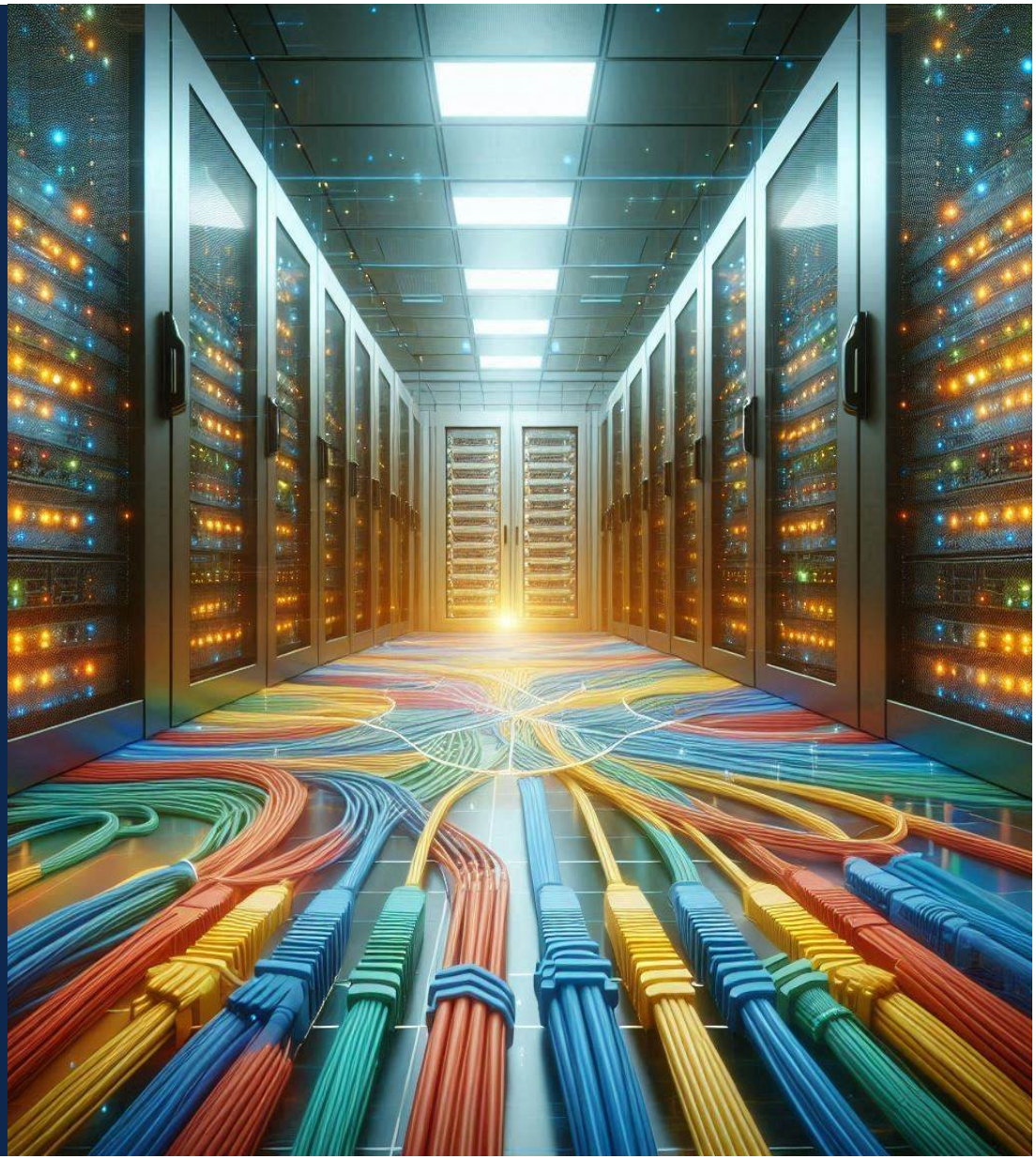
Tech Club webinář

Kubernetes, Openshift, Cilium a nový Cisco SmartSwitch: jak zapadají do nové koncepce bezpečnosti DC a aplikací

Martin Diviš, TSE, mdivis@cisco.com

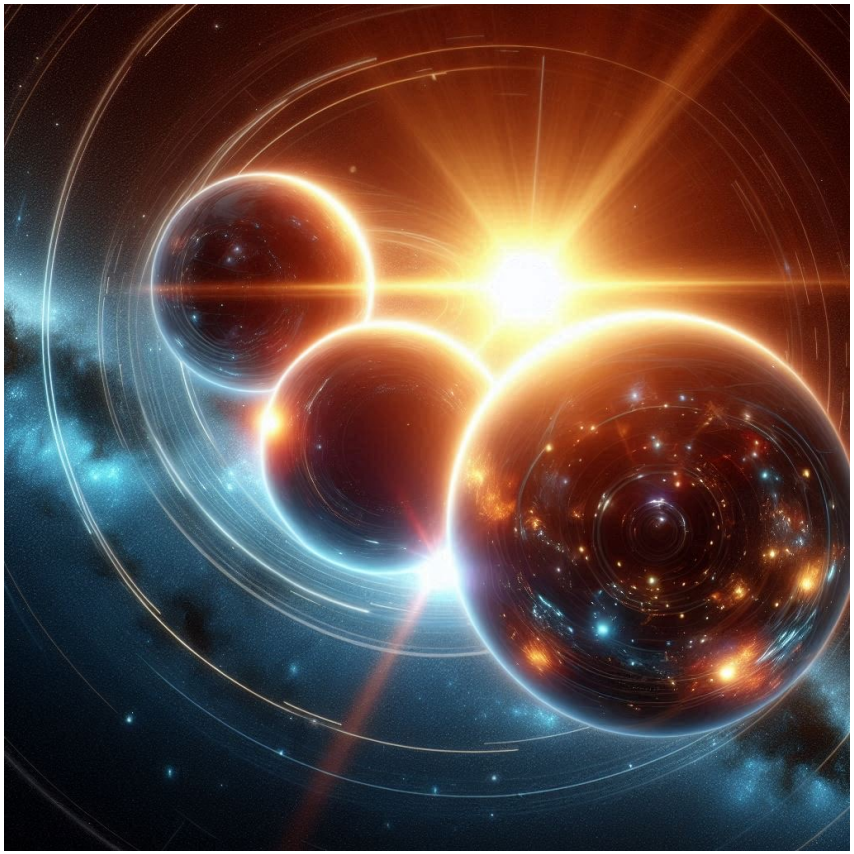
1. duben 2025

Segmentace – tradiční řešení bezpečnosti v DC



Segmentation – Policy vs. Enforcement

- The ”three-body problem”
  - **Security rules** – who can do what, talk to who, etc.
  - **Identity** – how to identify the subject of the rule?
  - **Observe and Enforce** – how to make sure we see the behavior of a subject and how to enforce the rules?



Segmentation in DC

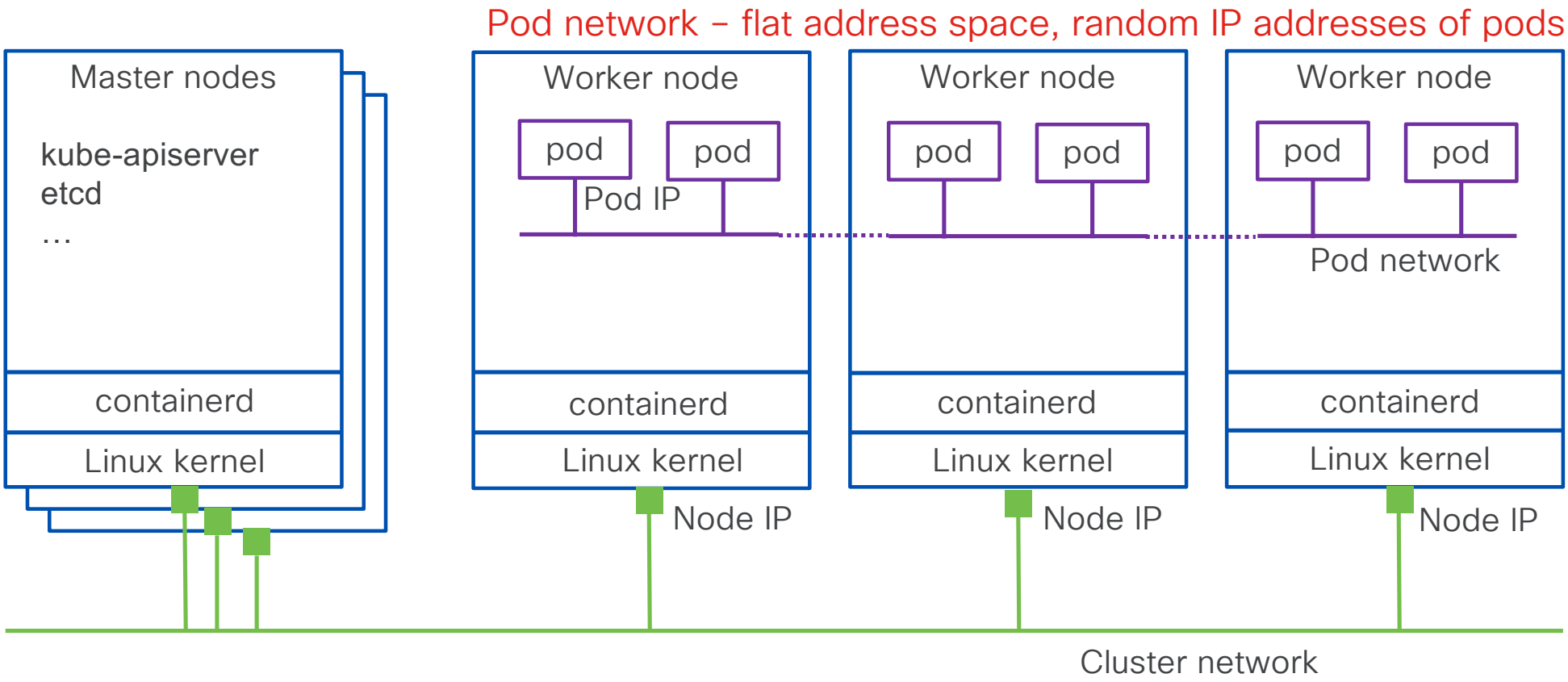
- **Identity** – IP address, IP subnet
- **Security rules** – ACL-like, requires a lot of knowledge and work, hard to maintain over time
- **Observe and Enforce** – make sure data flow is seen: network ACLs, Firewalls (often routing between subnets), ACI policies, host based firewalls

Segmentace v prostředí kontejnerových platforem

Kubernetes – the cloud native application platform

- Kubernetes (K8s) is an open-source system for automating deployment, scaling, and management of containerized applications.
- It provides a platform for managing resilient, scalable, and distributed systems.
- Key benefit: Abstracts away the complexities of managing individual containers.
- It allows for declarative configuration.

Kubernetes cluster



How are application really exposed

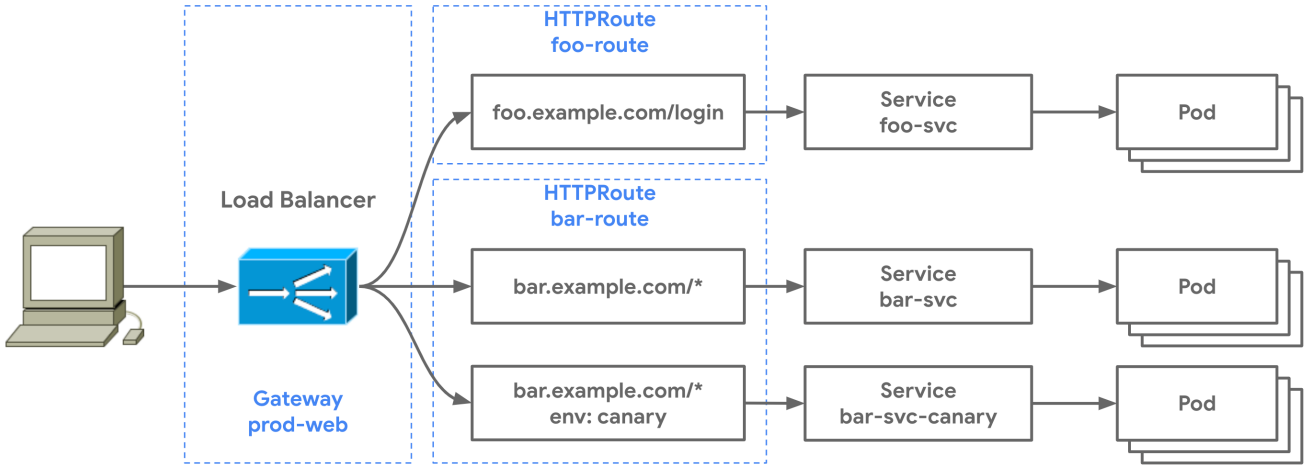
- Pods have ”random” IPs – direct communication is rather rare
- Pods have replicas – we need loadbalancing among the replicas
- Services
  - Expose applications in pods to other pods and to external world
  - Loadbalance traffic to pod replicas
  - Check availability of pods via probes (readiness, health)

Services types

```yaml
apiVersion: v1
kind: Service
metadata:
  name: my-loadbalancer-service
spec:
  type: LoadBalancer
  selector:
    app: my-app
  ports:
    - protocol: TCP
      port: 80
      targetPort: 80
```

- **Cluster IP** – not exposed outside of the cluster, for internal Pod-Service communication only
- **NodePort** – the same port assigned to an application on all nodes on node IPs. Externally, connect to any node on that port to reach the service
- **LoadBalancer** – single external entrypoint to the service, often relies on NodePort mappings. Not native part of Kubernetes, it can be external to the cluster (physical F5,...) or SW running on the cluster (MetalLB,...). External IP address provided by the LoadBalancer management and sent back to the K8S management.

Gateway API – Application Routing into K8S

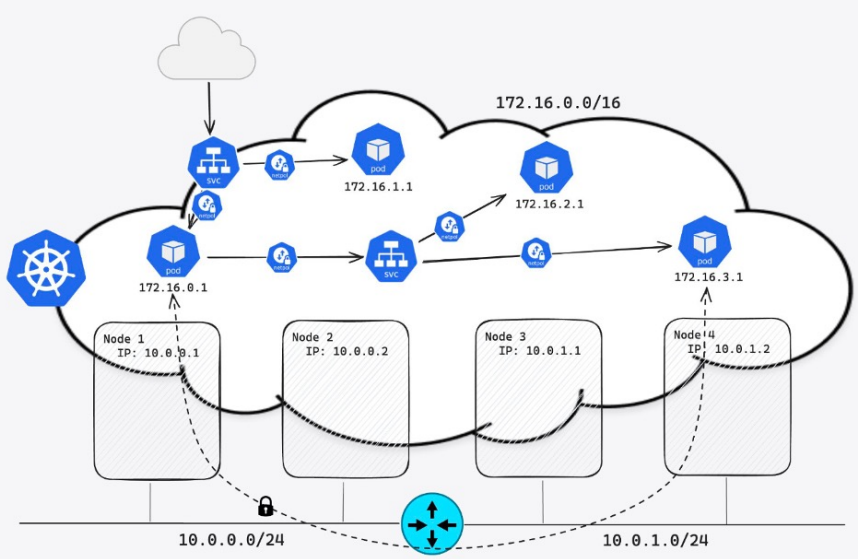



```yaml
apiVersion: gateway.networking.k8s.io/v1
kind: Gateway
metadata:
  name: example-gateway
spec:
  gatewayClassName: example-gateway-class
  listeners:
  - name: http
    protocol: HTTP
    port: 80
---
apiVersion: gateway.networking.k8s.io/v1
kind: HTTPRoute
metadata:
  name: example-route
spec:
  parentRefs:
  - name: example-gateway
  hostnames:
  - "example.com"
  rules:
  - backendRefs:
    - name: example-svc
      port: 80
```

Implementations:
https://gateway-api.sigs.k8s.io/implementations/

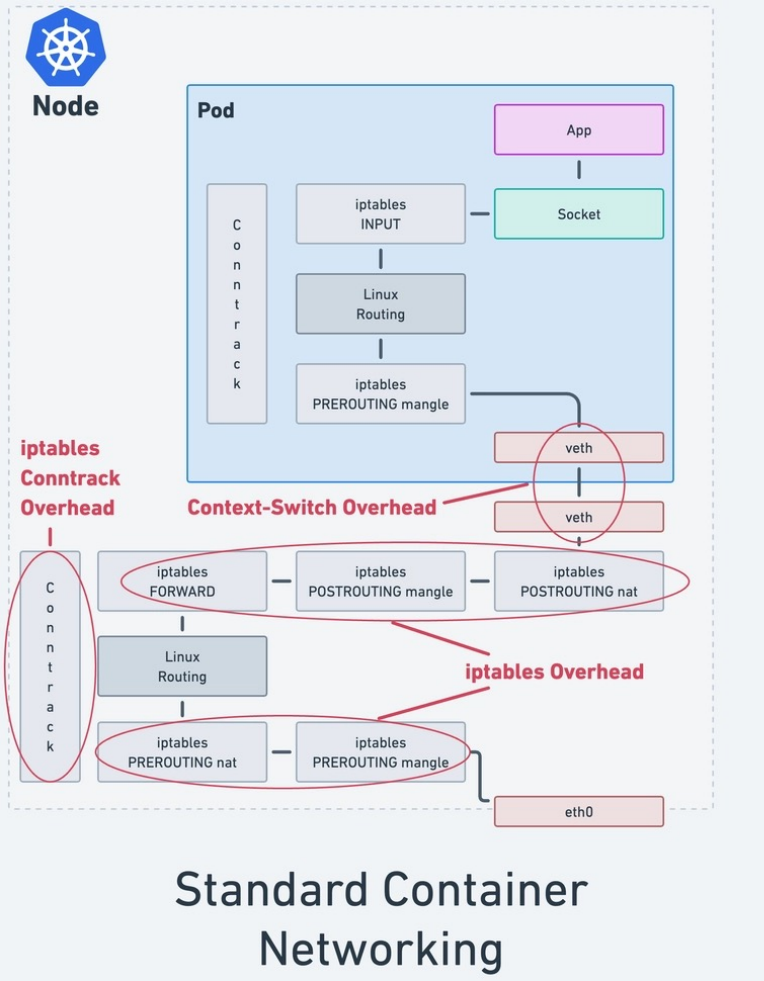
Kubernetes Networking - Summary

- IPAM – Dynamic (CNI)
- Routing – Dynamic (CNI)
- East-West connectivity
  - Service type: ClusterIP (kube-proxy)
- North-South connectivity
  - Service type: LoadBalancer (kube-proxy)
- Service discovery: Dynamic (CoreDNS)
- Security:
  - NetworkPolicy (CNI)
  - Transparent Encryption (CNI / Service Mesh)

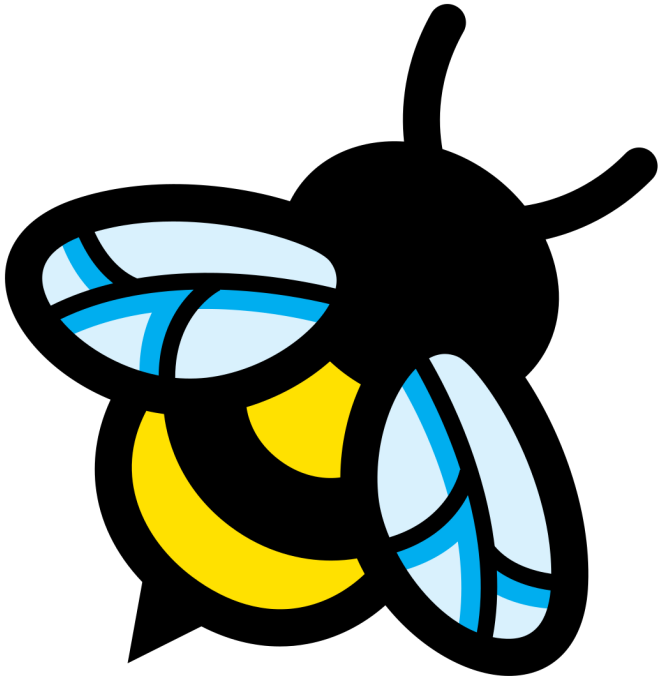


Challenges of iptables

- ACLs as sequential list of rules
- Updates all rules in single transaction
- Matches based on IP proto/addr/port only
- New IP/Port require rules added and chain changed
- Larger overhead
- Reduced performance and increased latency at scale

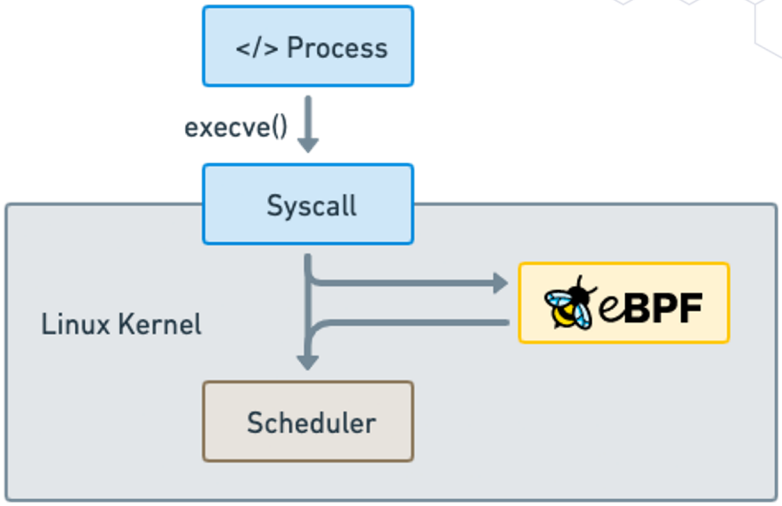


A Solution Arises



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

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

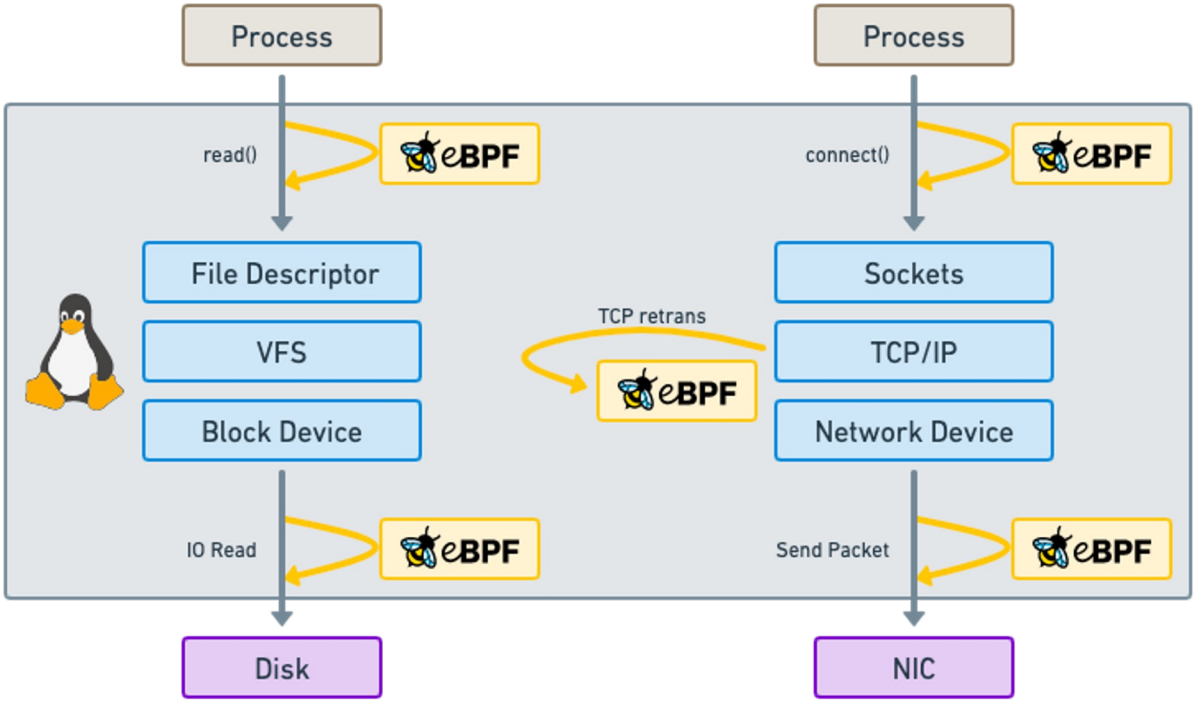



```
int syscall__ret_execve(struct pt_regs *ctx)
{
        struct comm_event event = {
                .pid = bpf_get_current_pid_tgid() >> 32,
                .type = TYPE_RETURN,
        };

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

        return 0;
}
```

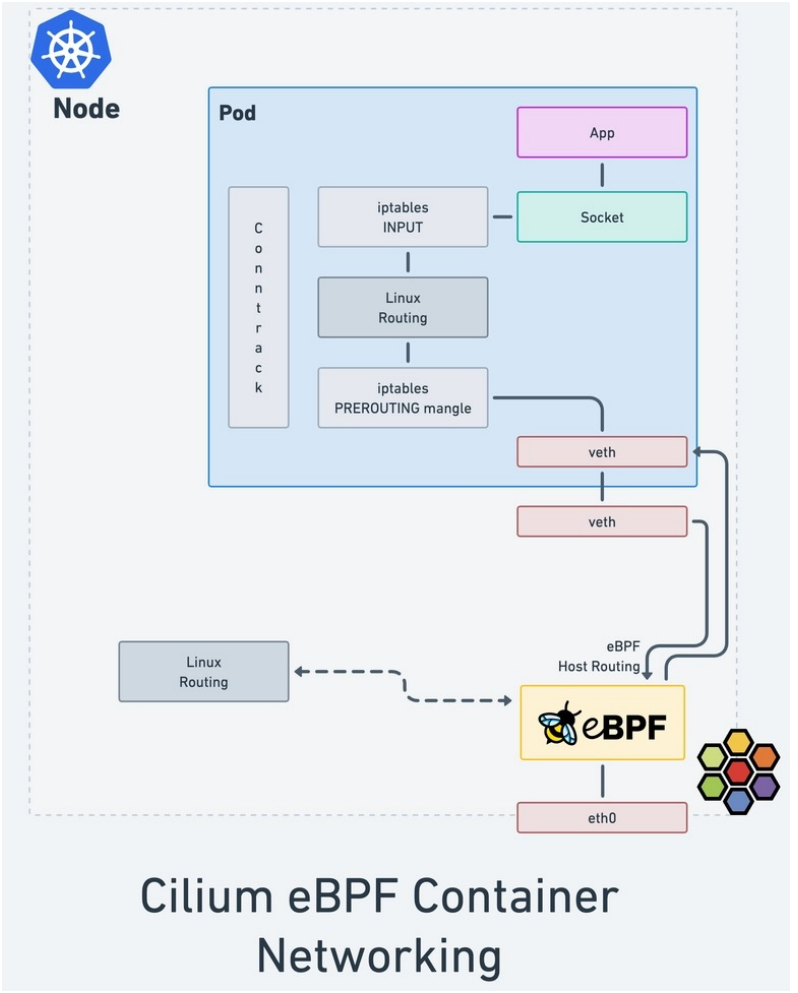
Run eBPF programs on events



Attachment points

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

Removing iptables

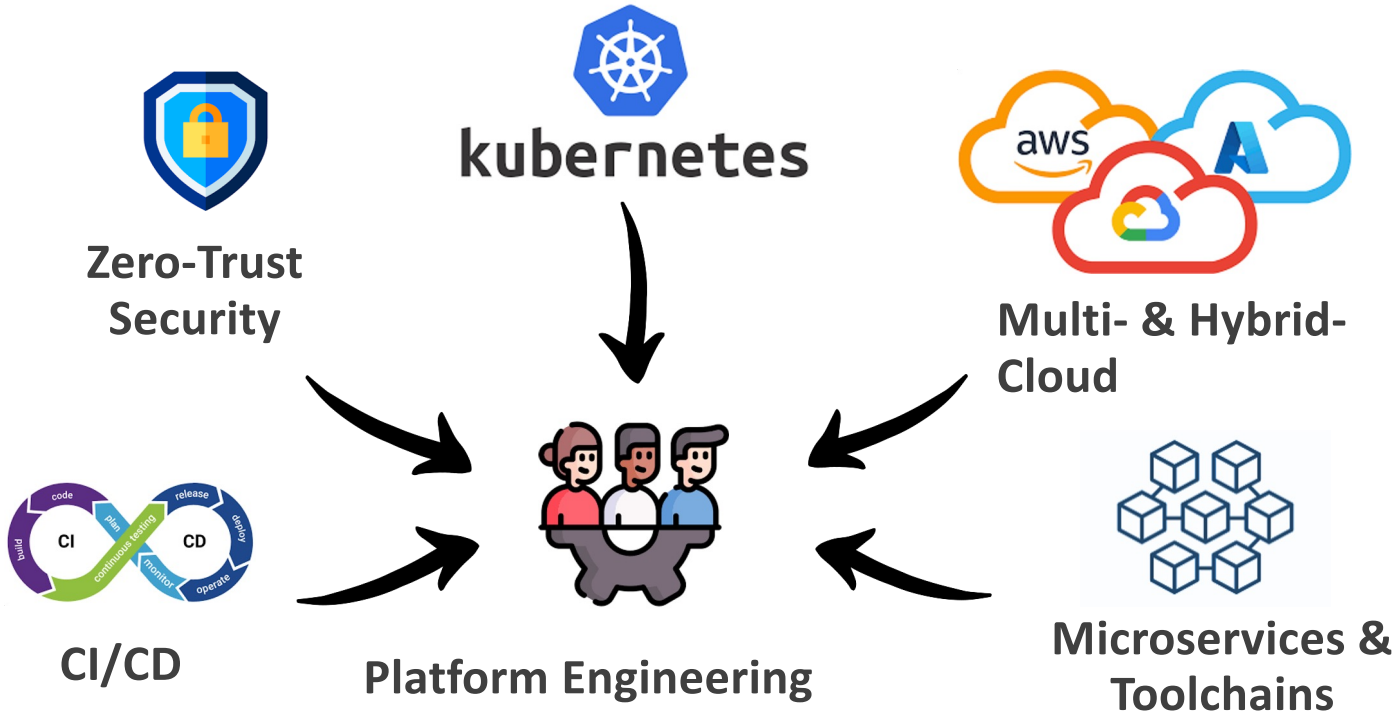


Cilium eBPF Container Networking

ISOVALENT

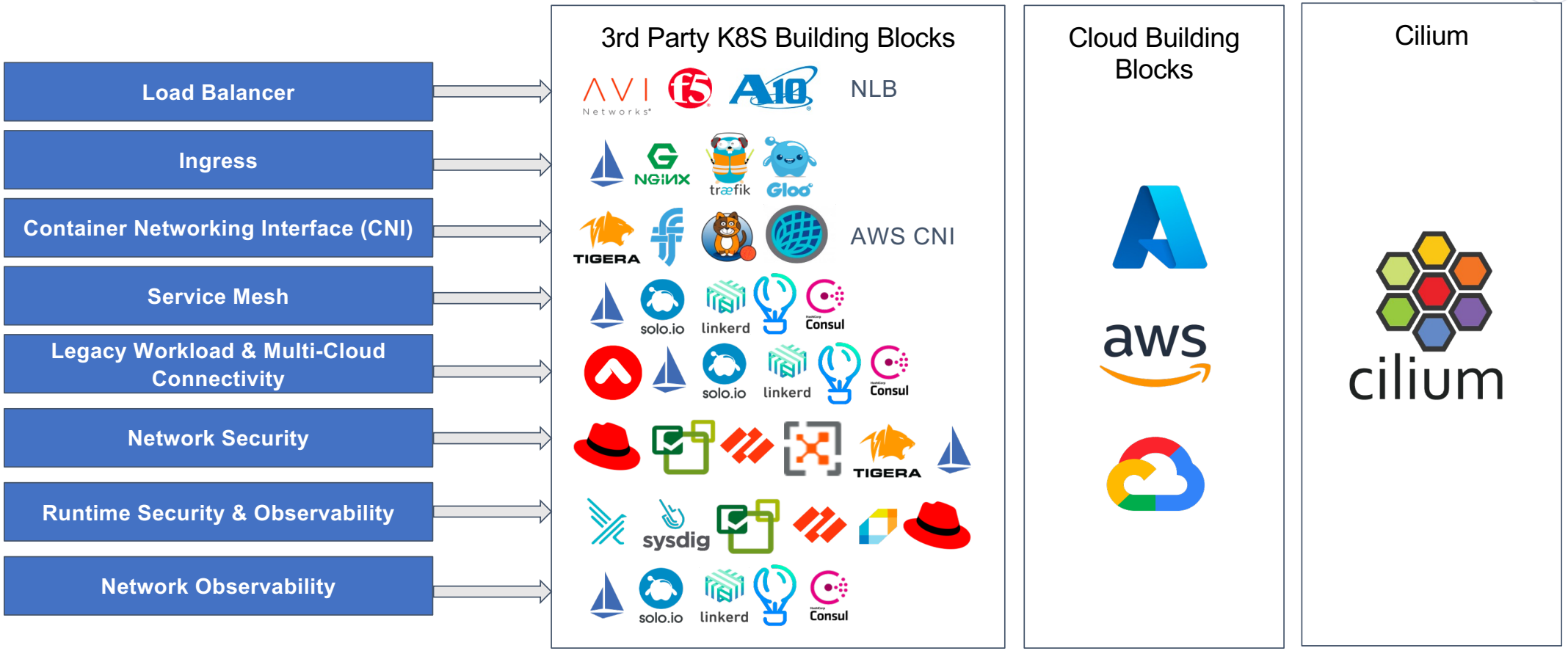
Platform Engineering

Key Infrastructure Trend

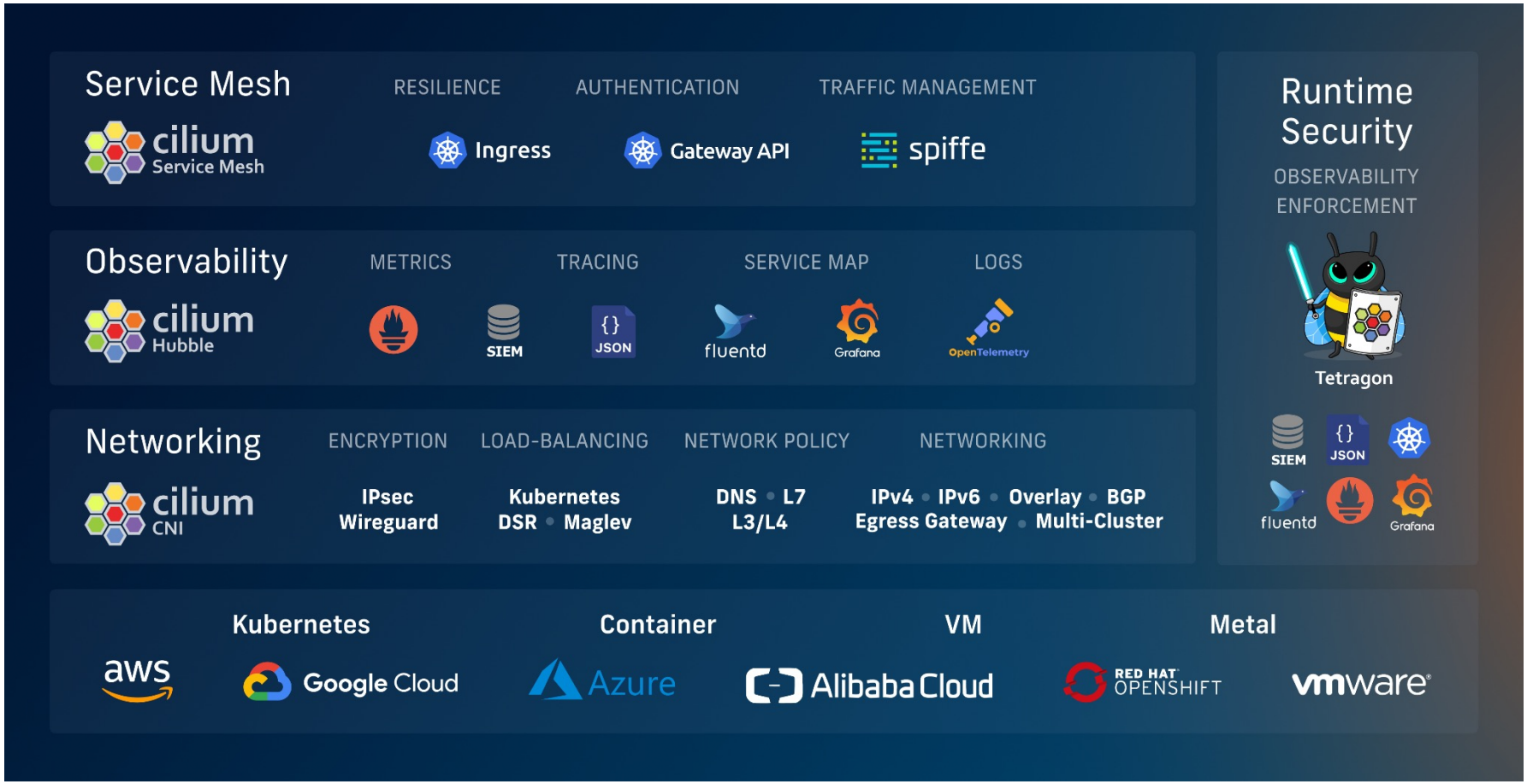


Platform engineering is the newest sociotechnical discipline to arise in response to the cloud native world. As the process of designing, building, and maintaining workflows and tools for software engineering organizations, platform engineering helps drive consistency and speed up common tasks.

Kubernetes Networking Tool Sprawl

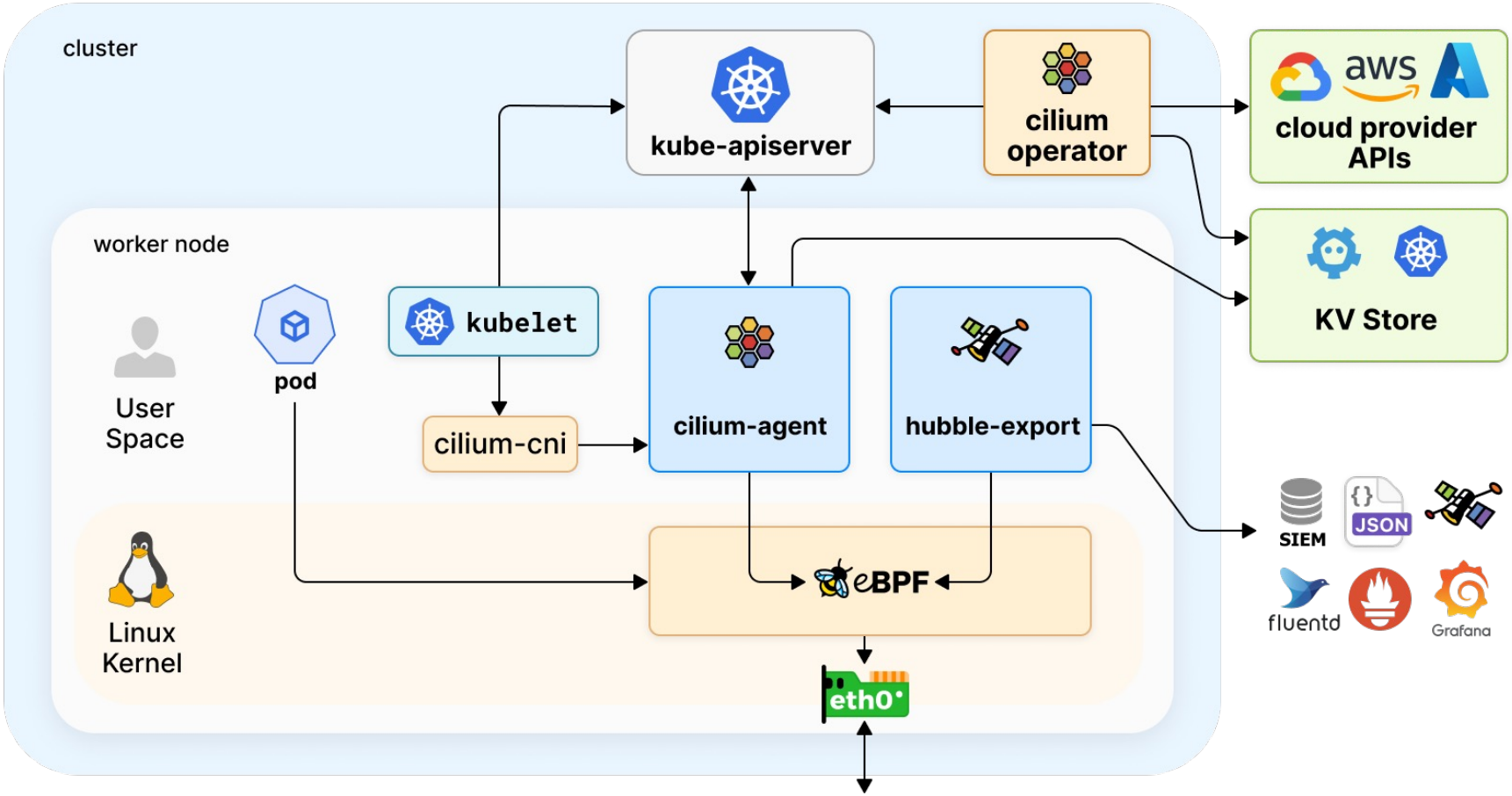


The Isovalent Overview

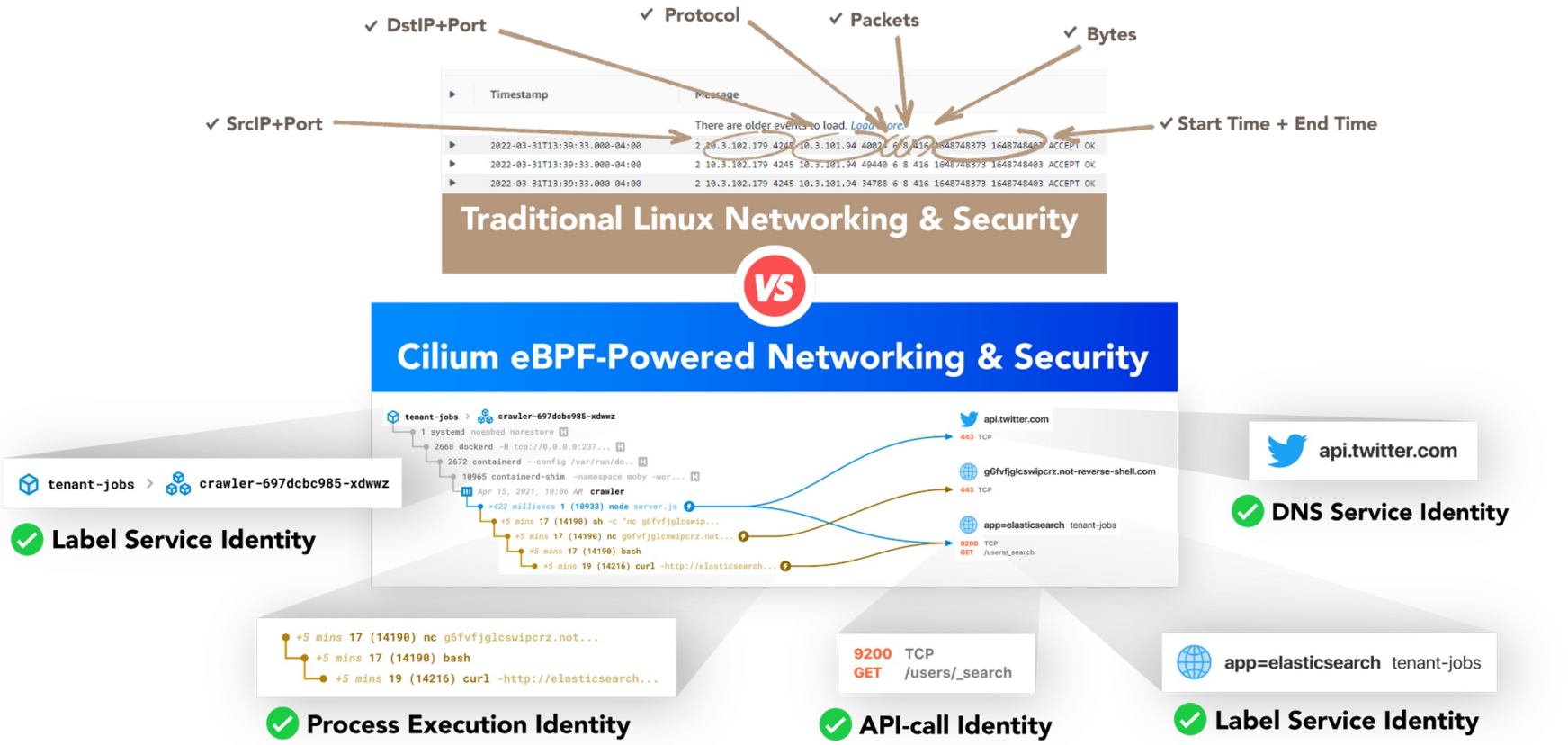


Cilium + Hubble

Cilium Architecture – More Detail

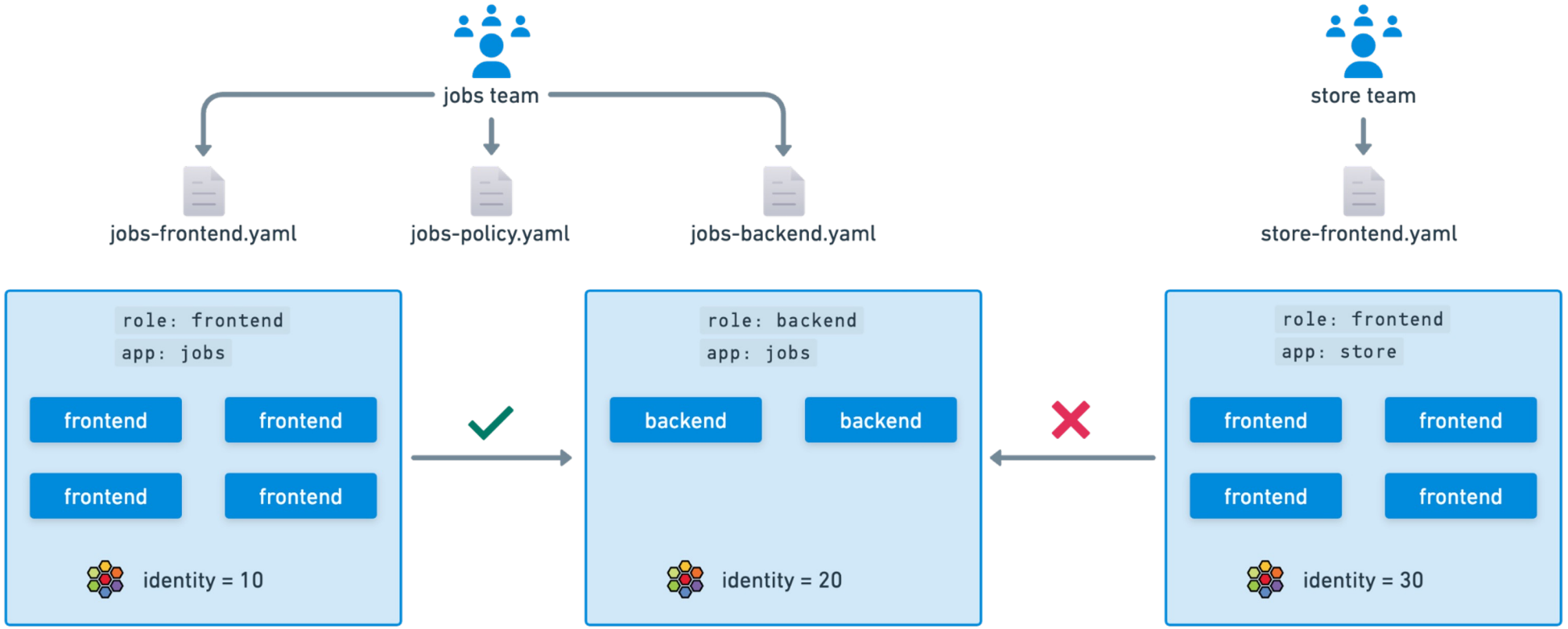


Identity-based Security



Micro-Segmentation

Label based East-West Application or Multi-tenant Security Enforcement



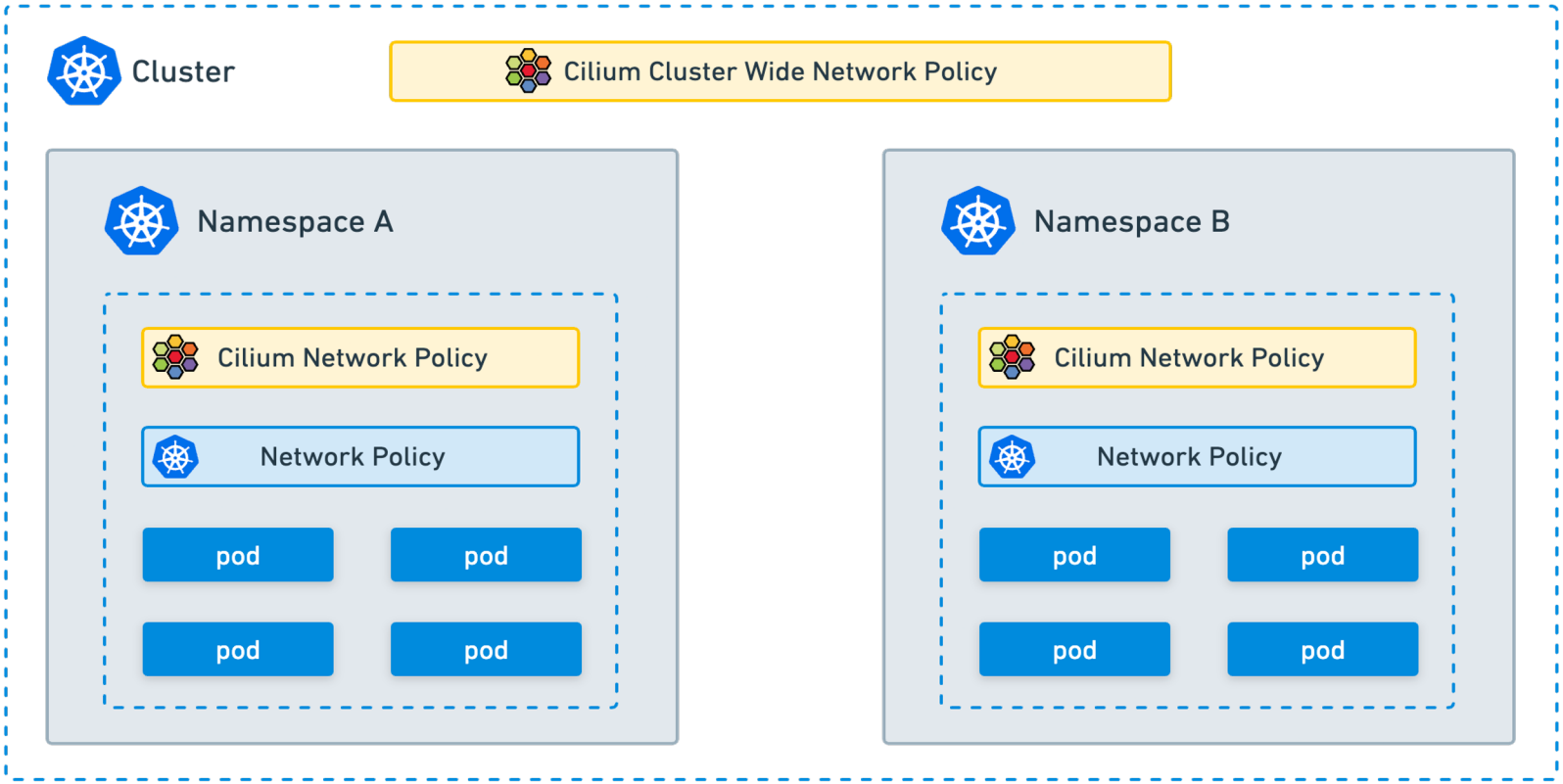
Example Layer 7 Rule

- CiliumNetworkPolicy
- Allows matching pods on labels
- Defines ingress from source labels
- Defined on destination range
- Applies to POST operations to specific URI

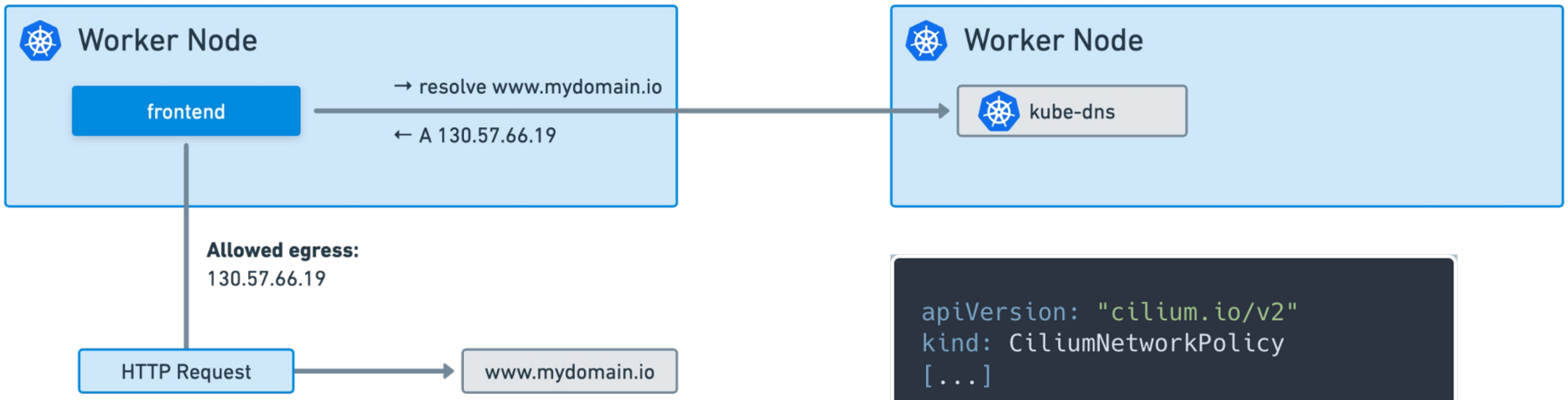
```yaml
apiVersion: "cilium.io/v2"
kind: CiliumNetworkPolicy
metadata:
  name: "rule1"
spec:
  description: "L7 policy to restrict access to specific HTTP call"
  endpointSelector:
    matchLabels:
      org: empire
      class: deathstar
  ingress:
  - fromEndpoints:
    - matchLabels:
        org: empire
    toPorts:
    - ports:
      - port: "80"
        protocol: TCP
      rules:
        http:
        - method: "POST"
          path: "/v1/request-landing"
```

Enforce Consistent Policies across Clusters

Simplify Network Management and set Guardrails for your Platform



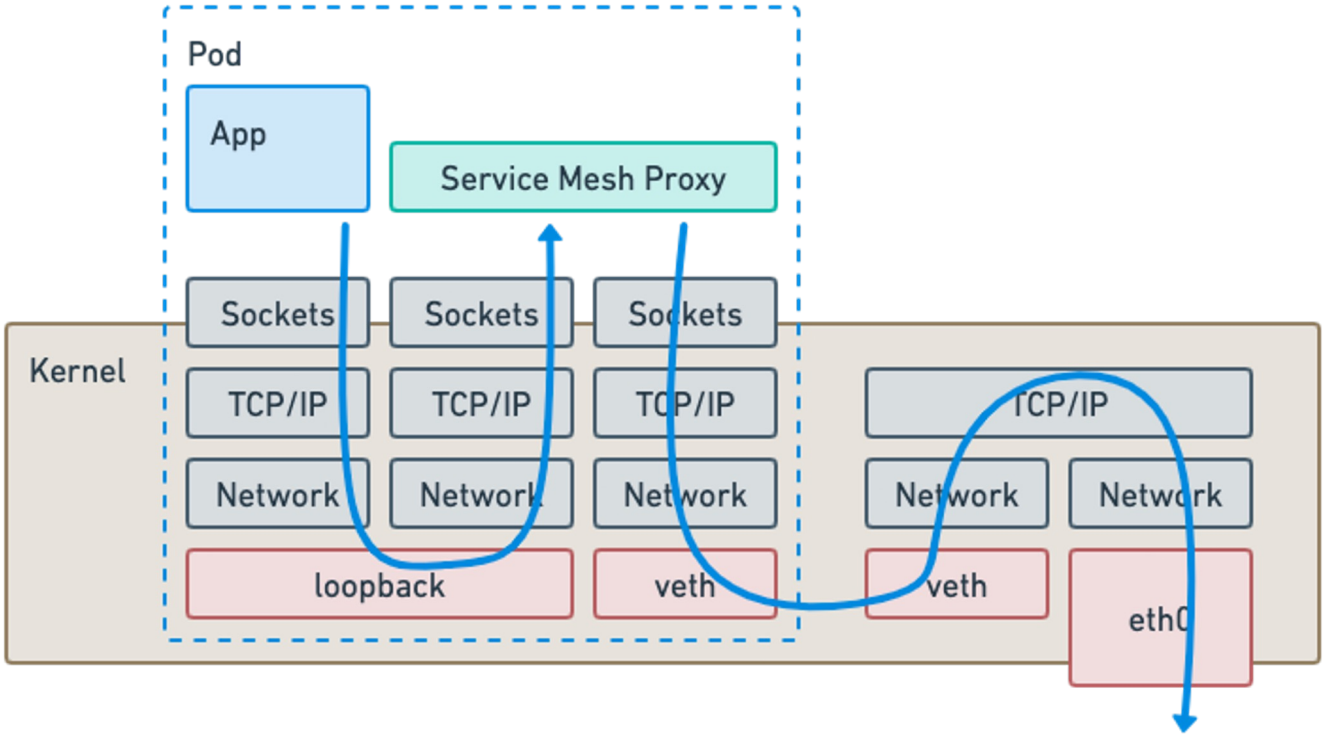
DNS-aware Cilium Network Policy




```
apiVersion: "cilium.io/v2"
kind: CiliumNetworkPolicy
[...]
specs:
- endpointSelector:
  matchLabels:
    app: frontend
  egress:
  - toFQDNs:
    - matchName: "*.mydomain.io"
    toPorts:
    - ports:
      - port: "443"
        protocol: TCP
```

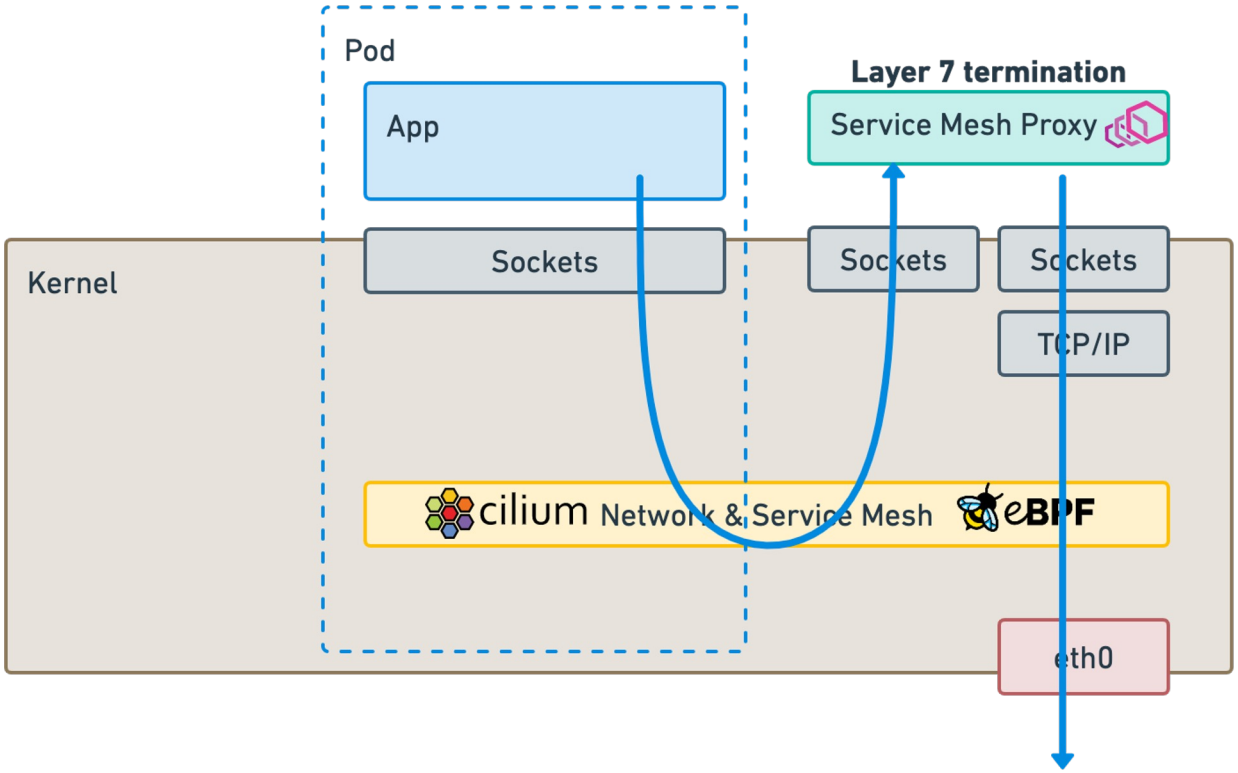
Service Mesh

Cost of sidecar injection

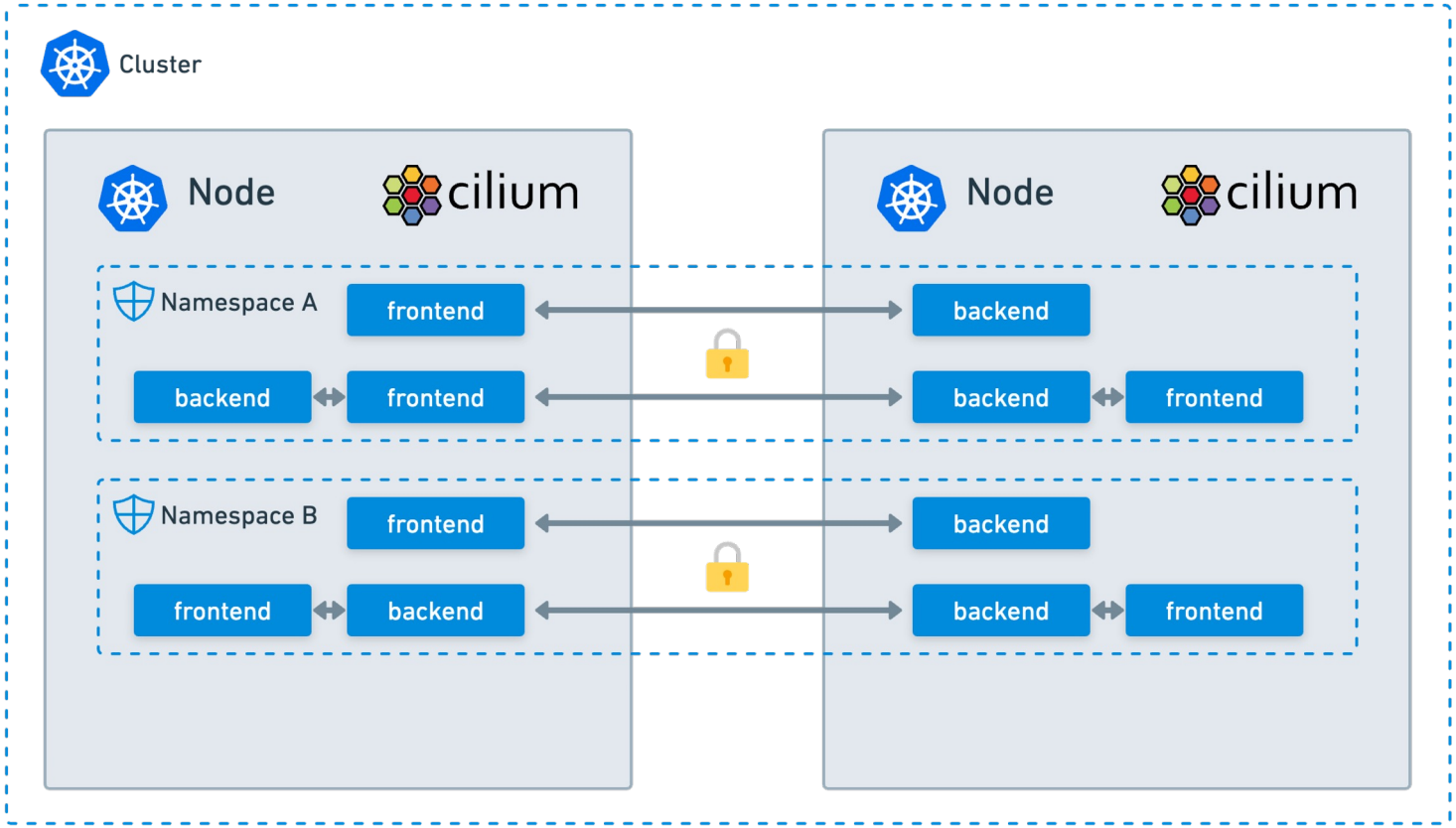


Service Mesh

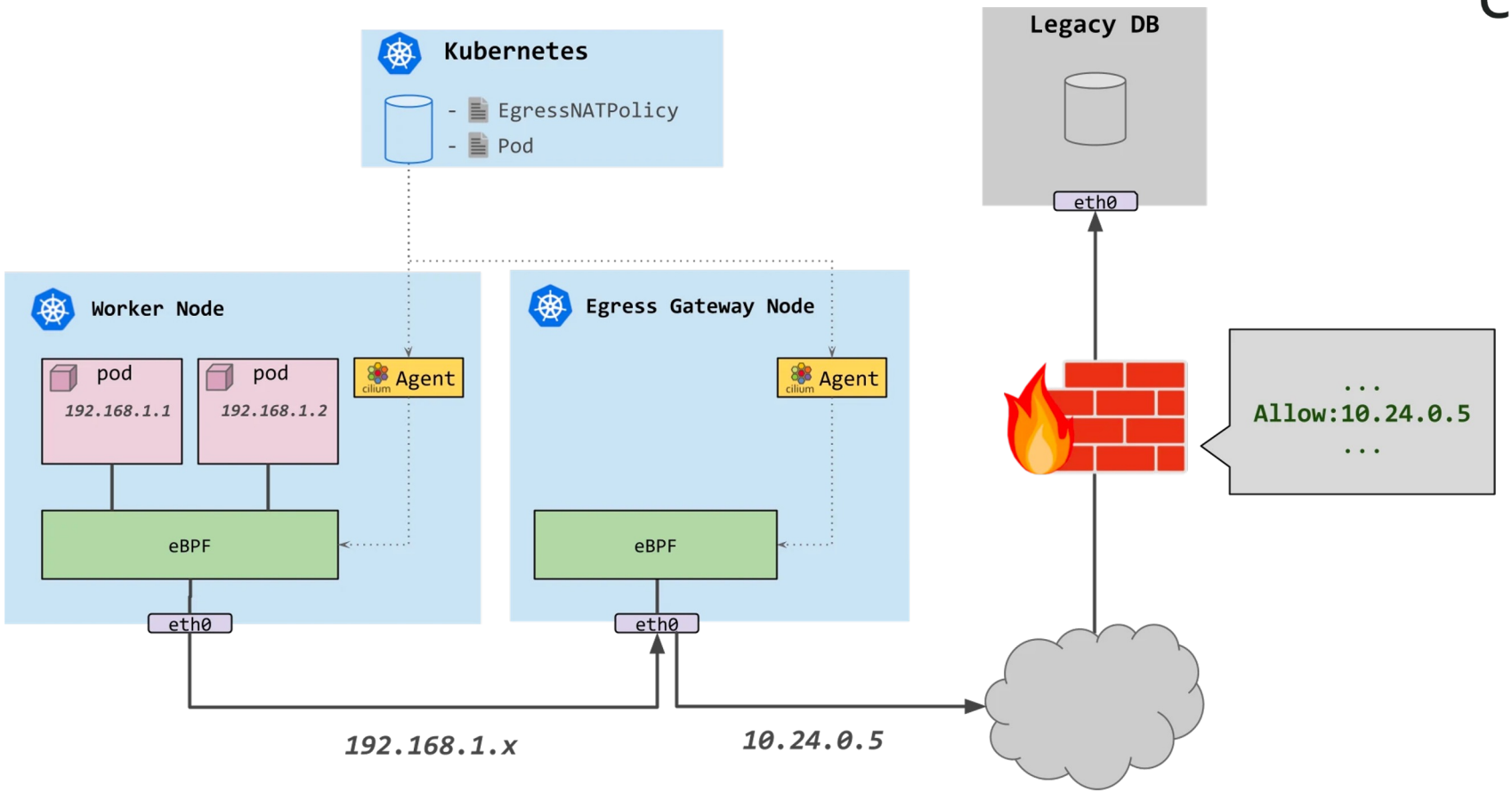
Envoy for Layer 7 termination when needed



Transparent Data Encryption

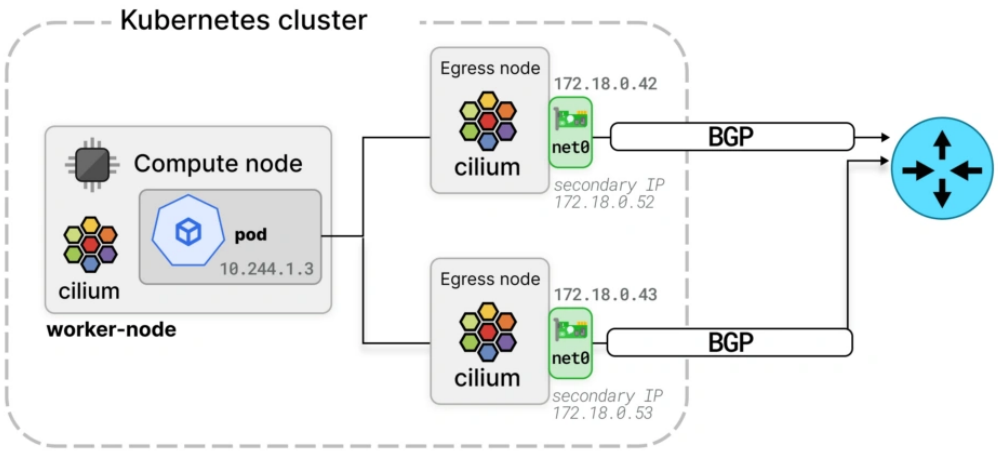


Egress Gateway



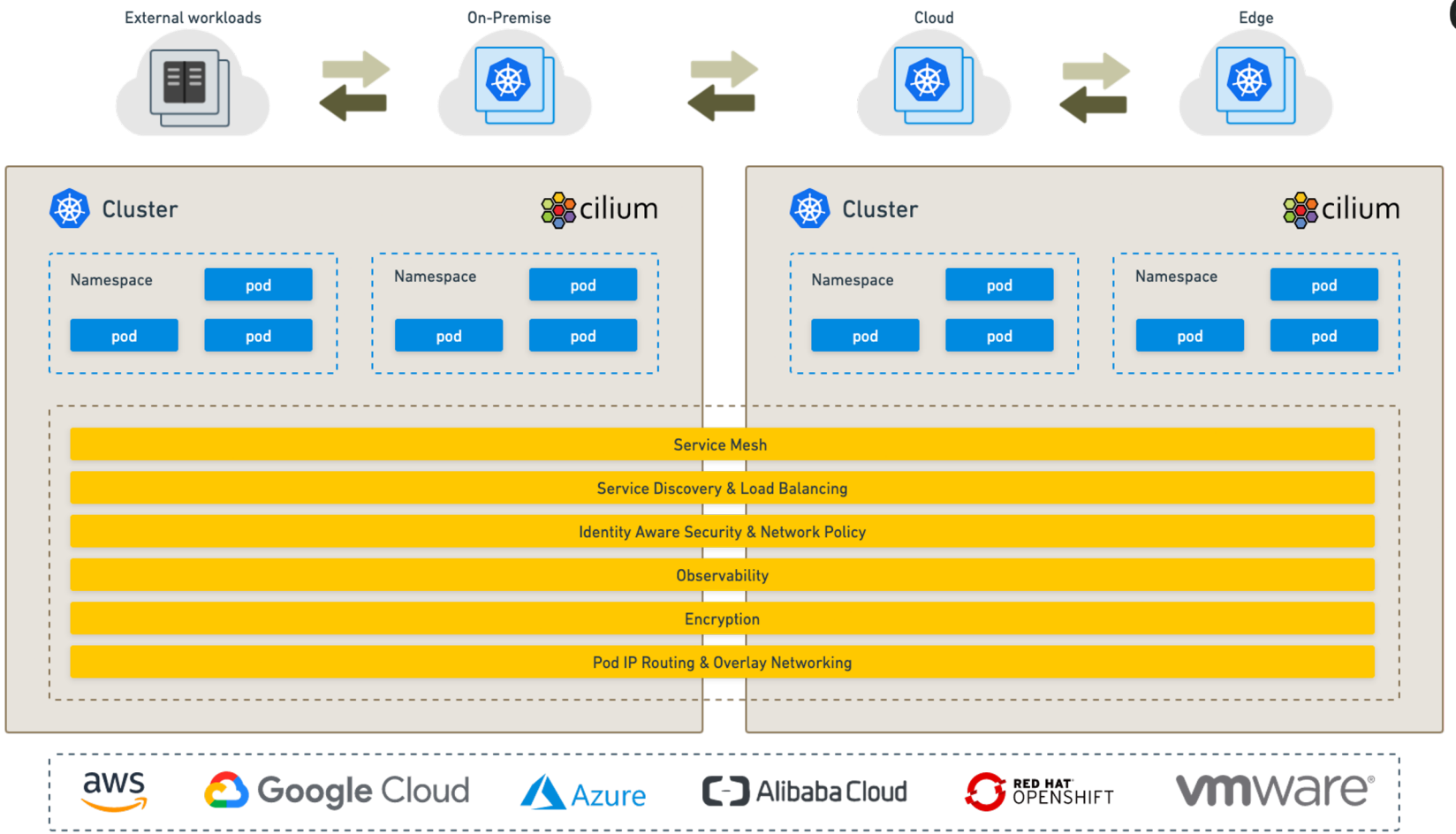
Consistent Peering

Peer into your existing spine and leaf network



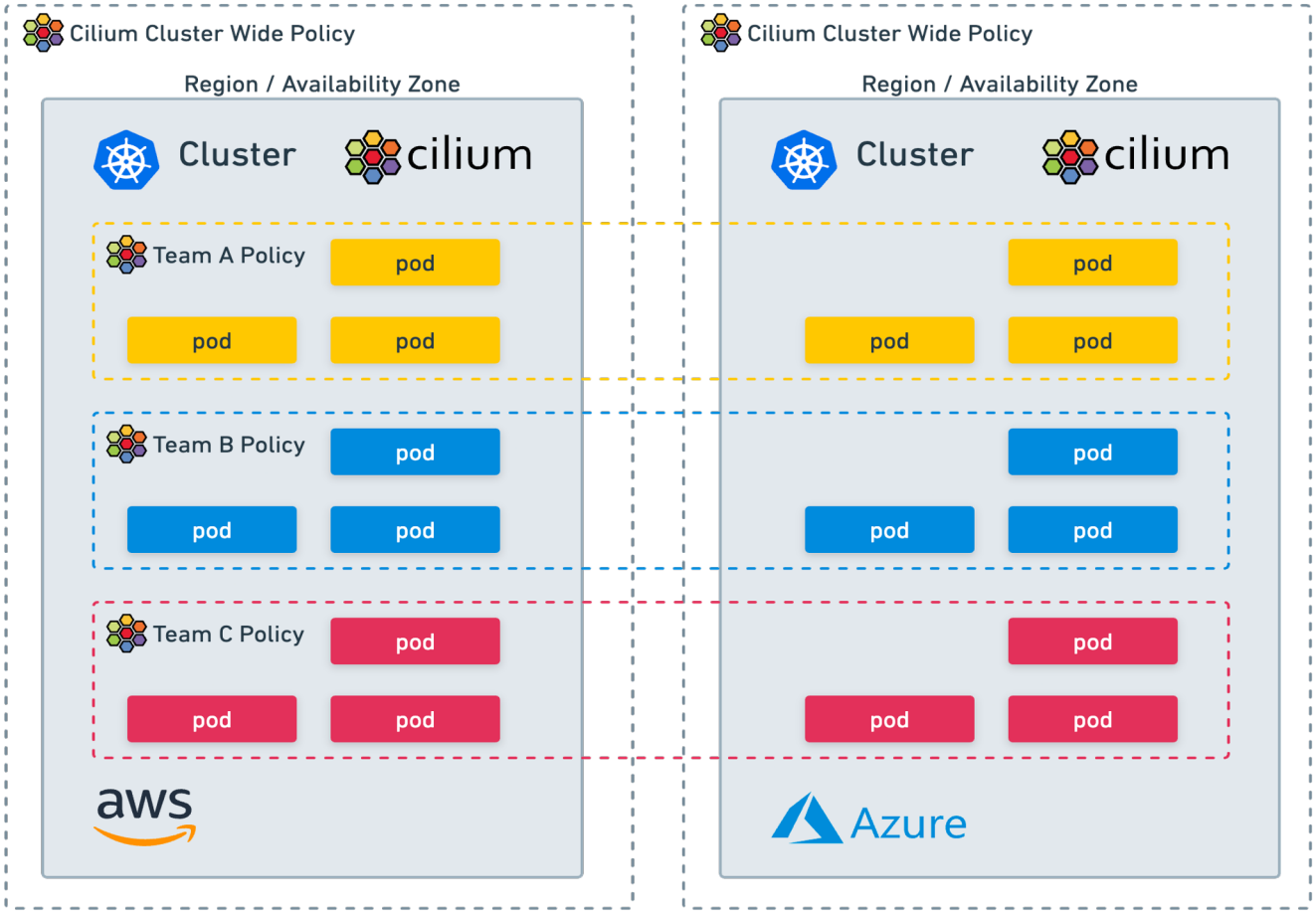
- BGP Support
  - Service Advertisement
  - BFD functionality
  - Supported by CiliumBGP CRD
  - LocalPref
  - Communities
  - IPAM
  - ClusterIP and ExternalIP advertisments

Cluster Mesh - Introduction

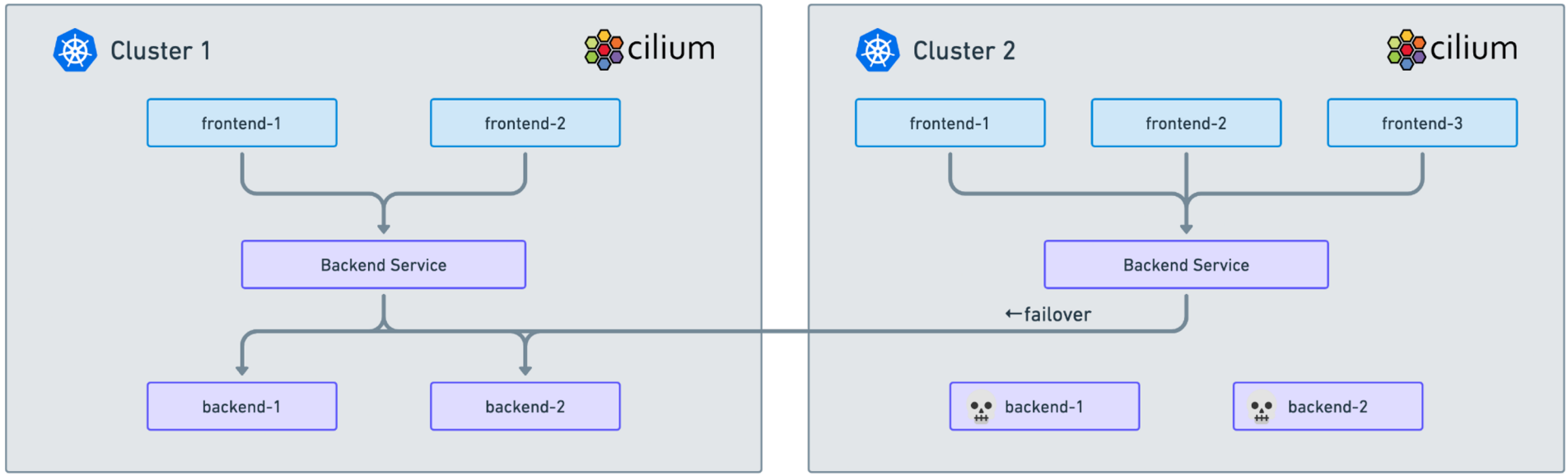


Multi-Cluster Security

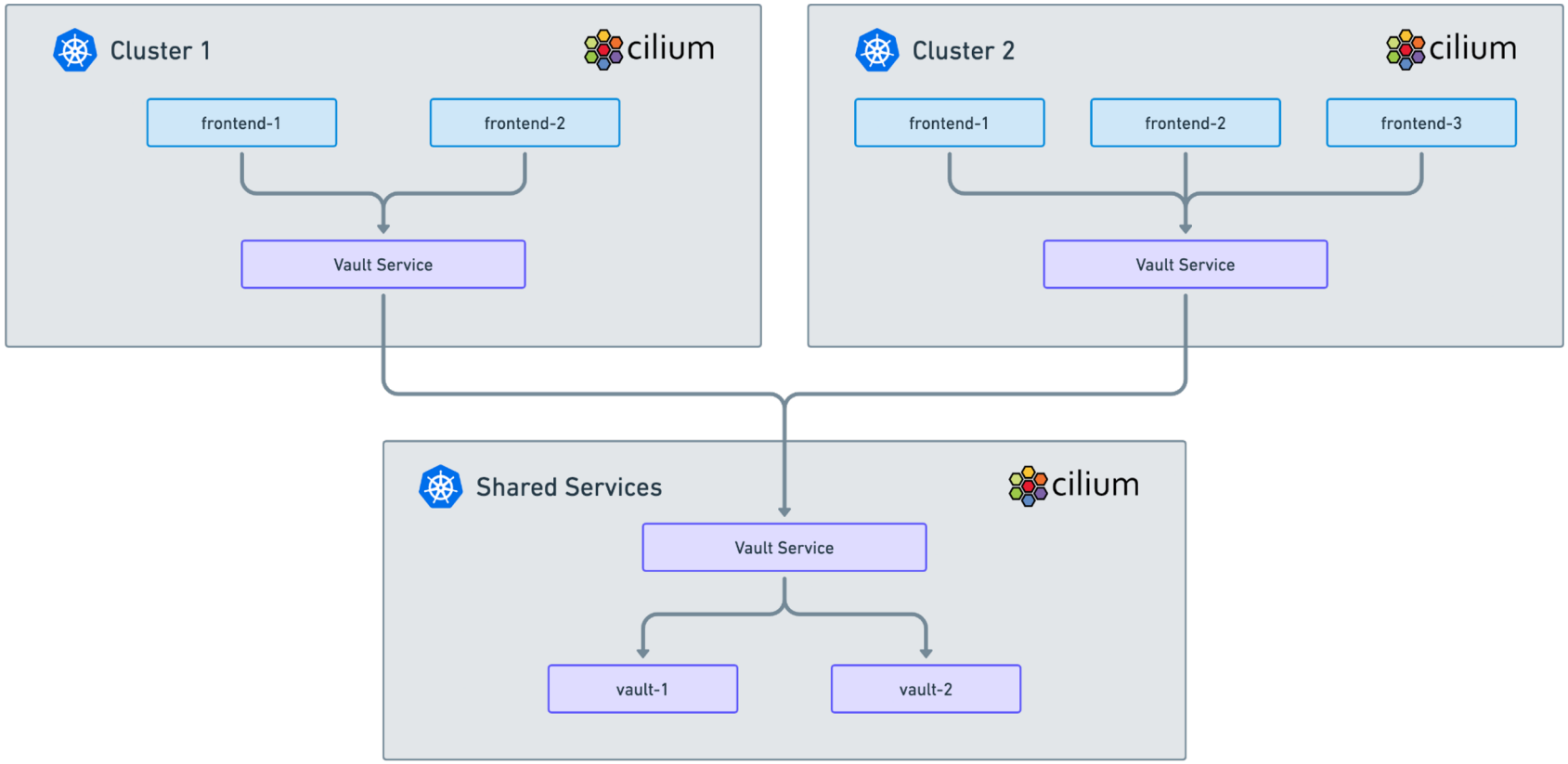
Policy Enforcement across Multiple Clusters



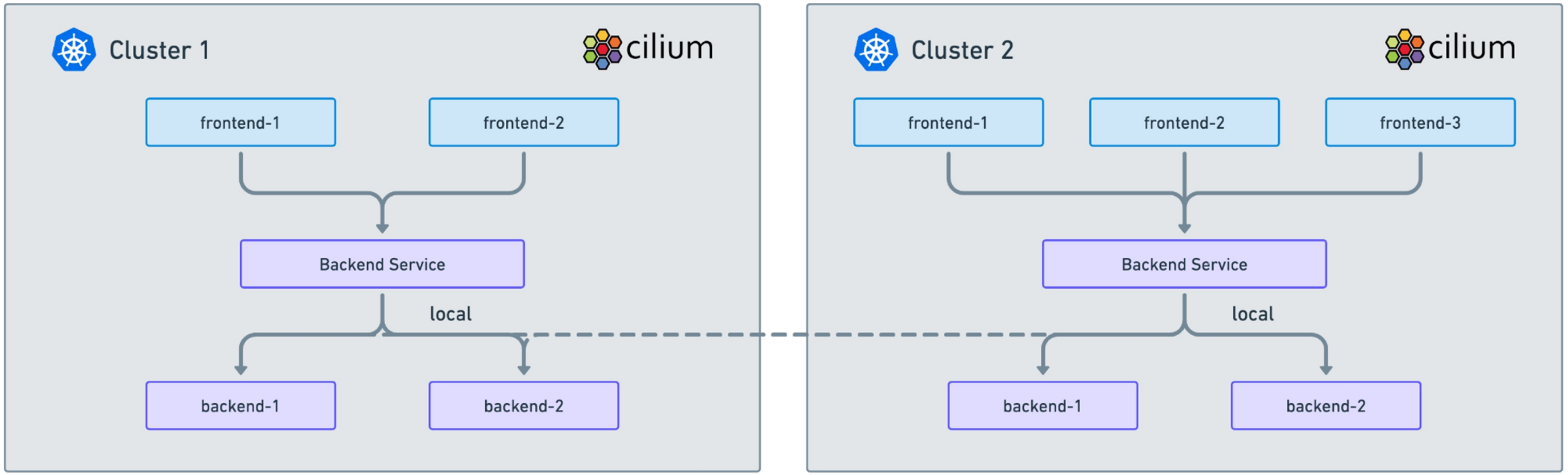
Cluster Mesh - High Availability



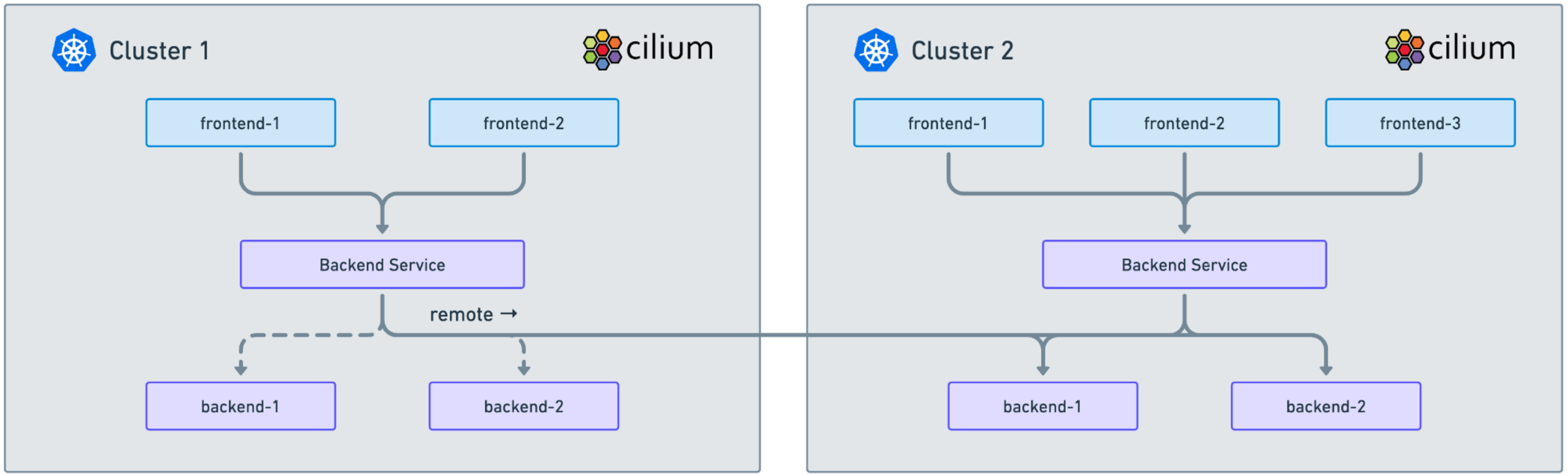
Cluster Mesh - Shared Services



Cluster Mesh - Local Service Affinity

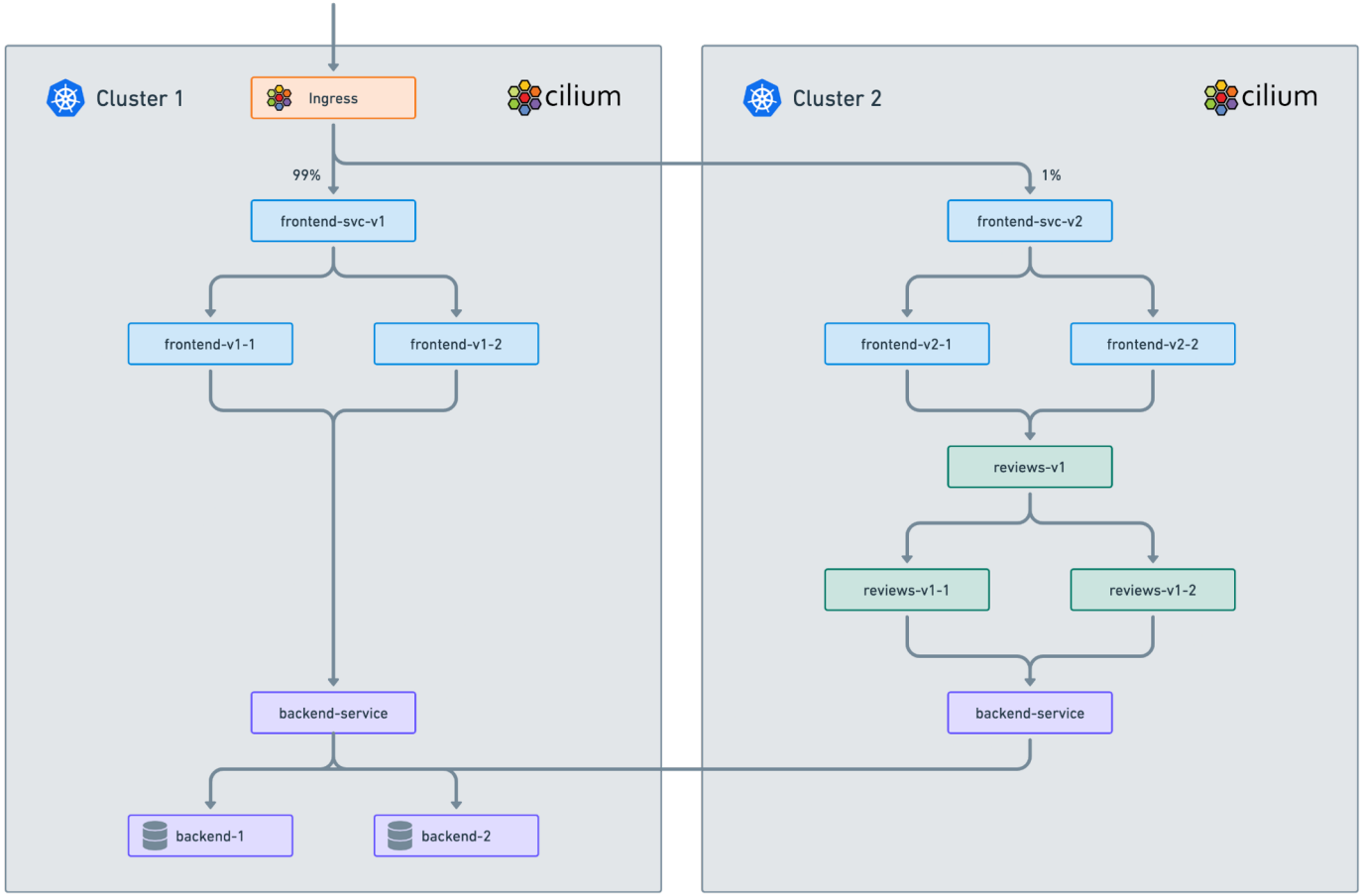


Cluster Mesh - Remote Service Affinity

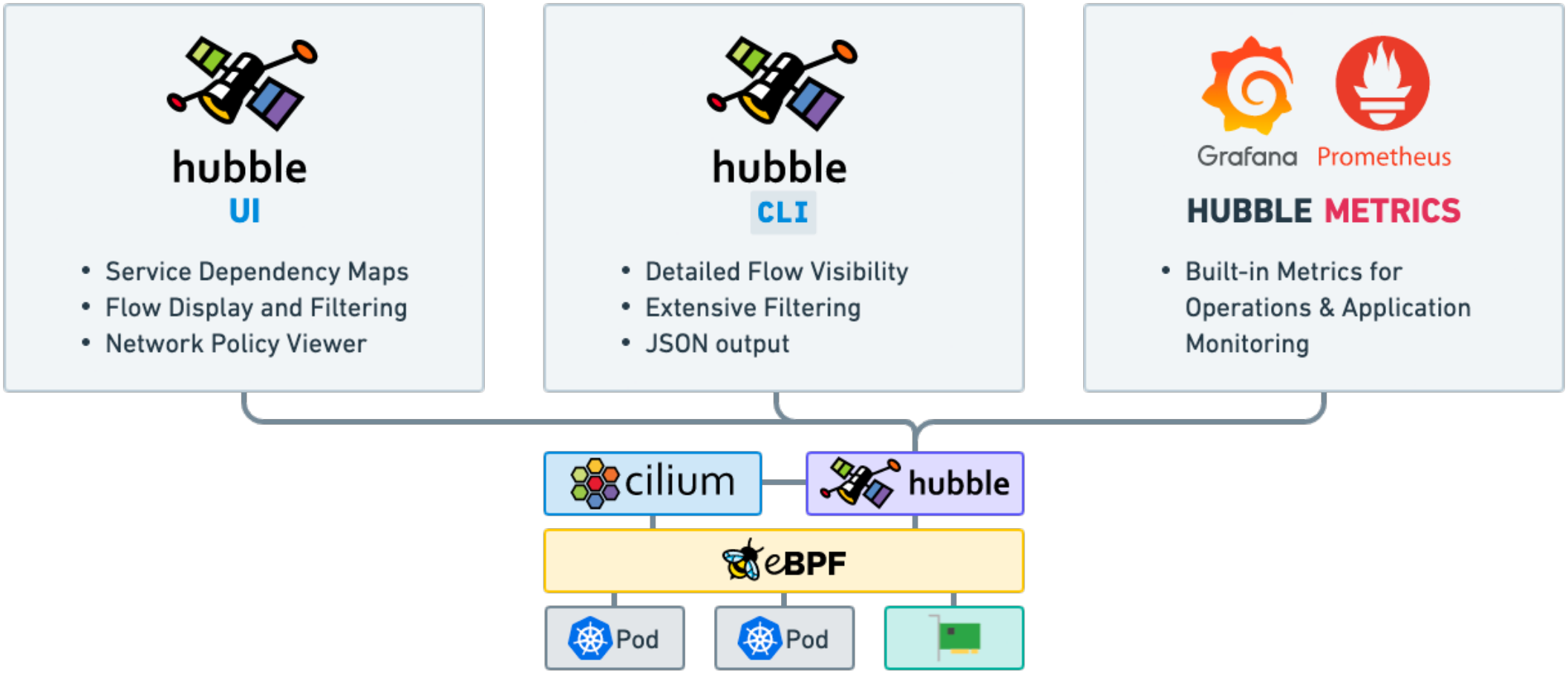


Cluster Mesh with Service Mesh

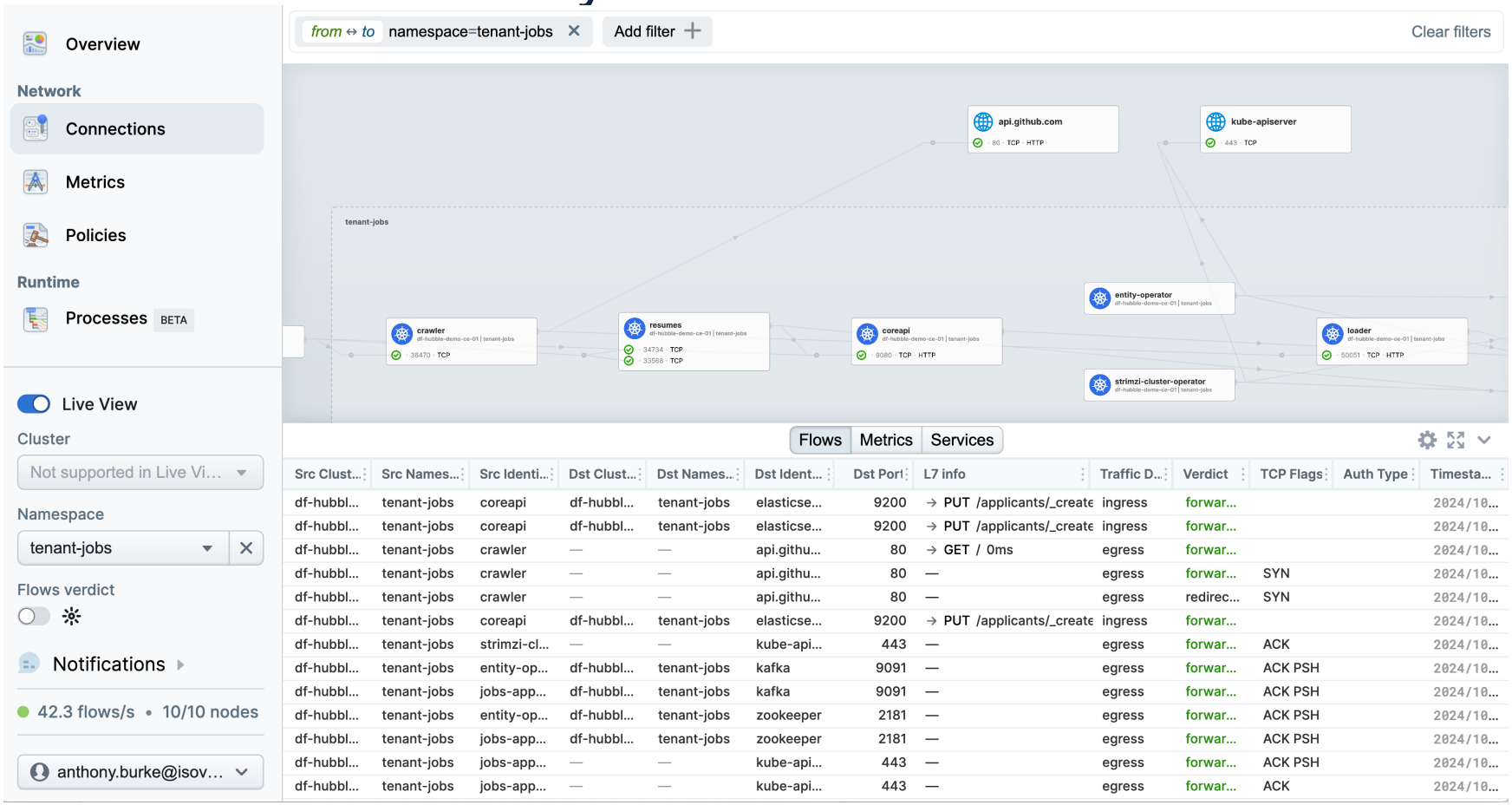
Canary Rollout to other Cluster



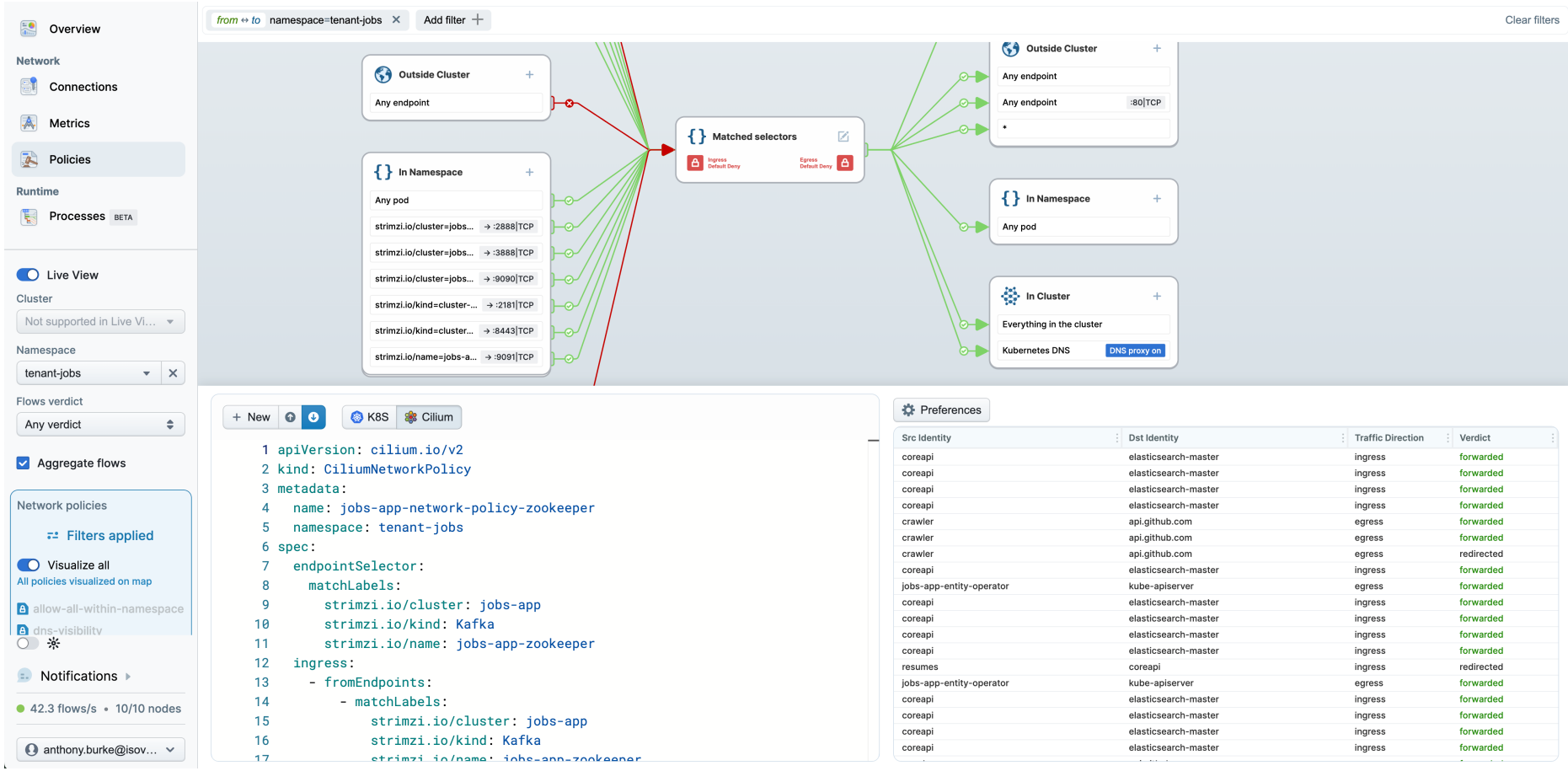
Hubble Overview



Hubble Observability



Hubble Network Policy Editor



Tetragon

Open Source

- Apache 2.0 (userspace) & GNU GPL (eBPF)
- Part of CNCF as a subproject of Cilium

eBPF-based

- Generic low level process events
- In-kernel filtering and enforcement

Kubernetes-native

- Kubernetes metadata in events
- Configuration via custom resources

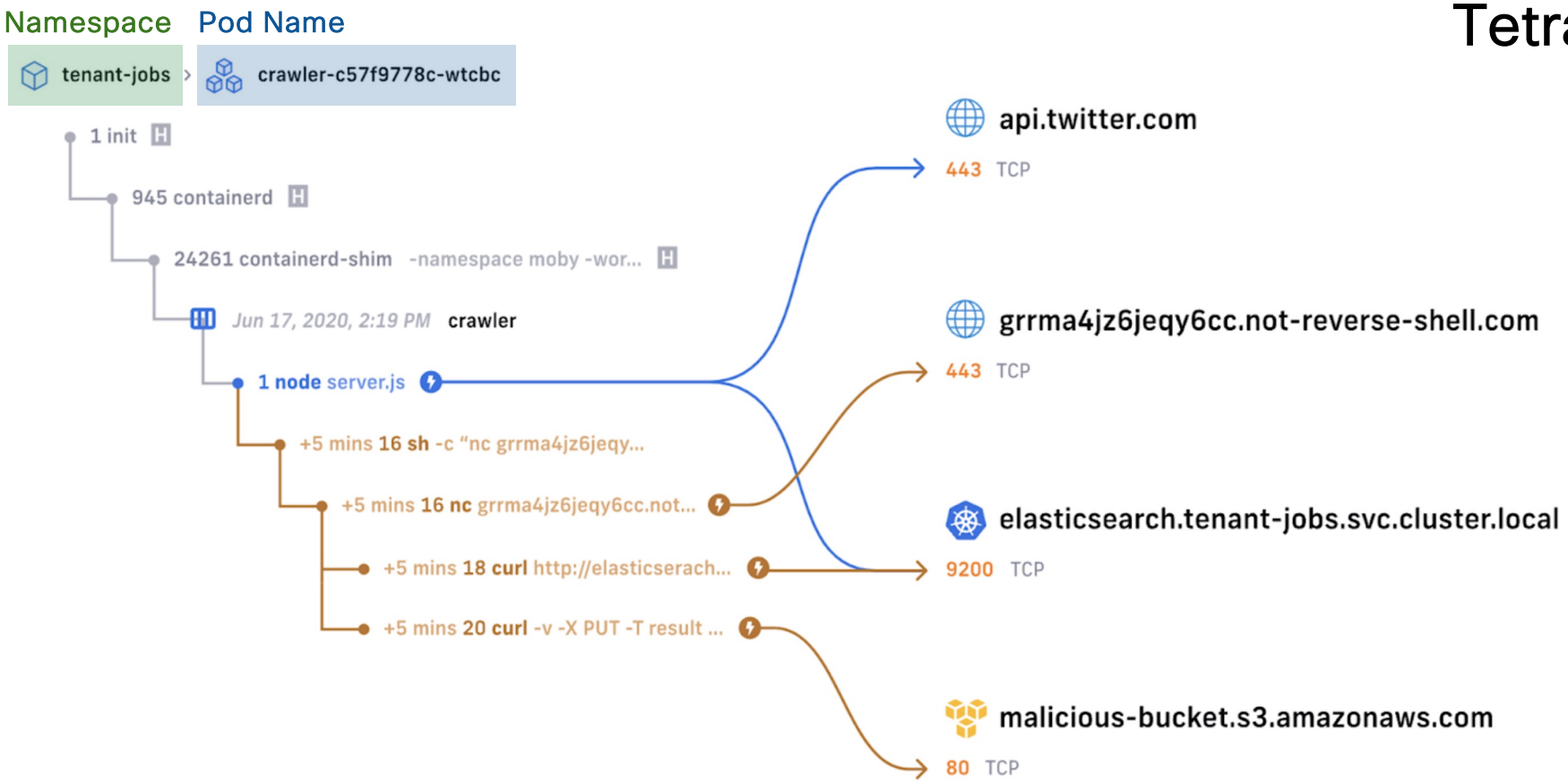
Tetragon Overview

What activities do we care about?

- Network traffic
- File & I/O traffic
- Running of executables
- System Call activity
- Changing privileges & namespace boundaries

Every malicious actor will do one or more of these thigs

Let’s Deep Dive into a Kubernetes Pod



Tetragon

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

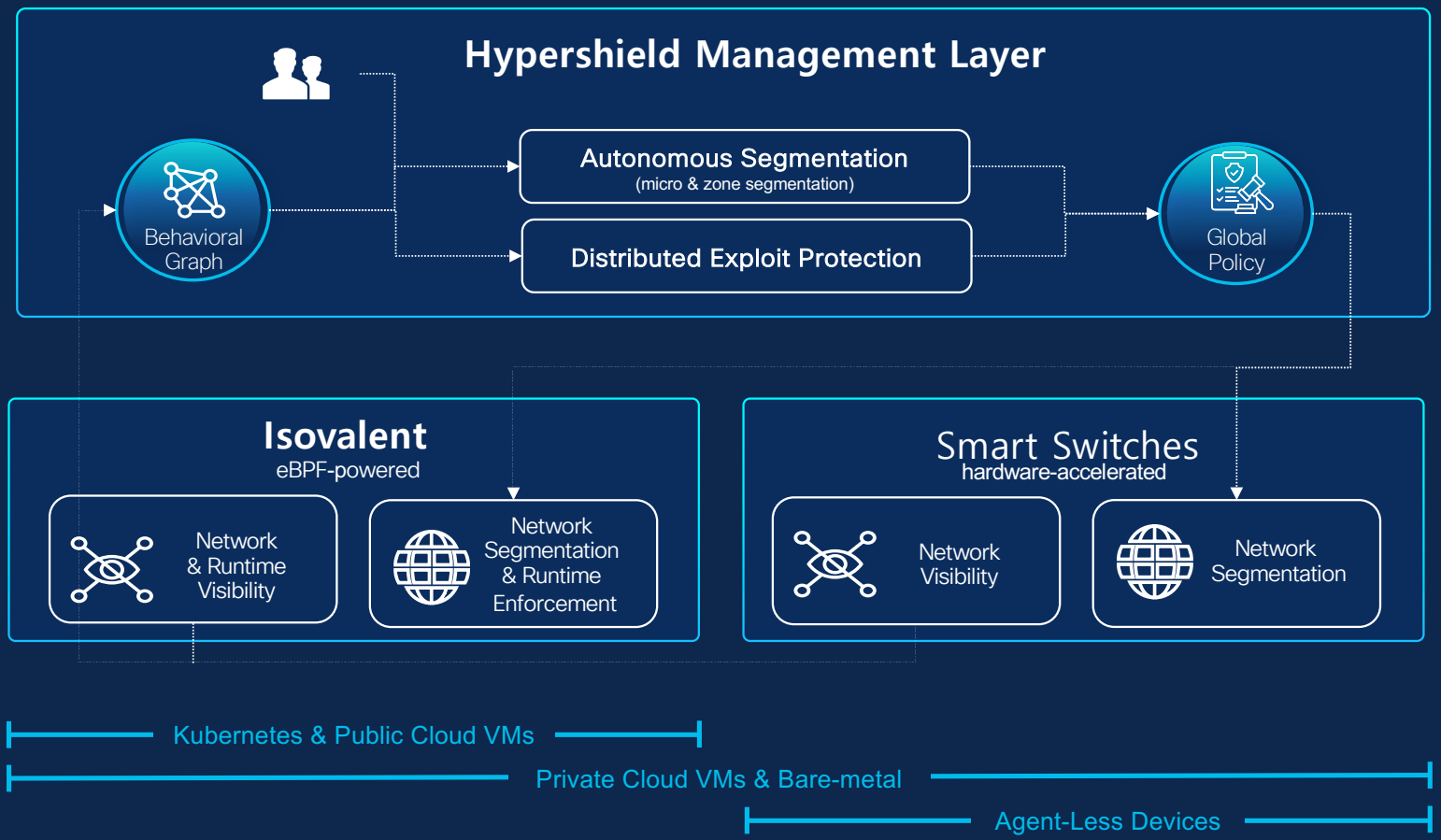
Base technology

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

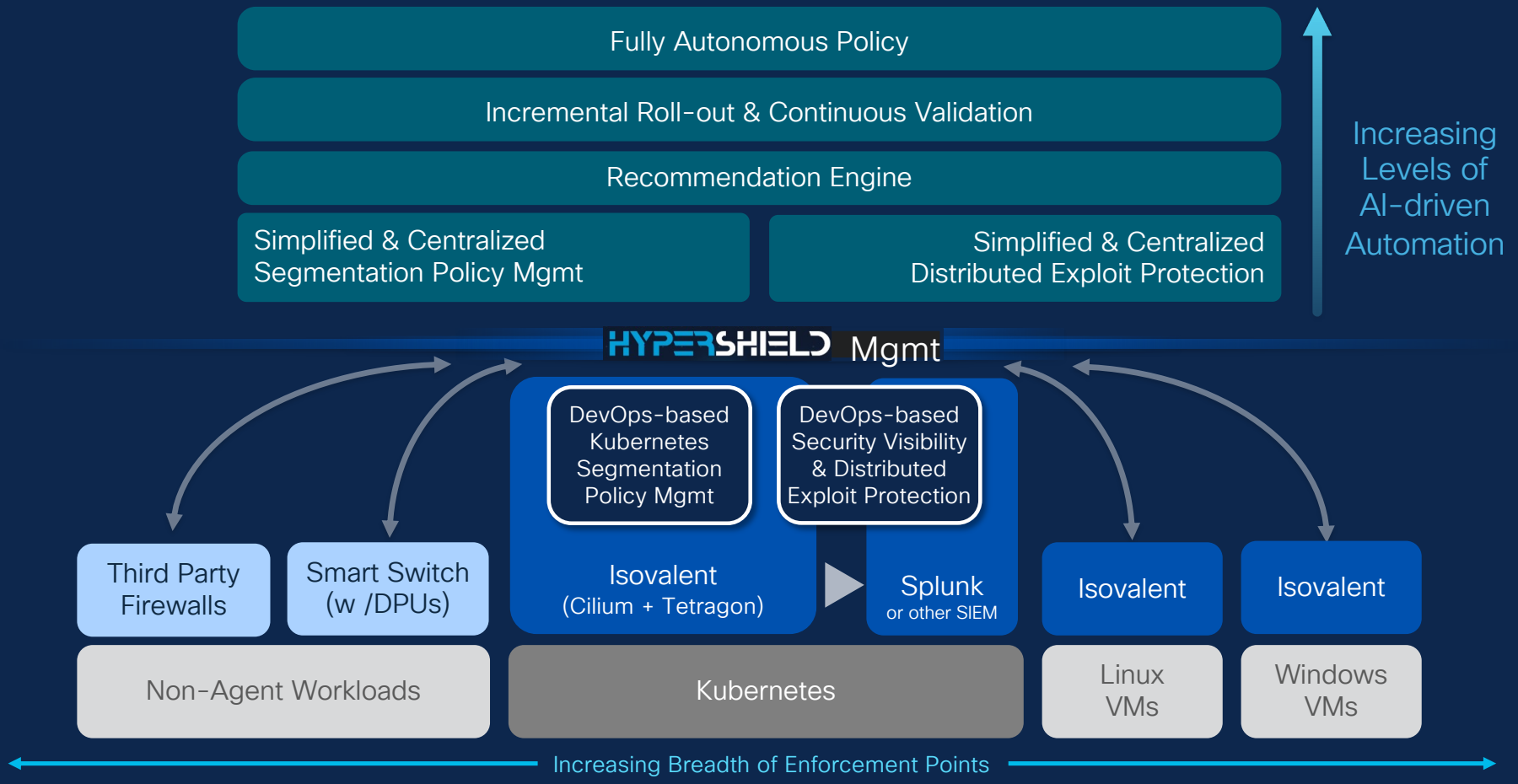
ebpf.io
What is eBPF? - ebook

From Cilium to Hypershield

Hypershield Architecture



Isovalent as a starting point for Hypershield adoption



Cisco N9300 Series Smart Switches

Best of breed platforms for data center services

Target General availability April 2025

N9324C-SE1U

24-port 100G

- Cloud Edge, Zone-Based segmentation, DCI, Top-of-Rack
- 800G Services Throughput
- Silicon One E100 ASIC + AMD DPUs

Target General availability July 2025

N9348Y2C6D-SE1U

48-port 25G, 6-port 400G, 2-port 100G

- DC Top-of-Rack
- 800G Services Throughput
- Silicon One E100 + AMD DPUs

Cisco and AMD – Better Together

Unmatched flexibility, performance, and efficiency

Silicon One

- Rich NX-OS Features and Services
- High-speed connectivity and scalable performance
- Optimized for latency and power efficiency

Routing Switching

EVPN/MPLS/ VXLAN/SR

Rich Telemetry

Line-rate Encryption

Power Efficiency

- Software-defined Stateful Services
- Programmable at all layers: add new services without HW change
- Scale-out services with wire-rate performance
- Power down DPU complex when not used

Large-Scale NAT

IPSEC Encryption

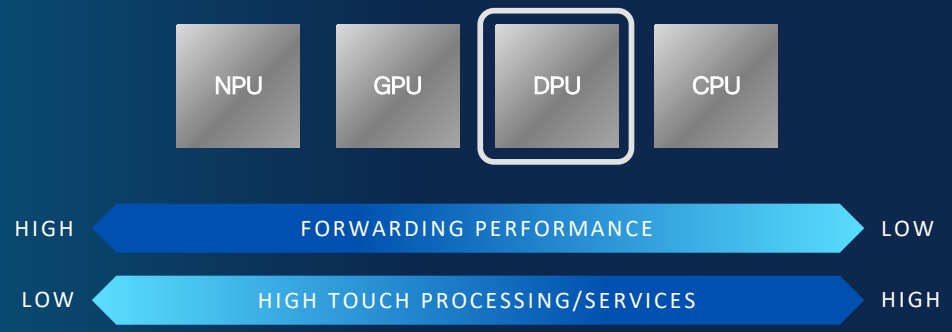
Distributed Firewall

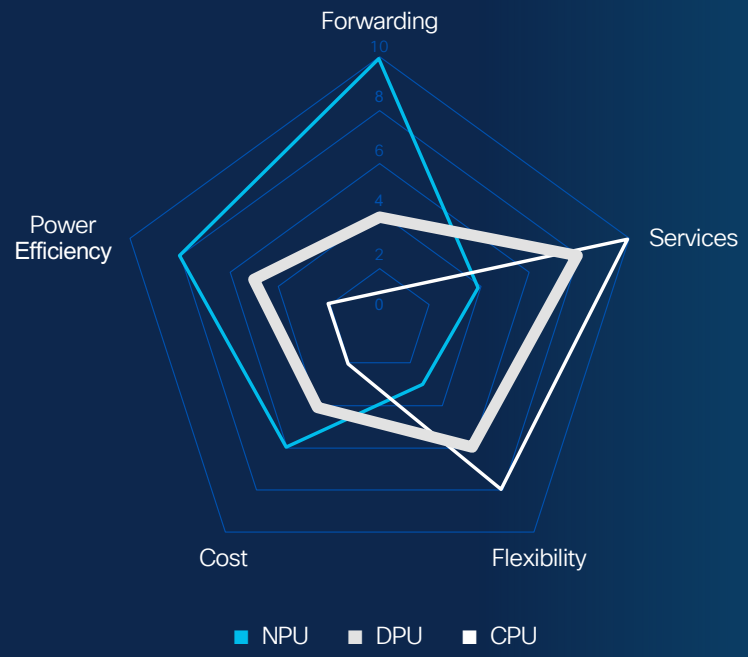
Event-Based Telemetry

DoS Protection

Data Processing Units

A game changer to deliver network services at scale





Higher scores are better

A Platform to Enable Stateful Services

Cisco N9300 Series Smart Switches

Telemetry and Analytics

- Packet inspection
- Timestamping
- DDoS offload

Performance and Optimization

- Virtual network bridging
- High scale traffic filtering

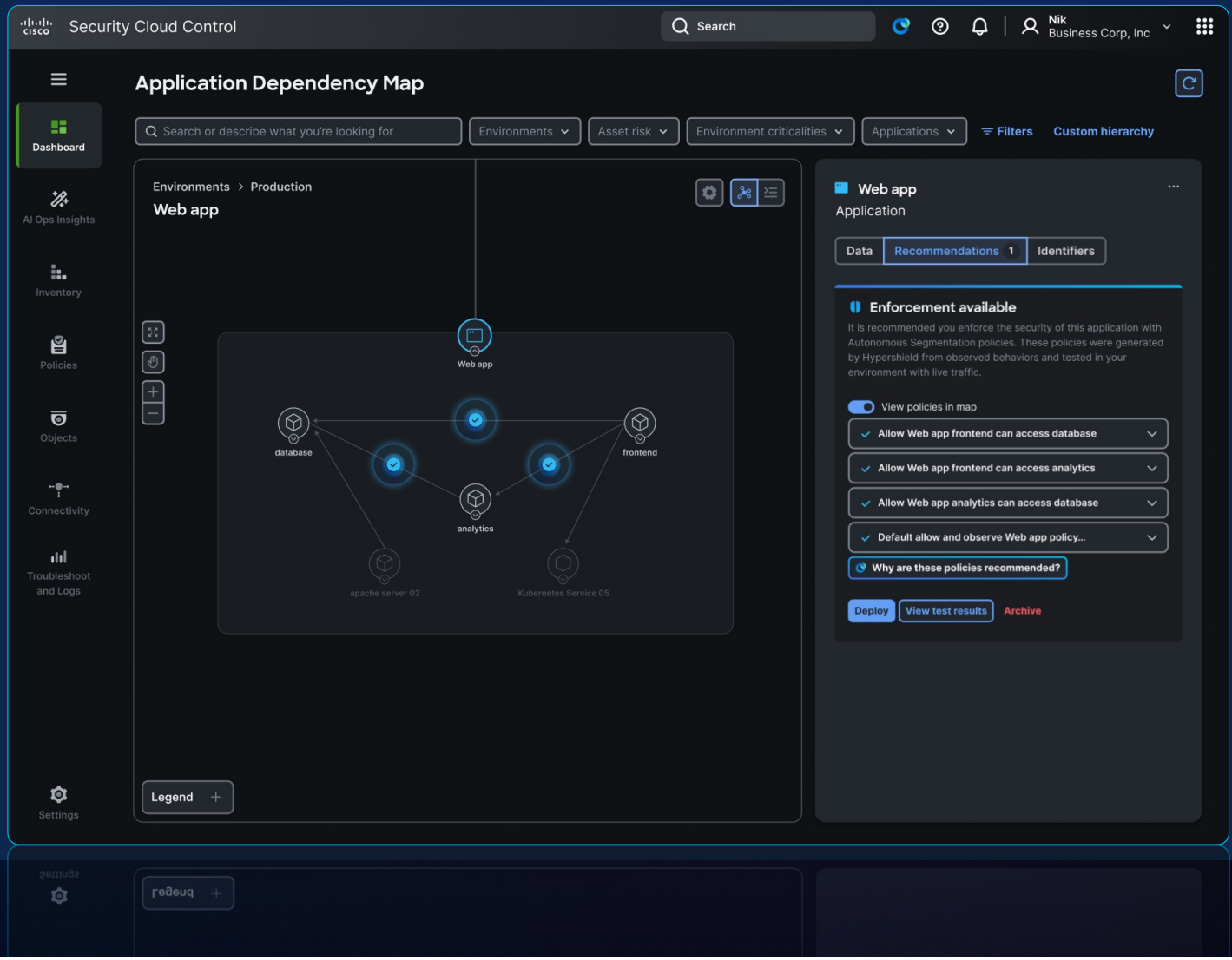
Security and Policy

- Stateful segmentation
- Service chaining/steering

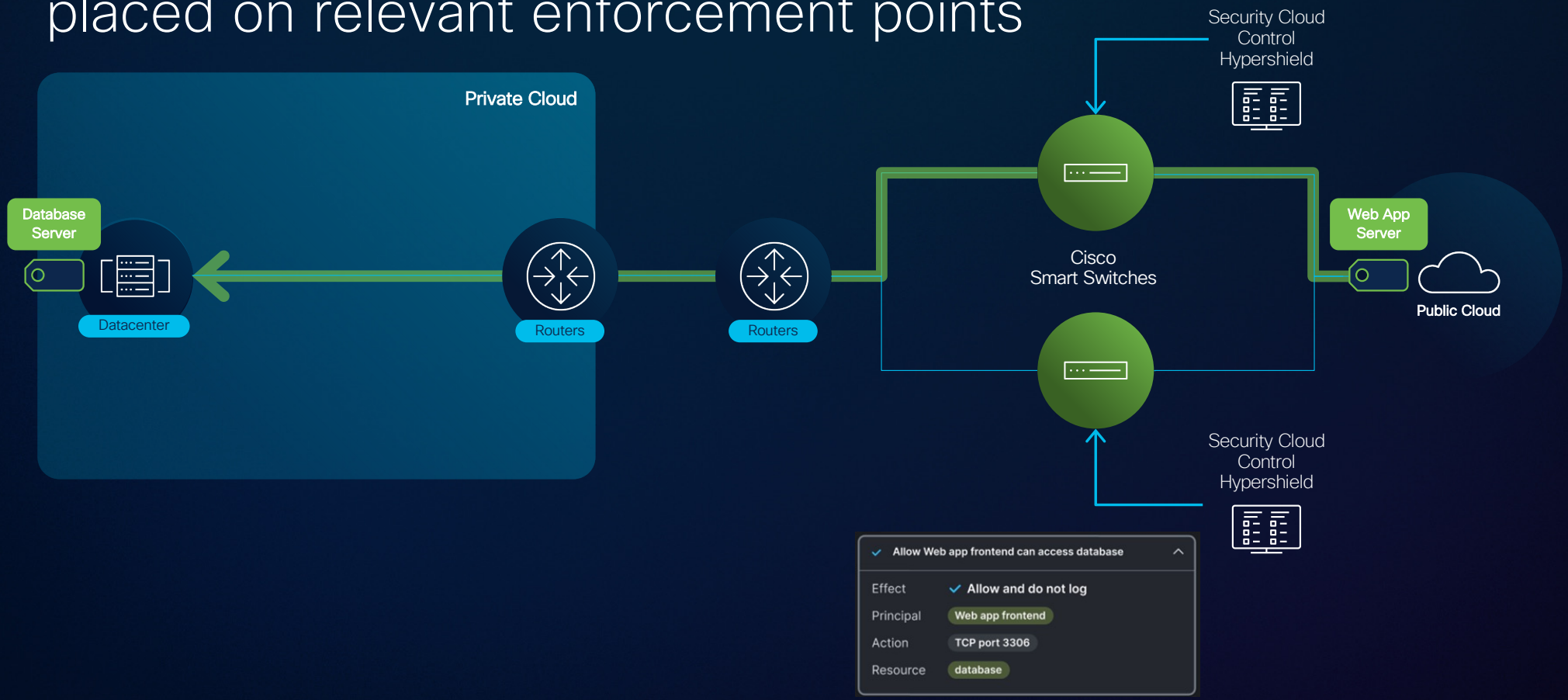
Network services offload

- Encryption
- Large scale NAT

Hypershield recommends segmentation policies based on observations



Hypershield policy is automatically placed on relevant enforcement points



Děkujeme za Vaši pozornost

Následující Tech Club webinář:

15.4. Co je nové v SP a ve světě rozlehlých sítí?

Přednášející: Peter Morvay

Registrovat se můžete na oficiálním webu **Cisco Tech Club webináře**


CISCO