

Service Mesh Meetup #3 深圳站

# Istio控制平面组件原理解析

朱经惠

2018.08.25

# 关于我

- 朱经惠，ETC车宝平台工程师。
- 喜欢开源，个人开源项目” Jaeger PHP Client”。
- 喜欢研究源码，对NSQ，Jaeger，Istio（控制平面）等go语言开源项目进行过研究。
- 除了代码还喜欢爬山和第二天睡醒后全身酸疼的感觉。

# 目录

## galley

Galley provides configuration management services for Istio.

## mixc

Utility to trigger direct calls to Mixer's API.

## ✓ pilot-agent

Istio Pilot agent.

## ✓ istio\_ca

Istio Certificate Authority (CA).

## ✓ mixs

Mixer is Istio's abstraction on top of infrastructure backends.

## ✓ pilot-discovery

Istio Pilot.

## istioctl

Istio control interface.

## node\_agent

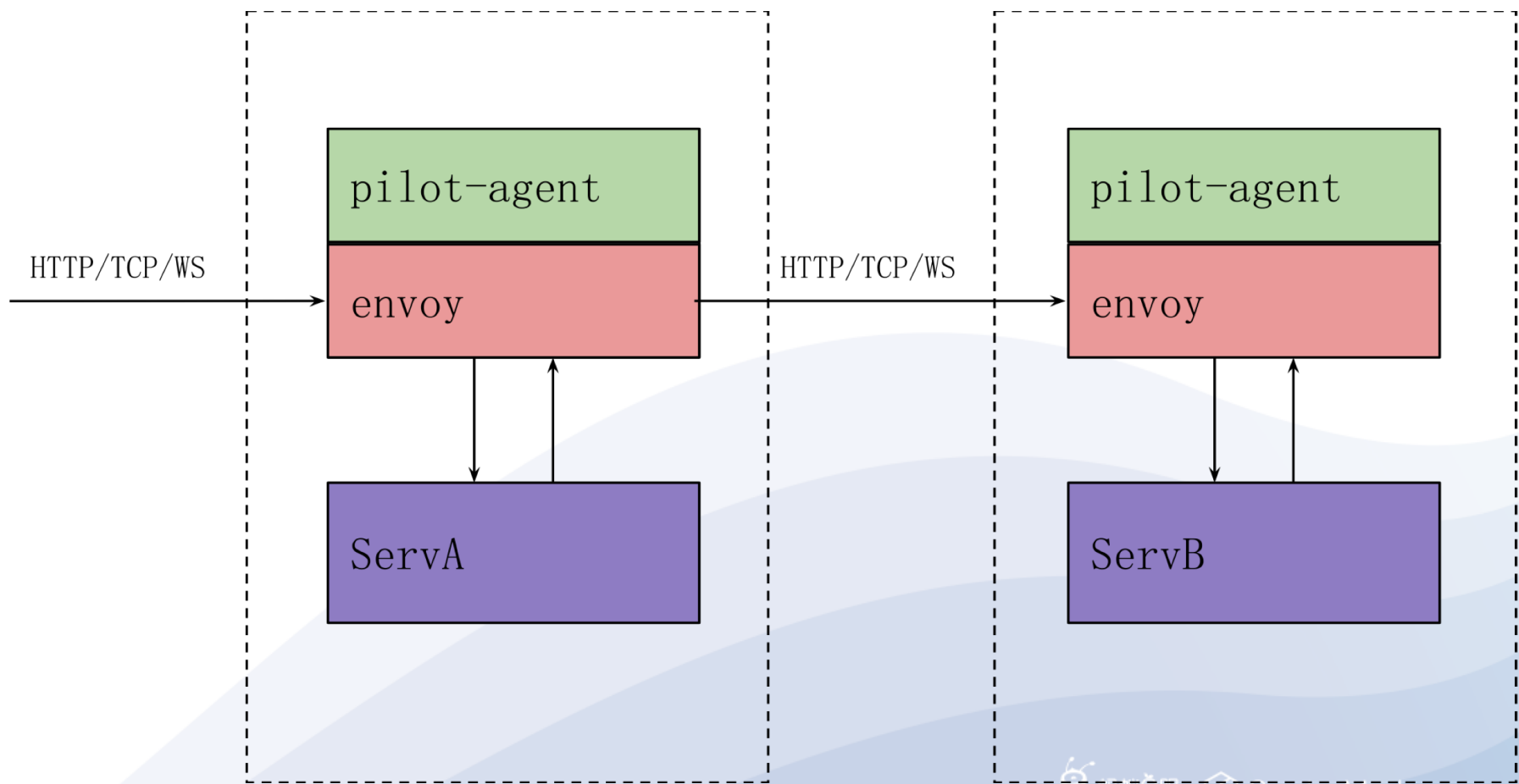
Istio security per-node agent.

## sidecar-injector

Kubernetes webhook for automatic Istio sidecar injection.

# Pilot-Agent——管理生命周期（PA）

- ◆启动envoy
- ◆热重启envoy
- ◆监控envoy
- ◆优雅关闭envoy

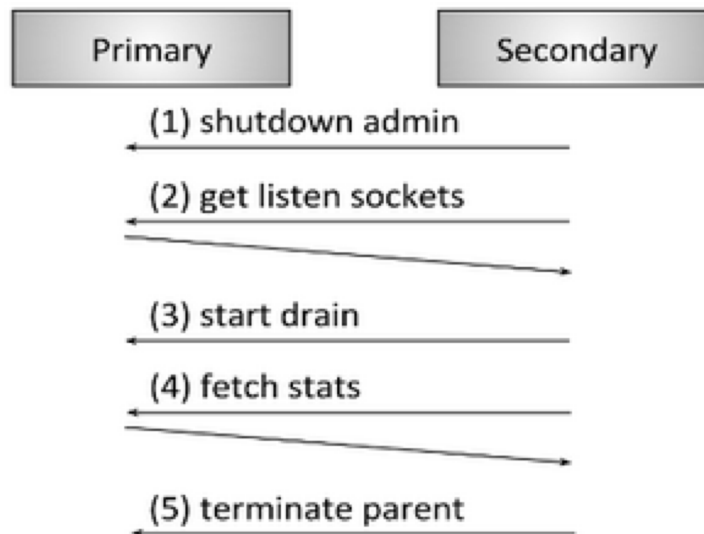
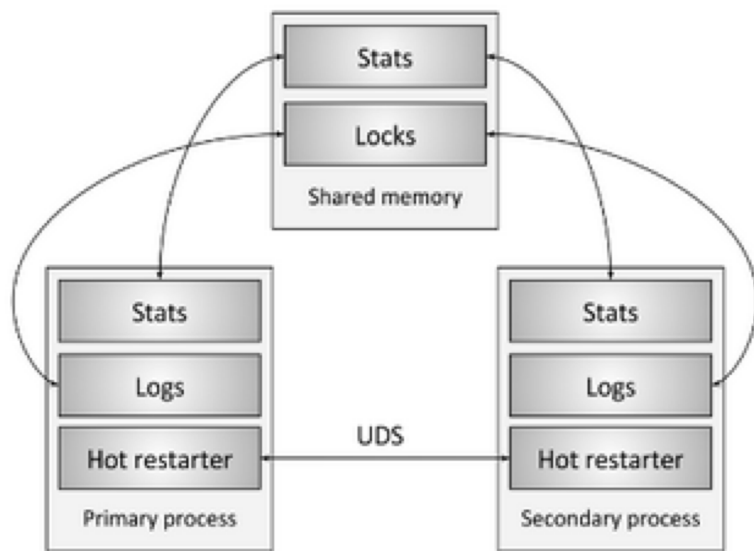


# 启动envoy

- ✓ 监听/etc/certs目录
- ✓ 生成envoy静态配置文件envoy-rev0.json
- ✓ 通过exec.Command启动 envoy并监听状态

- [文件配置文档](#)
- [启动参数文档](#)

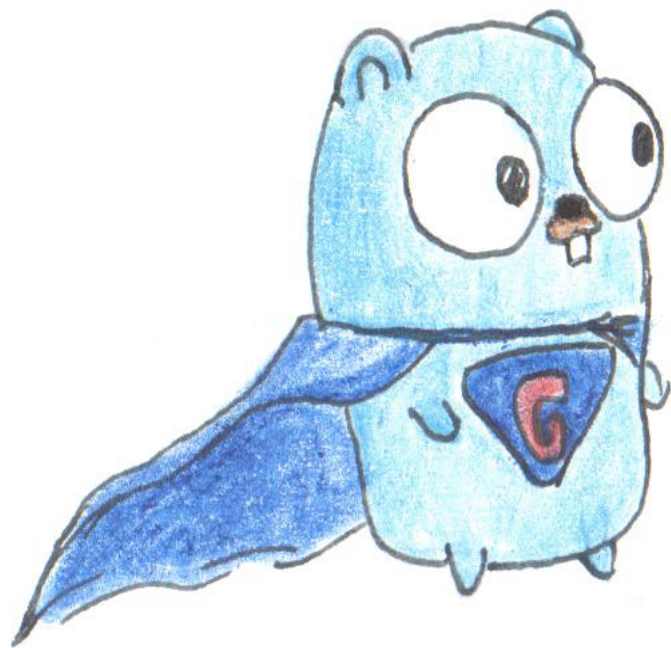
# 热重启envoy



# 热重启涉及以下步骤

- Pilot-Agent只是负责启动S，其他步骤由envoy完成。
  - 1. 启动另外一个S进程（Secondary process）
  - 2. S通知P（Primary process）关闭其管理的端口，由S接管
  - 3. S加载配置，开始绑定listen sockets，在这期间使用UDS从P获取合适的listen sockets
  - 4. S初始化成功，通知P停止监听新的链接并优雅关闭未完成的工作
  - 5. 在P优雅关闭过程中，S会从共享内存中获取stats
  - 5. 到了时间S通知P自行关闭
  - 6. S升级为P
- 
- 官方博客：[Envoy hot restart](#)

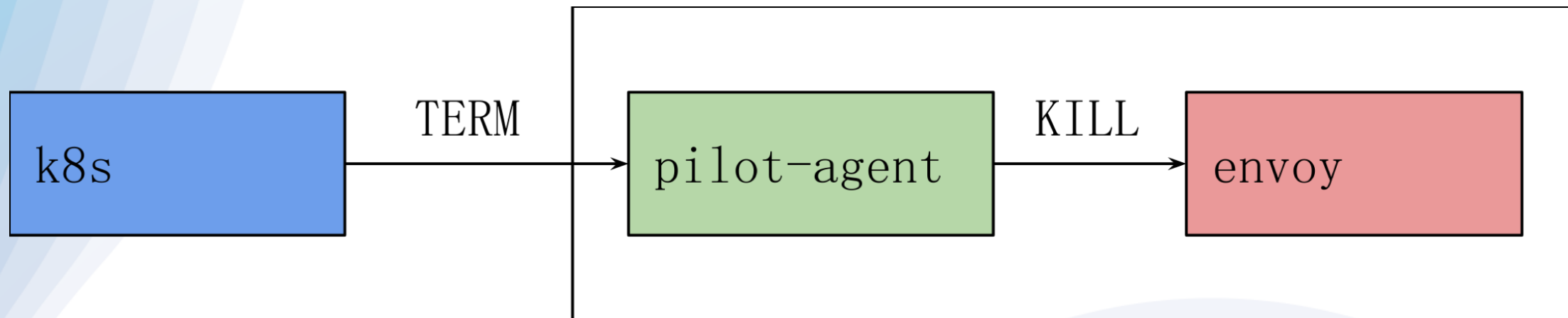
# 什么时候会进行热重启？



# 监控envoy

- ✓ 获取非正常退出状态
- ✓ 抢救机制触发
- ✓ 抢救令牌减少一个（总共10个）
- ✓ 在 $2^{(n-1)} * 200$ 毫秒后执行（为什么不立即执行）
- ✓ 失败再次触发抢救机制
- ✓ 10个令牌用完，没有抢救成功，放弃退出

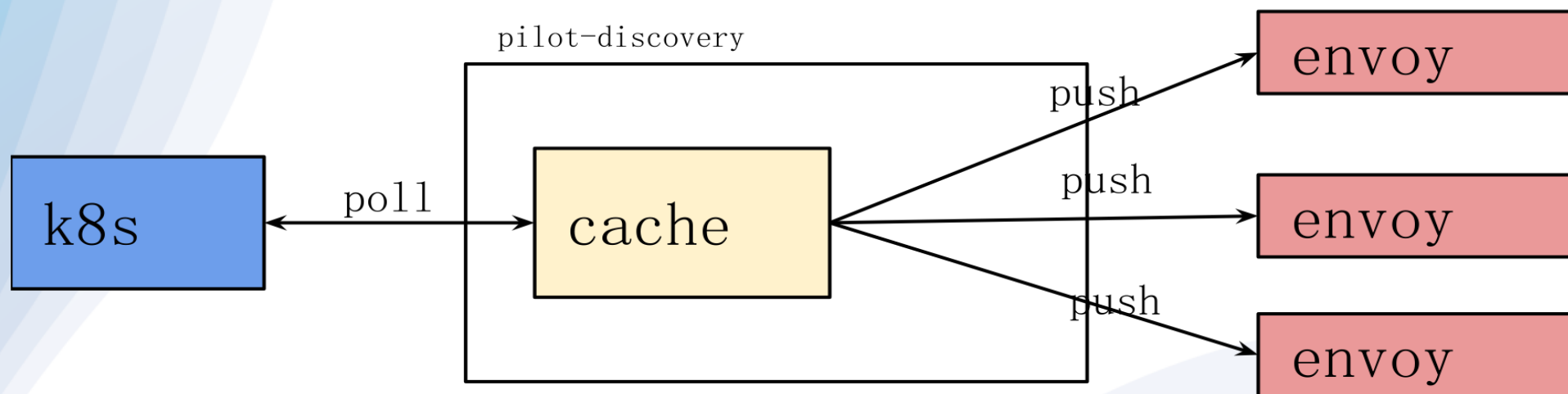
# 优雅关闭envoy



- ✓ K8s发送SIGTERM信号让容器优雅关闭
- ✓ Pilot-Agent接收信号通过context关闭子服务，发送SIGKILL关闭envoy
- ✓ Envoy不支持优雅关闭，需要通过金丝雀或蓝绿部署方式实现

Envoy优雅关闭实现方式讨论: [#3307](#) [#2920](#)

# Pilot-Discovery——配置中心（PD）



- ◆ v1版本和v2版本之间的区别
- ◆ 建立缓存配置
- ◆ 触发配置生效方式

# v1版本和v2版本之间的区别

## V1

HTTP1 REST

JSON/YAML 弱类型

轮询

SDS/CDS/RDS/LDS

奠定控制平面基础

## V2

HTTP2 GRPC

Proto3 强类型

Push

SDS/CDS/RDS/LDS/HDS/ADS/KDS

和Google强强联手

# 缓存Istio和k8s配置

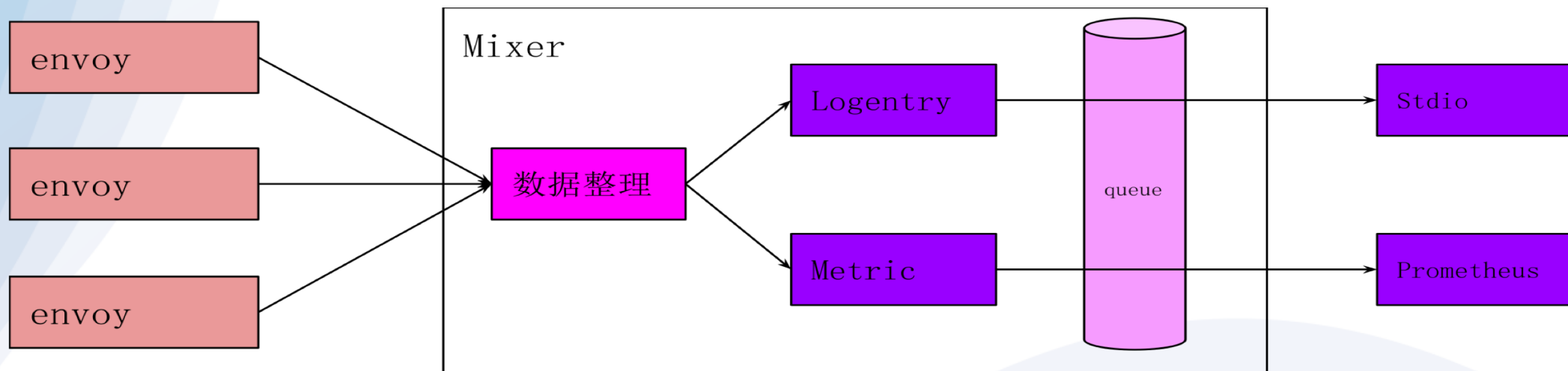
- ✓ 一个小型的非持久性key/value数据库
- ✓ 借助k8s.io/client-go建立缓存
- ✓ 缓存Istio: route-rule, virtual-service, gateway等
- ✓ 缓存k8s: node, Service, Endpoints等

# 触发配置生效方式

V2通过GRPC双向流，主动推送配置给envoy：

- ✓事件触发
  - 当配置有增/删/改事件
- ✓定时触发
  - 可配置环境变量 V2\_REFRESH，定时推送配置

# Mixer——遥测报告



- ◆上报的原始数据
- ◆异步Flush给Adapter

- ◆转换成属性词汇
- ◆问题讨论

# 属性词汇

| Name             | Type                | Description                                                                                            | Kubernetes Example                                      |
|------------------|---------------------|--------------------------------------------------------------------------------------------------------|---------------------------------------------------------|
| source.id        | string              | Platform-specific unique identifier for the source workload instance.                                  | kubernetes://redis-master-2353460263-lecey.my-namespace |
| source.ip        | ip_address          | Source workload instance IP address.                                                                   | 10.0.0.117                                              |
| source.labels    | map[string, string] | A map of key-value pairs attached to the source instance.                                              | version => v1                                           |
| destination.port | int64               | The recipient port on the server IP address.                                                           | 8080                                                    |
| request.time     | timestamp           | The timestamp when the destination receives the request. This should be equivalent to Firebase “now” . |                                                         |

# 上报的原始数据

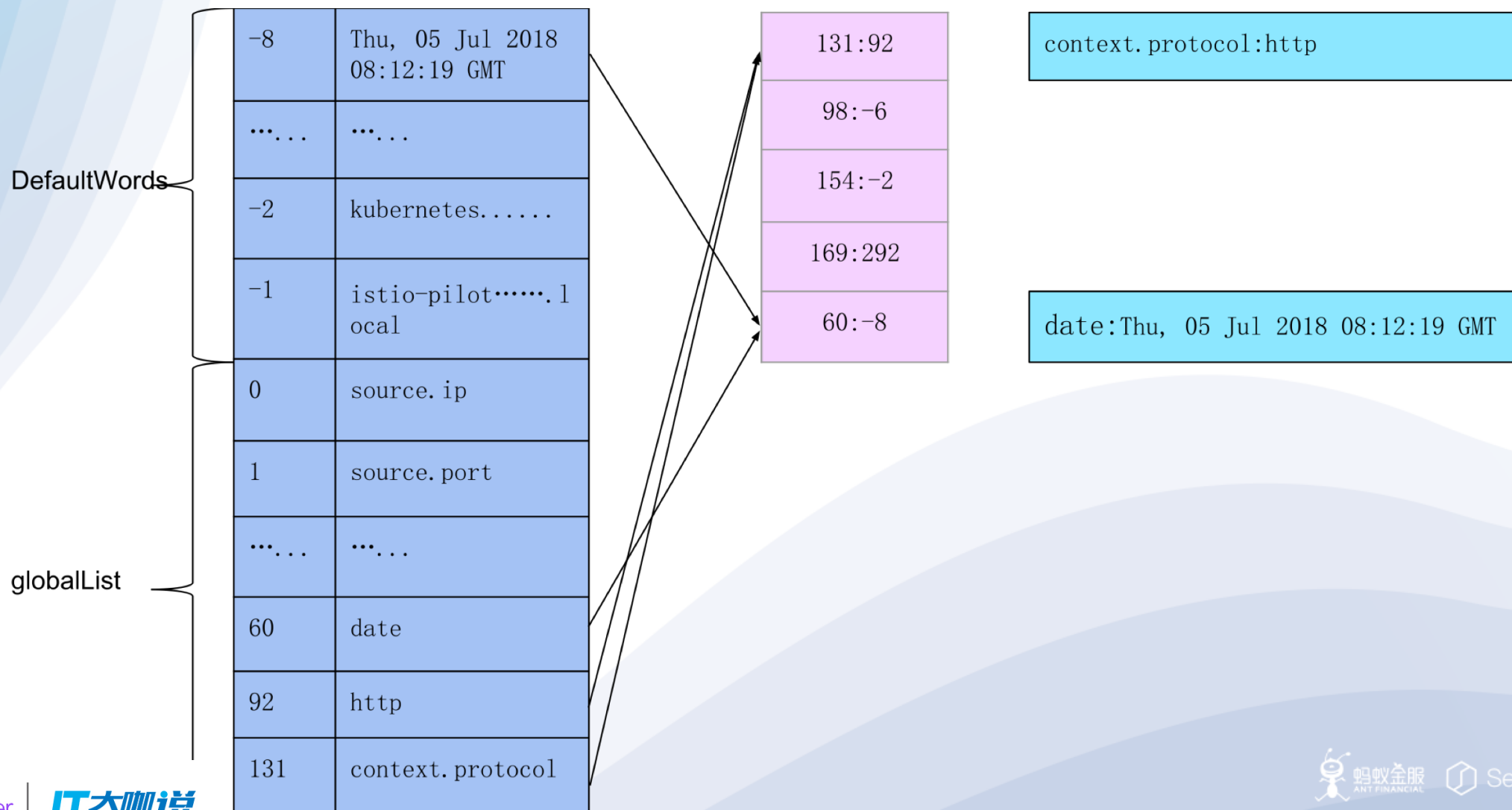
## ✓req.Attributes:

- "strings":{"131":92,"152":-1,"154":-2,"17":-7,"18":-4,"19":90,"22":92
- "int64s":{"1":33314,"151":8080,"169":292,"170":918,"23":0,"27":780,"30":200}
- "bools":{"177":false}
- "timestamps":{"24":"2018-07-05T08:12:20.125365976Z"},
- "durations":{"29":426699}
- "bytes":{"0":"rBQDuw==","150":"AAAAAAAAAAAAAP//rBQDqg=="}

## ✓req.DefaultWords :

- ["istio-pilot.istio-system.svc.cluster.local",
- "kubernetes://istio-pilot-8696f764dd-fqxtg.istio-system",
- "3a7a649f-4eeb-4d70-972c-ad2d43a680af",
- "172.00.00.000","Thu, 05 Jul 2018 08:12:19 GMT","780",
- "bc1f172f-b8e3-4ec0-a070-f2f6de38a24f","718"]

# 转换成属性词汇



# 异步Flush到Adapter

- ✓通过Template润色数据
- ✓使用Go的协程异步Flush到Adapter
- ✓Adapter展示数据
- ✓响应envoy数据处理完成

# 问题讨论

## Report流程回顾

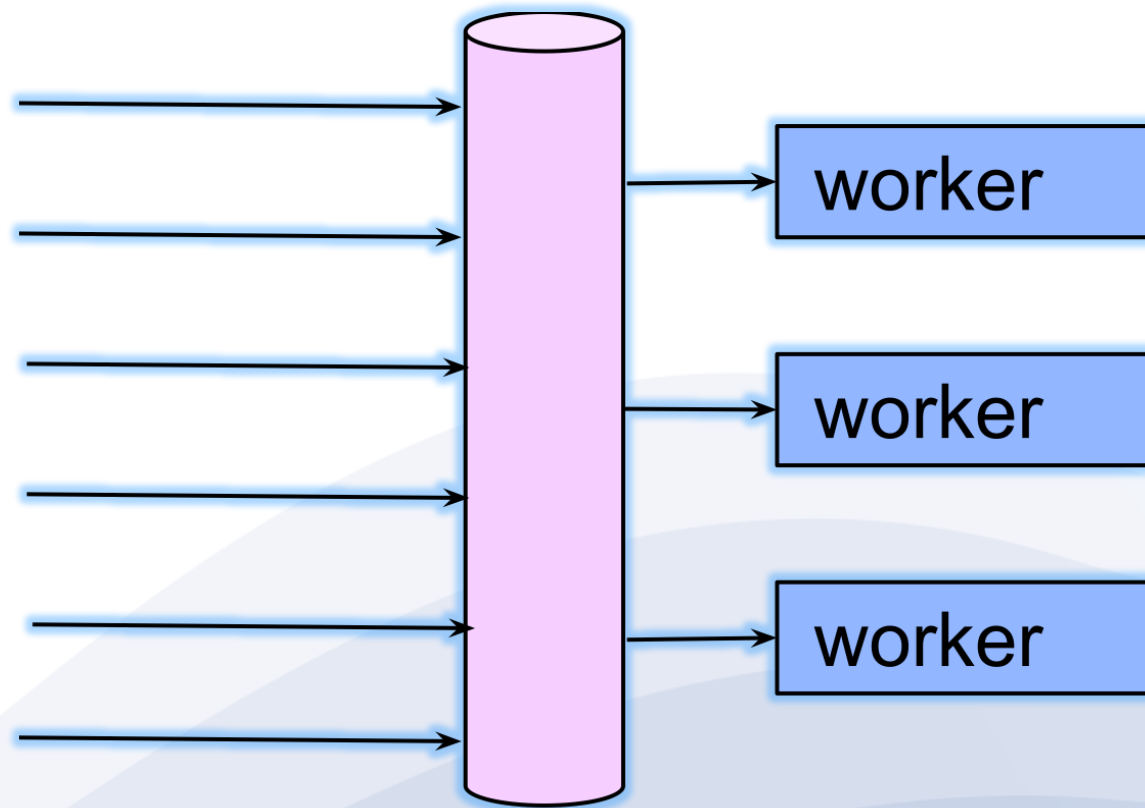
- ✓ Envoy异步批量发送数据给Mixer
- ✓ Mixer使用协程池处理Adapter
- ✓ 处理完成所有Adapter才响应Envoy

## 疑问

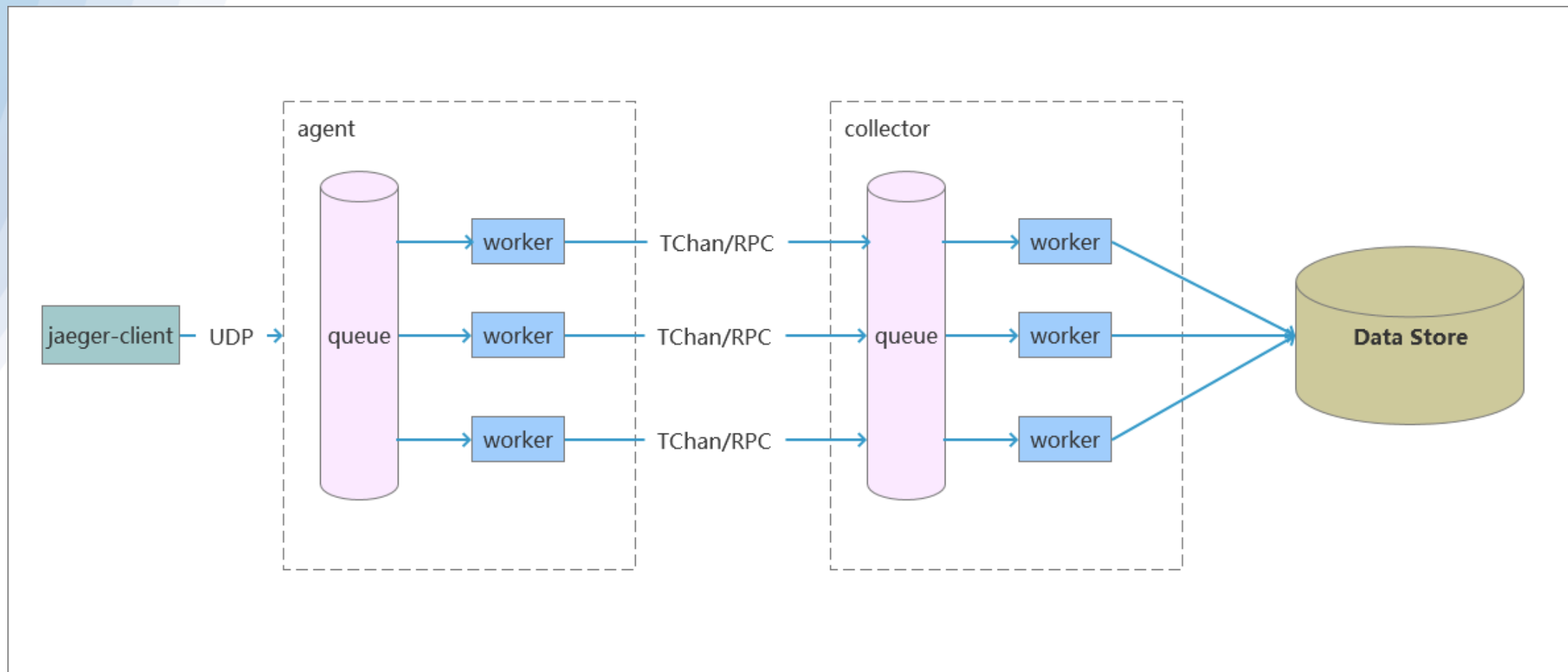
协程池堵塞是否会影响envoy性能？

# Mixer协程池

- ✓ 初始化一定量worker（协程）
- ✓ 监听同一队列
- ✓ 任务放入队列
- ✓ Worker处理任务



# Jaeger架构设计

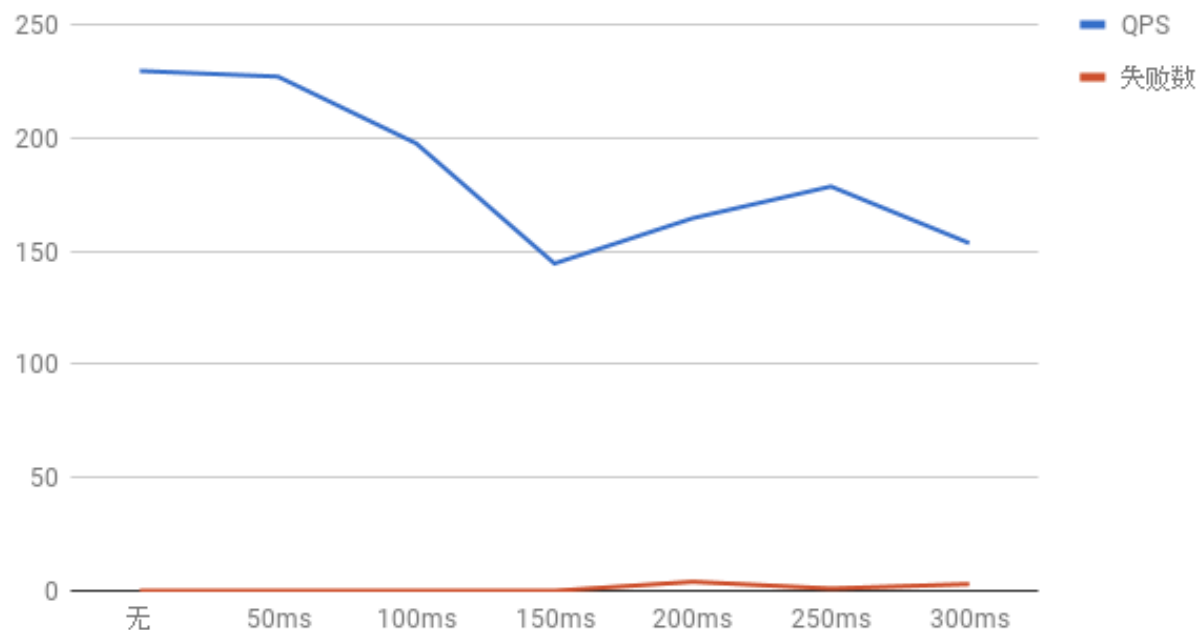


# Mixer阻塞对envoy的影响

压测环境:

- ✓ 模拟接口延迟响应
- ✓ 使用hey压力工具
- ✓ 相同压力
- ✓ 先用hey进行预热
- ✓ 从10份数据中取中位数

mixer阻塞影响



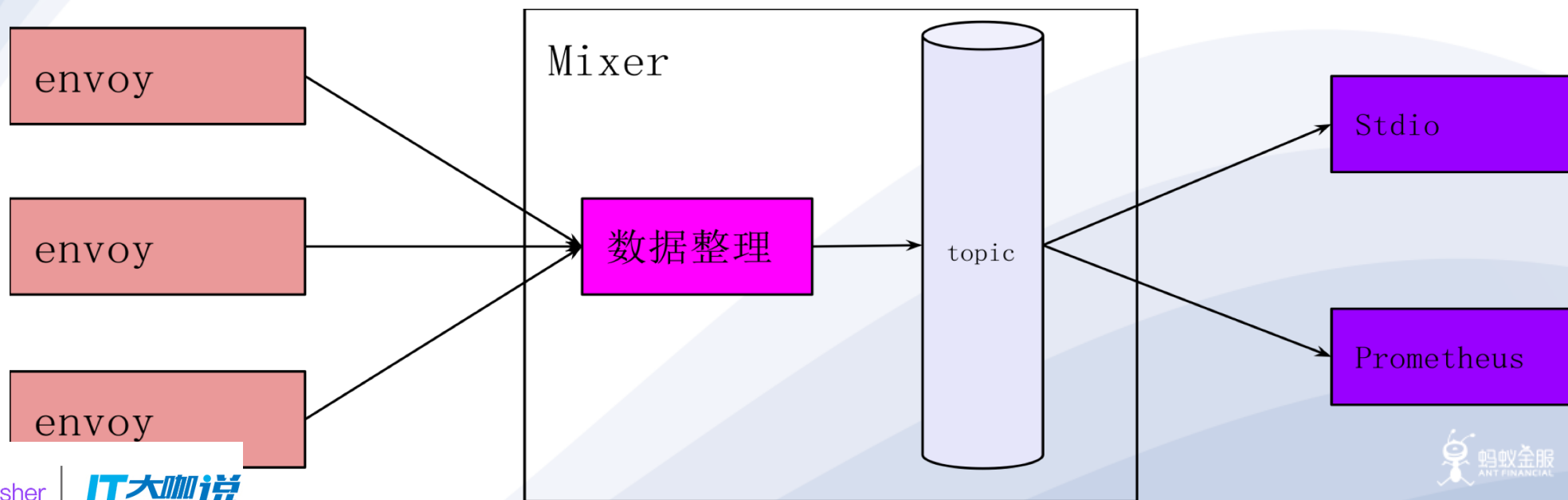
# 解决方案

## ✓方案一

- 业务性能和日志之间的选择，出现阻塞，丢弃日志保性能

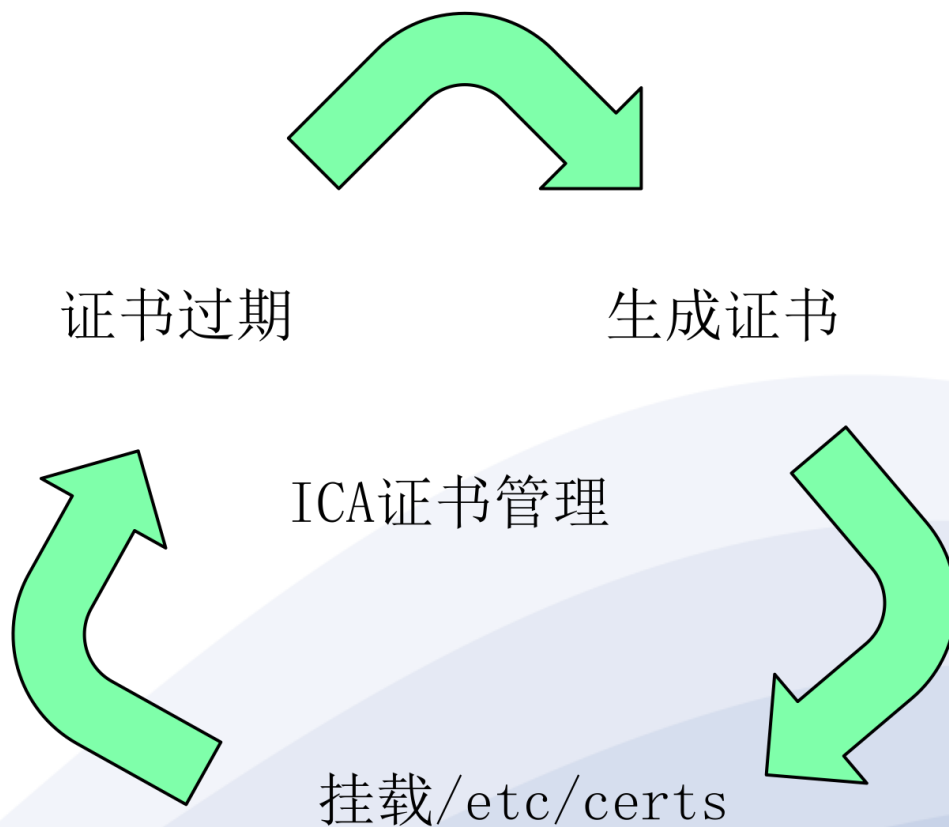
## ✓方案二

- 使用主题订阅模式，减少阻塞问题



# Istio\_Ca——安全证书管理（ICA）

- ◆证书生成
- ◆证书挂载
- ◆证书过期



# 证书生成

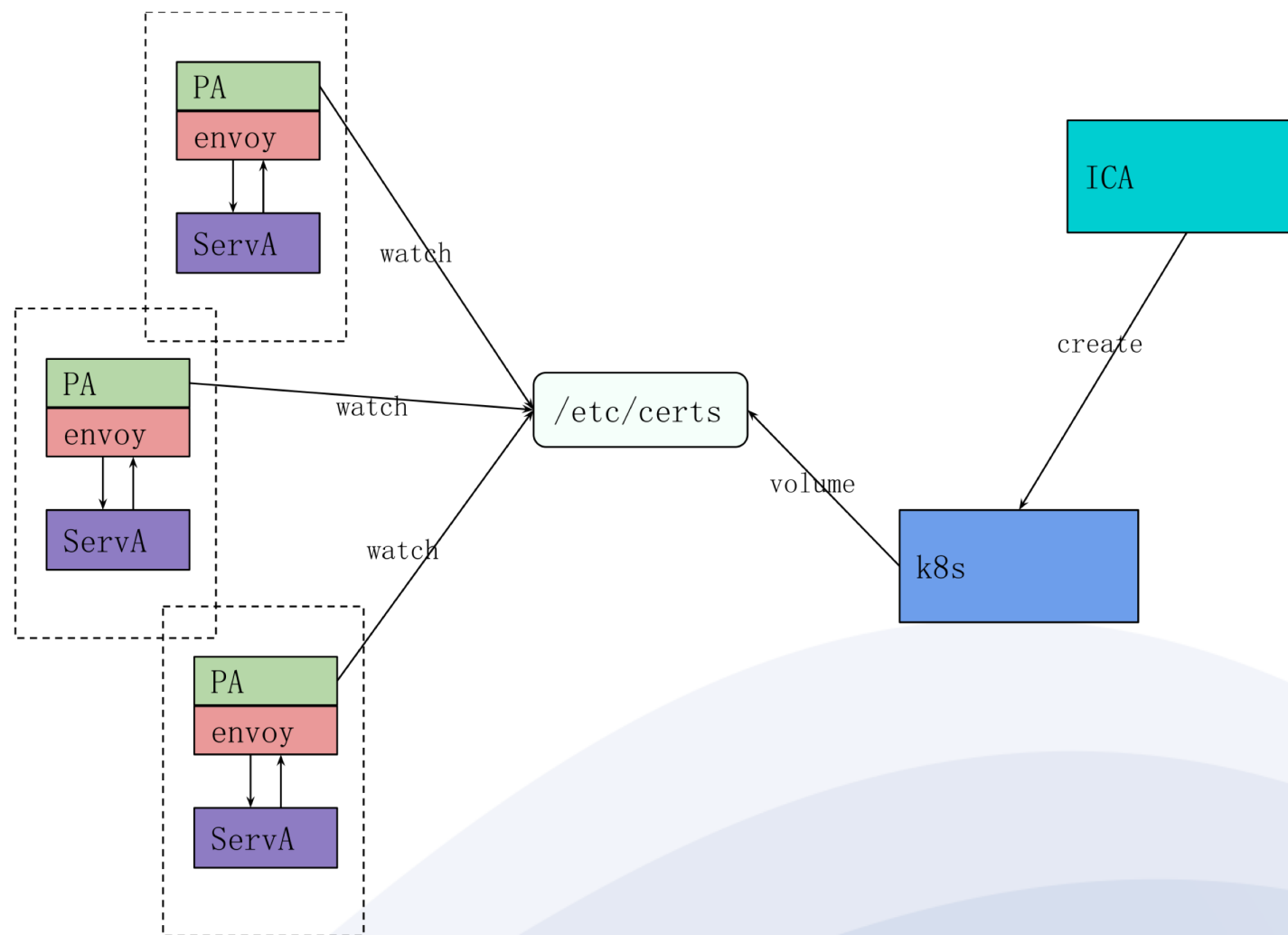
- ✓生成root-cert.pem
- ✓生成cert-chain.pem
- ✓生成key.pem

# 证书挂载

- ✓ ICA以Name为istio.default在k8s创建Secrets对象
- ✓ 应用服务获取Secrets对象证书，并挂载到/etc/certs
  - volumeMounts:
    - - mountPath: /etc/certs/
    - name: istio-certs
    - readOnly: true
  - volumes:
    - - name: istio-certs
    - secret:
      - optional: true
      - secretName: istio.default

# 证书过期

- ✓ root-cert.pem 实际有效期1年，没有找到更新方式，手动更新？
- ✓ cert-chain.pem 和 key.pem 实际有效期90天，程序控制有效期45天
- ✓ 证书过期会被重新生成并挂载到/etc/certs
- ✓ 触发envoy热启动



✓ 方案一:

- 把重新生成证书时间改为凌晨



金融级分布式架构



ServiceMesh

欢迎关注微信公众号，获取更多技术干货！

<http://www.servicemesh.com>