

## 背景

---

目前我行已有多数系统转而使用微服务架构，即把一个单体应用拆分成若干个小型的服务，协同完成系统功能的一种架构模式，在系统架构层面进行解耦合，将一个复杂问题拆分成若干个简单问题，开发、维护、部署的难度就降低了很多，吞吐量和稳定性大大增加，且可以自主选择合适的技术框架，提高了项目开发的灵活性。然而，在转型过程中，由于银行业系统的多样性、复杂性，全部加入微服务行列是不现实的，新老系统共存是一种最为常见的现象。我行C与Java并存的系统并不在少数，而共存系统间的治理、运维等非常困难，很难做到统一维护、治理、监控等，在过度时期往往需要多个团队分而管之，维护难度很大。除此之外还会面临过于绑定技术栈，在面对这种异构系统时，需要花费大量精力来进行代码的改造，且在改造过程中会面临各种问题。

对于上述问题，现有的微服务架构无法规避，在业界技术的不断探索中：新一代微服务架构--服务网格应运而生。

## Istio同业使用情况

---

中国工商银行从2019年开始服务网格技术的预研工作，通过对服务网格技术深入研究和实践后，于2021年建设了服务网格平台。服务网格与现有微服务架构融合发展，助力工行应用架构向分布式、服务化转型，承载了未来开放平台核心银行系统。工行服务网格目前已完结多言语、异构技能、边缘场景的事务试点，基本论证服务网格在流量管控、体系扩展性的优势，具有下沉服务管理才能到根底设施层，高度解耦中间件与事务体系的可行性，是目前金融同业中最成熟的实践之一。

中国光大银行基于服务网格模式应用服务框架打造了企业分布式服务平台项目，目的是通过服务网格技术方向下的应用服务框架，避免现有集中式企业应用集成架构的中心交换服务瓶颈风险，提供服务化应用注册、发现及通讯的基础功能，形成企业级去中心化服务框架标准，支持企业为推进服务化转型所需的服务治理能力。项目采用服务网格技术领域事实标准的Istio开源框架作为项目分布式服务框架原型，开展集成开发。该项目于2018年启动，2019年投产，2020年正式进入推广阶段，截至目前已有86个应用服务运行在该项目支撑的分布式服务体系，覆盖范围正在快速增长。

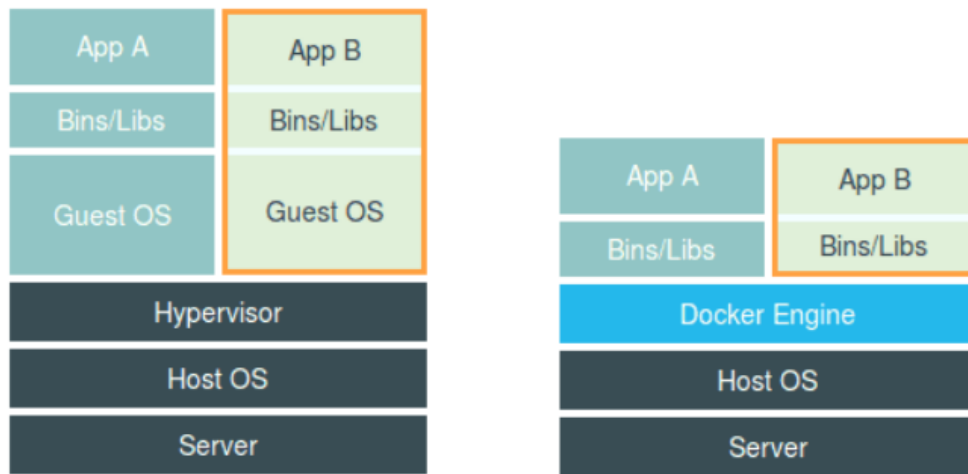
中国民生银行基于多年SOA、分布式、微服务和云原生研发领域的经验积累和沉淀，已打造了技术栈统一、组件丰富、平台健壮、弹性扩展的云原生技术平台和能力体系，推动应用架构和技术架构持续转型和科技治理、赋能业务交付方面的能力不断提升，最近也正在进行Istio服务网格相关技术的研究和探索。

## Istio技术介绍

---

### Docker浅析

#### 虚拟机与docker对比图



通过图片可以明显看到两边的差异在Hypervisor和Docker Engine。

**Hypervisor**通过硬件虚拟化功能，模拟出了运行一个操作系统需要的各种硬件，比如 CPU、内存、I/O 设备等等。然后，它在这些虚拟的硬件上安装了一个新的操作系统，即 Guest OS。

**Docker Engine**在Linux系统中可以理解为一个进程，没错每一个容器都是一个进程。docker通过 **Namespace 配置，配置Cgroups 参数，Change Root**调整根目录来实现一个容器。

**Namespace**：可以简单理解为访问权限控制。举个例子如果Linux系统为我们的软件研发中心，那么诺德中心、总部基地等为一个个进程，我们为每一个进程启用Namespace可以理解为禁止办公场地之间人员互相访问，这样诺德场地的人就只能调用诺德的资源，由此实现了资源的隔离。

**Cgroups**：通过Namespace我们知道容器之间是相互隔离的互相没有影响，但是容器就是Linux的一个进程，如果某一个进程疯狂吞噬操作系统的资源，可能会造其他容器无法运行。Cgroups是 Control Group。它最主要的作用，就是限制一个进程组能够使用的资源上限，包括 CPU、内存、磁盘、网络带宽等等。

**rootfs**：作为一个沙箱环境，资源应该也是隔离的，每当创建一个新容器时，希望容器进程看到的文件系统就是一个独立的隔离环境，而不是继承自宿主机的文件系统。通过mount namespace（特殊的 namespace，用于挂载）和容器镜像（rootfs）两种技术实现，mount namespace将根目录挂载到指定目录，rootfs用于恢复镜像内容。

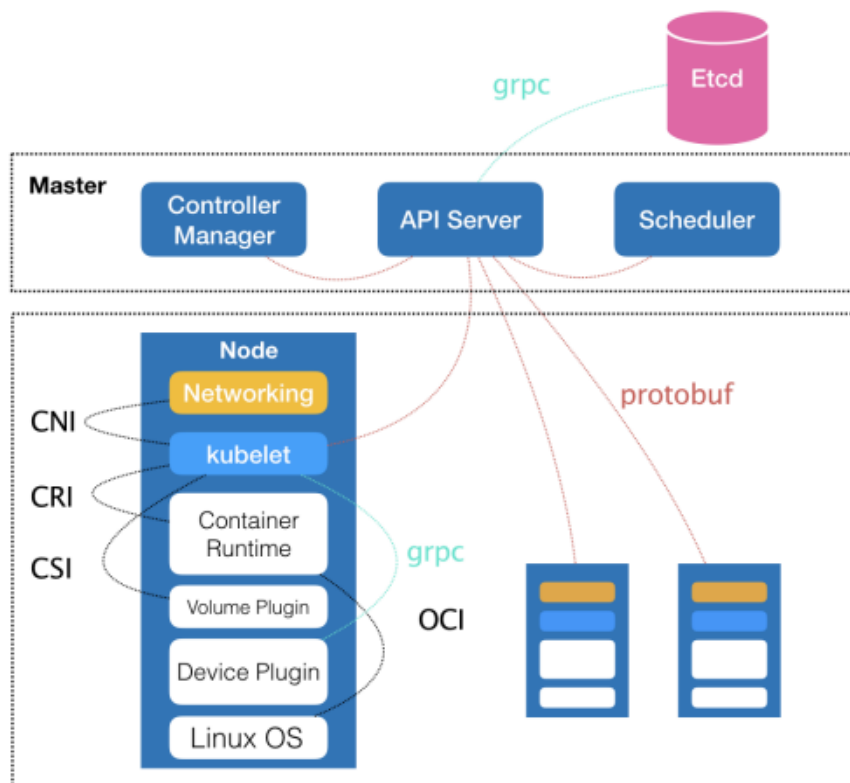
**问题：**

- 1.为什么通过docker启动的服务进程号都是1
- 2.为什么我们一个容器内只启动一个程序
- 3.为什么要使用docker

## Kubernetes浅析

### 为什么需要Kubernetes

用过docker的人都知道docker启动命令相当粗暴，比如说我们要启动一个容器就是docker run ...，从字面意思非常好理解。但是我们思考一下docker除了带给我们沙箱的环境和镜像的便利，其他的鞭长莫及。现在微服务动辄几百台节点，可能工作量并没有变化。2014年Kubernetes（下面简称k8s）发布第一个版本，集容器编排、调度、管理与一体。如果docker的交互方式是面向过程，k8s就是面向对象。那么k8s如何实现的呢。



k8s之所以可以管理docker容器，重点在于图中的CRI，它屏蔽了具体的容器实现，提供了一层抽象来管理容器。在用户输入一条指令后，API Server接收指令解析后存储到etcd，Scheduler监听到了变更后选取容器需要去往的node节点等信息存储到etcd中，Controller Manager通过监听数据变化，调用API Server发送请求到kubelet，kubelet执行具体和容器、操作系统的交互。我们只需要告诉k8s我们需要的，k8s会协调资源帮助我们实现，这就是“声明式API”。

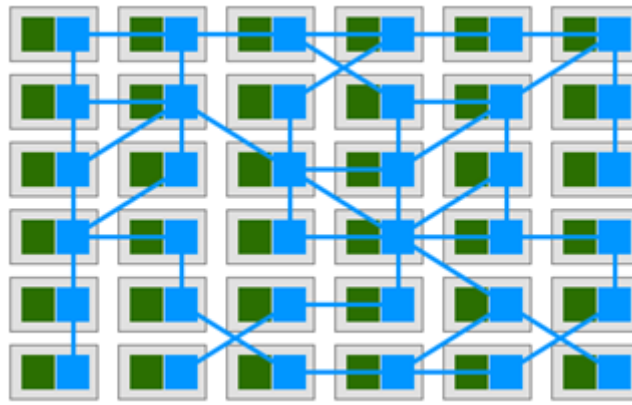
#### 问题：

- 1.k8s的“声明式API”的好处是什么
- 2.为什么要使用k8s

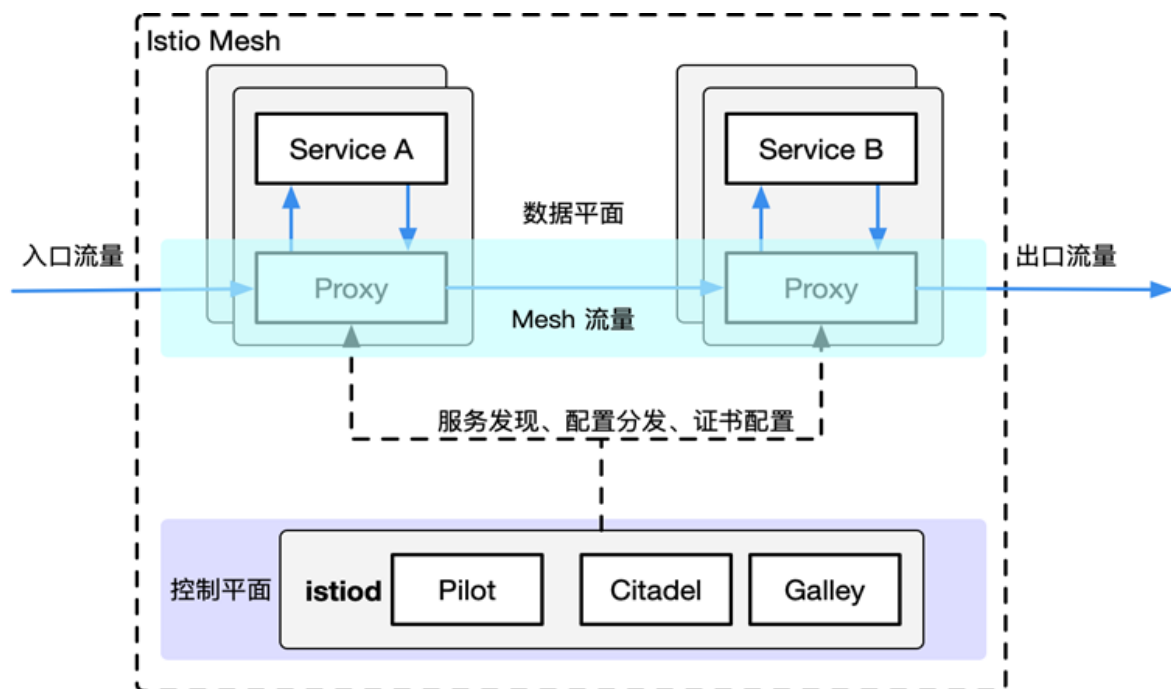
## Istio是什么

服务网格（Service Mesh）是一个专门处理服务通讯的基础设施层。它的职责是在由云原生应用组成服务的复杂拓扑结构下进行可靠的请求传送。其从总体架构上来讲比较简单：一组服务的用户代理，加上一组任务管理组件组成。

1. 用户代理在服务网格中被称为数据层或数据平面（data plane），直接处理入站和出站数据包，包括转发、路由、健康检查、负载均衡、认证、鉴权、产生监控数据等。
2. 管理组件被称为控制层或控制平面（control plane），负责与控制平面中的代理进行通信、下发策略和配置。



相较于传统微服务架构，服务网格不再需要服务网关、注册中心、负载均衡和流控等传统微服务架构所需的组件，以更轻量级的架构支持微服务间的网络通信和服务治理。其中Istio是服务网格的典型实现，目前需要依赖于k8s环境。以下是对Istio的架构及组件的简单介绍。



- Istio选择Envoy作为代理，其本质是一个为面向服务的架构而设计的代理和通信总线。
- Pilot是Istio实现流量管理的核心组件，它主要的作用是配置和管理 Envoy代理。
- Citadel是与安全相关的组件，主要负责密钥和证书的管理。
- Galley作为独立组件承担Istio的配置管理工作，负责配置的获取、处理和分发。

所以对于Istio的能力我们可以简单的用四个字概括（个人看法）：**流量治理**

## 为什么选择Istio

市面上流量治理的组件层出不穷，大众化的如hystrix和sentinel（java语言），每个组件都有自己特色，适用范围广API简单，那么为什么要学习转用Istio呢？Istio为什么成为划时代的组件呢？

### 1.性能好，功能强大

上面介绍了Istio的数据平面组件是**Envoy**，控制平面（Istiod）主要由Pilot、Citadel、Galley组成（精简后），由于Istiod的功能很容易理解所以我们重点说一下Envoy，Envoy是cncf基金会第三个毕业的项目，前两个是k8s和Prometheus。Envoy出身名门性能也非常优秀。其优势：

**L3/L4 过滤器架构：** Envoy 的核心是一个 L3/L4 网络代理。可插入的过滤器链机制允许编写过滤器来执行不同的 TCP/UDP 代理任务并插入到主服务器中。已经编写了过滤器以支持各种任务，例如原始TCP代理、UDP代理、HTTP代理、TLS 客户端证书身份验证、Redis、MongoDB、Postgres等。

**HTTP L7 过滤器架构：** HTTP 是现代应用程序架构的关键组件，Envoy支持额外的 HTTP L7 过滤器层。HTTP 过滤器可以插入 HTTP 连接管理子系统，执行不同的任务，如缓冲、速率限制、路由/转发、嗅探亚马逊的DynamoDB等。

**一流的 HTTP/2 支持：** 在 HTTP 模式下运行时，Envoy支持HTTP/1.1 和 HTTP/2。Envoy 可以作为透明的 HTTP/1.1 到 HTTP/2 双向代理运行。这意味着可以桥接 HTTP/1.1 和 HTTP/2 客户端和目标服务器的任意组合。推荐的服务到服务配置在所有 Envoy 之间使用 HTTP/2 来创建一个持久连接网格，请求和响应可以在该网格上多路复用。

**HTTP/3 支持（目前处于 alpha 阶段）：** 从 1.19.0 开始，Envoy 现在支持 HTTP/3 上游和下游，并在 HTTP/1.1、HTTP/2 和 HTTP/3 的任意组合之间进行双向转换。

**HTTP L7 路由：** 在 HTTP 模式下运行时，Envoy 支持 路由子系统，该子系统能够根据路径、权限、内容类型、运行时值等路由和重定向请求。当使用 Envoy 作为前端/边缘时，此功能最有用代理，但在构建服务网格服务时也会被利用。

**gRPC 支持：** gRPC是来自 Google 的 RPC 框架，它使用 HTTP/2 或更高版本作为底层多路复用传输。Envoy支持用作 gRPC 请求和响应的路由和负载平衡底层所需的所有 HTTP/2 功能。这两个系统非常互补。

**服务发现和动态配置：** Envoy 可选择使用一组分层的 动态配置 API进行集中管理。这些层为 Envoy 提供以下动态更新：后端集群中的主机、后端集群本身、HTTP 路由、侦听套接字和加密材料。对于更简单的部署，可以通过 DNS 解析（甚至 完全跳过）来完成后端主机发现，并将更多层替换为静态配置文件。

**健康检查：** 构建 Envoy 网络的推荐方法是将服务发现视为最终一致的过程。Envoy 包括一个健康检查子系统，它可以选择性地对上游服务集群执行主动健康检查。然后，Envoy 使用服务发现和健康检查信息的结合来确定健康的负载均衡目标。Envoy 还支持通过异常值检测子系统进行被动健康检查。

**高级负载均衡：** 分布式系统中不同组件之间的负载均衡是一个复杂的问题。因为 Envoy 是一个自包含的代理而不是一个库，所以它能够在 一个地方实现高级负载平衡技术，并让任何应用程序都可以访问它们。目前，Envoy 包括对自动重试、熔断、通过外部速率限制服务进行 全局速率限制、请求阴影和异常值检测的支持。

**前端/边缘代理支持：** 在边缘使用相同的软件有很大的好处（可观察性、管理、相同的服务发现和负载平衡算法等）。Envoy 具有一个功能集，使其非常适合作为大多数现代 Web 应用程序用例的边缘代理。这包括TLS终止、HTTP/1.1 HTTP/2 和 HTTP/3支持，以及 HTTP L7路由。

**一流的 可观察性：** 如上所述，Envoy 的主要目标是使网络透明。然而，问题出现在网络层面和应用层面。Envoy 包括对所有子系统的强大统计支持。statsd（和兼容的提供者）是当前支持的统计接收器，尽管插入不同的接收器并不困难。也可以通过管理端口查看统计信息。Envoy 还支持 通过第三方提供商进行分布式跟踪。

Istio可以无缝整合Jaeger、Kiali、Prometheus、Tracing、Zipkin等，这些组件提供了Istio的调用链、监控等功能，可以选择安装来完成完整的服务监控管理功能。

## 2. 契合容器云环境

上面介绍了Istio的相关组件和功能，在性能方面固然强大，但是部署和使用的难易程度也是我们需要考量的另一个重要指标。

假设我们已经基于k8s来部署服务（没意见吧！）

在k8s中我们想实现一些流量控制，我们可以继续使用hystrix和sentinel（java语言），但是我们需要为每一个服务嵌入这些组件，然后编写对应的逻辑处理。如果用sentinel的话虽然可以拥有远程下发配置的能力，但是我们需要额外部署控制台资源，还需要为这些交互打通网络关系。最主要的是我们需要熟悉这些组件才可以应用到系统中，那如果是多个语言组成的系统呢。。。

当然我们也可以自己编写组件，和应用服务部署在同一个pod里面，然后通过iptables来代理所有的流量，然后通过编写控制台动态下发配置。这个方式和Istio完全一致，并且我们不会比它做的更好。

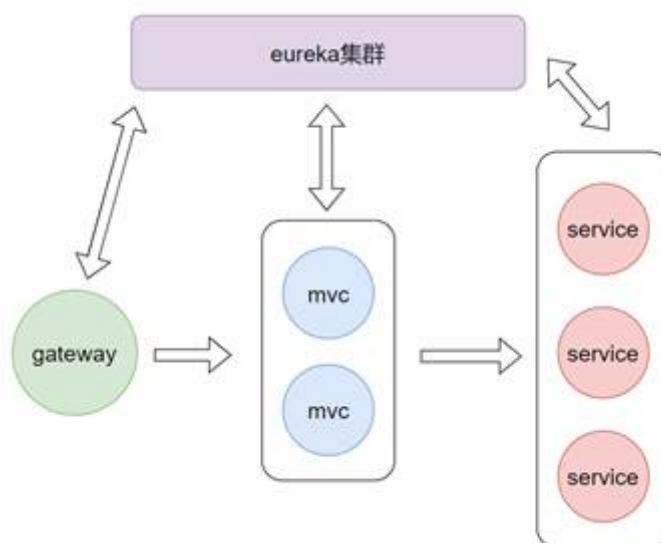
通过Istio应用可以无感知嵌入数据平面控制组件（Envoy），同时还可以代理外部应用在网格内的流量管理，动态刷新这些配置，完美的复用了k8s的能力，并且可以无需额外编码仅通过配置文件便可以实现流量的管理。

### 3.跨语言

Istio不关心语言实现，因为Istio通过iptables拦截了所有流量转发给Envoy，Envoy过滤流量后再转发给业务模块。

## 现有系统如何改造落地

以我行个人财富系统服务架构部分为例：



可以看到项目的一次请求路径简单抽象为gateway-mvc-service，所有节点都需要注册到eureka，保证了服务整体的高可用和弹性扩容的能力。通过hystrix做一些隔离、熔断策略。下面从代码层面我们看一下关键的部分：

```
@Bean
public RouteLocator routeLocator(RouteLocatorBuilder builder) {
    /**
     * 有的时候会持续调到一个节点，因为没刷新配置中心的节点
     * 注意: filter不支持传递特殊字符(例如: _ > < 等)
     */
    return builder.routes()
        .route(id: "gateway_filter_1", r -> r
            .path( ...patterns: "/filter/**")
            .filters(f -> f.addRequestParameter( param: "gateway-param", value: "gateway-filter-1:yes-i-got-it!"))
            .uri("lb://mvc"))
        .route(id: "gateway_filter_2", r -> r
            .path( ...patterns: "/hystrix/**")
            .filters(f -> f.hystrix(c -> c.setName("my-hystrix").setFallbackUri("forward:/hystrix/fail"))
                .addRequestParameter( param: "gateway-param", value: "gateway-filter-2:yes-i-got-it!"))
            .uri("lb://mvc"))
        .build();
}
```

这里gateway配置路由信息，为不同请求添加过滤器和uri等，实现路由分发。

```

@FeignClient(name = "service-user")
public interface UserService {

    @RequestMapping(value = "/user-drink")
    String drink(@RequestBody User user);
}

```

这里mvc通过feign调用service，通过ribbon实现负载均衡。

```

@RestController
public class FallBackController {

    @RequestMapping("/hystrix/fail")
    public String fail() { return "hystrix -> mvc fail"; }
}

```

利用hystrix处理异常，做兜底策略。

以docker为底层运行容器为例，需要编写dockerfile，将服务制作成镜像，然后启动镜像发布服务。以mvc为例：

```

#添加jdk8镜像依赖
FROM openjdk:8
#指定启动文件
ADD jar/mvc.jar mvc.jar
#设置端口
EXPOSE 2000

```

```

#添加jdk8镜像依赖
FROM openjdk:8
#指定启动文件
ADD jar/service.jar service.jar
#设置端口
EXPOSE 3000

```

将所有服务jar都打包成镜像后

```

docker build -f MvcDockerfile -t demomvc:v1 .
docker build -f ServiceDockerfile -t demoservice:v1 .

```

通过docker run启动便实现了全服务容器化。



## 编排容器

对个人财富项目来讲仅仅将服务进行容器化封装没有任何益处，反而还带来了运维成本加大以及服务观测性较差的问题。k8s来编排容器，网络控制来帮助我们进一步增加服务稳定性和降低运维及编码成本。

在k8s中，kube-proxy组件创建了外部入口，使我们可以直接访问到内部service（域名），service通过endpoints来寻找pod（简单理解也就是服务）。

所以如果要将个人财富系统通过Kubernetes部署，gateway和eureka这两个组件是可以去掉的，因为Kubernetes已经自带了这部分能力，但去掉的同时我们也失去了路由过滤、集中鉴权等能力（gateway），后文将会提出解决方案，仅仅通过调整一行代码就可以转成Kubernetes模式，无需再做其他配置。以下是代码需要调整的部分：

```
/**
 * @author dahua
 * @time 2022/4/21 13:37
 */
@FeignClient(name = "mvc", url = "service:3000")
public interface IstioService {

    @RequestMapping(value = "/istio-test-service")
    Map test();
}
```

去掉eureka的依赖和配置，为所有的mvc创建一个service（虚拟ip），再添加如上代码，即可获得和原来架构一样的能力。代码部分只在FeignClient注解内加上了url参数，在此处指定所有调用都会路由到authService:8080这个地址，Kubernetes会自动为其设置负载均衡策略。但如果url直接配置在代码中，每次调整地址时都需要重新打包，因此可以通过动态获取配置的方式：

```
@FeignClient(name = "mvc", url = "${service.address}")
```

通过动态获取配置，结合配置中心，经过这样特殊改造之后，可以做到不停机动态切换服务，在代码层面做到真正的动态路由。通过k8s的配置文件就可以通过对几个参数的配置，动态调整服务节点数和指定服务资源等，这些都是现有架构无法做到的。

## 创建命名空间

```
kubectl create namespace istio0322
```

k8s的配置文件示例如下，之后通过kubectl apply -f执行即可：

```
#service配置
apiVersion: v1
kind: Service
metadata:
  name: mvc-k8s-svc
  namespace: istio0322
spec:
  ports:
    #以下为端口映射设置
    #端口配置组名
    - name: http-mvc-k8s
      #service开放的端口
      port: 10001
      #容器内端口
      targetPort: 2000
```

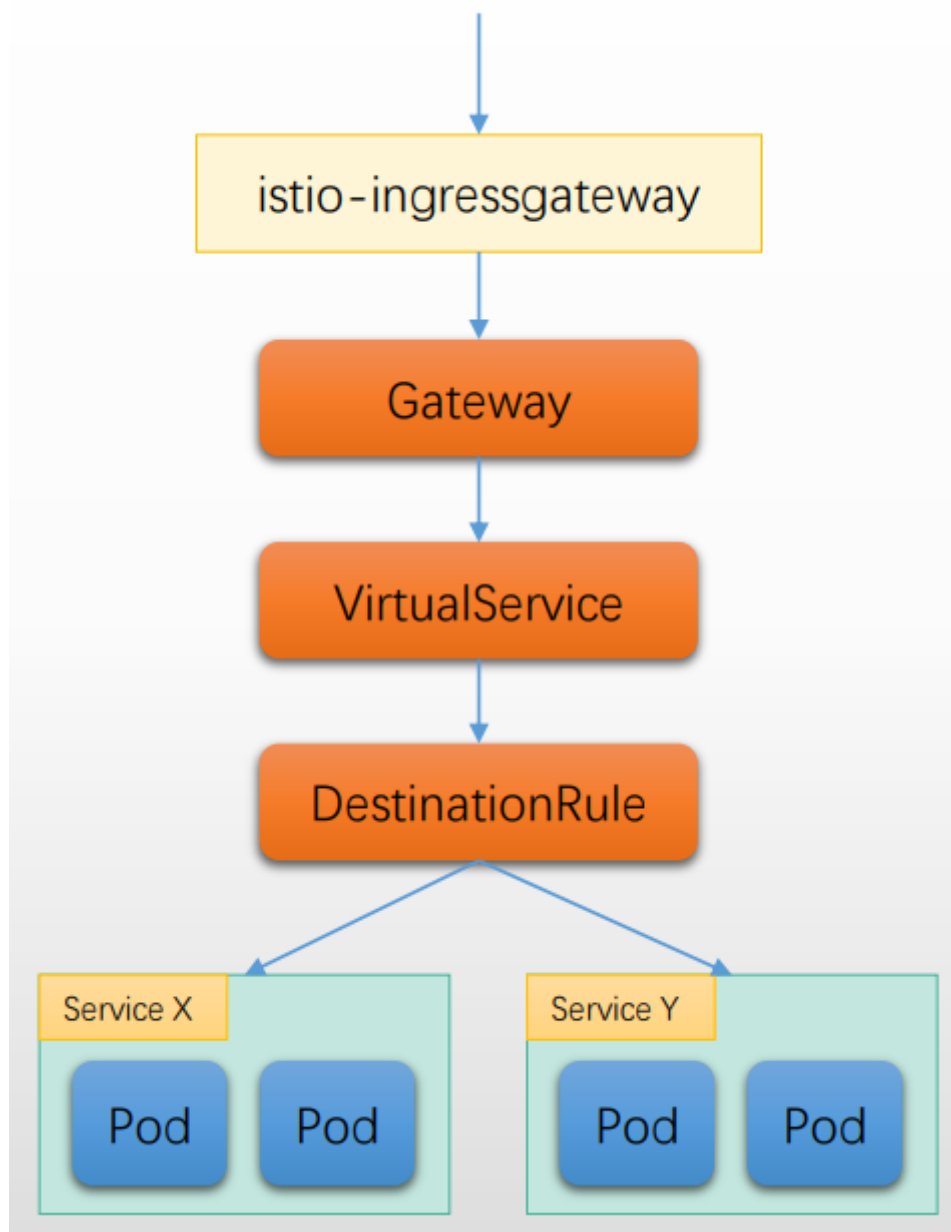
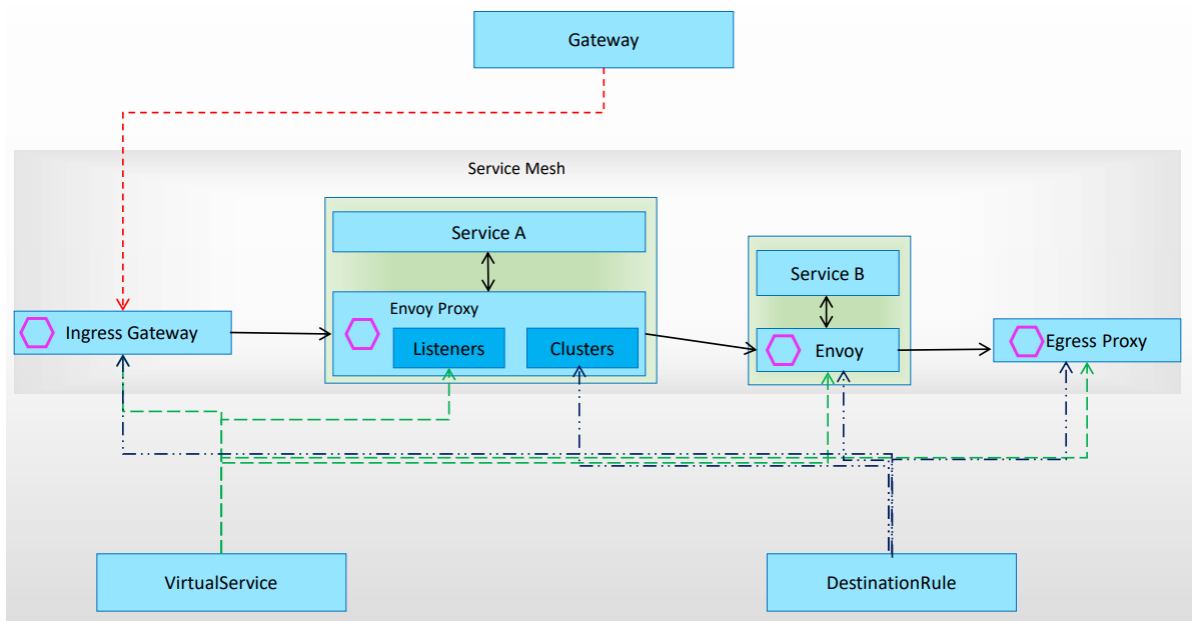


```

    #对集群外暴露的端口
selector:
  app: demo-mvc
#service类型 默认clusterIp
---
#以下为pod配置
apiVersion: apps/v1
#pod类型
kind: Deployment
#元数据信息
metadata:
  #Deployment名字
  name: mvc-deployment
  #所在命名空间
  namespace: istio0322
spec:
  #启动服务分片个数
  replicas: 1
  selector:
    #标签
    matchLabels:
      app: demo-mvc
  #模板
  template:
    metadata:
      #匹配service的标签
      labels:
        app: demo-mvc
        version: v1
      #以下为镜像库设置
    spec:
      containers:
        - image: demomvc:v1
          name: demomvc
          imagePullPolicy: IfNotPresent
          #设置pod资源和容器端口号
          resources:
            limits:
              cpu: "0.1"
              memory: 128Mi
            requests:
              cpu: "0.1"
              memory: 128Mi
          ports:
            - containerPort: 2000
          command: ["java", "-jar", "mvc.jar"]

```

## 服务网格化



上节中提到去掉gateway会失去路由过滤、集中鉴权等能力，而像个人财富这样安全性要求极高的系统，我们不仅需要上述能力，还应该有更多的能力如流控、灰度发布、故障注入等。在微服务体系，我们需要添加多种组件或者代码逻辑来实现等价功能，功能实现了却也为人极大地提高了维护难度，如果再遇到跨语言问题，情况会变得更加复杂。此时，Istio组件可以不用编码并且可以跨语言的能力便

能将这些问题轻松解决。只需在k8s中安装Istio。

自动注入边车指令

```
kubectl label namespace istio0322 istio-injection=enabled
```

调整istio监听网格外流量

```
kubectl get svc mvc-k8s-svc -nistio0322 -o yaml
```

添加需要的端口

```
- name: http-share
  nodePort: 30001
  port: 10001
  protocol: TCP
  targetPort: 2000
```

将配置文件调整为：

```
apiVersion: networking.istio.io/v1alpha3
kind: Gateway
metadata:
  name: mvc-istio-gateway
  namespace: istio0322
spec:
  selector:
    istio: ingressgateway
  servers:
    - port:
        number: 10001
        name: http
        protocol: HTTP
      hosts:
        - "*"

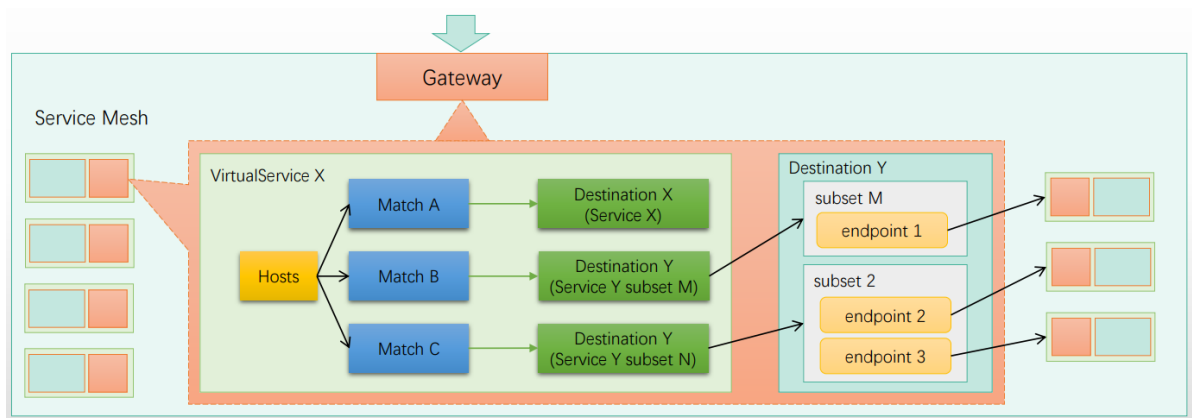
---
apiVersion: networking.istio.io/v1alpha3
kind: VirtualService
metadata:
  name: mvc-istio-vs
  namespace: istio0322
spec:
  hosts:
    - "*"
  gateways:
    - istio0322/mvc-istio-gateway
  http:
    - match:
        - queryParams:
            color:
              exact: white
        - queryParams:
            color:
              exact: red
      - queryParams:
```

```




    color:
      exact: yellow
  - queryParams:
      color:
        exact: blue
  route:
  - destination:
      host: mvc-k8s-svc
- match:
  - uri:
      prefix: /headerTest
  route:
  - destination:
      host: mvc-k8s-svc
  headers:
    request:
      set:
        test-header: "0322"

```

## 网络流量转发实例



## 设置三个版本的service服务

 ServiceDeploymentV1.yaml	2023/3/24 17:01	YAML 文件	1 KB
 ServiceDeploymentV2.yaml	2023/3/24 17:01	YAML 文件	1 KB
 ServiceDeploymentV3.yaml	2023/3/24 17:01	YAML 文件	1 KB

## 添加配置

```

apiVersion: networking.istio.io/v1alpha3
kind: VirtualService
metadata:
  name: service-istio-vs
  namespace: istio0322
spec:
  hosts:
  - service-k8s-svc
  http:
  - match:
    - headers:
        color:
          exact: red
    route:
    - destination:
        host: service-k8s-svc

```

```

      subset: sub-v1
- match:
  - headers:
      color:
        exact: yellow
    route:
  - destination:
      host: service-k8s-svc
      subset: sub-v2
- match:
  - headers:
      color:
        exact: blue
    route:
  - destination:
      host: service-k8s-svc
      subset: sub-v3
- route:
  - destination:
      host: service-k8s-svc
      subset: sub-v1
    weight: 33
  - destination:
      host: service-k8s-svc
      subset: sub-v2
    weight: 33
  - destination:
      host: service-k8s-svc
      subset: sub-v3
    weight: 34
---
apiVersion: networking.istio.io/v1alpha3
kind: DestinationRule
metadata:
  name: service-istio-rule
  namespace: istio0322
spec:
  host: service-k8s-svc
  subsets:
  - name: sub-v1
    labels:
      version: v1
  - name: sub-v2
    labels:
      version: v2
  - name: sub-v3
    labels:
      version: v3

```

## 实战

---

预计需要1-2个小时

## 结束语

---

综上所述，服务网格技术能够有效解决传统微服务架构存在维护成本过高、治理难度偏大、技术栈难以转型等痛点问题。与此同时，在诸多企业、机构的共同努力下，服务网格技术得以持续完善和优化，在银行业也有了较为充分的论证和应用。为稳步推进我行的信息系统转型任务，深度提升科技赋能实效，在当前信息化工作中，应当对服务网格技术的发展和应用提高重视，结合行业经验和我行现状，通过研讨、试验等方式，充分论证服务网格技术的可行性，深入挖掘应用潜力，为行内信息系统的技术转型探索出新的方向。