Alibaba Cloud

Alibaba Cloud Service Mesh
Product Introduction

Alibaba Cloud

Legal disclaimer

Alibaba Cloud reminds you to carefully read and fully understand the terms and conditions of this legal disclaimer before you read or use this document. If you have read or used this document, it shall be deemed as your total acceptance of this legal disclaimer.

1. You shall download and obtain this document from the Alibaba Cloud website or other Alibaba Cloud-authorized channels, and use this document for your own legal business activities only. The content of this document is considered confidential information of Alibaba Cloud. You shall strictly abide by the confidentiality obligations. No part of this document shall be disclosed or provided to any third party for use without the prior written consent of Alibaba Cloud.

2. No part of this document shall be excerpted, translated, reproduced, transmitted, or disseminated by any organization, company or individual in any form or by any means without the prior written consent of Alibaba Cloud.

3. The content of this document may be changed because of product version upgrade, adjustment, or other reasons. Alibaba Cloud reserves the right to modify the content of this document without notice and an updated version of this document will be released through Alibaba Cloud-authorized channels from time to time. You should pay attention to the version changes of this document as they occur and download and obtain the most up-to-date version of this document from Alibaba Cloud-authorized channels.

4. This document serves only as a reference guide for your use of Alibaba Cloud products and services. Alibaba Cloud provides this document based on the "status quo", "being defective", and "existing functions" of its products and services. Alibaba Cloud makes every effort to provide relevant operational guidance based on existing technologies. However, Alibaba Cloud hereby makes a clear statement that it in no way guarantees the accuracy, integrity, applicability, and reliability of the content of this document, either explicitly or implicitly. Alibaba Cloud shall not take legal responsibility for any errors or lost profits incurred by any organization, company, or individual arising from download, use, or trust in this document. Alibaba Cloud shall not, under any circumstances, take responsibility for any indirect, consequential, punitive, contingent, special, or punitive damages, including lost profits arising from the use or trust in this document (even if Alibaba Cloud has been notified of the possibility of such a loss).

5. By law, all the contents in Alibaba Cloud documents, including but not limited to pictures, architecture design, page layout, and text description, are intellectual property of Alibaba Cloud and/or its affiliates. This intellectual property includes, but is not limited to, trademark rights, patent rights, copyrights, and trade secrets. No part of this document shall be used, modified, reproduced, publicly transmitted, changed, disseminated, distributed, or published without the prior written consent of Alibaba Cloud and/or its affiliates. The names owned by Alibaba Cloud shall not be used, published, or reproduced for marketing, advertising, promotion, or other purposes without the prior written consent of Alibaba Cloud. The names owned by Alibaba Cloud include, but are not limited to, "Alibaba Cloud", "Aliyun", "HiChina", and other brands of Alibaba Cloud and/or its affiliates, which appear separately or in combination, as well as the auxiliary signs and patterns of the preceding brands, or anything similar to the company names, trade names, trademarks, product or service names, domain names, patterns, logos, marks, signs, or special descriptions that third parties identify as Alibaba Cloud and/or its affiliates.

6. Please directly contact Alibaba Cloud for any errors of this document.

Document conventions

| Style	Description	Example
⚠ Danger	A danger notice indicates a situation that will cause major system changes, faults, physical injuries, and other adverse results.	⚠ **Danger:** Resetting will result in the loss of user configuration data.
🔔 Warning	A warning notice indicates a situation that may cause major system changes, faults, physical injuries, and other adverse results.	🔔 **Warning:** Restarting will cause business interruption. About 10 minutes are required to restart an instance.
🔊 Notice	A caution notice indicates warning information, supplementary instructions, and other content that the user must understand.	🔊 **Notice:** If the weight is set to 0, the server no longer receives new requests.
❓ Note	A note indicates supplemental instructions, best practices, tips, and other content.	❓ **Note:** You can use Ctrl + A to select all files.
>	Closing angle brackets are used to indicate a multi-level menu cascade.	Click **Settings> Network> Set network type**.
Bold	Bold formatting is used for buttons , menus, page names, and other UI elements.	Click **OK**.
Courier font	Courier font is used for commands	Run the `cd /d C:/window` command to enter the Windows system folder.
Italic	Italic formatting is used for parameters and variables.	`bae log list --instanceid` *Instance_ID*
[] or [a|b]	This format is used for an optional value, where only one item can be selected.	`ipconfig [-all|-t]`
{} or {a|b}	This format is used for a required value, where only one item can be selected.	`switch {active|stand}`

Table of Contents

1.What is ASM?

Alibaba Cloud Service Mesh (ASM) is a fully managed service mesh platform. ASM is compatible with open source Istio. ASM allows you to manage services in a simplified manner. For example, you can use ASM to route and split inter-service traffic, secure inter-service communication with authentication, and observe the behavior of services in meshes. This greatly reduces your workload in development and O&M.

Service architecture

The following figure shows the architecture of ASM.
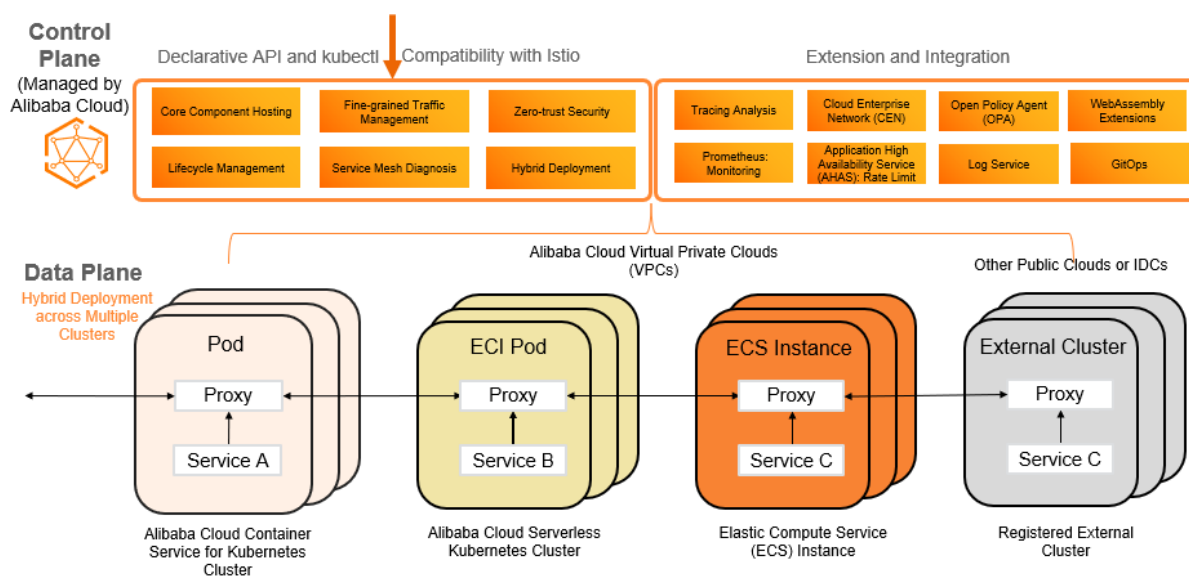
ASM integrates and manages all components on the Istio control plane to simplify your use of ASM. This way, you can focus on application development and deployment. In addition, ASM is compatible with open source Istio. You can use declarative parameters to define flexible routing rules, and centrally manage traffic between services in a mesh.

An ASM instance with the managed control plane supports application services from multiple Kubernetes clusters or application services that are run in pods of Elastic Container Instance. You can also deploy non-Kubernetes services, such as services that run on VMs or bare metal hosts, into the same service mesh.

Functions and features

ASM builds managed and unified service mesh capabilities in core scenarios, such as hybrid cloud, multi-cloud, multi-cluster, and non-containerized application migration. ASM provides the following benefits:

- Centralized management mode

  ASM manages application services that run in managed, dedicated, and serverless clusters of Container Service for Kubernetes (ACK) and registered clusters in hybrid cloud and multi-cloud environments in a centralized manner. This provides unified observability and throttling for application services.

- Centralized traffic management

  ASM centrally manages the traffic between application services that are deployed in a hybrid environment in which containers and virtual machines coexist.

- Managed core components of the control plane

  ASM manages core components of the Istio control plane. This helps minimize your resource overhead and O&M costs.

Instance editions

ASM provides four editions of instances: Standard Edition, Premium Edition, Enterprise Edition, and Ultimate Edition. Standard Edition is a free edition, and the other editions are commercial editions. Premium Edition, Enterprise Edition, and Ultimate Edition extend the features of Standard Edition to enhance multi-protocol support and dynamic extension capabilities, provide fine-grained service governance, and improve the zero-trust security system. In addition, the three commercial editions enhance performance, provide better support for large-scale clusters, and simplify the use of ASM instances in production environments. Premium Edition, Enterprise Edition, and Ultimate Edition are applicable to scenarios in which you require cross-language interoperability and fine-grained service governance and want to apply the service mesh technology in production environments on a large scale.

You can select the edition that you require based on the following feature comparison.

> ⓘ **Note**   ASM allows you to update your ASM instances. For more information, see 升级ASM实例.

Feature	Standard Edition	Premium Edition	Enterprise Edition	Ultimate Edition
Core features of the open source Istio, such as traffic routing, traffic governance, observability, and security	✔☺	✔☺	✔☺	✔☺
Centralized management of ACK clusters, Serverless Kubernetes (ASK) clusters, clusters deployed on Elastic Compute Service (ECS) instances, and Kubernetes clusters that are not on Alibaba Cloud	✔☺	✔☺	✔☺	✔☺
Visualized console	✔☺	✔☺	✔☺	✔☺
Envoy filter marketplace	✔☺	✔☺	✔☺	✔☺

Feature	Standard Edition	Premium Edition	Enterprise Edition	Ultimate Edition
Selective service discovery and automatic sidecar recommendation based on access log analysis for configuration optimization	✔☺	✔☺	✔☺	✔☺
Local throttling	✔☺	✔☺	✔☺	✔☺
Mesh diagnosis	✔☺	✔☺	✔☺	✔☺
Connection to a service registry	✔☺	✔☺	✔☺	✔☺
TLS acceleration based on Multi-Buffer	▯	✔☺	✔☺	✔☺
Advanced features of ASM gateways	▯	✔☺	✔☺	✔☺
Support for the Dubbo protocol	▯	✔☺	✔☺	✔☺
Graceful start and shutdown of services	▯	✔☺	✔☺	✔☺
End-to-End tag-based routing	▯	✔☺	✔☺	✔☺
Application High Availability Service (AHAS) for traffic protection	▯	✔☺	✔☺	✔☺
Support for Spring Cloud services	▯	✔☺	✔☺	✔☺

Feature	Standard Edition	Premium Edition	Enterprise Edition	Ultimate Edition
Supported scale	We recommend that you use this edition if you have a maximum of 50 pods in an ASM instance. We recommend that you do not use this edition in production environments.	You can use this edition if you have a maximum of 100 pods in an ASM instance.	You can use this edition if you have a maximum of 1,000 pods in an ASM instance.	You can use this edition if you have a maximum of 10,000 pods in an ASM instance.

References

- Istio
- ASM Workshop
- Course jointly developed by CNCF and Alibaba Cloud on cloud-native technologies
- ACP certification course for engineers who use Alibaba Cloud cloud-native containers

2.Benefits

This topic describes the benefits of Alibaba Cloud Service Mesh (ASM).

Ease of use

- You can create a service mesh instance with simple operations in the ASM console.
- You can upgrade a service mesh instance with simple operations in the ASM console.
- You can define virtual services, destination rules, and the ingress gateway resources of Istio.

Rich features

Feature	Description
Fully managed	- Allows you to install, deploy, and upgrade applications with simple operations. - Manages components of the control plane. - Allows you to focus on the development of business applications. - Is compatible with the specifications of the Istio community.
Traffic routing	- Supports custom traffic routing rules. - Supports managing traffic between application services in multiple Kubernetes clusters. - Provides fine-grained traffic management. - Provides out-of-the-box fault recovery and adopts chaos engineering.
Security	- Allows you to enable or disable mutual Transport Layer Security (mTLS) authentication for all services in a mesh. - Allows you to use mTLS authentication in different modes, such as the permissive mode and the strict mode. - Supports easy-to-use role-based access control (RBAC). - Allows you to connect third-party systems to it with simple operations.
Monitoring and diagnosis	- Provides the mesh diagnosis feature. - Supports Tracing Analysis, Application Real-Time Monitoring Service (ARMS), and Log Service in a managed manner. - Provides end-to-end visibility. - Builds an ecosystem with multiple Alibaba Cloud services.
High stability	Builds a dedicated support team to guarantee the stability of managed services.

Feature	Description
High availability	- Provides various control plane specifications and service level agreements (SLAs) to meet different business requirements. - Supports multi-zone deployment.
Technical support	- Manages the core components on the control plane to minimize O&M complexity. - Allows you to submit tickets for help.

3.Features

Alibaba Cloud Service Mesh (ASM) provides the following editions that support different features and capabilities: Standard Edition, Enterprise Edition, and Ultimate Edition. Standard Edition is a free edition, and the other editions are commercial editions. This topic describes the features supported by different ASM editions.

Mesh management

Feature	Open source edition	Standard Edition	Enterprise Edition	Ultimate Edition
Full lifecycle management of ASM instances such as instance deployment and upgrade management in the ASM console	☐	✔☺	✔☺	✔☺
Centralized management of Container Service for Kubernetes (ACK) clusters, Serverless Kubernetes (ASK) clusters, clusters deployed on elastic container instances (ECIs) and Elastic Compute Service (ECS) instances, and Kubernetes clusters that are not on Alibaba Cloud	☐	✔☺	✔☺	✔☺
Multi-cluster deployment across virtual private clouds (VPCs) and regions in production environments	☐	✔☺	✔☺	✔☺
Integration with the Resource Access Management (RAM) system to support various features such as RAM authorization	☐	✔☺	✔☺	✔☺
Mesh diagnosis	☐	✔☺	✔☺	✔☺
Advanced features of ASM gateways, including graceful shutdown for SLB instances, Horizontal Pod Autoscaler (HPA)-based scaling, rolling upgrade, and Transport Layer Security (TLS) acceleration	☐	☐	✔☺	✔☺
Multi-dimensional configurations of sidecar proxies in global, namespace, and workload levels	☐	✔☺	✔☺	✔☺

Feature	Open source edition	Standard Edition	Enterprise Edition	Ultimate Edition
Rollback of Istio resources to an earlier version		✔☺	✔☺	✔☺

Traffic management

Feature	Open source edition	Standard Edition	Enterprise Edition	Ultimate Edition
Core features of open source Istio, such as traffic routing, traffic governance, observability, and security	✔☺	✔☺	✔☺	✔☺
Local throttling (compatible with open source Istio)	✔☺	✔☺	✔☺	✔☺
Local throttling (exclusive for ASM commercial editions)			✔☺	✔☺
Support for the Dubbo protocol			✔☺	✔☺
Support for Spring Cloud services			✔☺	✔☺
Graceful start and shutdown of services			✔☺	✔☺
End-to-End tag-based routing			✔☺	✔☺
Application High Availability Service (AHAS) for traffic protection			✔☺	✔☺

Observability

Feature	Open source edition	Standard Edition	Enterprise Edition	Ultimate Edition
Integration with Alibaba Cloud services such as Log Service, Prometheus Service, and Tracing Analysis to improve usability		✔☺	✔☺	✔☺
Enhanced built-in common report dashboards		✔☺	✔☺	✔☺

Security

Feature	Open source edition	Standard Edition	Enterprise Edition	Ultimate Edition
Fine-grained access control by using an Open Policy Agent (OPA) policy	⬜	✔☺	✔☺	✔☺
ActionTrail	⬜	✔☺	✔☺	✔☺

Stability and supported scale

Feature	Open source edition	Standard Edition	Enterprise Edition	Ultimate Edition
Supported scale on the data plane	/	50 pods ⑦ **Note** We recommend that you do not use this edition in production environments.	1,000 pods	10,000 pods
Managed Istiod components on the control plane	/	Single replica	Multiple replicas	Multiple replicas

Scalability

Feature	Open source edition	Standard Edition	Enterprise Edition	Ultimate Edition
Envoy filter marketplace	⬜	✔☺	✔☺	✔☺
Connection to third-party registries such as Consul and Nacos	⬜	✔☺	✔☺	✔☺

Performance

Feature	Open source edition	Standard Edition	Enterprise Edition	Ultimate Edition
TLS acceleration based on Multi-Buffer	⬜	⬜	✔☺	✔☺

Feature	Open source edition	Standard Edition	Enterprise Edition	Ultimate Edition
Selective service discovery and automatic sidecar recommendation based on access log analysis for configuration optimization	🗙	✔☺	✔☺	✔☺

References for features of ASM commercial editions

Feature	References
Mesh management	Enable Multi-Buffer for TLS acceleration
ASM gateway	- Bind a certificate to a domain name - Enable data compression for the ingress gateway service of an ASM instance - Improve availability for the ingress gateway service of an ASM instance
Traffic management	- Use the local throttling feature of ASM - Enable graceful shutdown for the SLB instance of an ASM gateway to prevent traffic loss - Configure traffic routing for an ASM gateway - Traffic labeling and label-based routing - Manage Spring Cloud services - Use an ASM instance of a commercial edition to implement an end-to-end canary release

4.Scenarios

Alibaba Cloud Service Mesh (ASM) provides the following features for you to manage application services: traffic management, security management, fault recovery, observability, and microservices model. This topic describes the typical scenarios of ASM.

Traffic management

In ASM, you can set parameters to manage traffic:

- Traffic management is separated from infrastructure management. You can manage traffic without using the application code. Even if your application services scale up, you still can manage traffic with ease.
- You can configure service discovery, traffic routing, and load balancing for a mesh. The configuration applies to all services in the mesh. This dispenses with some settings that are required by each service, such as the timeout and retry settings.

Security management

In ASM, you can enable mutual Transport Layer Security (mTLS) authentication for services.

- You can enable mTLS authentication in different modes, such as the permissive mode and the strict mode. mTLS authentication can secure the communication between services and between users and services.
- When you enable mTLS authentication for services, you do not need to modify the service code. mTLS provides role-based authentication for each service so that they can access each other across clusters and cloud platforms.

In ASM, you can authorize services to access each other as required.

- Based on the authorization method of Istio, ASM only allows verified and authorized clients to access services that contain sensitive data.
- ASM supports role-based access control (RBAC) that provides namespace-, service-, and method-level access control for services in a mesh. RBAC includes role-based semantics and service-to-service and user-to-service authorization. In addition, RBAC allows you to define service roles and service role bindings with custom properties.

ASM supports key management.

- Based on the key and certificate management of Istio, ASM can automatically generate, distribute, rotate, and revoke keys and certificates.

Fault recovery

ASM provides out-of-the-box fault recovery.

- Distributed systems are complex, which brings risks to the stability of infrastructure, application logic, and O&M. This may lead to failures in business systems.
- Based on Istio, ASM supports chaos engineering, including circuit breaking based on connection pool settings and outlier detection, service retry, and fault injection.

Observability

ASM allows you to observe services in meshes with ease. Supported by powerful, reliable, and easy-to-use monitoring features, ASM can detect and resolve issues at the earliest opportunity.

ASM integrates Alibaba Cloud Tracing Analysis, which provides a wide range of tools to help developers identify performance bottlenecks of distributed applications. This helps developers improve the efficiency of developing and troubleshooting applications that use the microservices model. The provided tools can be used to map traces, offer trace topologies, analyze application dependencies, and calculate the number of requests.

Microservices model

ASM enables agile development and deployment to speed up the evolution of business models. Your workload in the production environment is divided into multiple microservice applications. These microservice applications are managed by Alibaba Cloud image repositories. Alibaba Cloud can schedule, orchestrate, deploy, and implement the canary releases of microservice applications. Therefore, you can focus on feature updates.

- ASM is integrated with Server Load Balancer (SLB) and a service discovery system. This allows ASM to forward Layer 4 and Layer 7 requests and bind services to backend containers.

- ASM provides a variety of scheduling and disaster recovery policies for you to schedule the affinity of services and implement high availability and disaster recovery across zones.

- ASM allows you to monitor microservices and containers. It also enables microservices to automatically scale in and out.

5.Terms

This topic introduces terms related to Alibaba Cloud Service Mesh (ASM).

ASM instance

An instance that you create in ASM to function as the Istio control plane. ASM instances allow you to focus on application development and deployment without the need to maintain the Istio control plane. ASM instances are easy to use and provide high availability at low cost.

control plane

The plane that manages and configures proxies to route traffic. An Istio service mesh is logically split into a data plane and a control plane.

data plane

The plane that is composed of a set of intelligent Envoy proxies that are deployed as sidecars. These proxies mediate and control all network communication between microservices, and collect and report telemetry on all mesh traffic.

namespace

A unit used in Kubernetes to divide cluster resources between users. By default, Kubernetes clusters start with three initial namespaces: default, kube-system, and kube-public. Administrators can create custom namespaces as required.

The namespaces that you create in an ASM instance, whether in the ASM console or by using the kubectl client, belong only to the ASM instance. They are independent of the Kubernetes clusters in the data plane managed by the ASM instance. Therefore, the namespaces in the ASM instance may be different from those of the Kubernetes clusters in the data plane managed by the ASM instance. When you create or delete namespaces for the ASM instance, the namespaces of the Kubernetes clusters in the data plane managed by the ASM instance are not affected.

virtual service

A custom resource of Istio that defines a set of routing rules for specific services in a service mesh. Each routing rule defines matching criteria for traffic of a specific protocol. The traffic that matches a routing rule is sent to a destination service, or a subset or version of the destination service defined in the service registry.

destination rule

A custom resource of Istio that defines policies that apply to traffic intended for a service after routing has occurred. These rules specify configuration for load balancing, connection pool size from the sidecar, and outlier detection settings to detect and evict unhealthy hosts from the load balancing pool.

Istio gateway

A custom resource of Istio that describes a load balancer running at the edge of the mesh for receiving incoming or outgoing HTTP/TCP connections. The specification describes a set of ports that need to be exposed, the type of protocol to use, and the server name indication (SNI) configuration for the load balancer.

service entry

A custom resource of Istio that is used to add a service to the abstract model or service registry of

Istio. Registered services are maintained in Istio. After you add an entry for an external service, the Envoy proxies can send traffic to the service as if the service was in the mesh.

ingress gateway service

A service that is used to manage access from applications in a Kubernetes cluster to applications outside the cluster. Ingress gateway services use Kubernetes resources instead of Istio resources. Ingress gateway services are supported by pods. When you create an ingress gateway service for a cluster in ASM, a Kubernetes service and a Kubernetes deployment are deployed in the cluster.

6.Supported regions

Regions are geographical areas where data centers reside. You cannot change the region to which a resource belongs after the resource is created. This topic describes the regions that are supported by Alibaba Cloud Service Mesh (ASM).

Cloud	Region	City	Region ID
Alibaba Cloud	China (Beijing)	Beijing	cn-beijing
	China (Zhangjiakou)	Zhangjiakou	cn-zhangjiakou
	China (Hohhot)	Hohhot	cn-huhehaote
	China (Hangzhou)	Hangzhou	cn-hangzhou
	China (Shanghai)	Shanghai	cn-shanghai
	China (Shenzhen)	Shenzhen	cn-shenzhen
	China (Heyuan)	Heyuan	cn-heyuan
	China (Guangzhou)	Guangzhou	guangzhou
	China (Chengdu)	Chengdu	cn-chengdu
	China (Hong Kong)	Hong Kong	cn-hongkong
	Singapore (Singapore)	Singapore	ap-southeast-1
	Malaysia (Kuala Lumpur)	Kuala Lumpur	ap-southeast-3
	Indonesia (Jakarta)	Jakarta	ap-southeast-5
	India (Mumbai)	Mumbai	ap-south-1
	US (Silicon Valley)	Silicon Valley	us-west-1
	US (Virginia)	Virginia	us-east-1
	Germany (Frankfurt)	Frankfurt	eu-central-1

7.Limits

Before you use Alibaba Cloud Service Mesh (ASM), you must bear in mind its usage limits. This topic describes the operation limits, supported regions, and resource quotas of ASM.

Limits

ASM is in the public preview stage. Before you use ASM, you must apply for the public preview qualification. The following table describes limits of ASM during the public preview stage.

Item	Limit
ASM instance creation	You are not allowed to perform the following operations on an ASM instance after it is created: • Change the VPC and VSwitch on which the service mesh depends. • Enable the Internet access to the API server if this feature is not enabled before. • Enable the Internet access to Istio Pilot if this feature is not enabled before. • Modify the configuration of Tracing Analysis.
Region	Currently, ASM is available only in the following regions: China (Zhangjiakou), China (Beijing), and China (Hangzhou).
Quota	• Currently, you can create a maximum of 10 ASM instances under an Alibaba Cloud account. • We recommend that you configure less than 1,000 Envoy proxies for each service mesh.

> ⑦ Note   Before you use ASM, make sure that you can create managed clusters in Container Service for Kubernetes. For more information about limits on Container Service for Kubernetes clusters, see Limits.

8.Road map

The Alibaba Cloud Service Mesh (ASM) roadmap lists the features that have been developed, the features under development, and the features that are to be developed for ASM. You can learn how to use ASM in the future from the roadmap. You can submit feedback on ASM features to communicate with the ASM team on the **Issues** tab of the ASM roadmap page at GitHub.

For more information, see ASM Roadmap.