《Istio大咖说》第5期

腾讯云服务网格生产落地最佳实践

主持人：宋净超（Tetrate）　嘉宾：钟华（腾讯云）

6月30日　晚8:00 – 9:00

联合主办方：tetrate　云原生社区 Cloud Native Community　ServiceMesher

扫码观看直播

tetrate

THE ENTERPRISE SERVICE MESH COMPANY

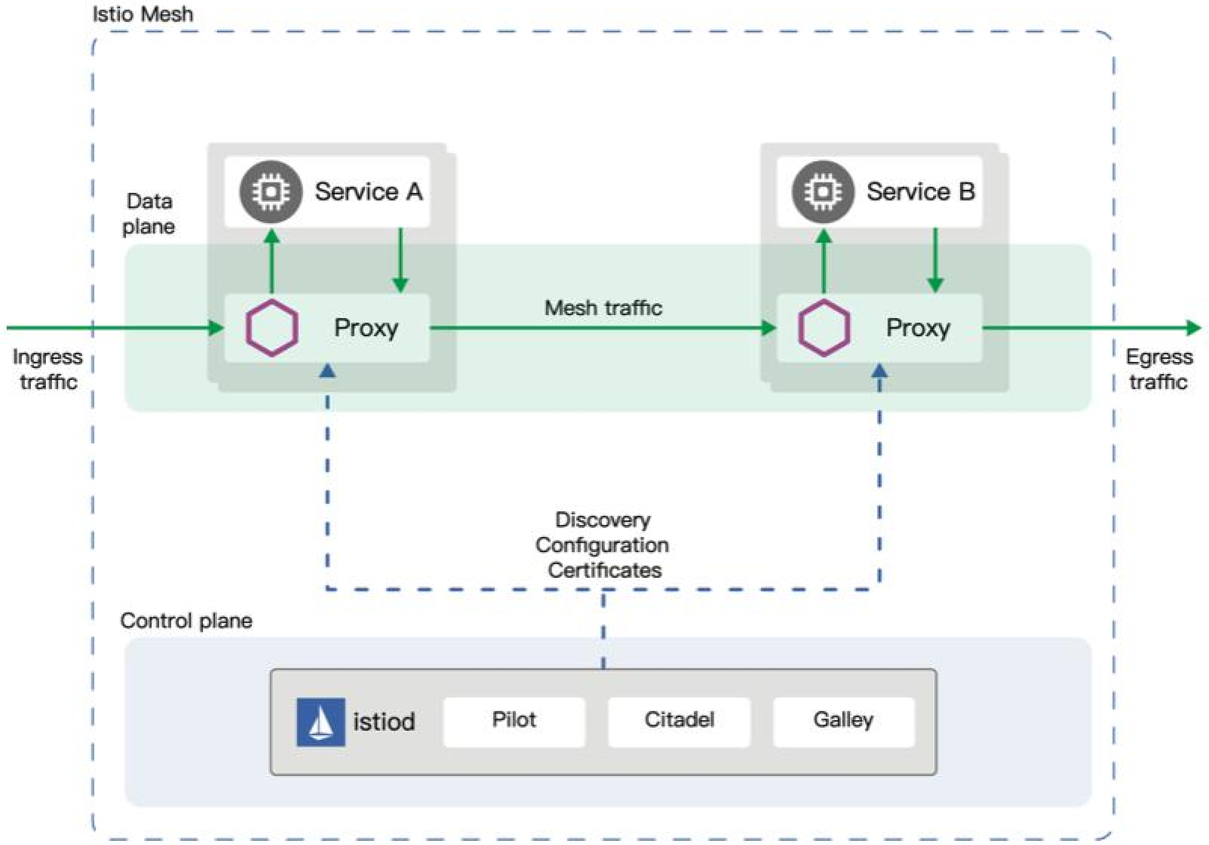
Agenda

Istio 大规模生产落地挑战

腾讯云服务网格 TCM 介绍

Istio 生产落地最佳实践

Istio 大规模生产落地挑战



性能开销

- Envoy 资源占用高
- 流量劫持，访问链路增长
- xDS 全量下发，缺乏分级

场景限制

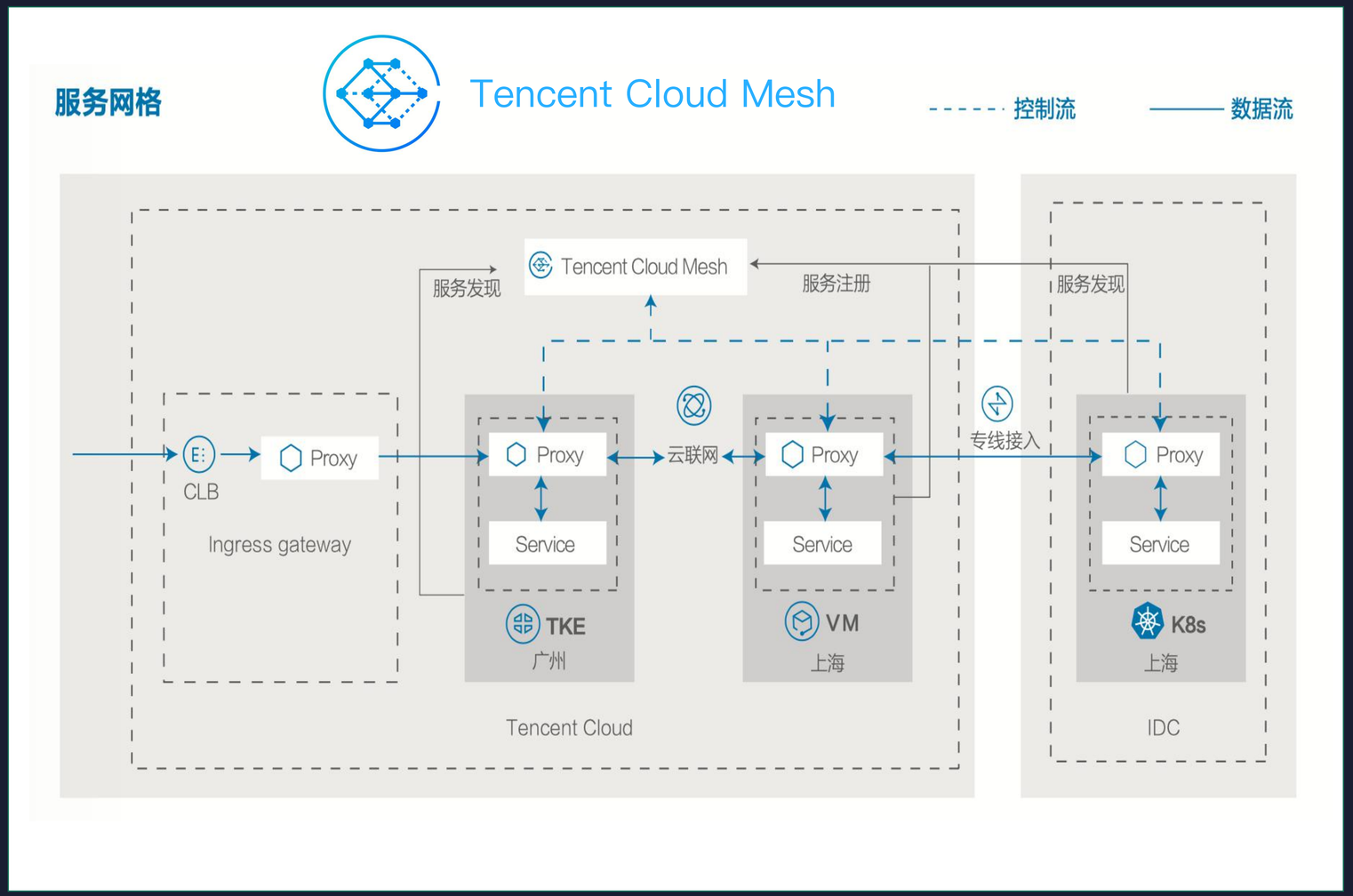
- 有限的协议支持
- 对非 K8S 平台支持有限
- 解决问题有限
- 扩展成本高

复杂度

- 管理复杂，版本迭代快
- 并非完全透明
- 学习成本高

腾讯云服务网格 TCM 介绍

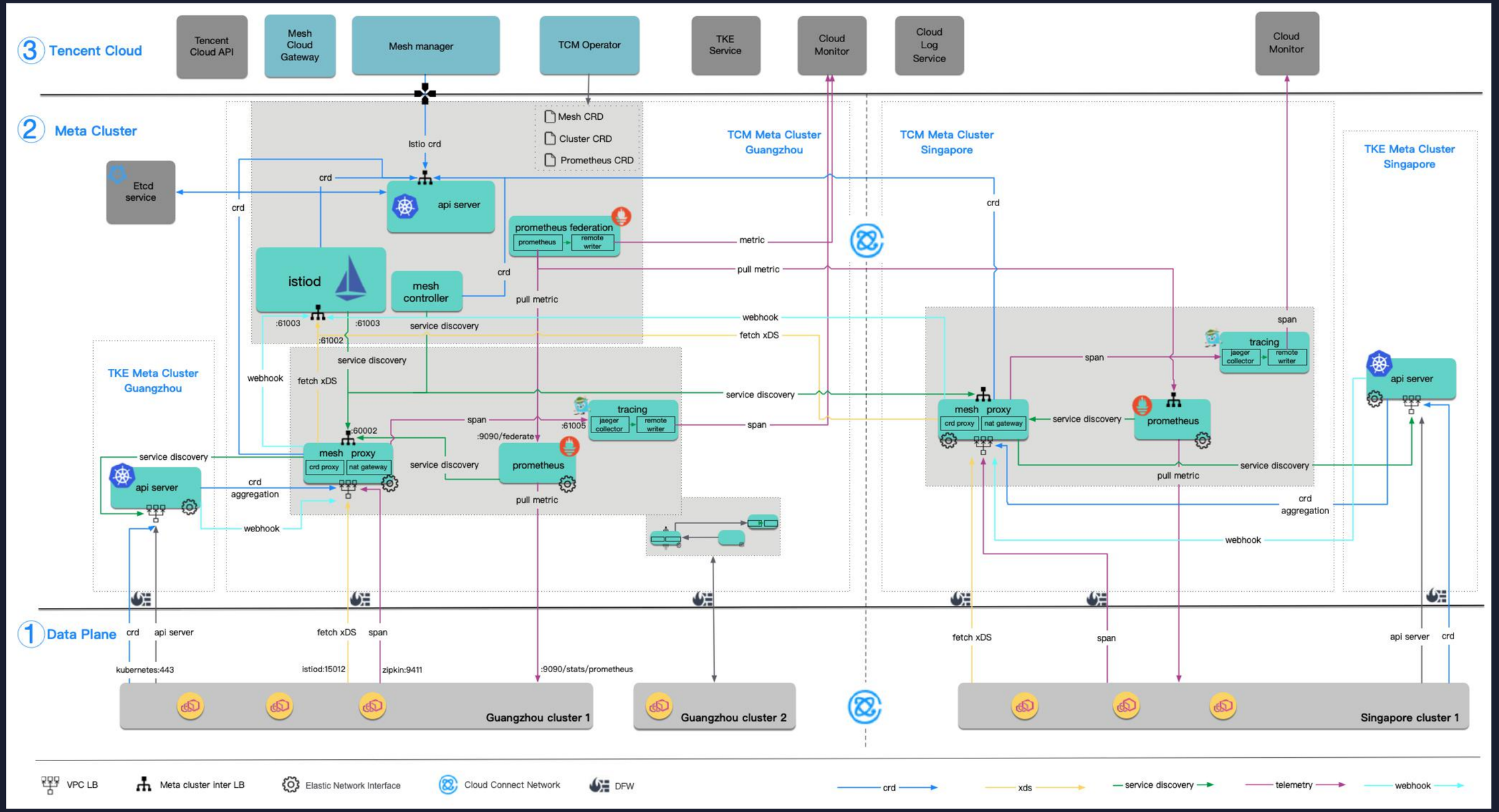
腾讯云服务网格 TCM



- 生态完善，兼容 Istio
- 全生命周期管理，控制面灰度升级
- 两种部署：独立版，托管版
- 全托管遥测系统
- 深度性能优化
- 协议和服务发现扩展，Aeraki（已开源）
- 能力拓展：边缘网关证书，跨地域互通...
- 和腾讯云深度集成：TKE，云监控，CLS ...

云原生应用网络管控基础平台

TCM 全托管架构


3 Tencent Cloud
Tencent Cloud API
Mesh Cloud Gateway
Mesh manager
TCM Operator
TKE Service
Cloud Monitor
Cloud Log Service
Cloud Monitor
2 Meta Cluster
Mesh CRD
Cluster CRD
Prometheus CRD
TCM Meta Cluster Guangzhou
TCM Meta Cluster Singapore
TKE Meta Cluster Singapore
TKE Meta Cluster Guangzhou
Istio crd
Etcd service
crd
api server
prometheus federation
prometheus
remote writer
metric
pull metric
istiod
mesh controller
:61003
:61002
service discovery
webhook
fetch xDS
span
tracing
jaeger collector
:61005
:60002
:9090/federate
mesh proxy
crd proxy
nat gateway
prometheus
api server
crd aggregation
1 Data Plane
kubernetes:443
istiod:15012
zipkin:9411
:9090/stats/prometheus
Guangzhou cluster 1
Guangzhou cluster 2
Singapore cluster 1
VPC LB
Meta cluster inter LB
Elastic Network Interface
Cloud Connect Network
DFW
crd
xds
service discovery
telemetry
webhook

tetrate

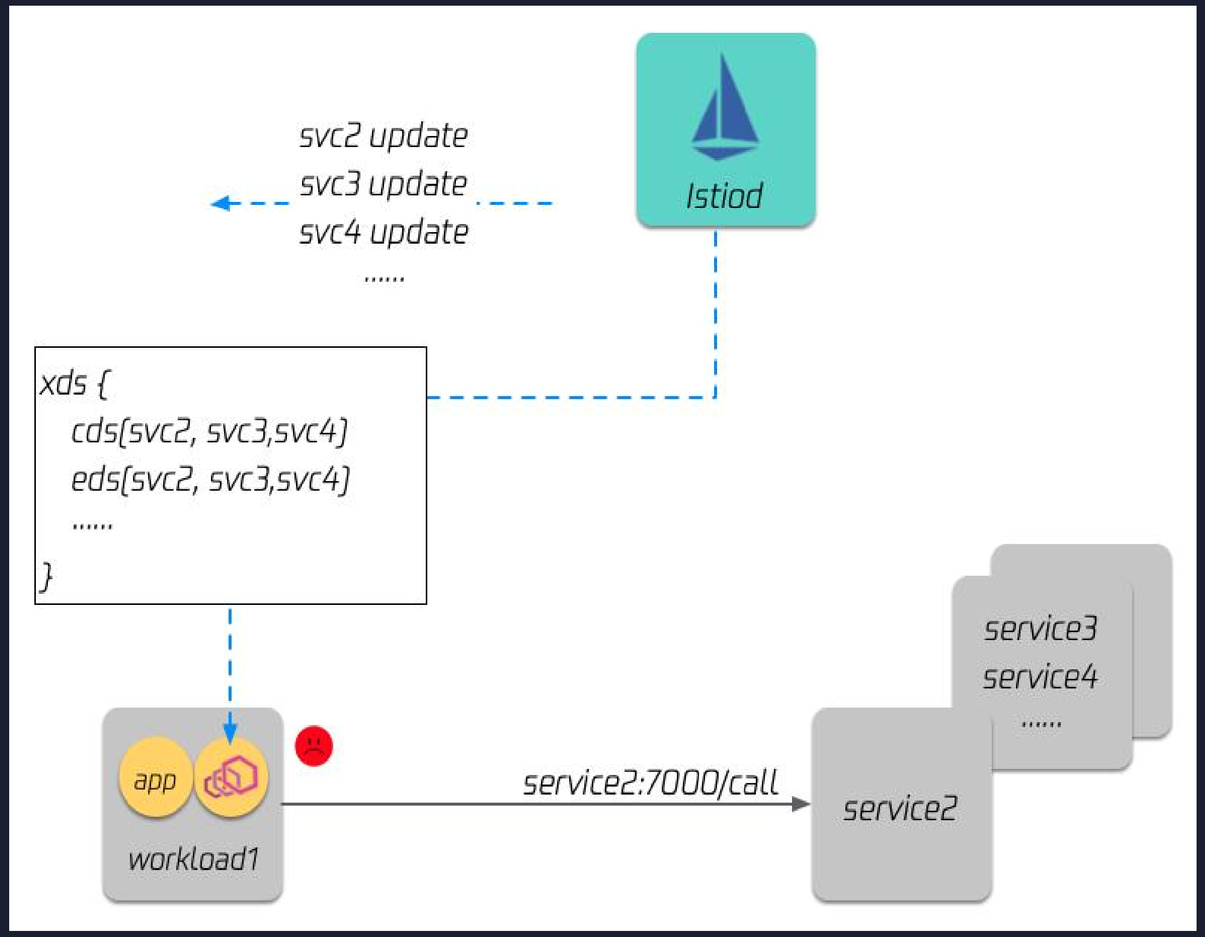
生产落地最佳实践

- xDS Lazy Loading
- 控制面灰度升级
- 控制面负载平衡
- Istio CRD 中心托管
- Istio 避坑指南

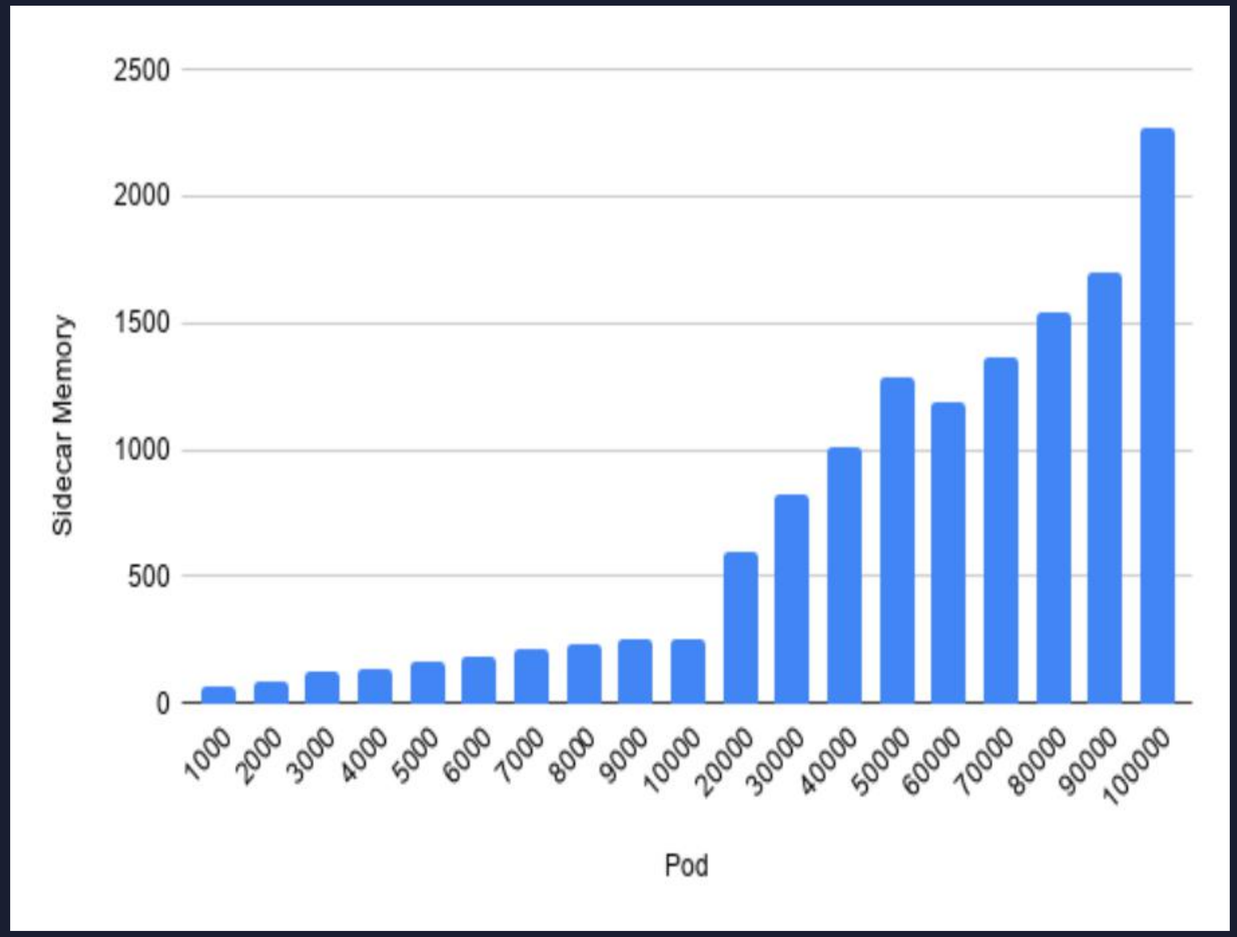
tetrate

最佳实践–xDS Lazy Loading

大规模场景下 xDS 下发瓶颈

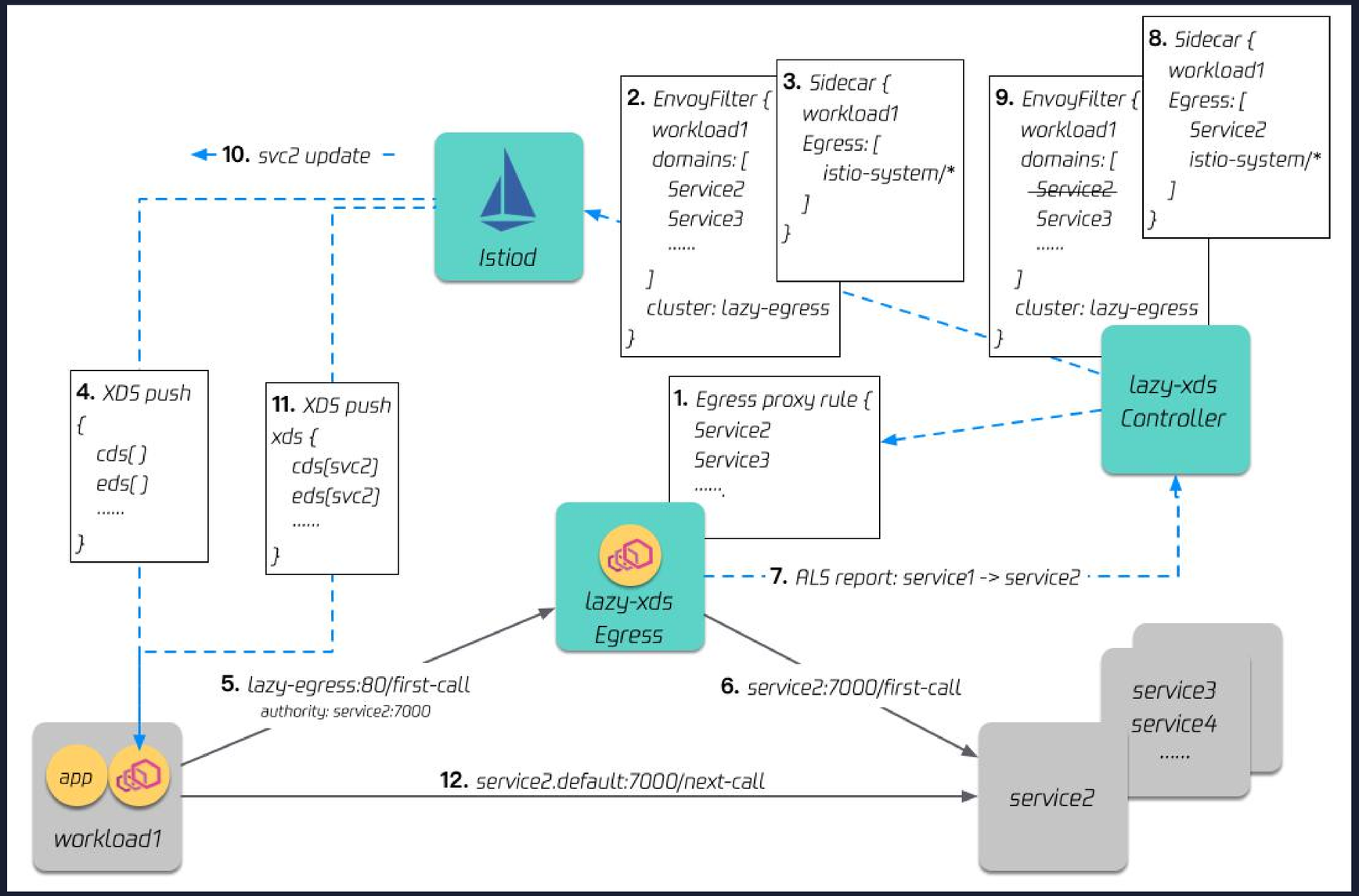


xDS 全量下发



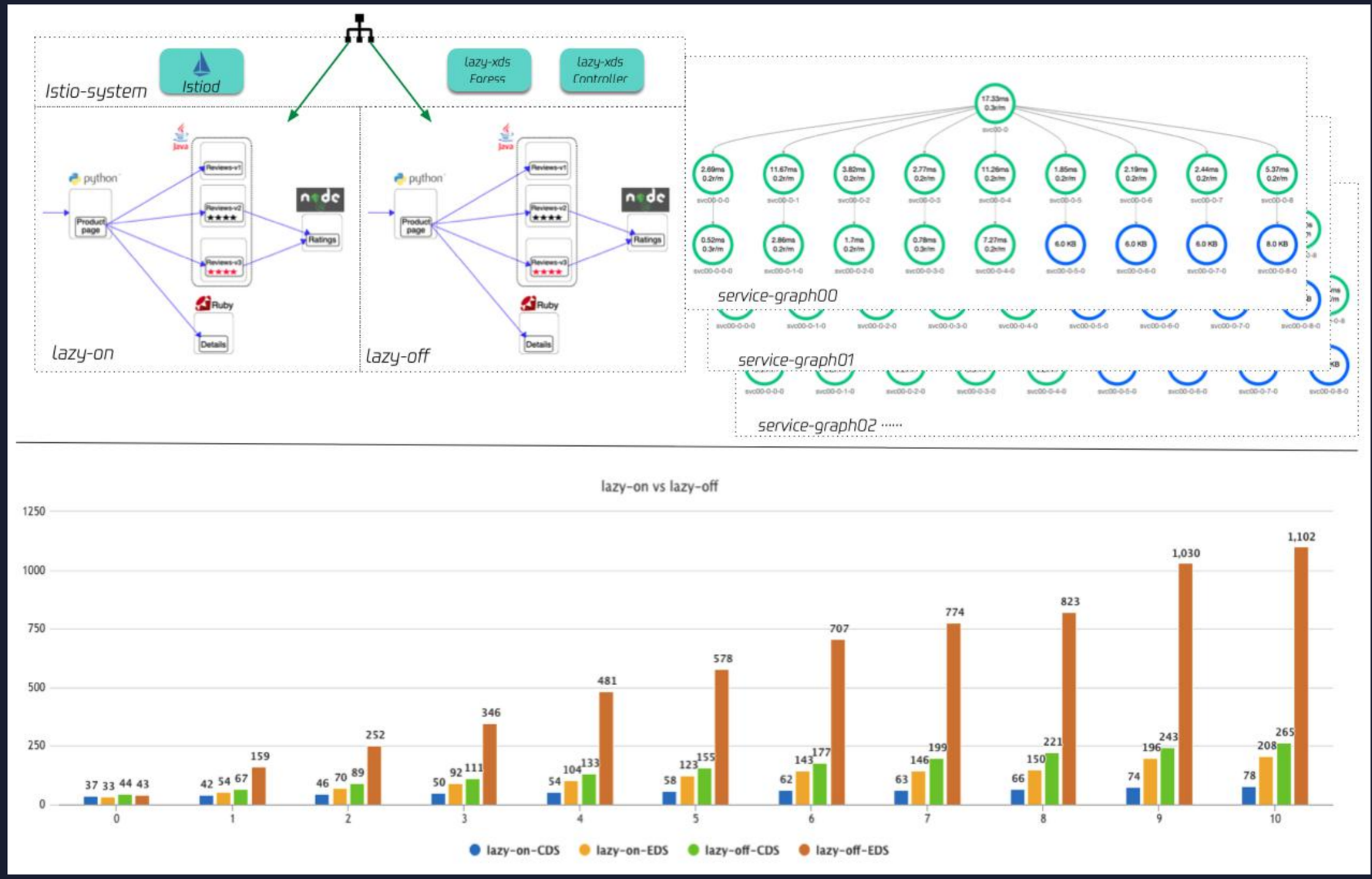
sidecar 内存随着网格规模增长

xDS Lazy Loading



- 无需提前配置服务依赖
- 允许服务依赖动态新增
- 只获取本身依赖服务数据
- 用户流量不阻塞
- 很小的性能损耗，且控制在首（几）次访问新服务

xDS Lazy Loading–对比方案



xDS Lazy Loading–优化效果

优化效果：

- 900 pods 规模 mesh，envoy 内存减少 14 M
- 10 万 pods 规模 mesh，envoy 内存降低约 1.5 G
- CDS/EDS 更新次数显著降低

Aeraki 项目简介

tetrate

服务发现扩展：

- Consul
- Nocos
- Eureka
- Zookeeper(dubbo)

……

协议扩展：

- Dubbo
- Redis
- Thrift
- ZK, Kafka

……

性能优化：

- Lazy xDS

低成本扩展任意七层协议：

- Meta Protocol

IstioCon2021: How to Manage Any Layer-7 Traffic in an Istio Service Mesh?

- Youtube：https://www.youtube.com/watch?v=sBS4utF68d8
- 中文：https://cloudnative.to/blog/istiocon-layer7-traffic/

项目地址：https://github.com/aeraki-framework/aeraki

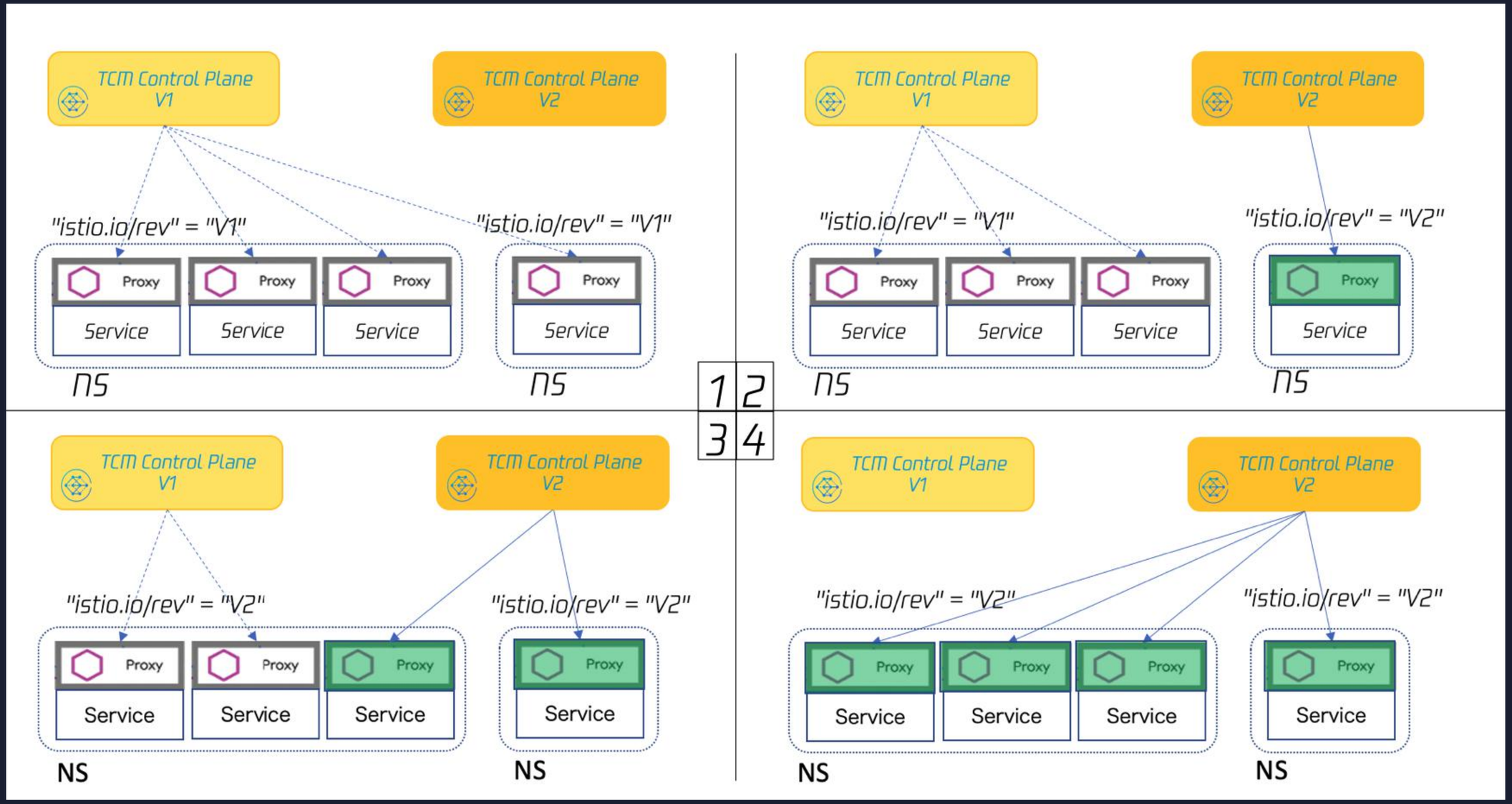
Lazy xDS：https://github.com/aeraki-framework/aeraki/blob/master/lazyxds/README.md

tetrate

最佳实践–控制面灰度升级

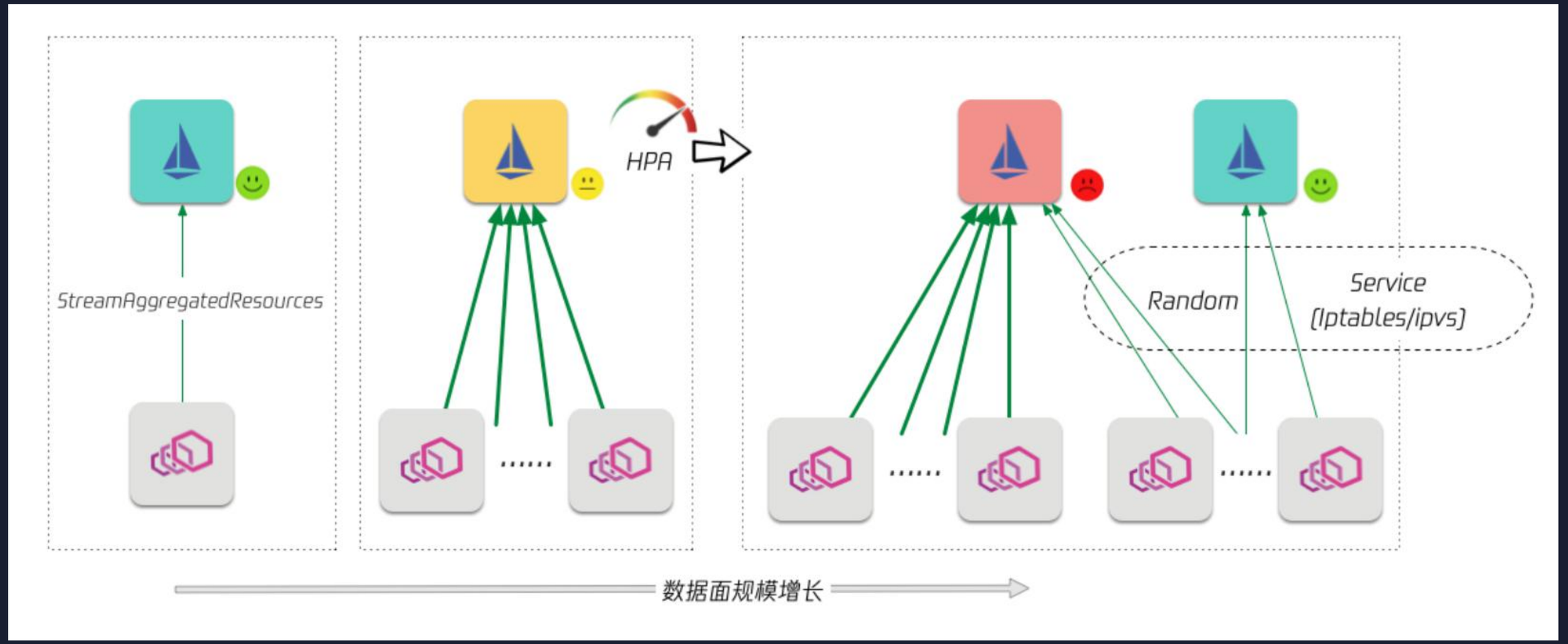
最佳实践–TCM 控制面灰度升级

tetrate



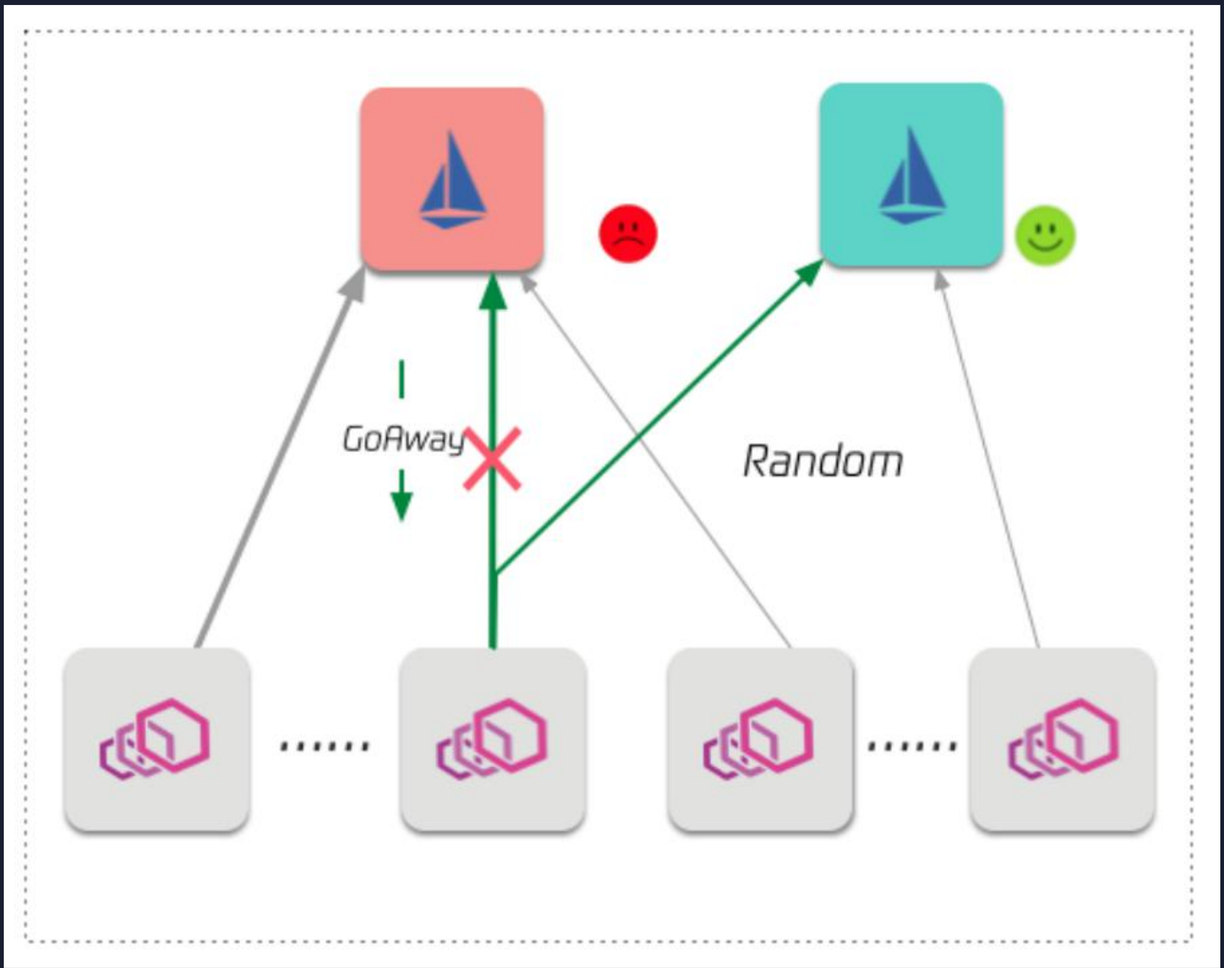
最佳实践–控制面负载平衡

控制面负载不均



控制面负载不均

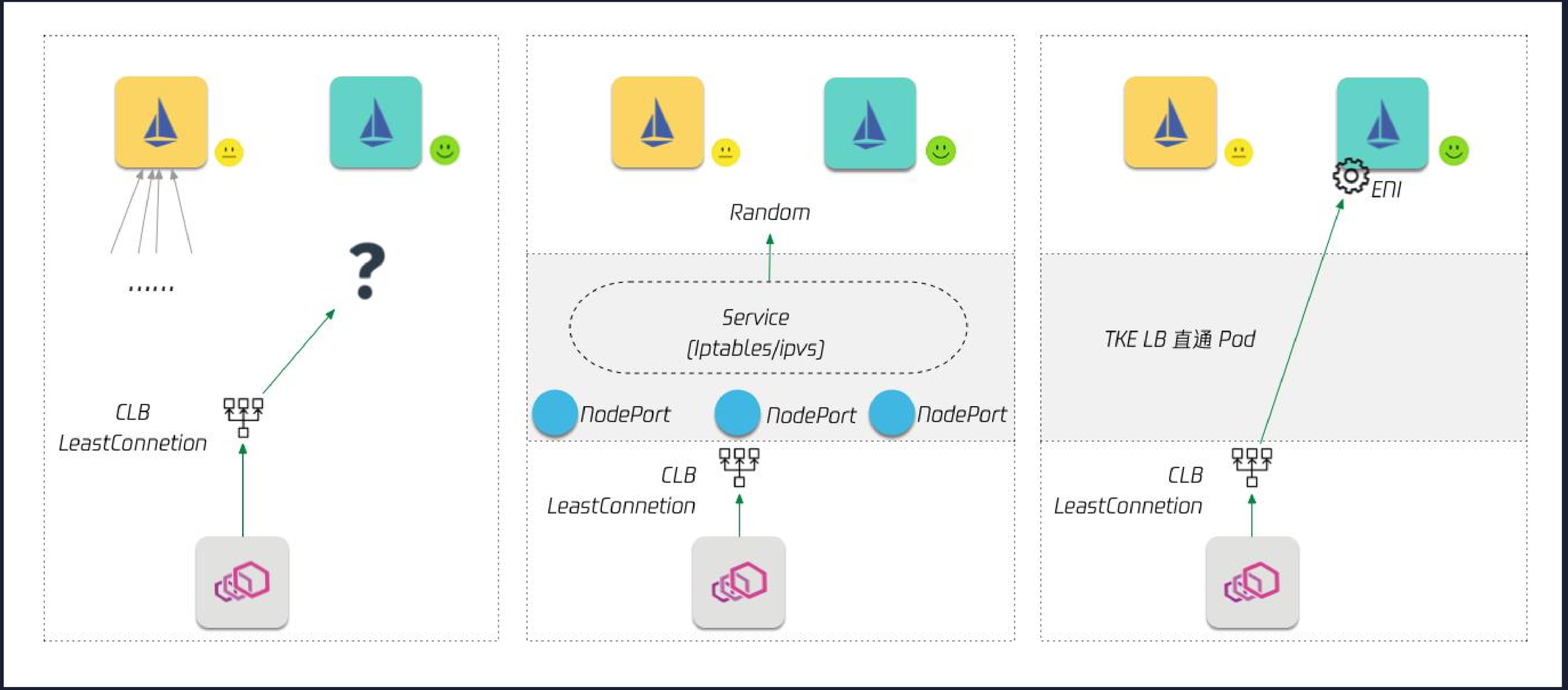
控制面负载平衡–社区方案



周期性断开 Server 和 Client 之间的长连接：

- MaxServerConnectionAge：
  连接最长连接时间
- MaxServerConnectionAgeGrace：
  强制断链前等待时间

TCM 控制面负载平衡



CLB +最小连接数算法+LB 直通Pod

控制面负载平衡效果

测试用例：

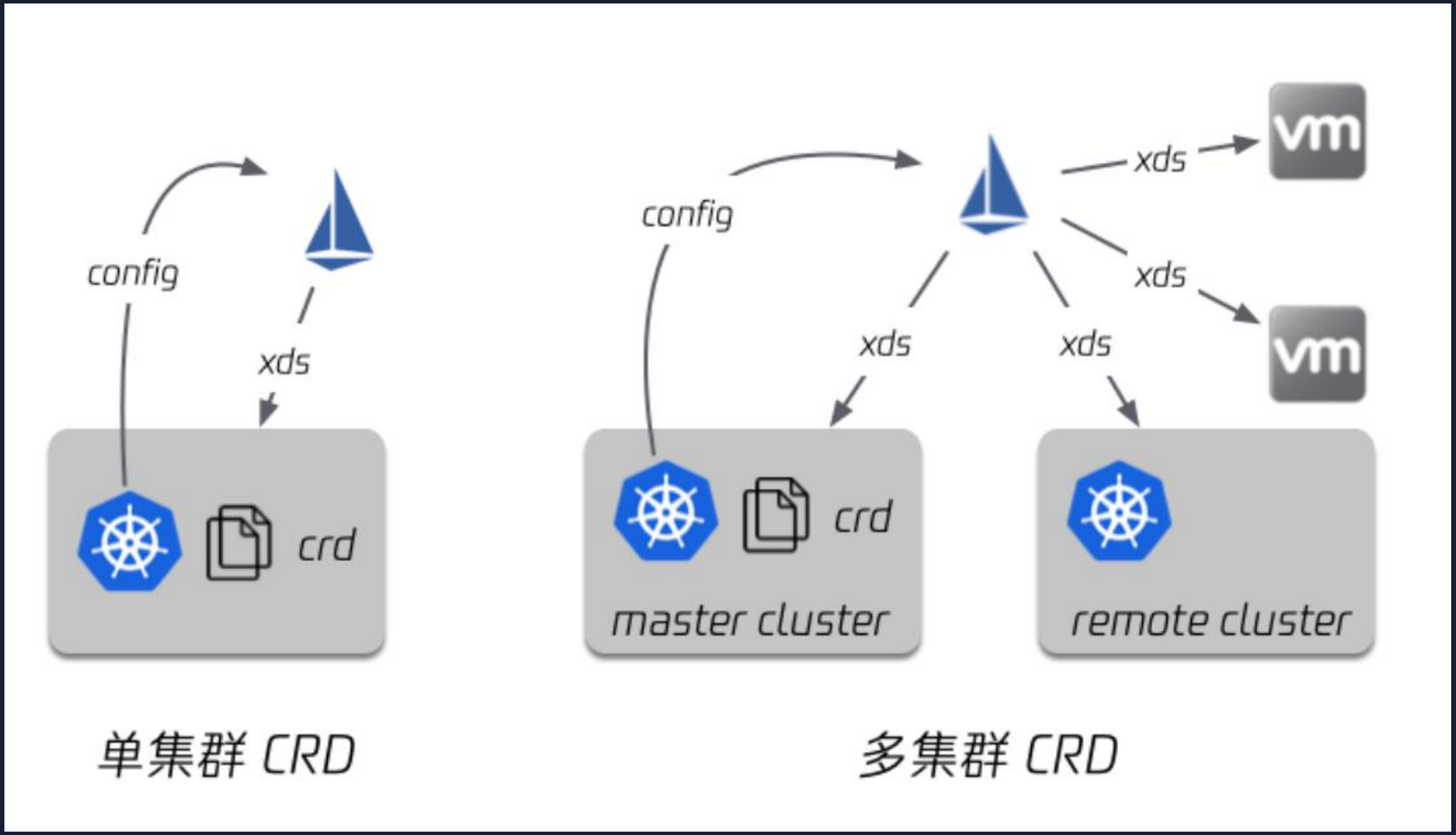
先启动 2个 istiod，60 个 数据面 pod

后续再将 istiod 扩容为3个，数据面扩容到90

- 默认情况： 标准差（43 38 9）= 18
- 优化后： 标准差（34 30 26）= 4

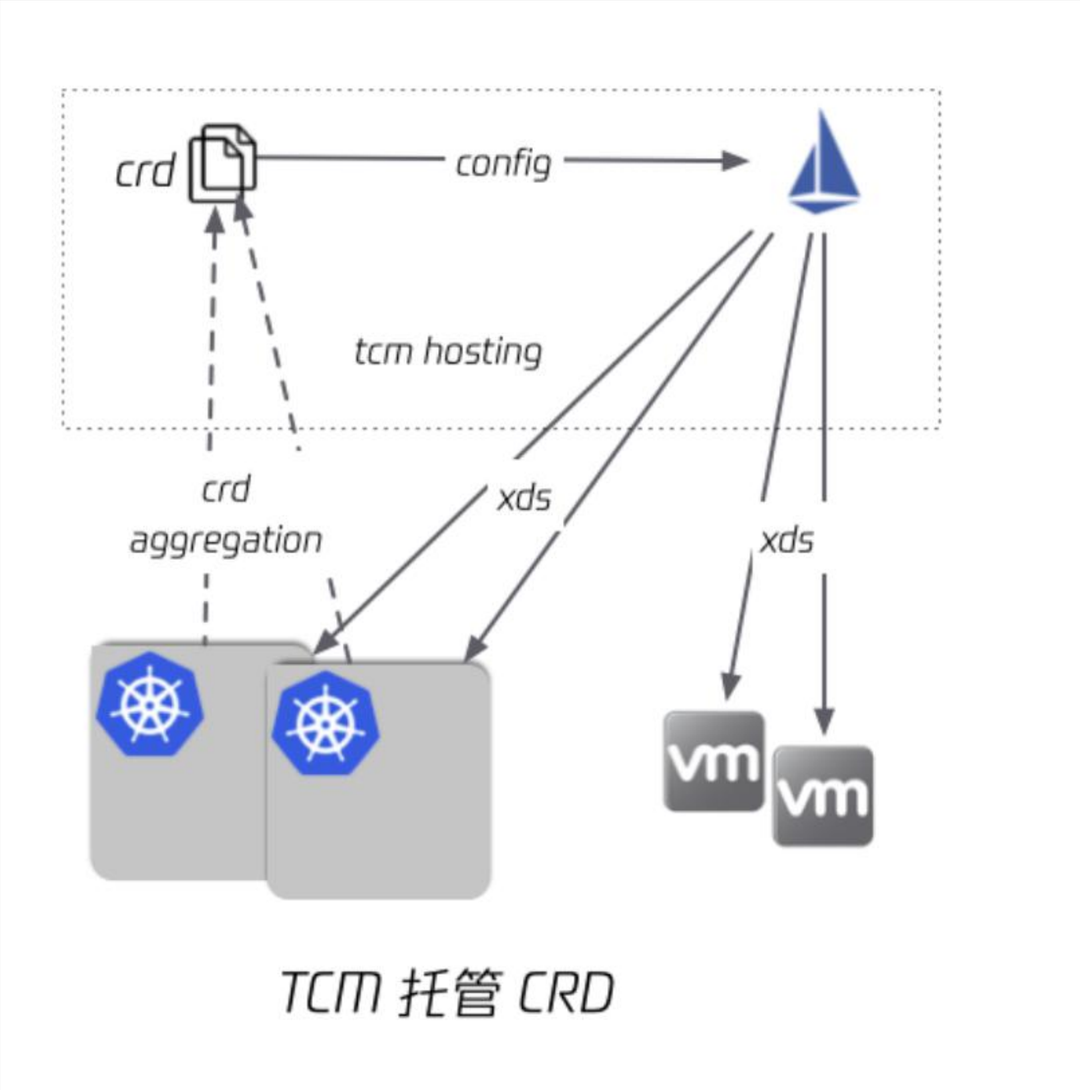
最佳实践–Istio CRD 中心托管

现有 Istio CRD 机制缺陷



- CRD 单点存储
- 主集群无法变更
- 子集群操作 CRD 困难
- 不支持空 mesh 持久化 CRD

多集群 Istio CRD 托管架构




```
~ %
~ % kubectl -n istio-system get pod
NAME                                    READY   STATUS
istio-ingressgateway-685944f55d-ssrnh   1/1     Running
~ %
~ % kubectl get crd | grep istio
~ %
~ % kubectl -nbase get virtualservice
NAME    GATEWAYS          HOSTS   AGE
mall    [mall-gateway]    [*]     80s
~ %
```

tetrate

Istio 避坑指南

Istio 避坑指南

tetrate

1. 业务代码应该对 HTTP Header 大小写不敏感
2. 保证业务接口幂等
3. 避免使用 istio 占用端口， 如 15090, 15021 等
4. 通过 service port name 前缀，显式指定端口协议
5. 保证一致的容器启停顺序
6. 停止执行 docker container 清理工作
7. 应用需要监听 0.0.0.0 （istio 1.10 之前）
8. 流控规则下发遵循 make before break

tetrate

Thanks

Tetrate 中国

云原生社区