
istio.io
service mesh for all
Mandar Jog
Google Cloud

Agenda

- What is a Service Mesh
- How we got here: A story
- Architecture and details
- Q & A

Google Cloud

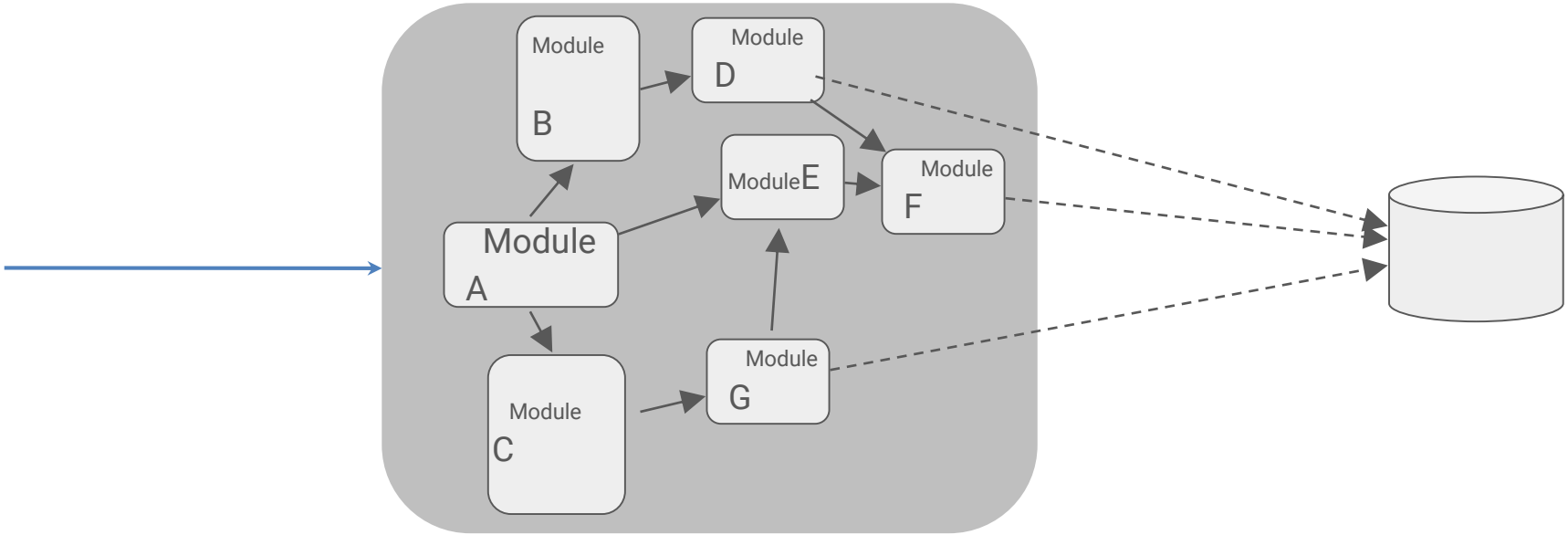
What is a ‘Service Mesh’ ?

A network for services, not bytes

- Traffic Control
- Visibility
- Resiliency & Efficiency
- Security
- Policy Enforcement

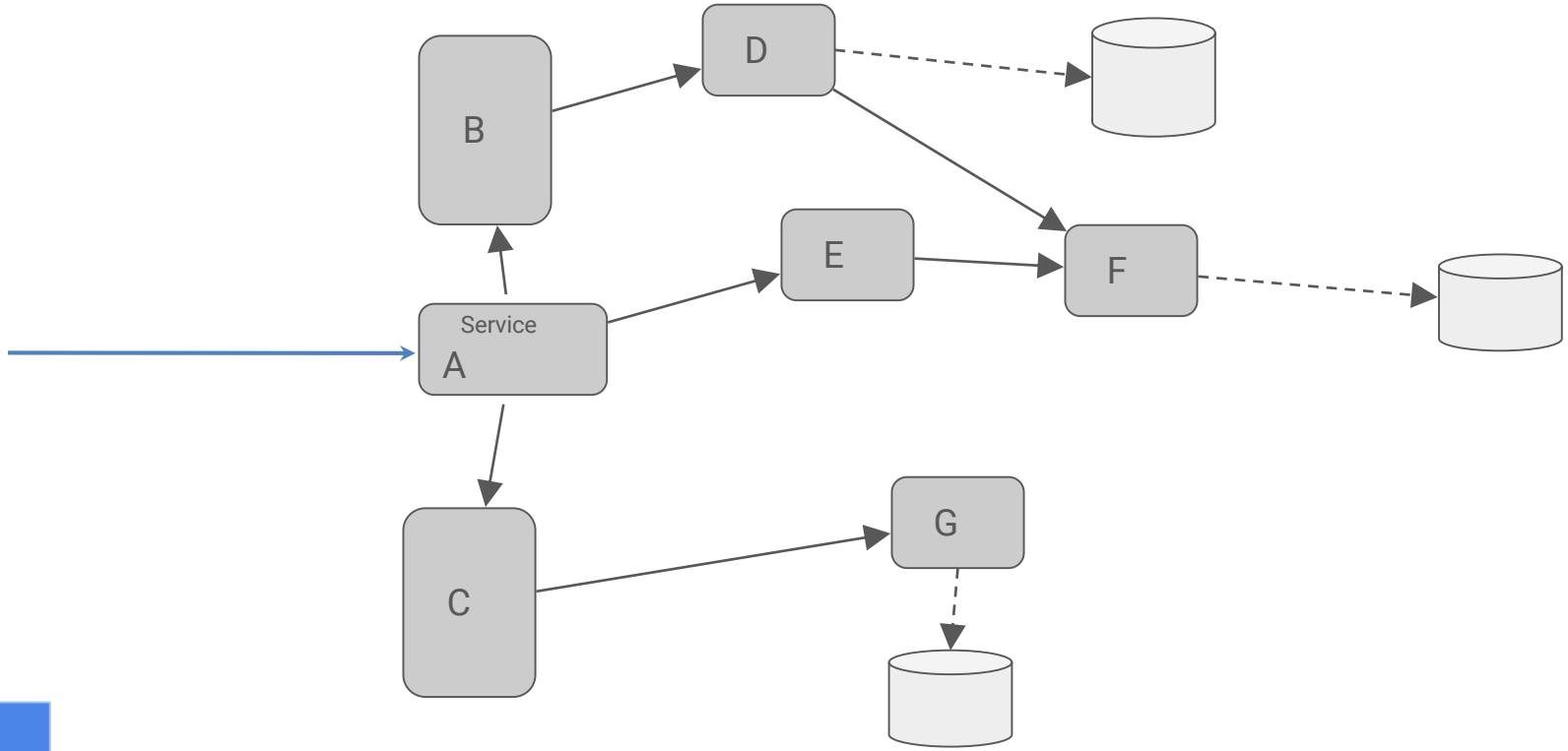
Google Cloud

Remember the Monolith?



Google Cloud

Micro Services



Google Cloud

Micro Services FTW

1. Gained development velocity !
2. Easy testing because of abstractions.
3. Scale services independently.

Problem Solved!

Google Cloud

Yay !

1. Gained development velocity !
2. Easy testing because of abstractions.
3. Scale services independently.

Problem Solved!

Google Cloud

What have we lost?

1. I replaced a **reliable** in-process call with an **unreliable rpc**.
2. Trivial single stepping replaced by … ?
3. Secure in-process communication is replaced by insecure network.
4. Access control within process was a NOOP
5. Latency went up

That abstraction was leaky ...

Google Cloud

Can we fix it?

1. Add retry logic to the application code.
2. Add entry-exit traces.
3. Secure inter service connections with strong authentication.

Now that we are adding code … choose the rpc endpoint intelligently

   a. Endpoints with low latency.
   b. Endpoints with warm caches.

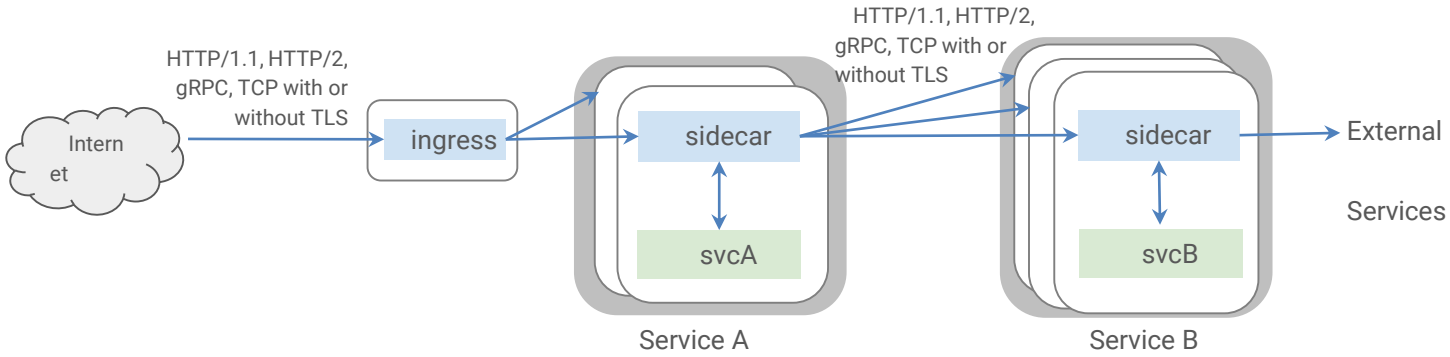
Google Cloud

Service Mesh

- Address service level concerns.
- Unlock the full power of microservices.

Google Cloud

Weaving the mesh



envoy by lyft

Outbound features:

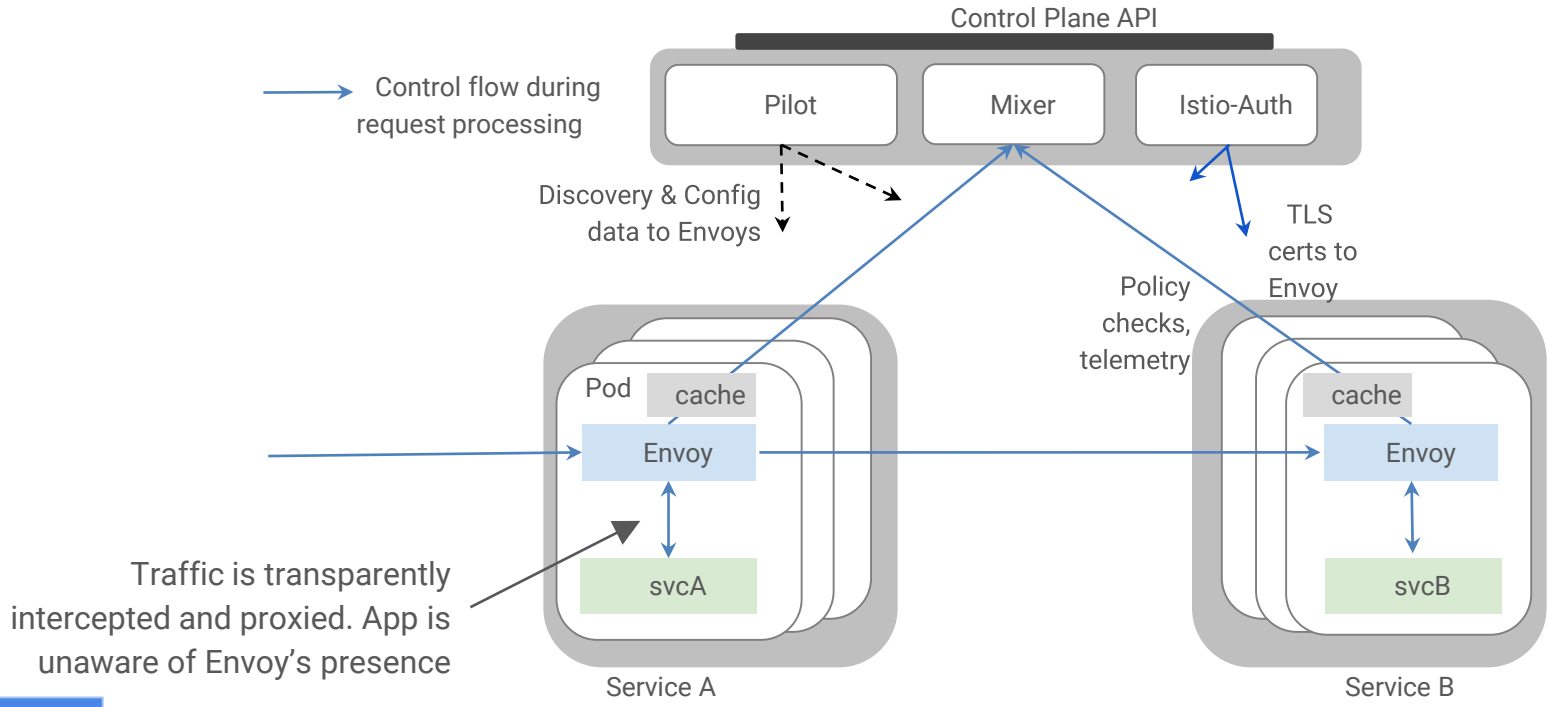
- Service authentication
- Load balancing
- Retry and circuit breaker
- Fine-grained routing
- Telemetry
- Request Tracing
- Fault Injection

Inbound features:

- Service authentication
- Authorization
- Rate limits
- Load shedding
- Telemetry
- Request Tracing
- Fault Injection

Google Cloud

Architecture



Google Cloud

istio Service Mesh

A network for services, not bytes

- **Traffic Control**
- Visibility
- Resiliency & Efficiency
- Security
- Policy Enforcement

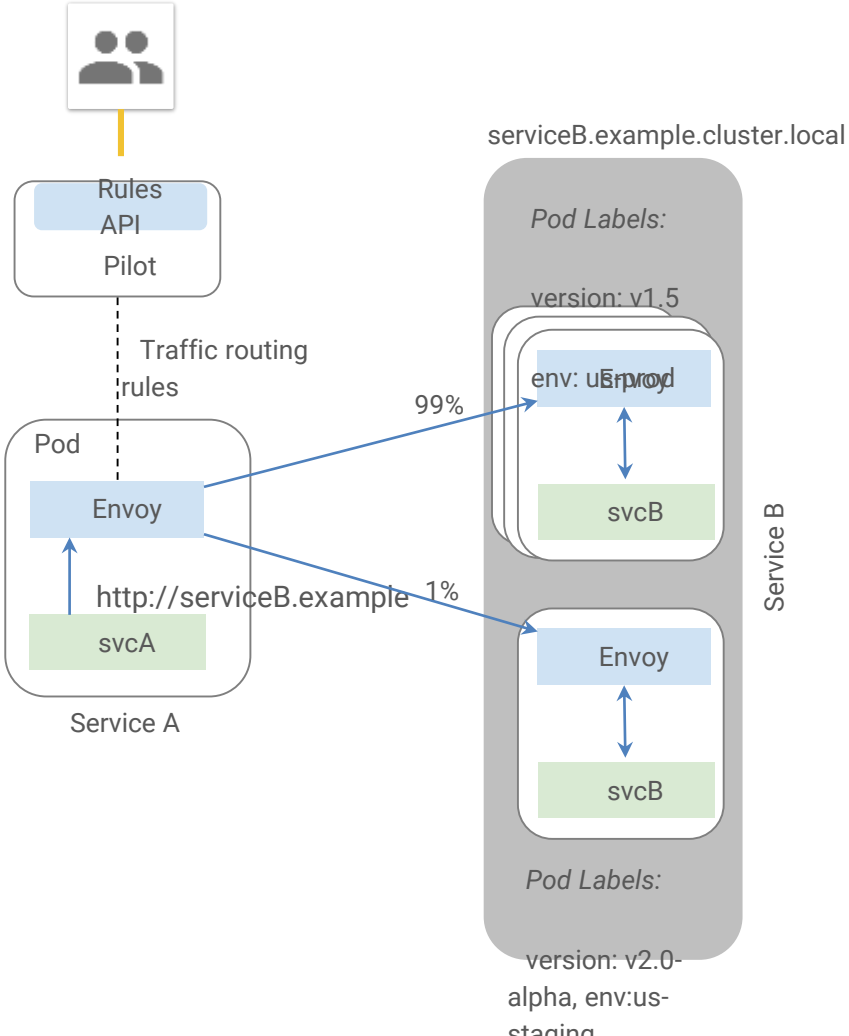
Google Cloud

Application Rollout

```
// A simple traffic splitting rule

destination: serviceB.example.cluster.local
match:
  source: serviceA.example.cluster.local
route:
- tags:
    version: v1.5
    env: us-prod
  weight: 99
- tags:
    version: v2.0-alpha
    env: us-staging
  weight: 1
```

Traffic control is decoupled from infrastructure scaling



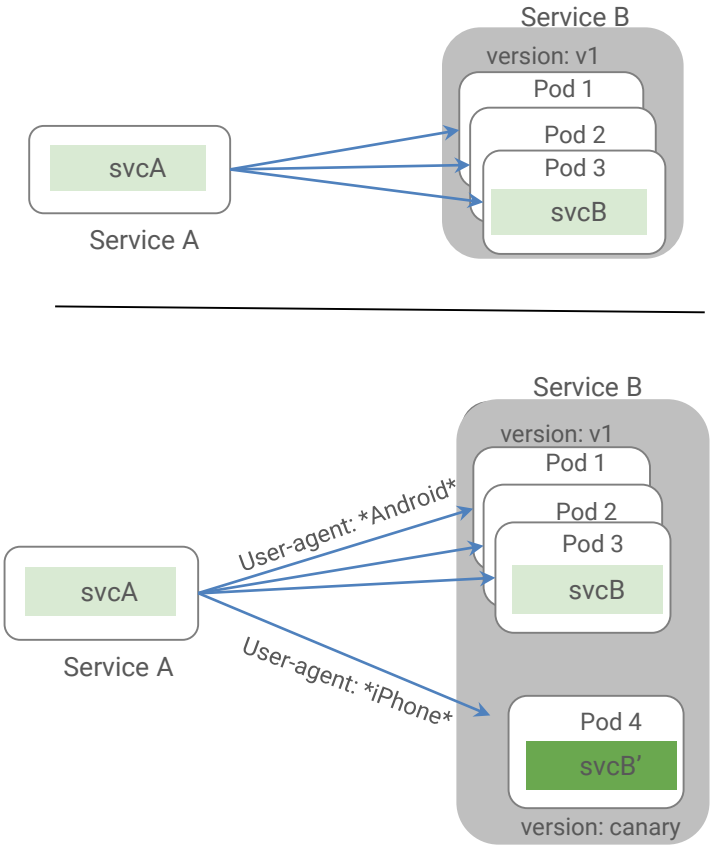
Google Cloud

Traffic Steering

```
// Content-based traffic steering rule

destination: serviceB.example.cluster.local
match:
  httpHeaders:
    user-agent:
      regex: ^(.*?;)?(iPhone)(;.*)?$
precedence: 2
route:
- tags:
    version: canary
```

Content-based traffic steering



Google Cloud

Istio Service Mesh

A network for services, not bytes

- Traffic Control
- **Visibility**
- Resiliency & Efficiency
- Security
- Policy Enforcement

Google Cloud

Visibility

Monitoring & tracing should not be an afterthought in the infrastructure

Goals

- Metrics without instrumenting apps
- Consistent metrics across fleet
- Trace flow of requests across services
- Portable across metric backend providers

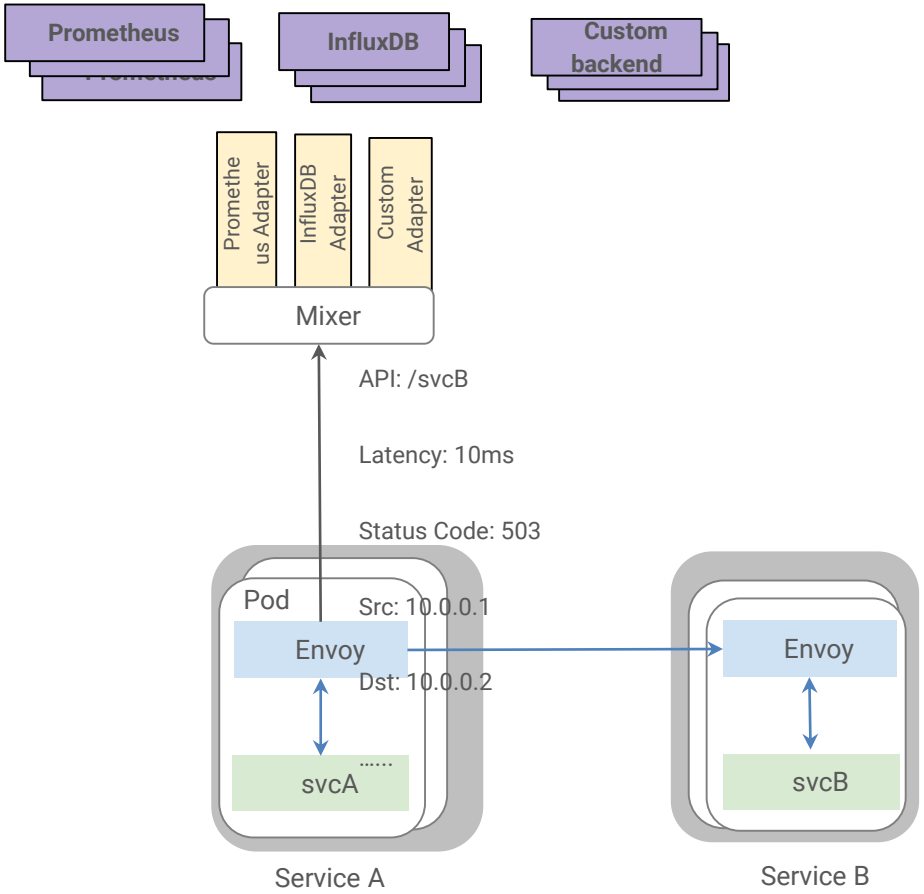
Istio - Grafana dashboard w/ Prometheus backend

Istio Zipkin tracing dashboard

Google Cloud

Metrics flow

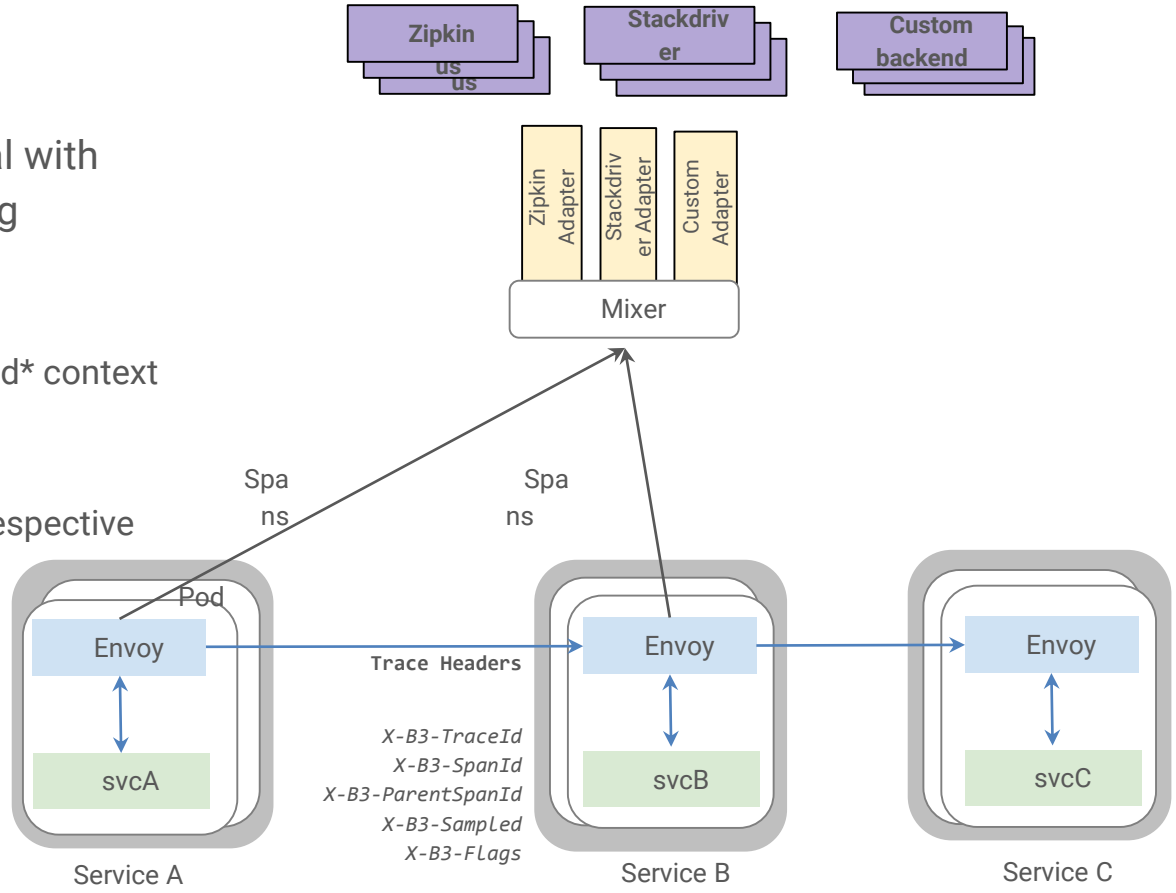
- Mixer collects metrics emitted by Envoys
- Adapters in the Mixer normalize and forward to monitoring backends
- Metrics backend can be swapped at runtime



Google Cloud

Visibility: Tracing

- Application do not have to deal with generating spans or correlating causality
- Envoys generate spans
  - Applications need to *forward* context headers on outbound calls
- Envoys send traces to Mixer
- Adapters at Mixer send traces to respective backends



Google Cloud

Istio Service Mesh

A network for services, not bytes

- Traffic Control
- Visibility
- **Resiliency & Efficiency**
- Security
- Control

Google Cloud

Resiliency

Istio adds fault tolerance to your application without any changes to code

```
 // Circuit breakers

 destination: serviceB.example.cluster.local
policy:
- tags:
    version: v1
  circuitBreaker:
    simpleCb:
      httpConsecutiveErrors: 7
      sleepWindow: 5m
      httpDetectionInterval: 1m
```

Resilience features

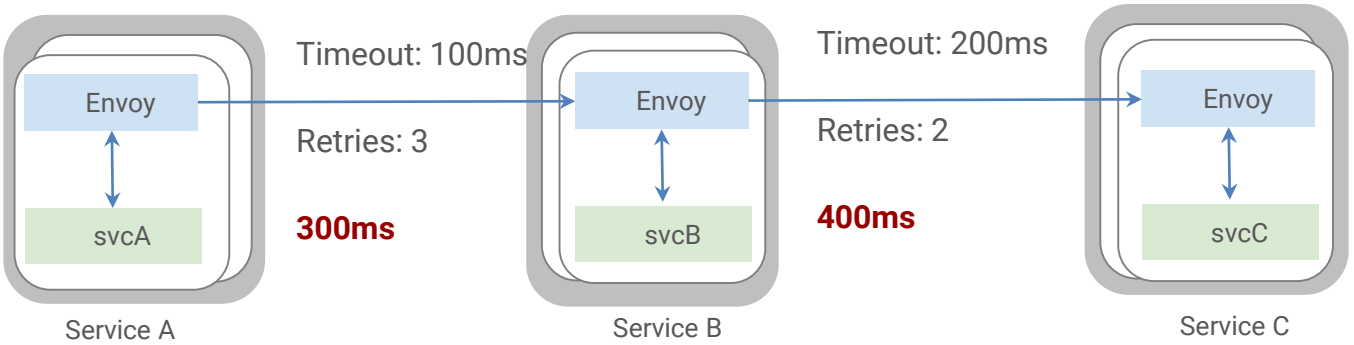
- Timeouts
- Retries with timeout budget
- Circuit breakers
- Health checks
- AZ-aware load balancing w/ automatic failover
- Control connection pool size and request load
- Systematic fault injection

Google Cloud

Resiliency Testing

Systematic fault injection to identify weaknesses in failure recovery policies

- HTTP/gRPC error codes
- Delay injection



Google Cloud

Efficiency

- L7 load balancing
  - Passive/Active health checks, circuit breaks
  - Backend subsets
  - Affinity
- TLS offload
  - No more JSSE or stale SSL versions.
- HTTP/2 and gRPC proxying

Google Cloud

Istio Service Mesh

A network for services, not bytes

- Traffic Control
- Visibility
- Resiliency & Efficiency
- **Security**
- Policy Enforcement

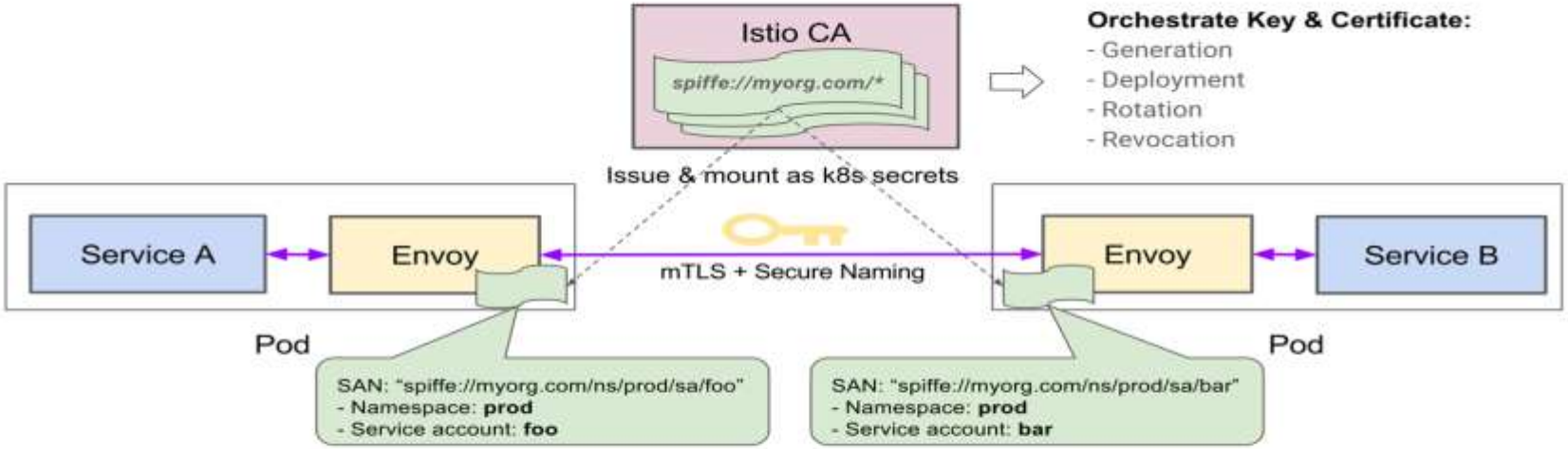
Google Cloud

Securing Microservices

- Verifiable identity
- Secure naming / addressing
- Traffic encryption
- Revocation

Google Cloud

Istio - Security at Scale



spiffe.io

Google Cloud

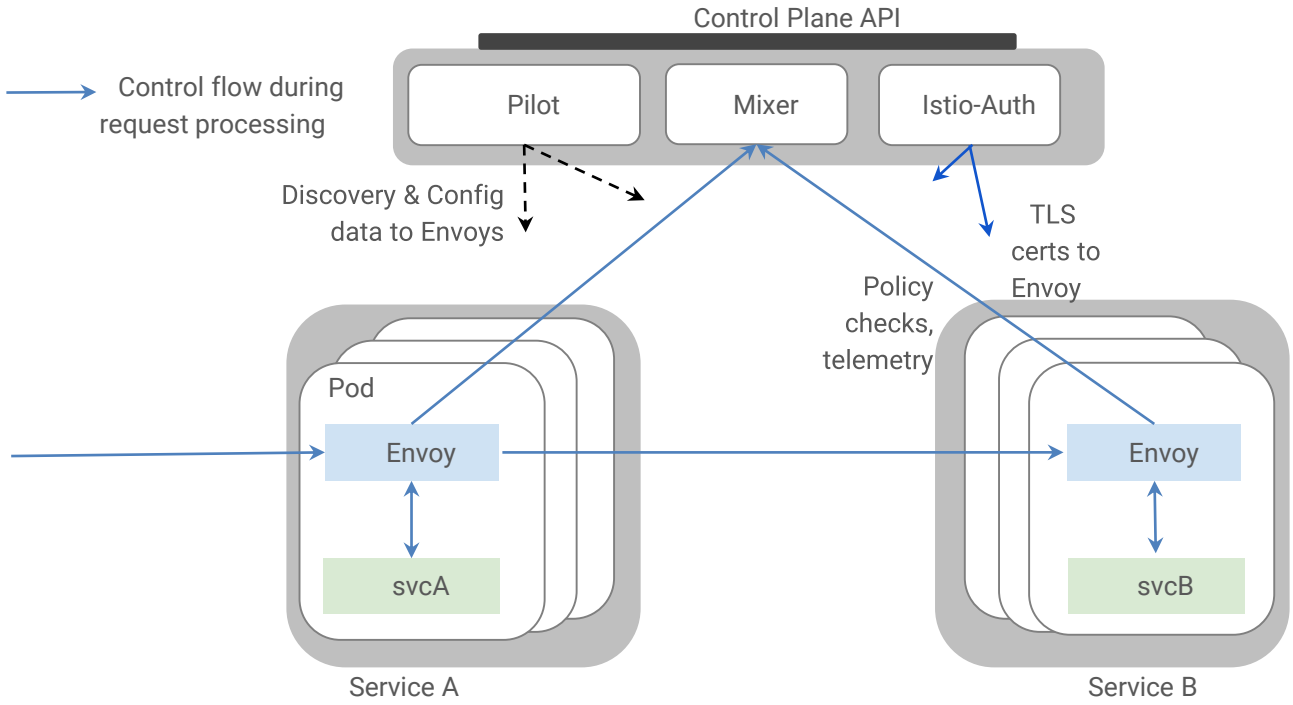
Istio Service Mesh

A network for services, not bytes

- Traffic Control
- Visibility
- Resiliency & Efficiency
- Security
- **Policy Enforcement**

Google Cloud

Putting it all together



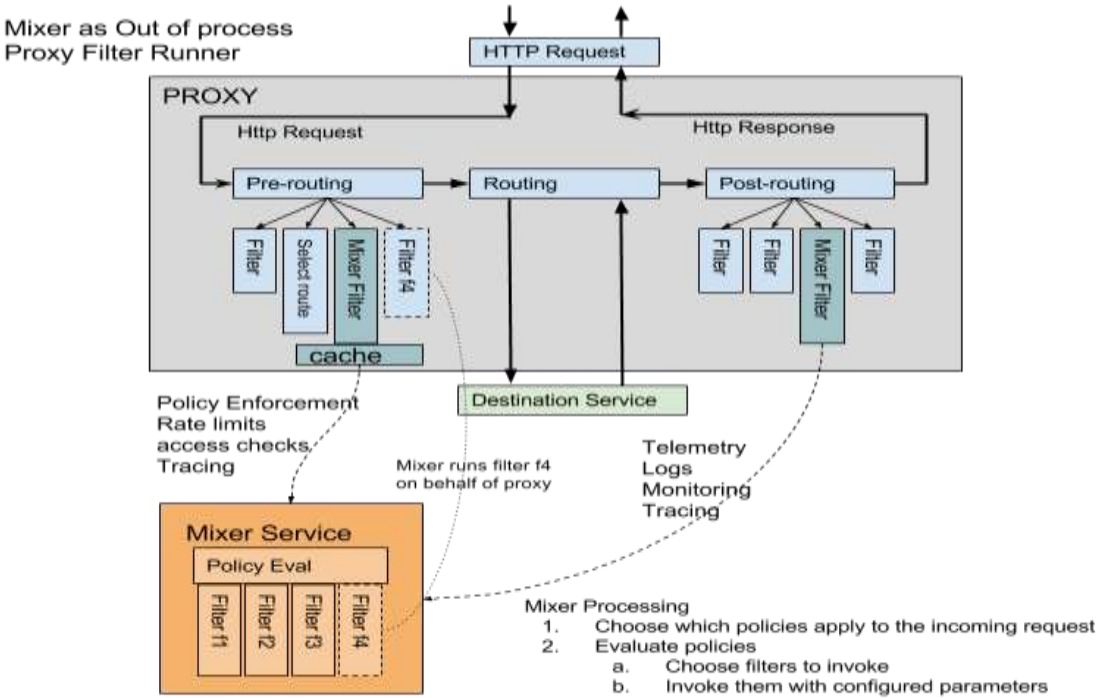
Google Cloud

What does Mixer do?

- Check()
  - Precondition checking
  - Quotas & Rate Limiting
- Report()
  - telemetry reporting
- Primary point of extensibility
- Enabler for platform mobility
- Operator-focused configuration model

Google Cloud

Mixer - Proxy Filter Extension



Google Cloud

Mixer Metrics

```
kind: metrics
metadata:
  name: requestcount
spec:
  value: “1”
  dimensions:
    Destination_service: destination.service
    Source_service: source.service
    Response_code: response.code
```

```
kind: rule
metadata:
  name: prometheus
spec:
  match: request.headers[“x-user”] != “admin”
  actions:
  - handler: handler.prometheus
    instances:
    - requestcount.metric
```

Google Cloud


istio.io
@mandarjog
Google Cloud