Building a Multi Cluster/Env Service Mesh at Airbnb

Weibo He & Stephen Chan, 10/13/2021, KubeCon + CloudNativeCon 2021

Agenda

- Introduction
- Support Multi Cluster
- Support Multi Environment
  - Multi Tier
  - Mesh Expansion
  - External Services
- Takeaways

Introduction

Airbnb's Istio Journey

2019

- Search of a modern Service Mesh started
- Evaluated and landed on Istio

2020

- Productionize Istio
- Feature Parity

2021

- Feature Parity cont.
- Migrate production workloads onto Istio

https://bit.ly/3z6i0gw

Multi Cluster

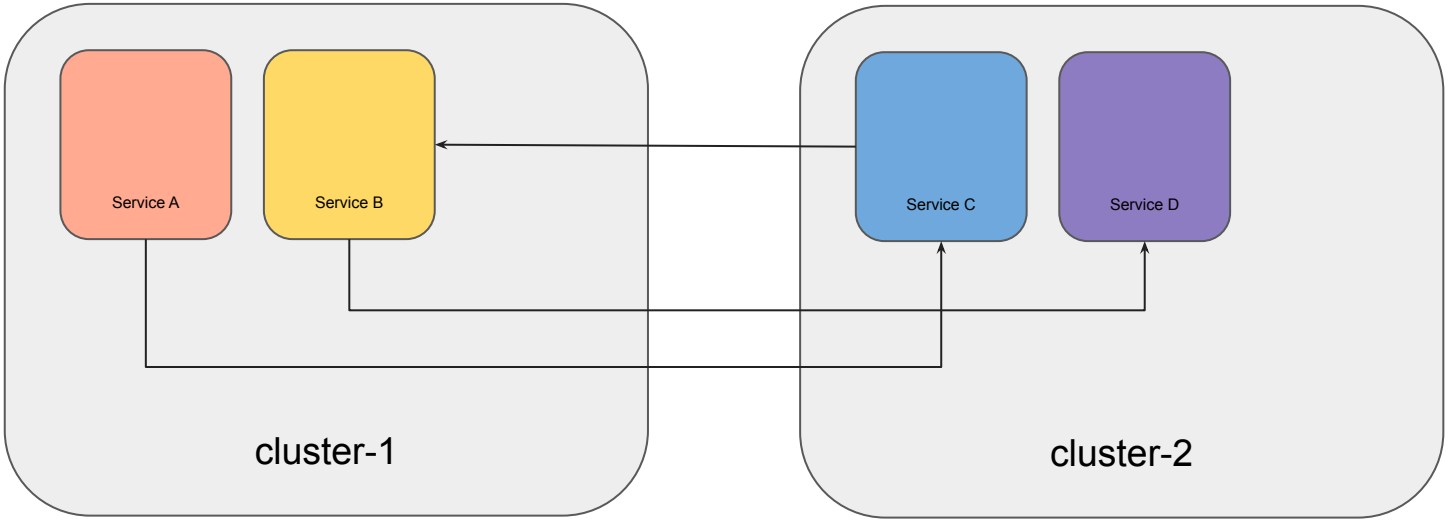
Multi Cluster

- Horizontally scale # of clusters instead of vertically scaling # of nodes in each cluster
- Keep each k8s cluster under 1k nodes and scale out by adding more clusters

https://bit.ly/38XJUkx

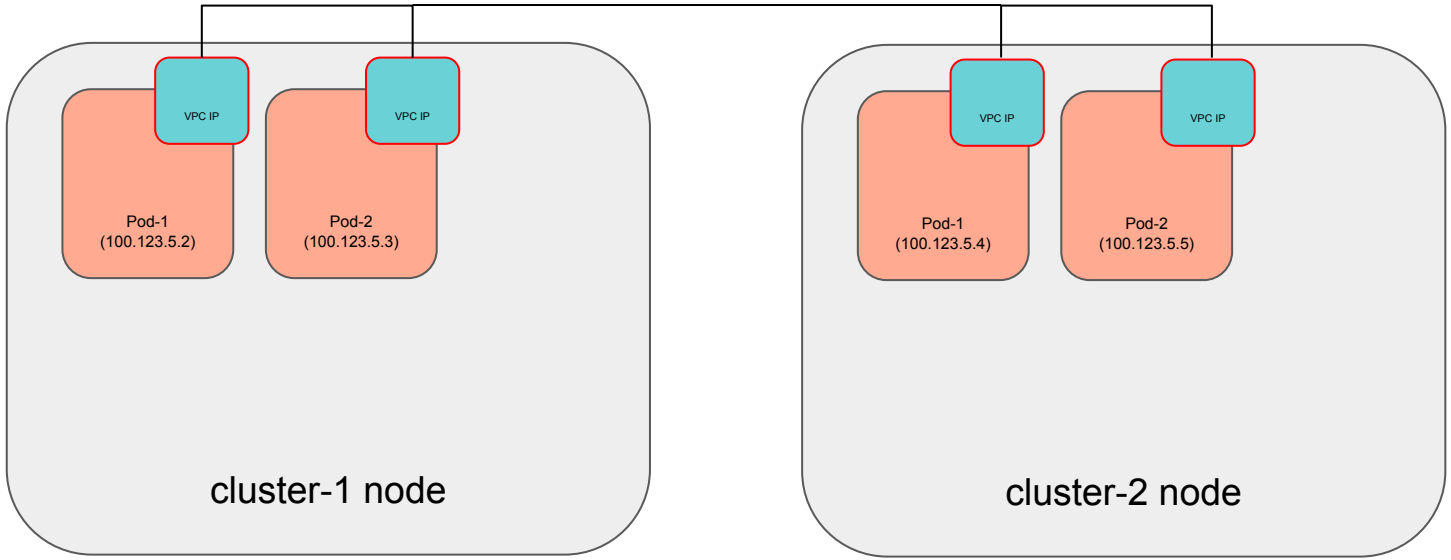
Workload Placement

- Clusters are just pools of compute and memory
- Workloads are randomly assigned to clusters



Flat Network

- AWS VPC CNI assign individually addressable VPC IPs to pods
- Direct pod to pod communication cross clusters



Flat Network Requirements

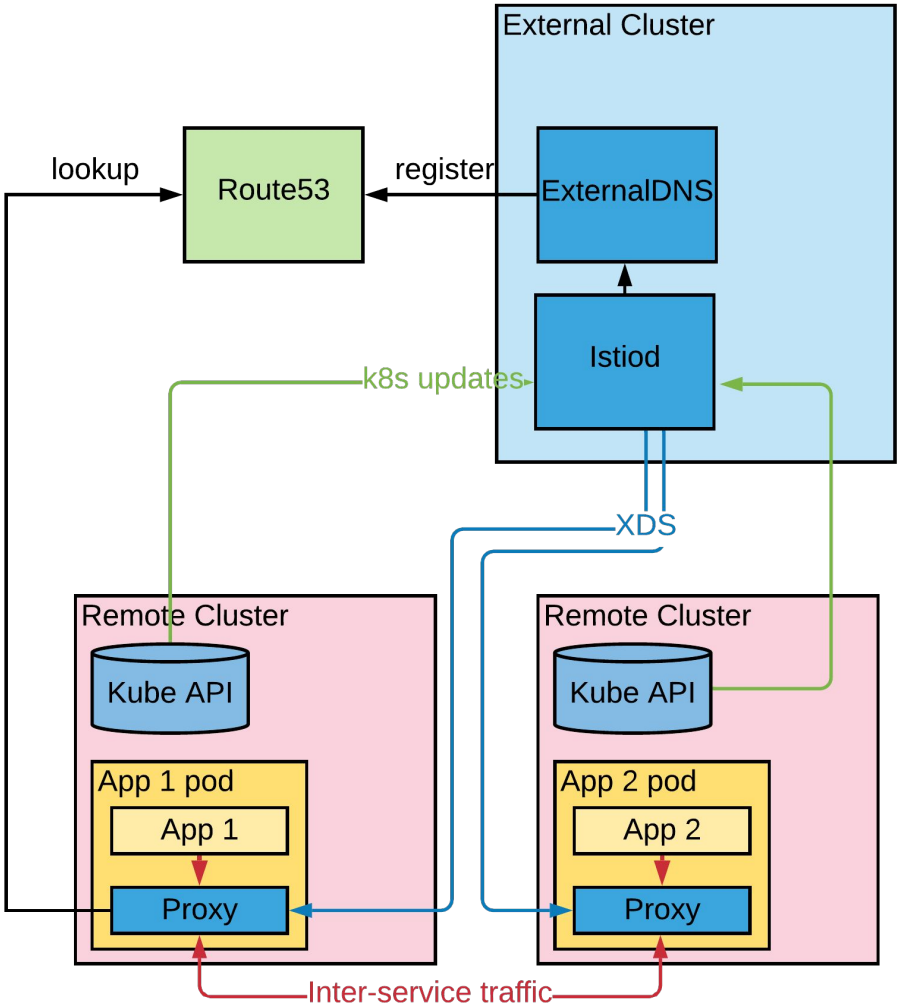
- Not relying on network boundary for security
  - Security Groups & mTLS
- Non-overlapping IP address space.
  - Centrally managed non-overlapping private IPv4 CIDRs
- Data plane scalability (VPC IP mapping limit)
  - With VPC CNI
    - 1 Node with 16 pods consumes **17** IP mappings (node IP + 16 pod IPs)
  - With VPC CNI + Prefix delegation
    - 1 Node with 16 pods consumes **2** IP mappings (node IP + 1 /28 prefix)

External Control Plane

- Clean separation of roles
  - Mesh Operator/Admin VS Mesh Users
- Tightened security
  - CA certs only installed on control plane cluster.
  - Tight access control for control plane cluster.
- Isolation from data plane workloads.
  - Bad deploys & outages does not affect control plane.
- Ease of operation.
  - Manage 1 deployment instead of N deployments.

Architecture

Single Mesh



Multi Environment

Multi Tier

Service Tiers

- Test
  - Unstable environment for testing
- Staging
  - Stable pre-prod environment
- Production
  - Powers Airbnb.com

Multi-Tiered Mesh & Release

Sandbox Mesh

Single cluster deployment

- Run functional tests
- Run performance tests

Test Mesh

Multi Cluster Setup

- Test integration with Airbnb systems
- Integration test workloads

Prod Mesh

Identical Setup as Test Mesh

- New Istio version tested on staging environment for 2 weeks
- Gradual release to production environment

Standalone Mesh

- Functional
  - Authorization
  - Locality based load balancing
  - Regression tests
- Performance
  - Validate latency & resource consumption of istio-proxy with Nighthawk

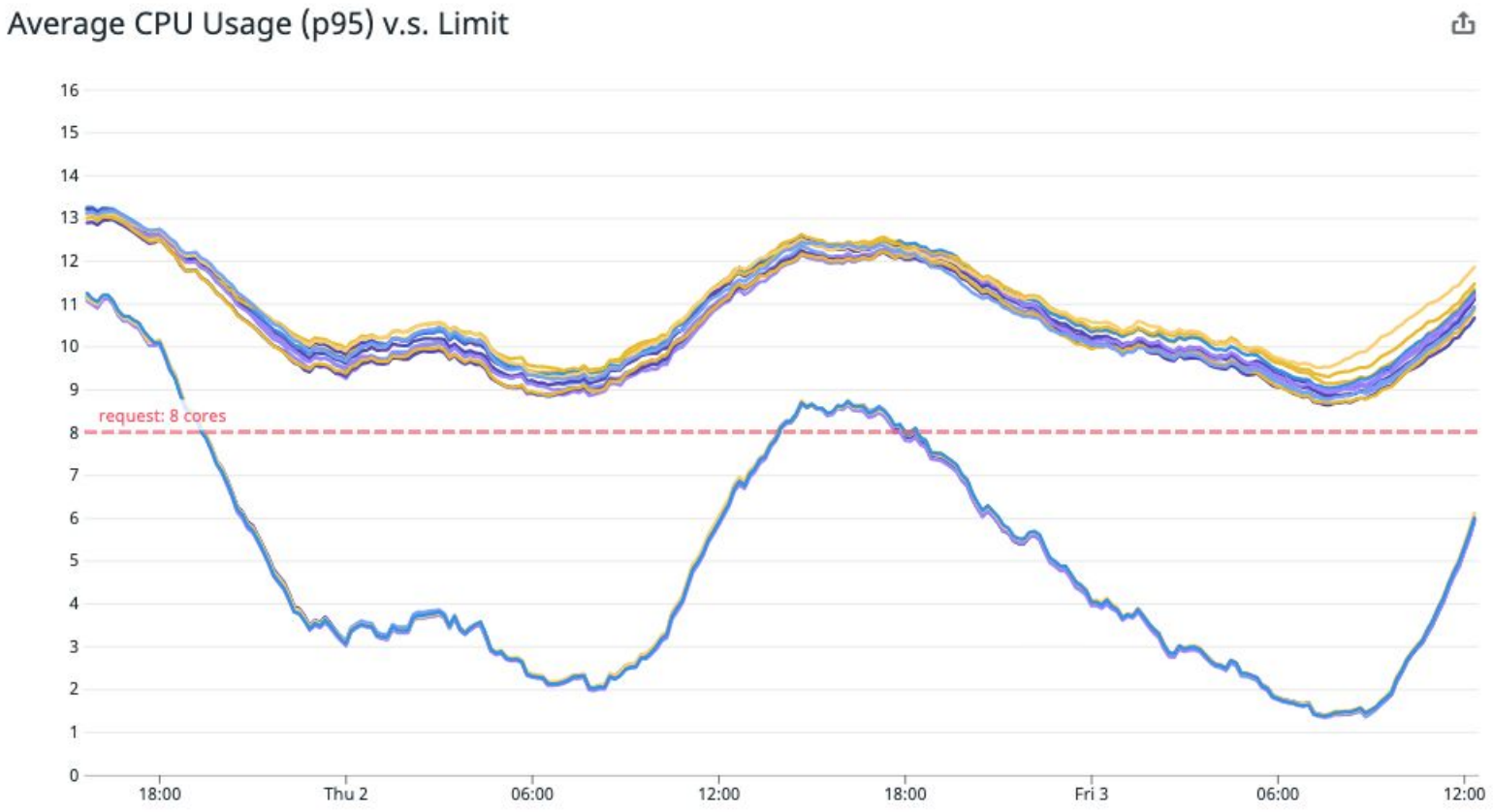
Test Mesh

- Integration with Airbnb systems
  - InternalCA
  - K8s clusters
  - Istio resources generation
  - Custom extension like Mesh expansion & External Services

Production Mesh

- Deploy new version of Istio with revision label
- Services pick up new istio-proxy & new control plane after a deploy
- Increasing scope of services connecting to the new version
  - Verification Fleet
  - Staging services
  - 25%, 50%, 100% of production services
- Clean up the old version

Production Mesh (Old vs New)

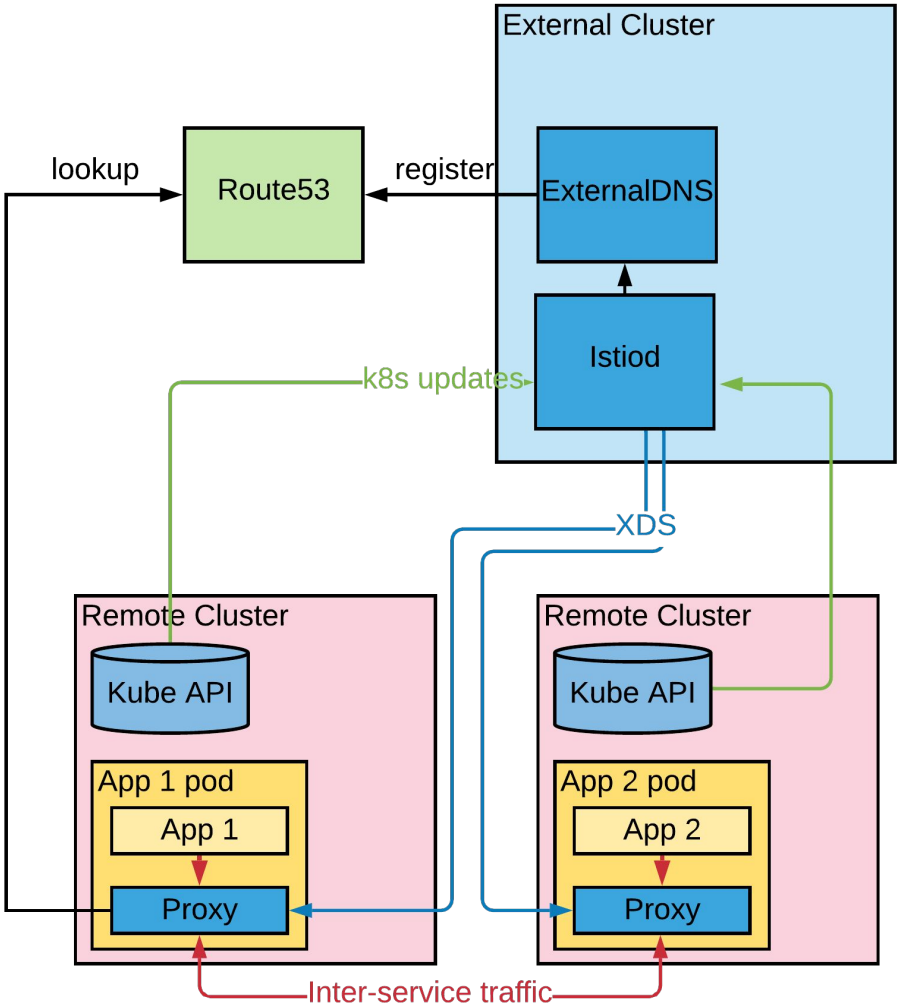


Multi Environment

Mesh Expansion

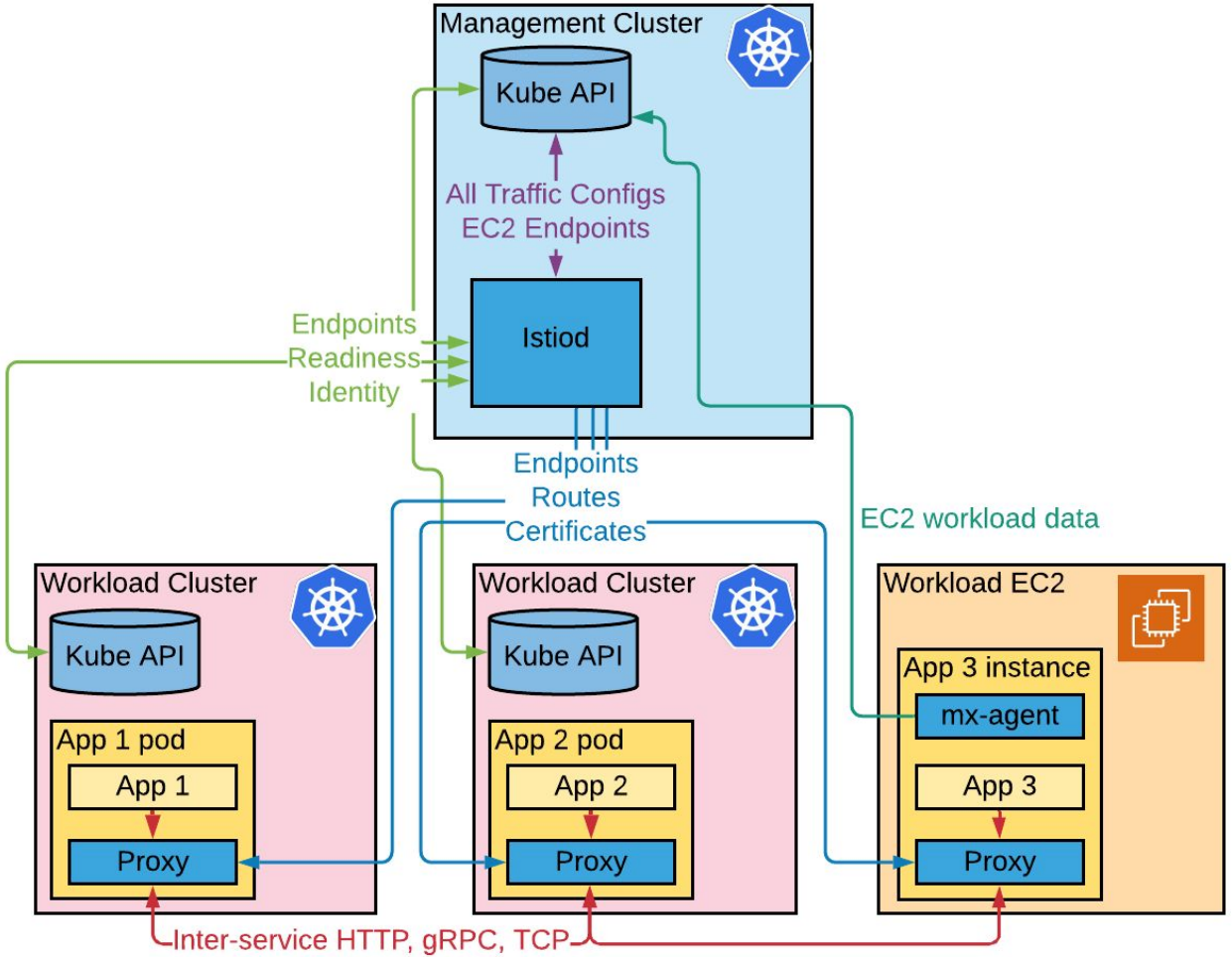
Architecture

Single Mesh



Architecture

Single Mesh



EC2 Support

Requirements

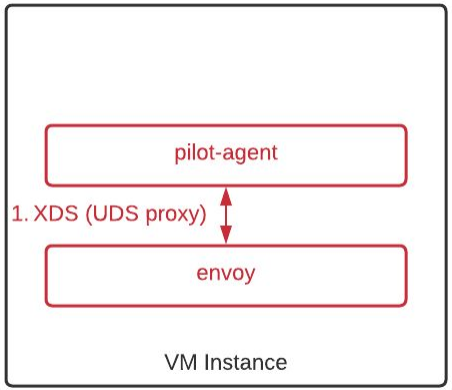
- Feature parity with Kubernetes
  - Automated endpoint registration
  - Server-side health checks
  - TLS
  - Transparent proxy injection
  - Gradual, automated Istio version upgrades
- Allow future migration to Kubernetes

EC2 Support

Endpoint Registration

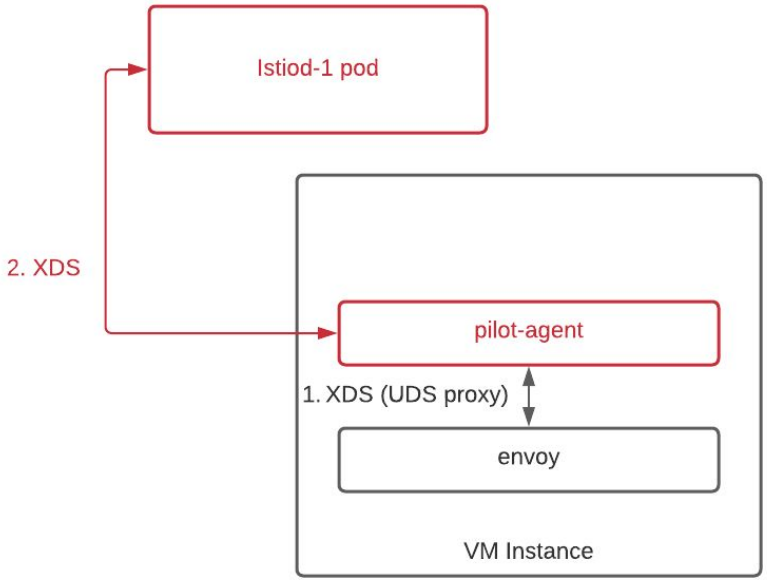
EC2 Support

Endpoint Registration



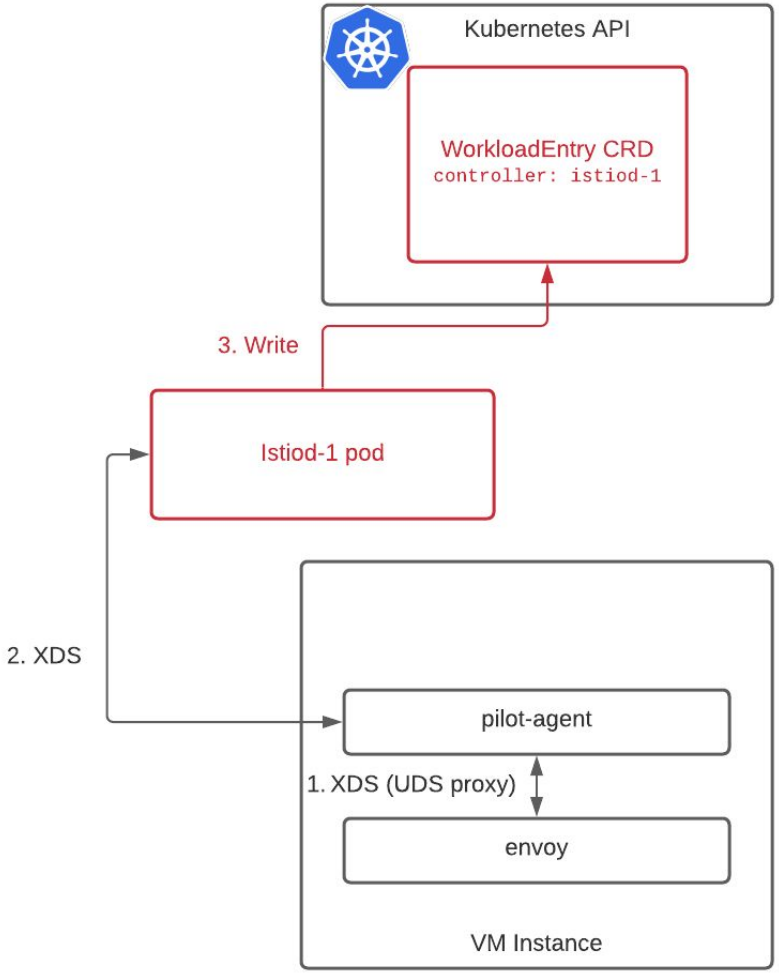
EC2 Support

Endpoint Registration



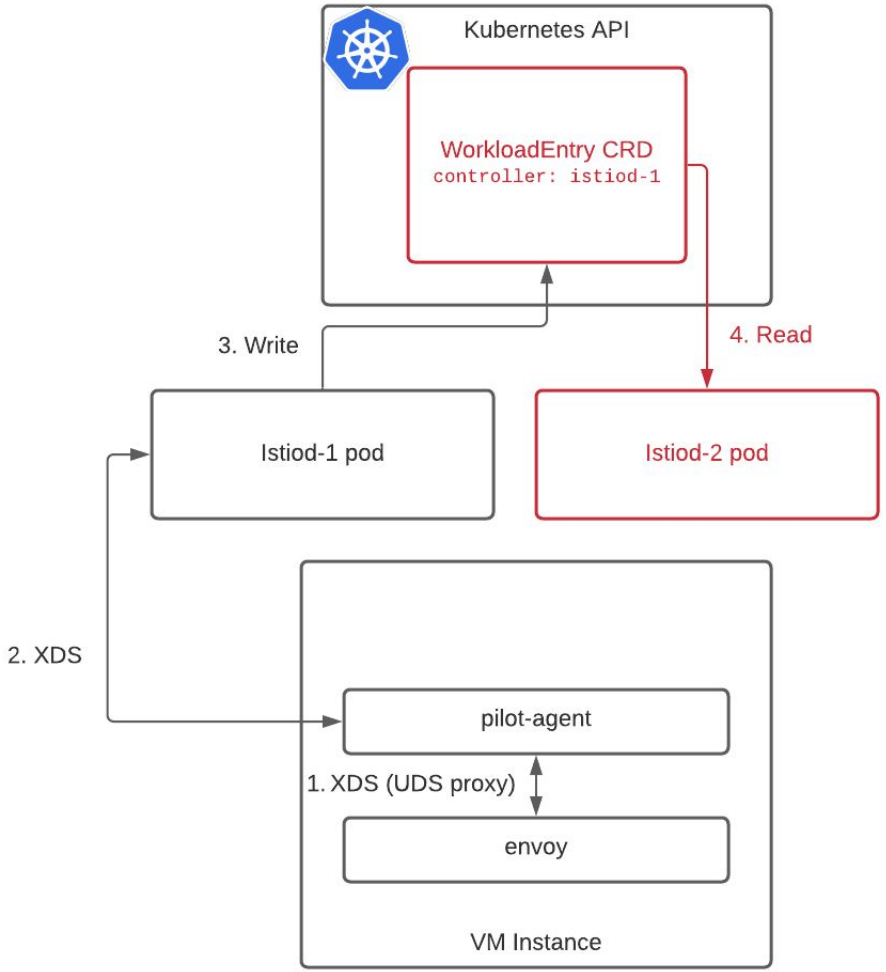
EC2 Support

Endpoint Registration



EC2 Support

Endpoint Registration

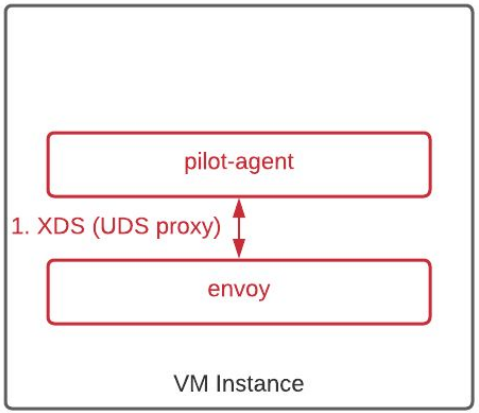


EC2 Support

Server-side Health Checks

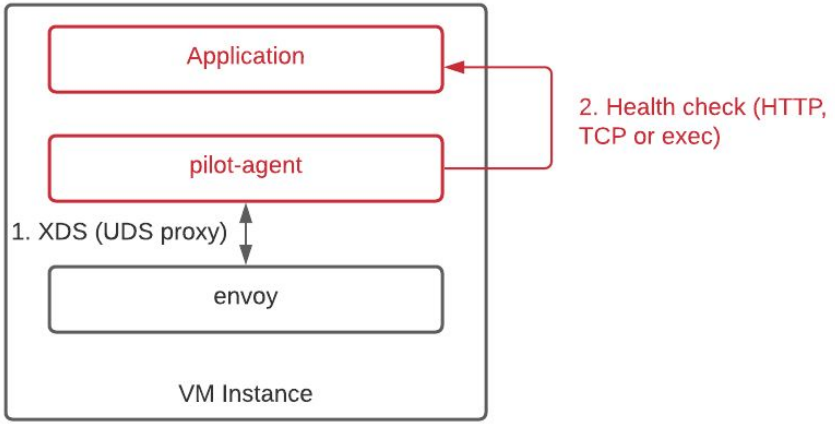
EC2 Support

Server-side Health Checks



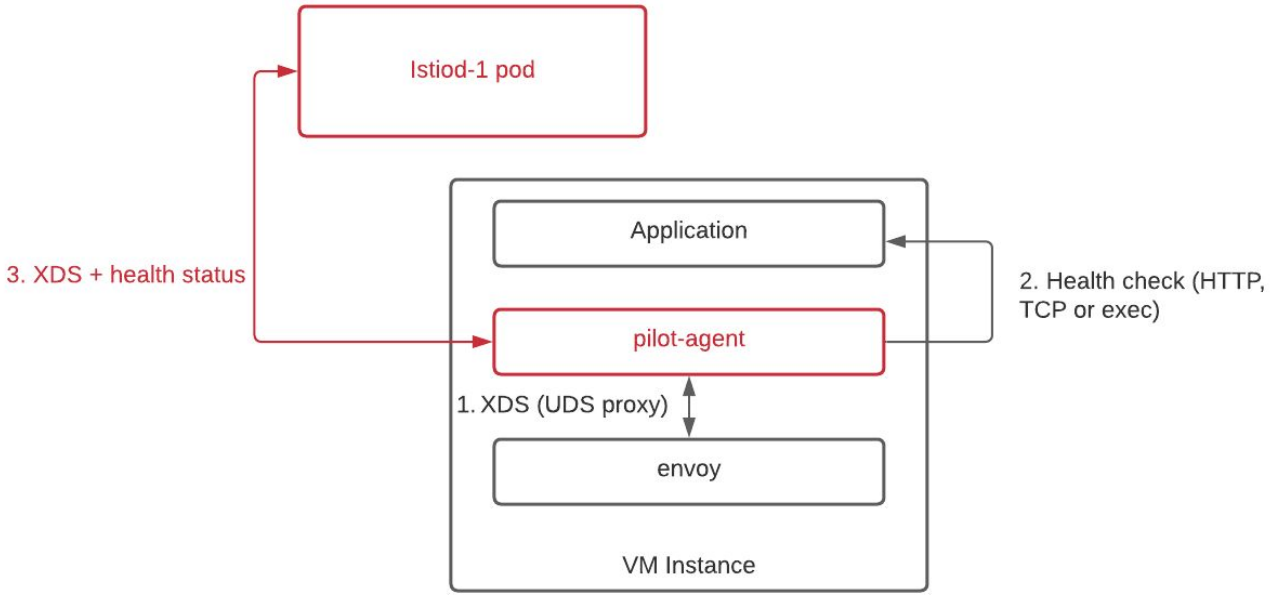
EC2 Support

Server-side Health Checks



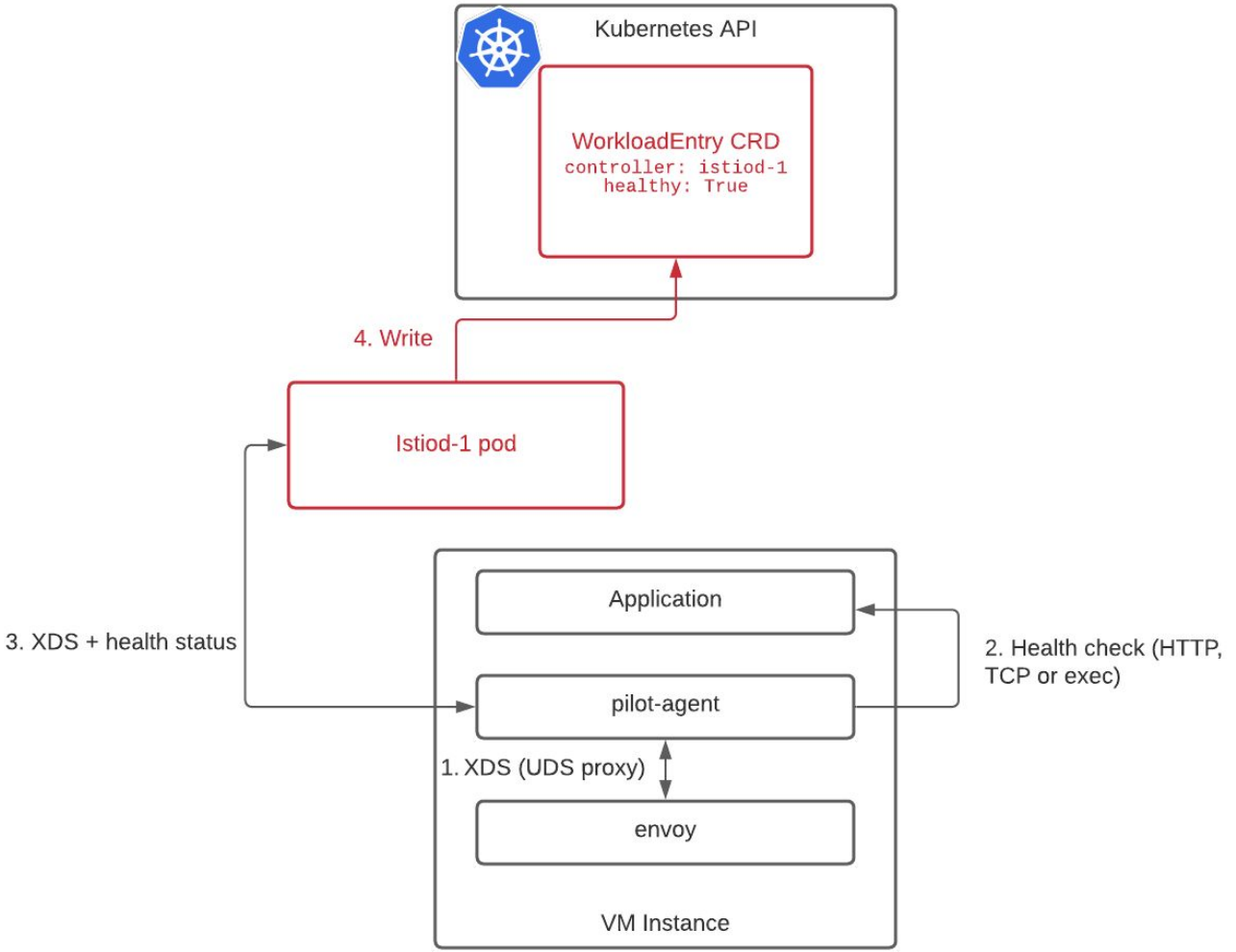
EC2 Support

Server-side Health Checks



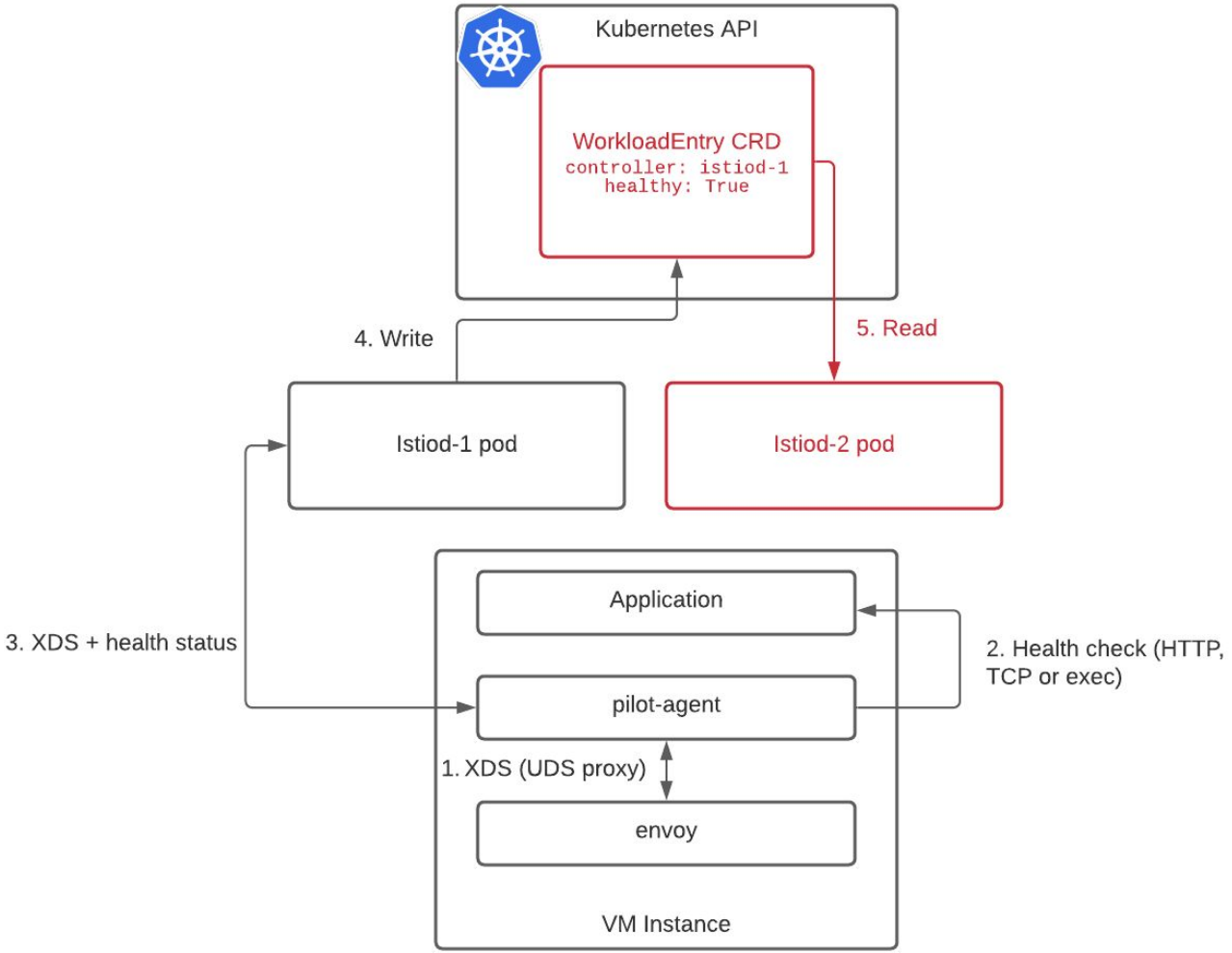
EC2 Support

Server-side Health Checks



EC2 Support

Server-side Health Checks

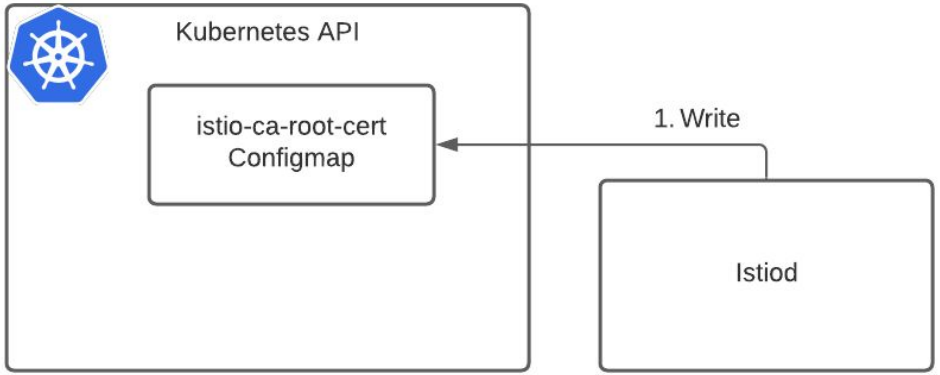


EC2 Support

TLS

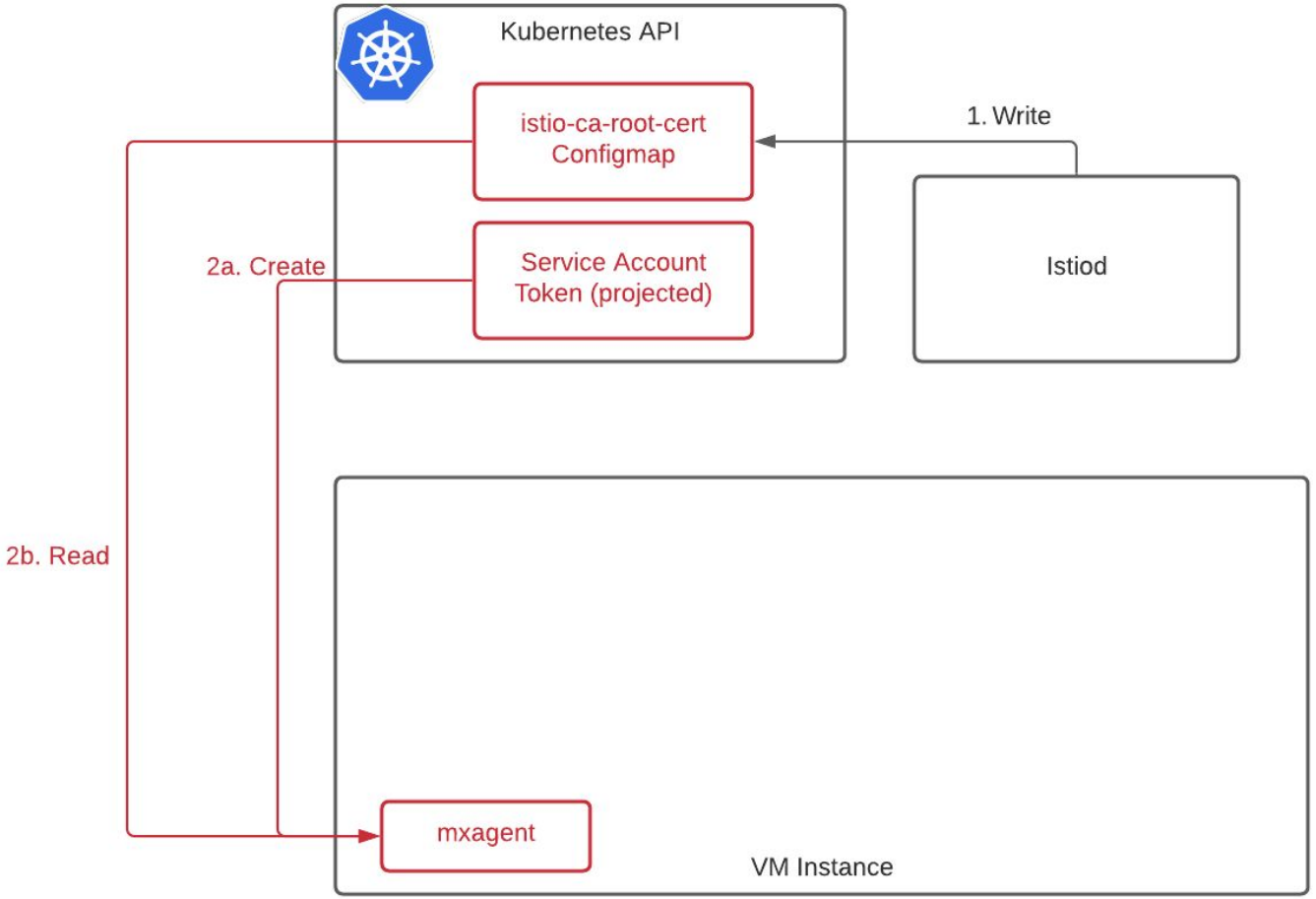
EC2 Support

TLS



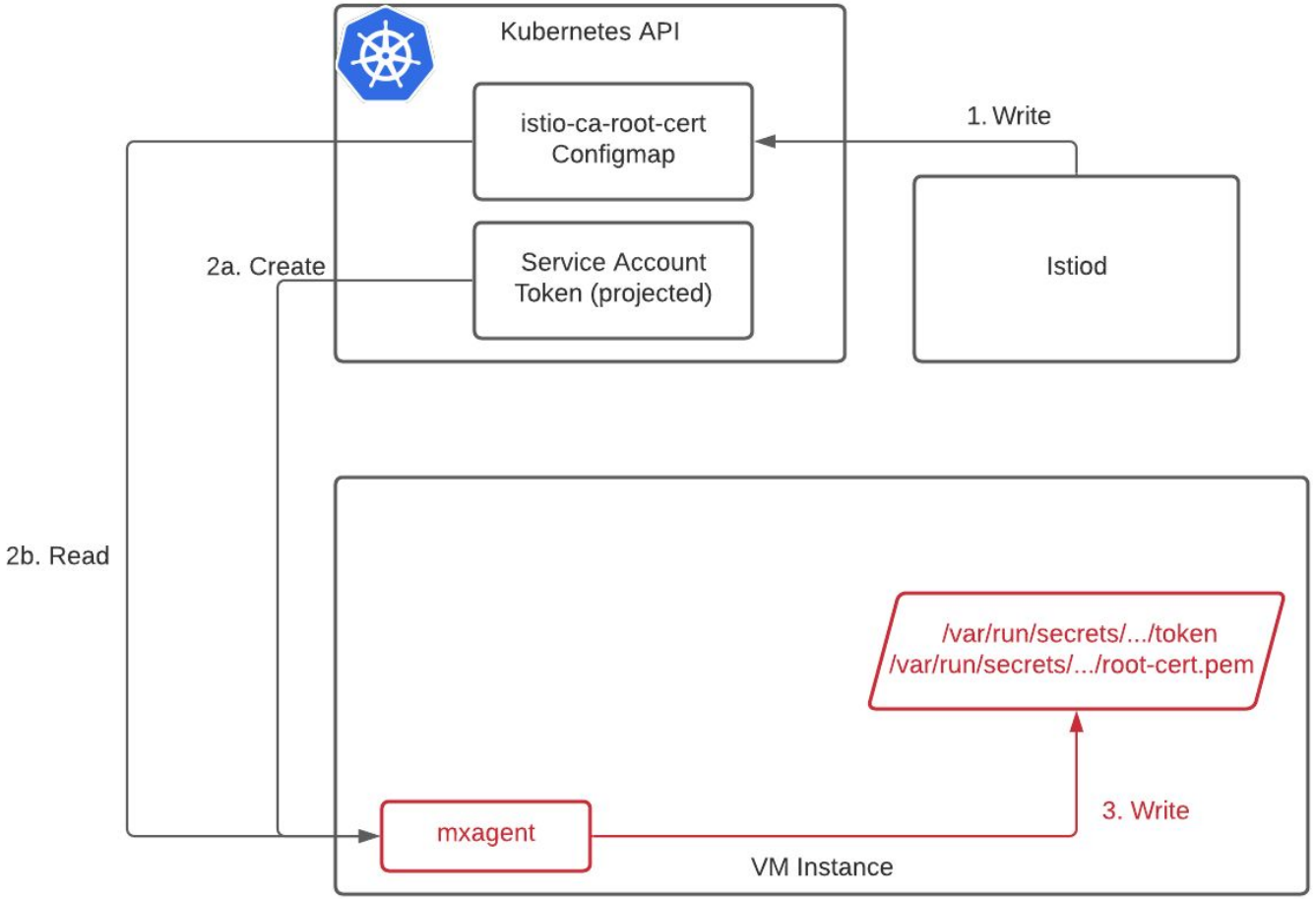
EC2 Support

TLS



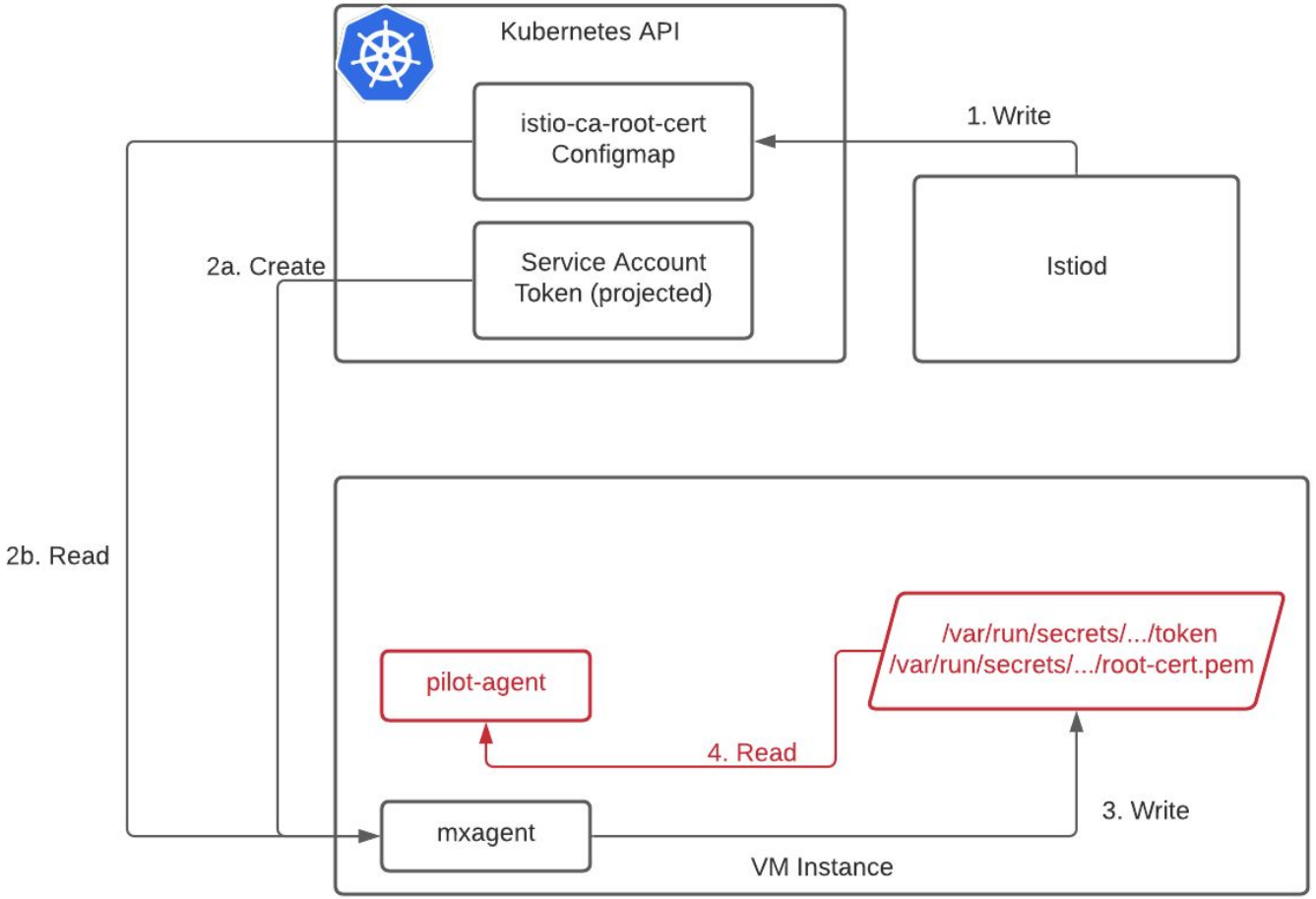
EC2 Support

TLS



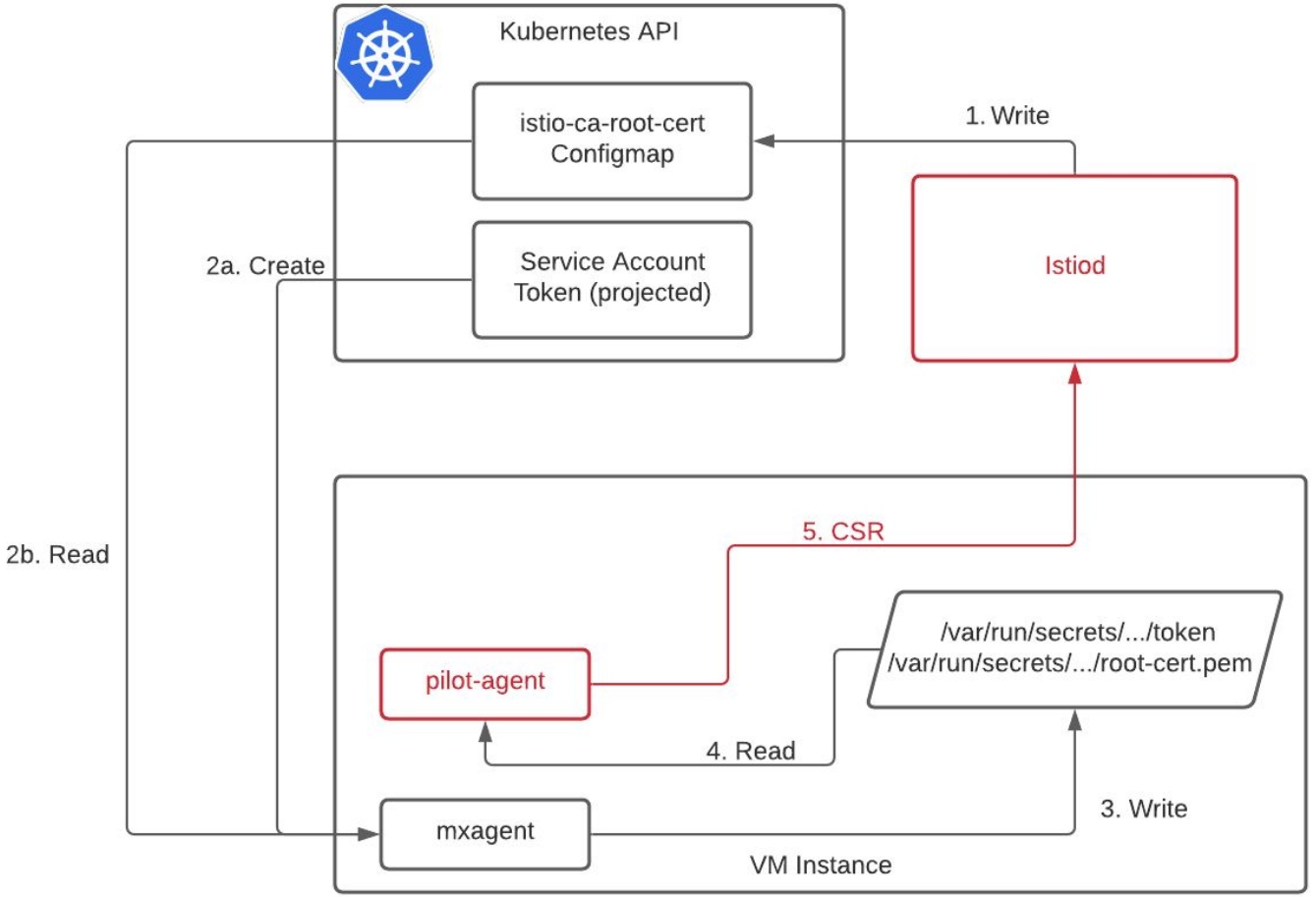
EC2 Support

TLS



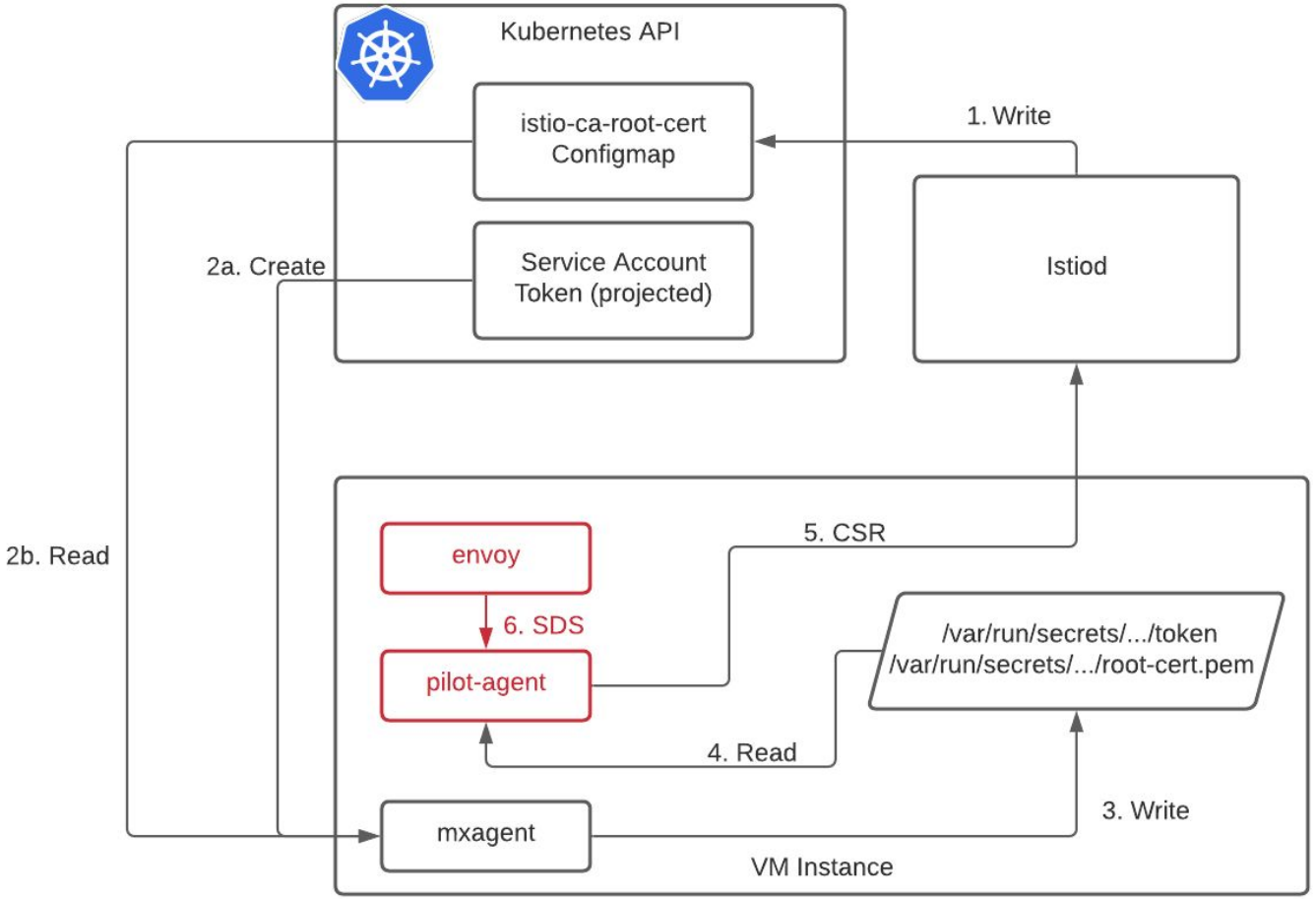
EC2 Support

TLS



EC2 Support

TLS

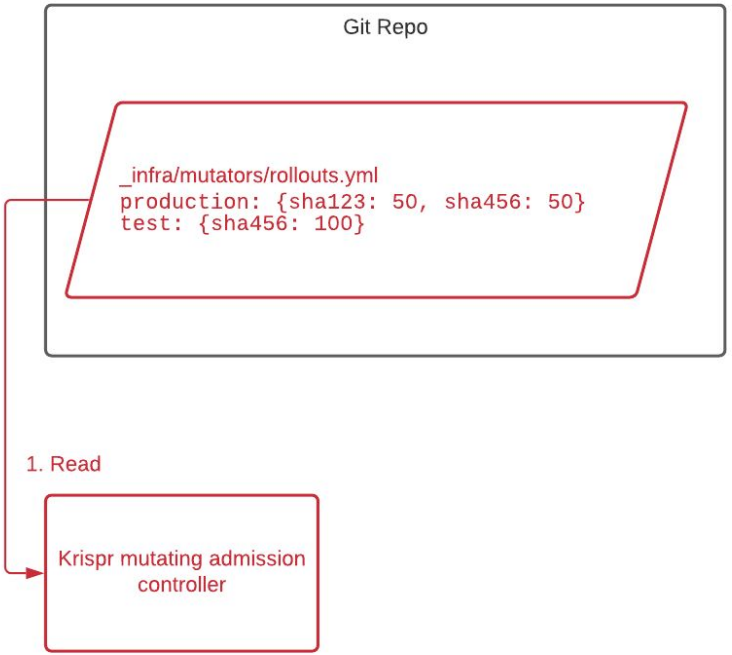


EC2 Support

Proxy Injection and Istio Version Upgrade

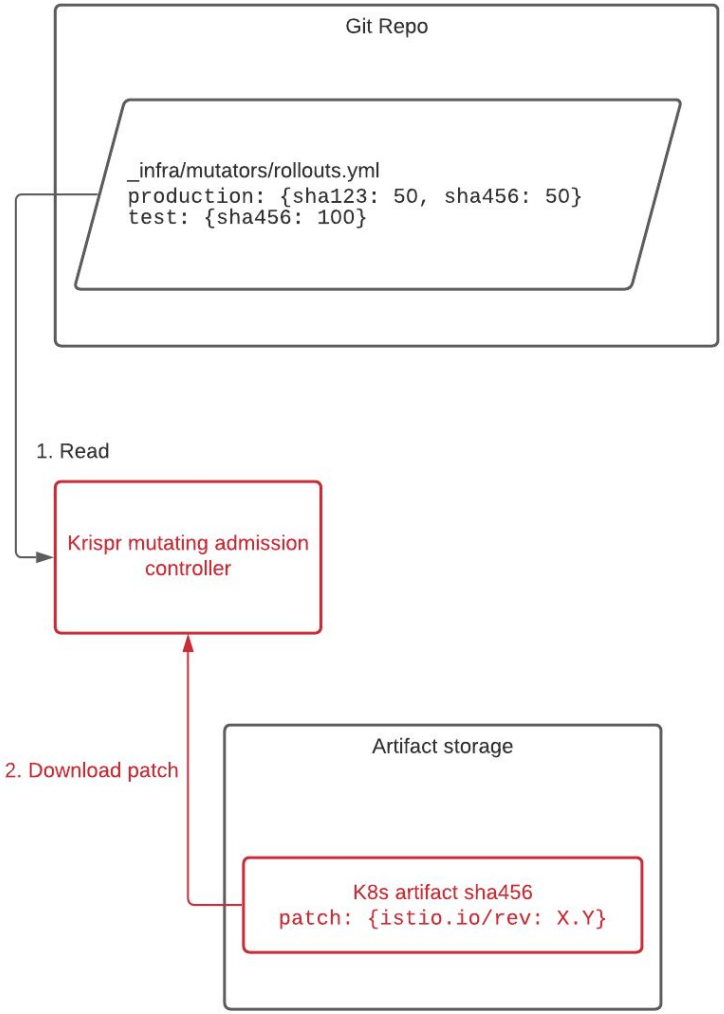
EC2 Support

Proxy Injection and Istio Version Upgrade



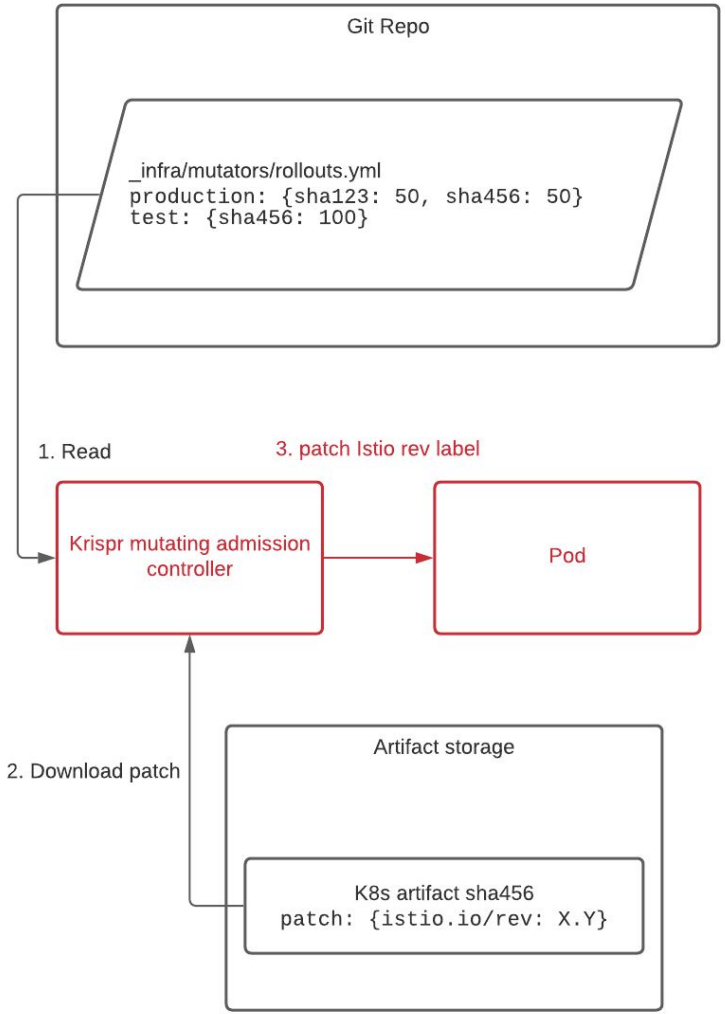
EC2 Support

Proxy Injection and Istio Version Upgrade



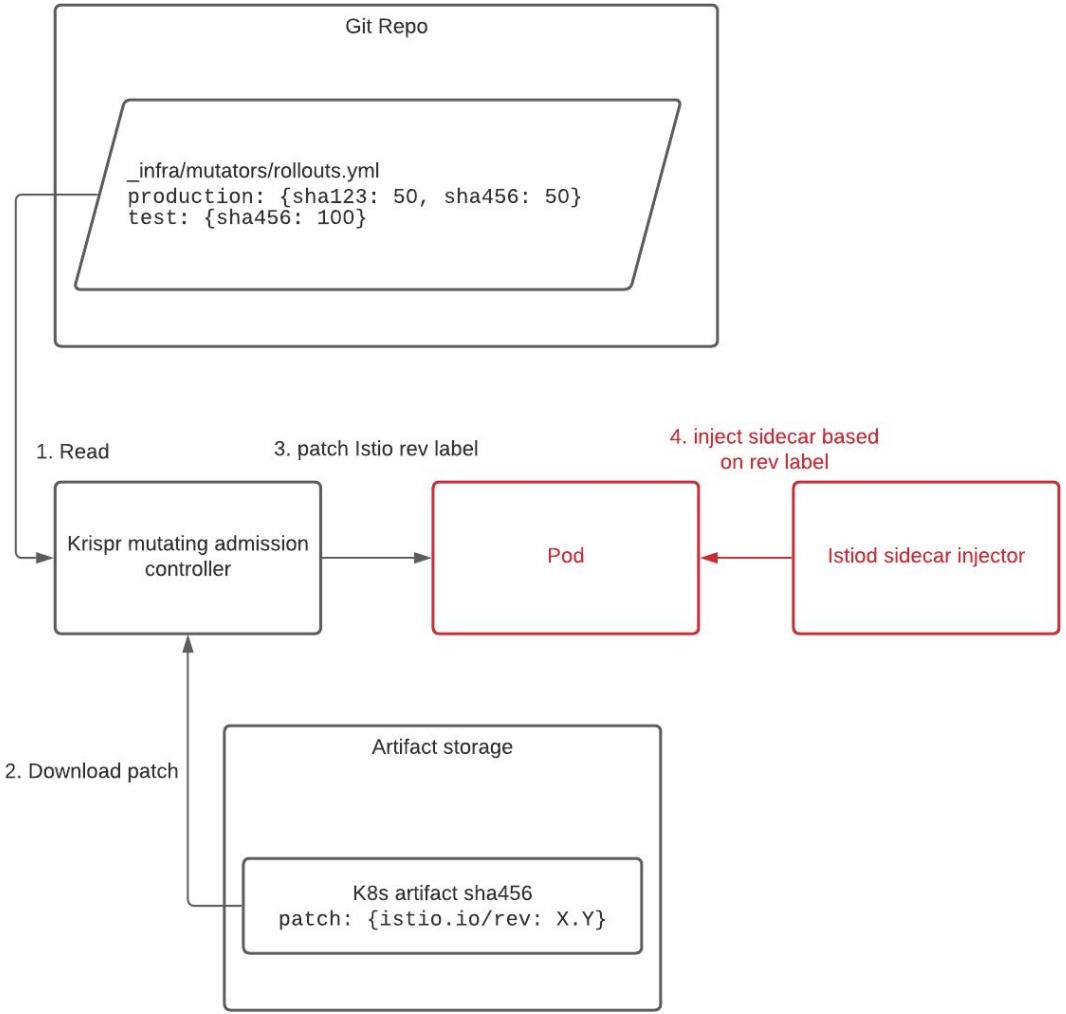
EC2 Support

Proxy Injection and Istio Version Upgrade



EC2 Support

Proxy Injection and Istio Version Upgrade

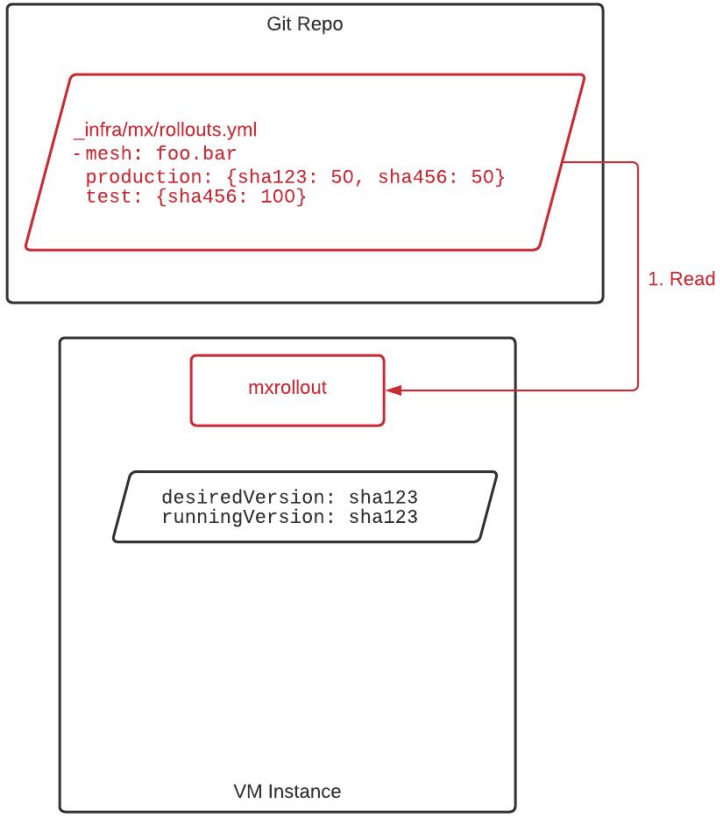


EC2 Support

Proxy Injection and Istio Version Upgrade

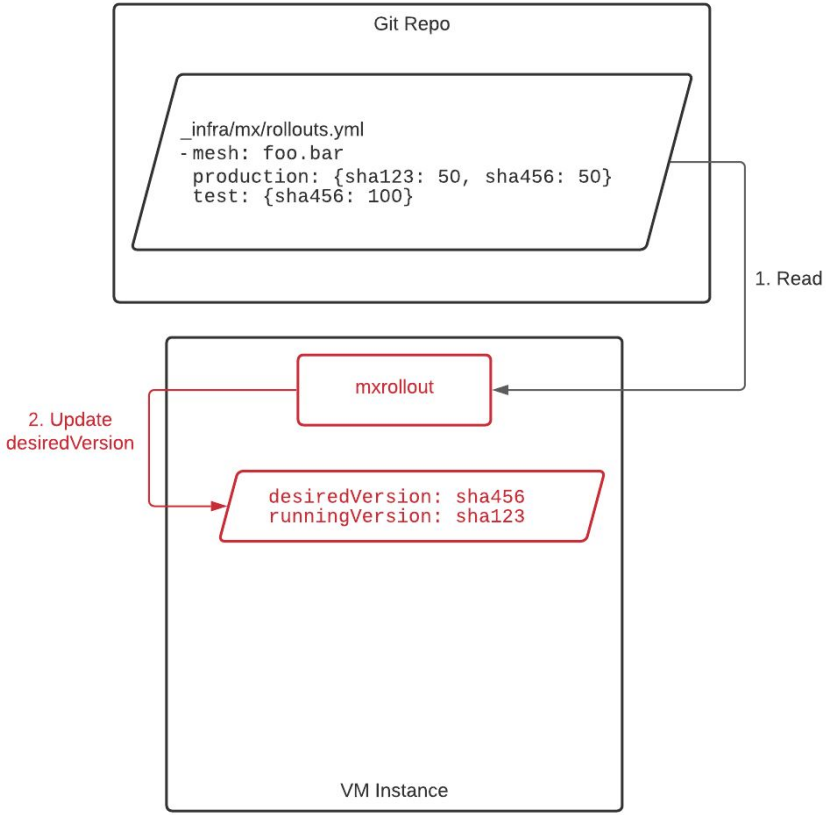
EC2 Support

Proxy Injection and Istio Version Upgrade



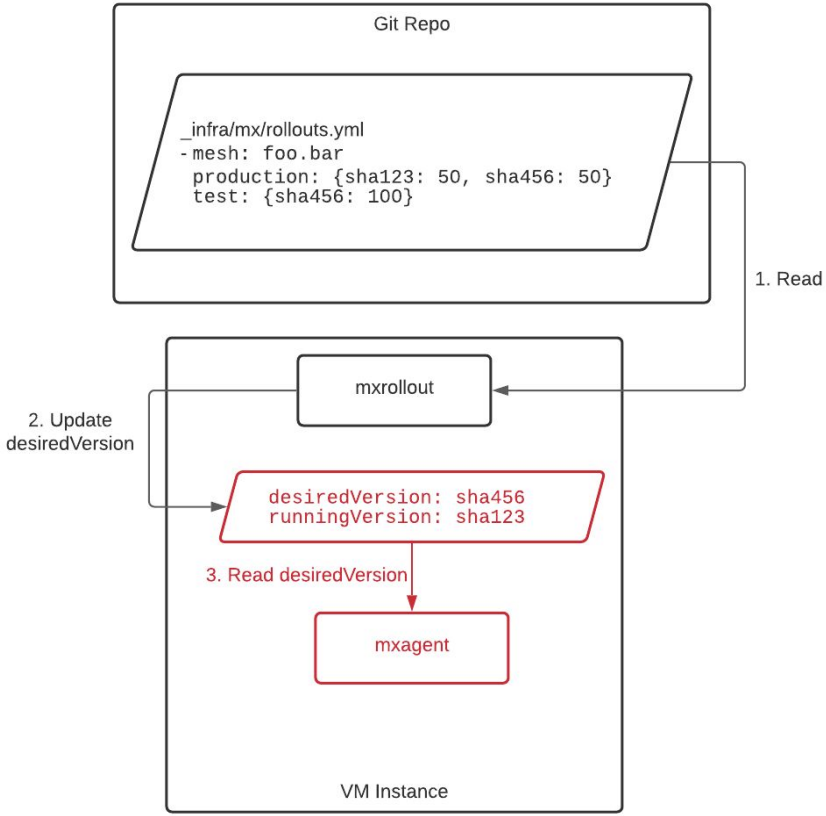
EC2 Support

Proxy Injection and Istio Version Upgrade



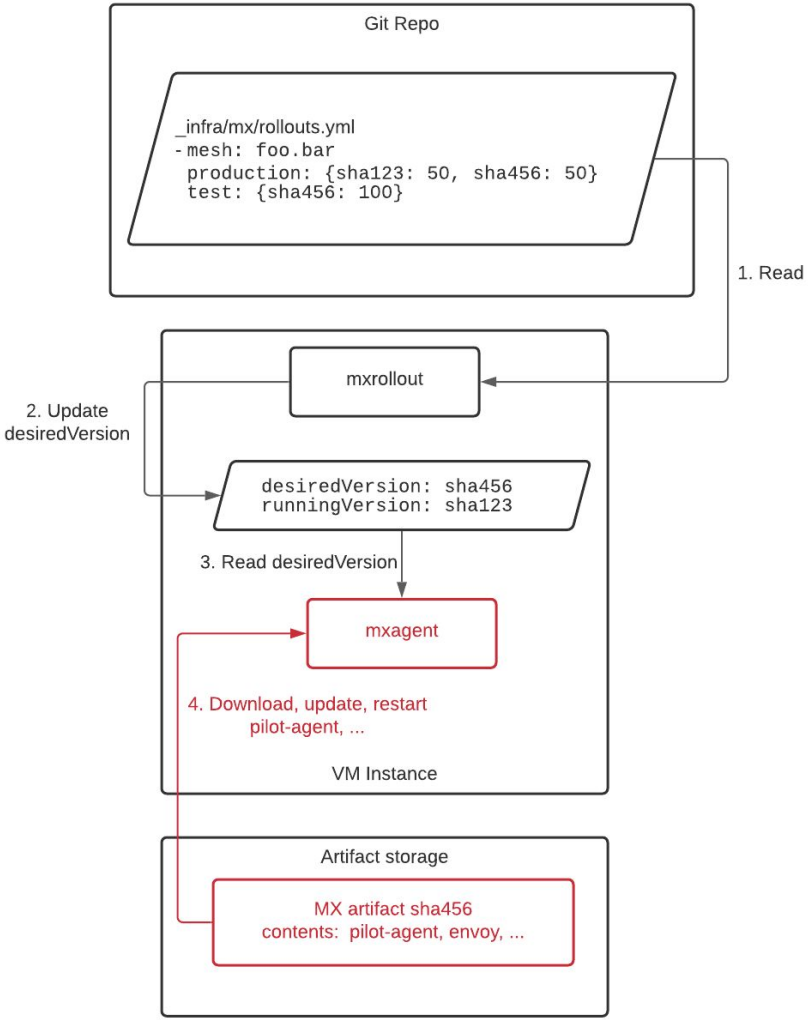
EC2 Support

Proxy Injection and Istio Version Upgrade



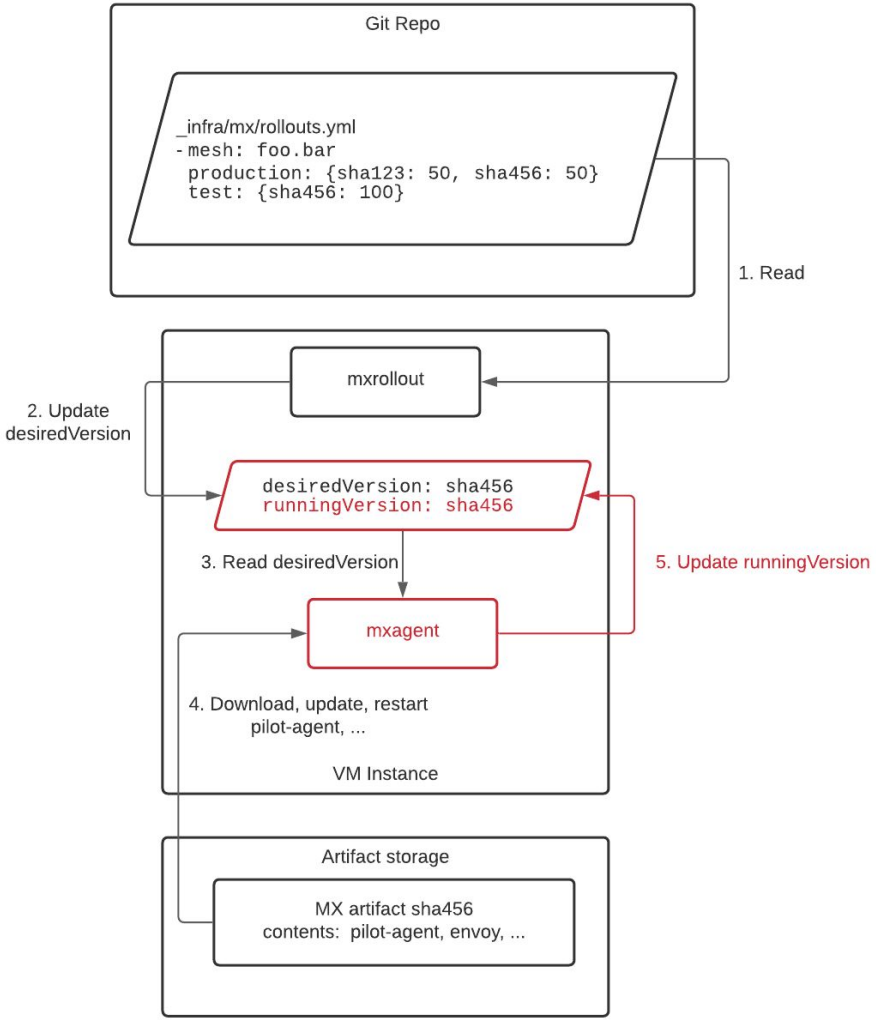
EC2 Support

Proxy Injection and Istio Version Upgrade



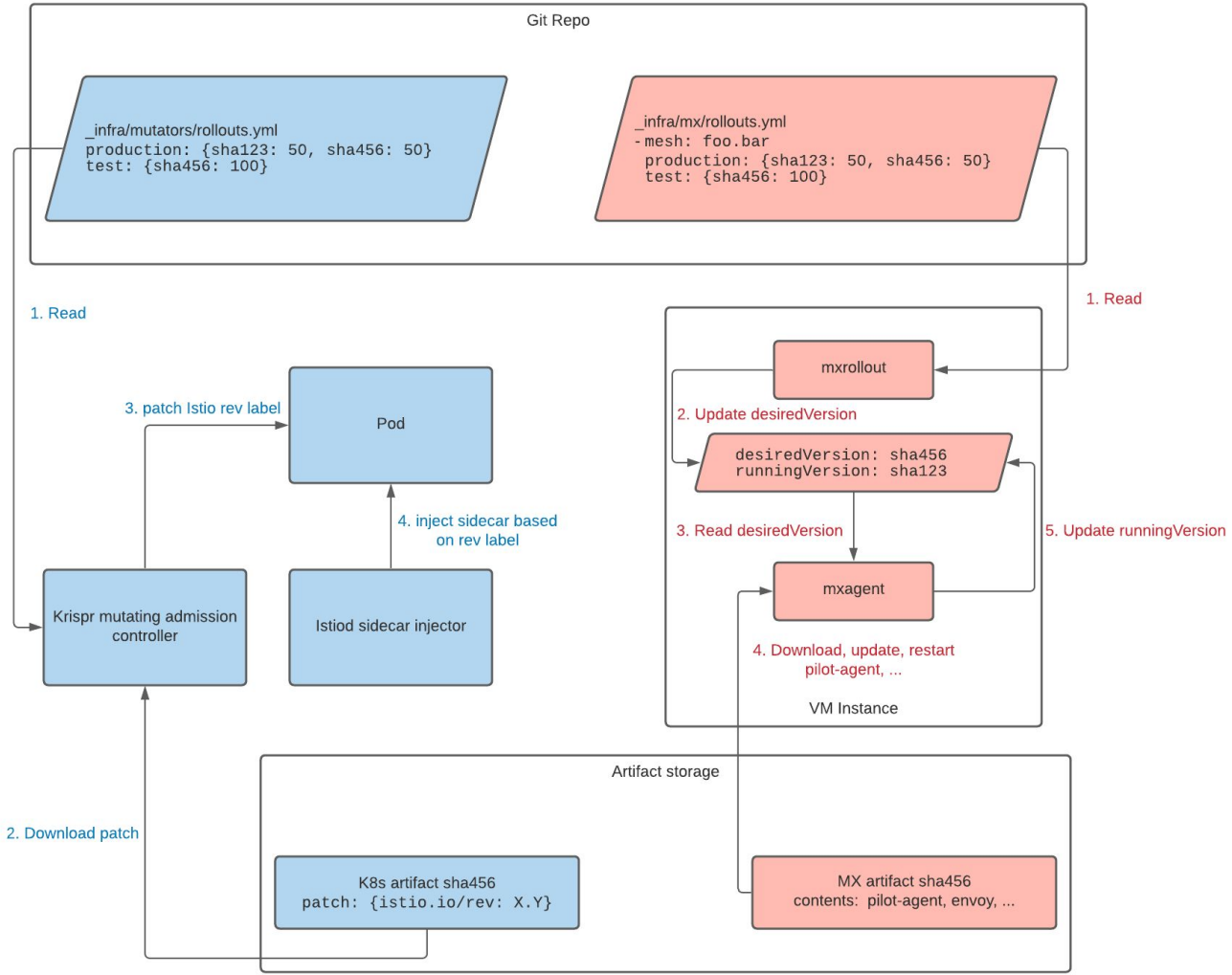
EC2 Support

Proxy Injection and Istio Version Upgrade



EC2 Support

Proxy Injection and Istio Version Upgrade



Multi Environment

External Services

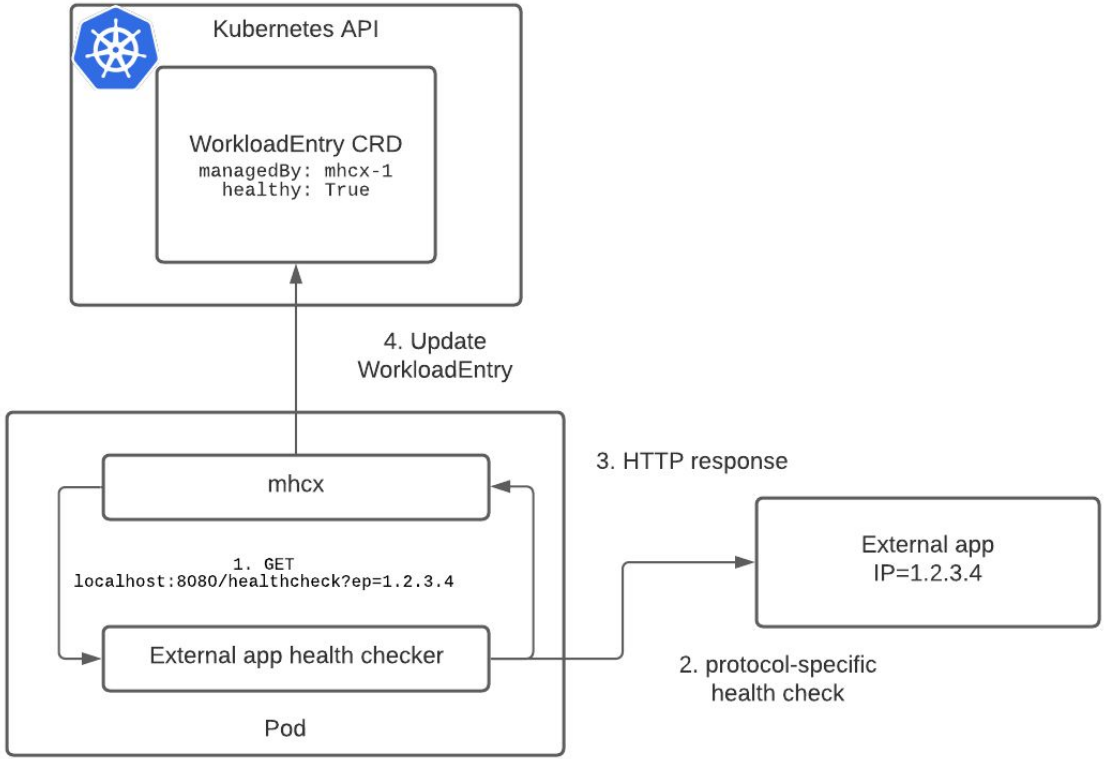
External Service Support

Requirements

- Server-side health checks
  - Clear ownership model for non-http health checks
- DNS name resolves to management cluster service IP
- Removal of stale endpoints

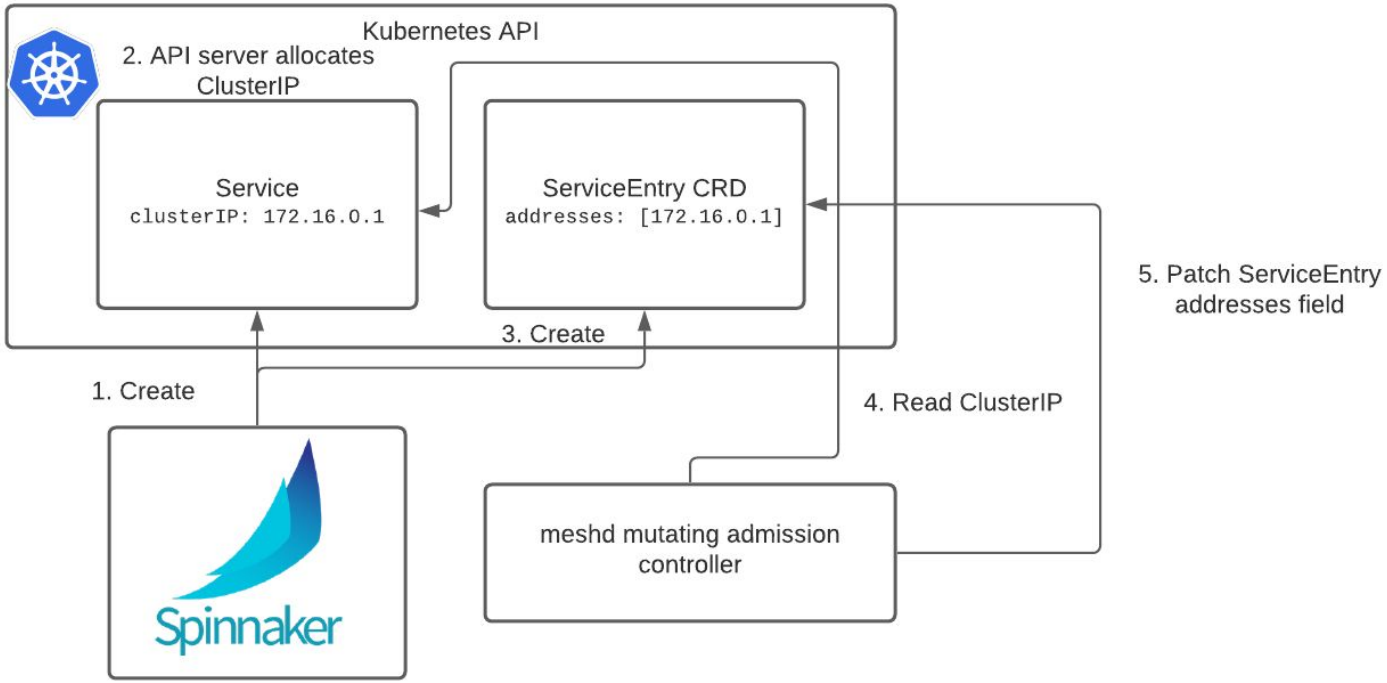
External Service Support

Server-side health checks



External Service Support

DNS ClusterIP setup



Key Concepts

Key Concepts

1. Globally flat IPv4 network
2. Single, external management cluster
3. Multiple tiers of pre-production meshes
4. Full feature parity for EC2, and path to migration to k8s
5. Generic mechanisms for adding non-http and external services

We’re hiring!

Open positions: airbnb.com/careers