Service Mesh

Ingress

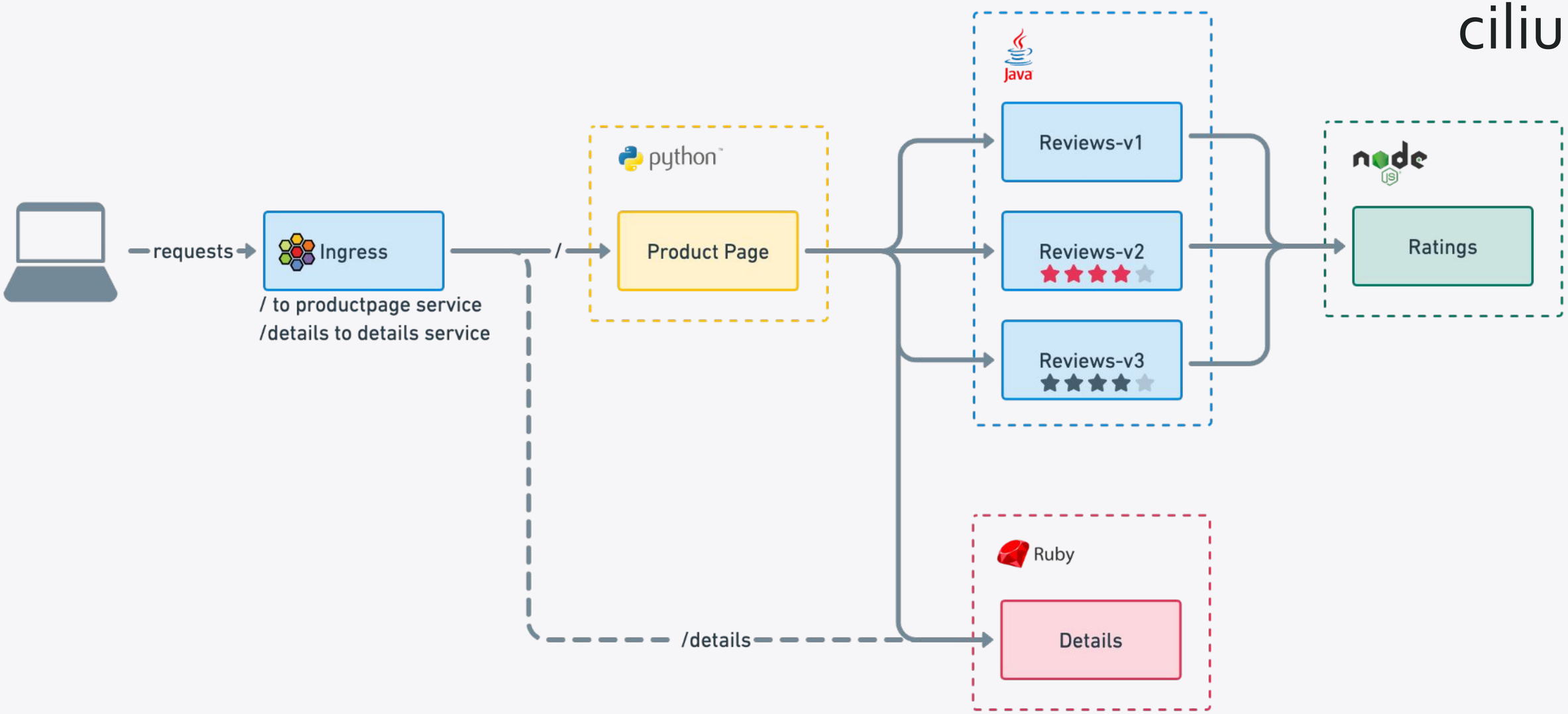
- Ingress can be used for path-based routing and TLS termination
- Cilium manages Ingress resources without external Ingress Controller
- Cilium Service Mesh Ingress Controller requires ability to create Service of Type LoadBalancer using either Cloud Provider integration or e.g. MetalLB
- Ingress CRD with `ingressClassName: cilium`

```yaml
apiVersion: networking.k8s.io/v1
kind: Ingress
metadata:
  name: basic-ingress
  namespace: default
spec:
  ingressClassName: cilium
  rules:
    - http:
        paths:
          - backend:
              service:
                name: details
                port:
                  number: 9080
            path: /details
            pathType: Prefix
          - backend:
              service:
                name: productpage
                port:
                  number: 9080
            path: /
            pathType: Prefix
```

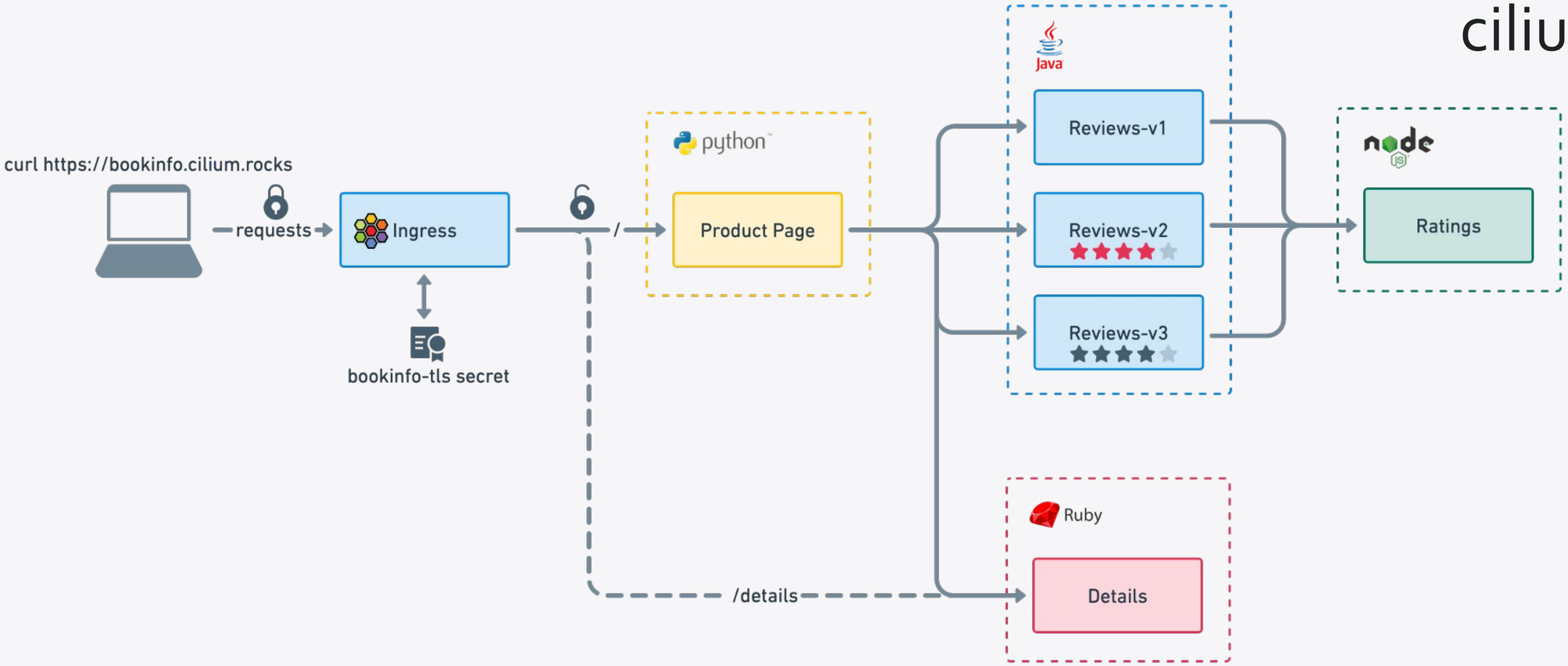
Ingress Overview



Ingress HTTP Example



TLS Termination



Ingress Limitations (some)

- Missing of expressivity
- Missing standardization
  - controller-specific annotations
  - controller-specific CRDs
- Limited Traffic Policies, missing:
  - retries
  - circuit breaking
  - rate limiting
  - header manipulation

- Time for a new standard?

Cilium Gateway API

The New L7 Kubernetes Ingress Standard

- Cilium & eBPF
- Kubernetes Services
- Ingress
- **Cilium Envoy Config**
- Gateway API
- Cilium Service Mesh

Cilium Envoy Config: Ingress Overview

CiliumEnvoyConfig Example

```yaml
apiVersion: cilium.io/v2alpha1
kind: CiliumEnvoyConfig
metadata:
  name: envoy-lb-listener
spec:
  services:
    - name: service-a
      namespace: app-test
    - name: service-b
      namespace: app-test
  resources:
    - "@type": type.googleapis.com/envoy.config.listener.v3.Listener
      name: envoy-lb-listener
(...)
      virtual_hosts:
        - name: "lb_route"
          domains: ["*"]
          routes:
            - match:
                prefix: "/"
              route:
                weighted_clusters:
                  clusters:
                    - name: "app-test/service-a"
                      weight: 75
                    - name: "app-test/service-b"
                      weight: 25
(...)
```

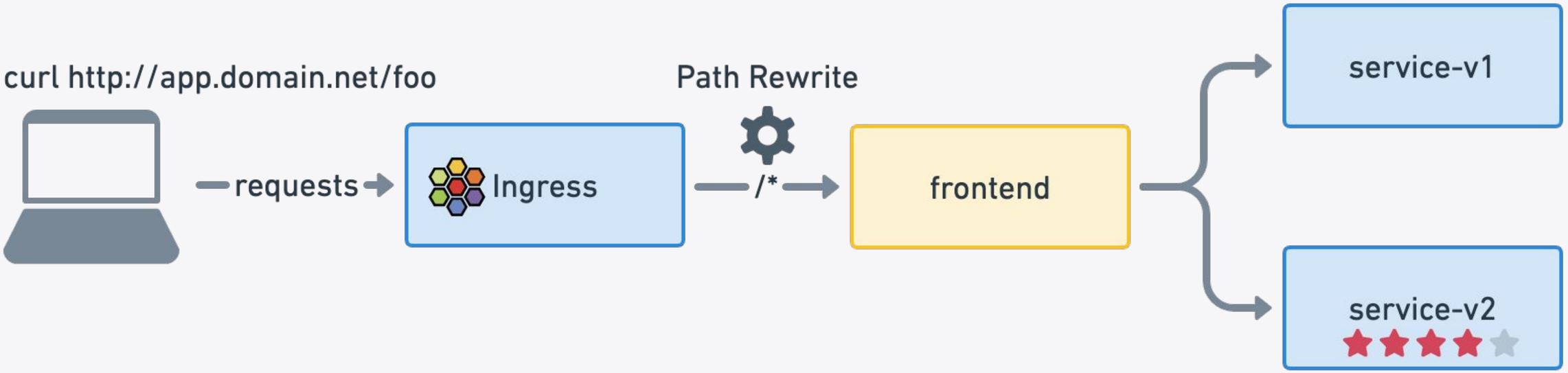
Cilium Envoy Config: Ingress Overview

CiliumEnvoyConfig Example

for Retries

```
apiVersion: cilium.io/v2alpha1
kind: CiliumEnvoyConfig
metadata:
  name: envoy-lb-listener
(...)
spec:
  services:
    - name: echo-other-node
      namespace: cilium-test
    - name: echo-same-node
      namespace: cilium-test
  resources:
    - "@type": type.googleapis.com/envoy.config.listener.v3.Listener
      name: envoy-lb-listener
      virtual_hosts:
        - name: "lb_route"
          domains: ["*"]
          routes:
          (...)
                retry_policy:
                  retry_on: 5xx
                  num_retries: 3
                  per_try_timeout: 1s
                regex_rewrite:
                  pattern:
                    google_re2: {}
                    regex: "^/foo.*$"
                  substitution: "/"
(...)
```

L7 Traffic Management
Path Rewrite

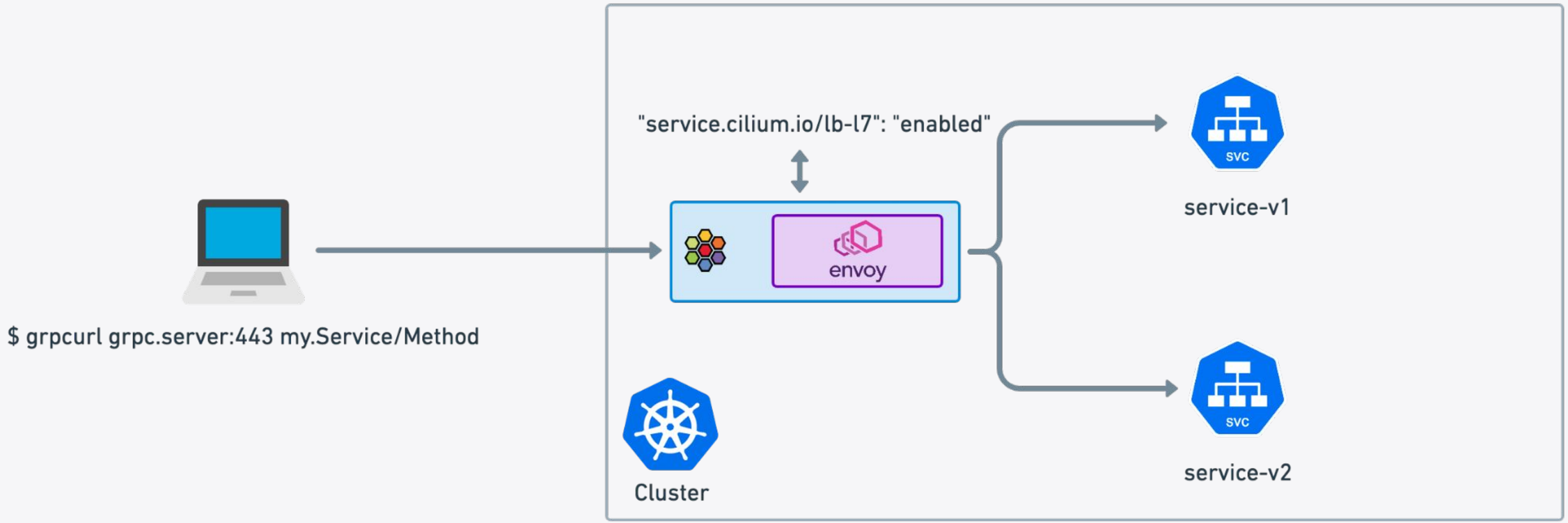


Cilium Envoy Config:
L7 Traffic Management Path Rewrite

CiliumEnvoyConfig Example

```yaml
apiVersion: cilium.io/v2alpha1
kind: CiliumEnvoyConfig
metadata:
  name: envoy-lb-listener
spec:
  services:
    - name: echo-other-node
      namespace: cilium-test
    - name: echo-same-node
      namespace: cilium-test
  resources:
(...)
      virtual_hosts:
        - name: "lb_route"
          domains: ["*"]
          routes:
            - match:
                prefix: "/"
              route:
(...)
                regex_rewrite:
                  pattern:
                    google_re2: {}
                    regex: "^/foo.*$"
                  substitution: "/"
```

L7 Load Balancing for Kubernetes Services with Annotations



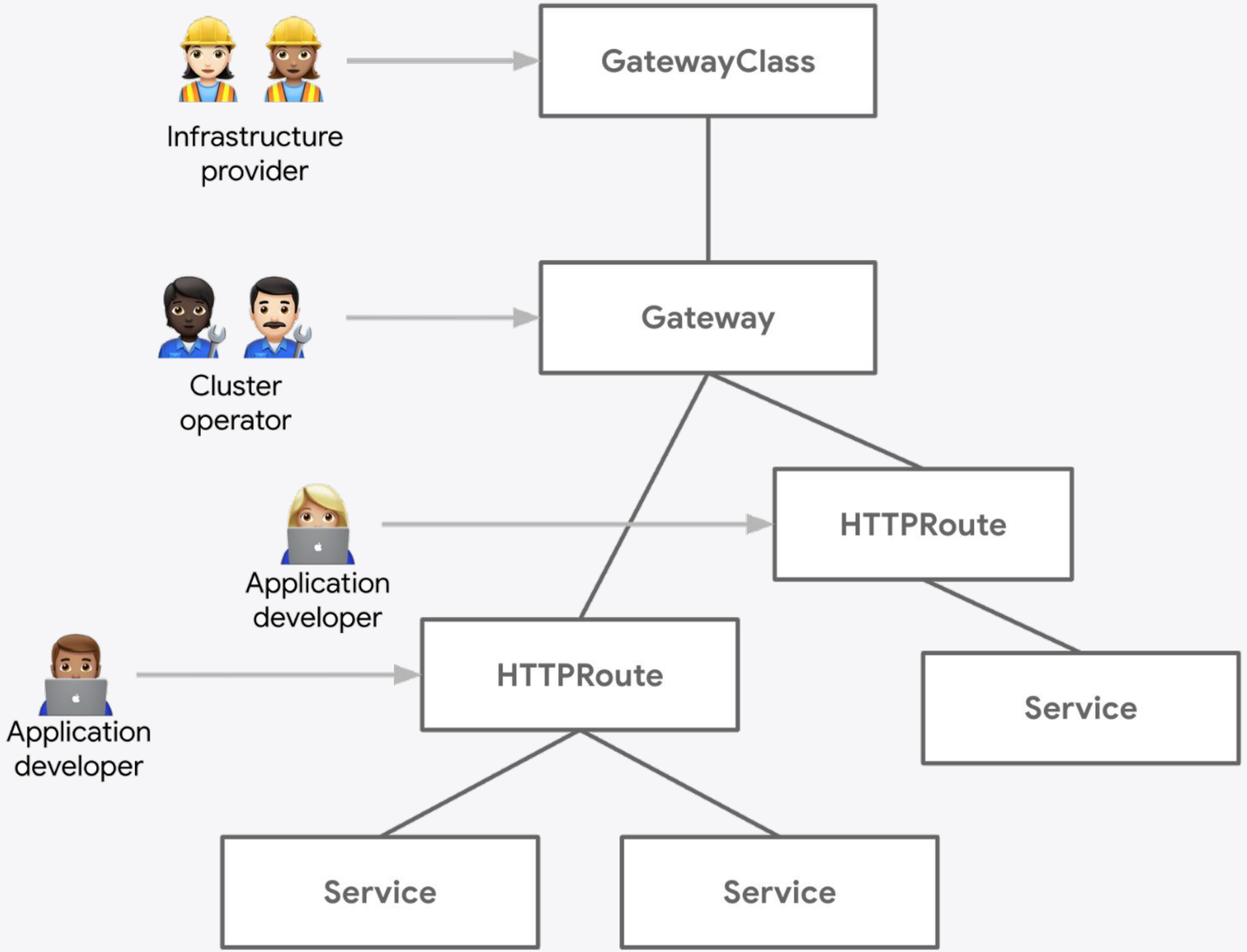
Cilium Gateway API

The New L7 Kubernetes Ingress Standard

- Cilium & eBPF
- Kubernetes Services
- Ingress
- Cilium Envoy Config
- **Gateway API**
- Cilium Service Mesh

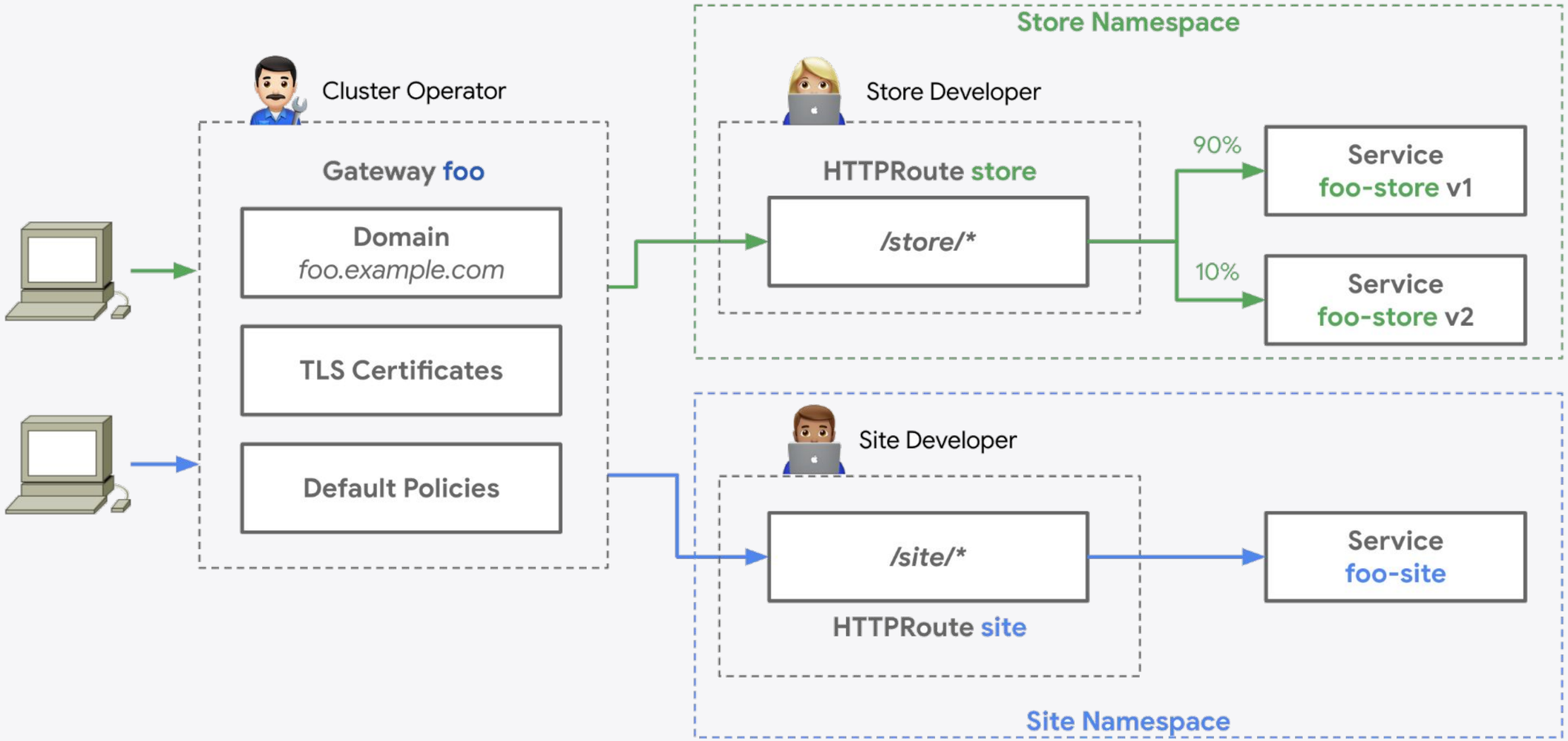
Gateway API

CRDs



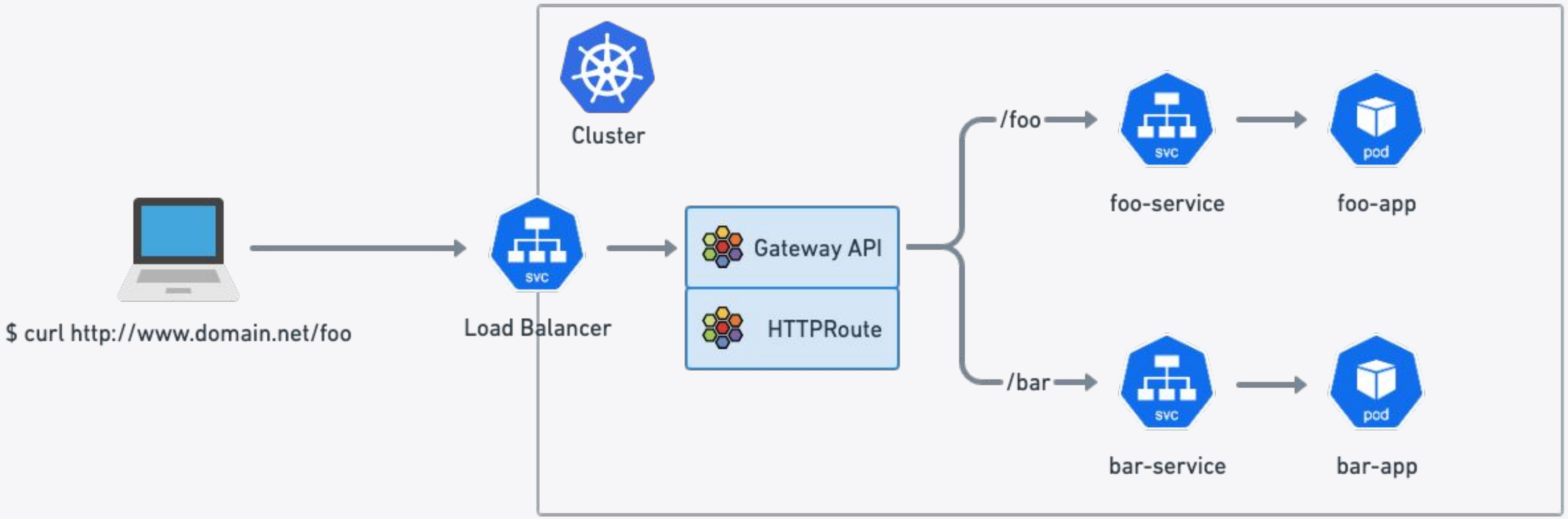
Gateway API

Principles



Gateway API

Cilium Overview



Gateway API

Use of Gateway and HTTPRoute objects for path-based routing


```yaml
apiVersion: gateway.networking.k8s.io/v1beta1
kind: Gateway
metadata:
  name: my-gateway
spec:
  gatewayClassName: cilium
  listeners:
  - protocol: HTTP
    port: 80
    name: web-gw
    allowedRoutes:
      namespaces:
        from: Same
```

```yaml
apiVersion: gateway.networking.k8s.io/v1alpha2
kind: HTTPRoute
metadata:
  name: http-app-1
spec:
  parentRefs:
  - name: my-gateway
    namespace: default
  rules:
  - matches:
    - path:
        type: PathPrefix
        value: /details
    backendRefs:
    - name: details
      port: 9080
```

Gateway API

Use of Gateway and HTTPRoute for TLS Termination

```yaml
apiVersion: gateway.networking.k8s.io/v1beta1
kind: Gateway
metadata:
  name: tls-gateway
spec:
  gatewayClassName: cilium
  listeners:
  - name: https
    protocol: HTTPS
    port: 443
    hostname: "bookinfo.cilium.rocks"
    tls:
      certificateRefs:
      - kind: Secret
        name: demo-cert
```

```yaml
apiVersion: gateway.networking.k8s.io/v1beta1
kind: HTTPRoute
metadata:
  name: https-app-route
spec:
  parentRefs:
  - name: tls-gateway
  hostnames:
  - "bookinfo.cilium.rocks"
  rules:
  - matches:
    - path:
        type: PathPrefix
        value: /details
    backendRefs:
    - name: details
      port: 9080
```

Gateway API

Traffic Splitting with Weighted Routes

cilium

```yaml
apiVersion: gateway.networking.k8s.io/v1alpha2
kind: HTTPRoute
metadata:
  name: example-weighted-route
spec:
  parentRefs:
  - name: my-gateway
  rules:
  - matches:
    - path:
        type: PathPrefix
        value: /echo
    backendRefs:
    - kind: Service
      name: echo-1
      port: 8080
      weight: 75
    - kind: Service
      name: echo-2
      port: 8090
      weight: 25
```

Layer 4/7 Traffic Management Options

Ingress

Original L7 load-balancing standard in K8s

Simple

Supported since Cilium 1.12

Services

Use of K8s services with annotations

Simple

Supported since Cilium 1.13

Gateway API

Originally labelled Ingress v2. Richer in features.

Simple

Supported for v0.5.1 since Cilium 1.13

EnvoyConfig

Raw Envoy Config via CustomResource

Advanced Users & Integrations

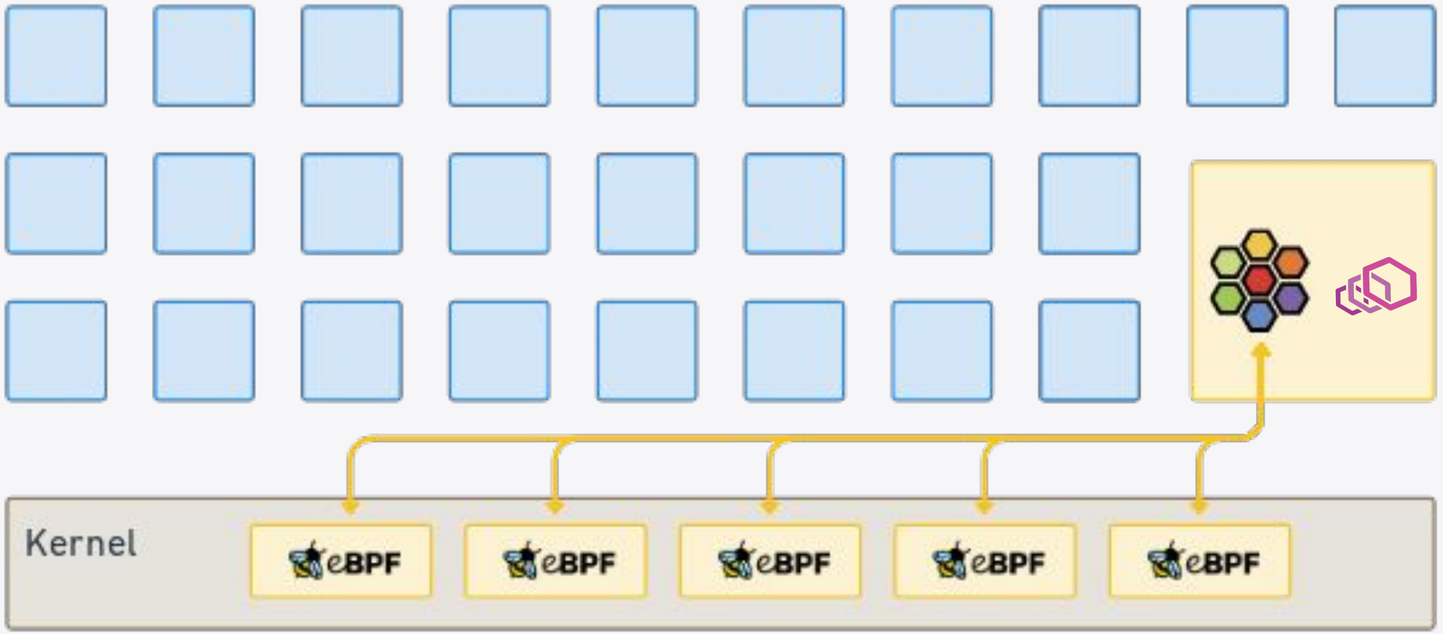
Supported since Cilium 1.12

Cilium Gateway API

The New L7 Kubernetes Ingress Standard

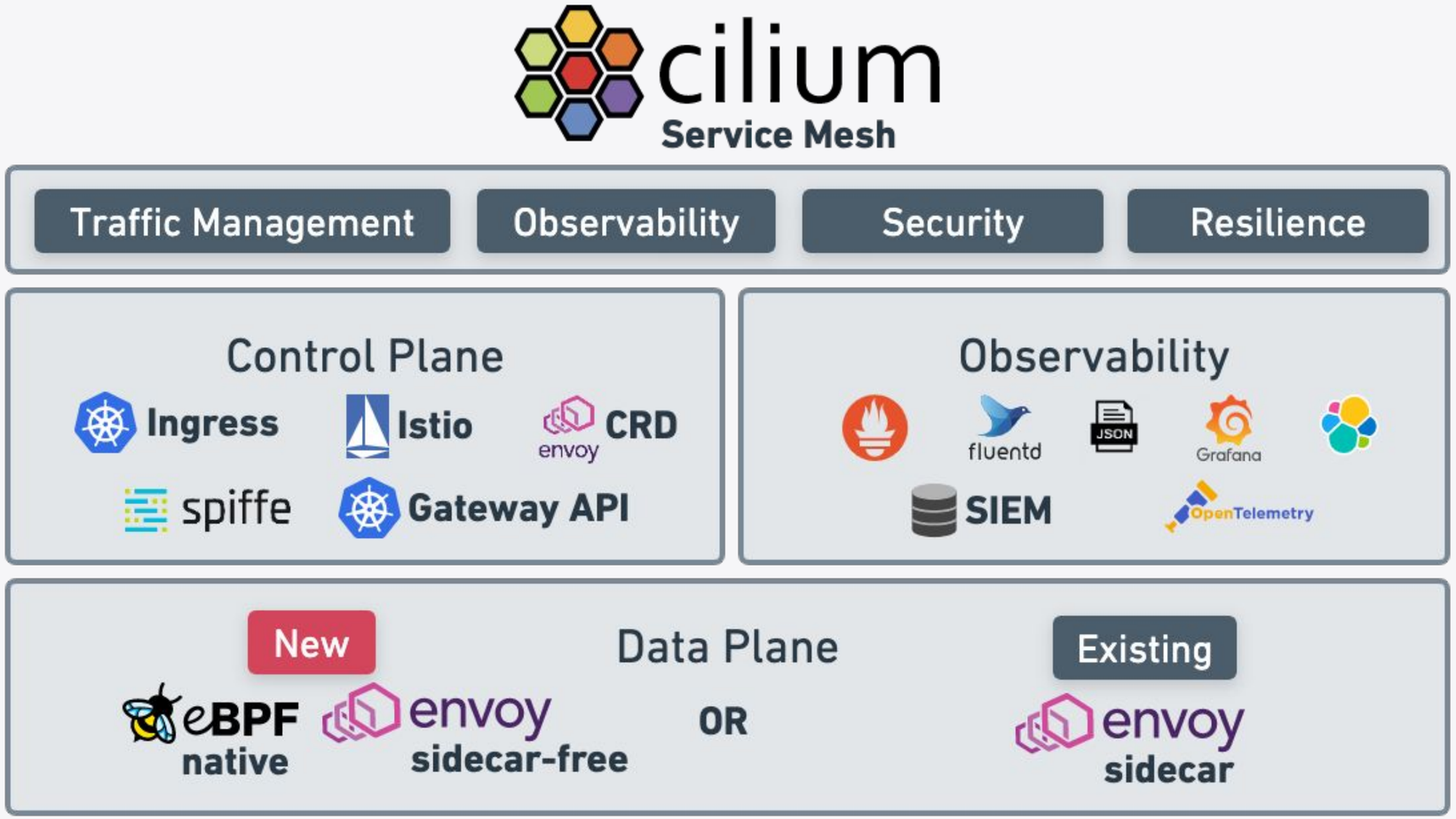
- Cilium & eBPF
- Kubernetes Services
- Ingress
- Cilium Envoy Config
- Gateway API
- **Cilium Service Mesh**

Cilium agent per node

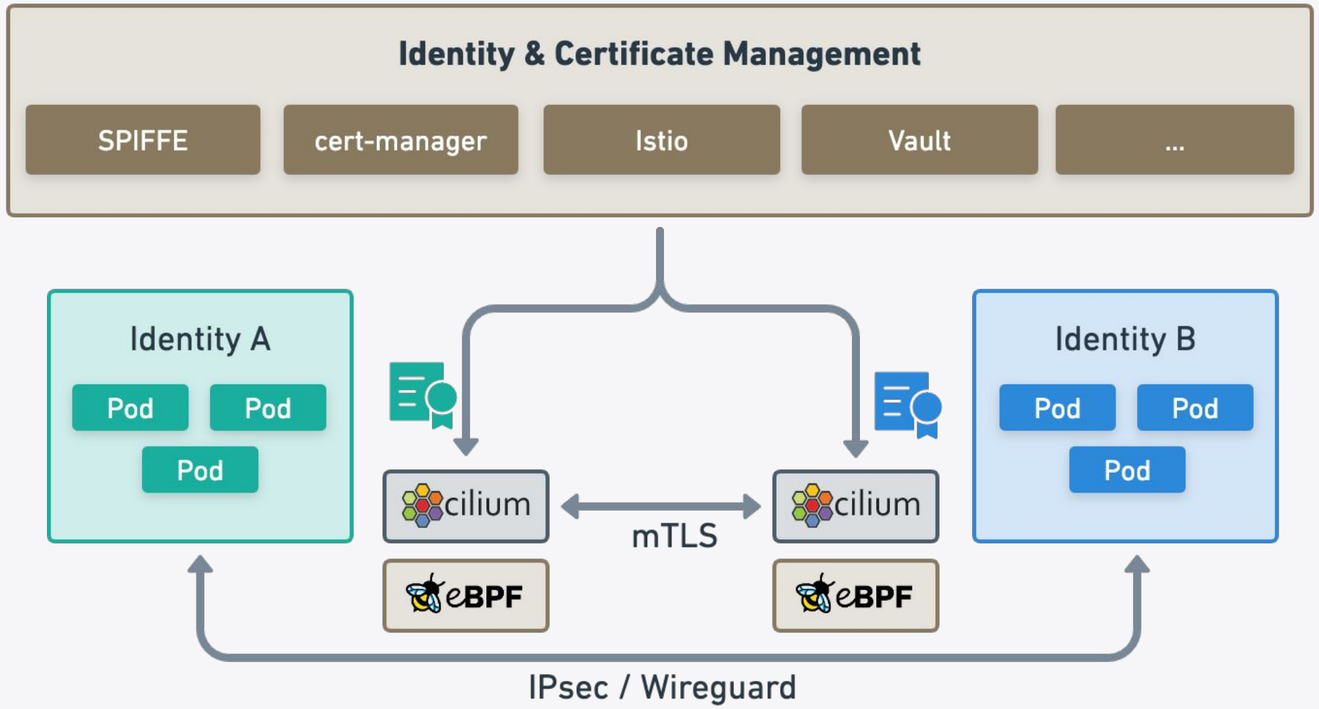


- Dynamic eBPF programs
- Envoy for L7 policies & observability

Cilium Service Mesh



Mutual Authentication



- User space mTLS authentication
- Proxy-free in-kernel datapath
- Keeps secrets out of L7 proxies
- Works for any protocol (UDP, SCTP, ...)
- IPsec/Wireguard can use TLS negotiated service-specific keys

More information: https://isovalent.com/blog/post/2022-05-03-servicemesh-security

Service + Annotations

Simple way to enable L7 load balancing

```yaml
apiVersion: v1
kind: Service
metadata:
  name: backend
  annotations:
    service.cilium.io/lb-l7: "enabled"
    service.cilium.io/lb-l7-algorithm: "least-request"
spec:
  type: ClusterIP
  ports:
    - port: 80
  selector:
    name: backend
```

Possible values: round_robin, least_request, random

Learn more!

ISOVALENT

For the Enterprise

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

isovalent.com/product
isovalent.com/labs

OSS Community

eBPF-based Networking, Observability, Security

cilium.io
cilium.slack.com
Regular news

Base technology

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

ebpf.io
What is eBPF? - ebook

Practical Labs

... to become a Cilium & eBPF Jedi

isovalent.com/labs

Practical Labs

... to become a Cilium & eBPF Jedi

isovalent.com/labs

ISOVALENT

Thank you!