Istio multi-cluster traffic management speed up automobile company new business dev, deploy and ops

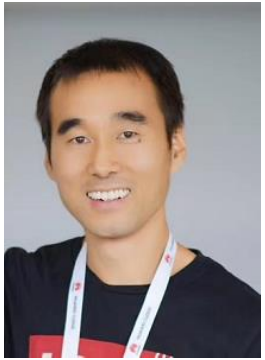
ken.liu@smart.com/zhangchaomeng@huawei.com

IstioCon

#IstioCon

Liu Kexing

smart Automobile | Head of IT development and architecture



Zhang Chaomeng

Chief architect of HUAWEI CLOUD Application Service Mesh(ASM).

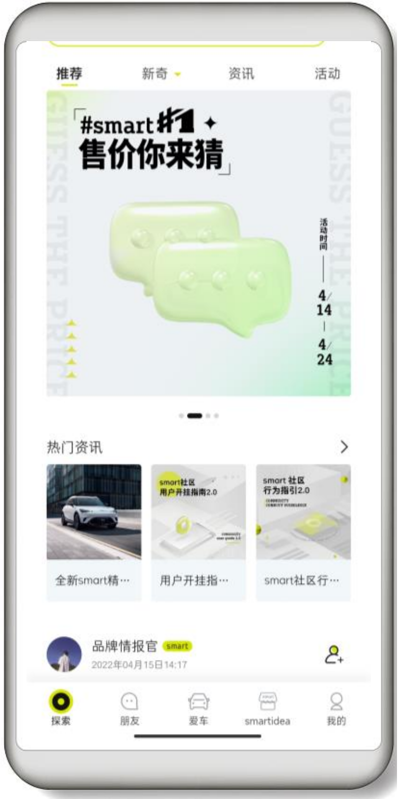
Liu Kexing has more than 10 years in automobile industry development and architecture experience. In recent 5 years he focuses on customer faced service development and connected car service architecture. Currently is developing the cloud native services for smart app, which is entire newly built under service meshed micro services for New Energy Vehicle sales model.

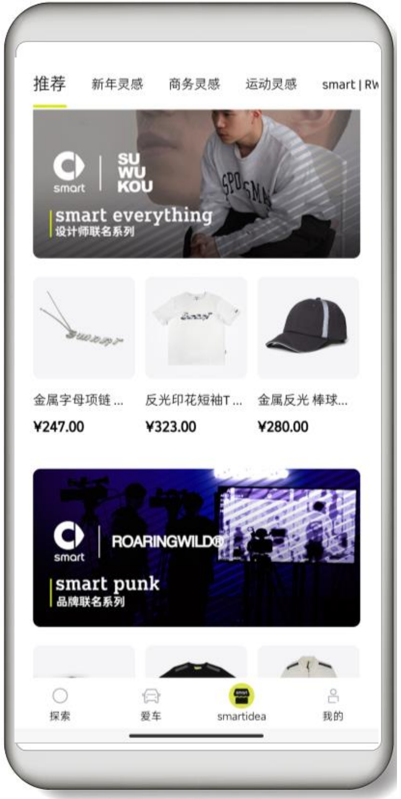
Chaomeng has been working on cloud native technologies in HUAWEI Cloud for more than 7 years, including Kubernetes, micro services, service catalog, APM, devops and service mesh for now.. He is Istio community member, author of book “Cloud Native Service Mesh Istio”(《云原生服务网格 Istio：原理、实践、架构与源码解析》).

Agenda

- Application background
- Service Mesh practice

Business Background

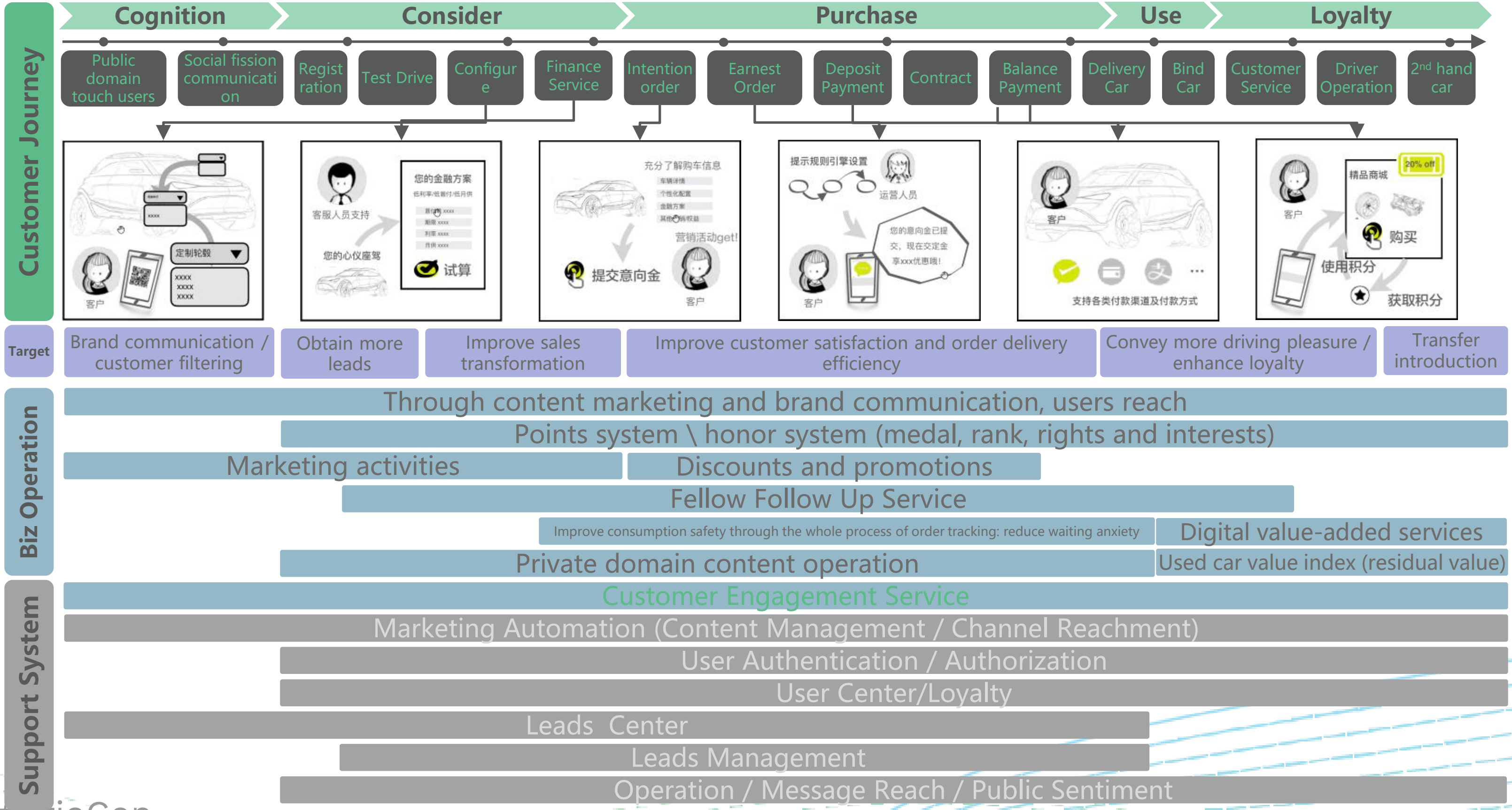




Since its birth, the smart brand has been a pioneer in future urban mobility. Today, smart is creating a more imaginative and innovative future with its progressive, premium style and futuristic technology.

- Entirely transformation
  from fuel vehicles to electric vehicles
- Intelligently social contact
  goes through customer journey
- Directedly to customer

Application Architecture



Customer Journey

Cognition: Public domain touch users; Social fission communication

Consider: Registration; Test Drive; Configure; Finance Service

Purchase: Intention order; Earnest Order; Deposit Payment; Contract; Balance Payment

Use: Delivery Car; Bind Car

Loyalty: Customer Service; Driver Operation; 2nd hand car

Target: Brand communication / customer filtering; Obtain more leads; Improve sales transformation; Improve customer satisfaction and order delivery efficiency; Convey more driving pleasure / enhance loyalty; Transfer introduction

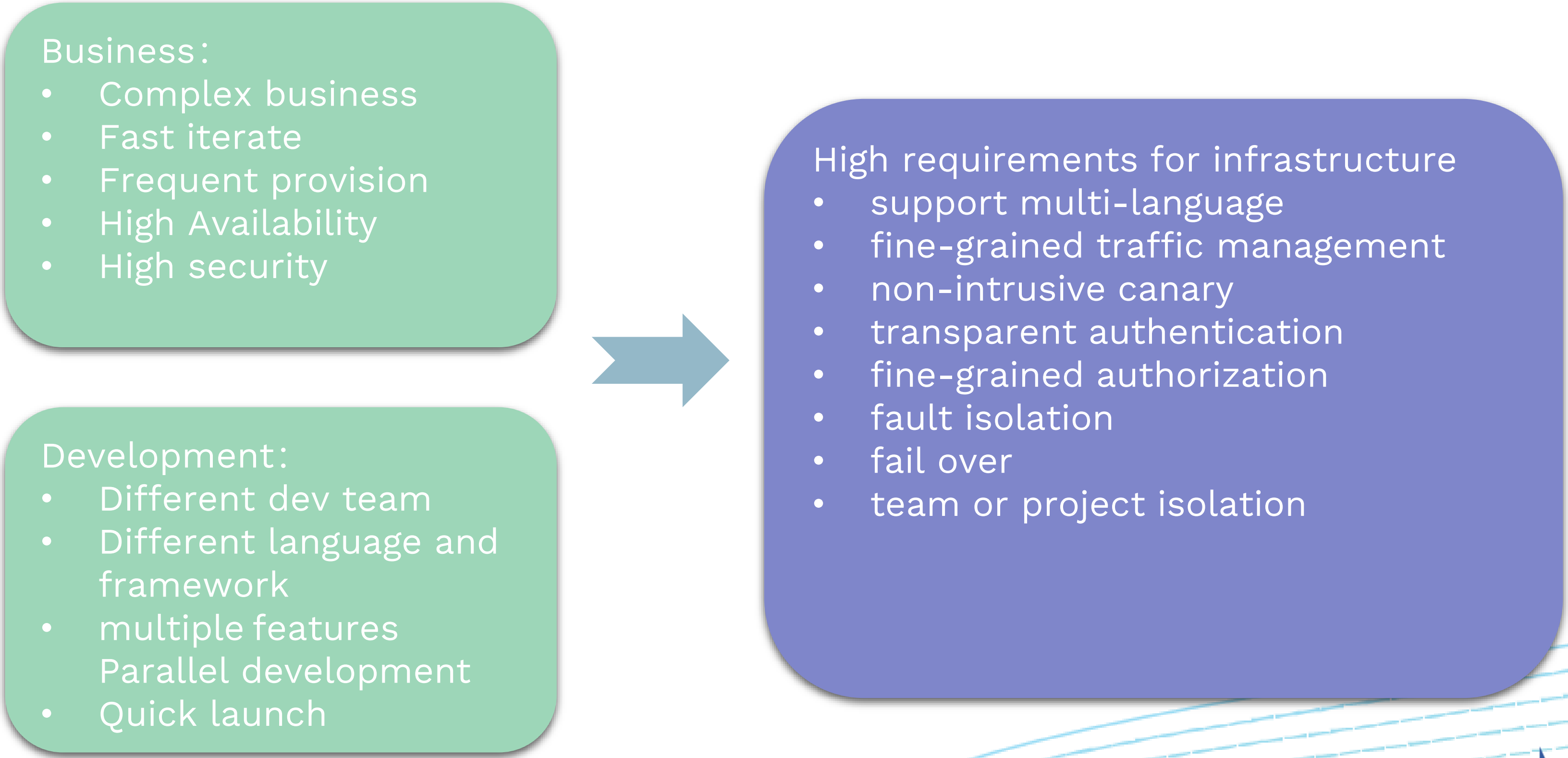
Biz Operation

- Through content marketing and brand communication, users reach
- Points system \ honor system (medal, rank, rights and interests)
- Marketing activities
- Discounts and promotions
- Fellow Follow Up Service
- Improve consumption safety through the whole process of order tracking: reduce waiting anxiety
- Digital value-added services
- Private domain content operation
- Used car value index (residual value)

Support System

- Customer Engagement Service
- Marketing Automation (Content Management / Channel Reachment)
- User Authentication / Authorization
- User Center/Loyalty
- Leads Center
- Leads Management
- Operation / Message Reach / Public Sentiment

IT Infrastructure Requirement



Business:

- Complex business
- Fast iterate
- Frequent provision
- High Availability
- High security

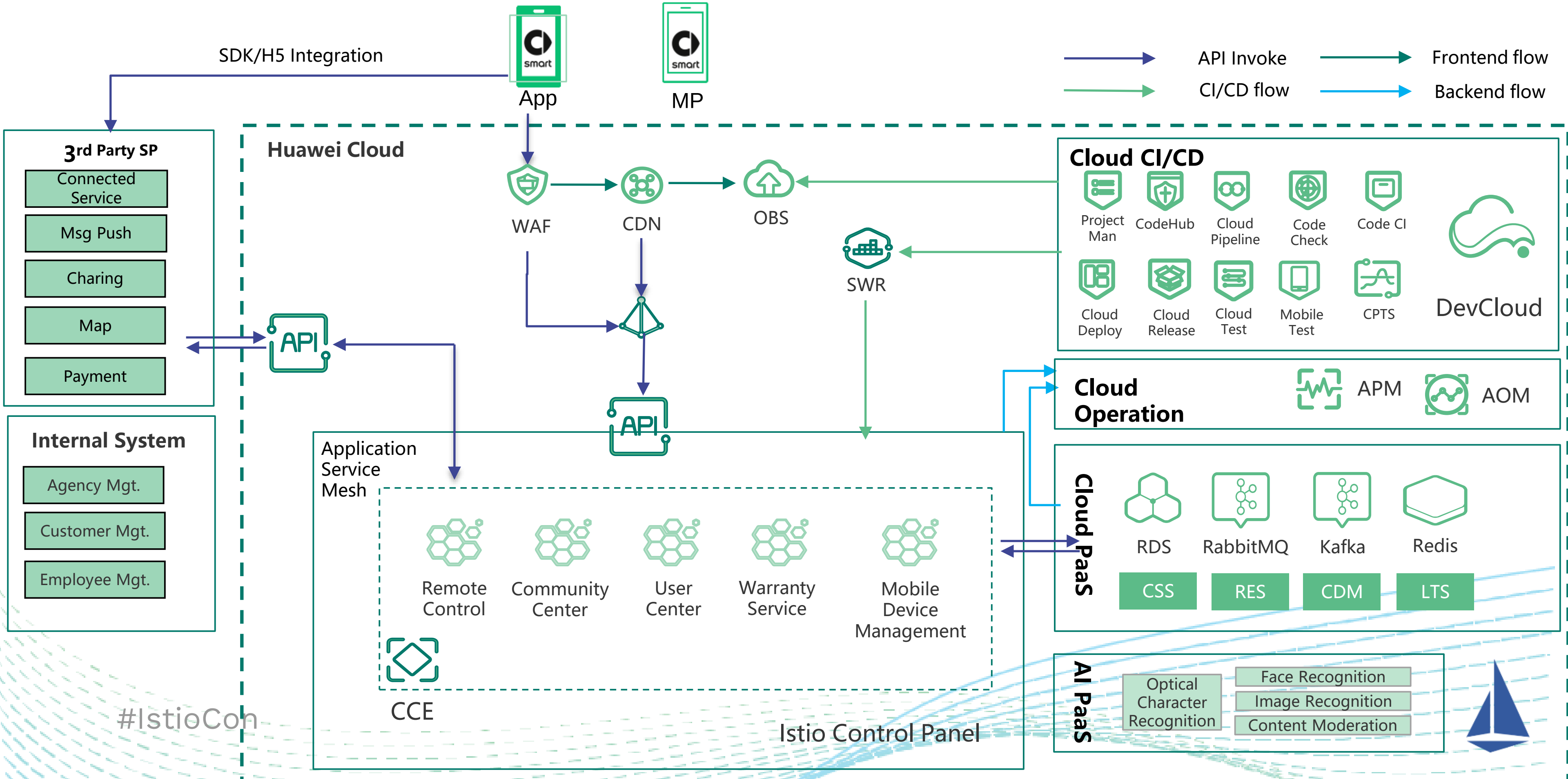
Development:

- Different dev team
- Different language and framework
- multiple features Parallel development
- Quick launch

High requirements for infrastructure

- support multi-language
- fine-grained traffic management
- non-intrusive canary
- transparent authentication
- fine-grained authorization
- fault isolation
- fail over
- team or project isolation

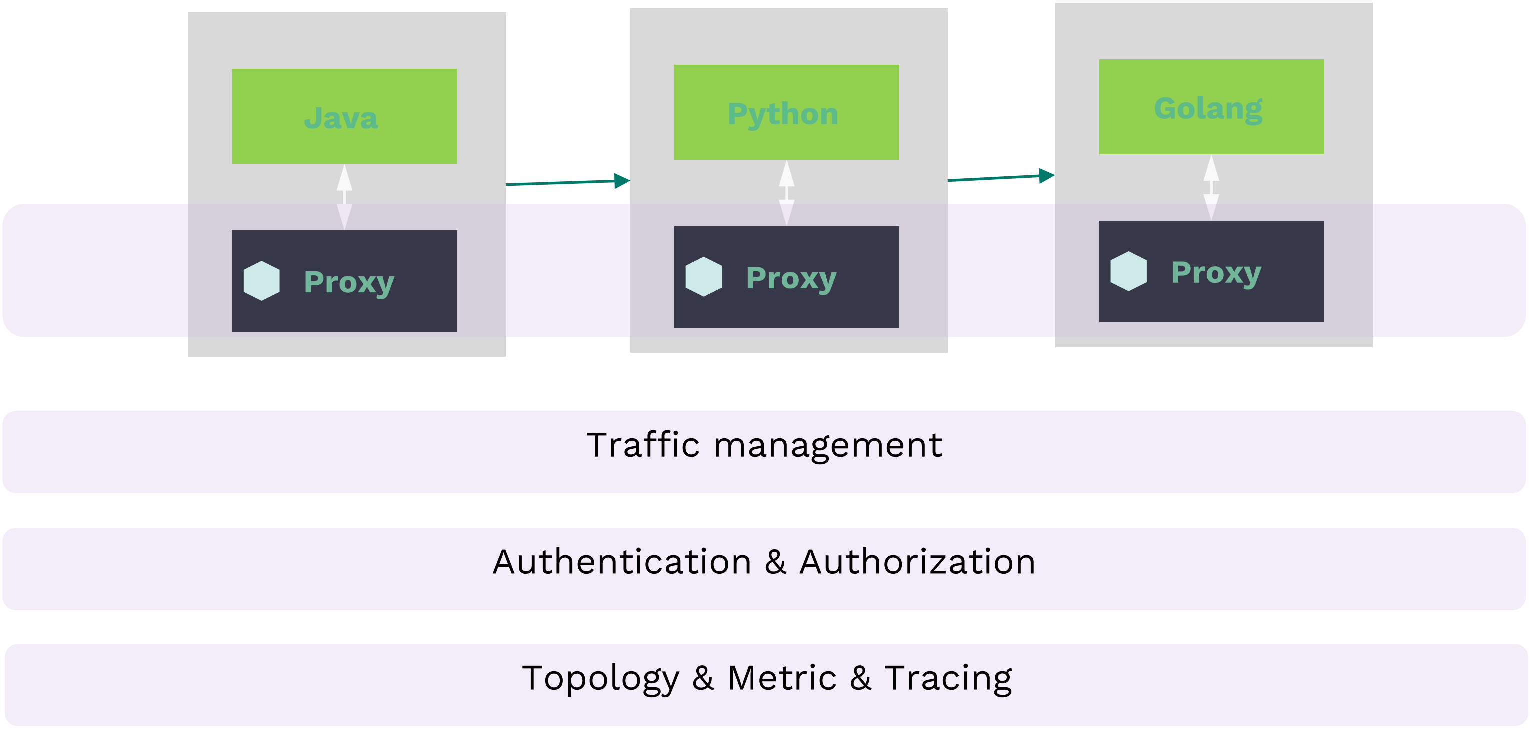
Cloud Native solution based on service mesh



Agenda

- Application background
- Service Mesh practice

Multi language app

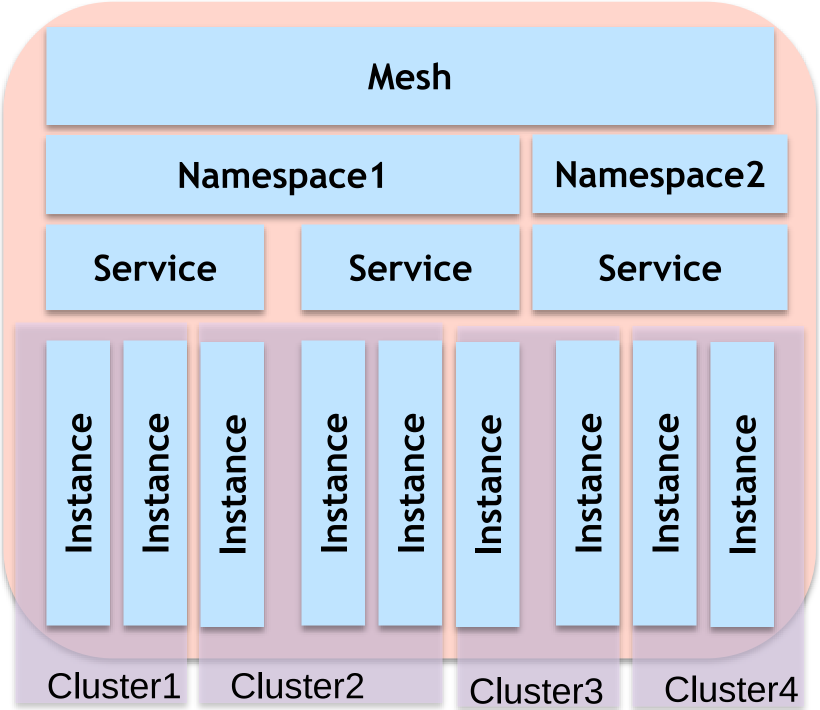


Traffic management

Authentication & Authorization

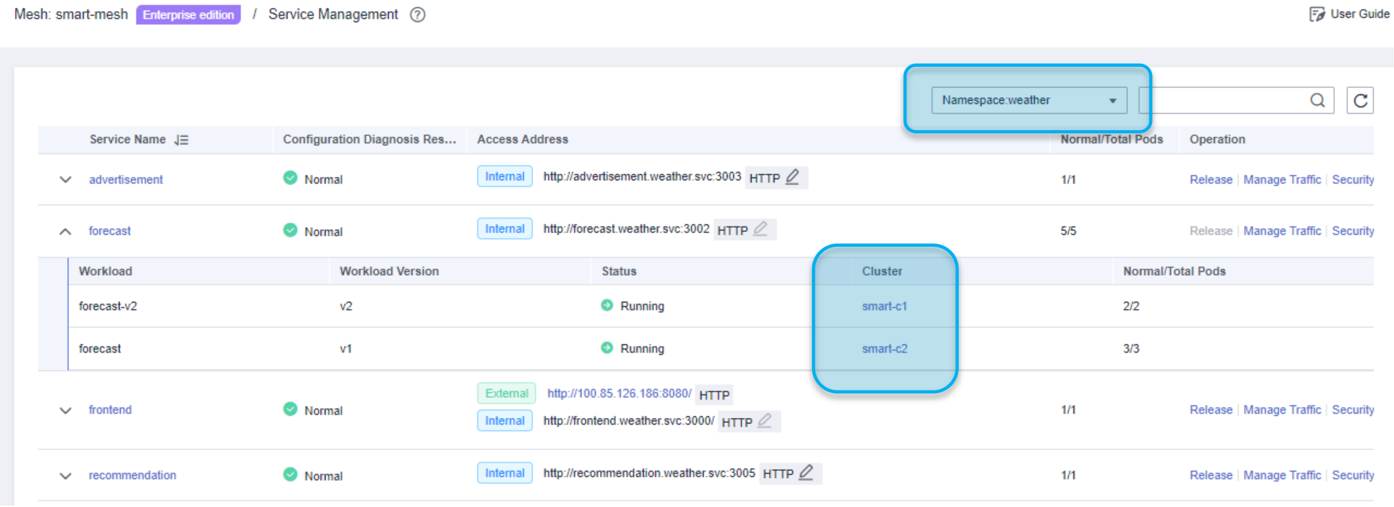
Topology & Metric & Tracing

Global service discovery across multi cluster



- Config and manage one service mesh to span several Kubernetes clusters.
- A Global namespace and service view across all clusters.
- All service endpoints are assumed to be reachable from any consumer in any of clusters of the mesh.
- Traffic is load-balanced across all clusters in the mesh for a given service.

Global service discovery across multi cluster

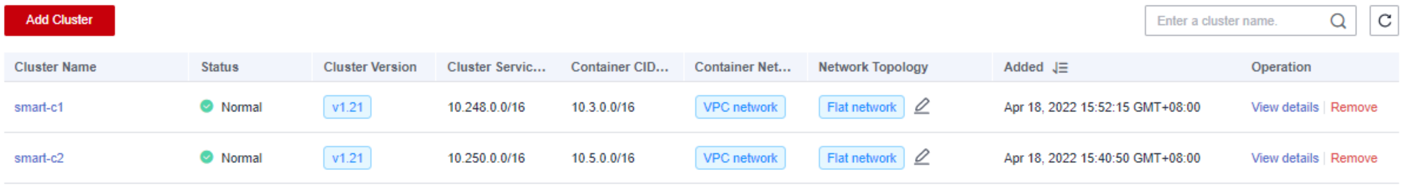


Mesh: smart-mesh Enterprise edition / Service Management

Namespace:weather

Service Name	Configuration Diagnosis Res...	Access Address	Normal/Total Pods	Operation
advertisement	Normal	Internal http://advertisement.weather.svc:3003 HTTP	1/1	Release | Manage Traffic | Security
forecast	Normal	Internal http://forecast.weather.svc:3002 HTTP	5/5	Release | Manage Traffic | Security
frontend	Normal	External http://100.85.126.186:8080/ HTTP Internal http://frontend.weather.svc:3000/ HTTP	1/1	Release | Manage Traffic | Security
recommendation	Normal	Internal http://recommendation.weather.svc:3005 HTTP	1/1	Release | Manage Traffic | Security

Workload	Workload Version	Status	Cluster	Normal/Total Pods
forecast-v2	v2	Running	smart-c1	2/2
forecast	v1	Running	smart-c2	3/3



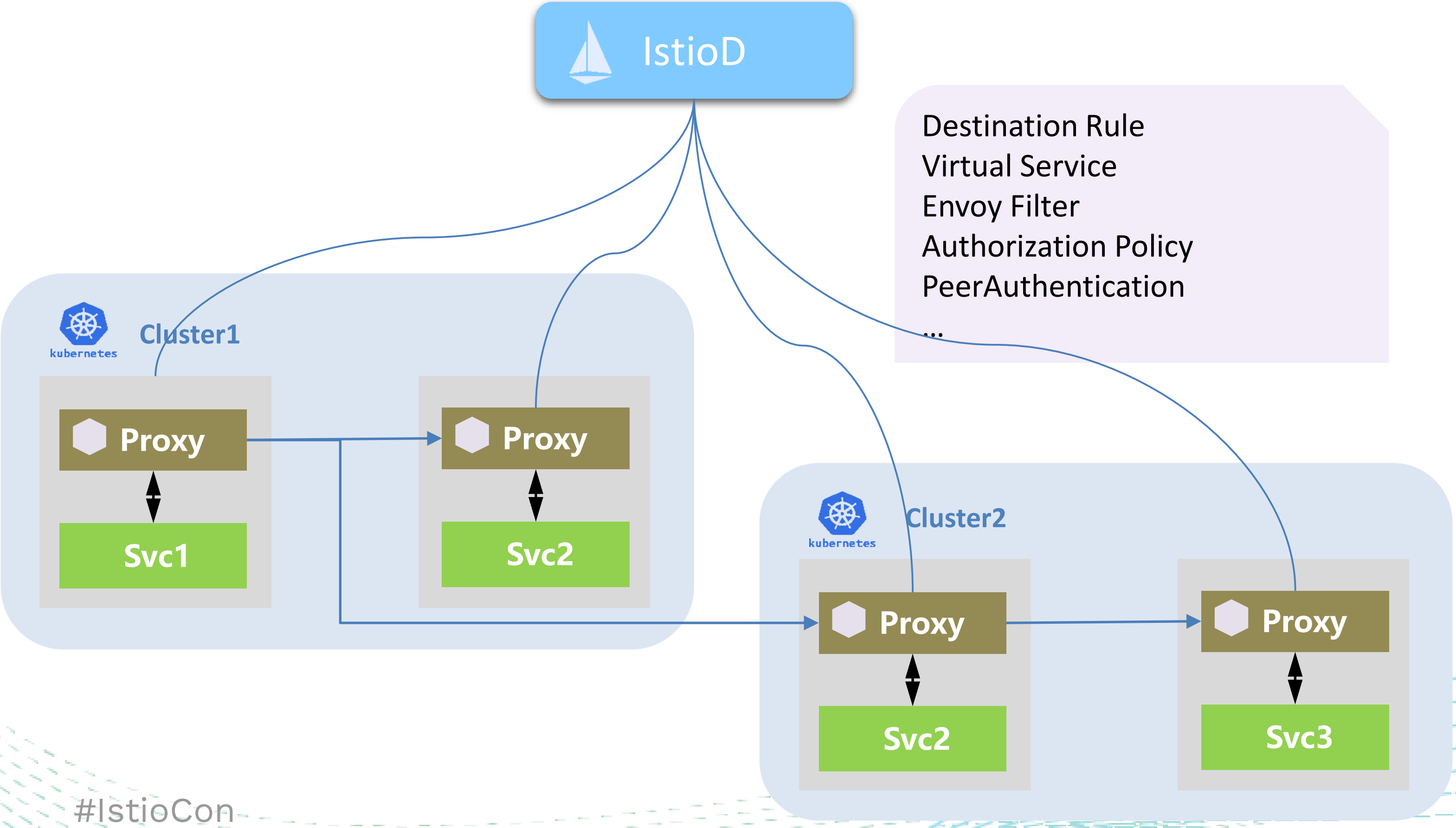
Cluster Name	Status	Cluster Version	Cluster Servic...	Container CID...	Container Net...	Network Topology	Added	Operation
smart-c1	Normal	v1.21	10.248.0.0/16	10.3.0.0/16	VPC network	Flat network	Apr 18, 2022 15:52:15 GMT+08:00	View details | Remove
smart-c2	Normal	v1.21	10.250.0.0/16	10.5.0.0/16	VPC network	Flat network	Apr 18, 2022 15:40:50 GMT+08:00	View details | Remove

(From HUAWEI CLOUD ASM)

- Services in one mesh are organized by namespaces rather than clusters.
- There is no difference in service management between single cluster and multi cluster.

#IstioCon

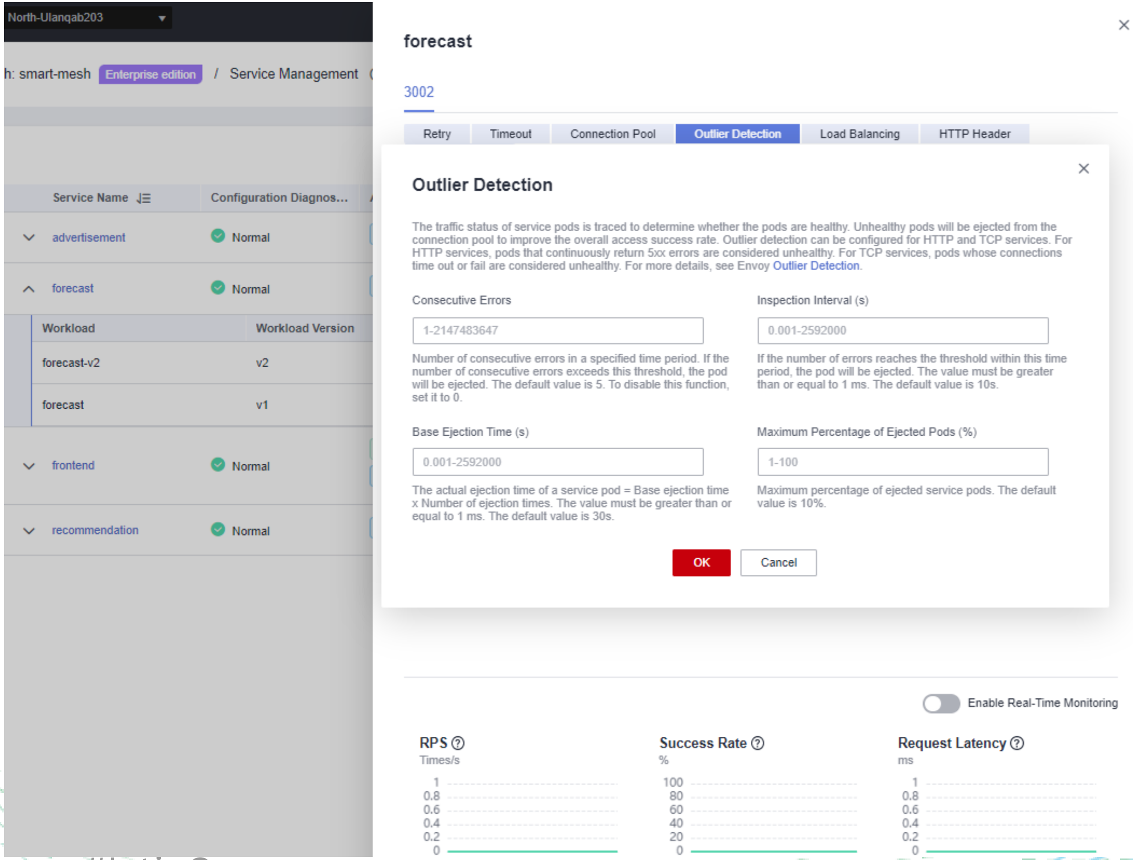
Global Policy Config across clusters



Destination Rule
Virtual Service
Envoy Filter
Authorization Policy
PeerAuthentication
...

#IstioCon

Global Policy Config across clusters



Service Name	Configuration Diagnos...
advertisement	Normal
forecast	Normal
frontend	Normal
recommendation	Normal

Workload	Workload Version
forecast-v2	v2
forecast	v1

Outlier Detection

The traffic status of service pods is traced to determine whether the pods are healthy. Unhealthy pods will be ejected from the connection pool to improve the overall access success rate. Outlier detection can be configured for HTTP and TCP services. For HTTP services, pods that continuously return 5xx errors are considered unhealthy. For TCP services, pods whose connections time out or fail are considered unhealthy. For more details, see Envoy Outlier Detection.

Consecutive Errors: 1-2147483647

Number of consecutive errors in a specified time period. If the number of consecutive errors exceeds this threshold, the pod will be ejected. The default value is 5. To disable this function, set it to 0.

Inspection Interval (s): 0.001-2592000

If the number of errors reaches the threshold within this time period, the pod will be ejected. The value must be greater than or equal to 1 ms. The default value is 10s.

Base Ejection Time (s): 0.001-2592000

The actual ejection time of a service pod = Base ejection time x Number of ejection times. The value must be greater than or equal to 1 ms. The default value is 30s.

Maximum Percentage of Ejected Pods (%): 1-100

Maximum percentage of ejected service pods. The default value is 10%.

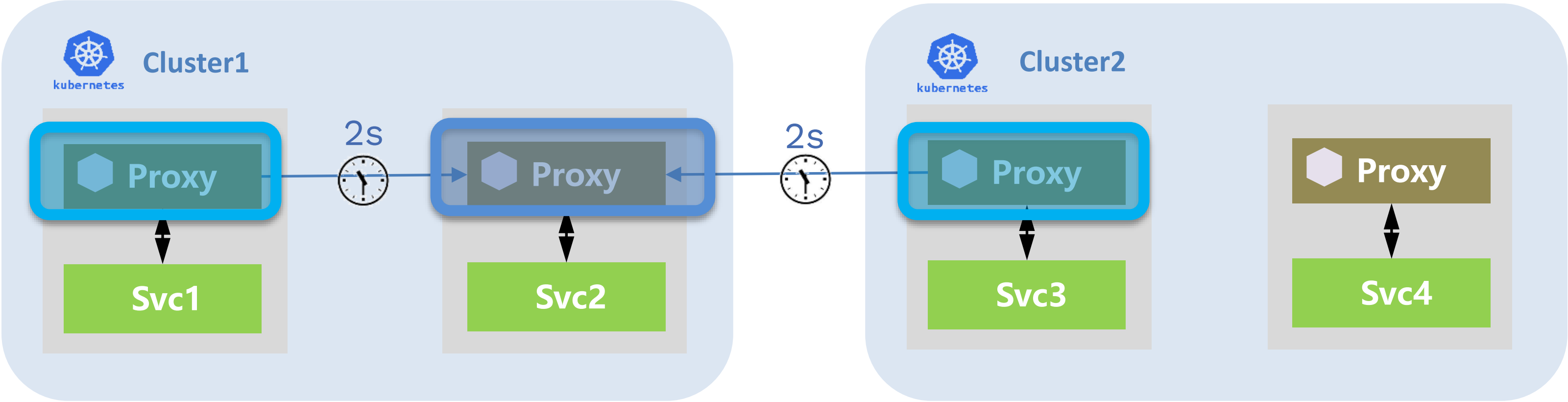
- Config traffic rule for specified service of mesh
- apply to all clusters of the mesh

(From HUAWEI CLOUD ASM)

Global Policy takes effects

Global policy for traffic FROM different cluster

```
spec:
  hosts:
  - svc2
  http:
  - route:
    - destination:
        host: svc2
    timeout: 2s
```

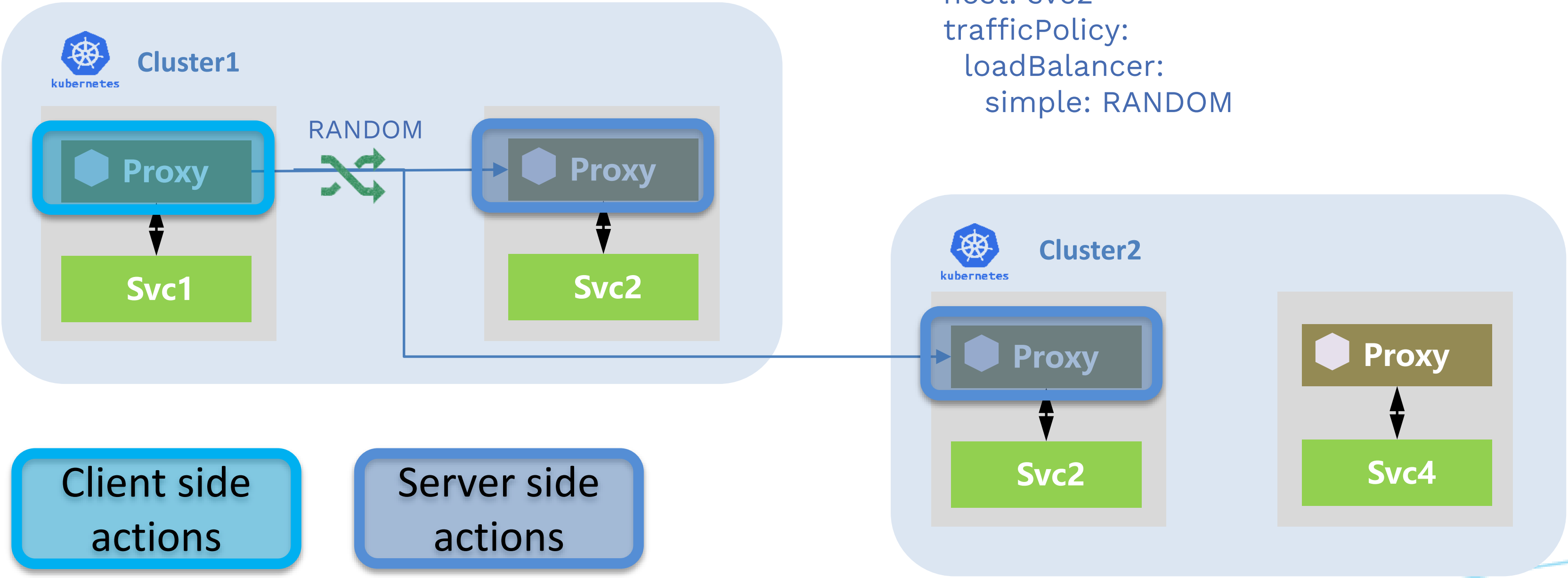


Client side actions

Server side actions

Global Policy takes effects

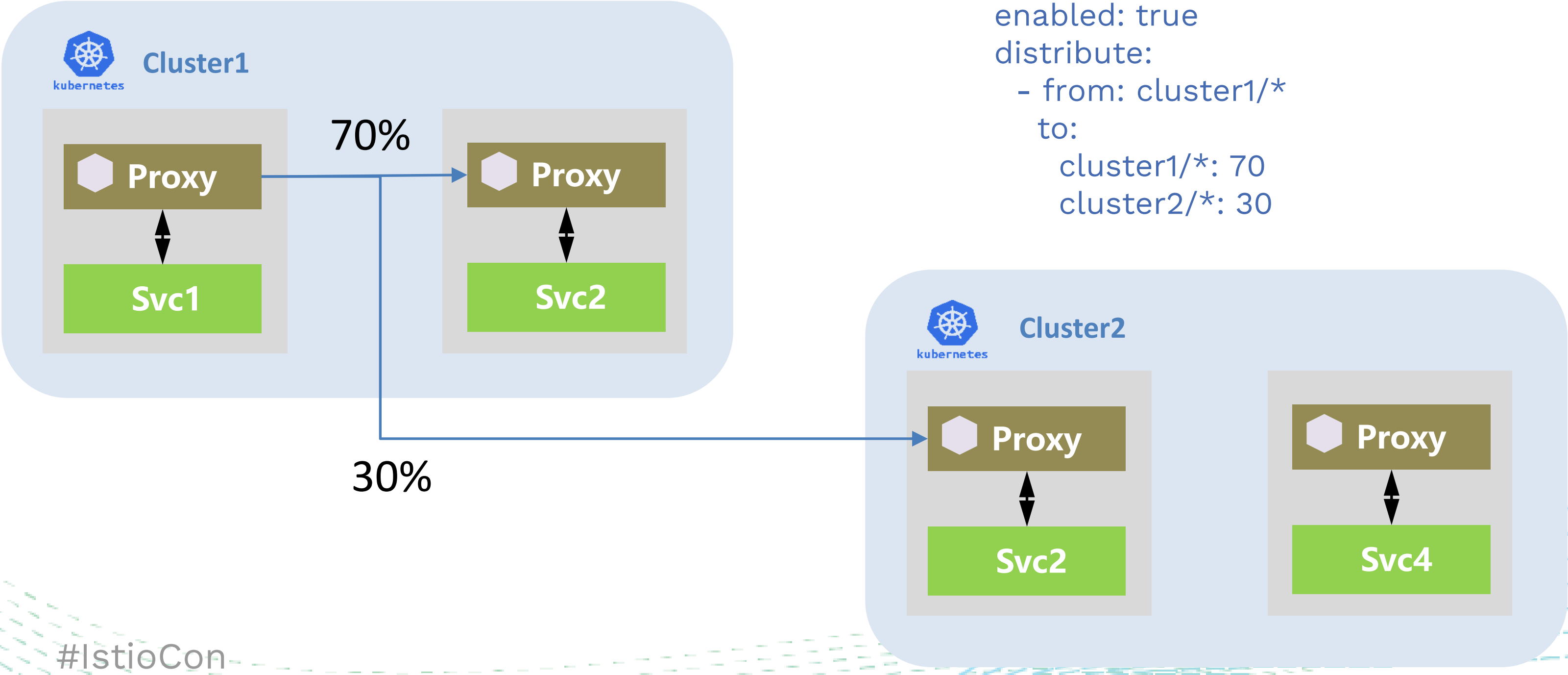
Global policy for traffic TO different cluster



```
spec:
  host: svc2
  trafficPolicy:
    loadBalancer:
      simple: RANDOM
```

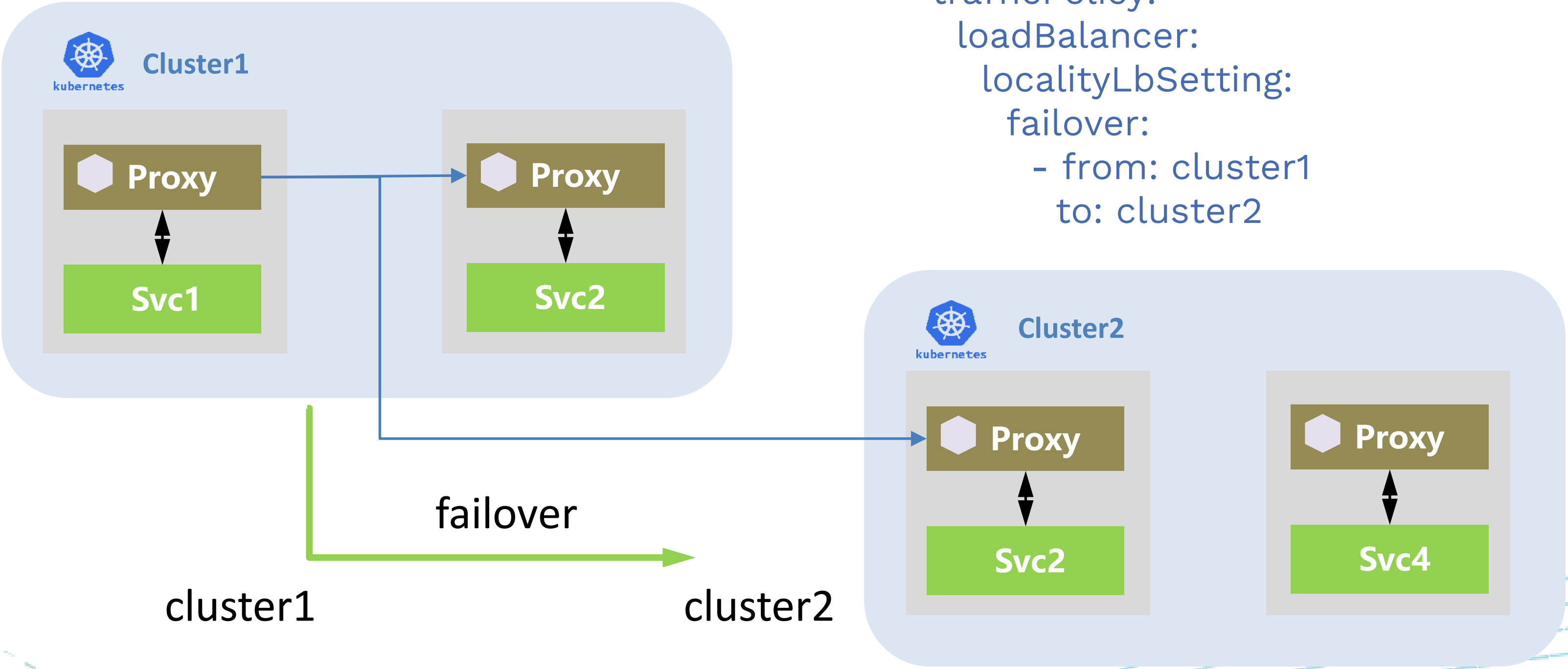
Locality load balance across clusters

```
spec:
  trafficPolicy:
    loadBalancer:
      localityLbSetting:
        enabled: true
        distribute:
          - from: cluster1/*
            to:
              cluster1/*: 70
              cluster2/*: 30
```



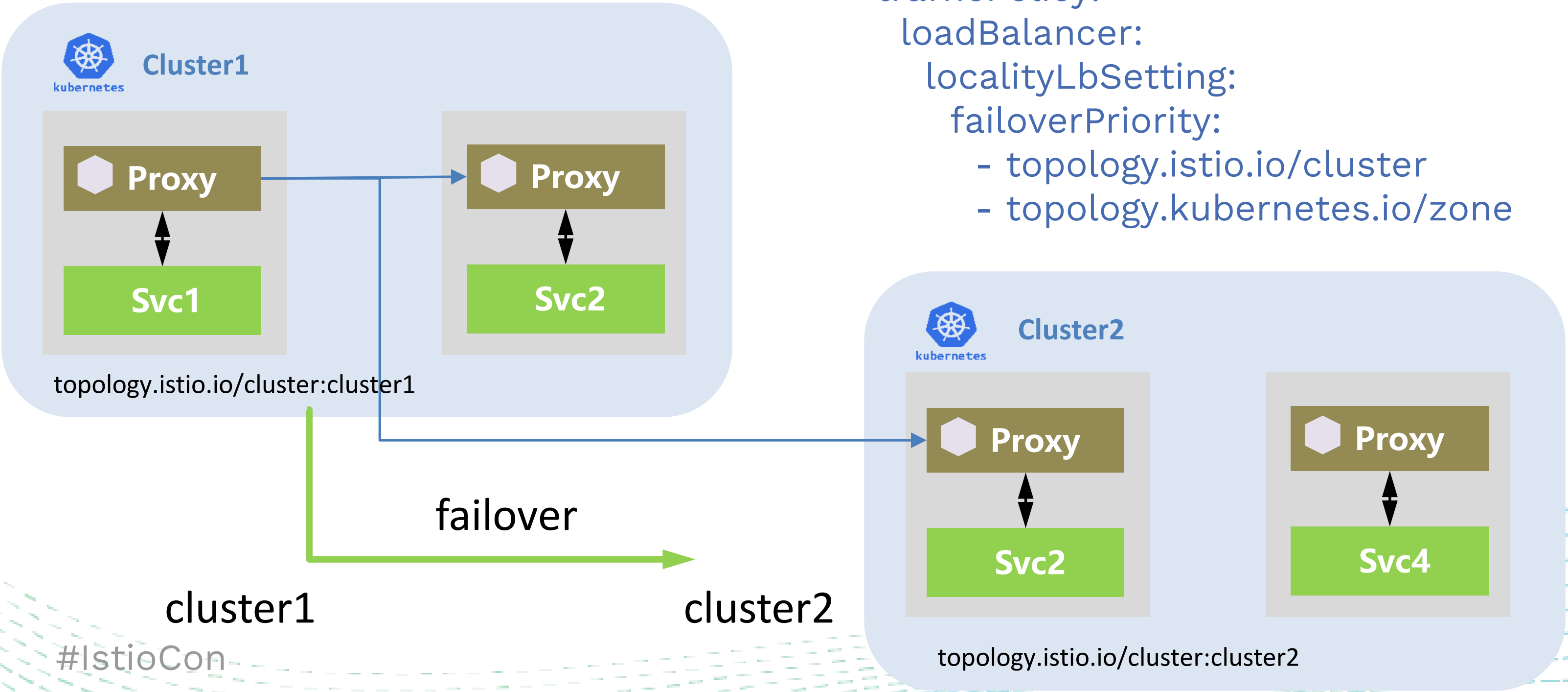
Locality failover across clusters

```
spec:
  trafficPolicy:
    loadBalancer:
      localityLbSetting:
        failover:
          - from: cluster1
            to: cluster2
```



Locality failover across clusters with priority

```
spec:
  trafficPolicy:
    loadBalancer:
      localityLbSetting:
        failoverPriority:
          - topology.istio.io/cluster
          - topology.kubernetes.io/zone
```



#IstioCon

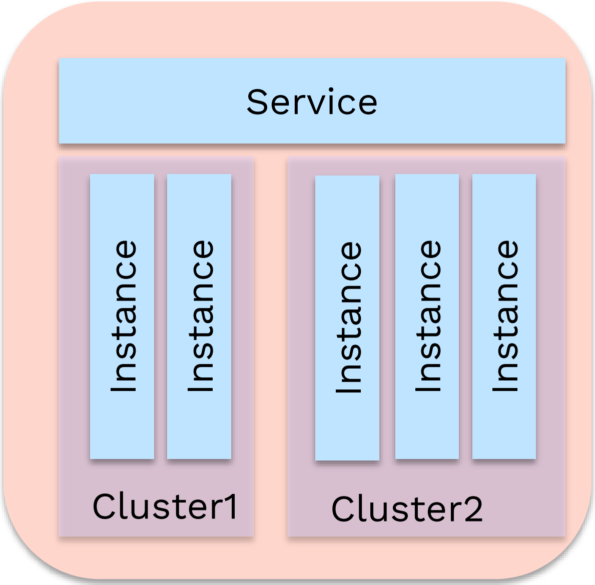
Locality failover across clusters with priority

Local instance healthy

P=0 healthy endpoints	P=1 healthy endpoints	P=2 healthy endpoints	Traffic to P=0	Traffic to P=1	Traffic to P=2
100%	100%	100%	100%	0%	0%
72%	72%	100%	100%	0%	0%
71%	71%	100%	99%	1%	0%
50%	50%	100%	70%	30%	0%
25%	100%	100%	35%	65%	0%
25%	25%	100%	35%	35%	30%
25%	25%	20%	36%	36%	28%

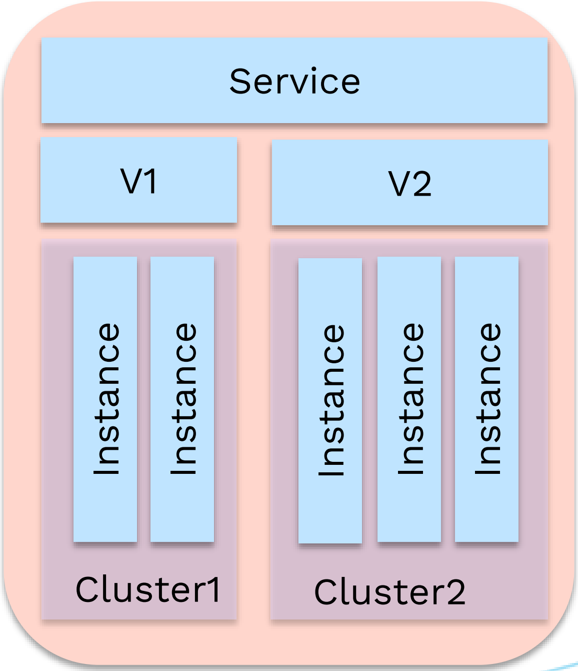
Partitioning Services by cluster

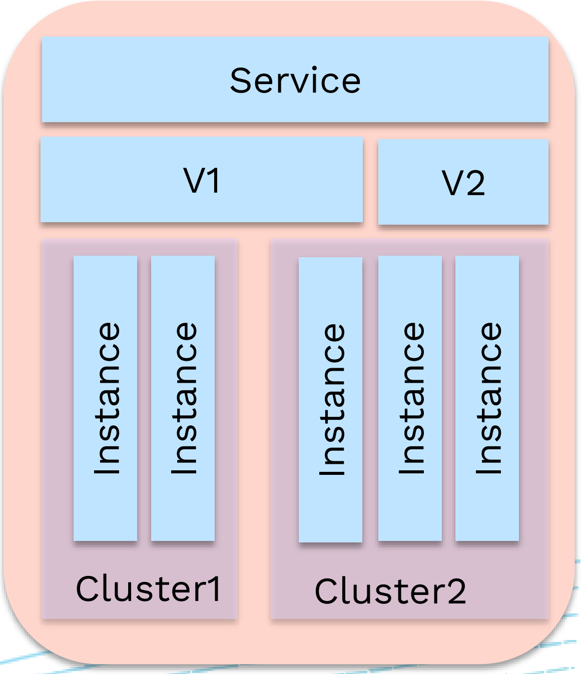
```
apiVersion: networking.istio.io/v1beta1
kind: DestinationRule
metadata:
  name: cas-dr
spec:
  host: cassvc.mss.svc.cluster.local
  subsets:
  - name: cluster1
    labels:
      topology.istio.io/cluster: cluster1
  - name: cluster2
    labels:
      topology.istio.io/cluster: cluster2
```



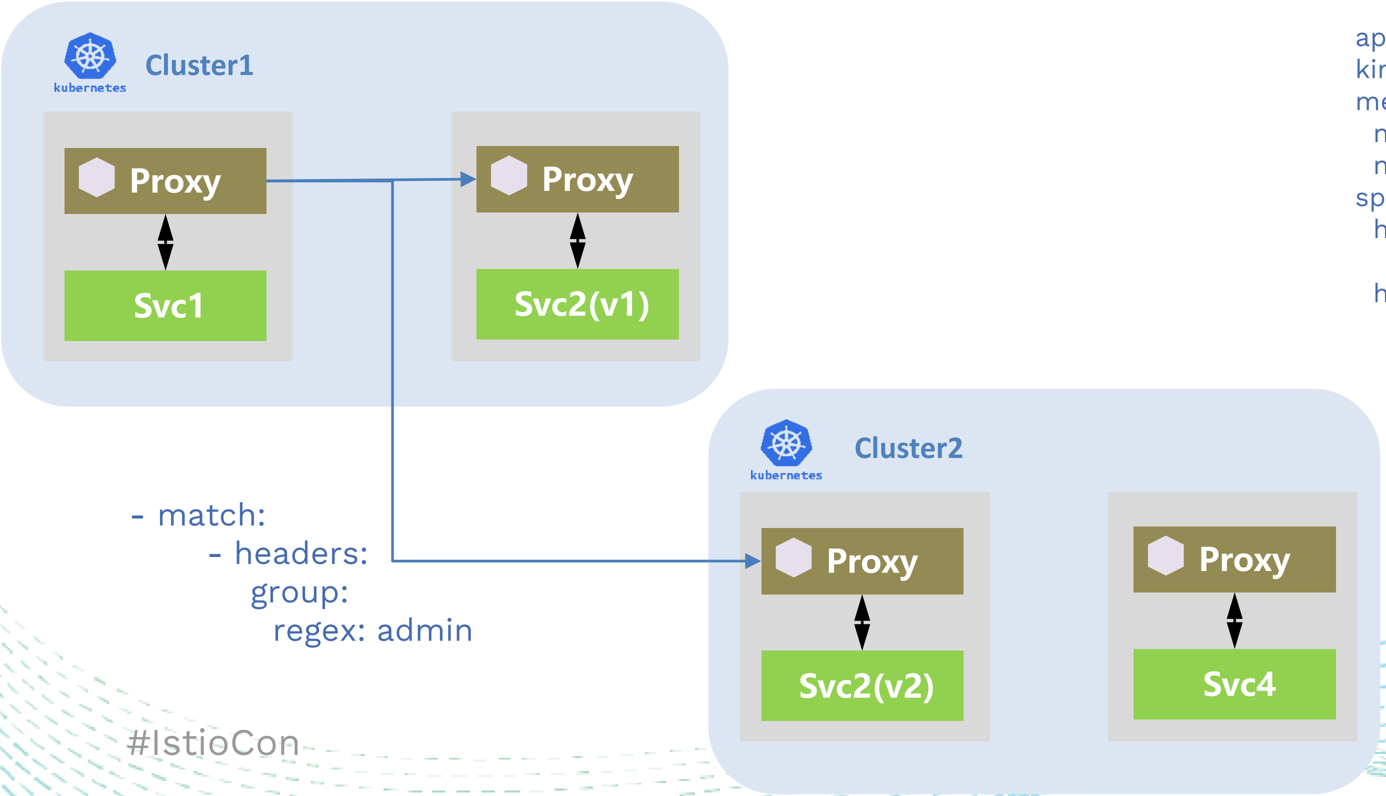
Partitioning Services by version

```
apiVersion: networking.istio.io/v1beta1
kind: DestinationRule
metadata:
  name: sms-dr
spec:
  host: smssvc.mss.svc.cluster.local
  subsets:
  - name: v1
    labels:
      version: V1
  - name: v2
    labels:
      version: v2
```





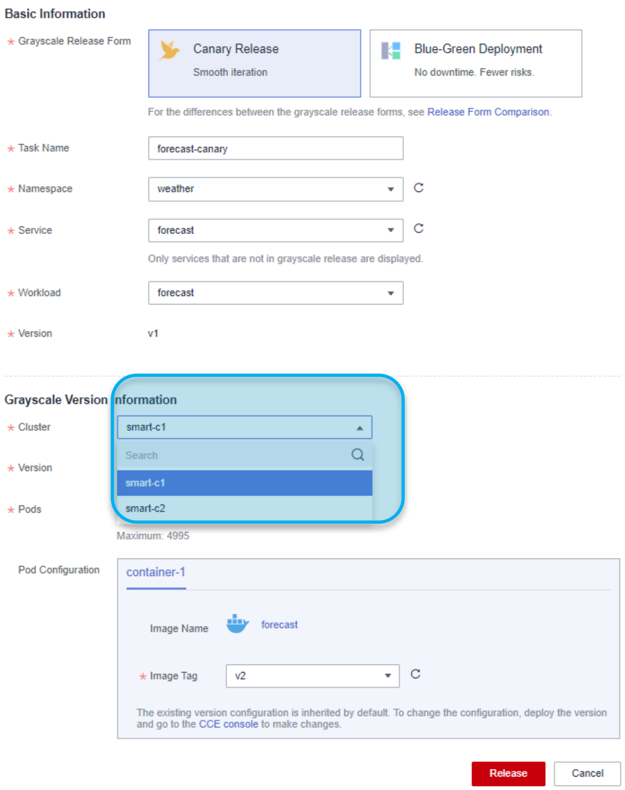
Canary enable split traffic between clusters



```
- match:
    - headers:
        group:
          regex: admin
```

```
apiVersion: networking.istio.io/v1beta1
kind: VirtualService
metadata:
  name: canary-vs
  namespace: mss
spec:
  hosts:
    - svc2
  http:
    - match:
        - headers:
            group:
              regex: admin
      route:
        - destination:
            host: svc2.mss.svc.cluster.local
            subset: v2
    - route:
        - destination:
            host: svc2.mss.svc.cluster.local
            subset: v1
```

Canary enable split traffic between clusters



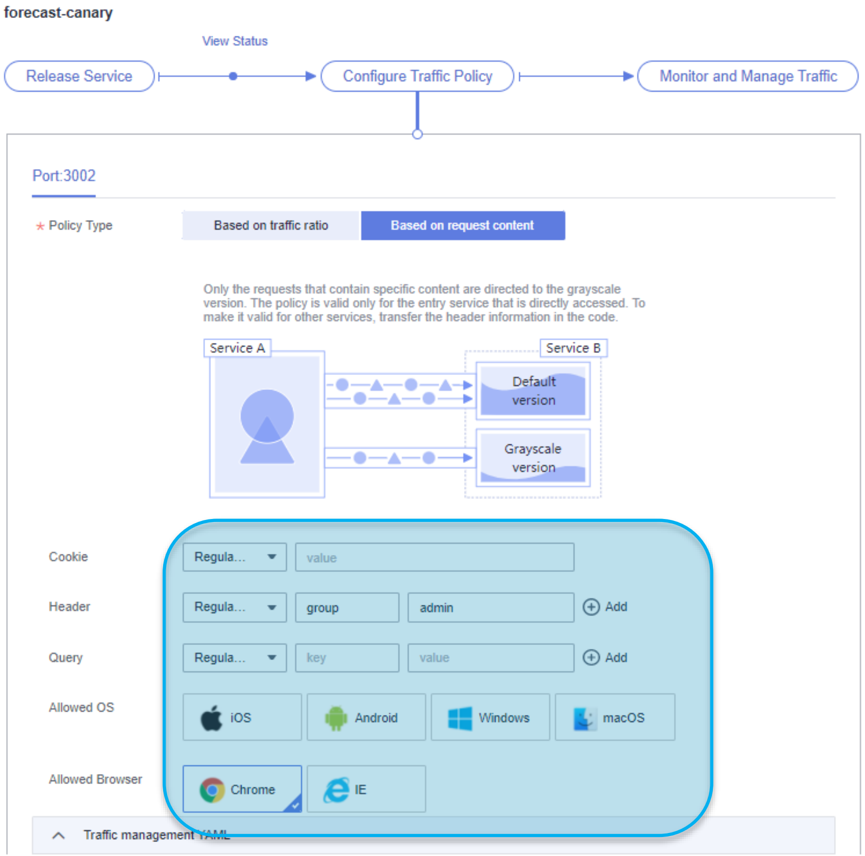
Basic Information

* Grayscale Release Form	Canary Release Smooth iteration Blue-Green Deployment No downtime. Fewer risks.
	For the differences between the grayscale release forms, see Release Form Comparison.
* Task Name	forecast-canary
* Namespace	weather
* Service	forecast
	Only services that are not in grayscale release are displayed.
* Workload	forecast
* Version	v1

Grayscale Version Information

* Cluster	smart-c1 Search smart-c1 smart-c2
* Version	
* Pods	
	Maximum: 4995
Pod Configuration	container-1
Image Name	forecast
* Image Tag	v2

The existing version configuration is inherited by default. To change the configuration, deploy the version and go to the CCE console to make changes.



forecast-canary

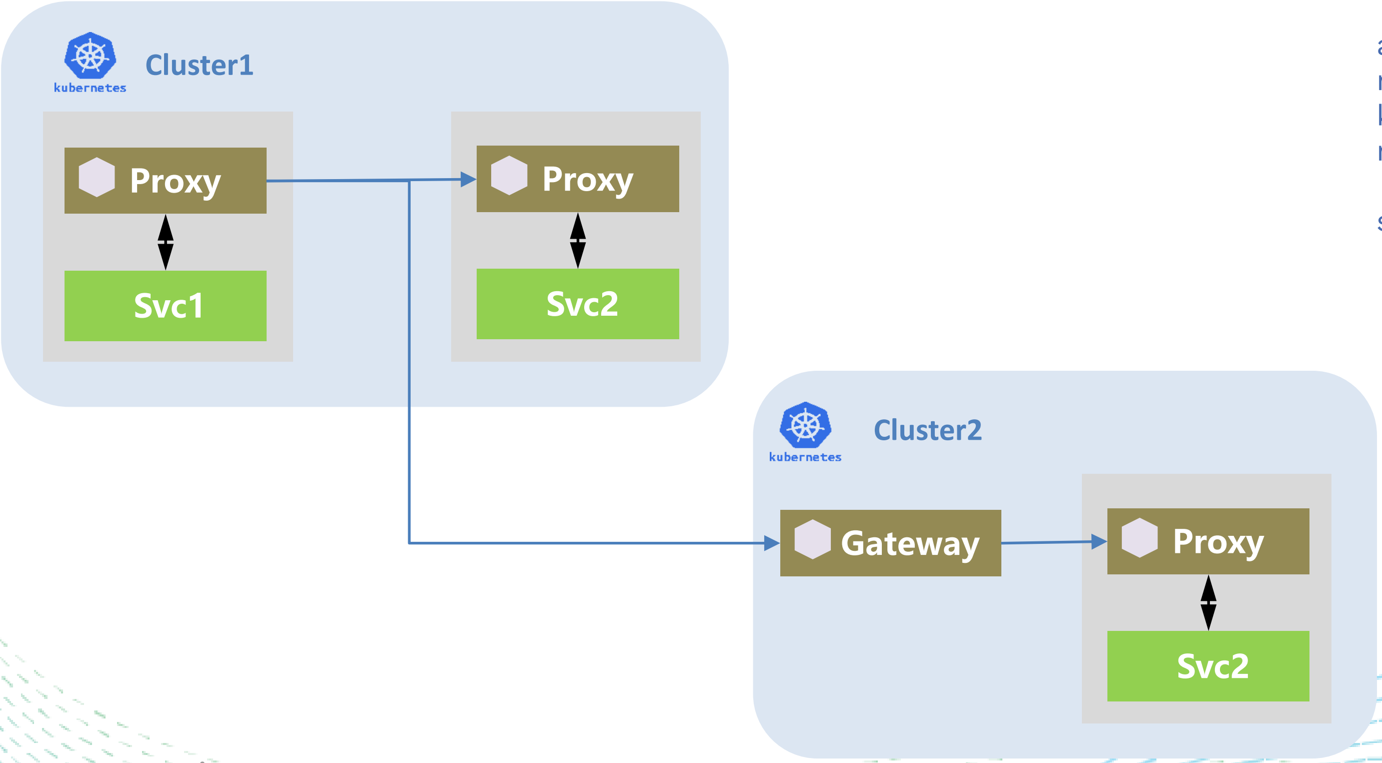
* Policy Type	Based on traffic ratio Based on request content

Only the requests that contain specific content are directed to the grayscale version. The policy is valid only for the entry service that is directly accessed. To make it valid for other services, transfer the header information in the code.

Cookie	Regula...	value	
Header	Regula...	group	admin
Query	Regula...	key	value
Allowed OS	iOS	Android	Windows macOS
Allowed Browser	Chrome	IE	

(From HUAWEI CLOUD ASM)

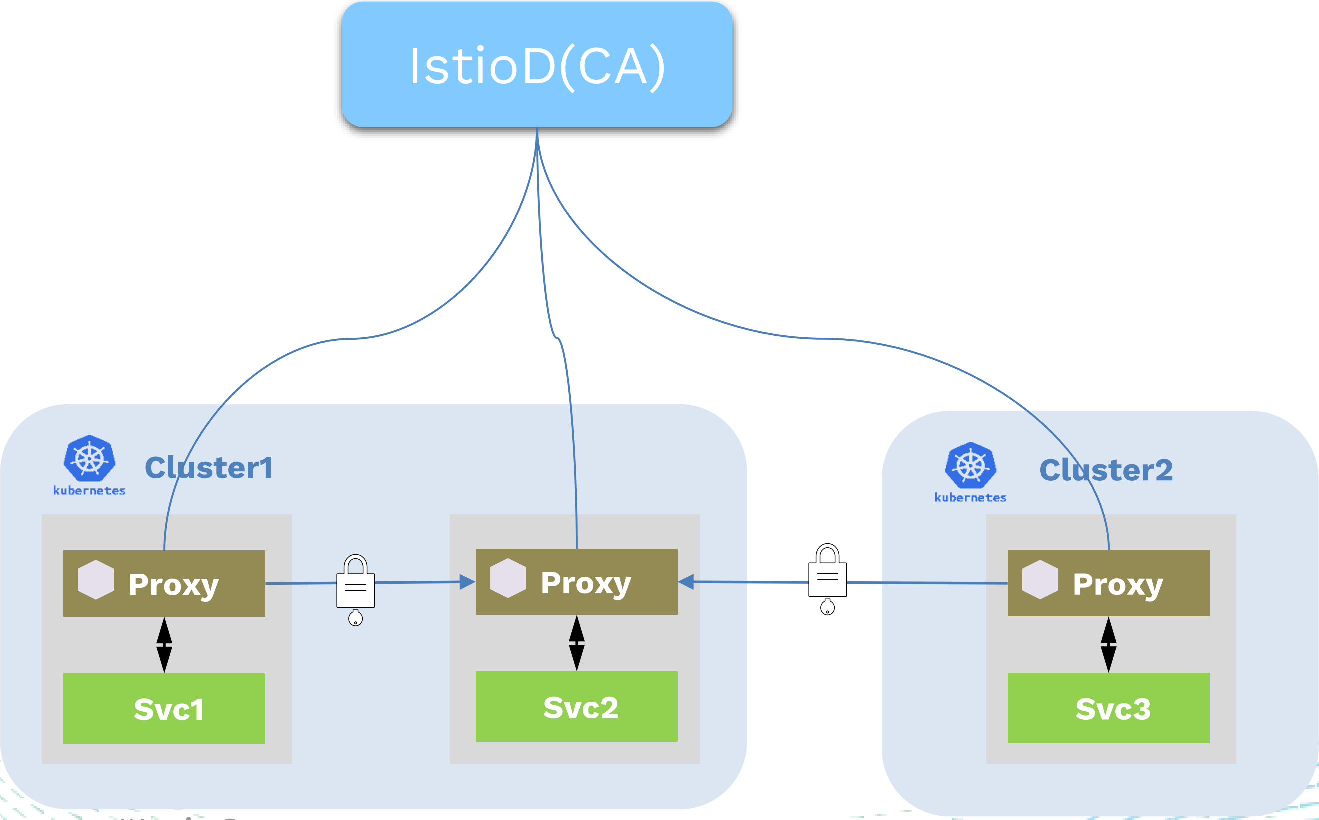
Multiple cluster on multiple networks



```
apiVersion:
networking.istio.io/v1alpha3
kind: Gateway
metadata:
  name: cross-network-gateway
spec:
  selector:
    istio: eastwestgateway
  servers:
    - port:
        number: 15443
        name: tls
        protocol: TLS
      tls:
        mode: AUTO_PASSTHROUGH
      hosts:
        - "*.local"
```

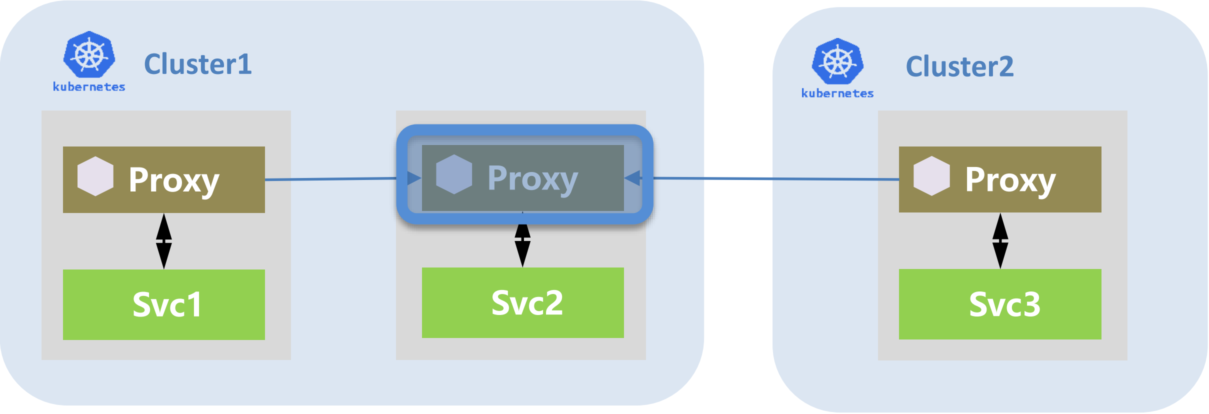
#IstioCon

Transparent authentication cross cluster



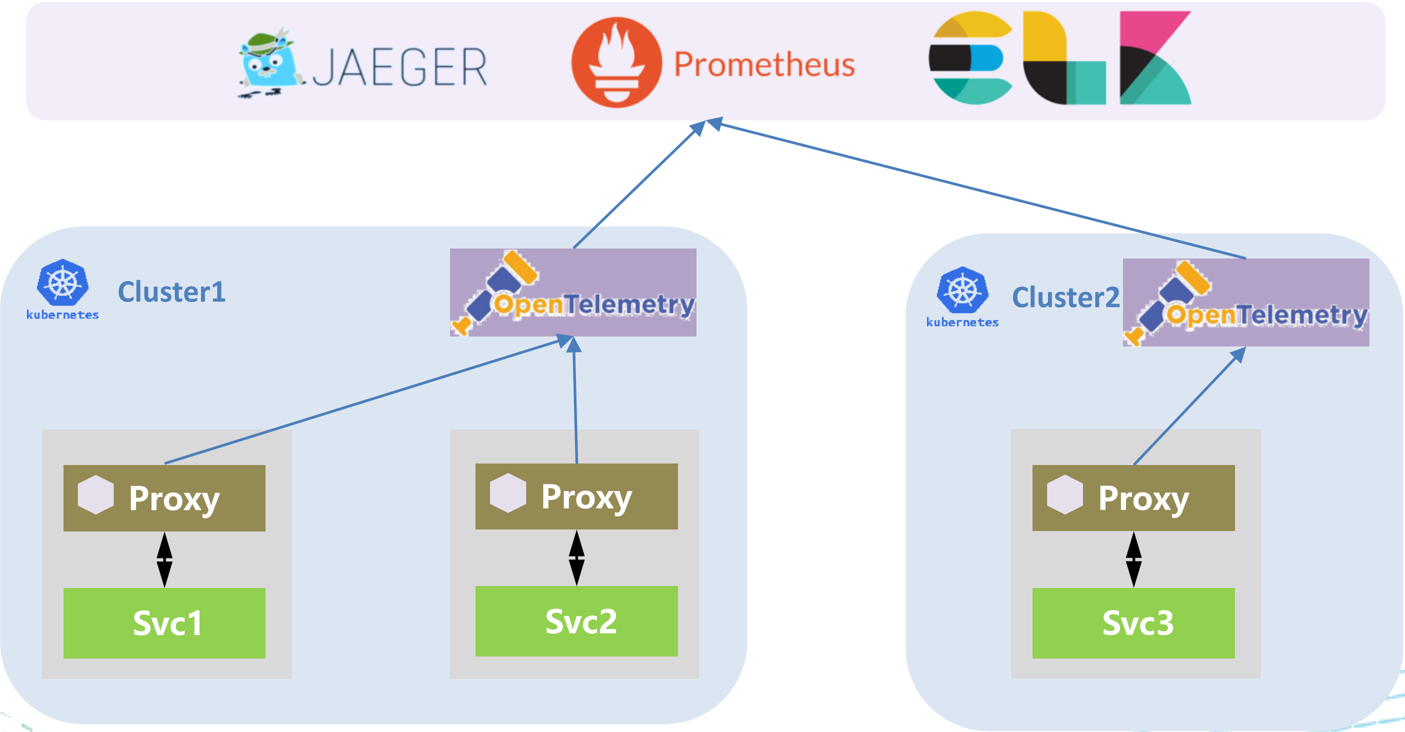
```
apiVersion: security.istio.io/v1beta1
kind: PeerAuthentication
metadata:
  name: "peer-authn"
  namespace: "mss"
spec:
  selector:
    matchLabels:
      app: svc2
  mtls:
    mode: STRICT
```

Global authority across clusters

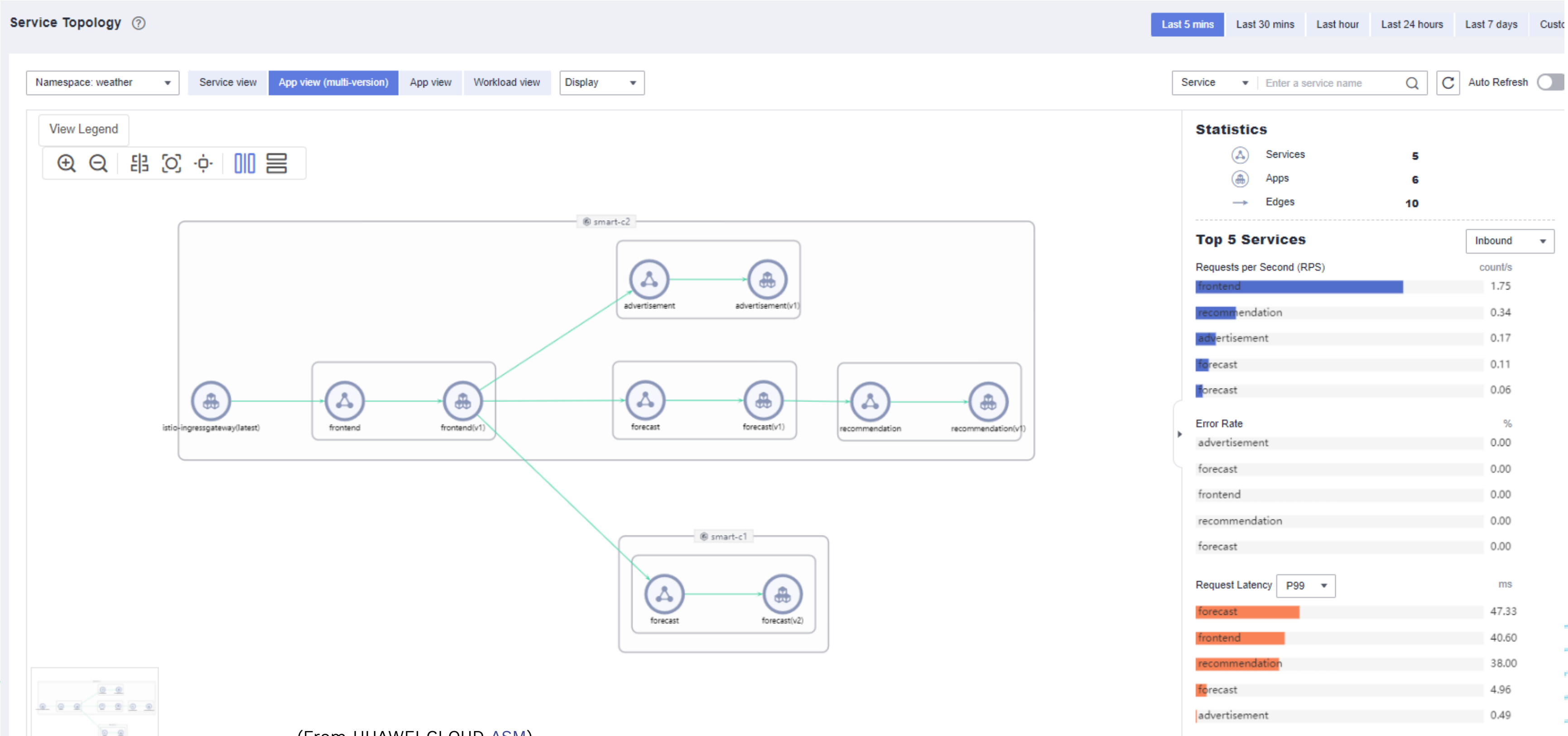


```
apiVersion: security.istio.io/v1beta1
kind: AuthorizationPolicy
metadata:
  name: svc2-auth
  namespace: mss
spec:
  selector:
    matchLabels:
      app: svc2
  action: ALLOW
  rules:
    - from:
        - source:
            principals:
              - cluster.local/ns/mss/sa/svc1
              - cluster.local/ns/mss/sa/svc3
      to:
        - operation:
            methods:
              - POST
```

Global observability



Global observability



Service Topology

Last 5 mins | Last 30 mins | Last hour | Last 24 hours | Last 7 days | Custo

Namespace: weather

Service view | App view (multi-version) | App view | Workload view | Display

Service | Enter a service name | Auto Refresh

View Legend

Statistics

Services	5
Apps	6
Edges	10

Top 5 Services Inbound

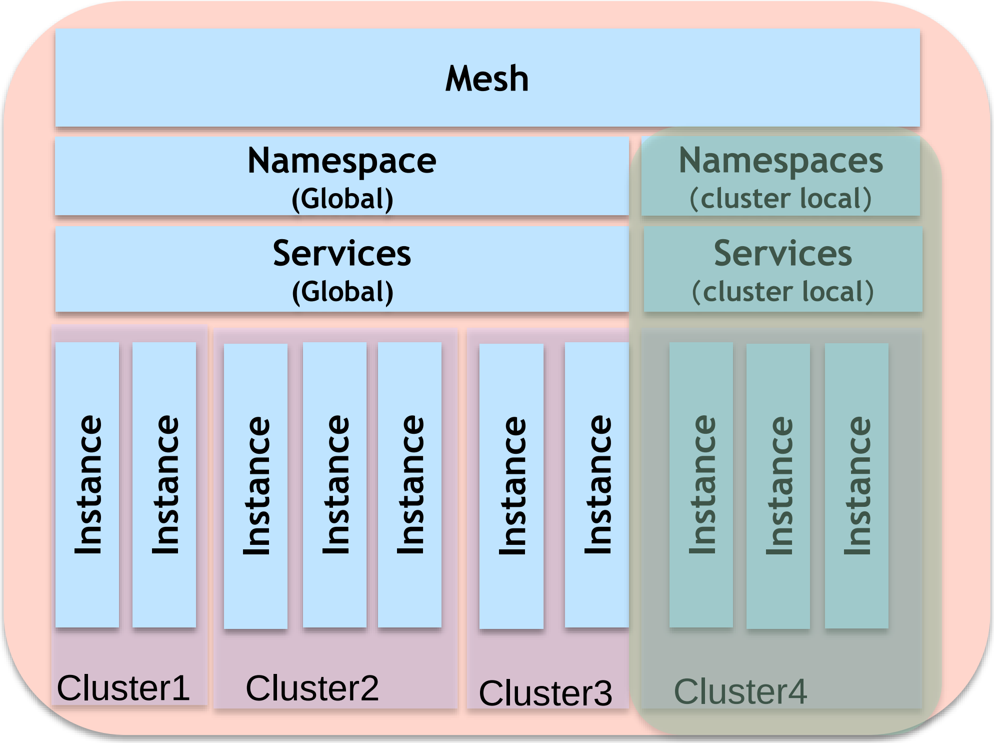
Requests per Second (RPS)	count/s
frontend	1.75
recommendation	0.34
advertisement	0.17
forecast	0.11
forecast	0.06

Error Rate	%
advertisement	0.00
forecast	0.00
frontend	0.00
recommendation	0.00
forecast	0.00

Request Latency P99	ms
forecast	47.33
frontend	40.60
recommendation	38.00
forecast	4.96
advertisement	0.49

(From HUAWEI CLOUD ASM)

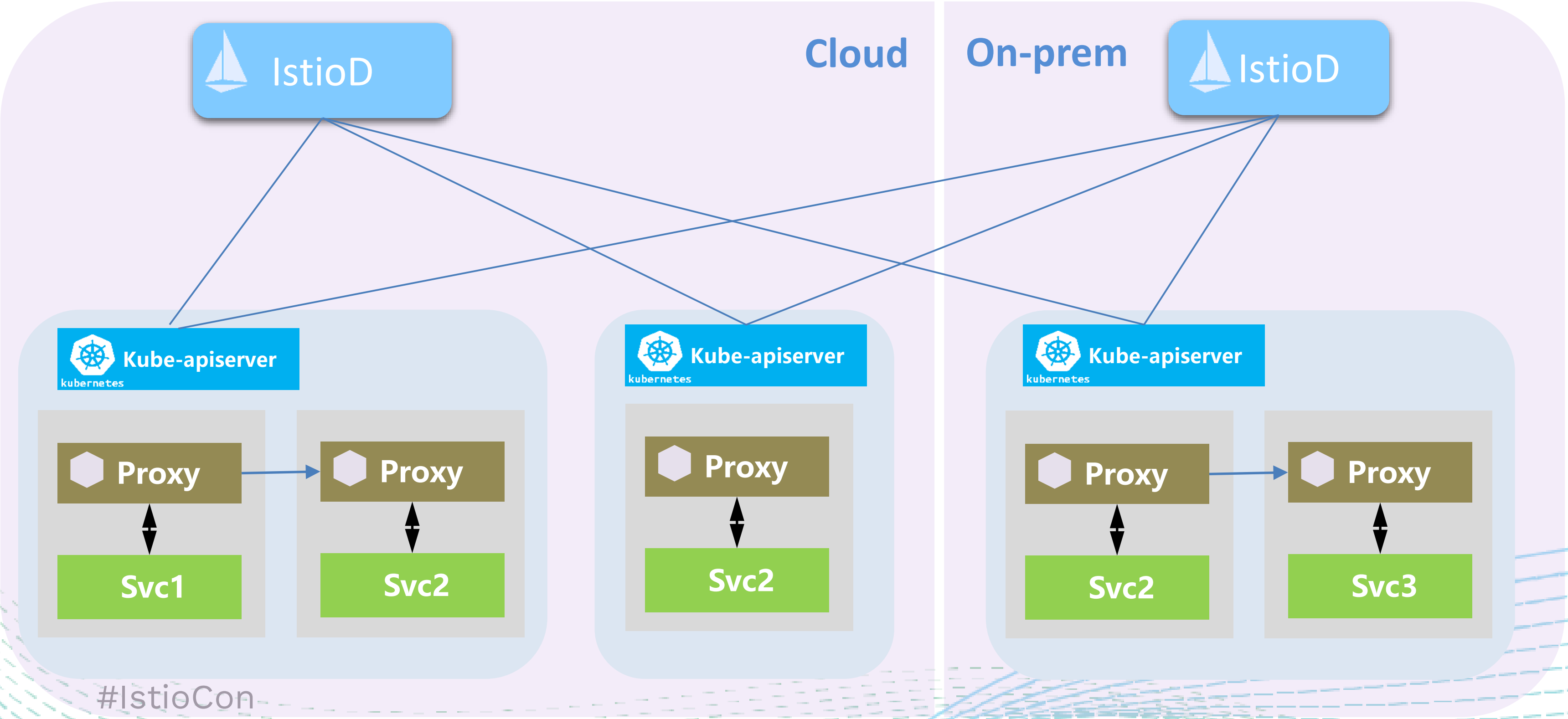
Cluster local Traffic



```
serviceSettings:
- settings:
    clusterLocal: true
  hosts:
  - "*.wechat.svc.cluster.local"
```

Limits the set of service endpoints visible to a client to be cluster scoped.

Next，Mesh for hybrid cloud



Thank you!