< HDC.Cloud >
华为开发者大会2021

华为云MCP
多云跨云的容器治理与实践

王泽锋　　华为云云原生开源负责人
徐中虎　　华为云高级工程师

HUAWEI

About Us

王泽锋

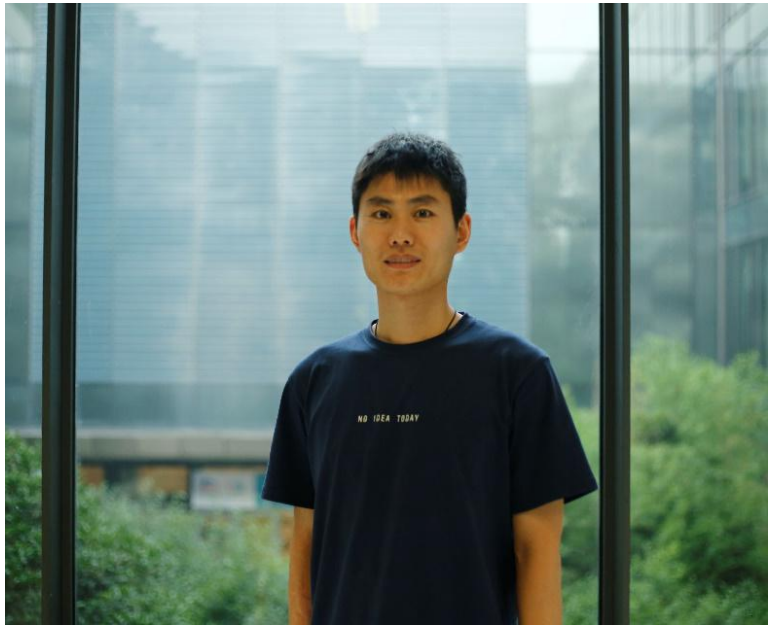
华为云·云原生开源负责人

Kubernetes社区资深Maintainer

KubeEdge和Volcano项目联合创始人

CNCF技术监督委员会贡献者

对云原生技术和开源生态有深入的见解



徐中虎

华为云·云原生开源团队核心成员

Istio社区Maintainer， Steering Committee委员

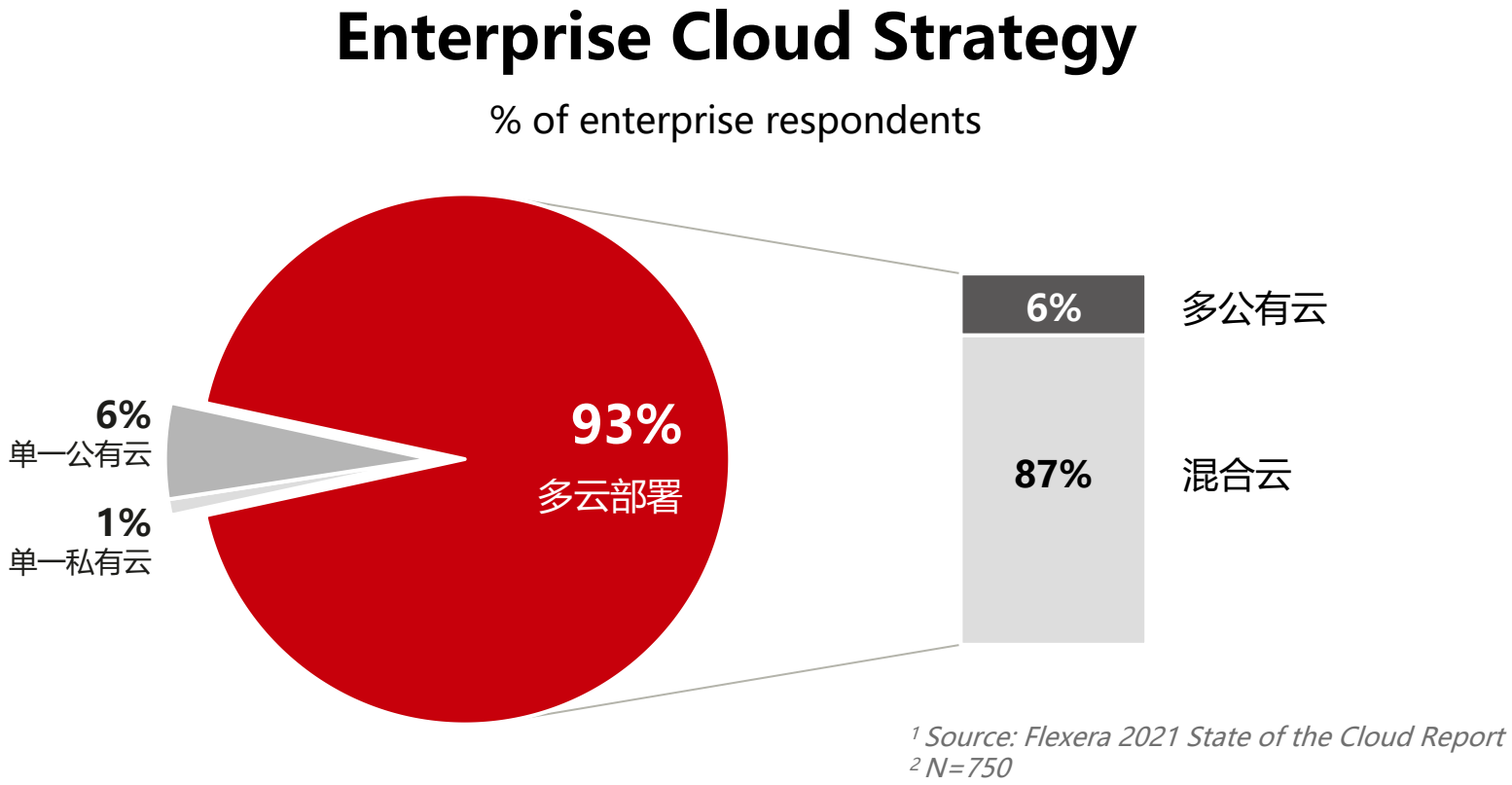
Kubernetes项目核心贡献者

聚焦在云原生Kubernetes, Docker以及服务网格等领域。

HUAWEI

目录

多云、多集群部署已经成为常态



[1] Source: Flexera 2021 State of the Cloud Report
[2] N=750

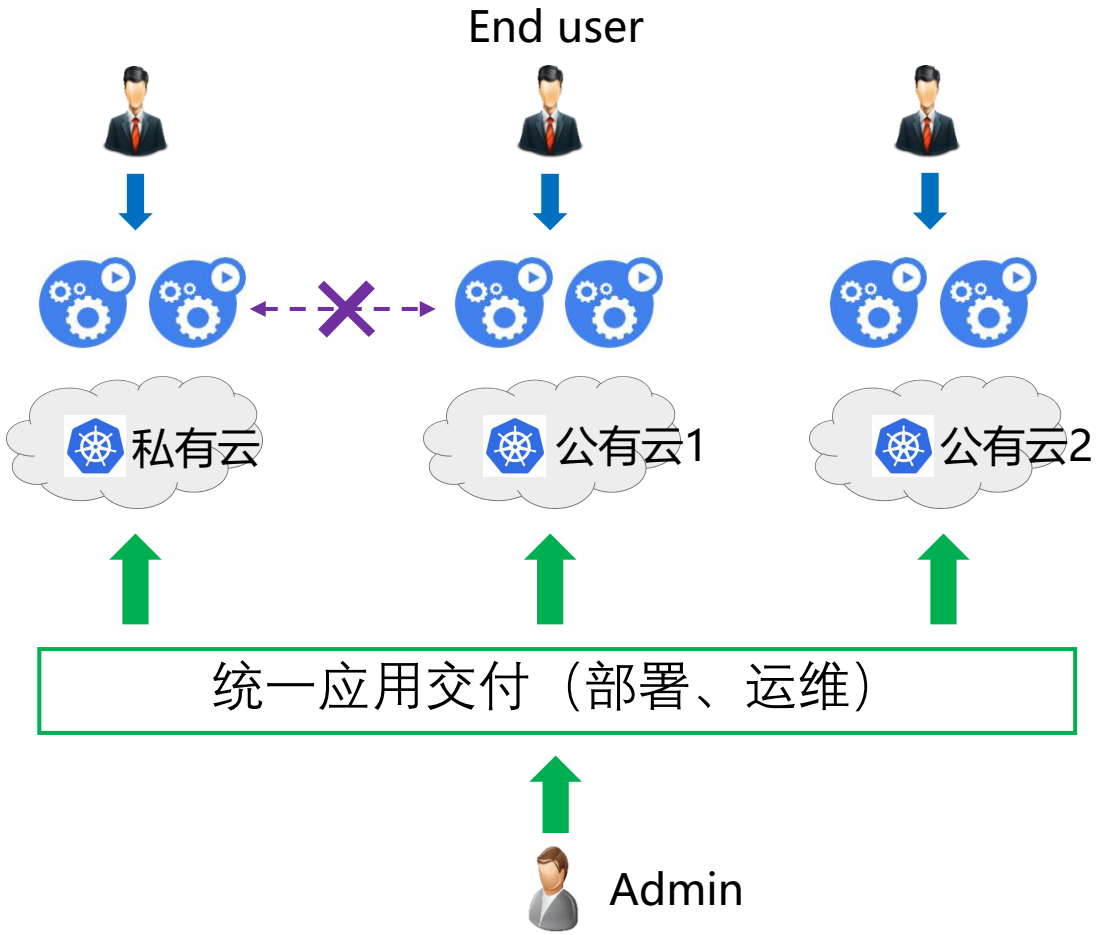
调查显示，超过93%的企业正同时使用多个云厂商的服务。

云原生技术和云市场不断成熟，未来将是编程式多云管理服务的时代。

HUAWEI

典型阶段1：多云多地部署，统一管理运维，减少重复劳动

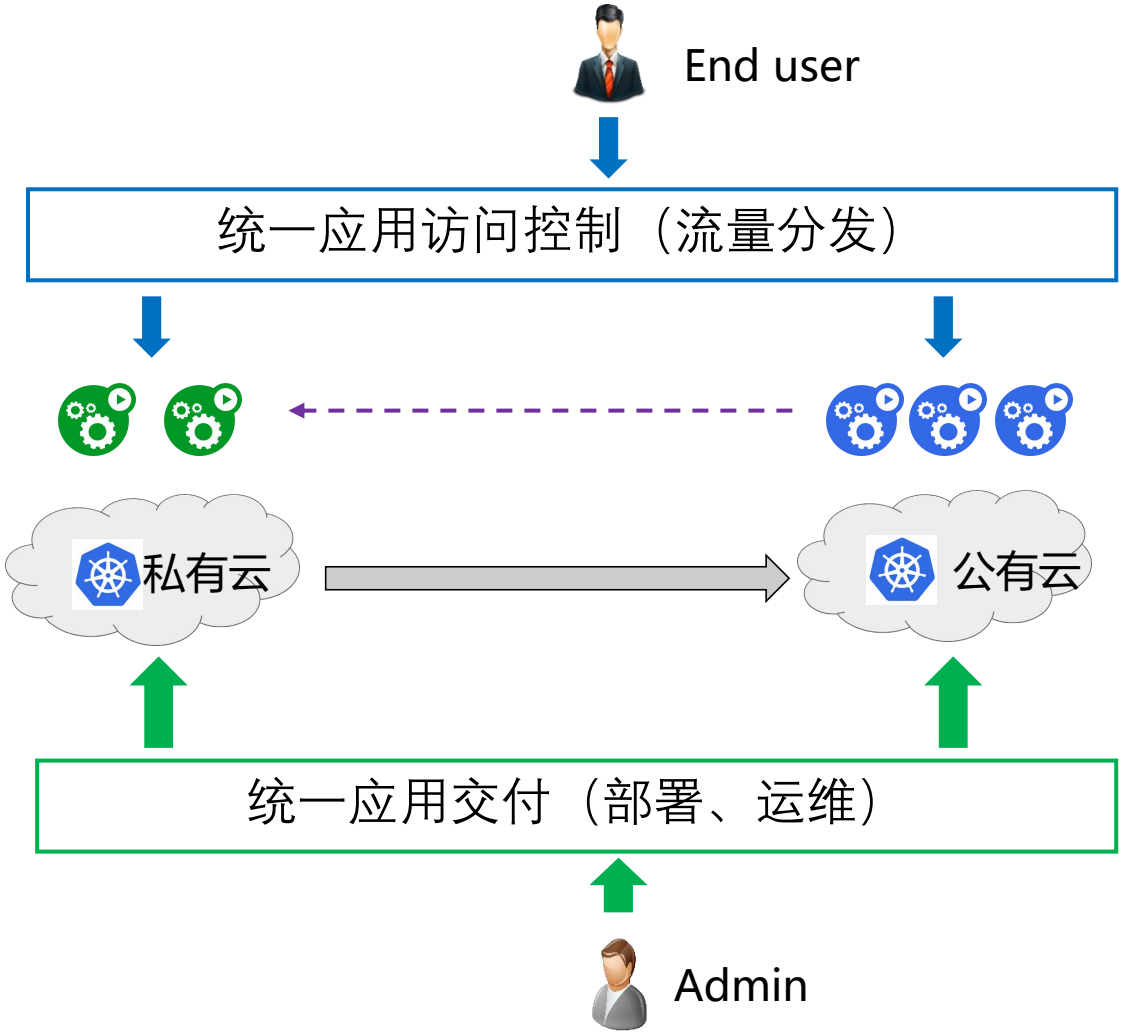
- **不同云上托管业务相同，无需跨云访问**
- **每朵云独立对外提供业务**
- **通过统一入口部署运维，减少重复劳动**
- **K8S版本、部署形态允许有一定差异**



HUAWEI

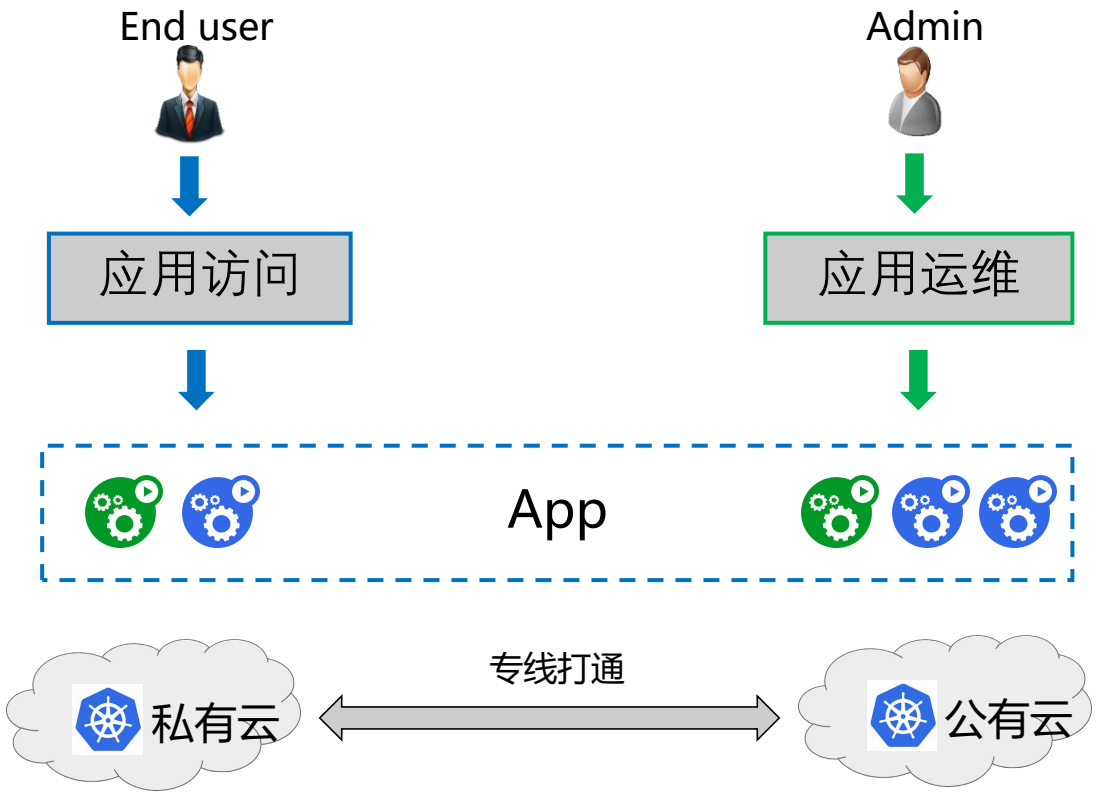
典型阶段2：多云统一资源池，应对业务压力波动

- **不同云上的托管业务可能不同**
  - 公有云为业务子集，或无状态应用
  - 私有云保留核心业务
- **基于业务特点设计流量分发策略**
  - 感知地域、云平台差异
  - 基于业务、作业类型
- **存在少量跨云访问，单向为主**
  - 多数情况下是公有云访问私有云
  - 需专线打通
- **通过统一平台控制业务跨云伸缩**
  - 业务忙时扩容到公有云，闲时收缩回私有云



典型阶段3：多云协同，统一应用平台，业务跨云部署

- **托管业务跨云部署，配置策略进行跨云调度**
  - 如配置实例数配比、跨云弹性伸缩
  - 私有云依然保留核心业务
- **应用平台统一管理应用访问、运维等**
  - 应用平台提供统一的应用访问控制（路由、负载均衡、灰度等）、应用运维（部署、监控、弹性等），无需定制
- **重度的跨云业务访问，需要云网络专线打通**



云原生的多云仍然充满挑战

多云容器集群管理的挑战

集群繁多

繁琐重复的集群配置

云厂商的集群管理差异

碎片化的API访问入口

业务分散

应用在各集群的差异化配置

业务跨云访问

集群间的应用同步

集群的边界限制

资源调度受限于集群

应用可用性受限于集群

弹性伸缩受限于集群

厂商绑定

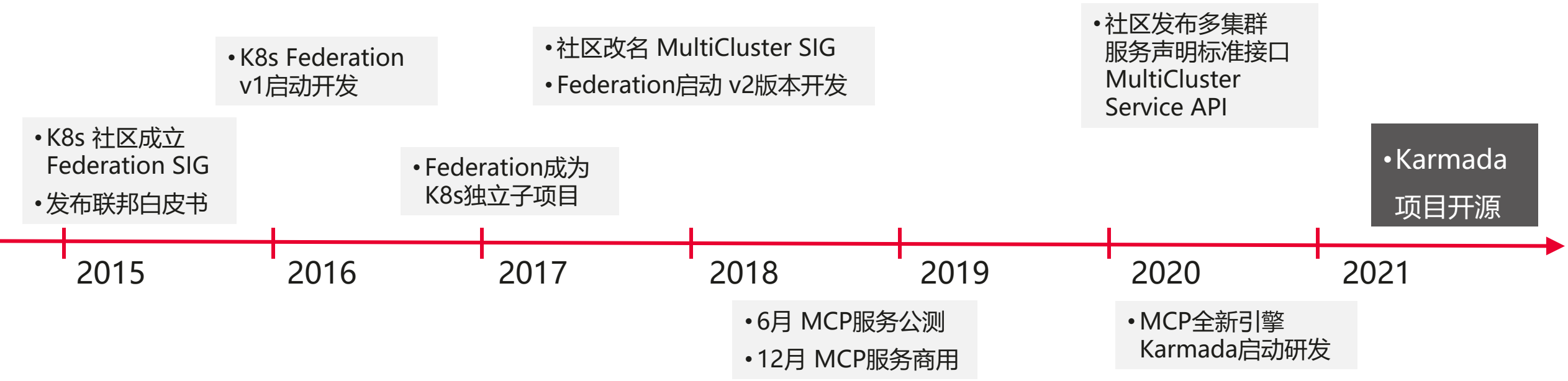
业务部署的“黏性”

缺少自动的故障迁移

缺少中立的开源多集群编排项目

HUAWEI

华为云MCP发展历程



HUAWEI

Karmada：开源的云原生多云容器编排引擎

使用Karmada构建无限可扩展的容器资源池

让开发者像使用单个K8s集群一样使用多云

联合发起单位

兼容K8s API

0代码改造升级多云架构

全网统一管理

公有云、混合云统一管理

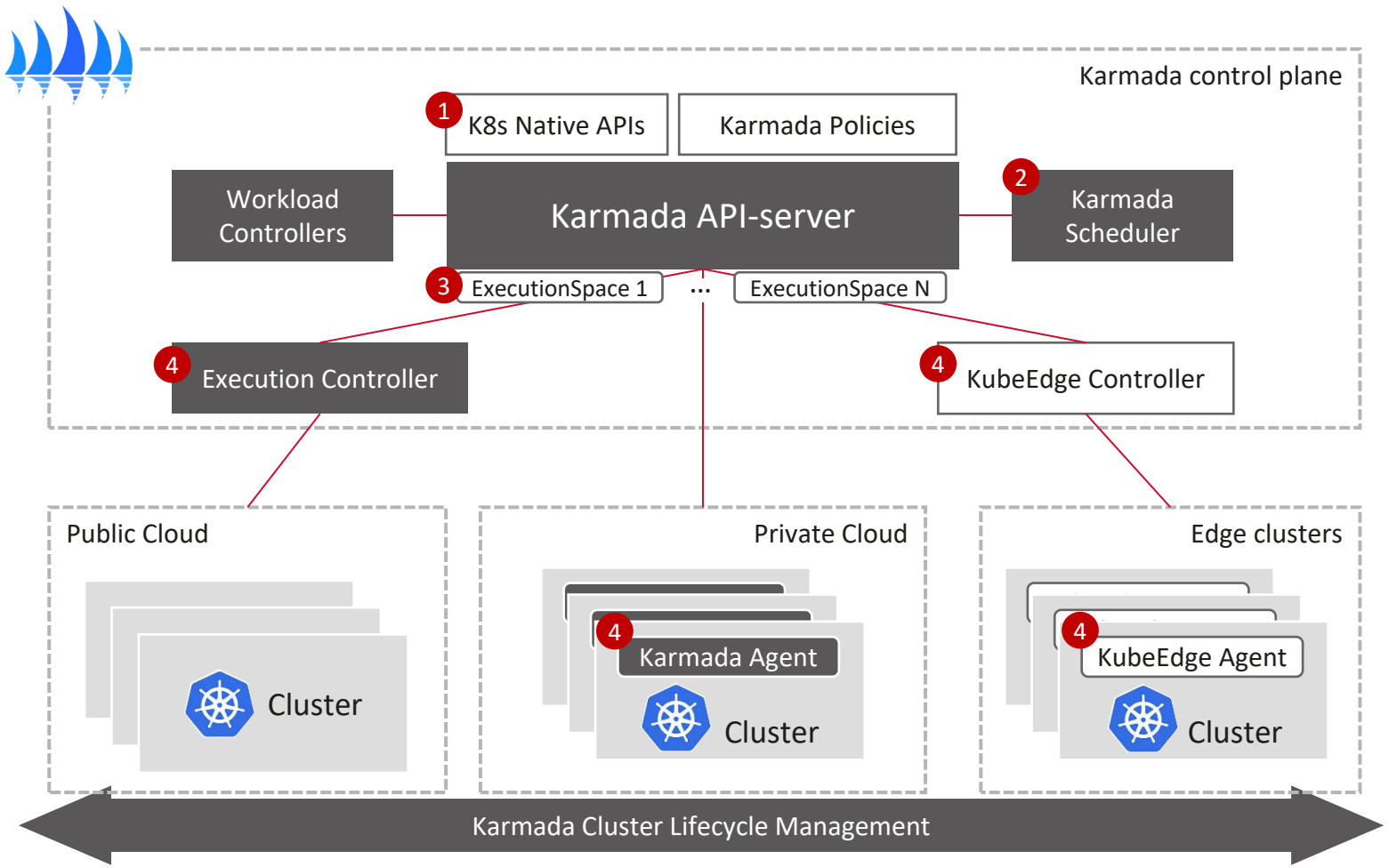
能力开箱即用

内置10+基于行业场景的调度能力插件

https://github.com/karmada-io/karmada

HUAWEI

Karmada：开源的云原生多云容器编排引擎



丰富云原生生态，开箱即用

K8s原生API + 扩展策略

零改造、开箱即用

丰富的多集群调度

支持集群亲和性、多维度HA等算法

支持静态、动态权重拆分引用

集群资源空间隔离

不同集群资源分别存放，隔离权限

多种模式集群同步

中心管理（Execution Controller）

分布式管理（Agent）

结合KubeEdge实现边缘集群管理

https://github.com/karmada-io/karmada

HUAWEI

多集群应用部署

零改造 — 使用K8s原生API部署一个多集群应用

可复用的分发策略

```yaml
apiVersion: policy.karmada.io/v1alpha1
kind: PropagationPolicy
metadata:
  name: multi-zone-replication
spec:
  resourceSelectors:
    - apiVersion: apps/v1
      kind: Deployment
      labelSelector:
        matchLabels:
          ha-mode: multi-zone-replication
  placement:
    spreadConstraints:
      - spreadByField: zone
        maxGroups: 3
        minGroups: 3
```

示例策略：为所有deployment配置多AZ的HA部署方案

```yaml
apiVersion: apps/v1
kind: Deployment
metadata:
  name: nginx-deployment
  app: nginx
  ha-mode: multi-zone-replication
spec:
  replicas: 3
  selector:
    matchLabels:
      app: nginx
  template:
    metadata:
      labels:
        app: nginx
    spec:
      containers:
        - name: nginx
          image: nginx
        ports:
        - containerPort: 80
```

原生的单集群API

使用标准的K8s API定义部署应用

```
kubectl create -f nginx-deployment.yaml
```

https://github.com/karmada-io/karmada

HUAWEI

Propagation Policy: 可重用的应用多集群调度策略

```yaml
apiVersion: policy.karmada.io/v1alpha1
kind: PropagationPolicy
metadata:
  name: example-policy
spec:
  resourceSelectors:
    - apiVersion: apps/v1
      kind: Deployment
      name: deployment-1
      labelSelector:          # standard labelSelector
  propagateDependensies: false
  placement:
    clusterAffinity:
      clusterNames:
        - cluster1
        - cluster3
    clusterTolerations:       # like pod tolerations
    spreadConstraints:
      - spreadByLabel: failuredomain.kubernetes.io/zone
        maxGroups: 3
        minGroups: 3
  schedulerName: default
```

resourceSelector

- 支持关联多种资源类型
- 支持使用 *name* 或 *labelSelector* 进行对象筛选

placement

- *clusterAffinity*:
  - 定义倾向调度的目标集群
  - 支持通过 names 或 labelselector 筛选
- *clusterTolerations*:
  - 类似单集群中Pod tolerations和 node taints
- *spreadConstraints*:
  - 定义应用分发的HA策略
  - 支持对集群动态分组：按Region、AZ、特性label分组，实现不同层级的HA

Override Policy: 跨集群可重用的差异化配置策略

```yaml
apiVersion: policy.karmada.io/v1alpha1
kind: OverridePolicy
metadata:
  name: example-override
  namespace: default
spec:
  resourceSelectors:
    - apiVersion: apps/v1
      kind: Deployment
  targetCluster:
    labelSelector:
      matchLabels:
        failuredomain.kubernetes.io/region: dc1
  overriders:
    imageOverrider:
    - component: prefix
      operator: replace
      value: "dc-1.registry.io"
```

resourceSelector

- 支持使用 *name* 或 *labelSelector* 进行对象筛选

overriders

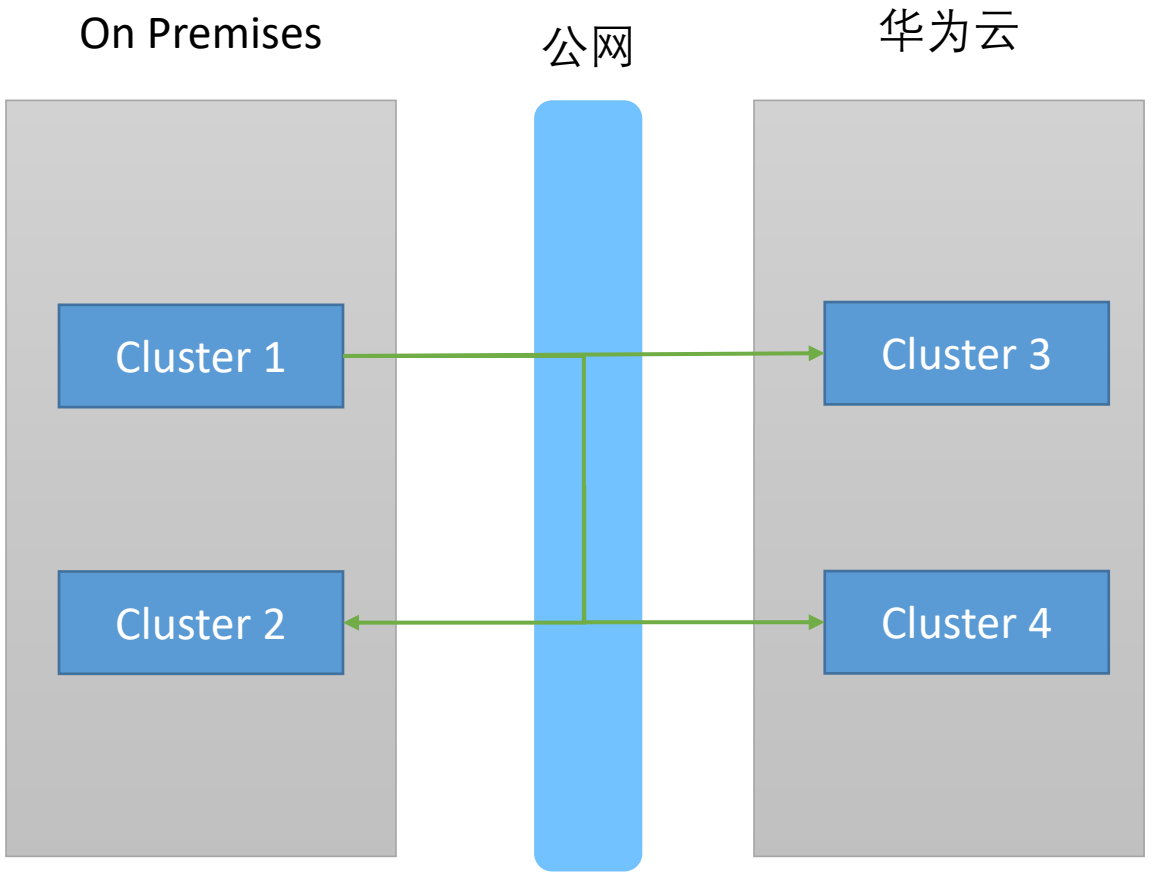
- 支持多种override插件类型
- *plainTextOverrider* :
  - 基础插件，纯文本操作替换
- *imageOverrider*:
  - 针对容器镜像的差异化配置插件

HUAWEI

多集群服务治理

多集群服务治理要解决的问题

- 服务发现
- DNS解析
- 负载均衡、熔断、故障注入，流量切分等高级流量治理
- 跨云的访问安全性



ServiceMesh的优势

流量治理

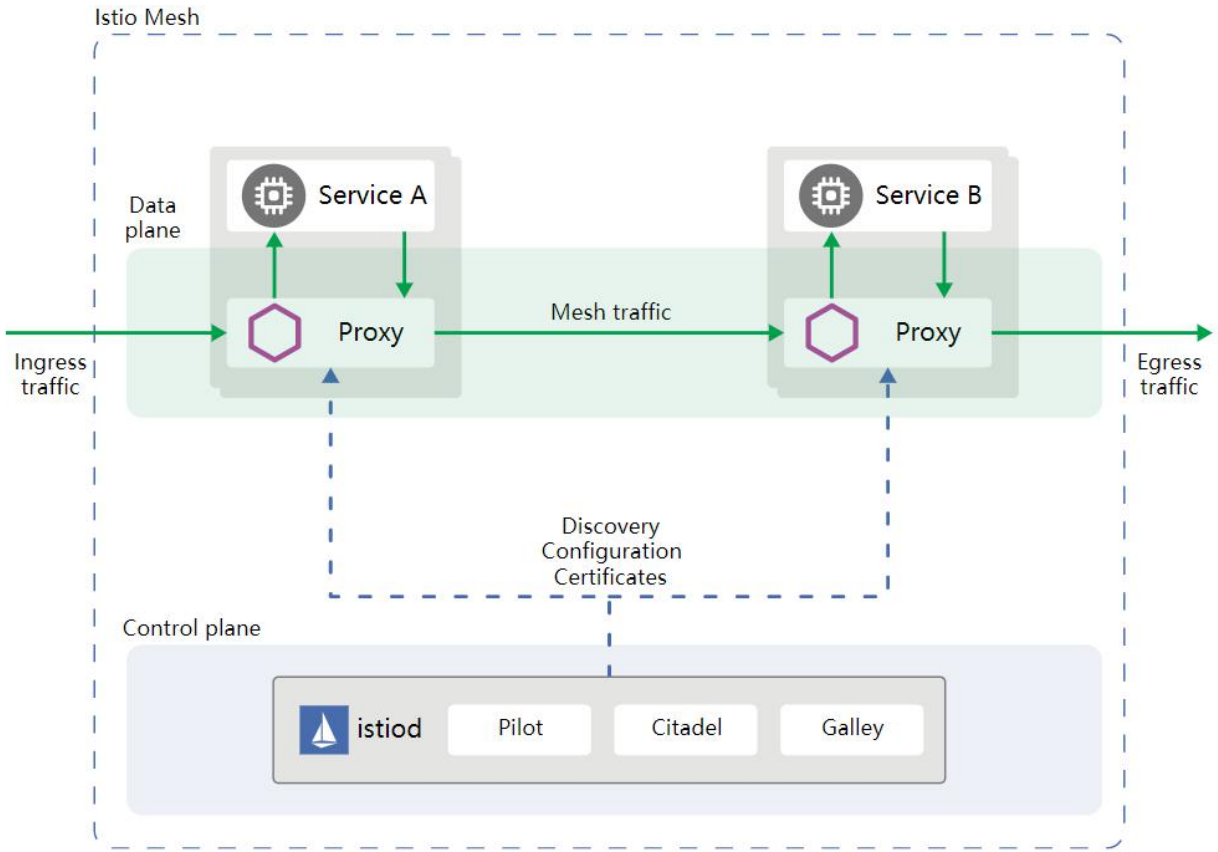
- Resilience（熔断、超时、重试、故障注入）
- 灰度发布
- 负载均衡、路由匹配

安全

- 数据加密、认证、鉴权

可观测

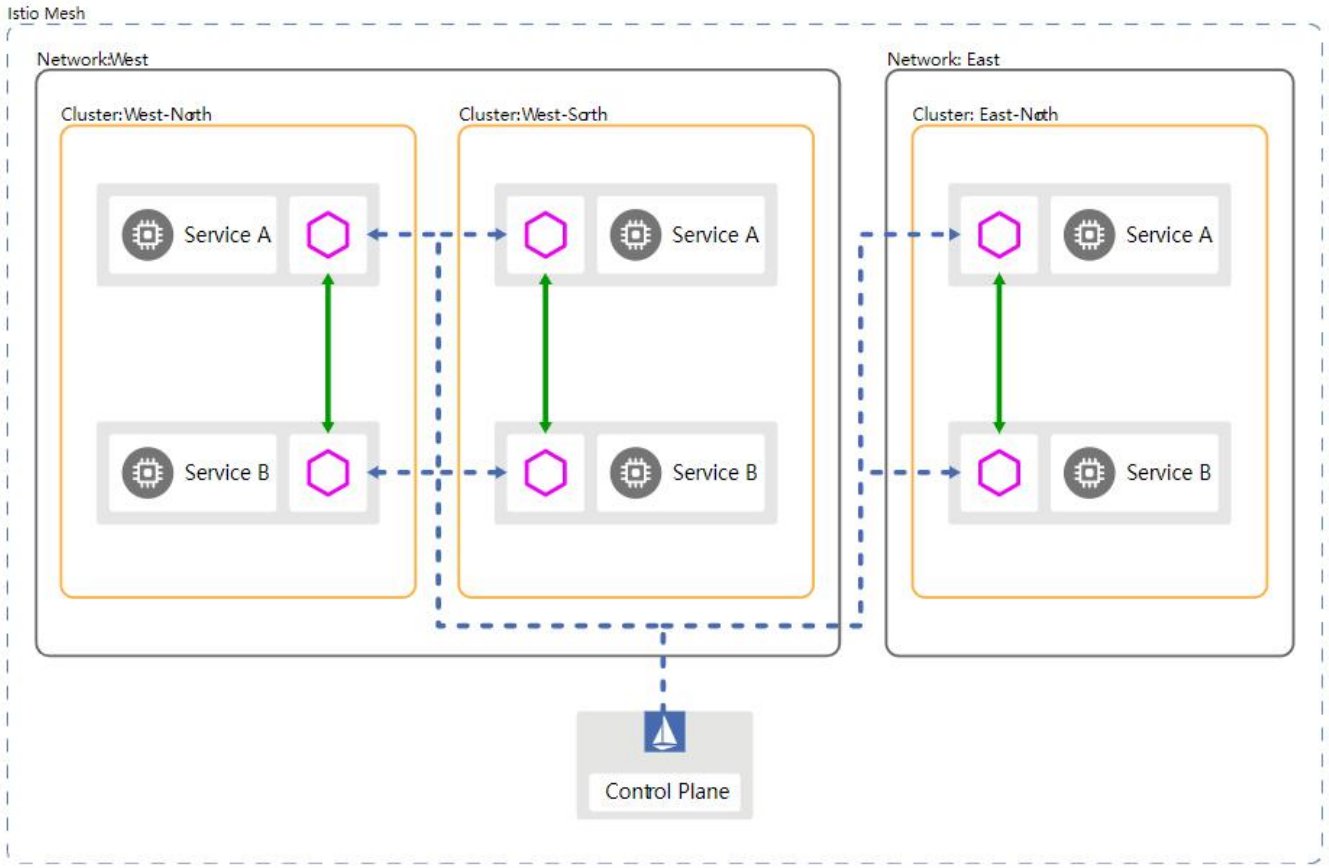
- Metrics
- Traces
- Access Log



HUAWEI

多集群服务网格技术模型

- 扁平网络 vs 非扁平网络
- 单服务网格 vs 多服务网格
- 单控制面 vs 多控制面



HUAWEI

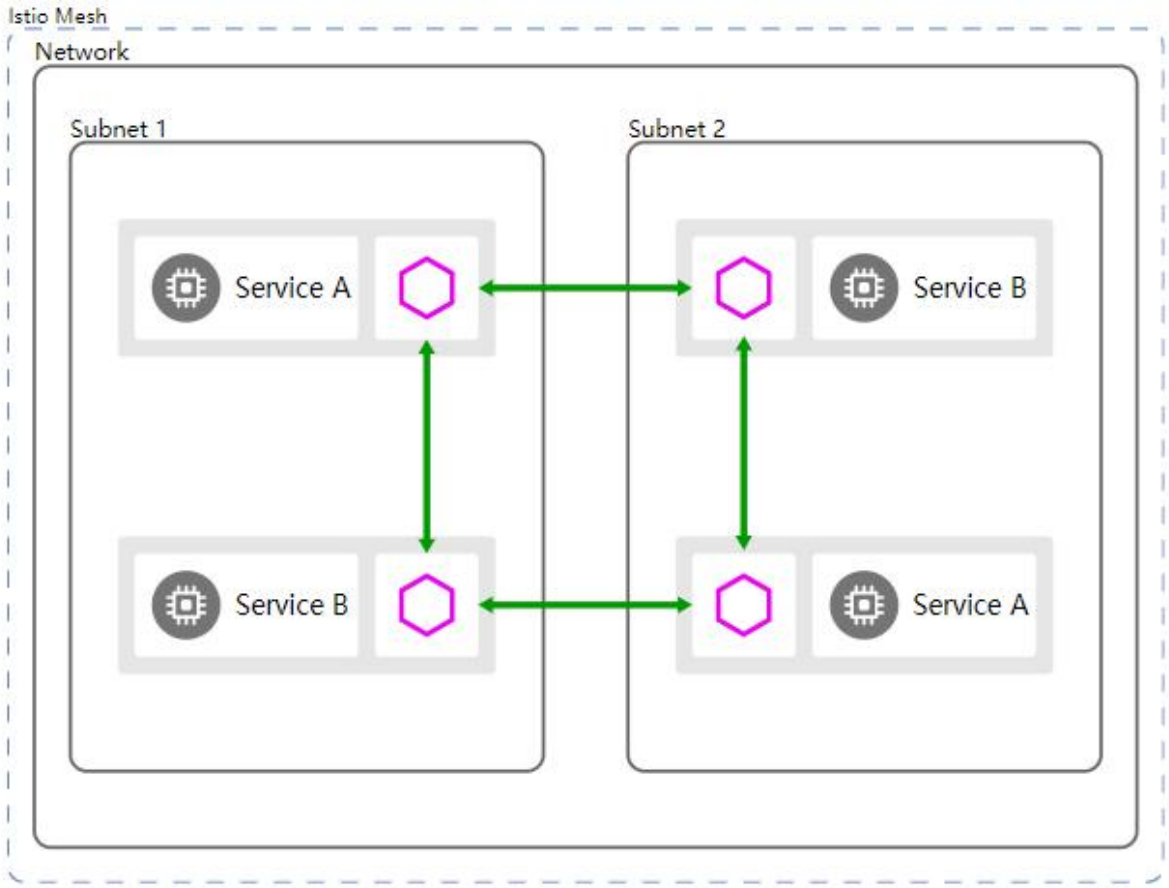
多集群服务网格-扁平网络

优势：

- 东西向服务访问延迟低

缺点：

- 组网复杂性
- 安全性：所有的工作负载在同一网络中.
- Scalability: pod、service IP地址不冲突



HUAWEI

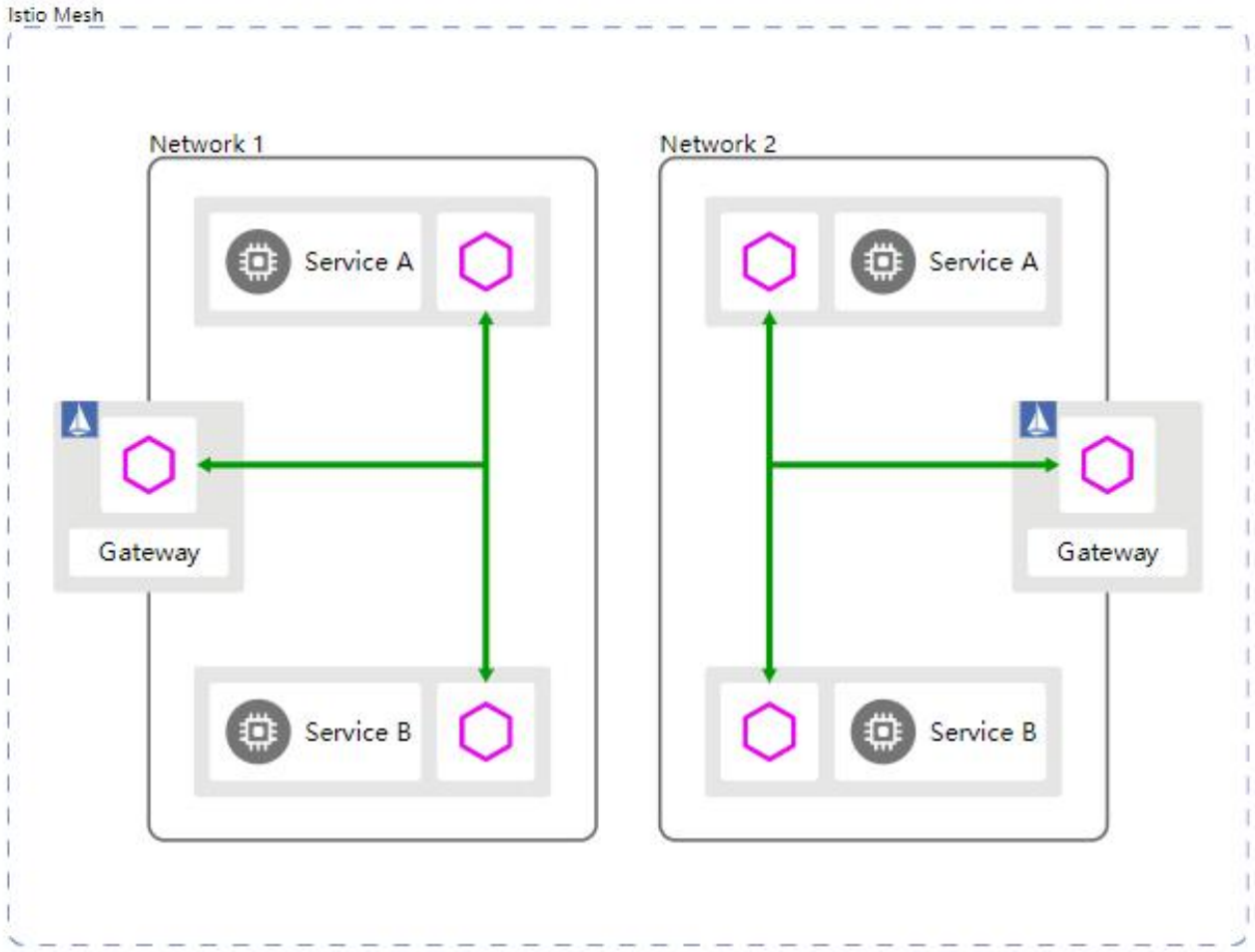
多集群服务网格-非扁平网络

优势：

- 网络隔离，安全性相对更高
- 组网简单
- Scalability: pod/service网络地址可以重叠

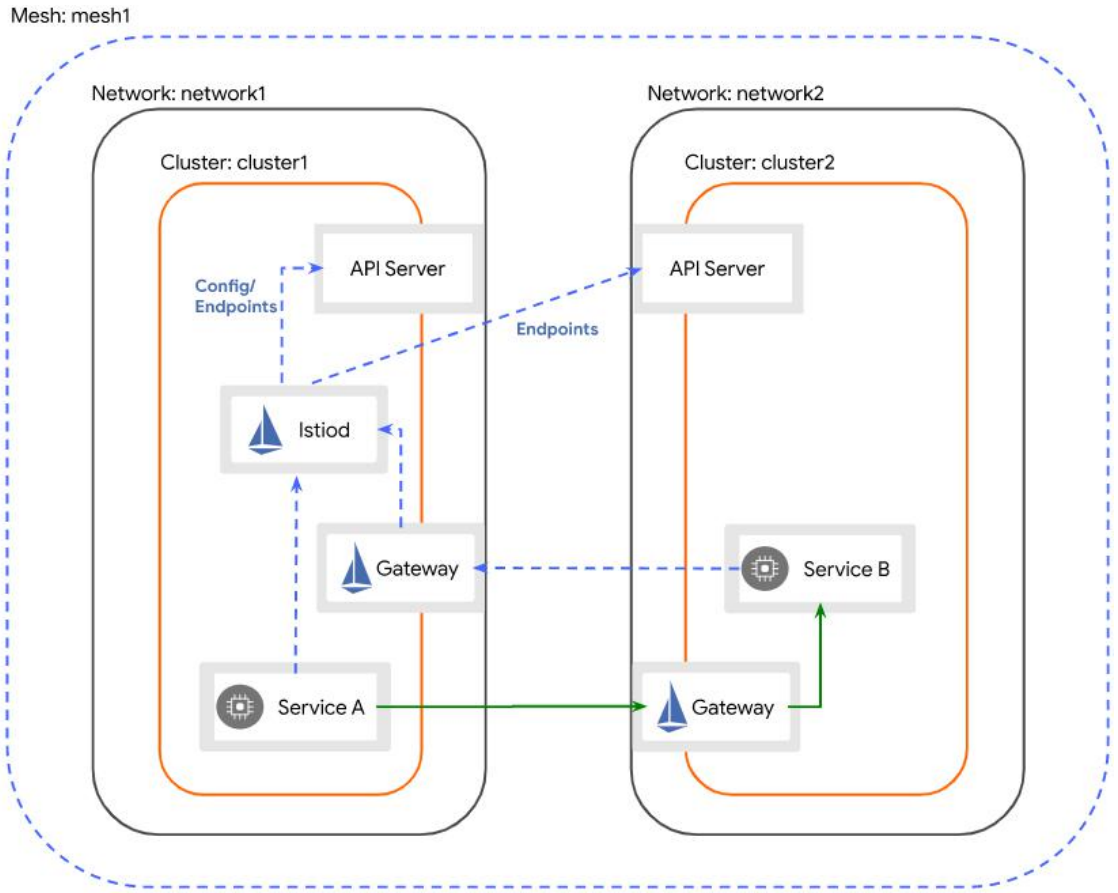
缺点：

- 跨集群服务访问需要通过东西向网关
- Gateway工作依赖TLS Auto Passthrough



非扁平网络-单控制面

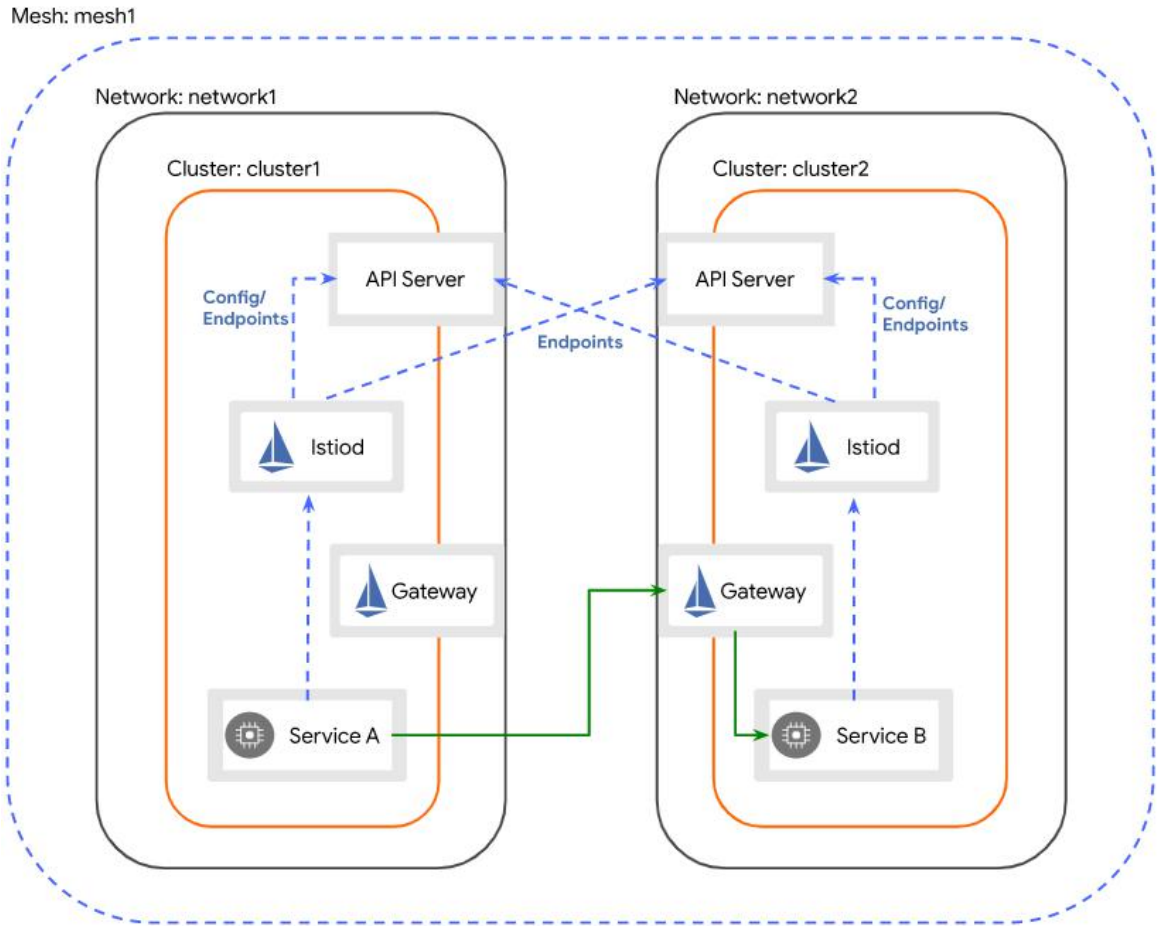
- 单个控制面（可以部署在用户集群或者完全托管）
- 服务发现
- 配置发现
- Split Horizon EDS
- 东西向网关



HUAWEI

非扁平网络-多控制面

- 控制面部署在每个集群
- 服务发现
- 配置发现
- Sidecar连接本集群内的Istio控制面，与单控制面相比，有更好的性能和可用性

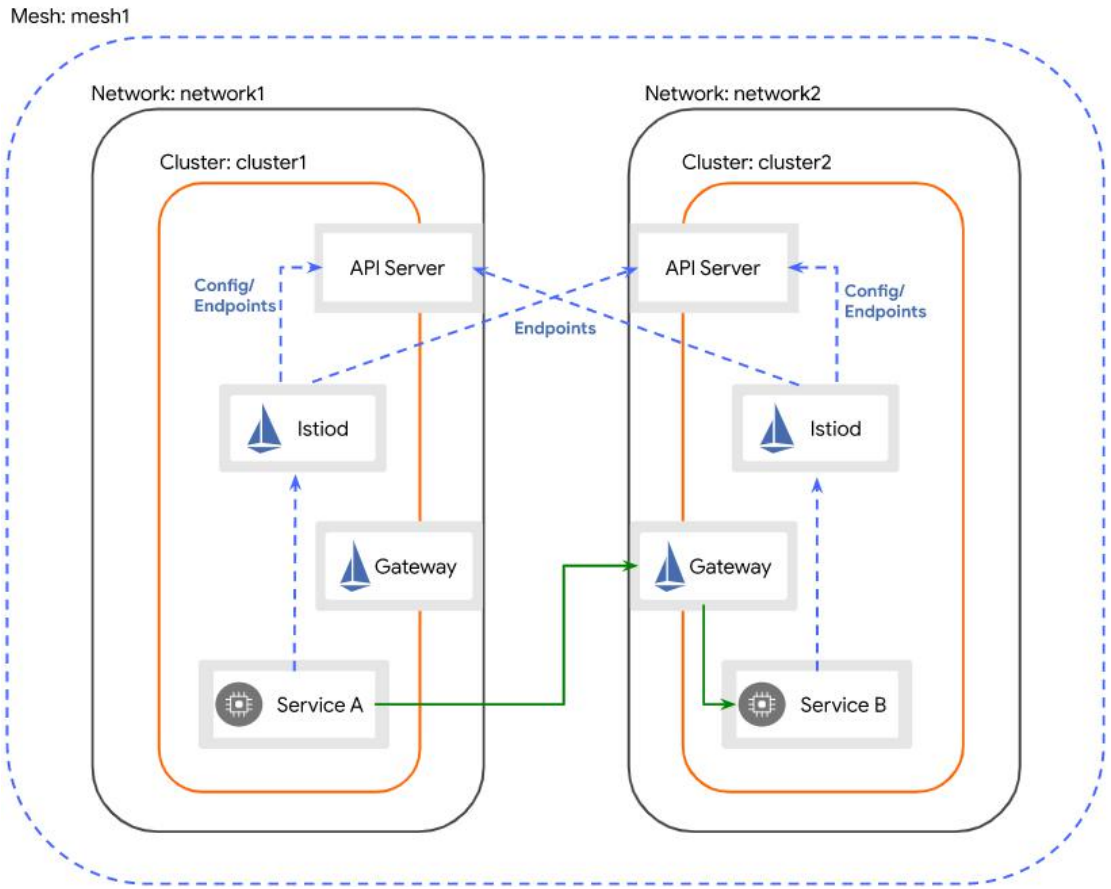


非扁平网络-东西向网关

- **网关地址获取**
- **Split horizon EDS:**

```
apiVersion: networking.istio.io/v1beta1
kind: Gateway
metadata:
  name: cross-network-gateway
  namespace: istio-system
spec:
  selector:
    istio: eastwestgateway
  servers:
  - hosts:
    - '*.local'
    port:
      name: tls
      number: 15443
      protocol: TLS
    tls:
      mode: AUTO_PASSTHROUGH
```

- **Network filter: "envoy.filters.network.sni_cluster"**



HUAWEI

Karmada社区路标

2021 Q1

- K8s原生API部署多集群应用
- 多集群应用状态聚合
- 多集群部署时，应用配置差异化
- 针对集群的亲和性调度
- 多集群高可用调度策略
- 纳管集群
- 集群目录

2021 Q2

- 多集群的应用故障迁移
- 根据多种策略动态拆分应用实例数
- 针对集群的Taint Toleration
- 多种集群同步模式：Push，Pull
- 集群生命周期管理

2021 Q3

- 多集群服务发现
- 多集群外部流量接入
- 多集群的应用自动伸缩
- 多集群的资源重平衡
- 更多高可用调度策略：多Region/多AZ/多厂商
- 聚合API访问

2021 Q4

- 多形态下多集群容器网络
- 多集群服务治理
- 多集群监控
- 多集群日志收集
- GitOps支持

HUAWEI

Take Away

- 多云已成必然
  - 云原生技术与多云诉求相互促进
- 多云的三个阶段
  - 标准技术栈，互操作的多个孤岛
  - 统一平台，云间统一调度，统一弹性
  - 多云无缝合一
- 云原生多云的典型挑战
  - 集群繁多
  - 业务分散、碎片化
  - K8s集群造成的边界
  - 厂商绑定
- Karmada项目核心价值
  - K8s原生API兼容，丰富云原生生态
  - 内嵌策略，开箱即用
  - 丰富的多集群调度支持
  - 集群资源空间隔离
  - 多种模式集群同步，屏蔽地域、网络限制
- 多集群服务治理
  - 扁平网络 vs 非扁平网络
  - 单服务网格 vs 多服务网格
  - 单控制面 vs 多控制面
  - 多种方案将持续并存
- Karmada后续计划
  - 整体技术栈Q4成型

HUAWEI

加入社区

https://github.com/karmada-io/karmada

https://karmada-io.slack.com

容器魔方公众号

每日推送图文

社区最新动态

直播课程、技术干货

扫码添加小助手

发送“karmada”加群

社区专家入驻

技术问题随时答疑

HUAWEI

华为云云原生专家团队倾情打造，免费推出云原生王者之路集训营！

即日起正式开启报名通道，云原生黄金集训营5月24号正式开课，线上全球直播！

报名链接及更多信息详情，扫描左侧二维码

HUAWEI

Thank you.

把数字世界带入每个人、每个家庭、
每个组织，构建万物互联的智能世界。

Bring digital to every person, home and organization for a fully connected, intelligent world.

HUAWEI