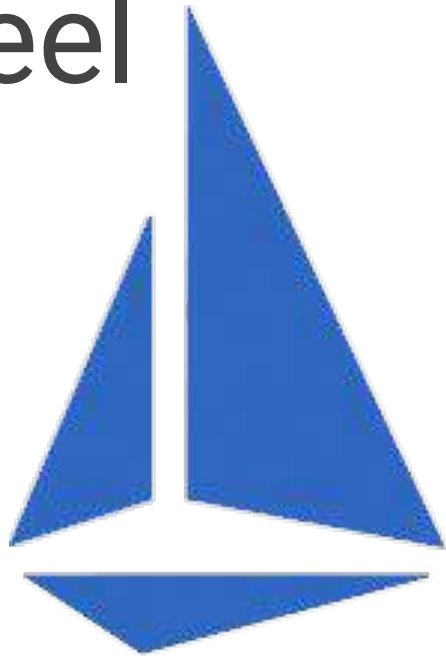
Stop reinventing the wheel with Istio

Mete Atamel
@meteatamel



Google Cloud

Agenda

1. The need for Istio

Containers, Kubernetes, Istio

2. What is Istio?

Istio at the high level, setup

3. Building Blocks

Envoy, Mixer, Pilot, Istio-Auth

4. Add-ons

Grafana, Prometheus, Zipkin, ServiceGraph

5. Traffic Management

Request Routing, Discovery & Load Balancing, Failure Recovery & Injection

The need for Istio

Containers, Kubernetes, Istio

Google Cloud

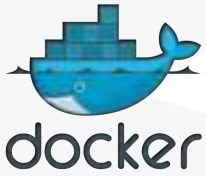
What is a container?

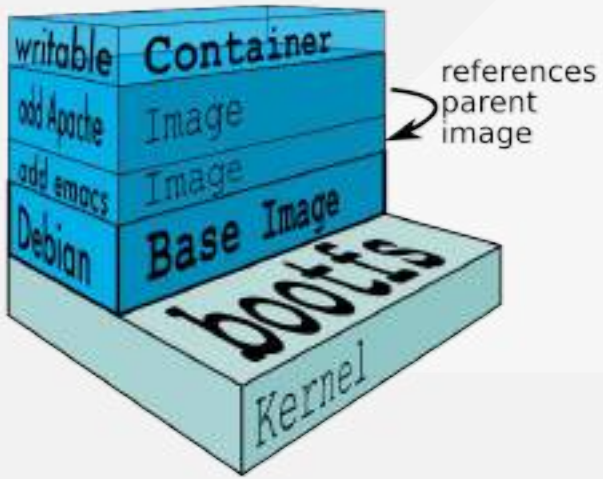
A **lightweight** way to virtualize applications

Linux (or Windows) processes

Lightweight
Hermetically sealed
Isolated

Easily deployable
Introspectable
Composable





Docker: Tooling for the masses

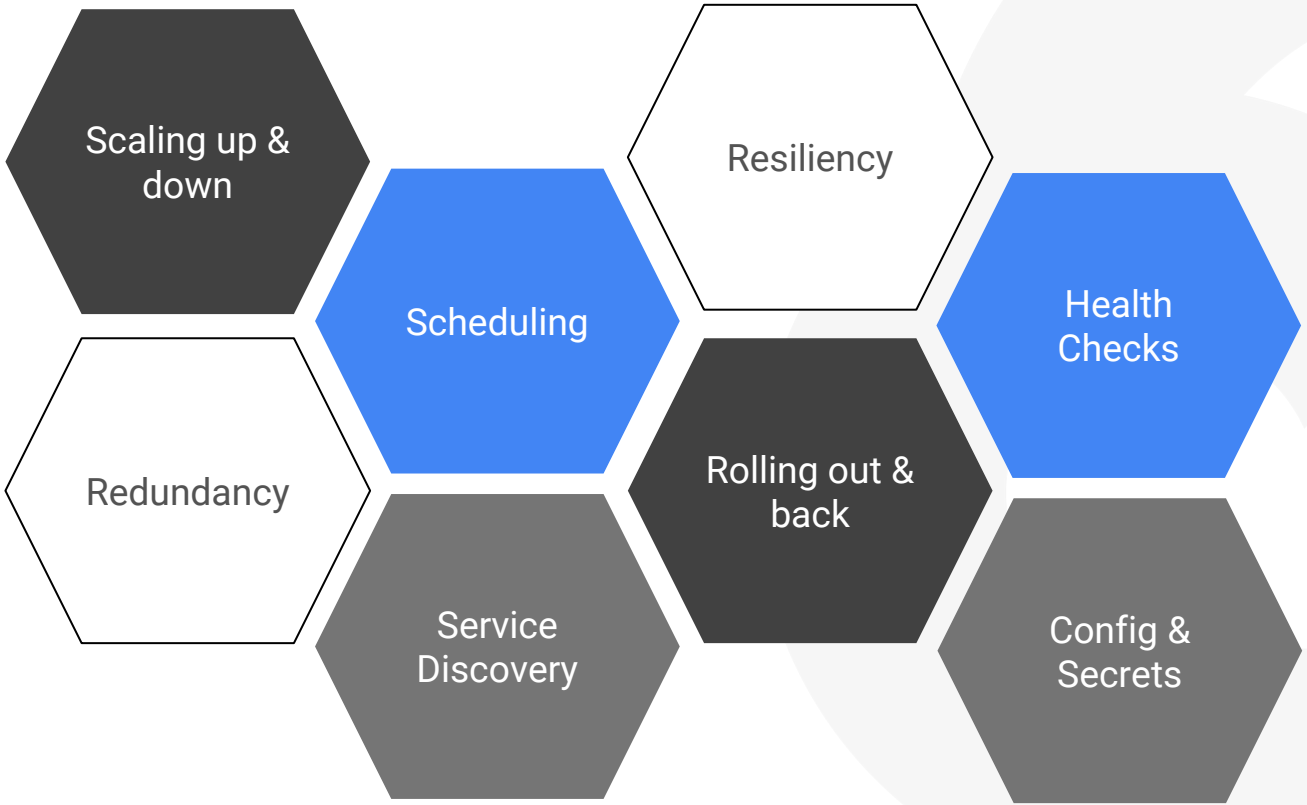
Docker is a container runtime and image format

Dockerfile defines the dependencies, environment and the code to run

Container is a consistent invocation of a Dockerfile


```
FROM debian:latest

RUN apt-get update
RUN apt-get install -y nginx

CMD [“nginx”,”-g”,”daemon off;”]

EXPOSE 80
```

Containers are not enough



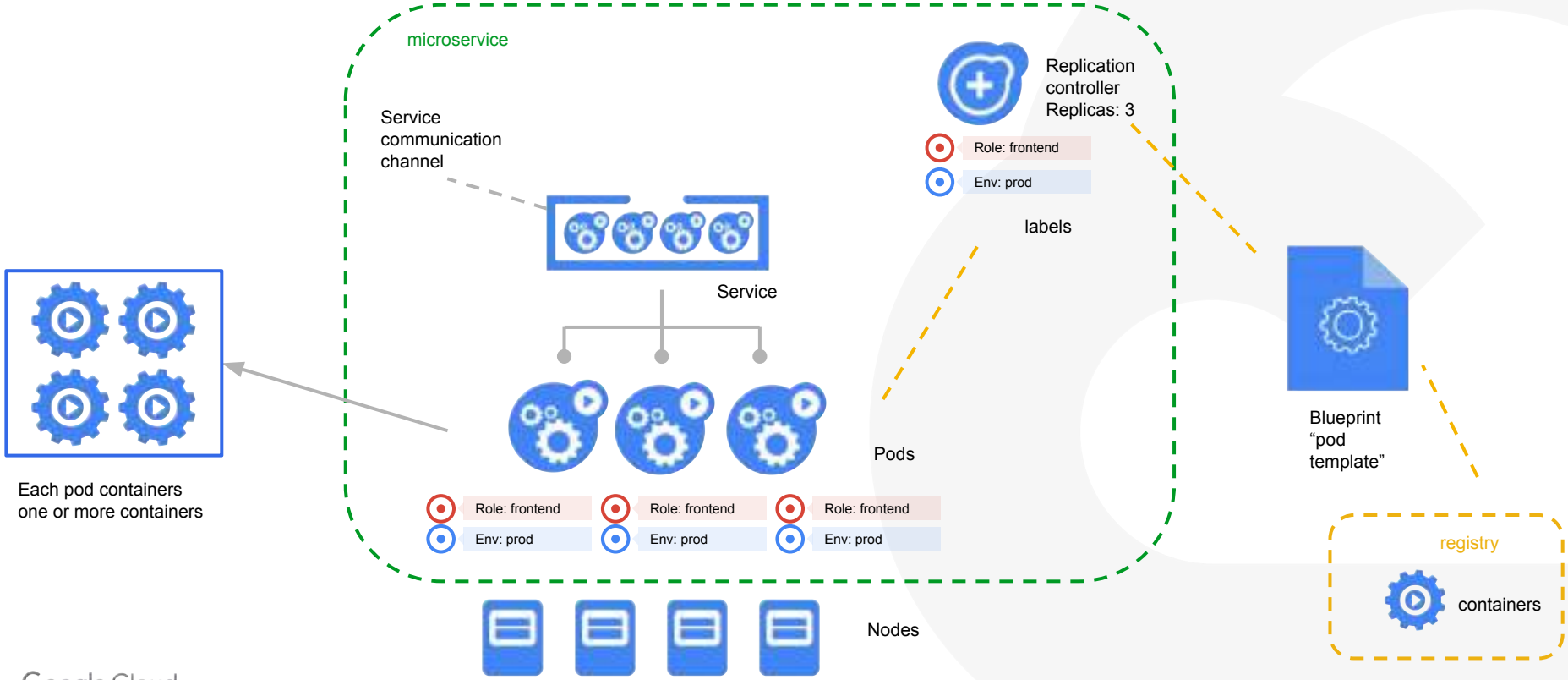
Kubernetes

Κυβερνήτης means *“governor”* in Greek

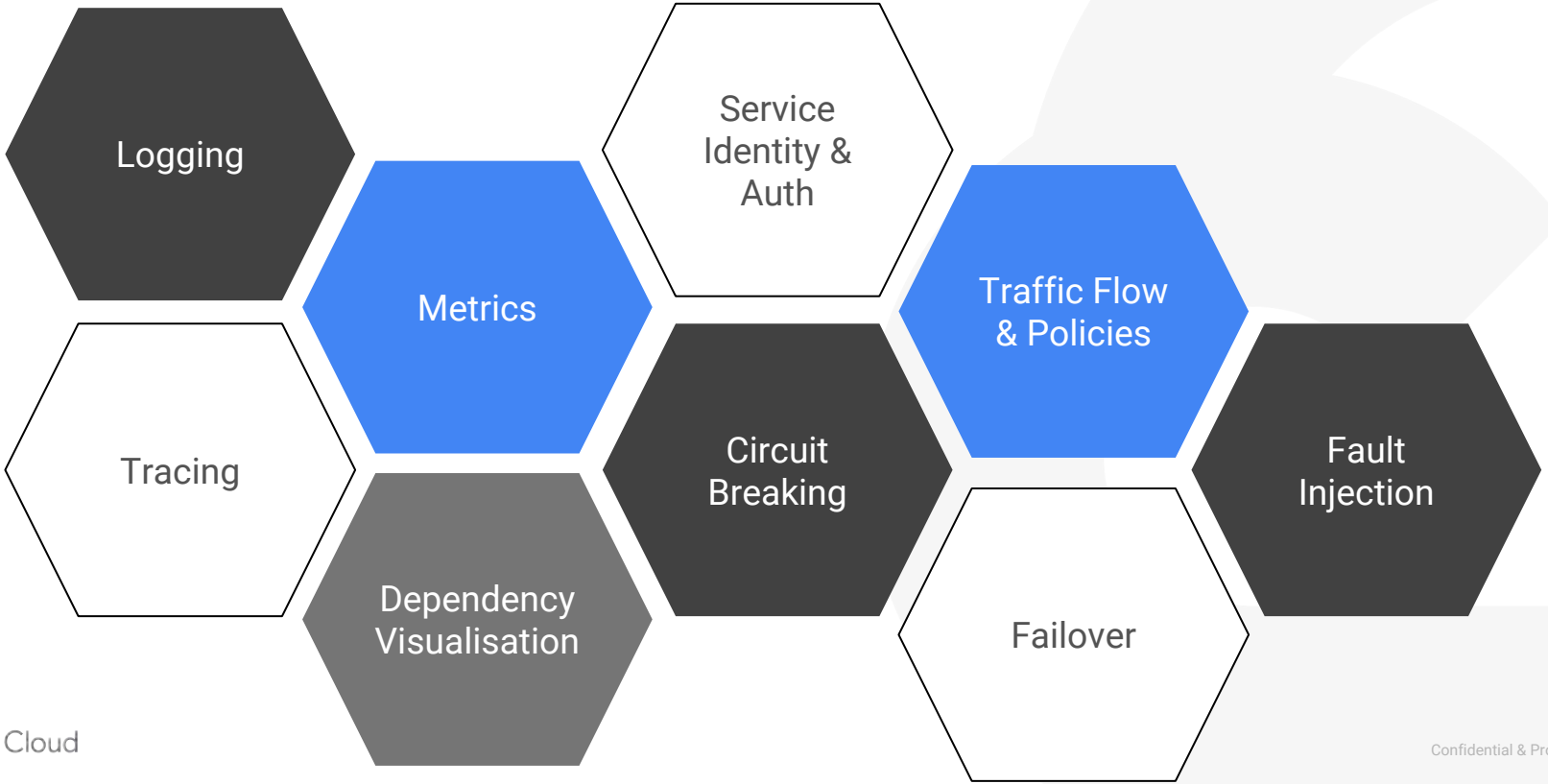
- Manages container clusters
- Inspired and informed by Google’s internal container system called Borg
- Supports multiple cloud and bare-metal environments
- **100% Open source**, written in Go

Manage applications, not machines

Microservices in Kubernetes world



Kubernetes is not enough either



What is Istio?

Istio at the high level, setup

02

Google Cloud

Istio: High level goals

Community maturing and gathering around common tools

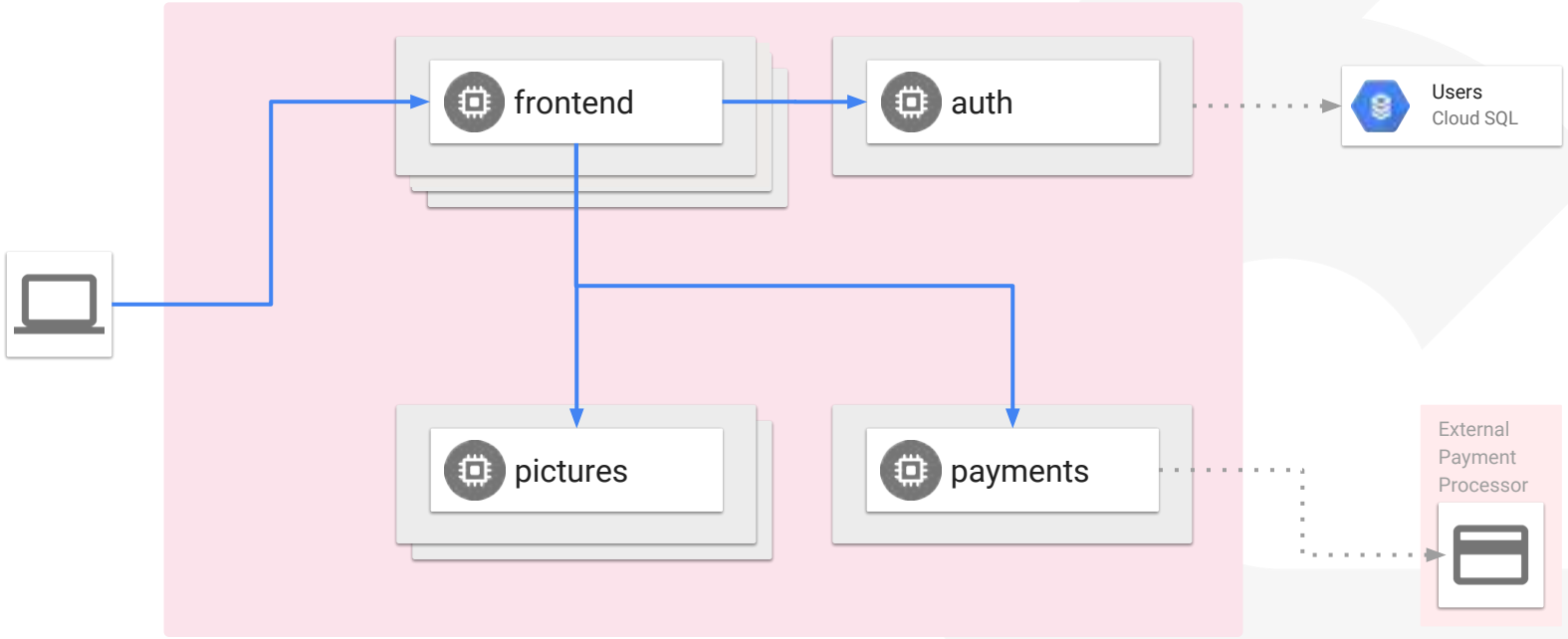
Decouple application code from underlying platform and policies

Istio

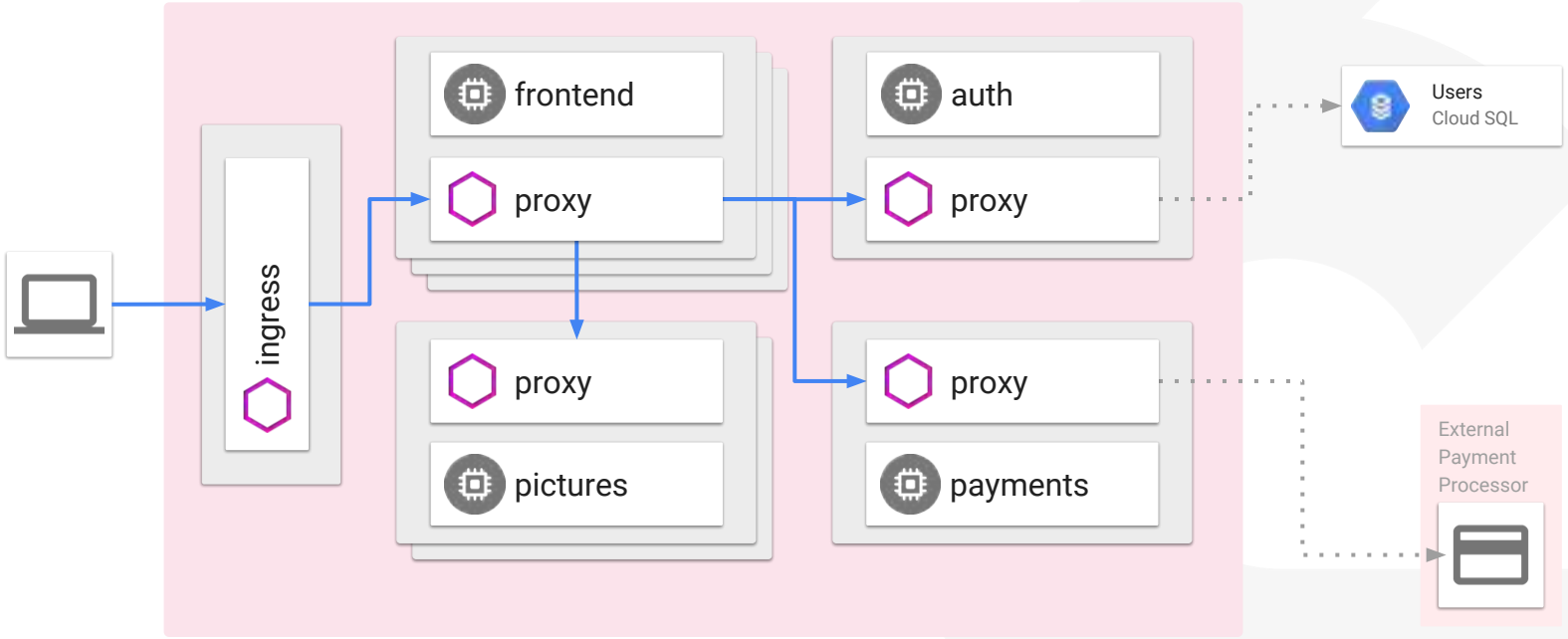
Ιστιο means *“sail”*. An open platform to connect, manage, and secure microservices.

- **Platform support:** Kubernetes, Mesos, Cloud Foundry
- **Observability**: Metrics, logs, traces, dependency visualisation
- **Service Identity & Security**: Provide verifiable identity to services, service-to-service authentication
- **Traffic Management**: Dynamically control traffic between services, ingress/egress routing, fault injection
- **Policy enforcement**: Precondition checking, quota management between services

Istio: At the very high level



Istio: At the very high level



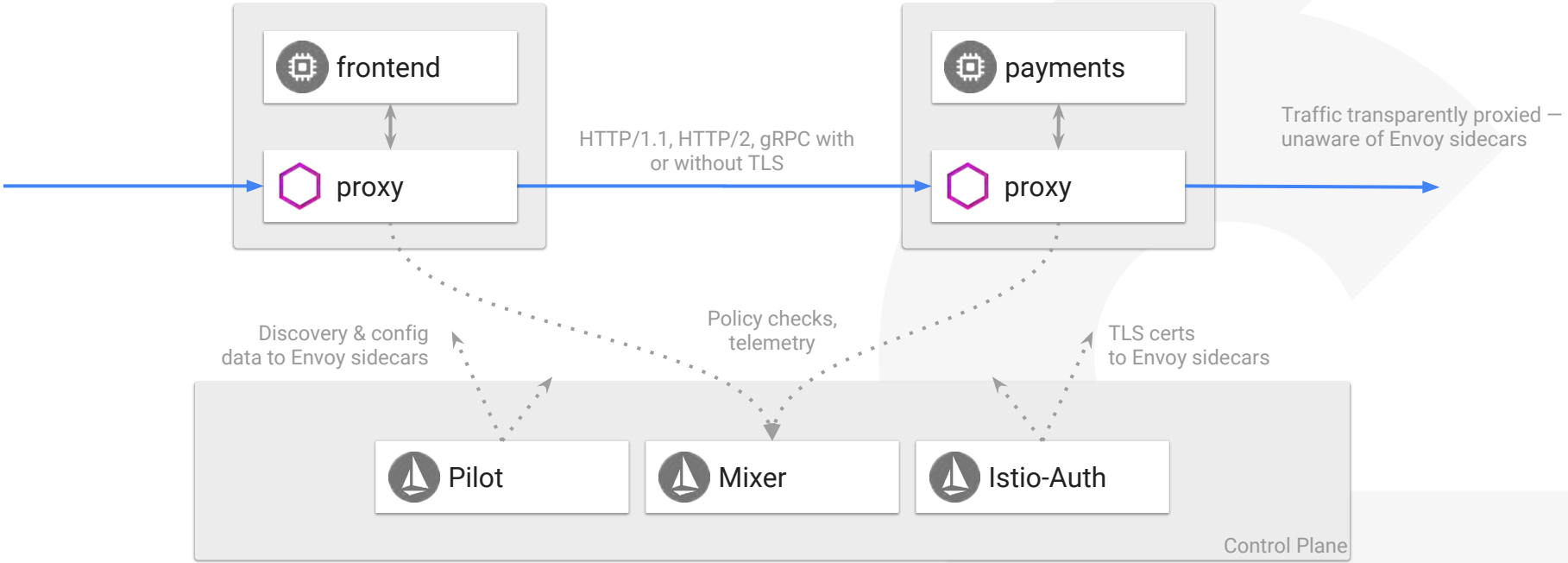
```
$ gcloud container clusters create hello-istio
     --enable-kubernetes-alpha
     --machine-type=n1-standard-2
     --num-nodes=4
     --no-enable-legacy-authorization
     --zone europe-west1-b

Creating cluster hello-istio...done.
Created [https://container.googleapis.com/v1/projects/dotnet-atamel/zones/europe-west1-b/clusters/hello-istio]
NAME         LOCATION       MASTER_VERSION      MASTER_IP       MACHINE_TYPE   NODE_VERSION  NUM_NODES  STATUS
hello-istio  europe-west1-b  1.7.12-gke.0 ALPHA  35.190.192.251  n1-standard-2  1.7.12-gke.0  4          RUNNING


$ kubectl create clusterrolebinding cluster-admin-binding \
    --clusterrole=cluster-admin \
    --user=$(gcloud config get-value core/account)

clusterrolebinding "cluster-admin-binding" created
```

Demo: Install Istio

Building Blocks

Envoy, Mixer, Pilot, Istio-Auth

Istio Architecture



Envoy Proxy

A high-performance proxy in C++, to mediate all inbound/outbound traffic

- Dynamic service discovery
- Load balancing, TLS termination
- HTTP/2 & gRPC proxying
- Circuit breakers, health checks, rich metrics

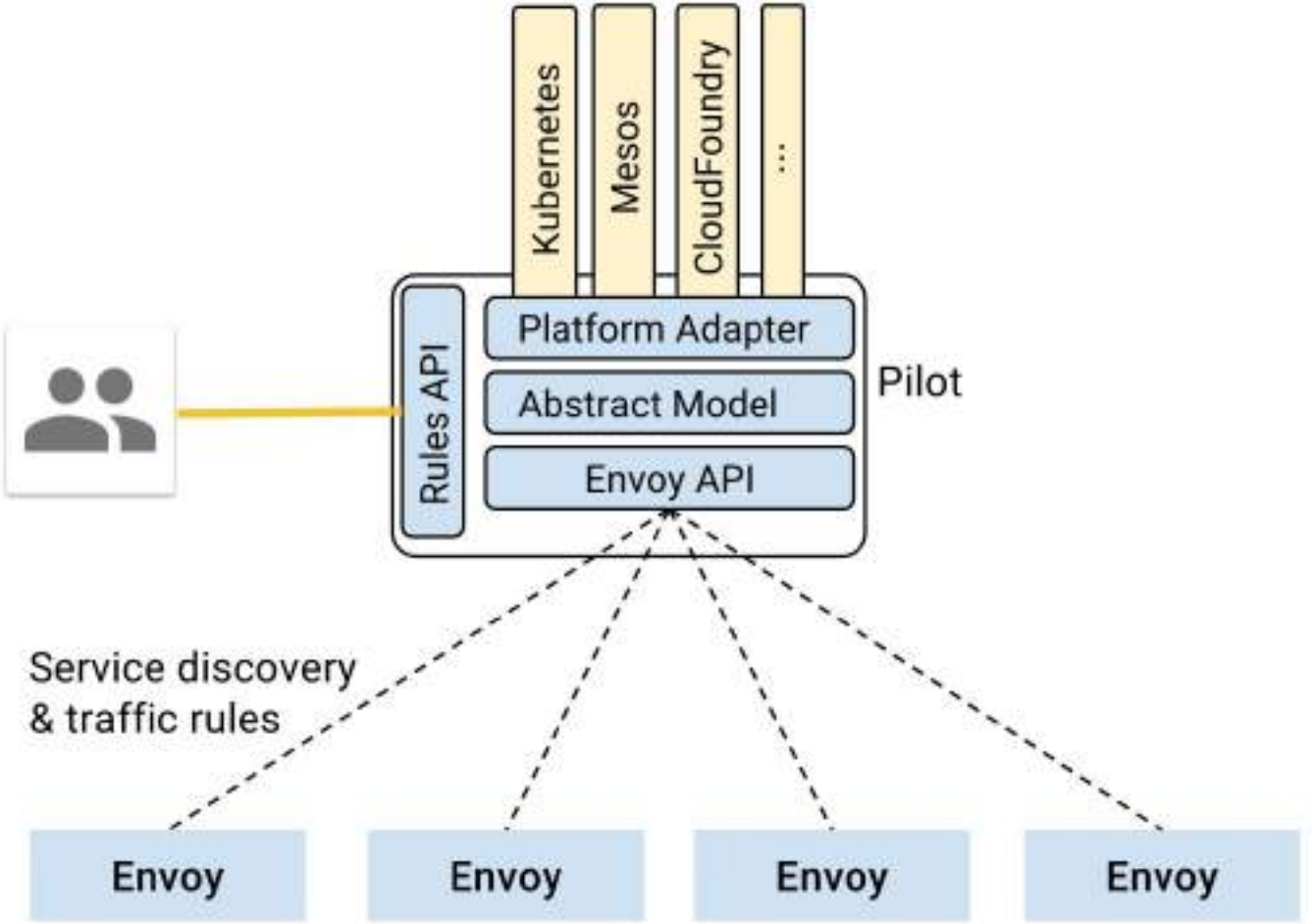
Deployed as a **sidecar** to the relevant service in the same Kubernetes pod

Pilot

Responsible for managing Envoy proxies in the service mesh.

- Service discovery for Envoy
- Traffic management capabilities for routing (A/B testing, canary deployments)
- Resiliency (timeouts, retries, circuit breakers)
- Converts high level routing rules into Envoy specific configurations and propogates them to sidecars at runtime

Pilot



Mixer

1. **Precondition Checking**. Enables callers to verify a number of preconditions before responding to an incoming request from a service consumer.

2. **Quota Management**. Enables services to allocate and free quota (eg. rate limits)

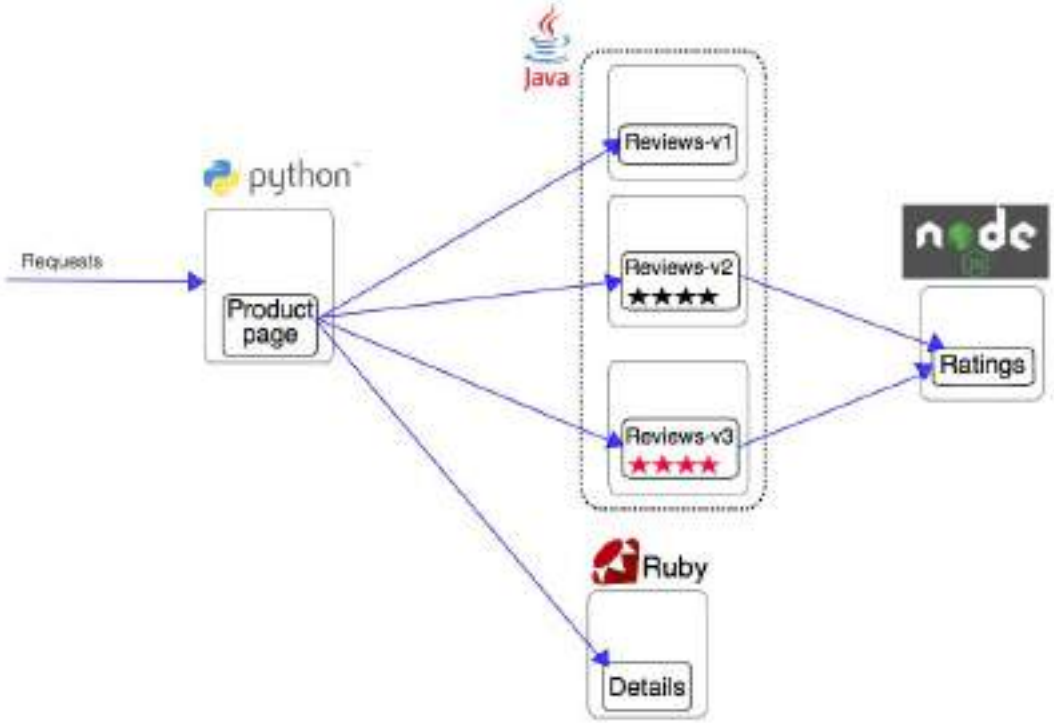
3. **Telemetry Reporting**. Enables services to report logging and monitoring

Istio-Auth

1. Provides each service with a **strong identity**

2. Provides **service-to-service** and **end-user authentication** using mutual TLS

3. Provides a **key management system** to automate key and certificate generation, distribution, rotation, and revocation



Bookinfo Application without Istio

Demo: Deploy App

Add-ons

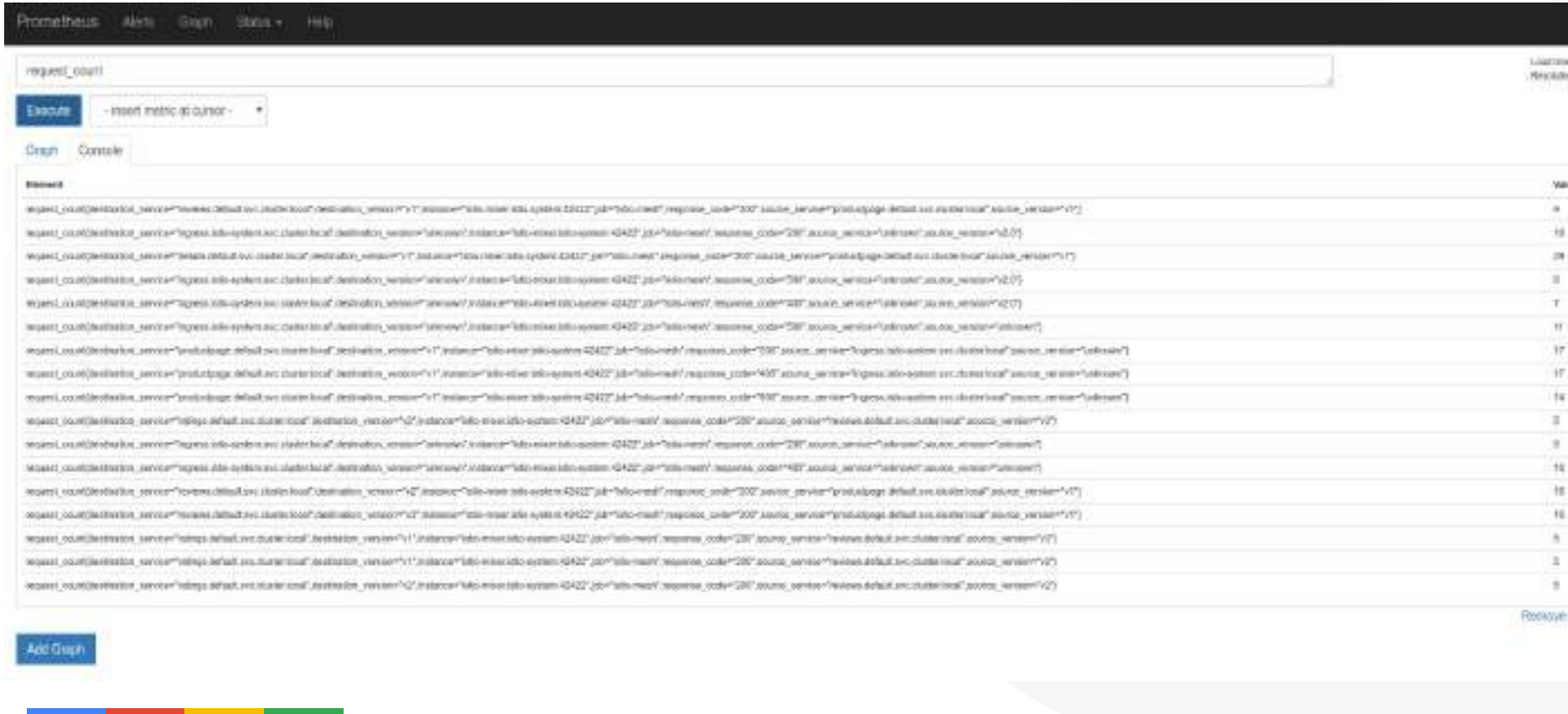
Grafana, Prometheus, Zipkin, ServiceGraph

04

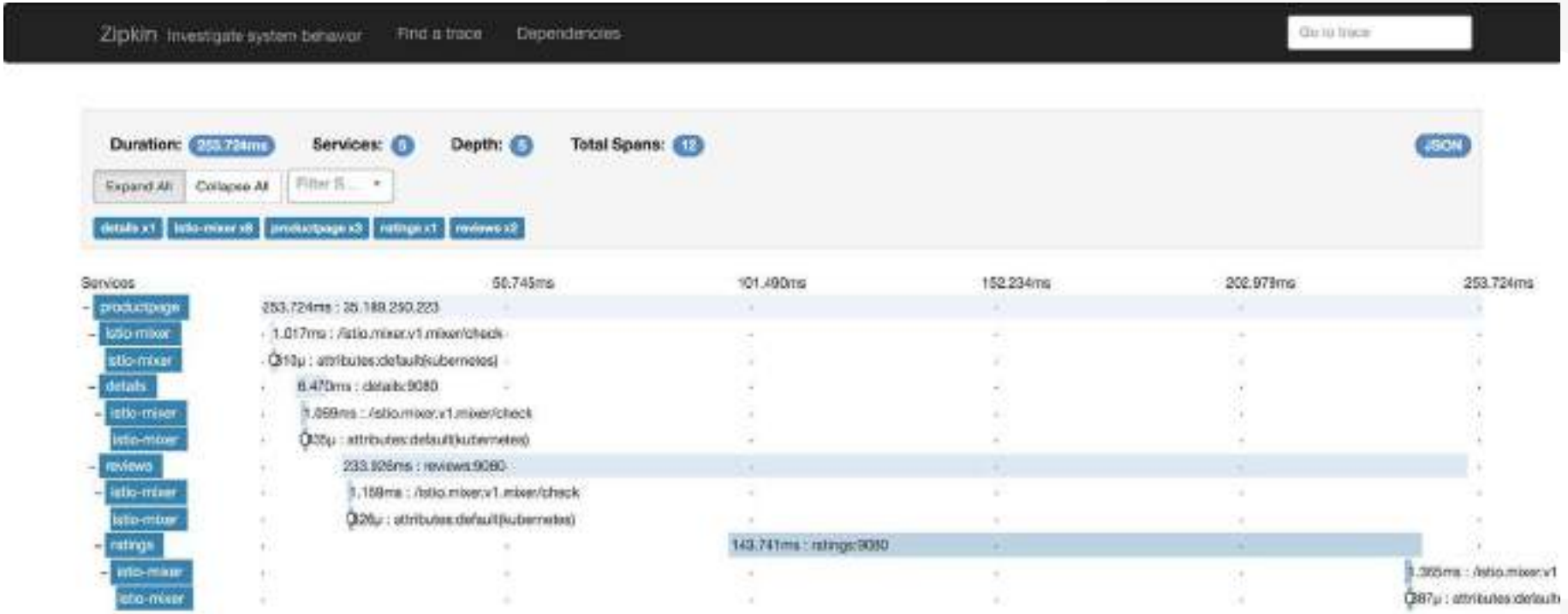
Google Cloud

Grafana: Analytics and monitoring

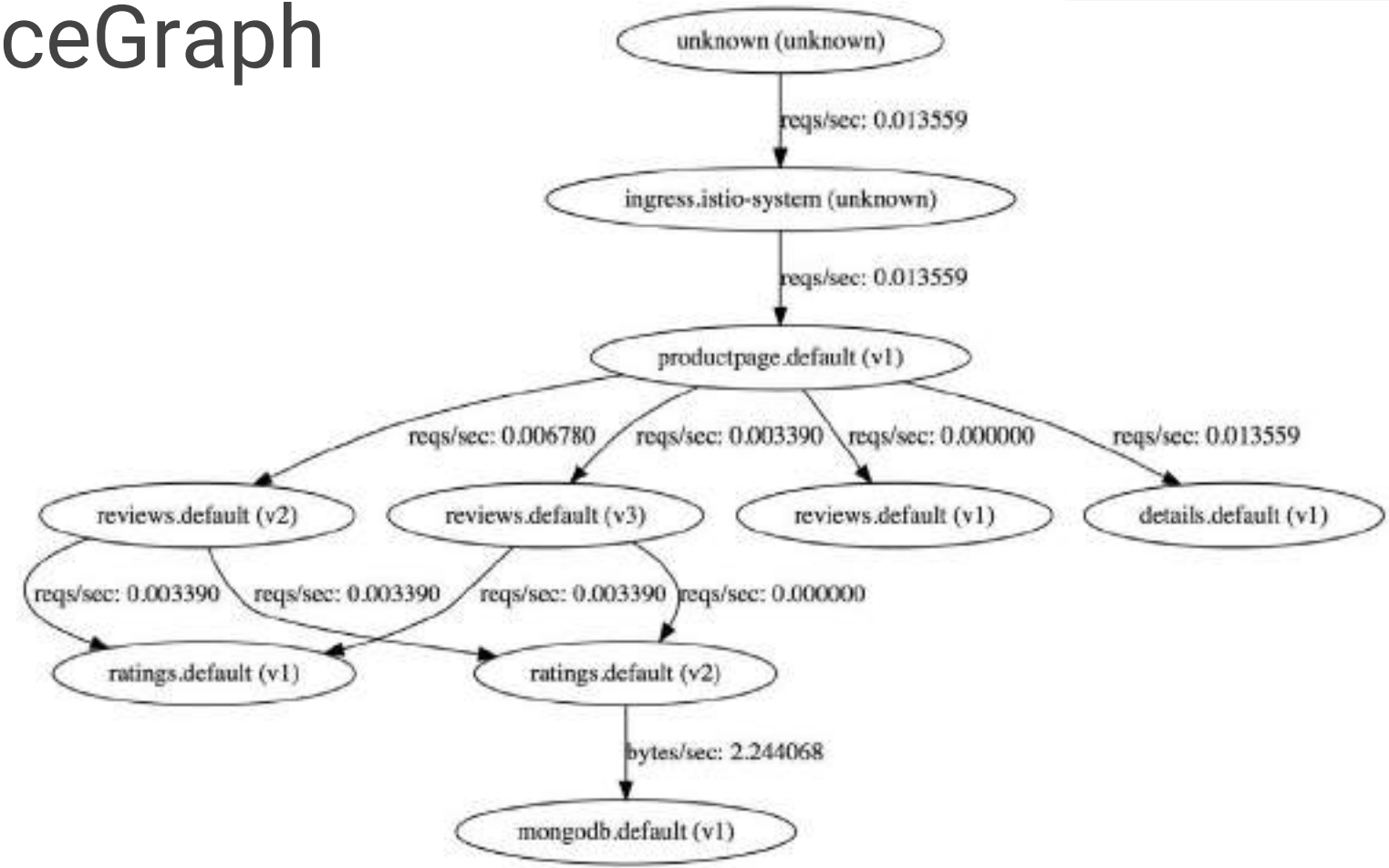
Prometheus: Query metrics



Zipkin: Tracing



ServiceGraph



Demo: Install add-ons

Traffic Management

Request Routing, Discovery & Load Balancing, Failure Recovery & Injection

05

Google Cloud

Traffic Management

Istio’s traffic management decouples traffic flow and infrastructure scaling

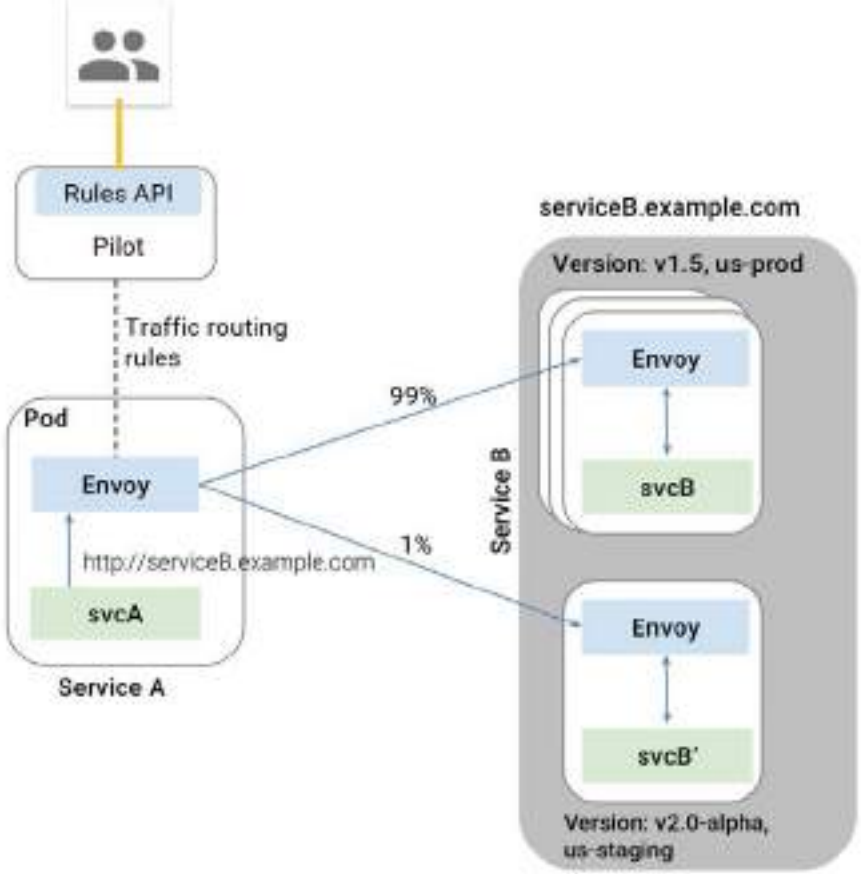
Dynamic request routing for A/B testing, gradual rollouts, canary releases

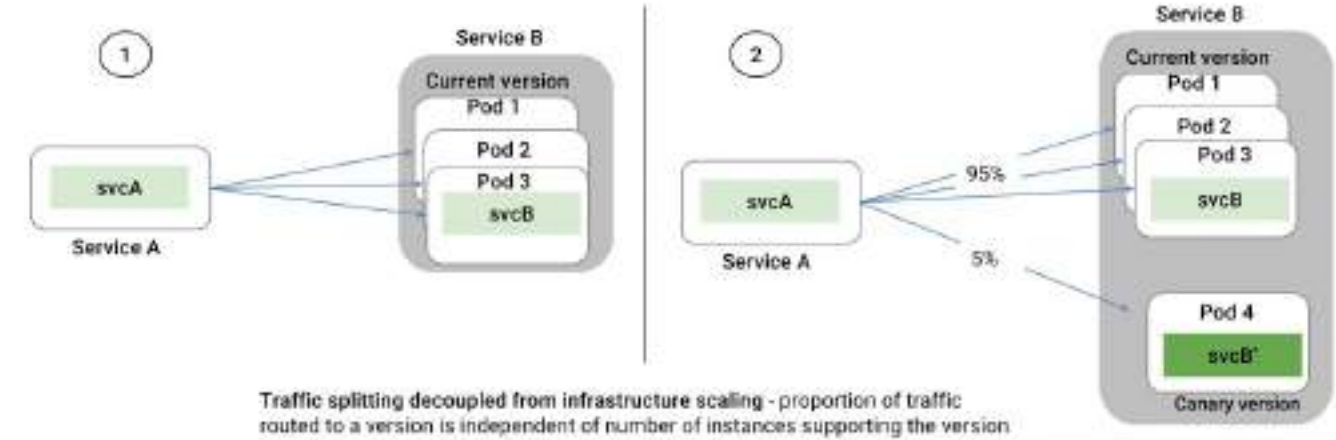
Discovery & load balancing across services

Failure recovery using timeouts, retries, and circuit breakers

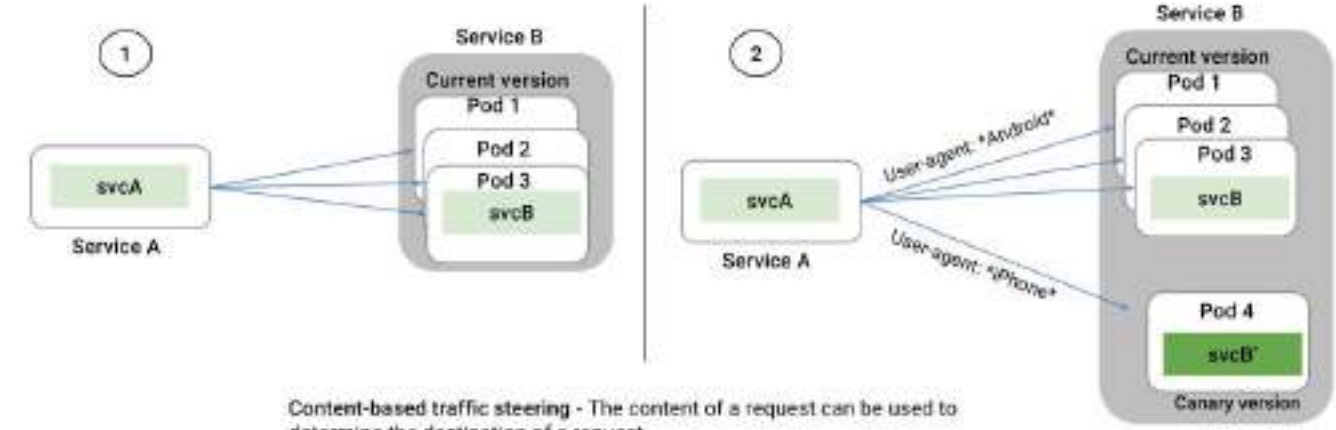
Fault injection to test the compatibility of recovery policies across services

Request Routing



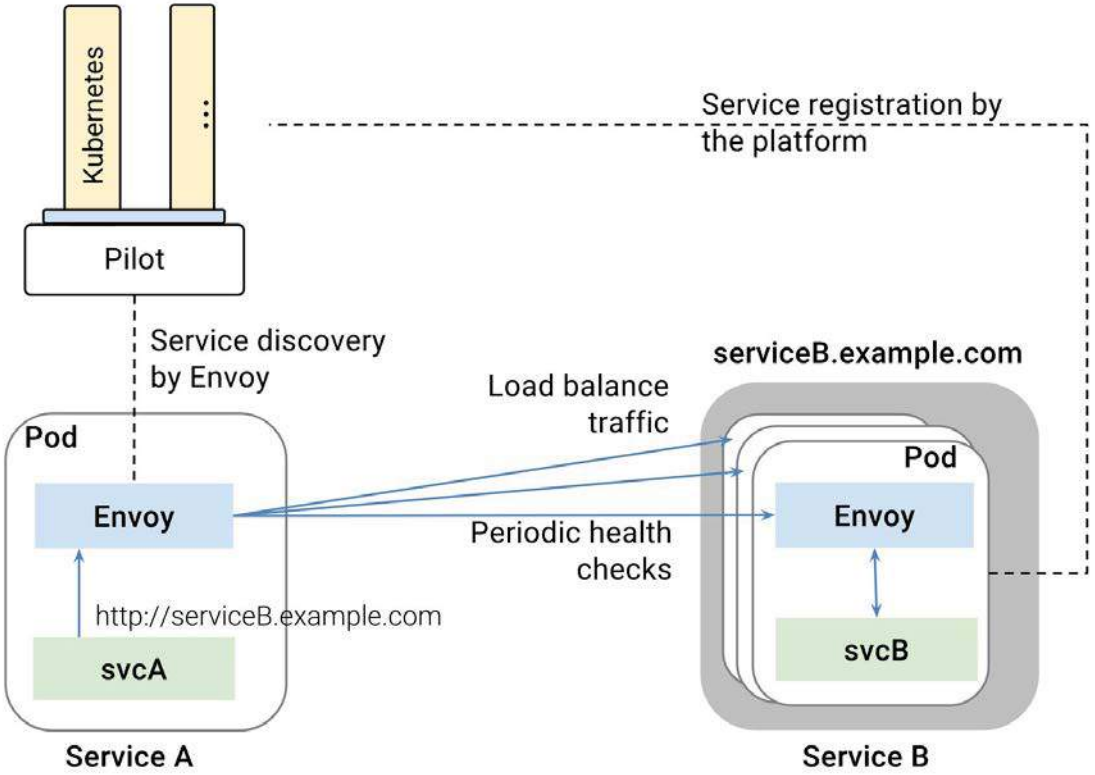


Traffic splitting decoupled from infrastructure scaling - proportion of traffic routed to a version is independent of number of instances supporting the version



Content-based traffic steering - The content of a request can be used to determine the destination of a request

Discovery & Load Balancing



Demo: Change routes

Failure Recovery

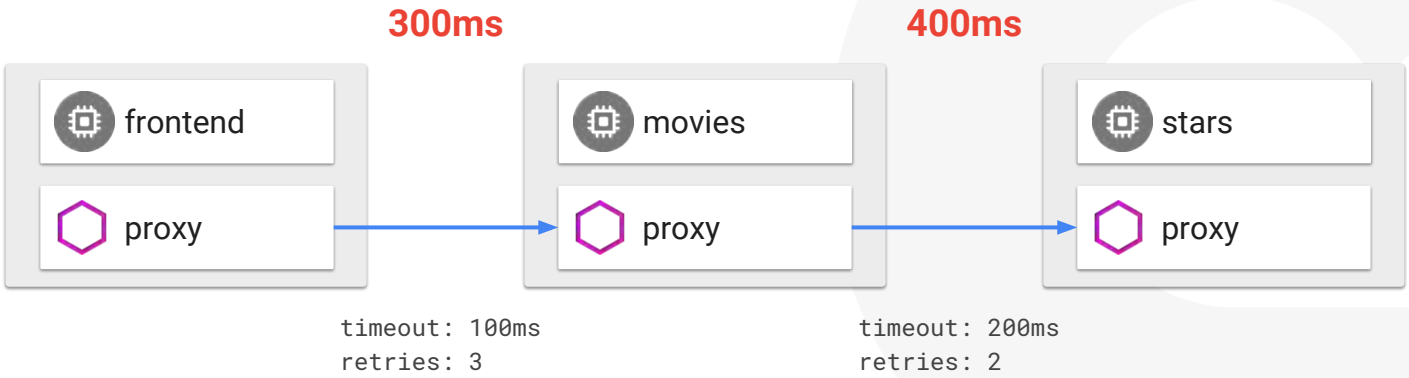
Out-of-the-box opt-in failure recovery features:

- Timeouts
- Bounded retries with timeout budgets and variable jitter between retries
- Limits on number of concurrent connections
- Periodic health checks on each member of the load balancing pool
- Fine-grained circuit breakers (passive health checks) – applied per instance in the load balancing pool

Fault Injection

Systematic fault injection to identify weaknesses in failure recovery policies

- HTTP/gRPC error codes
- Delay injection



Demo: Cleanup

Thank you!

Mete Atamel
@meteatamel

极客邦科技
企业培训与咨询

Geekbang.

扫码关注

获取更多培训信息