

Alibaba Cloud Apsara Stack Enterprise

Product Introduction

Version: 1901

Issue: 20190528

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 due to product version upgrades, adjustments, or other reasons. Alibaba Cloud reserves the right to modify the content of this document without notice and the updated versions of this document will be occasionally released through Alibaba Cloud-authorized channels. You shall 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 the document in the context that Alibaba Cloud products and services are provided on an "as is", "with all faults" and "as available" basis. 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 bear any liability for any errors or financial losses incurred by any organizations, companies, or individuals arising from their download, use, or trust in this document. Alibaba Cloud shall not, under any circumstances, bear responsibility for any indirect, consequential, exemplary, incidental, 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 contact Alibaba Cloud directly if you discover any errors in this document.

Document conventions

Style	Description	Example
	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.
	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.
	A caution notice indicates warning information, supplementary instructions, and other content that the user must understand.	 Note: If the weight is set to 0, the server no longer receives new requests.
	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 <code>cd /d C:/window</code> command to enter the Windows system folder.
<i>Italic</i>	Italic formatting is used for parameters and variables.	<code>bae log list --instanceid Instance_ID</code>
[] or [a b]	This format is used for an optional value, where only one item can be selected.	<code>ipconfig [-all -t]</code>
{ } or {a b}	This format is used for a required value, where only one item can be selected.	<code>switch {active stand}</code>

Contents

Legal disclaimer.....	I
Document conventions.....	I
1 What is Alibaba Cloud Apsara Stack?.....	1
2 The reasons to choose Apsara Stack.....	3
2.1 Hyper-scale distributed cloud operating system.....	3
2.2 Deployment and control system of Apsara Infrastructure Management Framework.....	4
2.3 High-reliability disaster recovery solutions.....	5
2.4 Centralized operations management and automated operations capabilities.....	6
2.5 Open cloud service interface.....	7
3 Product architecture.....	8
3.1 Types of private cloud architectures.....	8
3.2 System architecture.....	8
3.3 Network architecture.....	10
3.3.1 Network architecture overview.....	10
3.3.2 Business service area.....	12
3.3.3 Integrated access area.....	14
3.3.4 VPC leased line access.....	16
3.4 Security architecture.....	17
3.5 Base assembly.....	18
4 Scenarios.....	20
5 Compliance security solution.....	22
5.1 Overview.....	22
5.2 Interpretation on key points.....	22
5.3 Cloud-based classified protection compliance.....	24
5.4 Classified protection implementation process.....	27
5.5 Security compliance architecture.....	28
5.6 Solution benefits.....	29
6 Elastic Compute Service (ECS).....	30
6.1 What is ECS.....	30
6.2 Benefits.....	31
6.3 Architecture.....	34
6.4 Features.....	35
6.4.1 Instances.....	35
6.4.1.1 Overview.....	35
6.4.1.2 Instance type families.....	35
6.4.1.3 Instance types.....	45
6.4.1.4 Instance UserData.....	57
6.4.1.5 Instance lifecycle.....	57
6.4.1.6 ECS Bare Metal Instance.....	59

6.4.2 Cloud disks.....	61
6.4.2.1 Overview.....	61
6.4.2.2 Performance.....	62
6.4.2.2.1 Overview.....	62
6.4.2.2.2 Elastic block storage.....	62
6.4.2.2.3 Local storage.....	63
6.4.2.2.4 Performance test.....	63
6.4.2.3 Elastic block storage.....	65
6.4.2.3.1 Overview.....	65
6.4.2.3.2 Cloud disk.....	65
6.4.2.3.3 Shared block storage.....	66
6.4.2.3.4 Triplicate technology.....	67
6.4.2.4 ECS disk encryption.....	68
6.4.2.5 Local storage.....	69
6.4.3 Images.....	70
6.4.4 Snapshots.....	71
6.4.4.1 Overview.....	71
6.4.4.2 Incremental snapshot mechanism.....	71
6.4.4.3 ECS Snapshot 2.0.....	73
6.4.4.4 ECS Snapshot 2.0 vs. traditional storage products.....	74
6.4.5 Deployment sets.....	74
6.4.6 Network and security.....	76
6.4.6.1 IP address of a VPC.....	76
6.4.6.2 Elastic network interface.....	77
6.4.6.3 Intranet.....	78
6.4.6.4 Security group rules.....	79
6.5 Application scenarios.....	79
6.6 Usage limitations.....	80
6.7 Basic concepts.....	82
7 Auto Scaling (ESS).....	84
7.1 What is ESS.....	84
7.2 Benefits.....	84
7.3 Architecture.....	85
7.4 Features.....	86
7.5 Usage scenarios.....	87
7.6 Limits.....	87
7.7 Terms.....	88
8 Object Storage Service (OSS).....	90
8.1 What is OSS.....	90
8.2 Advantages.....	90
8.3 Architecture.....	91
8.4 Functions.....	93
8.5 Scenarios.....	94

8.6 Limits.....	95
8.7 Concepts.....	95
9 Table Store.....	98
9.1 What is Table Store.....	98
9.2 Benefits.....	99
9.3 Architecture.....	100
9.4 Features.....	101
9.5 Scenarios.....	104
9.6 Limits.....	110
9.7 Terms.....	112
10 Network Attached Storage (NAS).....	115
10.1 What is NAS?.....	115
10.2 Benefits.....	115
10.3 Architecture.....	116
10.4 Features.....	116
10.5 Scenarios.....	117
10.6 Limits.....	117
10.7 Terms.....	119
11 Distributed File System (DFS).....	120
11.1 What is DFS.....	120
11.2 Benefits.....	120
11.3 Architecture.....	120
11.4 Features.....	121
11.5 Scenarios.....	122
11.6 Limits.....	123
11.7 Terms.....	123
12 ApsaraDB for RDS.....	125
12.1 What is ApsaraDB for RDS?.....	125
12.2 Benefits.....	126
12.2.1 Ease of use.....	126
12.2.2 High performance.....	127
12.2.3 High security.....	127
12.2.4 High reliability.....	128
12.3 Architecture.....	129
12.4 Features.....	129
12.4.1 Data link service.....	130
12.4.2 High-availability service.....	131
12.4.3 Backup and recovery service.....	133
12.4.4 Monitoring service.....	134
12.4.5 Scheduling service.....	135
12.5 Scenarios.....	135
12.5.1 Diversified data storage.....	136
12.5.2 Read/write splitting.....	137

12.6 Usage limits.....	138
12.6.1 Usage limits of ApsaraDB RDS for MySQL.....	138
12.6.2 Usage limits of ApsaraDB RDS for PostgreSQL.....	140
12.6.3 Usage limits of ApsaraDB RDS for PPAS.....	141
12.7 Terms.....	141
12.8 Instance types.....	143
13 KVStore for Redis.....	151
13.1 What is KVStore for Redis.....	151
13.2 Benefits.....	151
13.3 Architecture.....	152
13.4 Features.....	153
13.5 Scenarios.....	155
13.6 Limits.....	156
13.7 Concepts.....	157
13.8 Instance types.....	158
14 ApsaraDB for MongoDB.....	163
14.1 What is ApsaraDB for MongoDB.....	163
14.2 Benefits.....	163
14.3 Architecture.....	164
14.4 Features.....	165
14.5 Scenarios.....	166
14.6 Limits.....	167
14.7 Glossary.....	168
14.8 Instance specifications.....	169
15 KVStore for Memcache.....	171
15.1 What is KVStore for Memcache.....	171
15.2 Benefits.....	171
15.3 Architecture.....	172
15.4 Features.....	174
15.5 Scenarios.....	175
15.6 Limits.....	175
15.7 Glossary.....	176
15.8 Instance specifications.....	176
16 Data Management Service (DMS).....	179
16.1 What is DMS?.....	179
16.2 Benefits.....	179
16.3 Architecture.....	179
16.4 Features.....	180
16.5 Scenarios.....	181
16.5.1 Convenient data operations.....	181
16.5.2 Prohibiting data export.....	181
16.5.3 SQL statement reuse.....	182
16.6 Limits.....	182

17 Server Load Balancer (SLB)	184
17.1 What is Server Load Balancer?	184
17.2 Benefits	185
17.3 Architecture	186
17.4 Features	188
17.5 Scenarios	189
17.6 Limits	190
17.7 Terms	190
18 Virtual Private Cloud (VPC)	192
18.1 What is VPC?	192
18.2 Benefits	193
18.3 Architecture	194
18.4 Features	195
18.5 Scenarios	196
18.6 Limits	196
18.7 Terms	197
19 Log Service	199
19.1 What is Log Service?	199
19.2 Benefits	199
19.3 Architecture	200
19.4 Key features	203
19.4.1 Core features	203
19.4.2 Other features	204
19.4.2.1 Log	204
19.4.2.2 Project	206
19.4.2.3 Logstore	207
19.4.2.4 Shard	207
19.4.2.5 Log topic	210
19.5 Scenarios	210
19.6 Limits	213
19.7 Glossary	214
20 Apsara Stack Security	216
20.1 What is Apsara Stack Security	216
20.2 Benefits	216
20.3 Architecture	219
20.4 Features	221
20.5 Restrictions	227
20.6 Concepts	228
21 Key Management Service (KMS)	229
21.1 What is KMS	229
21.2 Benefits	230
21.3 Architecture	230

21.4 Features.....	232
21.5 Scenarios.....	232
21.6 Limits.....	236
21.7 Terms.....	237
22 Domain Name System (DNS).....	238
22.1 What is Apsara Stack DNS.....	238
22.2 Benefits.....	238
22.3 Architecture.....	239
22.4 Features.....	239
22.5 Scenarios.....	240
22.6 Limits.....	241
22.7 Basic concepts.....	242
23 API Gateway.....	243
23.1 Product overview.....	243
23.2 Features.....	243
23.3 Benefits.....	244
23.4 Concepts.....	245

1 What is Alibaba Cloud Apsara Stack?

Private cloud

Private cloud is a cloud computing system that is built within enterprises by cloud computing service providers. It places cloud infrastructures, software and hardware resources within firewalls to allow departments within an organization or enterprise to share resources in their data centers. It can be managed by an organization or a third party and located within the organization or outside the organization. Compared with public cloud, private cloud provides better privacy and exclusivity.

Private cloud is divided into two types by the sizes of enterprises or business requirements:

- Multi-tenant comprehensive private cloud for industries and large groups: A full stack cloud system created in a top-down manner to run hyper-scale digital applications. It satisfies IT requirements, such as the continuous integration and development of DevOps applications and operation support of production environments.
- Single-tenant basic private cloud for small- and medium-sized enterprises and scenarios: A cloud system that hosts technical systems, including large-scale Software as a Service (SaaS) applications, industrial clouds, and large group clouds. It also performs local computing tasks.

Apsara Stack

During the evolution from enterprise IT architecture to clouds, more and more enterprises want to have the service experience that is brought by large-scale cloud computing in their own data centers, which is based on the construction requirements, such as security compliance, reuse of existing data centers, and localization experience.

Apsara Stack is an extension of Alibaba Cloud public cloud, which brings the technologies of public cloud to Apsara Stack. By helping enterprises deliver complete and customizable Alibaba Cloud software solutions in their own data centers, Apsara Stack allows you to have the same characteristics as the hyper-scale cloud computing and big data products provided by Alibaba Cloud public cloud in the local environment. Apsara Stack also provides enterprises with the consistent hybrid cloud experience where you can obtain IT resources as required and guarantee the business continuity.

Service values

Supported by various products and services, based on successful digital practice cases of Alibaba Group, and integrated with the mature solutions and rich experience in various industries, Apsara

Stack helps governments and enterprises digitally transform their businesses and services.

Apsara Stack provides service values in the following four aspects:

- **Elastic**

Combines all resources into a supercomputer and flexibly scales out resources to minimize costs and maximize performance and stability.

- **Agile**

Integrates business with Internet and micro services to speed up the innovation of traditional enterprises.

- **Data**

Uses digitalization to allow data to flow between vertical businesses and forms a data shared service to deal with large amounts of data.

- **Smart**

Allows smart transformation of businesses globally and helps reinvent business models.

Platform characteristics

As an enterprise-level cloud platform, Apsara Stack has the following characteristics:

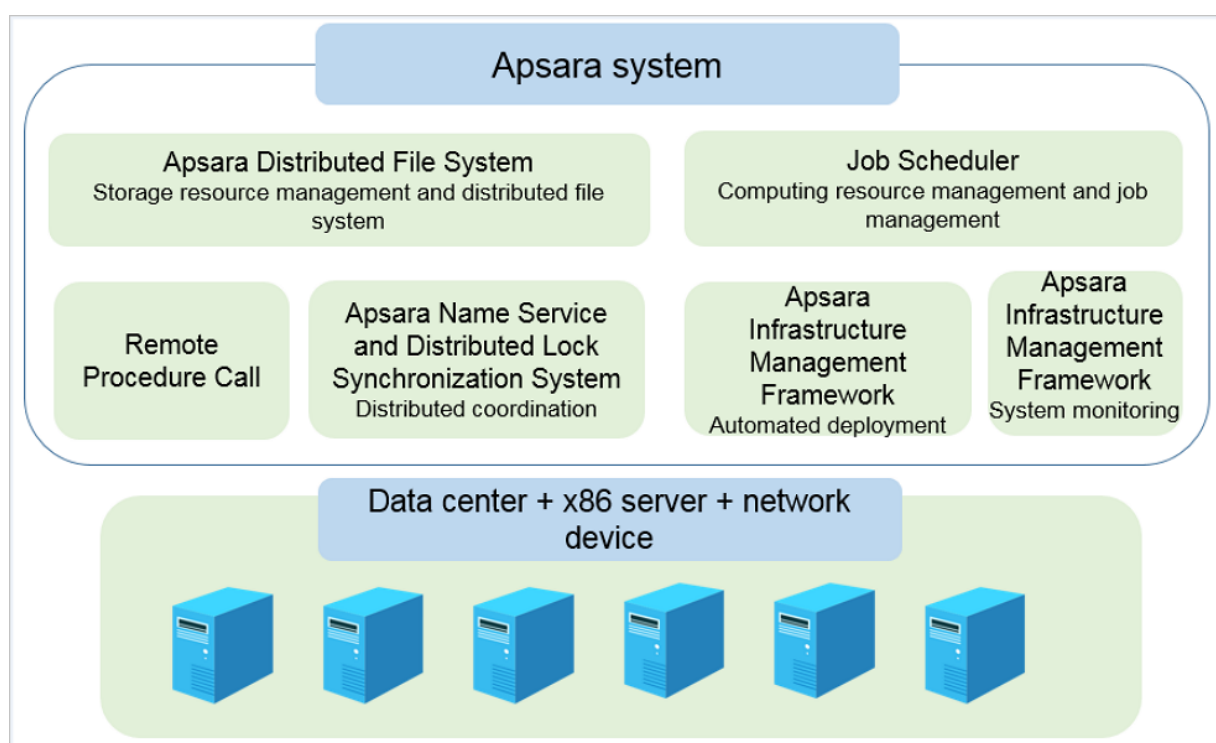
- Software-defined platform: masks underlying hardware differences, enables resources to scale up or out as required, and does not affect the performance of upper-layer applications.
- Production-level reliability and security compliance: guarantees the continuity and security of enterprise data.
- Centralized access management: isolates permissions of different roles for easy subsequent operations management.

2 The reasons to choose Apsara Stack

2.1 Hyper-scale distributed cloud operating system

Apsara Stack is based on the same underlying architecture (large-scale distributed computing system kernel of Apsara) as Alibaba Cloud public cloud. It provides underlying support for upper-layer services in terms of storage, computing, and scheduling. It is a hyper-scale and universal computing operating system that is independently developed by Alibaba Cloud for the global market. Apsara can connect millions of servers all over the world into a supercomputer and provide the community with computing capabilities in the form of online public services. The computing capabilities provided by Apsara are powerful, universal, and beneficial to everyone.

Figure 2-1: Apsara system kernel architecture



The modules of the Apsara platform kernel have the following primary functions:

- **Underlying services for distributed systems**

The modules provide the underlying services required in a distributed environment, such as coordination, remote procedure call, security management, and resource management services. These services provide support for the upper-layer modules, such as the distributed file system and job scheduling.

- **Distributed file system**

The modules aggregate storage capabilities from different nodes in a cluster to construct a massive, reliable, and scalable data storage service. The modules also protect against software and hardware faults automatically to guarantee uninterrupted data access. With the support for incremental expansion and automatic data balancing, the modules provide APIs that are similar to Portable Operating System Interfaces of UNIX (POSIX) for accessing the files in the user space. The modules also perform random read/write and append write operations.

- **Job scheduling**

The modules schedule jobs in cluster systems and support both online services that rely heavily on the response speed and offline jobs that require high data processing throughput. The modules detect faults and hot spots in systems automatically and guarantee a stable and reliable job completion in various methods, such as error retries and issuing concurrent backup jobs for long-tail jobs.

- **Cluster monitoring and deployment**

The modules monitor the running status and performance metrics of upper-layer application services and the status of clusters to send alert notifications and record exception events. The modules enable the operations personnel to manage the deployment and configuration of Apsara platform and upper-layer applications. The modules also support online cluster scaling and online update of application services.

2.2 Deployment and control system of Apsara Infrastructure Management Framework

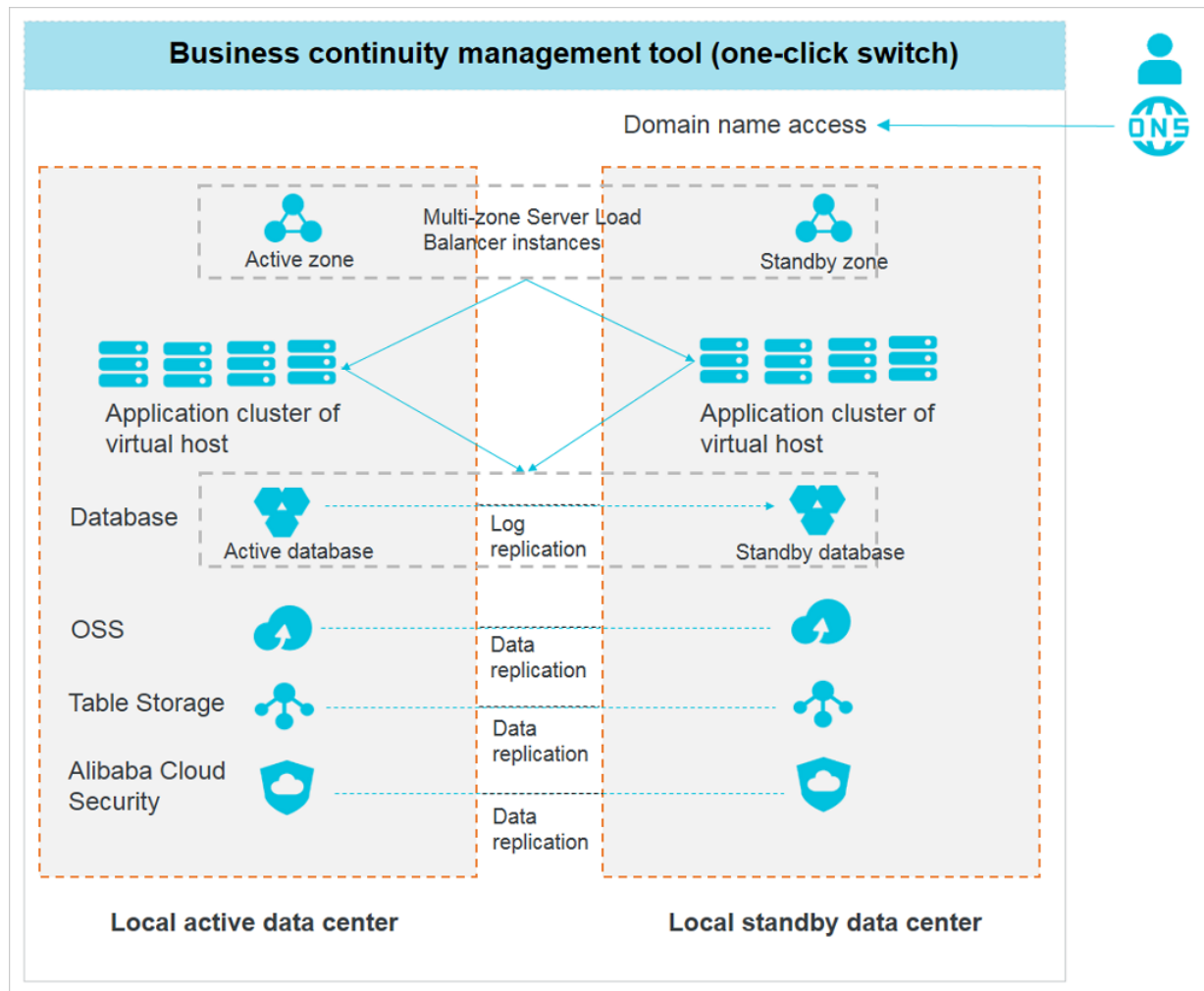
Apsara Infrastructure Management Framework provides the cloud services with basic support by providing the centralized deployment, authentication, authorization, and control for cloud service products. Apsara Infrastructure Management Framework contains various modules, including deployment framework, resource library, meta database, Alibaba Cloud Security, authentication and authorization component, interface gateway, Log Service, and control service module.

- The deployment framework provides all cloud services with unified access platform deployment and a management function that handles the dependencies among services.
- The resource library stores the execution files of all cloud services and their dependent components.
- Alibaba Cloud Security protects cloud services from Web attacks.

- The authentication and authorization component provides access control for cloud services and supports isolation of multiple tenants.
- The interface gateway provides a unified API management console for all cloud services.
- Log Service stores, retrieves, and obtains logs of cloud services.
- The control service module monitors the basic health of cloud services and supports the operations system of the cloud platform.

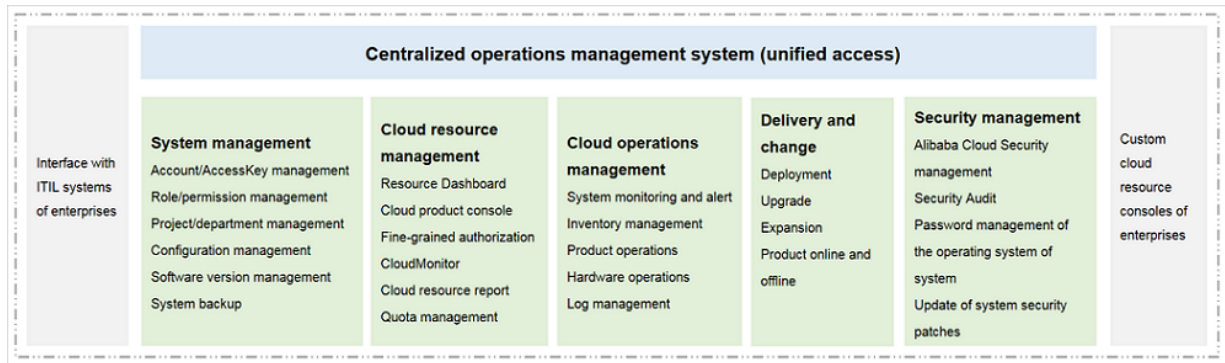
2.3 High-reliability disaster recovery solutions

Apsara Stack disaster recovery solutions are designed and developed based on the cloud computing capabilities of Alibaba Cloud. The solutions comply with common international disaster recovery standards. The standby data center must be within a 50-kilometer radius from the active data center in the same city, with a network latency of less than 0.6ms. The Apsara Stack platform deploys the network access layer and user application layer in active-active mode and the data persistence layer in active-standby mode.

Figure 2-2: Local disaster recovery

2.4 Centralized operations management and automated operations capabilities

Apsara Stack provides a centralized operations management system to configure different management permissions for different user roles. You can gain access to operations management capabilities by using APIs and customize your own cloud resource consoles. To interface and integrate with existing IT systems of various enterprises synchronously, Apsara Stack can interface with the Information Technology Infrastructure Library (ITIL) systems of enterprises.

Figure 2-3: Centralized operations management

2.5 Open cloud service interface

Cloud services provide a wide variety of SDKs and RESTful APIs on an API platform. You can use the open interfaces to flexibly access various cloud services provided by Apsara Stack. You can also obtain basic control information about the cloud platform by using these APIs and connect the Apsara Stack platform to your unified control system.

3 Product architecture

3.1 Types of private cloud architectures

Private cloud has two types of architectures: native cloud architecture and integrated cloud architecture.

- **Native cloud architecture**

The native cloud architecture evolves from the open architecture of Internet and is based on the distributed system framework. It is initially designed to handle big data and host Web applications, and subsequently expands to run basic services.

- **Integrated cloud architecture**

The integrated cloud architecture focuses on virtualization of computing services. As a breakthrough from the traditional architecture, it is open-sourced by the OpenStack and becomes the mainstream private cloud architecture.

Apsara Stack adopts the native cloud architecture and is based on self-developed distributed technologies and products of Alibaba Cloud. The single system supports all cloud products and services, and enables complete openness of the cloud platform. It comes with comprehensive service features for enterprises, a complete backup capabilities, and full autonomous control capabilities.

3.2 System architecture

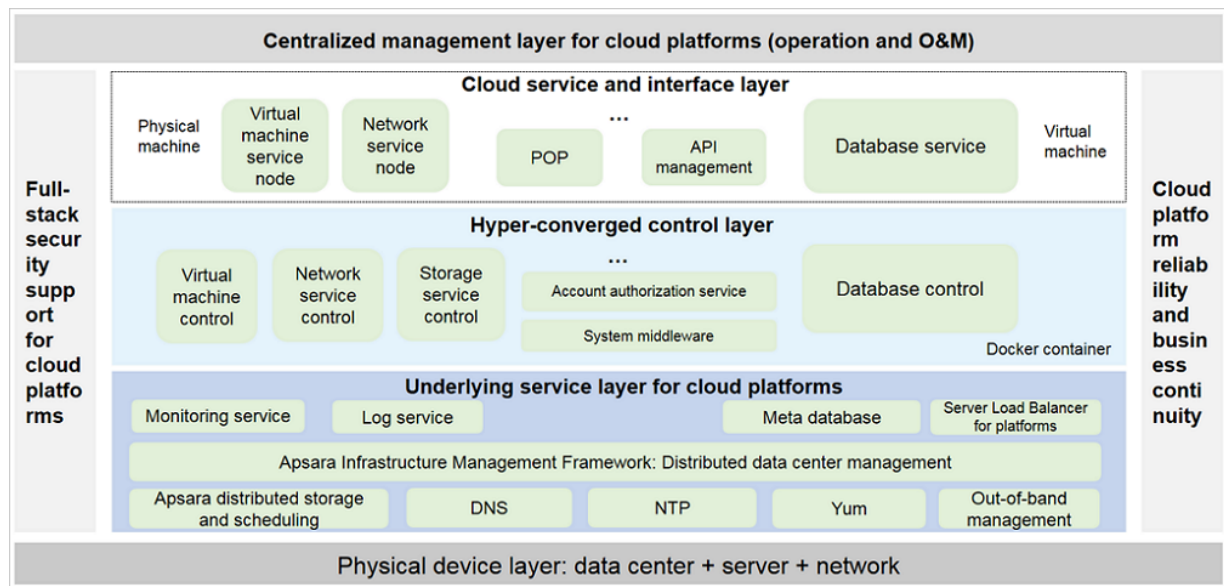
The Apsara Stack system architecture consists of the following parts, as shown in [Figure 3-1: Apsara Stack system architecture](#):

- Physical device layer: includes hardware devices for cloud computing, such as physical data centers, servers, and network.
- Underlying service layer for cloud platforms: bases on the underlying physical environment to provide underlying services for upper-layer applications.
- Hyper-converged control layer: provides centralized schedule for upper-layer applications or services by using the hyper-converged control architecture.
- Cloud service and interface layer: provides centralized management and Operation and Maintenance (O&M) for virtual machines and physical machines by using converged service nodes management, and uses the API platform to unify the interfaces and support custom development.

- Centralized management layer for cloud platforms: provides centralized operation and O&M management.

Apsara Stack also provides full-stack security support and guarantees the reliability of cloud platforms and business continuity.

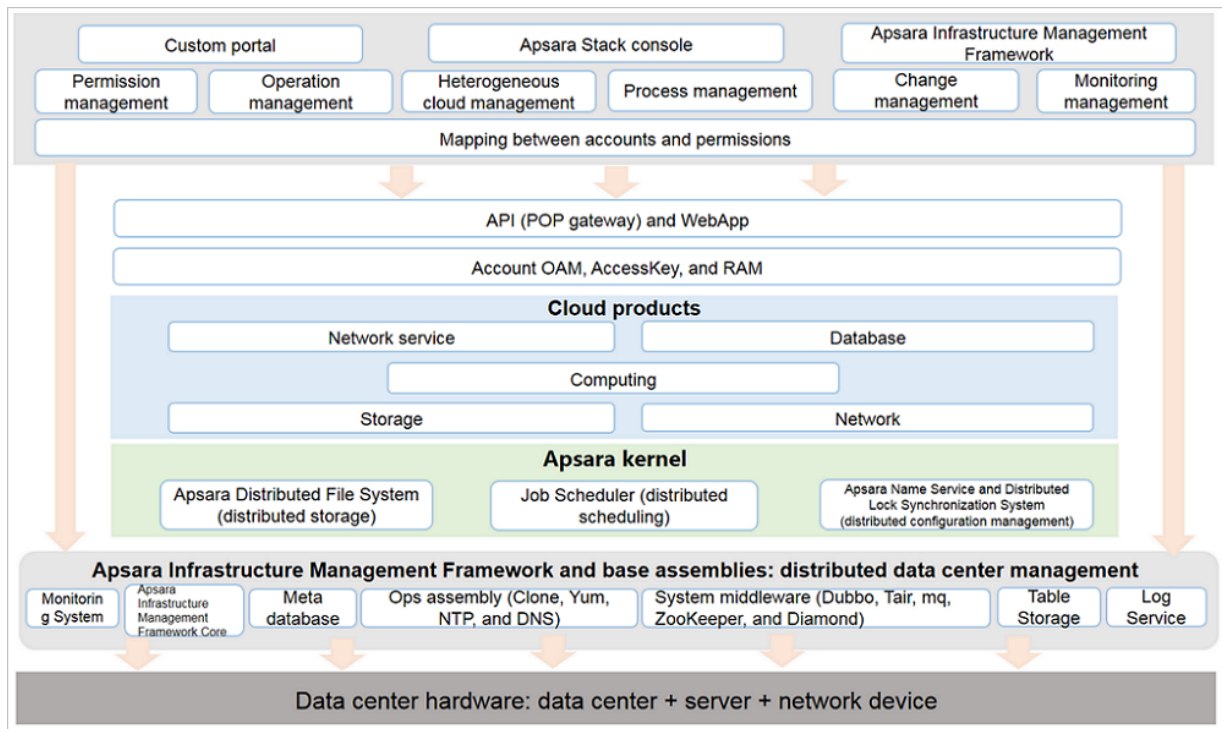
Figure 3-1: Apsara Stack system architecture



Logical architecture

Apsara Stack virtualizes the computing and storage capabilities of physical servers and network devices to achieve virtual computing, distributed storage, and software-defined networks. On this basis, Apsara Stack provides ApsaraDB and big data processing services. Apsara Stack also provides the supporting capabilities of underlying IT services for your applications, and can be interconnected with your existing account systems and monitoring operations systems. The logical architecture of Apsara Stack has the following characteristics:

- With data center + x86 server + network device as the hardware basis
- Based on the Apsara kernel (distributed engine) to provide various cloud products
- All cloud products are required to follow a unified API framework, management and O&M system (accounts, authorization, monitoring, and logs), and security system.
- Make sure that all cloud products have a consistent user experience.

Figure 3-2: Apsara Stack logical architecture

3.3 Network architecture

3.3.1 Network architecture overview

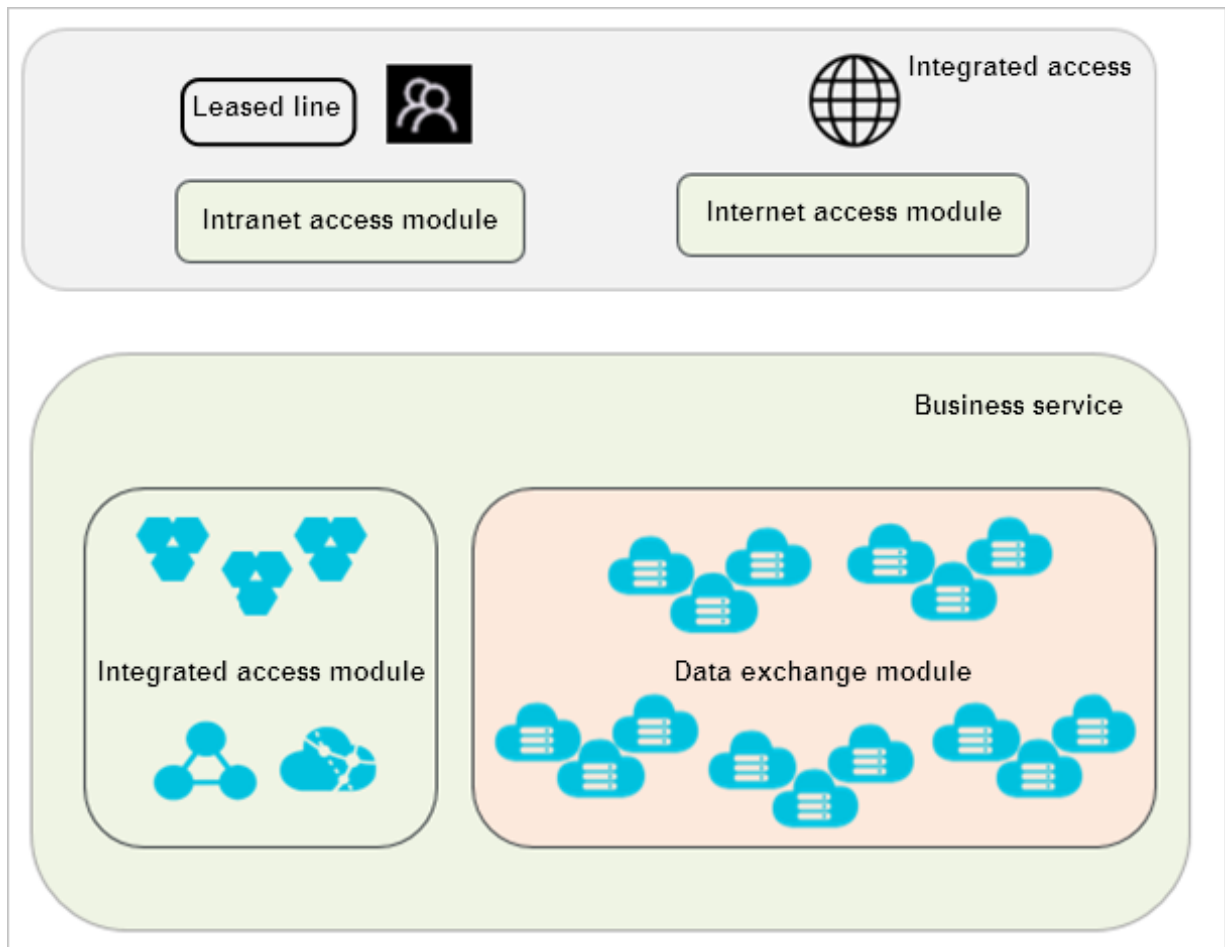
The Apsara Stack network architecture defines two logical areas, the business service area and the integrated access area, as shown in [Figure 3-3: Logical areas](#).

- Business service area

This area provides the networks of all cloud services and all cloud service systems exchange traffic in this area. This is the core area of Apsara Stack networks.

- Integrated access area

This area can be tailored based on the actual deployment requirements. As an extension of the business service area, the integrated access area provides a channel for user management, user private networks, and the access to Apsara Stack networks of Internet.

Figure 3-3: Logical areas

The roles and purposes of the switches in each area are as follows:

Role	Module	Purpose
Internet switch (ISW)	Internet access module	ISW is an egress switch and provides access to Internet service providers (ISPs) or users' backbone networks.
Customer switch (CSW)	Intranet access module	CSW facilitates the access to users' internal backbone networks. It performs route distribution and interaction between the inside and outside of cloud networks, including the access to VPC instances by using leased lines.

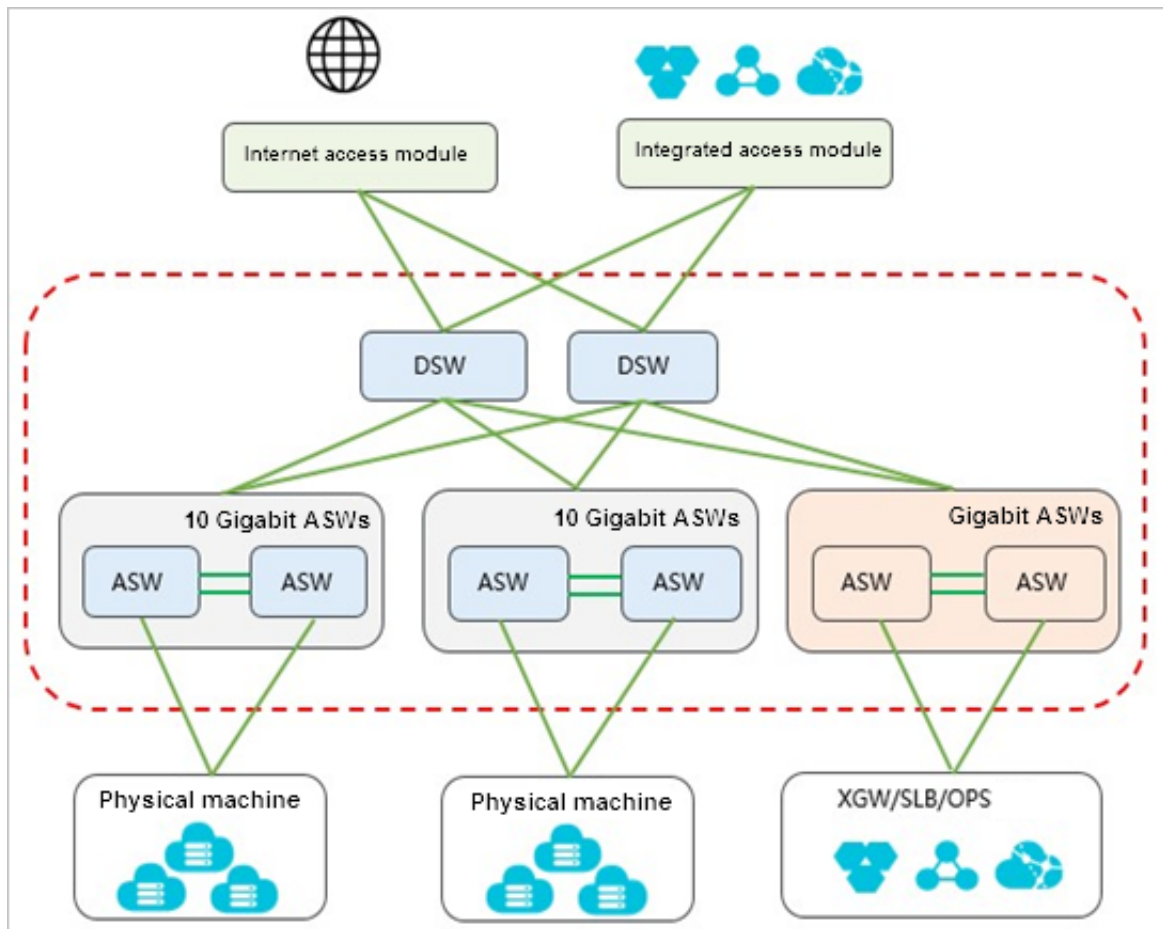
Role	Module	Purpose
Distributed switch (DSW)	Data exchange module	DSW functions as a core switch to connect all access switches.
Access switch (ASW)	Data exchange module	ASW provides access to cloud servers and is uplinked with the core switch DSW.
Integrated access switch (LSW)	Integrated access module	LSW provides access to cloud products, such as VPC and Server Load Balancer (SLB).

3.3.2 Business service area

The business service area consists of the data exchange module and the integrated service module.

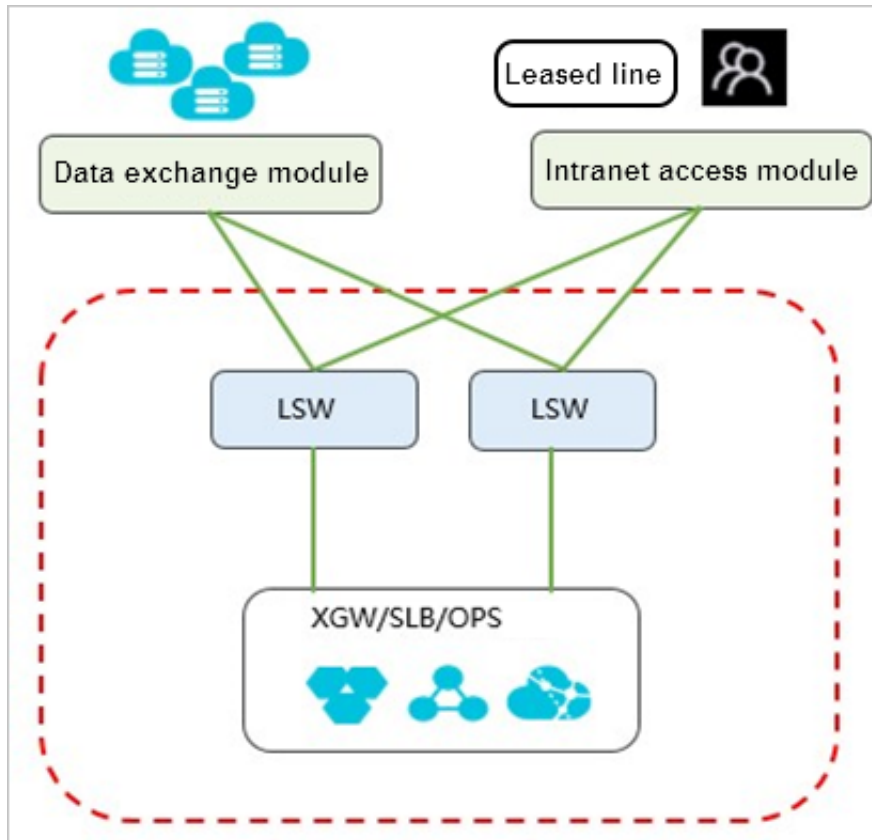
- Data exchange module

The data exchange module has a layer-2 CLOS architecture that consists of DSWs and ASWs . Each ASW pair forms a stack as a leaf node. According to the network sizes, this node can select data exchange models that have different applicable scopes. All cloud service servers are uplinked with the devices on the ASW stacks. ASWs are connected to DSWs by using External Border Gateway Protocol (EBGP). The DSWs are isolated from each other. The data exchange module is connected to other modules by using EBGP, receives the Internet routes from ISWs, and releases the Classless Inter-Domain Routing (CIDR) block of cloud products to the ISWs.

Figure 3-4: Data exchange module

- **Integrated service module**

Each cloud service server (XGW/SLB/OPS) is connected to two LSWs. These servers exchange routing information by using Open Shortest Path First (OSPF). The two LSWs exchange routing information between each other by using Internal Border Gateway Protocol (IBGP), and LSW exchange routing information with DSWs and CSWs by using EBGp.

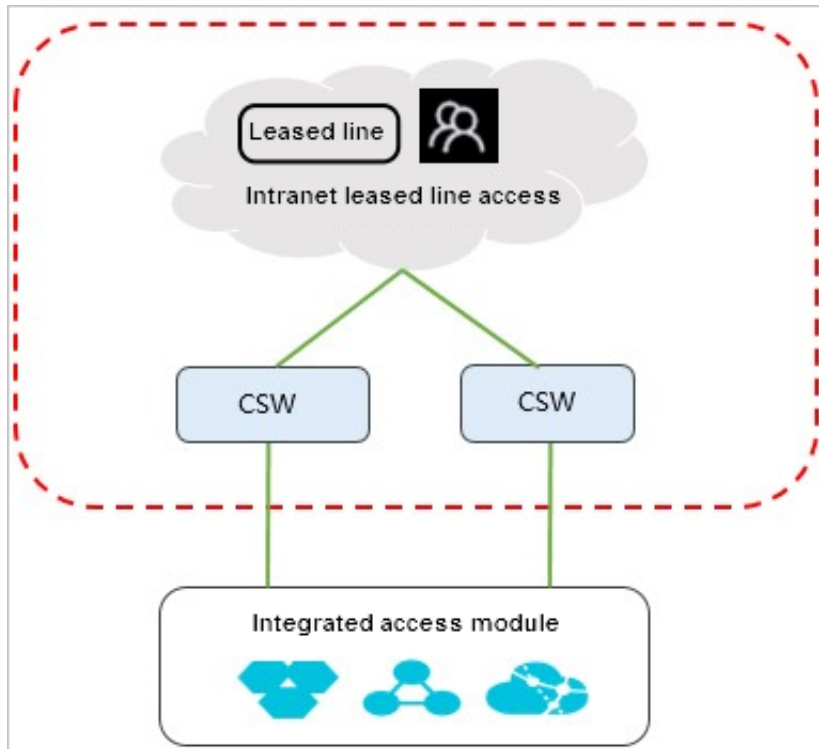
Figure 3-5: Integrated service module

3.3.3 Integrated access area

The integrated access area consists of the intranet access module and Internet access module.

- Intranet access module

In the intranet access module, two CSWs provide internal users with access to VPC instances and general cloud services. For access to VPC instances, CSWs set up a map from internal users to VPC instances and import these users into different VPC instances. Different user groups are isolated from each other on CSWs. For access to general cloud services, CSWs are connected to the integrated service module by using External Border Gateway Protocol (EBGP) and allow direct access to all resources in the business service area.

Figure 3-6: Intranet access module

- Internet access module

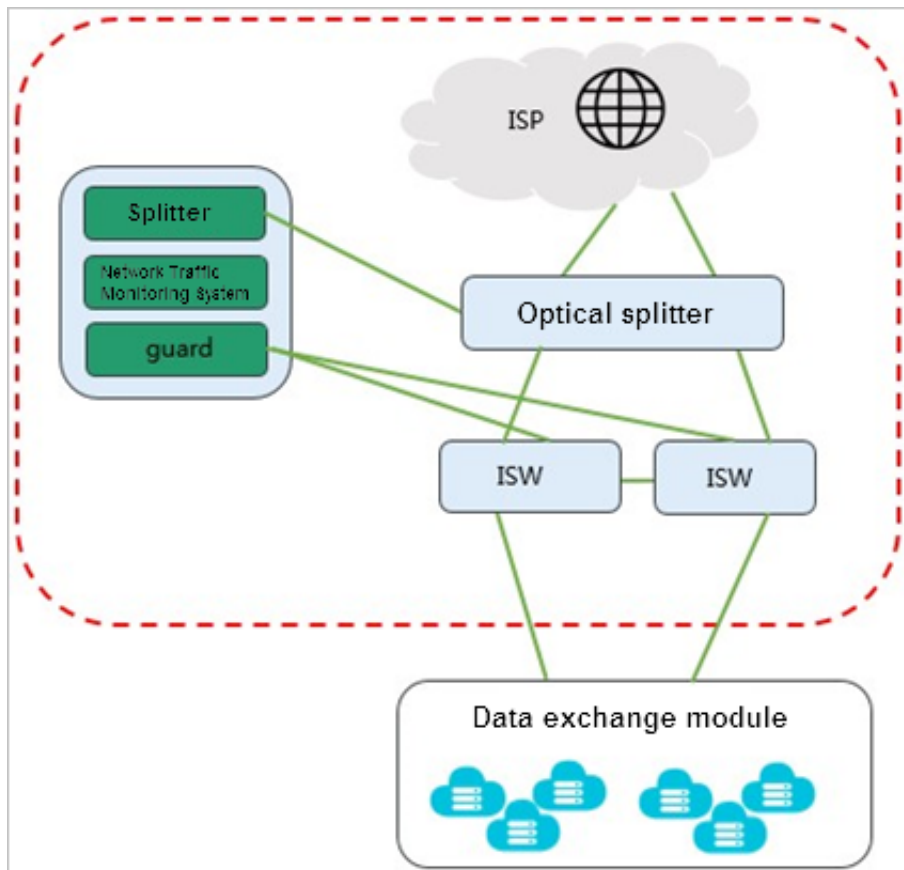
The Internet access module consists of two ISWs. It facilitates the access to ISPs or users' public backbone networks. It performs route distribution and interaction between the inside and outside of cloud networks. The two ISWs run Internal Border Gateway Protocol (IBGP) to back up routes between each other. Based on actual conditions, ISWs can use static routing or EBGP to uplink with Internet service providers (ISPs) or users' public backbone networks. The link bandwidth is defined based on the size of users' Alibaba Cloud networks and the bandwidth of their public backbone networks. We recommend that ISWs can use BGP to connect with multiple carriers to improve the reliability. Each carrier has 2×10 GE lines.

The Internet access module also uses EBGP to exchange routes with the data exchange module, releases Internet routes to the data exchange module, and receives the internal cloud service routes that are sent by the data exchange module to implement the interaction between the inside and outside of cloud networks.

The Internet access module is parallel to an Alibaba Cloud security protection system. The traffic generated by the Internet to cloud networks is diverted to Network Traffic Monitoring System by using an optical splitter. When Network Traffic Monitoring System detects malicious traffic, it releases the corresponding route by using Alibaba Cloud Security to divert the

malicious traffic to Alibaba Cloud Security for scrubbing. The scrubbed traffic is injected back into the Internet access module.

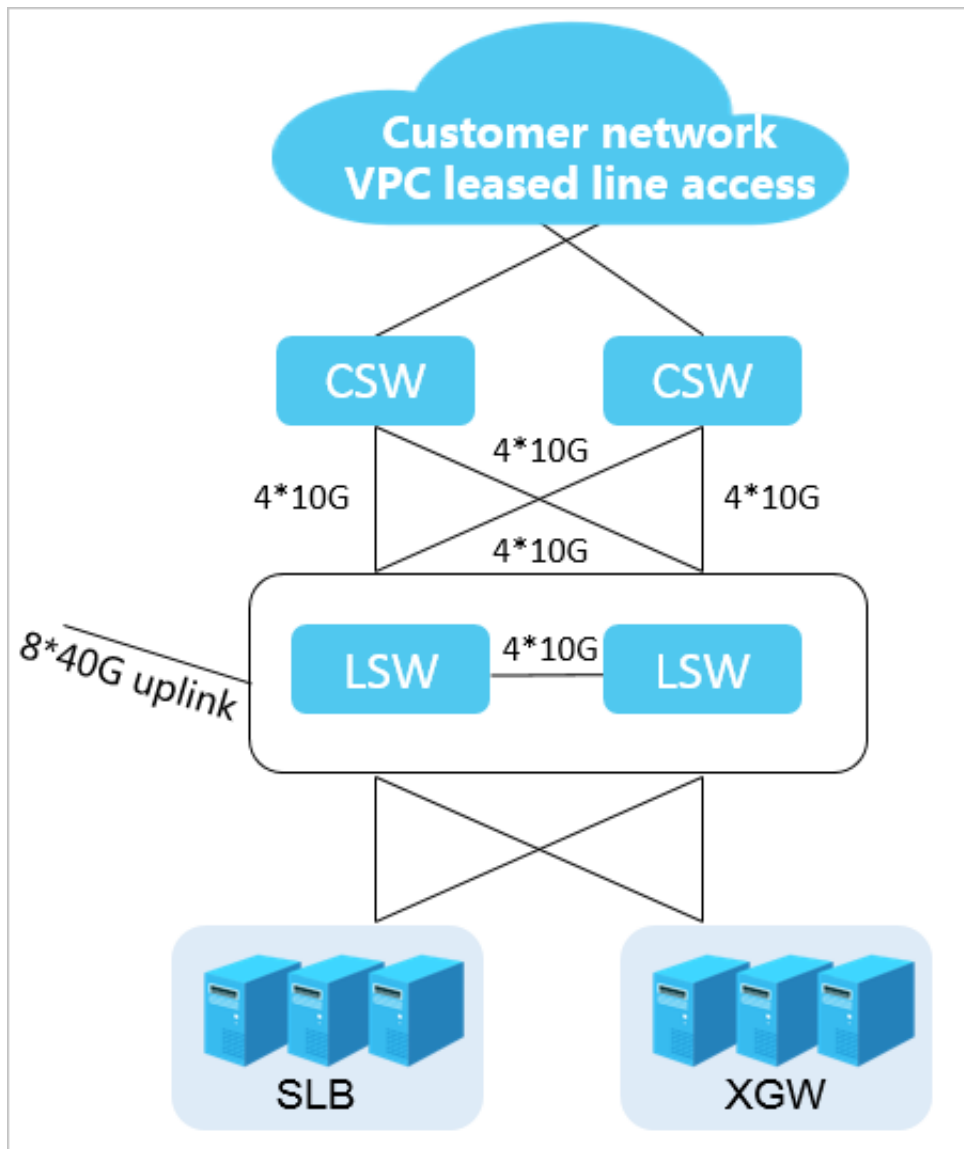
Figure 3-7: Internet access module



3.3.4 VPC leased line access

The Virtual Private Cloud (VPC) leased line access solution gives you full control over your own virtual networks, such as selecting your own IP address ranges and configuring the route tables and gateways. You can also connect your VPC instances to a traditional data center by using leased lines or VPN connections to create a customized network environment. This enables smooth migration of applications to the cloud.

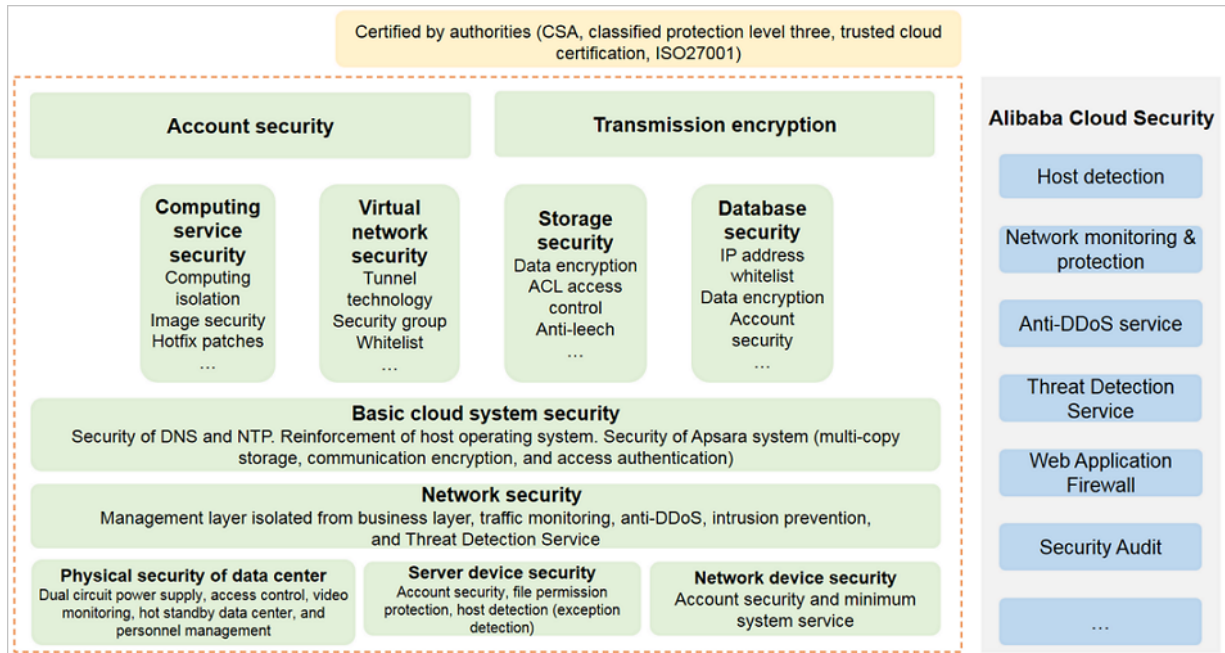
Each cloud service server (XGW/SLB) is connected to two LSWs. These servers exchange routing information by using Open Shortest Path First (OSPF). Two LSWs exchange routing information between each other by using Internal Border Gateway Protocol (IBGP), and LSWs exchange routing information with CSWs by using External Border Gateway Protocol (EBGP).

Figure 3-8: VPC leased line access

3.4 Security architecture

Apsara Stack provides all-around security capabilities from underlying communication protocols to upper-layer applications to guarantee security of your access and data. Access to every console in Apsara Stack is allowed only with HTTPS certificates. Apsara Stack provides a comprehensive role authorization mechanism to guarantee the secure and controllable access to resources in multi-tenant mode. It supports different security roles, such as security administrators, system administrators, and security auditors.

Apsara Stack has incorporated Alibaba Cloud Security since the third version and provides you with a multi-level and integrated cloud security protection solution.

Figure 3-9: Hierarchical security architecture of Apsara Stack

3.5 Base assembly

Apsara Stack base consists of three types of assemblies, which provides support for the deployment and operations of the cloud platform.

Table 3-1: Base assembly

Assembly		Function
Ops assemblies	Yum	Install package The software source is deployed during the initial installation phase. This package is mainly used to install the operating system and deploy application software packages and their dependent components of Apsara Stack, such as the Apsara platform and ECS, on physical machines.
	Clone	Machine cloning service
	NTP	Clock source service The physical machines deployed on Apsara Stack synchronize time from a standard NTP time source and provide the time to other hosts.
	DNS	Domain name resolution service DNS provides forward and reverse resolution of domain names for the internal Apsara Stack

Assembly		Function
		environment. It runs a bind instance on each of the two OPS machines and uses keepalived to provide high-availability services. When one machine fails, the other machine automatically takes over its work.
Base middleware	Dubbo	Distributed RPC service
	Tair	Cache service
	mq	Message Queue service
	ZooKeeper	Distributed collaboration
	Diamond	Configuration management service
	SchedulerX	Timing task service
Basic base assemblies	Apsara Infrastructure Management Framework	Data center management
	Monitoring System	Data center monitoring
	OTS-inner	Table Storage service
	SLS-inner	Cloud platform Log Service
	Meta database	Meta database
	POP	APIs on the cloud platform
	OAM	Account system
	RAM	Authentication and authorization system
	WebApps	Support for the Apsara Stack Operations console

4 Scenarios

Apsara Stack provides flexible and scalable industrial solutions for you who are from different scales in the same sector. Based on the business traits of different sectors, such as industry, agriculture, transportation, government, finance, and education, Apsara Stack creates custom solutions to provide you with one-stop products and services. This topic focuses on introducing the following two scenarios:

City Brain

Urban management is a field that involves one of the largest volumes of data in China. This marks the transition of governmental information from a closed-flow model to an open-flow online model. With more time and space to flow in, urban data has a higher value. Cloud computing becomes an urban infrastructure, data becomes a new means of production and a strategic resource, and AI technology becomes the nerve center of a smart city. All of these forms the City Data Brain.

Values and features

- A breakthrough of urban governance mode. With the urban data as a resource, City Brain improves the government management capabilities, resolves prominent issues of urban governance, and achieves an intelligent, intensive, and humane form of governance.
- A breakthrough of urban service mode. City Brain provides services for enterprises and individuals more accurately and conveniently, makes the urban public services more efficient, and saves more public resources.
- A breakthrough of urban industrial development. City Brain lays down an industrial AI layout, takes open urban data as an important fundamental resource, drives the development of industries, and promotes the transformation and upgrade of traditional industries.

Finance Cloud

Finance Cloud is an industrial cloud that serves financial organizations, such as banks, security agencies, insurance companies, and funds. It relies on a cluster of independent data centers to provide cloud products that meet the regulatory requirements of the People's Bank of China, China Banking Regulatory Commission (CBRC), China Securities Regulatory Commission (CSRC), and China Insurance Regulatory Commission (CIRC). It also provides more professional and comprehensive services for financial customers. Enterprises can build Finance Cloud independently or with Alibaba Cloud. Finance Cloud meets the requirements of large- and medium-sized financial organizations for independent cloud data centers that are completely physically isolated. It can also output the cloud computing and big data platforms to customers' data centers.

Values and features

- Independent resource clusters
- Stricter data center management
- Better disaster recovery capability
- Stricter requirements for network security isolation
- Stricter access control
- Compliance with the security supervision requirements and compliance requirements of banks
- Dedicated security operation team, security compliance team, and security solution team of the Finance Cloud sector
- Dedicated Finance Cloud account managers and cloud architects
- Stricter user access mechanism

5 Compliance security solution

5.1 Overview

On June 1, 2017, the *Cybersecurity Law of the People's Republic of China* was officially implemented, which has made clear provisions for classified protection compliance. To help enterprise users quickly meet the requirements of the classified protection compliance, Alibaba Cloud integrates the technical advantages of Alibaba Cloud Security to establish the Classified Protection Compliance Ecology. Alibaba Cloud works with its cooperative assessment agencies and security consulting manufacturers in various places to provide you with an one-stop classified protection assessment. The complete attack protection, data audit, encryption, and security management help you quickly and easily pass the classified protection compliance assessment.

5.2 Interpretation on key points

Network and communication security

Interpretation on clauses

- The network is divided into different security domains by server role and server importance.
- An access control policy is set at the security domain boundary between the intranet and Internet, which must be configured to specific ports.
- Intrusion prevention means must be deployed at the network boundary to prevent against and record intrusion behaviors.
- Information of security events and logs of user behaviors in network must be recorded and audited.

Coping strategies

- We recommend that you use VPC and security group of Alibaba Cloud to divide a network into different security domains and control the access reasonably.
- Web Application Firewall is used to prevent against network intrusion.
- Use the logging function to record, analyze, and audit security events and logs of user behaviors.
- If a system is frequently threatened by DDoS, you can use Anti-DDoS Pro to filter and scrub abnormal traffic.

Device and computing security

Interpretations on clauses

- It is the basic security requirement to record and audit operations actions and avoid sharing accounts.
- Necessary security measures are taken to guarantee the security of the system layer and prevent against intrusion to servers.

Coping strategies

- Audit the actions on servers and data. Create an independent account for each operations personnel to avoid sharing accounts.
- Use Server Guard to conduct complete management of server vulnerabilities, baseline check, and intrusion prevention.

Application and data security

Interpretations on clauses

- An application is the direct implementation of specific business. Unlike network and system, applications do not have the relative standard characteristics. The functions of most applications, such as identity authentication, access control, and operation audit, are difficult to be replaced by third-party products.
- Encryption is the most effective method to secure data integrity and confidentiality except security prevention methods at other levels.
- Remote data backup is one of the most important requirements that distinguish the third level of classified protection from the second level of classified protection. It is also the most foundational technical safeguard for business continuity.

Coping strategies

- At the beginning of the application development, application functions such as identity authentication, access control, and security audit must be considered.
- For online systems, functions such as account authentication, user permission classification, and log audit are designed to satisfy classified protection requirements.
- For data security, use HTTPS to make sure that data is encrypted in transmission.
- For data backup, we recommend that you use an ApsaraDB for Relational Database Service (RDS) remote disaster tolerance instance to automatically back up data. You can also manually synchronize the database backup files to Alibaba Cloud servers in other regions.

Security management policies

Interpretations on clauses

- Security policies, regulations, and management personnel are foundational for sustainable security. The policy guides a security direction, the regulation identifies a security process, and persons fulfill security responsibilities.
- Classified protection requirements provide a methodology and best practice. Security can be continuously constructed and managed according to the classified protection methodology.

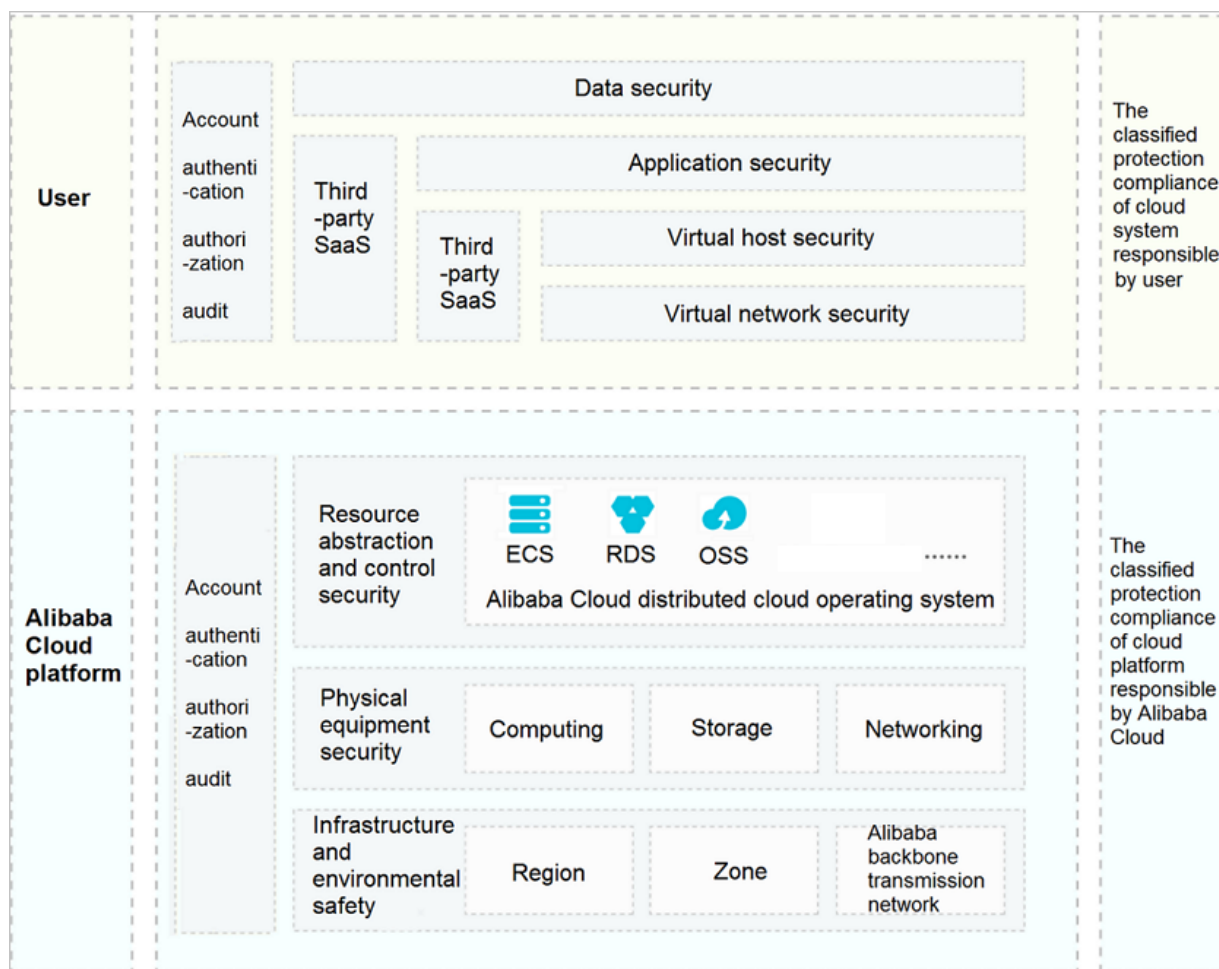
Coping strategies

- Customer management level must arrange, prepare, and fulfill the security policies, regulations, and management personnel according to actual conditions of enterprises and prepare specific documents.
- For technical measures required in the process of vulnerability management, we recommend that you use Alibaba Cloud Server Guard to quickly detect system vulnerabilities on the cloud and resolve them in time.

5.3 Cloud-based classified protection compliance

Shared compliance responsibilities

The Alibaba Cloud platform and cloud tenant systems must be rated and assessed respectively. Assessment conclusions of the Alibaba Cloud platform can be reused by tenant systems in assessment.

Figure 5-1: Shared compliance responsibilities

Alibaba Cloud provides the following contents:

- Classified protection archival filing certification of the Alibaba Cloud platform
- Key pages of the Alibaba Cloud assessment report
- Sales license of Alibaba Cloud Security
- Description of partial assessment items of Alibaba Cloud

Detailed interpretations on shared responsibility are as follows:

- Alibaba Cloud is the unique cloud service provider in China that participates in and passes the pilot demonstration of cloud computing classified protection standard. Public cloud and e-government cloud pass the archival filing and assessment of the third level of classified protection. The Finance Cloud passes the archival filing and assessment of the fourth level of classified protection.
- According to the conclusion reuse rules issued by the supervision authorities, the tenant systems of Alibaba Cloud can reuse conclusions of physical security, partial network security,

and security management when passing classified protection assessment. Alibaba Cloud can provide explanations.

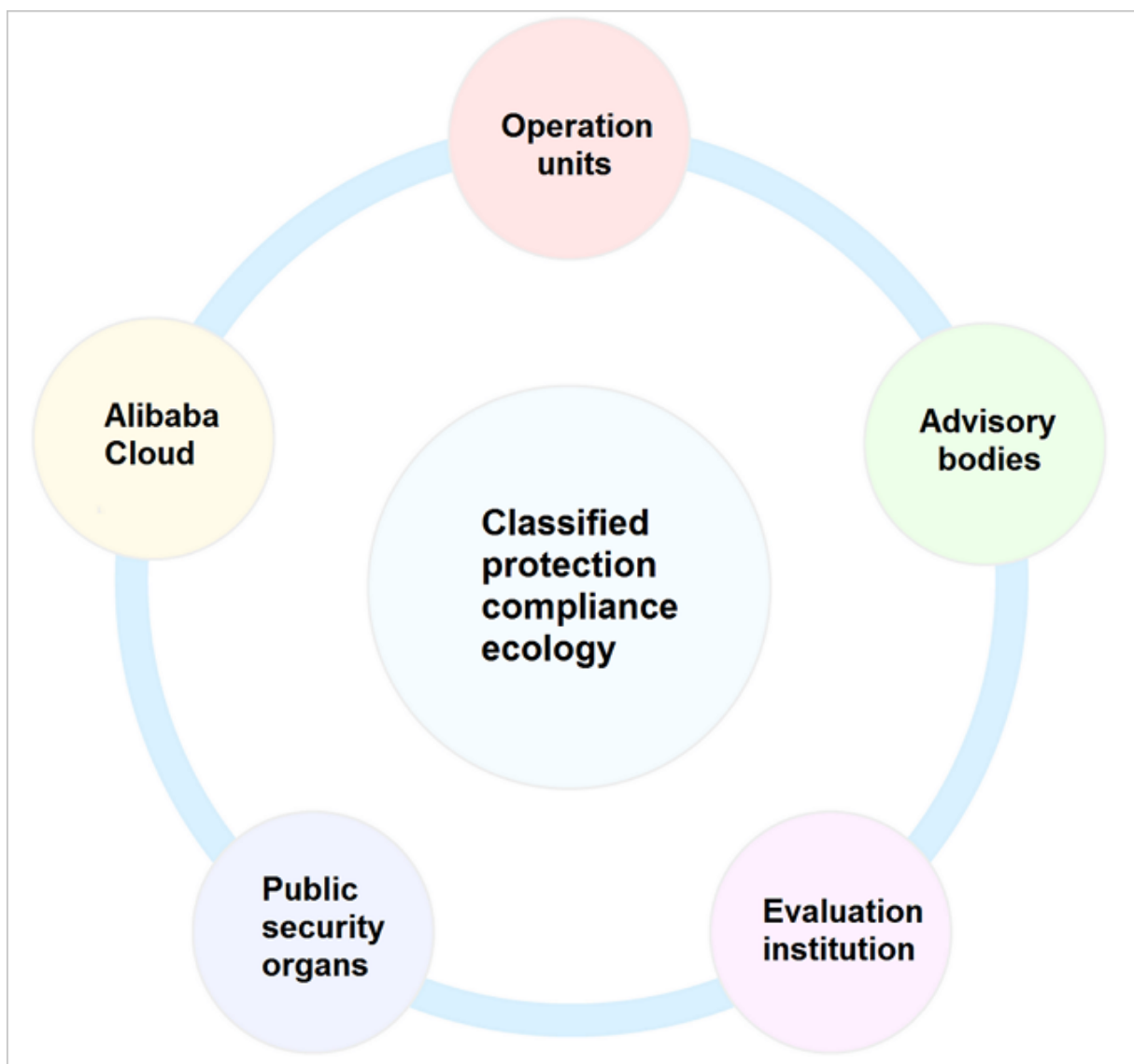
- Complete security technologies and management architecture of the Alibaba Cloud platform and Alibaba Cloud Security protection system facilitate tenants to pass classified protection assessment.

Classified protection compliance ecology

Current conditions of cloud-based classified protection are as follows:

- Most tenants do not know classified protection.
- Most tenants do not know how to start with classified protection.
- Most tenants are not good at communicating with supervision authorities.
- Security systems lag behind business development.

To facilitate cloud-based systems to quickly pass classified protection assessment, Classified Protection Compliance Ecology is established by Alibaba Cloud to provide one-stop classified protection compliance solution.

Figure 5-2: Classified protection compliance ecology

Work division of classified protection:

- Alibaba Cloud: integrates capabilities of service agencies and provides security products.
- Consulting firm: provides technical support and consulting services in the whole process.
- Assessment agency: provides assessment services.
- Public security organ: takes charge of archival filing review, supervision, and inspection.

5.4 Classified protection implementation process

The classified protection implementation process is as shown in [Figure 5-3: Classified protection implementation process](#).

Figure 5-3: Classified protection implementation process

	Operating unit	Alibaba Cloud	Consulting or evaluation agency	Public security organ
System rating	Determine the level of security protection, and write rating report.	Coordinate the third party agencies to provide counseling services for operating units.	Counseling the operating unit to prepare the rating materials and organize expert review. (Level 3 of classified protection)	None.
System filing	Prepare and present the filing materials to the local public security organ.	Coordinate the third party agencies to provide counseling services for operating units.	Counseling the operating unit to prepare the filing materials and file.	None.
Construction rectification	Construct the safety technology and management system in line with grade requirements.	Provide the obligatory security products and services that meet the grade requirements.	Counseling the operating unit to carry out system security reinforcement and develop safety management system.	The local public security organ reviews and accepts the filing materials.
Rating assessment	Prepare for and accept the evaluation from the evaluation agencies.	Provide the cloud service provider's security qualification and the proof that the cloud platform has passed the classified protection.	The evaluation agency evaluates the system level conformity.	None.
Supervision & inspection	Accept the regular inspection of public security organs	None.	None.	Supervise and inspect the operating unit to carry out the class protection work.

5.5 Security compliance architecture

Quickly get access to Alibaba Cloud Security and complete security correction. Satisfy technical requirements for foundational compliance in classified protection with minimal security investment S.

Basic requirements of classified protection are as follows:

- Physical and environmental security: includes measures such as data center power supply, temperature and humidity control, wind prevention, rain prevention, and thunder prevention. Assessment conclusions of Alibaba Cloud can be directly reused.
- Network and communication security: includes network architecture, boundary protection, access control, intrusion prevention, and communication encryption.
- Device and computing security: includes intrusion prevention, malicious code prevention, identity authentication, access control, centralized management and control, and security audit.
- Application and data security: includes security audit, data integrity, and data confidentiality.

5.6 Solution benefits

One-stop classified protection assessment service

Select and cooperate with local consulting and assessment agencies that have high-quality services, provide one-stop and whole-process compliance, and greatly reduce investments of operating units.

- Avoid multi-point communications and repeated work to reduce investments of operating units.
- Greatly improve efficiency and complete assessment in minimal two weeks.
- Alibaba Cloud provides cloud security and compliance best practices.

A complete security prevention system

With a complete Alibaba Cloud Security architecture, an operating unit can locate corresponding products on Alibaba Cloud, correct non-conformances, and completely satisfy classified protection requirements.

6 Elastic Compute Service (ECS)

6.1 What is ECS

Elastic Compute Service (ECS) is a type of computing service that features elastic processing capabilities. As compared with the physical servers, ECS is more user-friendly and can be managed more efficiently. You can create instances, resize disks, and add or release any number of ECS instances any time according to your business demands.

As a virtual computing environment made up of the basic components such as CPU, memory, and storage, an ECS instance is provided by ECS for you to carry out relevant operations. It is the core concept of ECS and you can perform actions on ECS instances on the ECS console. As for other resources such as block storage, images, and snapshots, they cannot be used until being integrated with ECS instances. [Figure 6-1: Concept of an ECS instance](#) illustrates the services supported by an ECS instance.

Figure 6-1: Concept of an ECS instance

6.2 Benefits

Compared to the traditional Internet Data Centers (IDCs) or servers, ECS has the following advantages:

- *High availability*
- *Security*
- *Elasticity*

High availability

Alibaba Cloud adopts more stringent IDC standards, server access standards, and O&M standards to guarantee data reliability and high availability of cloud computing infrastructure and cloud servers.

When even higher availability is needed, you can build active/standby or active/active services in multiple zones. For a finance-oriented solution with three IDCs in two regions, you can deliver services of higher availability with multiple regions and zones. For such services as disaster tolerance and backup, mature solutions are readily available in Alibaba Cloud.

Alibaba Cloud provides you with the following support services:

- Products and services for availability improvement, including cloud servers, server load balancers, multi-backup databases, and Data Transport Services (DTS).
- Industry partners and ecosystem partners that help you build a more advanced and stable architecture and guarantee service continuity.
- Diverse training services that enable you to deliver high availability from the business level to the underlying service level.

Security

Users of cloud computing are most concerned about security and stability. Alibaba Cloud has recently passed a host of international information security certifications, including ISO 27001 and MTCS, which demand strict confidentiality of user data and user information and user privacy protection.

- **Alibaba Cloud VPC offers more business possibilities.** You only need to perform simple configuration to connect your business environment to global IDCs, making your business more flexible, stable, and extensible.
- **You can build a more flexible business** with the powerful network functions from Alibaba Cloud's various hybrid cloud solutions and network products. A superior business ecosystem is possible with Alibaba Cloud's ecosystem.
- **Alibaba Cloud VPC is more stable and secure.**
 - **Stable:** After building your business on VPC, you can update your network architecture and functions on a daily basis as the network infrastructure evolves constantly, allowing your business to run steadily. You can divide, configure, and manage your network on VPC according to your needs.
 - **Secure:** VPC is endowed with traffic isolation and attack isolation to protect your services from endless attack traffic on the Internet. After you build your business on VPC, the first line of defense is established immediately.

VPC provides a stable, secure, fast-deliverable, self-managed, and controllable network environment. With the capability and architecture of VPC hybrid cloud, the technical advantages of cloud computing are open to all industries and enterprises.

Elasticity

Elasticity is the biggest advantage of cloud computing.

- **Elastic computing**

- **Vertical scaling.** Vertical scaling involves modifying the configuration of an individual server, which is hard for traditional IDCs. This, however, is just an easy task for Alibaba Cloud as you can scale up or down ECS or storage capacity based on your transaction volume.

- **Horizontal scaling.** During peak hours for gaming or live video streaming apps, your hands may be tied when a request for additional resources arises in the traditional IDC mode. On the contrary, cloud computing can leverage elasticity to tide you over that period. When the period ends, you can release unnecessary resources to reduce your business cost. With horizontal scaling and auto-scaling, you can determine how and when to scale your resources or implement scaling based on business loads.

- **Elastic storage**

Alibaba Cloud has a powerful elastic storage. When more storage space is required, you can only add servers in the traditional IDC mode, which has a limit on the number of servers that can be added. In the cloud computing mode, however, the sky is the limit. You can order as needed to guarantee sufficient storage space.

- **Elastic network**

Alibaba Cloud also features an elastic network. When you purchase the Alibaba Virtual Private Cloud (VPC), you can have the same network configuration as that of IDCs. In addition, you can have the following benefits:

- Interconnection between data centers
 - Secure domains isolated among data centers
 - Flexible network configuration and planning within the VPC

The elasticity of Alibaba Cloud is reflected in computing, storage, network, and the ability to redesign business architecture. By using Alibaba Cloud, you can work out your business portfolio as you wish.

34

34

34

34



6.4 Features

6.4.1 Instances

6.4.1.1 Overview

An ECS instance is equivalent to a virtual machine that includes CPU, memory, operating system, bandwidth, disks, and other basic computing components. You can easily customize and change the configuration of an instance and you have full control over such a virtual machine.

Unlike a local server, you can use ECS instances and perform such operations as independent management and top-level configuration so long as you log in to Alibaba Cloud.

6.4.1.2 Instance type families

An ECS instance is the minimal unit that can provide computing capabilities and services for your business. ECS instances are available in several type families based on their configuration and business purposes they serve.



Note:

All instance type families listed in this document are for reference purpose only. The specific configurations of your instances are determined by the physical servers that the instances are hosted on.

Type	Feature	Ideal for
n4, general entry-level instance type family	<ul style="list-style-type: none"> vCPU : Memory = 1:2 2.5 GHz Intel Xeon E5-2680 v3 (Broadwell) processors The latest DDR4 memory I/O optimized by default 	<ul style="list-style-type: none"> Small and medium-sized Web servers Batch processing Distributed analysis Advertisement services
mn4, balanced entry-level instance type family	<ul style="list-style-type: none"> vCPU : Memory = 1:4 2.5 GHz Intel Xeon E5-2680 v3 (Broadwell), E5-2680 v4 (Haswell), E5-2682 v4 (Broadwell), or E5-2650 v2 (Haswell) processors The latest DDR4 memory I/O optimized by default 	<ul style="list-style-type: none"> Medium-sized Web servers Batch processing Distributed analysis Advertisement services Hadoop clusters
xn4, compact entry-level instance type family	<ul style="list-style-type: none"> vCPU : Memory = 1:1 	<ul style="list-style-type: none"> Small-sized Web applications Small-sized databases

Type	Feature	Ideal for
	<ul style="list-style-type: none"> 2.5 GHz Intel Xeon E5-2680 v4 (Haswell) or E5-2682 v4 (Broadwell) processors The latest DDR4 memory I/O optimized by default 	<ul style="list-style-type: none"> Applications for development or testing environments Code repositories
e4, memory instance type family	<ul style="list-style-type: none"> vCPU : Memory = 1:8 2.5 GHz Intel Xeon E5-2680 v4 (Broadwell), E5-2680 v3 (Broadwell), E5-2650 v2 (Haswell), or E5-2682 v4 (Broadwell) processors I/O optimized by default 	Applications that involve numerous operations in the memory, searching and computing, for example, Cache/Redis, searching, in-memory database, and so on
sn1ne, compute optimized type family with enhanced network performance	<ul style="list-style-type: none"> vCPU : Memory = 1:2 Ultra high packet forwarding rate 2.5 GHz Intel Xeon E5-2682 v4 (Broadwell) or E5-2680 v4 (Haswell) processors Higher computing specifications matching higher network performance 	<ul style="list-style-type: none"> Scenarios that require receiving and transmitting a large volume of packets, such as live commenting on videos, retransmission of telecommunication services Web front-end servers Massively Multiplayer Online (MMO) game front-ends Data analysis, batch compute, and video coding High performance science and engineering applications
sn2ne, general-purpose type family with enhanced network performance	<ul style="list-style-type: none"> vCPU : Memory = 1:4 Ultra high packet forwarding rate 2.5 GHz Intel Xeon E5-2682 v4 (Broadwell) or E5-2680 v4 (Haswell) processors Higher computing specifications matching higher network performance 	<ul style="list-style-type: none"> Scenarios that require receiving and transmitting a large volume of packets, such as live commenting on videos, retransmission of telecommunication services Enterprise-level applications of various types and sizes Small and medium database systems, caches, and search clusters Data analysis and computing

Type	Feature	Ideal for
		<ul style="list-style-type: none"> Computing clusters and data processing depending on the memory
se1ne, memory optimized type family with enhanced network performance	<ul style="list-style-type: none"> vCPU : Memory = 1:8 Ultra high packet receiving and forwarding rate 2.5 GHz Intel Xeon E5-2682 v4 (Broadwell) or E5-2680 v4 (Haswell) processors Higher computing specifications matching higher network performance 	<ul style="list-style-type: none"> Scenarios that require receiving and transmitting a large volume of packets, such as live commenting on videos, retransmission of telecommunication services High performance databases and high memory databases Data analysis and mining, and distributed memory caches Hadoop, Spark, and other enterprise-level applications with large memory requirements
se1, memory optimized type family	<ul style="list-style-type: none"> vCPU : Memory = 1:8 2.5 GHz Intel Xeon E5-2682 v4 (Broadwell), or E5-2680 v4 (Haswell) processors The latest DDR4 memory Higher computing specifications matching higher network performance I/O optimized by default 	As an instance that uses the memory exclusively, SE1 features a greater ratio of memory to vCPU. It is intended for the scenarios that require fixed performance of computing, such as Cache/Redis, searching, memory databases, high I/O databases (e.g., Oracle, MongoDB), Hadoop clusters, and so on
ebmg5, general-purpose ECS Bare Metal Instance type family	<ul style="list-style-type: none"> vCPU : Memory = 1:4 2.5 GHz Intel Xeon Platinum 8163 (Skylake) processors, 96-core vCPU, up to 2.9 GHz Turbo Boost High network performance: 4.5 million pps packet forwarding rate Supports SSD cloud disks and Ultra cloud disks 	<ul style="list-style-type: none"> Deployment of OpenStack, ZStack, and other private cloud services Deployment of Docker containers and other services Scenarios that require receiving and transmitting a large volume of packets, such as live commenting on

Type	Feature	Ideal for
		videos, retransmission of telecommunication services <ul style="list-style-type: none"> Enterprise-level applications of various types and sizes Medium and large database systems, caches, and search clusters Data analysis and computing Computing clusters and data processing depending on memory
i2, type family with local SSD disks	<ul style="list-style-type: none"> High-performance local NVMe SSD disks with high IOPS, high I/O throughput, and low latency vCPU : Memory = 1:8, designed for high performance databases 2.5 GHz Intel Xeon Platinum 8163 (Skylake) processors Higher computing specifications matching higher network performance 	<ul style="list-style-type: none"> OLTP and high performance relational databases NoSQL databases, such as Cassandra and MongoDB Search applications, such as Elasticsearch
d1, big data type family	<ul style="list-style-type: none"> High-volume local SATA HDD disks with high I/O throughput and up to 17 Gbit/s of bandwidth for a single instance vCPU : Memory = 1:4, designed for big data scenarios 2.5 GHz Intel Xeon E5-2682 v4 (Broadwell) processors Higher computing specifications matching higher network performance 	<ul style="list-style-type: none"> Hadoop MapReduce, HDFS, Hive, HBase, and so on Spark in-memory computing, MLlib, and so on Enterprises that require big data computing and storage analysis to store and compute massive data, for example, companies in the Internet and finance industries Elasticsearch, logs, and so on
sccg5ib, geneneral-purpose Super Computing Cluster	<ul style="list-style-type: none"> vCPU : Memory = 1:8 Ultra high packet forwarding rate 	<ul style="list-style-type: none"> Data analysis and computing AI computing Manufacturing emulation

Type	Feature	Ideal for
(SCC) instance type family	<ul style="list-style-type: none"> 2.5 GHz Intel Xeon Platinum 8163 (Skylake) processors 100G IB network with ultra high bandwidth and ultra low latency 	<ul style="list-style-type: none"> High-performance computing clusters Genetic analysis Pharmaceutical analysis
scch5ib, Super Computing Cluster (SCC) instance type family with high clock speed	<ul style="list-style-type: none"> vCPU : Memory = 1:6 Ultra high packet forwarding rate 3.1 GHz Intel Xeon Gold 6149 (Skylake) processors 100G IB network with ultra high bandwidth and ultra low latency 	<ul style="list-style-type: none"> Data analysis and computing AI computing Manufacturing emulation High-performance computing clusters Genetic analysis Pharmaceutical analysis
re5, type family with enhanced memory	<ul style="list-style-type: none"> Optimized for memory-intensive enterprise applications that involve high-performance databases and in-memory databases 2.5 GHz Intel Xeon Platinum 8163 (Skylake) processors vCPU : Memory = 1:16, up to 2970 GiB of memory 	<ul style="list-style-type: none"> High performance databases and in-memory databases Memory intensive applications Big data engines like Apache Spark and Presto
sn1, compute optimized type family	<ul style="list-style-type: none"> vCPU : Memory = 1:2 2.5 GHz Intel Xeon, E5-2682 v4 (Broadwell), or E5-2680 v3 (Haswell) processors Higher computing specifications matching higher network performance 	<ul style="list-style-type: none"> Web front-end servers Front ends of Massively Multiplayer Online (MMO) games Data analysis, batch compute, and video coding High performance science and engineering applications
sn2, general purpose type family	<ul style="list-style-type: none"> vCPU : Memory = 1:4 2.5 GHz Intel Xeon, E5-2682 v4 (Broadwell), or E5-2680 v3 (Haswell) processors Higher computing specifications matching higher network performance 	<ul style="list-style-type: none"> Enterprise-class applications of various types and sizes Medium and small database systems, cache, and search clusters Data analysis and computing Computing clusters, and data processing depending on memory

Type	Feature	Ideal for
f1, compute optimized type family with FPGA	<ul style="list-style-type: none"> Intel ARRIA 10 GX 1150 FPGA vCPU : Memory = 1:7.5 2.5 GHz Intel Xeon E5-2682 v4 (Broadwell) processors Higher computing specifications matching higher network performance 	<ul style="list-style-type: none"> Deep learning and reasoning Genomics research Financial analysis Picture transcoding Computational workloads, such as real-time video processing and security
f3, compute optimized type family with FPGA	<ul style="list-style-type: none"> Self-developed compute cards based on Xilinx Virtex UltraScale+ VU9P vCPU : Memory = 1:4 2.5 GHz Intel Xeon Platinum 8163 (Skylake) processors Higher computing specifications matching higher network performance 	<ul style="list-style-type: none"> Deep learning Genomics research Video coding and decoding Chip prototype verification Database acceleration

Type	Feature	Ideal for
gn5, compute optimized type family with GPU	<ul style="list-style-type: none"> NVIDIA P100 GPU processors Various ratios of vCPU to memory High-performance NVMe SSD disks 2.5 GHz Intel Xeon E5-2682 v4 (Broadwell) processors Higher computing specifications matching higher network performance 	<ul style="list-style-type: none"> Deep learning Scientific computing, such as computational fluid dynamics, computational finance, genomics, and environmental analysis High performance computing, rendering, multimedia coding and decoding, and other server-side GPU compute workloads
gn4, compute optimized type family with GPU	<ul style="list-style-type: none"> NVIDIA M40 GPU processors Various ratios of CPU to memory 	<ul style="list-style-type: none"> Deep learning Scientific computing, such as computational fluid dynamics, computational

Type	Feature	Ideal for
	<ul style="list-style-type: none"> 2.5 GHz Intel Xeon E5-2680 v4 (Haswell) processors Higher computing specifications matching higher network performance 	<ul style="list-style-type: none"> finance, genomics, and environmental analysis High performance computing, rendering, multimedia coding and decoding, and other server-side GPU compute workloads
ga1, visualization compute type family with GPU	<ul style="list-style-type: none"> AMD S7150 GPU processors vCPU : Memory = 1:2.5 2.5 GHz Intel Xeon E5-2682 v4 (Broadwell) processors High-performance local NVMe SSD disks Higher computing specifications matching higher network performance 	<ul style="list-style-type: none"> Rendering, multimedia coding and decoding Machine learning, high-performance computing, and high performance databases Other server-end business scenarios that require powerful concurrent floating-point compute capabilities
gn5i, compute optimized type family with GPU	<ul style="list-style-type: none"> NVIDIA P4 GPU processors vCPU : Memory = 1:4 2.5 GHz Intel Xeon E5-2682 v4 (Broadwell) or E5-2680 v4 (Haswell) processors Higher computing specifications matching higher network performance 	<ul style="list-style-type: none"> Deep learning Multi-media coding and decoding and other server-side GPU compute workloads
gn5e, GPU compute instance type family	<ul style="list-style-type: none"> I/O-optimized instances 2.5 GHz Intel Xeon Platinum 8163 (Skylake) processors Nvidia P4 GPU processors Higher computing specifications matching higher network performance 	<ul style="list-style-type: none"> Deep learning Video image processing, such as noise reduction, and coding and decoding

The following instance types are applicable for environments upgraded from V2 to V3.

Type	Feature	Ideal for
n1, general entry-level instances	<ul style="list-style-type: none"> vCPU : Memory = 1:2 2.5 GHz Intel Xeon E5-2680 v3 (Haswell) processors Higher computing specifications matching higher network performance I/O-optimized instances Supporting SSD cloud disks and Ultra cloud disks 	<ul style="list-style-type: none"> Small and medium-sized web servers Batch processing Distributed analysis Advertisement services
n2, balanced entry-level instances	<ul style="list-style-type: none"> vCPU : Memory = 1:4 2.5 GHz Intel Xeon E5-2680 v3 (Haswell) processors Higher computing specifications matching higher network performance I/O-optimized instances Supporting SSD cloud disks and Ultra cloud disks 	<ul style="list-style-type: none"> Medium-sized Web servers Batch processing Distributed analysis Advertisement services Hadoop clusters
e3, memory entry-level instances	<ul style="list-style-type: none"> vCPU : Memory = 1:8 2.5 GHz Intel Xeon E5-2680 v3 (Haswell) processors Higher computing specifications matching higher network performance I/O-optimized instances Supporting SSD cloud disks and Ultra cloud disks 	<ul style="list-style-type: none"> Cache/Redis Search Memory databases Databases with high I/O. For example, Oracle and MongoDB Hadoop clusters Computing scenarios that involve massive data processing
c1, instance types of Generation I	<ul style="list-style-type: none"> 1.9 GHz Intel Xeon E5-2420 processors or higher The latest DDR3 memory I/O-optimized and non I/O-optimized at your choice I/O-optimized instances support SSD cloud disks and Ultra cloud disks. 	These instance types are legacy shared instance. They are still categorized by the number of cores (1, 2, 4, 8, and 16 cores) and are not sensitive to type families.

Type	Feature	Ideal for
	<ul style="list-style-type: none"> Non I/O-optimized instances only support basic cloud disks. 	
c2, instance types of Generation I	<ul style="list-style-type: none"> 1.9 GHz Intel Xeon E5-2420 processors or higher The latest DDR3 memory I/O-optimized and non I/O-optimized at your choice I/O-optimized instances support SSD cloud disks and Ultra cloud disks. Non I/O-optimized instances only support basic cloud disks. 	These instance types are legacy shared instance. They are still categorized by the number of cores (1, 2, 4, 8, and 16 cores) and are not sensitive to type families.
m1, instance types of Generation I	<ul style="list-style-type: none"> 1.9 GHz Intel Xeon E5-2420 processors or higher The latest DDR3 memory I/O-optimized and non I/O-optimized at your choice I/O-optimized instances support SSD cloud disks and Ultra cloud disks. Non I/O-optimized instances only support basic cloud disks. 	These instance types are legacy shared instance. They are still categorized by the number of cores (1, 2, 4, 8, and 16 cores) and are not sensitive to type families.
m2, instance types of Generation I	<ul style="list-style-type: none"> 1.9 GHz Intel Xeon E5-2420 processors or higher The latest DDR3 memory I/O-optimized and non I/O-optimized at your choice I/O-optimized instances support SSD cloud disks and Ultra cloud disks. Non I/O-optimized instances only support basic cloud disks. 	These instance types are legacy shared instance. They are still categorized by the number of cores (1, 2, 4, 8, and 16 cores) and are not sensitive to type families.
s1, instance types of Generation I	<ul style="list-style-type: none"> 1.9 GHz Intel Xeon E5-2420 processors or higher 	These instance types are legacy shared instance. They

Type	Feature	Ideal for
	<ul style="list-style-type: none"> The latest DDR3 memory Non I/O-optimized instances Only supporting basic cloud disks 	are still categorized by the number of cores (1, 2, 4, 8 , and 16 cores) and are not sensitive to type families.
s2, instance types of Generation I	<ul style="list-style-type: none"> 1.9 GHz Intel Xeon E5-2420 processors or higher The latest DDR3 memory I/O-optimized and non I/O-optimized at your choice I/O-optimized instances support SSD cloud disks and Ultra cloud disks. Non I/O-optimized instances only support basic cloud disks. 	These instance types are legacy shared instance. They are still categorized by the number of cores (1, 2, 4, 8 , and 16 cores) and are not sensitive to type families.
s3, instance types of Generation I	<ul style="list-style-type: none"> 1.9 GHz Intel Xeon E5-2420 processors or higher The latest DDR3 memory I/O-optimized and non I/O-optimized at your choice I/O-optimized instances support SSD cloud disks and Ultra cloud disks. Non I/O-optimized instances only support basic cloud disks. 	These instance types are legacy shared instance. They are still categorized by the number of cores (1, 2, 4, 8 , and 16 cores) and are not sensitive to type families.
t1, instance types of Generation I	<ul style="list-style-type: none"> 1.9 GHz Intel Xeon E5-2420 processors or higher The latest DDR3 memory Non I/O-optimized instances Only supporting basic cloud disks 	These instance types are legacy shared instance. They are still categorized by the number of cores (1, 2, 4, 8 , and 16 cores) and are not sensitive to type families.

6.4.1.3 Instance types

Instance is the minimum unit for providing computing services, and its type reflects the computing capacity.

For an ECS instance, its type specifies two attributes, its CPU (including model and clock speed), and memory. To determine the scenario, however, you must select the image, disk, and network service at the same time.

Instance type family	Instance type	vCPU (Core)	Memory (GiB)	Local storage (GiB)	Bandwidth (Gbit/s)	Packet forwarding rate (Thousand pps)	NIC queues	ENI (including 1 primary elastic NIC)
n4	ecs.n4.small	1	2.0	N/A	0.5	50	1	1
	ecs.n4.large	2	4.0	N/A	0.5	100	1	1
	ecs.n4.xlarge	4	8.0	N/A	0.8	150	1	2
	ecs.n4.2xlarge	8	16.0	N/A	1.2	300	1	2
	ecs.n4.4xlarge	16	32.0	N/A	2.5	400	1	2
	ecs.n4.8xlarge	32	64.0	N/A	5.0	500	1	2
mn4	ecs.mn4.small	1	4.0	N/A	0.5	50	1	1
	ecs.mn4.large	2	8.0	N/A	0.5	100	1	1
	ecs.mn4.xlarge	4	16.0	N/A	0.8	150	1	2
	ecs.mn4.2xlarge	8	32.0	N/A	1.2	300	1	2
	ecs.mn4.4xlarge	16	64.0	N/A	2.5	400	1	2
	ecs.mn4.8xlarge	32	128.0	N/A	5.0	500	2	8

Instance type family	Instance type	vCPU (Core)	Memory (GiB)	Local storage (GiB)	Bandwidth (Gbit/s)	Packet forwarding rate (Thousand pps)	NIC queues	ENI (including 1 primary elastic NIC)
xn4	ecs.xn4.small	1	1.0	N/A	0.5	50	1	1
e4	ecs.e4.small	1	8.0	N/A	0.5	50	1	1
	ecs.e4.large	2	16.0	N/A	0.5	100	1	1
	ecs.e4.xlarge	4	32.0	N/A	0.8	150	1	2
	ecs.e4.2xlarge	8	64.0	N/A	1.2	300	1	3
	ecs.e4.4xlarge	16	128.0	N/A	2.5	400	1	8
sn1ne	ecs.sn1ne.large	2	4.0	N/A	1.0	300	2	2
	ecs.sn1ne.xlarge	4	8.0	N/A	1.5	500	2	3
	ecs.sn1ne.2xlarge	8	16.0	N/A	2.0	1,000	4	4
	ecs.sn1ne.3xlarge	12	24.0	N/A	2.5	1,300	4	6
	ecs.sn1ne.4xlarge	16	32.0	N/A	3.0	1,600	4	8
	ecs.sn1ne.6xlarge	24	48.0	N/A	4.5	2,000	6	8

Instance type family	Instance type	vCPU (Core)	Memory (GiB)	Local storage (GiB)	Bandwidth (Gbit/s)	Packet forwarding rate (Thousand pps)	NIC queues	ENI (including 1 primary elastic NIC)
	ecs.sn1ne.8xlarge	32	64.0	N/A	6.0	2,500	8	8
sn2ne	ecs.sn2ne.large	2	8.0	N/A	1.0	300	2	2
	ecs.sn2ne.xlarge	4	16.0	N/A	1.5	500	2	3
	ecs.sn2ne.2xlarge	8	32.0	N/A	2.0	1,000	4	4
	ecs.sn2ne.3xlarge	12	48.0	N/A	2.5	1,300	4	6
	ecs.sn2ne.4xlarge	16	64.0	N/A	3.0	1,600	4	8
	ecs.sn2ne.6xlarge	24	96.0	N/A	4.5	2,000	6	8
	ecs.sn2ne.8xlarge	32	128.0	N/A	6.0	2,500	8	8
	ecs.sn2ne.14xlarge	56	224.0	N/A	10.0	4,500	14	8
se1ne	ecs.se1ne.large	2	16.0	N/A	1.0	300	2	2

Instance type family	Instance type	vCPU (Core)	Memory (GiB)	Local storage (GiB)	Bandwidth (Gbit/s)	Packet forwarding rate (Thousand pps)	NIC queues	ENI (including 1 primary elastic NIC)
	ecs.se1ne.xlarge	4	32.0	N/A	1.5	500	2	3
	ecs.se1ne.2xlarge	8	64.0	N/A	2.0	1,000	4	4
	ecs.se1ne.3xlarge	12	96.0	N/A	2.5	1,300	4	6
	ecs.se1ne.4xlarge	16	128.0	N/A	3.0	1,600	4	8
	ecs.se1ne.6xlarge	24	192.0	N/A	4.5	2,000	6	8
	ecs.se1ne.8xlarge	32	256.0	N/A	6.0	2,500	8	8
	ecs.se1ne.14xlarge	56	480.0	N/A	10.0	4,500	14	8
se1	ecs.se1.large	2	16.0	N/A	0.5	100	1	2
	ecs.se1.xlarge	4	32.0	N/A	0.8	200	1	3
	ecs.se1.2xlarge	8	64.0	N/A	1.5	400	1	4
	ecs.se1.4xlarge	16	128.0	N/A	3.0	500	2	8
	ecs.se1.8xlarge	32	256.0	N/A	6.0	800	3	8

Instance type family	Instance type	vCPU (Core)	Memory (GiB)	Local storage (GiB)	Bandwidth (Gbit/s)	Packet forwarding rate (Thousand pps)	NIC queues	ENI (including 1 primary elastic NIC)
	ecs.se1.14xlarge	56	480.0	N/A	10.0	1,200	4	8
ebmg5	ecs.ebmg5.24xlarge	96	384.0	N/A	10.0	4,000	8	32
i2	ecs.i2.xlarge	4	32.0	1 * 894	1.0	500	2	3
	ecs.i2.2xlarge	8	64.0	1 * 1,788	2.0	1,000	2	4
	ecs.i2.4xlarge	16	128.0	2 * 1,788	3.0	1,500	4	8
	ecs.i2.8xlarge	32	256.0	4 * 1,788	6.0	2,000	8	8
	ecs.i2.16xlarge	64	512.0	8 * 1,788	10.0	4,000	16	8
d1	ecs.d1.2xlarge	8	32.0	4 * 5,500	3.0	300	1	4
	ecs.d1.3xlarge	12	48.0	6 * 5,500	4.0	400	1	6
	ecs.d1.4xlarge	16	64.0	8 * 5,500	6.0	600	2	8
	ecs.d1.6xlarge	24	96.0	12 * 5,500	8.0	800	2	8
	ecs.d1-c8d3.8xlarge	32	128.0	12 * 5,500	10.0	1,000	4	8
	ecs.d1.8xlarge	32	128.0	16 * 5,500	10.0	1,000	4	8
	ecs.d1-c14d3.14xlarge	56	160.0	12 * 5,500	17.0	1,800	6	8

Instance type family	Instance type	vCPU (Core)	Memory (GiB)	Local storage (GiB)	Bandwidth (Gbit/s)	Packet forwarding rate (Thousand pps)	NIC queues	ENI (including 1 primary elastic NIC)
	ecs.d1.14xlarge	56	224.0	28 * 5, 500	17.0	1,800	6	8
scch5ib	ecs.scch5ib.16xlarge	64	192.0	N/A	10.0	4,500	8	32
sccg5ib	ecs.sccg5ib.24xlarge	96	384.0	N/A	10.0	4,500	8	32
re5	ecs.re5.15xlarge	60	990.0	N/A	10.0	1,000	16	8
	ecs.re5.30xlarge	120	1,980.0	N/A	15.0	2,000	16	15
	ecs.re5.45xlarge	180	2,970.0	N/A	30.0	4,500	16	15
sn1	ecs.sn1.medium	2	4.0	N/A	0.5	100	1	2
	ecs.sn1.large	4	8.0	N/A	0.8	200	1	3
	ecs.sn1.xlarge	8	16.0	N/A	1.5	400	1	4
	ecs.sn1.3xlarge	16	32.0	N/A	3.0	500	2	8
	ecs.sn1.7xlarge	32	64.0	N/A	6.0	800	3	8
sn2	ecs.sn2.medium	2	8.0	N/A	0.5	100	1	2
	ecs.sn2.large	4	16.0	N/A	0.8	200	1	3
	ecs.sn2.xlarge	8	32.0	N/A	1.5	400	1	4

Instance type family	Instance type	vCPU (Core)	Memory (GiB)	Local storage (GiB)	Bandwidth (Gbit/s)	Packet forwarding rate (Thousand pps)	NIC queues	ENI (including 1 primary elastic NIC)
	ecs.sn2.3xlarge	16	64.0	N/A	3.0	500	2	8
	ecs.sn2.7xlarge	32	128.0	N/A	6.0	800	3	8
	ecs.sn2.14xlarge	56	224.0	N/A	10.0	1,200	4	8

Instance type family	Instance type	vCPU (Core)	Memory (GiB)	Local storage (GiB)	Bandwidth (Gbit/s)	Packet forwarding rate (Thousand pps)	NIC queues	Instance type family	FPGA
f1	ecs.f1-c8f1.2xlarge	8	60.0	N/A	3.0	400	4	4	Intel ARRIA 10 GX 1150
	ecs.f1-c8f1.4xlarge	16	120.0	N/A	5.0	1,000	4	8	2 * Intel ARRIA 10 GX 1150
	ecs.f1-c28f1.7xlarge	28	112.0	N/A	5.0	2,000	8	8	Intel ARRIA 10 GX 1150
	ecs.f1-c28f1.14xlarge	56	224.0	N/A	10.0	2,000	14	8	2 * Intel ARRIA 10

Instance type family	Instance type	vCPU (Core)	Memory (GiB)	Local storage (GiB)	Bandwidth (Gbit/s)	Packet forwarding rate (Thousand pps)	NIC queues	Instance type family	FPGA
									GX 1150
f3	ecs.f3-c16f1.4xlarge	16	64.0	N/A	5.0	1,000	4	8	1 * Xilinx VU9P
	ecs.f3-c16f1.8xlarge	32	128.0	N/A	10.0	2,000	8	8	2 * Xilinx VU9P
	ecs.f3-c16f1.16xlarge	64	256.0	N/A	20.0	2,500	16	8	4 * Xilinx VU9P

Instance type family	Instance type	vCPU (Core)	Memory (GiB)	Local storage (GiB)	Bandwidth (Gbit/s)	Packet forwarding rate (Thousand pps)	NIC queues	Instance type family	GPU
gn5	ecs.gn5-c4g1.xlarge	4	30.0	440	3.0	300	1	3	1 * NVIDIA P100
	ecs.gn5-c8g1.2xlarge	8	60.0	440	3.0	400	1	4	1 * NVIDIA P100
	ecs.gn5-c4g1.2xlarge	8	60.0	880	5.0	1,000	2	4	2 * NVIDIA P100
	ecs.gn5-c8g1.4xlarge	16	120.0	880	5.0	1,000	4	8	2 * NVIDIA P100
	ecs.gn5-c28g1.7xlarge	28	112.0	440	5.0	1,000	8	8	1 * NVIDIA P100

Instance type family	Instance type	vCPU (Core)	Memory (GiB)	Local storage (GiB)	Bandwidth (Gbit/s)	Packet forwarding rate (Thousand pps)	NIC queues	Instance type family	GPU
	ecs.gn5-c8g1.8xlarge	32	240.0	1,760	10.0	2,000	8	8	4 * NVIDIA P100
	ecs.gn5-c28g1.14xlarge	56	224.0	880	10.0	2,000	14	8	2 * NVIDIA P100
	ecs.gn5-c8g1.14xlarge	54	480.0	3,520	25.0	4,000	14	8	8 * NVIDIA P100
gn4	ecs.gn4-c4g1.xlarge	4	30.0	N/A	3.0	300	1	3	1 * NVIDIA M40
	ecs.gn4-c8g1.2xlarge	8	30.0	N/A	3.0	400	1	4	1 * NVIDIA M40
	ecs.gn4.8xlarge	32	48.0	N/A	6.0	800	3	8	1 * NVIDIA M40
	ecs.gn4-c4g1.2xlarge	8	60.0	N/A	5.0	500	1	4	2 * NVIDIA M40
	ecs.gn4-c8g1.4xlarge	16	60.0	N/A	5.0	500	1	8	2 * NVIDIA M40
	ecs.gn4.14xlarge	56	96.0	N/A	10.0	1,200	4	8	2 * NVIDIA M40
ga1	ecs.ga1.xlarge	4	10.0	1 * 87	1.0	200	1	3	0.25 * AMD S7150

Instance type family	Instance type	vCPU (Core)	Memory (GiB)	Local storage (GiB)	Bandwidth (Gbit/s)	Packet forwarding rate (Thousand pps)	NIC queues	Instance type family	GPU
	ecs.ga1.2xlarge	8	20.0	1 * 175	1.5	300	1	4	0.5 * AMD S7150
	ecs.ga1.4xlarge	16	40.0	1 * 350	3.0	500	2	8	1 * AMD S7150
	ecs.ga1.8xlarge	32	80.0	1 * 700	6.0	800	3	8	2 * AMD S7150
	ecs.ga1.14xlarge	56	160.0	1 * 1,400	10.0	1,200	4	8	4 * AMD S7150
gn5i	ecs.gn5i-c2g1.large	2	8.0	N/A	1.0	100	2	2	1 * NVIDIA P4
	ecs.gn5i-c4g1.xlarge	4	16.0	N/A	1.5	200	2	3	1 * NVIDIA P4
	ecs.gn5i-c8g1.2xlarge	8	32.0	N/A	2.0	400	4	4	1 * NVIDIA P4
	ecs.gn5i-c16g1.4xlarge	16	64.0	N/A	3.0	800	4	8	1 * NVIDIA P4
	ecs.gn5i-c16g1.8xlarge	32	128.0	N/A	6.0	1,200	8	8	2 * NVIDIA P4

Instance type family	Instance type	vCPU (Core)	Memory (GiB)	Local storage (GiB)	Bandwidth (Gbit/s)	Packet forwarding rate (Thousand pps)	NIC queues	Instance type family	GPU
	ecs.gn5i-c28g1.14xlarge	56	224.0	N/A	10.0	2,000	14	8	2 * NVIDIA P4
gn5e	ecs.gn5e-c11g1.3xlarge	10	58.0	N/A	2.0	150	1	6	1 * NVIDIA P4
	ecs.gn5e-c11g1.5xlarge	22	116.0	N/A	4.0	300	1	8	2 * NVIDIA P4
	ecs.gn5e-c11g1.11xlarge	44	232.0	N/A	6.0	600	2	8	4 * NVIDIA P4
	ecs.gn5e-c11g1.22xlarge	88	464.0	N/A	10.0	1,200	4	15	8 * NVIDIA P4

The following instance types are only applicable for environments that upgrade from V2 to V3.

Instance type family	Instance type	vCPU (Core)	Memory (GiB)
n1	ecs.n1.tiny	1	1.0
	ecs.n1.small	1	2.0
	ecs.n1.medium	2	4.0
	ecs.n1.large	4	8.0
	ecs.n1.xlarge	8	16.0
	ecs.n1.3xlarge	16	32.0
	ecs.n1.7xlarge	32	64.0

Instance type family	Instance type	vCPU (Core)	Memory (GiB)
n2	ecs.n2.small	1	4.0
	ecs.n2.medium	2	8.0
	ecs.n2.large	4	16.0
	ecs.n2.xlarge	8	32.0
	ecs.n2.3xlarge	16	64.0
	ecs.n2.7xlarge	32	128.0
e3	ecs.e3.small	1	8.0
	ecs.e3.medium	2	16.0
	ecs.e3.large	4	32.0
	ecs.e3.xlarge	8	64.0
	ecs.e3.3xlarge	16	128.0
c1	ecs.c1.small	8	8.0
	ecs.c1.large	8	16.0
c2	ecs.c2.medium	16	16.0
	ecs.c2.large	16	32.0
	ecs.c2.xlarge	16	64.0
m1	ecs.m1.medium	4	16.0
	ecs.m1.xlarge	8	32.0
m2	ecs.m2.medium	4	32.0
s1	ecs.s1.small	1	2.0
	ecs.s1.medium	1	4.0
	ecs.s1.large	1	8.0
s2	ecs.s2.small	2	2.0
	ecs.s2.large	2	4.0
	ecs.s2.xlarge	2	8.0
	ecs.s2.2xlarge	2	16.0
s3	ecs.s3.medium	4	4.0
	ecs.s3.large	4	8.0
t1	ecs.t1.small	1	1.0

6.4.1.4 Instance UserData

As the basis of personalized customization of ECS instances, the UserData function of Alibaba Cloud allows you to customize the startup behaviors of an ECS instance and to pass data into an ECS instance.

UserData is mainly implemented via different types of user-defined scripts. Before this function was introduced, the initially started ECS instances can be thought as having the same environment and configuration. With UserData, enterprises or individuals can enter effective UserData as needed and the initially started instances have the configuration you need.

How to use it

- **UserData-Scripts:** suitable for users who need to initialize instances by running shell scripts, starting with `#!/bin/sh`. In practice, most of the users use this method to enter UserData and it is fit for relatively complex deployments.
- **Cloud-Config:** a unique format supported by cloud-init. With this method, the common personalized configuration is packed in a YAML file, thus implementing the common configuration more conveniently. The first line is `#cloud-config`, followed by an associative array. This method provides such keys as `ssh_authorized_keys`, `hostname`, `write_files`, `manage_etc_hosts`, and so on.

Ideal for

- SSH authentication
- Updating and configuring software resources
- DNS configuration
- Installing and configuring applications

6.4.1.5 Instance lifecycle

The lifecycle of an instance begins with creation and ends with release. This section introduces such information of an instance as its status, status attributes, and corresponding API status.

[Table 6-1: Lifecycle description](#) shows the different states of an instance during its entire lifecycle.

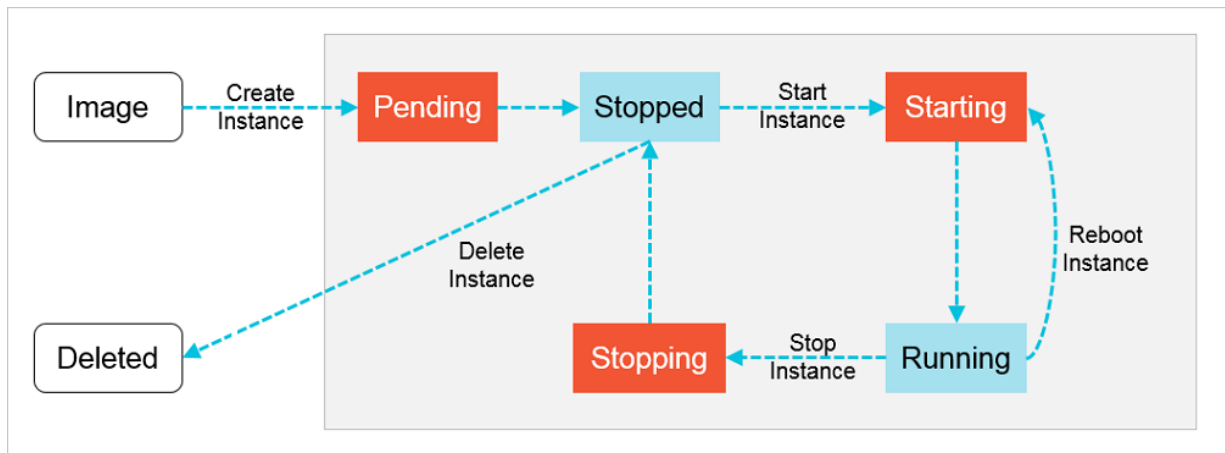
Table 6-1: Lifecycle description

Status	Status attribute	Description	Corresponding API status
Creating task	Intermediate status	Instance creation in progress. Waiting for start. If an instance is	Pending

Status	Status attribute	Description	Corresponding API status
		in this status for a long time, an exception occurs.	
Starting	Intermediate status	It is the state entered by an instance prior to the Running state before you perform a restart or start operation for that instance on the console or via an API. If an instance is in this status for a long time, an exception occurs.	Starting
Running	Stable status	Indicates that the instance is running smoothly. The instance in this state can accommodate your business needs.	Running
Stopping	Intermediate status	An instance is in this status after the Stop operation is performed on the console or using an API but before the instance enters the stop state. If an instance is in this status for a long time, an exception occurs.	Stopping
stop	Stable status	Indicates the instance has been stopped normally. In this status, the instance cannot accommodate external services.	Stopped
Re-initializing	Intermediate status	An instance is in this status after the system disk and/or data disk is re-initialized in the console or using an API until it is Running . If an instance is in this status for a long time, an exception occurs.	Stopped
Replacing System Disk	Intermediate status	An instance is in this status after the operating system is replaced or another such operation is performed on the console using an API until, and before the instance enters the Running state. If an instance is in this status for a long time, an exception occurs.	Stopped

Table 6-1: Lifecycle description describes mappings between console statuses and API statuses. The API status chart is shown in *Figure 6-3: API status chart*.

Figure 6-3: API status chart



6.4.1.6 ECS Bare Metal Instance

ECS Bare Metal Instance is a new type of computing product that features both the elasticity of virtual machines and the performance and characteristics of physical machines. ECS Bare Metal Instance is based on next-generation virtualization technology that is independently developed by Alibaba Cloud.

Compared with the previous generation, next-generation virtualization technology supports both typical ECS instances and nested virtualization. It retains the resource elasticity of common ECS instances and uses nested virtualization to retain the user experience of physical machines.

Benefits

ECS Bare Metal Instance has the following benefits:

- **Exclusive computing resources**

As a cloud-based elastic computing service, ECS Bare Metal Instance provides better performance and isolation than typical physical machines. It can enable exclusive computing resources without virtualization performance overheads and feature loss. ECS Bare Metal Instance supports 8-core, 16-core, 32-core, and 96-core CPU instances and ultrahigh frequency instances. For example, an ECS Bare Metal Instance with 8 cores supports an ultrahigh frequency of up to 3.7 to 4.1 GHz. Compared with similar products, ECS Bare Metal Instance can provide better performance and responsiveness for gaming and financial businesses.

- **Encrypted computing**

To ensure that encrypted data is computed in a safe and trusted environment, ECS Bare Metal Instance uses a chip-level trusted execution environment (Intel® SGX) in addition to the physical server isolation. The chip-level hardware security keeps your data safely isolated on the cloud and gives you control over the entire data encryption and key protection process.

- **Any Stack on Alibaba Cloud ECS Bare Metal Instance**

combines the performance strengths of physical machines with the ease-of-use of ECS instances. This not only meets your requirements for high-performance computing, but also helps you build new hybrid clouds. ECS Bare Metal Instance is also capable of re-virtualization. Local private clouds can be smoothly migrated to Alibaba Cloud without concern for the performance overhead that comes with nested virtualization.

- **Heterogeneous instruction set processor support**

ECS Bare Metal Instance uses the virtualization 2.0 technology developed by Alibaba Cloud to provide support for ARM and other instruction set processors at zero cost.

Features

The following table describes ECS Bare Metal Instance features.

Table 6-2: Features

Configuration	Description
CPU	Supports the EBMG5 family of instance types.
Memory	Allows you to scale up the memory capacity from 32 GiB to 384 GiB. We recommend that you maintain a CPU-to-memory ratio of 1:2 or 1:4.
Storage	Allows you to boot VM images and cloud disks to deploy instances within seconds.
Network	Uses the Virtual Private Cloud (VPC) to connect to other instances , such as ECS and GPU instances. The VPC provides the same performance and stability as physical machine networks.
Image	Allows you to use ECS images.
Security	Provides the same security policies and flexibility as ECS instances.

6.4.2 Cloud disks

6.4.2.1 Overview

Block Storage provides a variety of storage types, including distributed elastic block storage and local storage.

Elastic block storage and local storage are described as follows:

- **Elastic block storage** is a low-latency, persistent, and high-reliability random block-level data storage service provided for ECS users. It uses a triplicate distributed system to ensure ECS instance data reliability. It can be created, released, and scaled up at any time.
- **Local storage** refers to a local disk mounted on the physical machine (host) where the ECS instance resides. It is designed for scenarios with high storage I/O performance requirements. This storage service provides block-level data access for instances, featuring low latency, high random IOPS, and high I/O throughput.

Block storage, OSS, and NAS

Alibaba Cloud provides three types of data storage services, which are Block Storage, Object Storage Service (OSS), and Network Attached Storage (NAS).

Table 6-3: Comparison between storage services

Type	Feature	Scenario
Block Storage	Block Storage is a high-performance and low-latency block storage service provided for ECS users. It allows you to randomly read and write data and format and create file systems just like on physical disks.	It can be used for data storage in many common business scenarios.
OSS	OSS is a big data storage space for large amounts of unstructured data such as images, audios, and videos. You can use APIs to access data stored in OSS anytime, anywhere.	It is typically used for business scenarios such as website construction, separation between dynamic and static resources, and CDN acceleration.
NAS	Similar to OSS, NAS is used to store large amounts of	It is applicable to business scenarios such as cross-

Type	Feature	Scenario
	unstructured data. However, data must be accessed by using standard file access protocols such as Network File System (NFS) in Linux and Common Internet File System (CIFS) in Windows. You can set permissions to allow different clients to access the same file at the same time.	department file sharing, non-linear audio and video editing, high-performance computing, and Docker virtualization.

6.4.2.2 Performance

6.4.2.2.1 Overview

The following contents describe the key measures and performance of block storage products.

6.4.2.2.2 Elastic block storage

Elastic block storage includes cloud disks and shared block storage, whose performance is detailed in the following contents.

Cloud disks



Note:

the following data is obtained with standard tests and configuration.

Table 6-4: Performance data

Block storage	SSD cloud disk	Ultra cloud disk	Basic cloud disk
Capacity of a single disk	32,768 GiB	32,768 GiB	2,000 GiB
Max. IOPS	20,000	3,000	Several hundreds
Max. throughput	300 MBps	80 MBps	20 ~ 40 MBps
Formulas to calculate performance of a single disk	$\text{IOPS} = \min\{30 * \text{capacity}, 20,000\}$ $\text{Throughput} = \min\{50 + 0.5 * \text{capacity}, 300\} \text{ MBps}$	$\text{IOPS} = \min\{1,000 + 6 * \text{capacity}, 3,000\}$ $\text{Throughput} = \min\{50 + 0.1 * \text{capacity}, 80\} \text{ MBps}$	N/A
API name	cloud_ssd	cloud_efficiency	cloud

Block storage	SSD cloud disk	Ultra cloud disk	Basic cloud disk
Typical scenarios	<ul style="list-style-type: none"> I/O-intensive applications Big and medium -sized relational databases NoSQL databases 	<ul style="list-style-type: none"> Small and medium -sized relational databases Large-scale development and testing Web server log 	Applications with occasional access requests or low I/O loads

Shared block storage

Table 6-5: Performance Data

Parameters	SSD shared block storage	Ultra shared block storage
Capacity	<ul style="list-style-type: none"> Single disk: 32,768 GiB All disks attached to an instance: up to 128 TiB 	<ul style="list-style-type: none"> Single disk: 32,768 GiB All disks attached to an instance: up to 128 TiB
Max. random read/write IOPS	30,000	5,000
Max. sequential read/write throughput	512 MBps	160 MBps
Formulas to calculate performance of a single disk	$\text{IOPS} = \min\{40 * \text{capacity}, 30,000\}$	$\text{IOPS} = \min\{1,000 + 6 * \text{capacity}, 5,000\}$
	$\text{Throughput} = \min\{50 + 0.5 * \text{capacity}, 512\} \text{ MBps}$	$\text{Throughput} = \min\{50 + 0.15 * \text{capacity}, 160\} \text{ MBps}$
Typical scenarios	<ul style="list-style-type: none"> Oracle RAC SQL Server Failover clusters High-availability of servers 	<ul style="list-style-type: none"> High-availability of servers High-availability of development and testing databases

6.4.2.2.3 Local storage

For the performance of local disks, see [Local storage](#).

6.4.2.2.4 Performance test

Different tools can be used to test Block Storage performance depending on which operating system the ECS instances are running on.

- Linux: DD, fio, or sysbench

- Windows: fio or Iometer

This topic describes how to use the fio tool to test the performance of Block Storage on an ECS instance deployed in Linux. Before testing, make sure that Block Storage is 4 KiB aligned.



Note:

Although you can obtain more accurate performance data by testing raw devices, there is a risk of damaging the file system structure. Make sure that you back up your data before testing. We recommend that you test Block Storage performance only on ECS instances that contain no data to avoid the risk of losing data.

- Test the random write IOPS:

```
fio -direct=1 -iodepth=128 -rw=randwrite -ioengine=libaio -bs=4k -size=1G -numjobs=1 -runtime=1000 -group_reporting -filename=iotest -name=Rand_Write_Testing
```

- Test the random read IOPS:

```
fio -direct=1 -iodepth=128 -rw=randread -ioengine=libaio -bs=4k -size=1G -numjobs=1 -runtime=1000 -group_reporting -filename=iotest -name=Rand_Read_Testing
```

- Test the write throughput:

```
fio -direct=1 -iodepth=64 -rw=write -ioengine=libaio -bs=64k -size=1G -numjobs=1 -runtime=1000 -group_reporting -filename=iotest -name=Write_PPS_Testing
```

- Test the read throughput:

```
fio -direct=1 -iodepth=64 -rw=read -ioengine=libaio -bs=64k -size=1G -numjobs=1 -runtime=1000 -group_reporting -filename=iotest -name=Read_PPS_Testing
```

The following table describes the fio-related parameters involved in the preceding tests.

Parameter	Description
-direct=1	Ignores the I/O cache and directly writes data when performing tests.
-rw=randwrite	Indicates the read/write policy, which can be randread (random read), randwrite (random write), read (sequential read), write (sequential write), and randrw (random read and write).
-ioengine=libaio	Uses the Linux asynchronous I/O (libaio) method to perform tests. There are two types of I/O synchronization: synchronous and asynchronous. In synchronous I/O, a thread sends only one I/O request to the kernel and waits for the I/O operation to complete. In this case, the iodepth of a single thread is always less than 1. However, you can

Parameter	Description
	use 16 to 32 concurrent threads to fill up the iodepth. In asynchronous I/O, a job uses the libaio method to send multiple I/O requests to the kernel and waits for the I/O operations to complete. Asynchronous I/O reduces the number of interactions and increases operation efficiency.
-bs=4k	Indicates that a single I/O operation uses block files of 4 KB. This parameter is 4 KB by default.
-size=1G	Indicates that the tested file is 1 GB in size.
-numjobs=1	Indicates that the number of tested jobs is 1.
-runtime=1000	Indicates that the test duration is 1,000 seconds. If this parameter is not configured, the file of the size specified by -size is written in blocks of the size specified by -bs.
-group_reporting	Indicates the test result display mode. group_reporting summarizes the statistics of each process, but does not show information of different jobs.
-filename=iotest	Indicates the path and name of the output files. After the test is complete, delete unnecessary files to prevent high disk space usage.
-name=Rand_Write_Testing	Indicates the name of the testing task.

6.4.2.3 Elastic block storage

6.4.2.3.1 Overview

Based on whether they can be attached to multiple ECS instances, elastic block storage products fall into:

- **Cloud disk:** One cloud disk can only be attached to one ECS instance in the same zone.
- **Shared block storage:** One shared block storage can be attached to 4 ECS instances in the same zone at the same time.

6.4.2.3.2 Cloud disk

There are two ways to categorize cloud disks:

- **By performance**

Cloud disks can be categorized by performance into general cloud disks, ultra cloud disks, and SSD cloud disks.

- General cloud disks can provide only hundreds of random IOPS to ECS instances. This makes them ideal for less I/O intensive application scenarios.
- Ultra cloud disks can provide a performance of up to 3,000 random IOPS to ECS instances. This makes them ideal for moderately I/O intensive application scenarios.
- SSD cloud disks can provide stable performance with high random IOPS to ECS instances. This makes them ideal for highly I/O intensive application scenarios.

- **By function**

Cloud disks can be categorized by function into system disks and data disks.

- System disks have the same lifecycle as the ECS instances to which they are mounted. System disks are created and released at the same time as ECS instances. System disks cannot be shared.
- Data disks can be created at the same time as ECS instances or independently. Data disks cannot be shared. A data disk that is created at the same time as the ECS instance has the same lifecycle as the instance. It is also released at the same time as the instance. Data disks that are created independently from the ECS instance can be released at the same time as the ECS instance or independently. The capacity of a data disk is determined by the cloud disk type. For more information, see [Performance](#).

6.4.2.3.3 Shared block storage

Shared block storage is a block level data storage service with high level of concurrency, performance, and reliability. It supports concurrent reads/writes to multiple ECS instances. One shared block storage can be attached to 4 ECS instances at the same time.

Shared block storage can only be used as data disks and created separately. Shared access is allowed. Shared block storage can be configured to be released with the ECS instances.

Based on their performance, the shared block storage can be divided into:

- **SSD shared block storage:** It adopts SSD as the storage medium to provide stable and high-performance storage with enhanced random I/O and data reliability.
- **Ultra shared block storage:** It adopts the hybrid media of SSD and HDD as the storage media.

When used as data disks, shared block storage shares the data disk quota with cloud disks, that is , up to 16 data disks can be attached to one ECS instance.

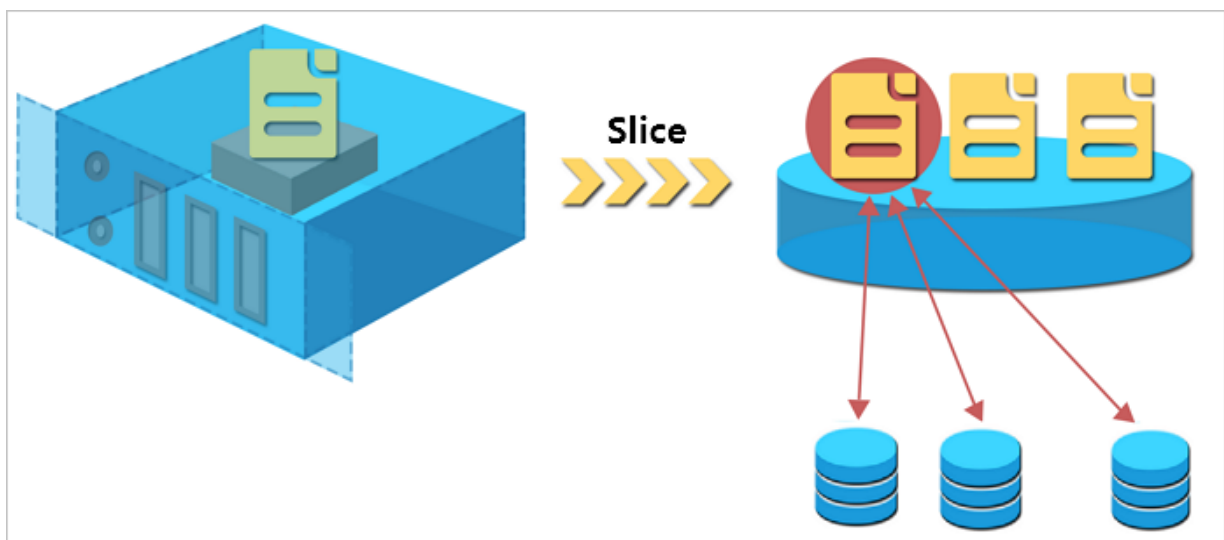
6.4.2.3.4 Triplicate technology

The Alibaba Cloud Distributed File System provides stable and efficient data access and reliability for ECS.

Chunks

ECS users' reads/writes to virtual disks are mapped to the reads/writes to the files on the file platform of Alibaba Cloud. The Distributed File System of Alibaba Cloud uses a flat design in which a linear address space is divided into slices, also called chunks. Each chunk has three copies stored on different server nodes on different racks, thus guaranteeing data reliability.

Figure 6-4: Triple replication of data



How triplicate technology works

Triplicate technology involves three key components: Master, Chunk Server, and Client. To demonstrate how triplicate technology works, in this example, the write operation of an ECS user undergoes several conversions before being executed by the Client. The process is as follows:

1. The Client determines the location of a chunk corresponding to one of your write operations.
2. The Client sends a request to the Master to query the storage locations (that is, the Chunk Servers) of the three copies of the chunk.
3. The Client sends write requests to the corresponding three Chunk Servers according to the results returned from the Master.
4. The Client returns a message to the user indicating whether the operation was successful.

The distribution strategy of the Master takes into account such factors as the disk usage of all the Chunk Servers in a rack, how they are distributed in different racks, availability of power supply and machine workloads, thereby guaranteeing that all the copies of a chunk are distributed on different Chunk Servers on different racks. This approach effectively reduces the potential of total data loss caused by failure of a Chunk Server or a rack.

Data protection

If a system failure occurs because of a corrupted node or hard drive failure, some chunks may lose one or more of the three valid chunk copies associated with them. If this occurs and triplicate technology is enabled, the Master replicates data between Chunk Servers to restore the missing chunk copies across different nodes.

Figure 6-5: Auto sync. of data



As described above, whenever users add, modify or delete data on the cloud disks, their operations are synchronized to the three copies. By doing so, the reliability and consistency of users' data is guaranteed.

For data loss in ECS instances caused by virus, operational mistakes or hackers, solutions include backup, snapshots, and so on. However, there is no single technology that solves all the issues and it is important to take appropriate measures to protect your data based on the actual situation.

6.4.2.4 ECS disk encryption

As a simple and secure encryption method, ECS disk encryption encrypts newly created cloud disks. You do not have to create, maintain, or protect your own key management infrastructure, nor change any of your existing applications or O&M processes. In addition, no extra encryption /decryption operations are required, so your business is not impacted by the disk encryption function.

After an encrypted cloud disk is created and attached to an ECS instance, the data in the following list can be encrypted:

- Data on the cloud disk.
- Data transmitted between the cloud disk and the instance. However, data in the instance operating system is not encrypted.
- All snapshots created from the encrypted cloud disk, which are called encrypted snapshots.

Encryption and decryption are performed on the host that runs the ECS instance, so the data transmitted from the ECS instance to the cloud disk is encrypted.

ECS disk encryption supports all available cloud disks (basic cloud disks, ultra cloud disks, and SSD cloud disks) and shared block storage (ultra and SSD) in a VPC.

6.4.2.5 Local storage

Local storage, also called local disks, refers to the disks attached to the physical servers (host machines) where ECS instances are hosted. They provide temporary block level storage for instances, featuring low latency, high random IOPS, and high I/O throughput. They are designed for business scenarios requiring high storage I/O performance.

Because a local disk is attached to a single physical server, the data reliability depends on the reliability of the physical server, which is subject to the single point failure. We recommend that you implement data redundancy at the application layer to guarantee data availability.



Note:

Using a local disk for data storage comes with the risk of losing your data in some cases, such as when the host machine is down. Therefore, never store any business data that requires long-term persistence on a local disk. If no data reliability architecture is available for your application, we strongly recommend that you build your ECS with cloud disks or shared block storage.

Categories

Currently, Alibaba Cloud provides two types of local disks:

- **Local NVMe SSD:** This disk is used together with instances of the following type families: gn5 and gal.
- **Local SATA HDD:** This disk is used together with instances of the d1ne and d1 type families. It is applicable to the Internet, finance, and other allied industries that require big data computing and storage analysis for massive data storage and offline computing. It fully meets the needs

of distributed computing business models represented by Hadoop regarding instance storage performance, capacity, and Intranet bandwidth.

SATA HDD

The following table lists the performance of local SATA HDD of a d1ne or d1 ECS instance.

Table 6-6: Performance

Parameters	SATA HDD
Capacity	<ul style="list-style-type: none">Single disk: 5,500 GiBTotal capacity per instance: 154,000 GiB
Throughput	<ul style="list-style-type: none">Single disk: 190 MBpsTotal throughput per instance: 5,320 MBps
Access latency	In milliseconds

6.4.3 Images

An image is a running environment template for ECS instances. It includes an operating system and in some cases, pre-installed software. An image file is like a duplicate that contains all the data of one or more disks. For ECS, such disks can be a single system disk or the combination of a system disk and data disks. You can use an image to create an ECS instance or change the system disk of an ECS instance.

Image types

ECS provides various types of images for you to easily access image resources.

Table 6-7: Image types

Type	Description
Public image	<p>Public images officially provided by Alibaba Cloud support nearly all mainstream Windows and Linux versions. Including:</p> <ul style="list-style-type: none">WindowsCentOSCoreOSDebianGentooFreeBSDOpenSUSE

Type	Description
	<ul style="list-style-type: none"> • SUSE Linux • Ubuntu
Custom image	Custom images are created based on your existing physical server, virtual machine, or cloud host. These images can meet your personalized needs flexibly.

How to obtain an image

You can obtain an image for your ECS instance by:

- Creating a custom image based on an existing ECS instance.
- Choosing an image shared from other accounts.
- Importing a local image file into an ECS cluster to generate a custom image.
- Copying a custom image to other regions to maintain a consistent deployment of environment and application across multiple regions.

Image format

Currently, ECS supports images in VHD and RAW formats. You must convert other formats into VHD or RAW to use them in ECS. For how to convert a format, please refer to **How to convert the image format** in *ECS User Guide*.

6.4.4 Snapshots

6.4.4.1 Overview

A snapshot is a copy of data on a disk created at a specific point in time. In reality, you may have the following needs:

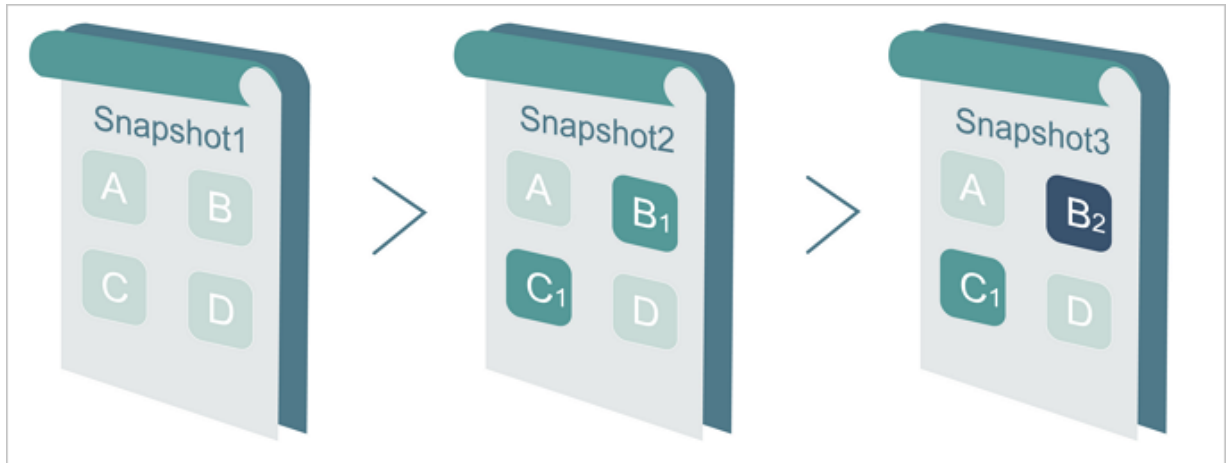
- When writing or storing data on multiple disks, you want to use snapshot data from one disk as the basis for other disks.
- While cloud disks represent a secure way to store data, the data on them may be subject to errors (for example, data errors due to application errors or malicious read/write by hackers), which requires other mechanism to safeguard your data. For this reason, you may want to use snapshots to restore data to a previous point in time even if cloud disks are used.

6.4.4.2 Incremental snapshot mechanism

A snapshot is a copy of data in a certain point in time. You can create snapshots as scheduled to guarantee business continuity.

Snapshots are created in an incremental mode, that is, only data changes between two snapshots are copied. *Figure 6-6: Schematic diagram of snapshot* shows the incremental snapshot process.

Figure 6-6: Schematic diagram of snapshot



As shown above, Snapshot1, Snapshot2, and Snapshot3 are the first, second, and third snapshot of a disk respectively. During the snapshot creation process, the Distributed File System checks the disk data by blocks. Only the blocks with changed data are copied to the snapshot. In the preceding figure:

- In Snapshot 1, all data on the disk is copied because it is the first disk snapshot.
- Snapshot 2 only copies the changed data blocks B₁ and C₁. Data blocks A and D only reference their counterparts in Snapshot 1.
- Snapshot 3 copies the changed data block B₂ but references data blocks A and D from Snapshot 1 and data block C₁ from Snapshot 2.
- When you roll back the disk to Snapshot 3, blocks A, B₂, C₁, and D are copied to the disk to replicate Snapshot 3.
- When you delete Snapshot 2, block B₁ is deleted, but block C₁ is retained because blocks that are referenced by other snapshots cannot be deleted. When you roll back a disk to Snapshot 3, block C₁ is recovered.



Note:

Snapshots are stored in the Object Storage Service (OSS), but are invisible to users. They do not consume the bucket space in OSS. Snapshot operations can only be performed by using the ECS console or APIs.

6.4.4.3 ECS Snapshot 2.0

Built on original basic snapshot features, ECS Snapshot 2.0 data backup service provides a higher snapshot quota and more flexible automatic snapshot policies, further reducing its impact on business I/O. The features of ECS Snapshot 2.0 are described in the following table.

Table 6-8: Snapshot 2.0 vs. original Snapshot

Feature	Original snapshot specifications	Snapshot 2.0 specifications	Benefits of using ECS Snapshot 2.0	Comments
Snapshot quota	Number of disks * 6 + 6.	64 snapshots for each disk.	Longer protection circle; smaller protection granularity.	<ul style="list-style-type: none"> Snapshot backup of a data disk for non-core business occurs at 00:00 every day. The backup data is retained for over two months. Snapshot backup of a data disk for core business occurs every four hours. The backup data is retained for over 10 days.
Automatic snapshot policy	Triggered once daily by default. It cannot be modified.	Customizable weekly snapshot day, time of day, and snapshot retention period. You can query the quantity and details of disks associated with an automatic snapshot policy.	More flexible protection policy.	<ul style="list-style-type: none"> You can take snapshots on the hour and several times in a day. You can choose any day of the week as the recurring day for taking snapshots. You can specify the snapshot retention period or choose to retain it permanently (when the maximum number of automatic snapshots is reached, the oldest automatic snapshot is automatically deleted).
Implementation	COW (Copy-On-Write).	ROW (Redirect-On-Write).	Mitigated performance impact of the	No interruption to your business, allowing snapshots to be taken at any time.

Feature	Original snapshot specifications	Snapshot 2.0 specifications	Benefits of using ECS Snapshot 2.0	Comments
			snapshot task on business I/O writes.	

6.4.4.4 ECS Snapshot 2.0 vs. traditional storage products

Alibaba Cloud ECS Snapshot 2.0 has many advantages as compared with the snapshot feature of traditional storage products, as described in the following table.

Table 6-9: Comparison of technical advantages

Item	ECS Snapshot 2.0	Traditional snapshot products
Capacity	Unlimited capacity that meets the requirements of protecting ultra-large-scale business data.	Limited capacity, often determined by the initial storage device capacity, which meets the requirements of protecting the core business data.

6.4.5 Deployment sets

You may require higher reliability and performance when buying multiple ECS instances in the same zone, for example:

- **Improve business reliability**

To avoid interrupting business when a physical host, rack or switch malfunctions, you may want the same instance to be distributed on different hosts, racks or switches.

- **Improve business network performance**

There are many interactions among instances in some business scenarios. You may want to have a lower network latency and higher network bandwidth. In this case, it is necessary to put instances together on one switch to meet your needs.

ECS provides deployment sets for you to perceive the physical topology of hosts, racks and switches, and to choose appropriate deployment policies according to your business types to improve overall business reliability and performance.

Deployment granularities and policies

- **Deployment granularities**

- Host: means that the minimum scheduling granularity is a physical server. It is also the default value.
- Rack: means that the minimum scheduling granularity is a rack.
- Switch: means that the minimum scheduling granularity is a network switch.

- **Deployment policies**

- LooseAggregation
- StrictAggregation
- LooseDispersion
- StrictDispersion

Where, LooseAggregation and StrictAggregation are intended for higher performance, while LooseDispersion and StrictDispersion are intended for higher reliability.

For deployment policies and business scenarios that correspond to each deployment granularities, see the [Table 6-10: Granularities and policies](#).

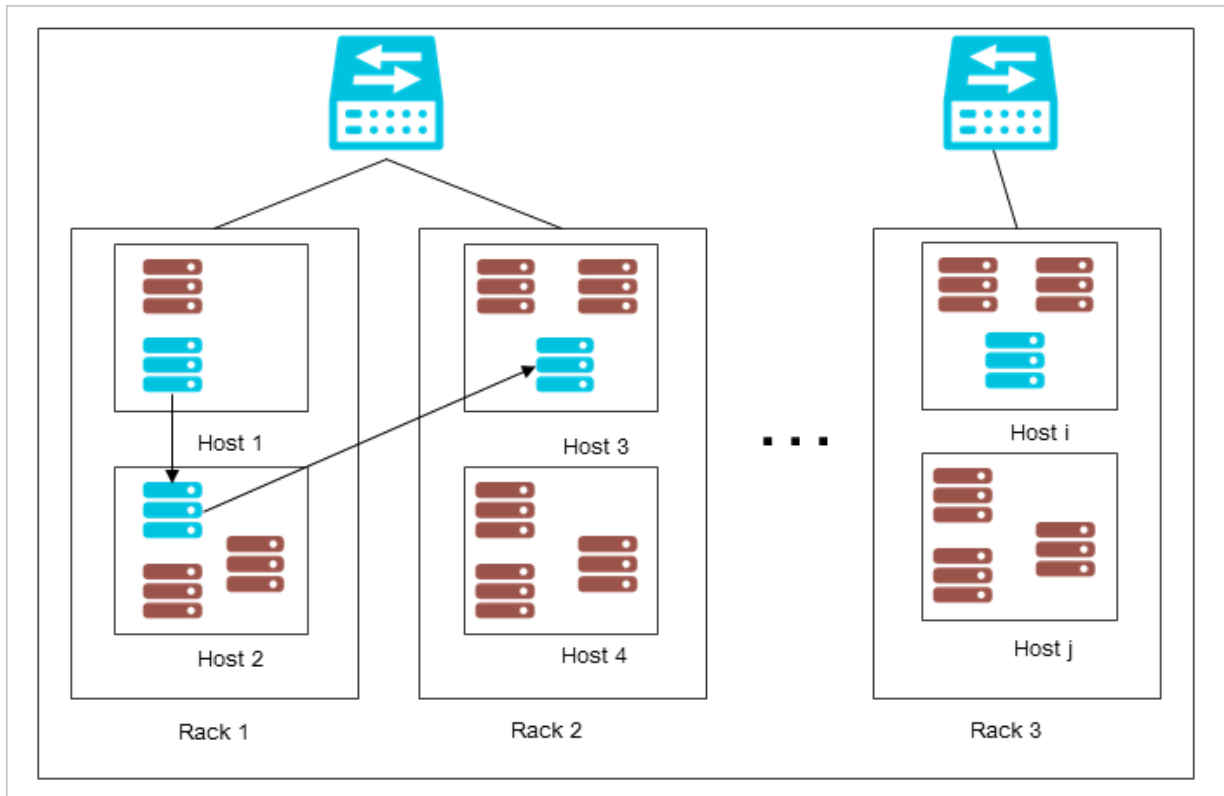
Table 6-10: Granularities and policies

Deployment granularity	Deployment policy	Business scenario
Host	StrictDispersion	General
	LooseDispersion	
Rack	Strict Distribution	Big data, Databases
	LooseDispersion	Game clients
Switch	StrictDispersion	VPN
	LooseDispersion	Game clients
	StrictAggregation	Big data, Databases
	LooseAggregation	Game clients

Example

The figure below shows a typical example that improves business reliability through deployment sets. The three ECS instances of the user are distributed on three physical hosts on at least two racks.

Figure 6-7: Example



Note:

For specific APIs related to deployment sets, see **Deployment sets** in *ECS User Guide*.

6.4.6 Network and security

6.4.6.1 IP address of a VPC

This section introduces the supported type of IP addresses and the specific application scenario.

IP address type

ECS instances have the following IP address type:

- **Private IP**

A private IP is assigned when you create an ECS instance according to the VPC and switch segment that the instance belongs to.

Application scenario

- **Private IP:** A private IP is used to access intranet. You can directly configure the private IP when you create an ECS instance.

**Note:**

The system automatically assigns a private IP if it is not configured.

6.4.6.2 Elastic network interface

This section briefly introduces the concept and application scenarios of elastic network interface.

Elastic network interface (ENI), also known as the auxiliary network interface, is a kind of virtual network interface that can be attached to ECS instances in a VPC. It can help you build highly available clusters, and implement low-cost failover and fine network management.

ENIs are applicable for the three scenarios:

- **Build highly available clusters**

Meet the highly available architecture's demands for a single instance with multiple network interfaces.

- **Low-cost failover**

By separating ENIs from an ECS instance and attaching them to another, you can rapidly migrate business traffic on a malfunctioning instance to a standby instance to resume the service.

- **Fine network management**

An instance can have multiple ENIs such as ENIs for internal management and those for Internet business access, thus separating management data from business data. You can configure accurate security group rules for each ENI to ensure that their traffic is under secure access control.

ENI properties

The table below shows the ENI information.

Table 6-11: Property description

Property	Number
Master private IP address	1
MAC address	1
Security group	1 ~ 5
Description	1
Name of network interface	1

Limitations

ENIs can be created and then attached to or detached from instances. However, ENIs have the following limitations:

- For one account, the maximum number of ENIs that can be created in one region is 100.
- ECS instances and ENIs must be in the same zone of the same VPC, but can belong to different switches.
- For instance types that support ENIs and the number of ENIs that they support, see [Instance types](#).
- The instance bandwidth will not increase by attaching multiple ENIs.

**Note:**

The instance bandwidth is determined by instance type.

6.4.6.3 Intranet

Currently, Alibaba Cloud servers communicate through the Intranet using 1 Gbit/s shared bandwidth for non-I/O optimized instances, and 10 Gbit/s shared bandwidth for I/O optimized instances. However, because the servers communicate over a shared network, the bandwidth may fluctuate.

**Note:**

Currently, most mainstream instances are I/O optimized, and the actual bandwidths are related to physical hardware.

If you need to transmit data between two ECS instances in the same region, Intranet communication is recommended. Intranet communication is also recommended for RDS, Server Load Balancer, and OSS as they communicate with the 1 Gbit/s shared bandwidth in Intranet as well.

Currently, RDS, Server Load Balancer, and OSS can communicate with ECS through Intranet directly so long as they are in the same region.

For ECS instances in a **VPC** of Intranet:

- For instances in the same region and VPC and under the same account, Intranet communication is enabled by default if they are in the same security group. If instances are in different security groups, Intranet communication can be enabled via authorization among security groups.
- For instances under the same account and region while in different VPCs, Intranet communication can be enabled via Express Connect.
- Private IP addresses of instances can be modified or replaced.
- Private and public IP addresses of instances do not support virtual IP (VIP) configuration.
- Instances of different network types cannot communicate with each other in Intranet.

6.4.6.4 Security group rules

Security group rules can allow or deny inbound or outbound traffic of the Internet and Intranet for ECS instances.

You can authorize or cancel security group rules at any time. Changes to security group rules are automatically applied to ECS instances associated with security groups.

When configuring security group rules, make sure the rules are concise. If you associate an instance with multiple security groups, hundreds of rules may apply to the instance, which may cause connection errors when you access the instance.

6.5 Application scenarios

ECS is a highly flexible solution since it can be used either independently as a simple web server, or with other Alibaba Cloud products, such as OSS and CDN, to provide advanced multimedia solutions.

ECS can be used in the following scenarios:

Corporate websites and simple web applications

In the initial stage, corporate websites have low traffic volumes and require only low-configuration ECS instances to run applications, databases, storage files, and other resources. As your business expands, you can upgrade the ECS configuration and increase the number of ECS instances at any time. You no longer need to worry about insufficient resources during peak traffic.

Multimedia and large-traffic apps or websites

ECS can be used with OSS to store static images, videos, and downloaded packages, reducing storage fees. In addition, ECS can be used with CDN or Server Load Balancer to greatly reduce response time and bandwidth fees, thus improving availability.

Apps or websites with significant traffic fluctuations

Some applications like the 12306 website may encounter large traffic fluctuations within a short period. When ECS is used with Auto Scaling, the number of ECS instances is automatically adjusted based on traffic. This feature allows you to meet resource requirements at a low cost. ECS can be used with Server Load Balancer to implement a high availability architecture.

Databases

ECS supports databases with high I/O requirements. An I/O-optimized ECS instance with high configuration can be used with an SSD cloud disk to deliver high I/O concurrency and higher data reliability. Alternatively, multiple I/O-optimized ECS instances with lower configuration can be used with Server Load Balancer to build a highly available architecture.

6.6 Usage limitations

ECS has the following limitations:

- For a cloud server with 4 GiB or more RAM, select the 64-bit operating system (the 32-bit operating system has the 4GiB RAM addressing limitation).
- Windows 32-bit Operating System supports up to 4 vCPUs.
- Windows does not support instance types that have more than 64 vCPUs.
- ECS does not support virtual application installation or re-virtualization (for example, VMware).
- ECS does not support sound card applications (currently, only GPU instances support analog sound card) or directly loading external hardware device, such as hardware dongle, USB memory, external hard disk and bank U key.
- ECS does not support multicast protocol. If multicasting services are required, we recommend that you use unicast point-to-point method.

Besides the preceding limitations, others are mentioned in the following table.

Table 6-12: Other limitations

Type	Limitation description
Instance type	For specific limitations, see Instance type families and Instance types .

Type	Limitation description
Block storage	<p>Type limitation</p> <ul style="list-style-type: none"> • Quota of system disks for one ECS instance: 1. • Quota of data disks for one ECS instance: 16. • Instances to which one shared block storage can be attached: 4. • Capacity of the system disk: 40 GiB ~ 500 GiB. • Capacity of one Basic Cloud Disk: 5 GiB ~ 2,000 GiB. • Capacity of one SSD Cloud Disk: 20 GiB ~ 32,768 GiB. • Capacity of one Ultra Cloud Disk: 20 GiB ~ 32,768 GiB. • Total capacity of one Ultra Block Storage: 32,768 GiB. <p>Usage limitations</p> <ul style="list-style-type: none"> • Only data disks can be encrypted, while system disks cannot. • Existing non-encrypted disks cannot be directly converted into encrypted disks. • Existing encrypted disks cannot be converted into non-encrypted disks either. • Snapshots generated from existing non-encrypted disks cannot be directly converted into encrypted snapshots. • Encrypted snapshots cannot be converted into non-encrypted snapshots either. • Images with encrypted snapshots cannot be shared. • Images with encrypted snapshots cannot be exported.
Quota of snapshots	Number of disks × 64.
Images	<ul style="list-style-type: none"> • Quota of accounts to share one custom image: 50. • Requirements of images for instance types: 32-bit images are not supported on an instance with 4 GiB or more RAM.
Security groups	<ul style="list-style-type: none"> • The number of instances within one security group cannot exceed 1,000. If more than 1,000 instances need to access each other in Intranet, they can be divided into several security groups and access each other via mutual authorization. • Each instance can join up to 5 security groups. • Each account can have up to 100 security groups. • Each security group can have up to 100 security group rules. • Adjustment operation on security group will not interrupt your service continuity. • Security groups are stateful. If data packets are permitted in the outbound direction, so are the packets in the inbound direction.

Type	Limitation description
ENI	For the number of ENIs that can be bound to different instance type families, see Instance types .
Instance UserData	Currently, the ECS UserData feature only supports VPC+I/O optimized instance. Besides, cloud-init should be installed in the image as UserData depends on the cloud-init service. For details, see Install cloud-init in <i>ECS User Guide</i> .

6.7 Basic concepts

Cloud server

A simple and efficient cloud computing service with elastic processing capacity that is supported in Linux, Windows, and other operating systems.

Instance

An independent virtual machine that includes basic cloud computing components such as CPU, memory, operating system, bandwidth, disks, and so on.

Security group

A security group is a virtual firewall that has such functions as detecting the state, filtering packets, and setting up network access control for one or more cloud servers. Instances within one security group are interconnected. Instances between different security groups can only access each other through authorization between the two security groups.

Image

A running environment template for ECS instances. It generally includes an operating system and preinstalled software. There are three types of images: public images, custom images, and shared images. You can use an image to create an ECS instance or change the system disk of an ECS instance.

Snapshot

Data backup of a disk at a time point. Snapshots are classified into Auto Snapshot and Manual Snapshot.

Cloud disk

A kind of independent disk that can be attached to any ECS instance in the same region and zone. Cloud disks are classified into Ultra Cloud Disk, SSD Cloud Disk and Basic Cloud Disk according to their performances.

Alibaba Cloud Block Storage (Block Storage)

A low-latency, persistent, high-reliability, block-level, random storage for ECS instances.

Throughput

The amount of data successfully transmitted through a network, device, port, virtual circuit, or other facilities within a unit time.

Performance Testing Service (PTS)

A world-class, powerful testing platform that simulates real-world business scenarios involving massive users to observe real world capabilities and identify limitations.

Alibaba Cloud Virtual Private Cloud (VPC)

An Alibaba Cloud Virtual Private Cloud (VPC) is a private network built and customized based on Alibaba Cloud. Full logical isolation is achieved between Alibaba VPCs. Users can create and manage cloud product instances, such as ECS, Intranet Server Load Balancer, and RDS in their own VPCs.

Private IP address

A connection address used to access the host on a private network.

GPU cloud server

Computing service based on GPU applications that is applicable for video decoding, graphics rendering, deep learning, scientific computing, and other scenarios. GPU cloud server features real-time, high-speed, concurrent computing and powerful floating-point computing capacity.

7 Auto Scaling (ESS)

7.1 What is ESS

Auto Scaling (ESS) is a management service that automatically adjusts the number of elastic computing resources based on your business demands and strategies.

Based on user-defined scaling rules, ESS automatically adds ECS instances as business loads increase to ensure sufficient computing capabilities. When your business loads decrease, ESS automatically removes ECS instances to reduce running costs.

ESS provides the following functions:

- **Elastic scale-out**

When business loads surge, ESS automatically increases underlying resources. This helps maintain access speed and ensure that resources are not overloaded. For example, if the CPU utilization of ECS instances exceeds 80%, ESS scales out ECS resources based on the rules you defined. During the scale-out process, ESS automatically creates and adds ECS instances to a scaling group, and adds the new instances to the SLB instance and RDS whitelist.

- **Elastic scale-in**

When business loads decrease, ESS automatically releases underlying resources. This prevents resource wastage and helps to reduce cost. For example, if the CPU utilization of ECS instances in a scaling group falls below 30%, ESS scales in ECS resources based on the rules you defined. During the scale-in process, ESS removes the ECS instances from the scaling group, the SLB instance, and RDS whitelist.

- **Elastic recovery**

The health status of ECS instances in a scaling group is determined based on the life cycle of the instances. If an ECS instance is in an unhealthy state, ESS automatically releases the instance and creates a new one. ESS adds the new instance to the SLB instance and RDS whitelist. This process is called elastic recovery. It ensures that the number of healthy ECS instances in a scaling group will not fall below the threshold that you defined.

7.2 Benefits

ESS has the following benefits:

- **Automatic scaling of instances on-demand**

ESS can automatically add ECS instances during peak traffic hours, and remove ECS instances during off-peak hours to scale with actual business needs. This helps to lower infrastructure costs because you only pay for what you actually use.

- **Real-time instance monitoring and automatic replacement of unhealthy instances**

ESS performs real-time monitoring on instances and automatically replaces unhealthy instances that are discovered, reducing operations and maintenance (O&M) overheads.

- **Intelligent whitelist management and control, no user intervention required**

ESS is integrated with Server Load Balancer (SLB) and ApsaraDB for Relational Database Service (RDS). It automatically manages SLB backend servers and RDS whitelists, eliminating the need to perform manual O&M.

- **Various scaling modes for you to mix and match**

ESS allows you to schedule, customize, fix the minimum number of instances, and configure automatic replacement of unhealthy instances. It also provides APIs to allow you to monitor instances through external monitoring systems.

7.3 Architecture

Table 7-1: Architecture description

Component	Description
OpenAPI Gateway	Provides basic services such as authentication and parameter passthrough.
Coordinator	Serves as the ingress of the ESS architecture. It provides external management and control for services, processes API calls, and triggers tasks.
Trigger	Obtains information from the health checks of instances and scaling groups, scheduled tasks, and CloudMonitor to perform tasks scheduling.
Worker	Functions as the core part of ESS. After receiving a task, it handles the entire life cycle of the task, including splitting, executing, and returning the execution results.
Database	Includes the business database and workload database.
Middleware layer	ZooKeeper: ensures consistency by implementing distributed locks for Server Controller.
	Tair: provides caching services for Server Controller.

Component	Description
	Message Queue (MQ): provides message queuing services of VM statuses.
	Diamond: manages persistent configurations.

7.4 Features

ESS has the following features:

- **Automatically adding or removing ECS instances based on your business demands**

You can use the following scaling modes to adjust the number of ECS instances:

- Scheduled mode: Configure periodic tasks to add or remove ECS instances at a specified point in time, such as 13:00 every day.
- Custom mode: Call APIs to manually adjust the number of ECS instances based on monitoring system statistics.
 - You can manually implement scaling rules.
 - You can manually add or remove existing ECS instances.
 - After you have manually adjust MinSize (the minimum number of instances) and MaxSize (the maximum number of instances), ESS automatically creates or releases ECS instances to ensure that the number of instances remains within the MinSize and MaxSize range.
- Fixed-number mode: Maintain a fixed number of healthy ECS instances by specifying the MinSize attribute. This mode can be used to ensure day-to-day business availability.
- Health mode: Automatically remove or release ECS instances when they are detected as unhealthy (such as they are not in the running state).
- Multimode: Combine any of the preceding modes to meet your own business requirements . For example, if you predict that business peak hours are between 13:00 to 14:00, you can configure a scaling mode that creates 20 ECS instances at the scheduled time. If you are not sure whether the actual demand during peak hours will exceed the number of scheduled resources (for example, the actual load requires 40 ECS instances), another scaling mode can be configured to handle unexpected business loads.

- **Automatically adding or removing ECS instances to or from the SLB backend server group**

The health status of an ECS instance in a scaling group is determined based on the life cycle of the instances. If an ECS instance is in an unhealthy state, ESS automatically removes the instance and creates a new one. ESS then adds the new instance to the SLB instance and RDS whitelist.

**Note:**

ECS instances used for automatic scaling can be removed. Therefore, these instances cannot be used to store application status information (such as sessions) and related data (such as databases and logs). If applications deployed on these ECS instances require data to be saved, you can save the status information to independent ECS instances, databases to RDS, and logs to Log Service.

- **Automatically adding or removing IP addresses of ECS instances to or from the RDS whitelist**

When an ECS instance is automatically added to or removed from an SLB backend server group, the IP address of the ECS instance is also automatically added or removed from the RDS whitelist. This mechanism automatically maintains the RDS whitelist and effectively controls access to the RDS instance.

7.5 Usage scenarios

ESS can be used in the following scenarios:

- Video streaming: Traffic loads surge during holidays and festivals. Cloud computing resources must be automatically scaled out to meet the increased demands.
- Live streaming and broadcast: Traffic loads are ever-changing and difficult to predict. Cloud computing resources must be scaled based on CPU utilization, application load, and bandwidth usage.
- Gaming: Traffic loads increase during lunch and after work hours. Cloud computing resources must be scaled out in advance.


7.6 Limits

ESS has the following limits:

- Applications on ECS instances deployed in a scaling group must be stateless and horizontally scalable.

- Instances created by ESS cannot be automatically added to the instance access whitelist of ApsaraDB for Memcache. You must manually add the instances to the whitelist. For more information, see *ApsaraDB for Memcache Product Introduction*.
- ESS does not support vertical scaling. It can only scale the number of ECS instances. The CPU, memory, and bandwidth configurations of the ECS instances cannot be automatically adjusted.
- Scaling configurations, scaling rules, and scaling activities are dependent on the life cycle of a scaling group. If a scaling group is deleted, all scaling group configurations, rules, and activities associated with this group are also deleted.
- Scheduled tasks are independent from scaling groups. Deleting a scaling group does not affect the scheduled tasks.
- Each user can create a limited number of scaling groups, scaling configurations, scaling rules, ECS instances for scaling, and scheduled tasks. For more information, see [Table 7-2: Quantity restrictions](#).

Table 7-2: Quantity restrictions

Item	Description
Scaling group	You can create up to 20 scaling groups.
Scaling configuration	You can create up to 10 scaling configurations in a scaling group.
Scaling rule	You can create up to 10 scaling rules in a scaling group.
ECS instance for scaling	<p>You can configure up to 100 ECS instances for automatic scaling in a scaling group.</p> <div>  Note: The limit applies to the ECS instances that are automatically created, but does not apply to manually added ones. </div>
Scheduled task	You can create up to 20 scheduled tasks.

7.7 Terms

Auto Scaling

Auto Scaling (ESS) is a management service that automatically adjusts the number of elastic computing resources based on your business demands and strategies. It automatically creates

ECS instances during high business loads, and automatically releases ECS instances during low business loads.

Scaling group

A scaling group is a group of ECS instances that is dynamically scaled based on the configured scenario. You can specify the maximum and minimum number of ECS instances in a scaling group, as well as SLB and RDS instances associated with the group.

Scaling configuration

A scaling configuration specifies the configurations of ECS instances in ESS.

Scaling rule

A scaling rule defines the specific scaling activity, for example, the number of ECS instances to be added or removed.

Scaling activity

After a scaling rule is triggered, a scaling activity is performed. A scaling activity shows the changes to the ECS instances in a scaling group.

Scaling trigger task

A scaling trigger task is a task that triggers a scaling rule, such as scheduled tasks.

Cooldown period

The cooldown period indicates a period of time after the completion of a scaling activity in a scaling group. During this period, no other scaling activities can be executed.

8 Object Storage Service (OSS)

8.1 What is OSS

Alibaba Cloud Object Storage Service (OSS) is a massive, secure, low-cost, and highly reliable cloud storage service provided by Alibaba Cloud.

It can be considered as an out-of-the-box storage solution with unlimited storage capacity.

Compared with the user-created server storage, OSS has many outstanding advantages in reliability, security, cost, and data processing capabilities. Using OSS, you can store and retrieve a variety of unstructured data files, such as text files, images, audios, and videos, over the network at any time.

OSS uploads data files as objects to buckets. OSS is an object storage service that uses a key-value pair format. You can retrieve object content based on unique object names (keys).

On OSS, you can:

- Create a bucket and upload objects to the bucket.
- Obtain an object URL from OSS to share or download an object.
- Complete the ACL settings of a bucket or object by modifying its properties or metadata.
- Perform basic and advanced OSS tasks through the OSS console.
- Perform basic and advanced OSS tasks using the Alibaba Cloud SDKs or directly calling the RESTful APIs in your application.

8.2 Advantages

Advantages of OSS over user-created server storage

Item	OSS	User-created server storage
Reliability	<ul style="list-style-type: none"> • The capacity is automatically expanded without affecting external services. • Offers automatic redundant data backup. 	<ul style="list-style-type: none"> • Prone to errors due to low hardware reliability. If a disk has a bad sector, data may be irretrievably lost. • Manual data restoration is complex and requires a lot of time and technical resources.
Security	<ul style="list-style-type: none"> • Provides hierarchical security protection for enterprises. 	<ul style="list-style-type: none"> • Additional scrubbing and black hole equipment is required.

Item	OSS	User-created server storage
	<ul style="list-style-type: none"> User resource isolation mechanisms and local disaster recovery Provides various authentication and authorization mechanisms, as well as whitelisting, hotlinking protection, and RAM. It also provides Security Token Service (STS) for temporary access. 	<ul style="list-style-type: none"> A separate security mechanism is required.
Data processing	Image processing capabilities	Image processing capabilities must be purchased and deployed separately.

More benefits of OSS

- Ease of use

Provides standard RESTful APIs (some compatible with Amazon S3 APIs), a wide range of SDKs and client tools, and a management console. You can easily upload, download, retrieve, and manage large amounts of data for websites and applications, similar to regular files systems.

- There is no limit on the number and size of objects. Therefore, you can easily expand your buckets in OSS as required.
- Supports streaming writing and reading, which is suitable for business scenarios where you need to simultaneously read and write videos and other large objects.
- Supports lifecycle management. You can delete expired data in batches.

- Powerful and flexible security mechanisms

Flexible authentication and authorization mechanisms are available. OSS provides STS and URL authentication and authorization, as well as whitelisting, hotlinking protection, and RAM.

- Rich image processing functions

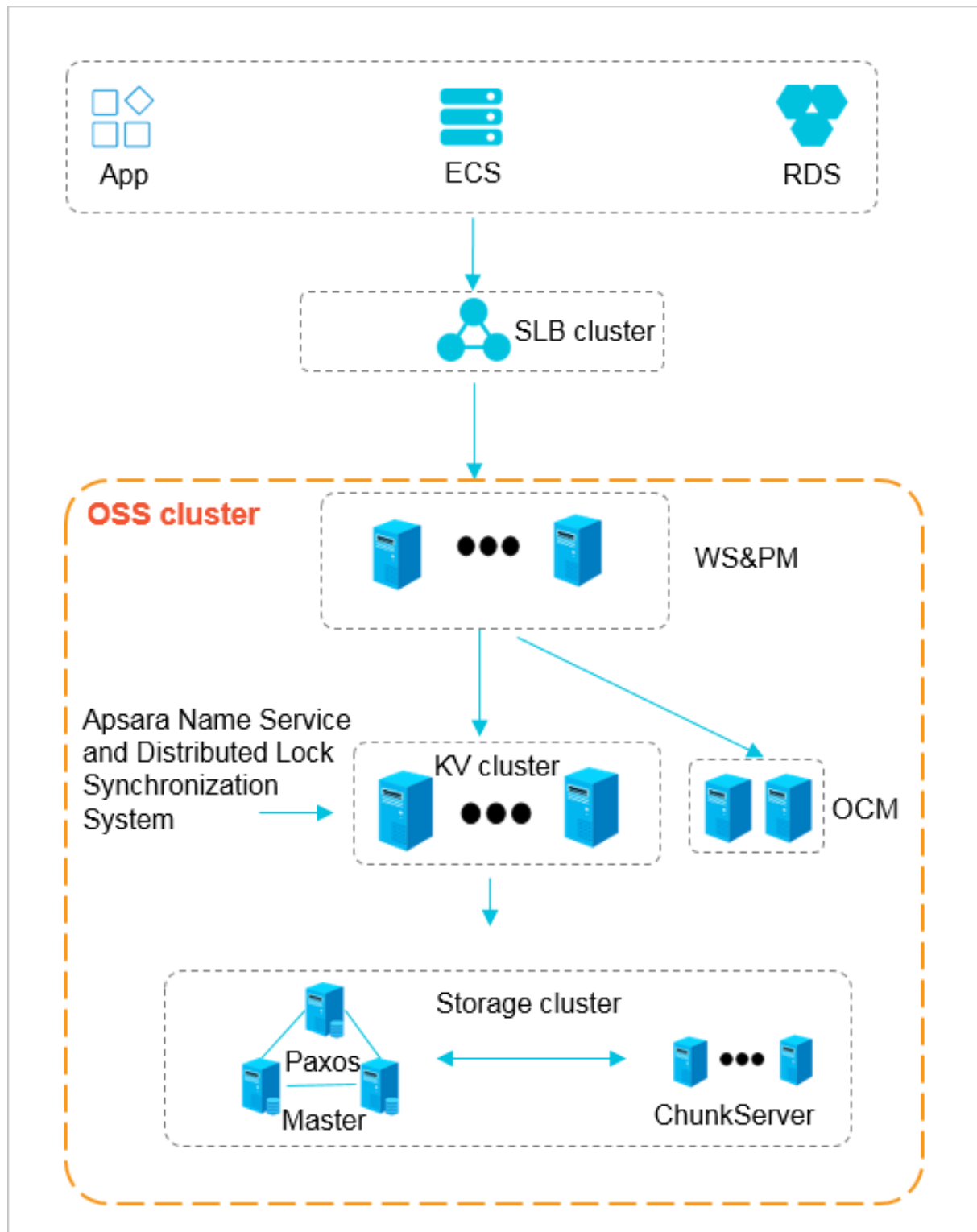
Supports format conversion, thumbnails, cropping, watermarking, resizing for object formats such as JPG, PNG, BMP, GIF, WEBP, and TIFF.

8.3 Architecture

Object Storage Service (OSS) is a storage solution built on the Alibaba Cloud Apsara platform. It is based on infrastructure such as Apsara Distributed File System and SchedulerX. Such

infrastructure provides OSS and other Alibaba Cloud services with distributed scheduling, high-speed networks, and distributed storage features. The following figure shows the OSS architecture

Figure 8-1: OSS architecture



- **WS & PM (the protocol layer):** is used for receiving users' requests sent through the REST protocol and performing authentication. If the authentication succeeds, users' requests are forwarded to the key-value engine for further processing. If the authentication fails, an error message is returned.
- **KV cluster:** is used for processing structured data, including reading and writing data based on keys. The KV cluster also supports large-scale concurrent requests. When a service has to operate on a different physical server due to a change in the service coordination cluster, the KV cluster can quickly coordinate and find the access point.
- **Storage cluster:** Metadata is stored on the master node. A distributed message consistency protocol (Paxos) is adopted between master nodes to ensure the consistency of metadata. This ensures efficient distributed storage and access of objects.

8.4 Functions

OSS offers the following functions:

Table 8-1: OSS functions

Category	Function	Description
Bucket	Create a bucket	Before uploading an object to OSS, you need to create a bucket to store the objects.
	Delete a bucket	If you are no longer using a bucket, delete it to avoid incurring further fees.
	Modify ACL settings for a bucket	OSS provides an ACL for permission control. You can configure an ACL when creating a bucket and modify it after creating the bucket.
	Configure static website hosting	You can configure static website hosting for your bucket and access this static website through the bucket endpoint.
	Configure hotlinking protection	To prevent fees incurred by hotlinked OSS data, OSS supports hotlinking protection based on the referer field in the HTTP header.
	Manage CORS	OSS provides Cross-Origin Resource Sharing (CORS) settings in the HTML5 protocol to help you achieve cross-region access.
	Configure lifecycle	You can define and manage the lifecycle of all objects in a bucket or a subset of objects. Lifecycle

Category	Function	Description
		is generally set for batch object management and automatic part deletion.
Object	Upload an object	You can upload all types of objects to a bucket.
	Create a folder	You can manage OSS folders in the same way you manage Windows folders.
	Search for objects	You can search for objects with the same prefix in a bucket or folder.
	Obtain an object URL	You can obtain an object URL from OSS to share or download an object.
	Delete an object	You can delete a single object or several objects in batches.
	Delete a folder	You can delete a folder or delete folders in batches.
	Modify ACL settings for an object	You can configure ACL settings when uploading an object and modify them after uploading the object.
	Manage parts	You can delete all or some parts from a bucket.
Image processing	Image processing	You can perform operations such as format conversion , cropping, scaling, rotating, watermarking, style encapsulation on images stored in OSS.
OSS access control for VPC	Single tunnel	You can establish single tunnels to access OSS resources from VPC.
API	API	Provides RESTful API operations supported by OSS and relevant examples.
SDK	SDK	Provides SDK development operations and relevant examples in commonly used languages.

8.5 Scenarios

Massive storage for image, audio, and video applications

OSS can be used to store large amounts of data, such as images, audios, videos, and logs. OSS supports various devices. Websites and mobile applications can directly read or write OSS data. OSS supports file writing and streaming writing.

Dynamic and static content separation for websites and mobile applications

OSS leverages the BGP bandwidth to achieve ultra-low latency of direct data download.

Offline data storage

OSS is cheap and highly available, enabling enterprises to store data that needs to be archived offline for a long time to OSS.

8.6 Limits

Item	Description
Bucket	<ul style="list-style-type: none"> You can create a maximum of 10 buckets. Once a bucket is created, its name and region cannot be modified.
Upload objects	<ul style="list-style-type: none"> Objects uploaded through the management console, simple upload, form upload, and append upload cannot exceed 5 GB. To upload an object greater than 5 GB, you must use multipart upload. The size of each object uploaded through multipart upload cannot be greater than 48.8 TB. You can upload objects with the same name, but the existing objects are overwritten.
Delete objects	<ul style="list-style-type: none"> Deleted objects cannot be restored. You can delete up to 50 objects in batches through the management console. To delete more objects in batches, you must use APIs or SDKs.
Lifecycle	You can configure up to 1,000 lifecycle rules for each bucket.
Image processing	<ul style="list-style-type: none"> For a source image: <ul style="list-style-type: none"> Only JPG, PNG, BMP, GIF, WEBP, and TIFF files are supported. The file size cannot exceed 20 MB. For image rotation, the width or height of the image cannot exceed 4096 pixels. For a thumbnail: <ul style="list-style-type: none"> The product dimensions cannot exceed 4096 x 4096 pixels. The length of each side cannot exceed 4096 pixels.

8.7 Concepts

This topic describes several basic concepts of OSS.

Object

Objects, also known as OSS files, are the basic entities stored in OSS. An object is composed of metadata, data, and key. An object is identified by a unique key in the bucket. Metadata defines the properties of an object, such as the last modification time and the object size. You can also specify custom metadata for an object.

The lifecycle of an object starts when it is uploaded, and ends when it is deleted. During the lifecycle of an object, the metadata cannot be changed. Unlike the file system, OSS does not allow you to modify objects directly. If you want to modify an object, you must upload a new object with the same name as the existing one to replace it.

**Note:**

Unless otherwise stated, objects and files mentioned in OSS documents are collectively called objects.

Bucket

A bucket is a container for objects. All objects must be stored in a bucket. You can set and modify the properties of a bucket for object access control and lifecycle management. These properties apply to all objects in the bucket. Therefore, you can create different buckets to perform different management functions.

- OSS does not have the hierarchical structure of directories and subfolders as in a file system. All objects are directly related to their corresponding buckets.
- You can have multiple buckets.
- A bucket name must be globally unique within OSS and cannot be changed once a bucket is created.
- A bucket can contain an unlimited number of objects.

Strong consistency

Object operations in OSS are atomic, which means operations are either successful or failed. There is no intermediate state. OSS will never write corrupted or partial data.

Object operations in OSS are strongly consistent. For example, once you receive a successful upload (PUT) response, the object can be read immediately, and the data has already been written in triplicate. Therefore, OSS avoids the situation where no data is obtained when you perform the read-after-write operation. When you delete an object, the object also has no intermediate state. Once you delete an object, that object no longer exists.

This feature allows OSS to be operated similar to traditional storage devices. Modifications are immediately visible, and consistency is guaranteed.

Comparison between OSS and the file system

OSS is a distributed object storage service that uses a key-value pair format. You can retrieve object content based on unique object names (keys). Although you can use names like test1/test.jpg, this does not necessarily indicate that the object is saved in a directory named test1. In OSS, test1/test.jpg is only a string, which is no different from a.jpg. Therefore, similar amounts of resources are consumed when you access objects that have different names.

A file system uses a typical tree index structure. Before accessing a file named test1/test.jpg, you must access directory test1 and then locate the file named test.jpg. This makes it easy for a file system to support folder operations, such as renaming, deleting, and moving directories, as these operations are only directory node operations. System performance depends on the capacity of a single device. The more files and directories that are created in the file system, the more resources are consumed, and the lengthier your process becomes.

You can simulate similar functions in OSS, but this operation is costly. For example, if you want to rename the test1 directory test2, the actual OSS operation would be to replace all objects whose names start with test1/ with copies whose names start with test2/. Such an operation would consume a large amount of resources. Therefore, when using OSS, try to avoid such operations.

You cannot modify objects stored in OSS. A specific API must be called to append an object, and the generated object is of a different type from that of normally uploaded objects. Even if you only want to modify a single byte, you must re-upload the entire object. A file system allows you to modify files. You can modify the content at a specified offset location or truncate the end of a file. These features make file systems suitable for more general scenarios. However, OSS supports massive concurrent access, whereas the performance of a file system is subject to the performance of a single device.

Therefore, mapping OSS objects to file systems is very inefficient, which is not recommended. If attaching OSS as a file system is required, we recommended that you perform only the operations of writing new files, deleting files, and reading files. We recommend that you take full advantage of OSS features, such as its massive data processing capabilities to store large amounts of unstructured data, such as images, videos, and documents.

9 Table Store

9.1 What is Table Store

Table Store is a NoSQL database service independently developed by Alibaba Cloud. Table Store is a copyrighted software program that is certified by the relevant authority in China. Table Store is built on Alibaba Cloud's Apsara system, and can store and access large volumes of structured data in real time.

Table Store provides the following features:

- Supports a minimum of 10 PB of data in each cluster, and a minimum of 1 PB of data or 1 trillion records in each table. Table Store offers schema-free data structure storage. You do not need to define attribute columns before you use them. You do not require table-level changes to add or reduce attribute columns. You can enable Time To Live (TTL) on a table to delete expired data from the table.
- Adopts the triplicate technology to keep three copies of data on three servers placed on three different racks. Each cluster supports either pure SSD instances or mixed storage instances to meet different budget and performance requirements.
- Adopts a fully redundant architecture that prevents single point of failure (SPOF). With support for hot cluster upgrades and automatic data migration, you can dynamically add or delete nodes without service interruptions. The concurrent read and write throughput and storage capacity can be linearly scaled. Each cluster can have no less than 500 nodes.
- Supports highly concurrent read/write operations. Concurrent read/write operations scale with the number of hosts. Read/write performance is not directly related to the amount of data in a single table.
- Supports identity authentication and multi-tenancy; provides comprehensive permission authentication and isolation mechanisms to safeguard your data; supports VPC networks and access through HTTPS; supports RAM and account authorization; provides multiple authentication and authorization mechanisms so that you can define access permissions for individual tables and APIs.

9.2 Benefits

Scalability

There is no upper limit to the amount of data that can be stored in Table Store tables. As data increases, Table Store adjusts partitions to provide more storage space for tables and improve the capability of handling access request bursts.

High performance

High-performance instances provide the single-digit millisecond average access latency for single rows. The read/write performance is not affected by the size of data in a table.

Reliability

Table Store stores multiple data copies. If one backup fails, servers with copied data will immediately restore services, achieving high data reliability.

High availability

Through automatic failure detection and data migration, Table Store shields applications from host - and network-related hardware faults to achieve high availability.

Ease of management

Table Store automatically performs complex O&M tasks, such as the management of data partitions, software and hardware upgrades, configuration updates, and cluster scale-out.

Access security

Table Store provides multiple permission management mechanisms. It performs identity authentication and authorization for each application request to prevent unauthorized data access and ensure the security of data access.

High consistency

Table Store ensures high data consistency for data writes. After a write operation succeeds, three replicas are written to a disk. Applications can read the latest data immediately.

Flexible data models

Table Store tables do not require a fixed format. Each row can contain a different number of columns. Table Store supports multiple data types, such as integer, boolean, double, string, and binary.

Pay-As-You-Go

Table Store has low minimum charges. You are only charged based on the actual Table Store resources you reserve and use.

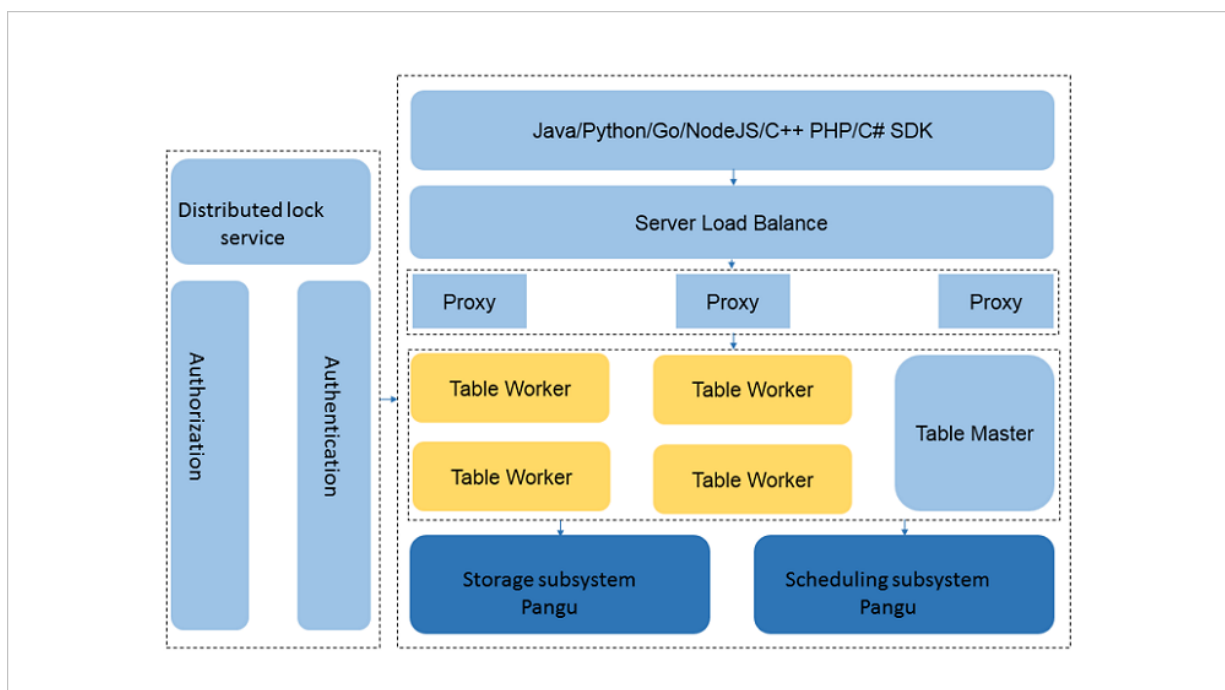
Monitoring integration

You can log on to the Table Store console to obtain monitoring information in real time, including the requests per second and average response latency.

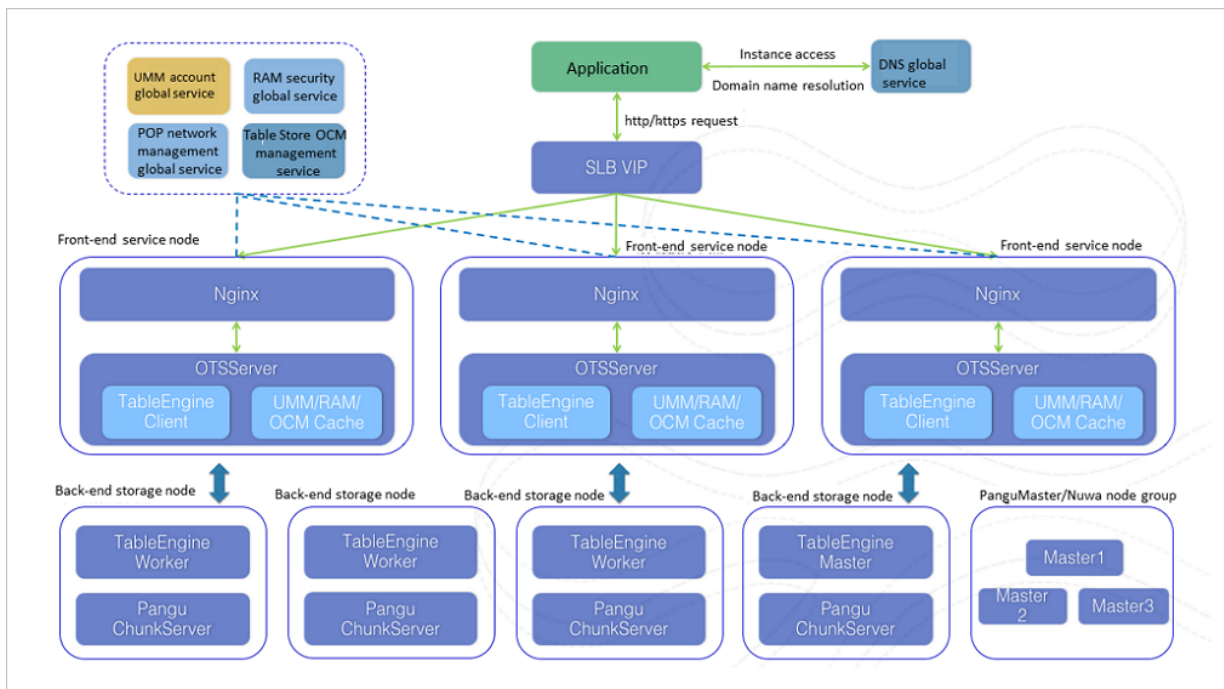
9.3 Architecture

The architecture of Table Store references the BigTable (one of the three core technologies of Google) and uses the log-structured merge-tree (LSM) storage engine to provide high performance writes. The performance of primary key-based single-row queries and range queries is stable and predictable. The performance is not affected by the volume of data and access concurrency.

The following figure shows the basic architecture of Table Store.



The following figure shows the detailed architecture of Table Store.



- The top layer is the protocol access layer. SLB distributes user requests to various proxy nodes. The proxy nodes receive requests that are sent through the RESTful protocol and implement security authentication. If the authentication succeeds, the user requests are forwarded to the corresponding data engine based on the value of the first primary key for further operations. If the authentication fails, the error information is directly returned to the user.
- Table Worker is the data engine layer that processes structured data through a primary key to search for or store data. Table Worker supports large-scale access request bursts.
- The bottom layer is the persistent storage layer. At this layer, the large-scale Apsara Distributed File System is deployed. Metadata is stored in Masters. The distributed message consistency protocol Paxos is adopted between Masters to ensure metadata consistency. In this scenario, efficient distributed file storage and access are achieved. This method guarantees three replicas of data stored in the system and system recovery from any hardware or software fault.

9.4 Features

Data partition and load balancing

The first primary key column in each row of a table is called the partition key. The system partitions a table into multiple partitions based on the range of the partition key. When the data in a partition exceeds a certain size, the partition is automatically split into two smaller partitions. The data and access loads are scattered to two partitions. The partitions are scheduled to different nodes. Eventually, the linear scalability of the single-table data scale and access loads

is achieved. A partitions is a logical organization of data based on the shared storage mechanism . No migration of physical data is involved when a partition is split. The theoretical impact of load balancing on the partition is that the partition fails to provide services within 100 milliseconds.

Automatic recovery upon a single node failure

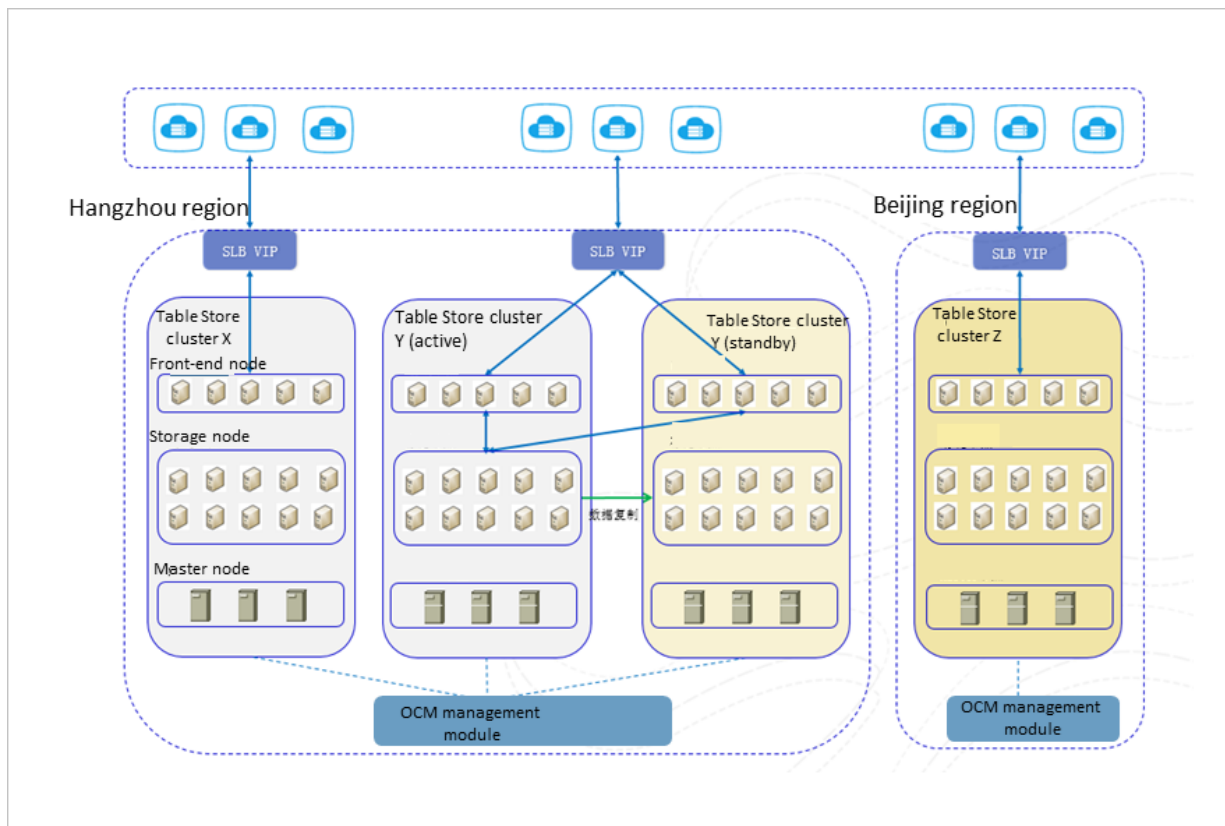
In the storage engine of Table Store, each node serves a number of data partitions in different tables. A master node monitors partition distribution and dispatching, and the health of each service node. If a service node fails, the master node migrates data partitions from this faulty node to other healthy nodes. The migration is logically performed, and does not involve physical entities, so services can recover from the single point of failure (SPOF) within several minutes.

Zone-disaster recovery and geo-disaster recovery

To meet business security and availability requirements, Table Store provides active-standby cluster-based zone-disaster recovery and geo-disaster recovery. Disaster recovery supports instance-based recovery. Any table operation on the primary instance, including insertion, update , or deletion, is synchronized to the table of the same name on the secondary instance. The duration of data synchronization between the primary and secondary instances depends on the network environment of the active and standby clusters. In the ideal network environment, the synchronization latency reaches the millisecond level. Before the manual failover, you must stop resource access to the active cluster and wait for all data to be completely backed up. After the failover, do not perform any failover operation within one hour. Clear original cluster data and reset the standby cluster.

In the active-standby cluster-based zone-disaster recovery scenario, the endpoints remain unchanged when applications access Table Store in the active and standby clusters. In other words, the application endpoints do not need to be changed after the failover. In the active-standby cluster-based geo-disaster recovery scenario, the endpoints of the active and standby clusters are different. After the failover, endpoints need to be changed for applications.

The following figure shows the multi-cluster-based disaster recovery architecture (including active and standby clusters). Table Store cluster Y (active) and Table Store cluster Y (standby) are used to achieve active-standby-mode based disaster recovery.



In the single-cluster based architecture, all data of Table Store has three replicas. Data is returned to users after the three copies are written to the disk. This method ensures that single clusters have high reliability.

In the zone-disaster recovery and geo-disaster recovery mechanisms, data is sent from service nodes of Table Store to the frontend service nodes of the standby cluster asynchronously based on the data write time. After the frontend service nodes receive the replication request, the frontend service nodes write data to the standby cluster based on the normal write process.

In this disaster recovery mode, data synchronization and failover are two most important terms. When the active and standby clusters are first created, data synchronization includes full data synchronization and incremental data synchronization. After the failover, the standby cluster begins to provide services. After the active cluster recovers, data is synchronized from the standby cluster to the active cluster. Eventually, services are failed over to the active cluster.

Note that the data synchronization uses the async mode. A piece of data is returned to the client after successfully written to the active cluster, but before it has been synchronized to the standby cluster. Asynchronous replication results in data inconsistency. The business-side must specify the impact of the data inconsistency and take corresponding measures.

Specifically, if some data is not synchronized to the standby cluster during the failover when a fault occurs,

- The original active cluster retains all original data, and the data that has not synchronized to the standby cluster before the failover.
- Aside from the data that is not synchronized before the failover, the new active cluster (or the standby cluster) has all written data.

Table Store provides the Recovery Point Objective (RPO) and Recovery Time Objective (RTO) in normal network environments. The maximum value of RPO is one minute. The maximum value of RTO is five minutes.

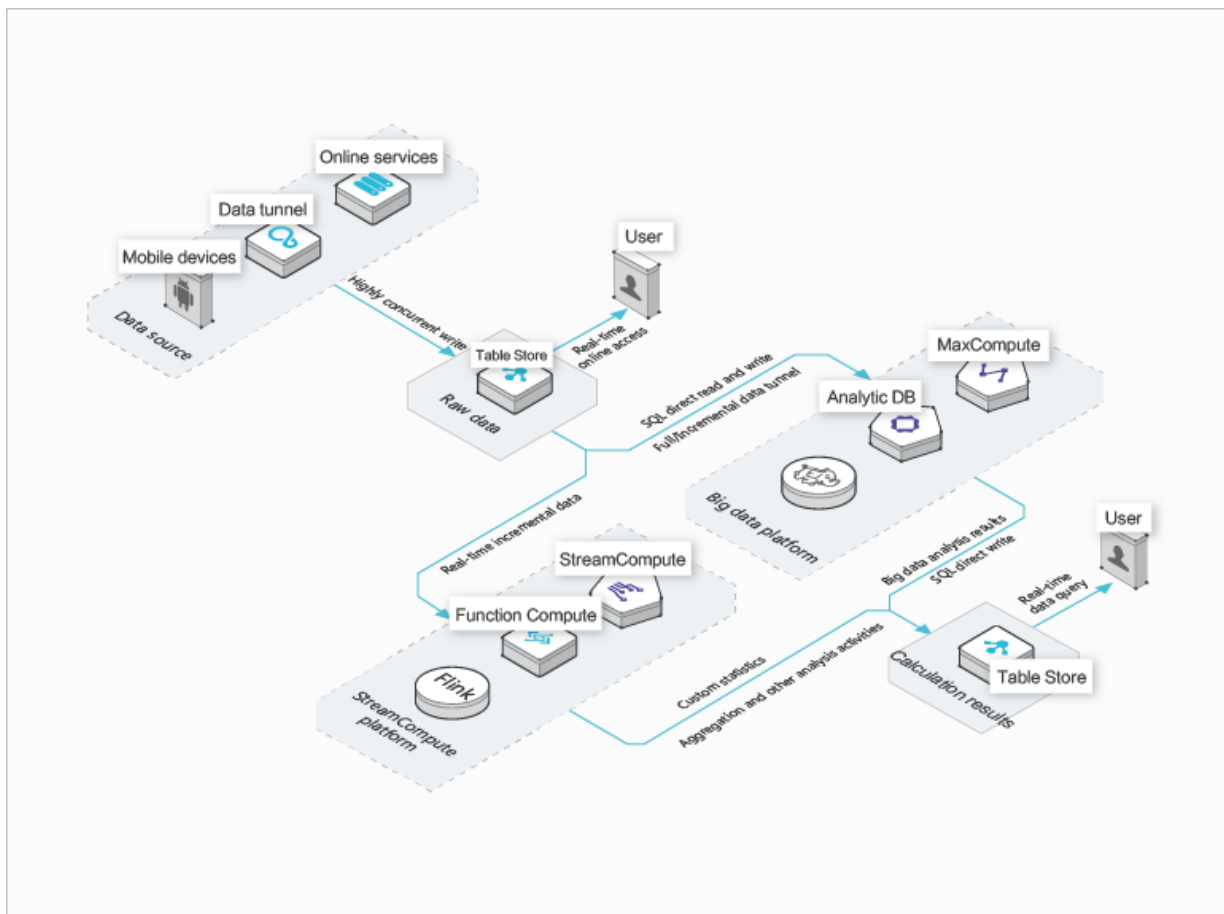
9.5 Scenarios

Scenario 1: Big data storage and analysis

Table Store provides cost-friendly, highly concurrent low-latency storage, and online access to a large amount of data. It provides full and incremental data tunnels and supports direct SQL read and write for various big data analytics platforms such as MaxCompute. An efficient incremental streaming read interface is provided for easy computing of real-time data streams.

Specific features are as follows:

- A table can process 10 PB of data or 1 trillion data records.
- Various big data computing platforms, stream computing, and real-time computing services are provided.
- Pay-As-You-Go for elastic resources provides instances with high performance and high capacity to meet the requirements of different services.

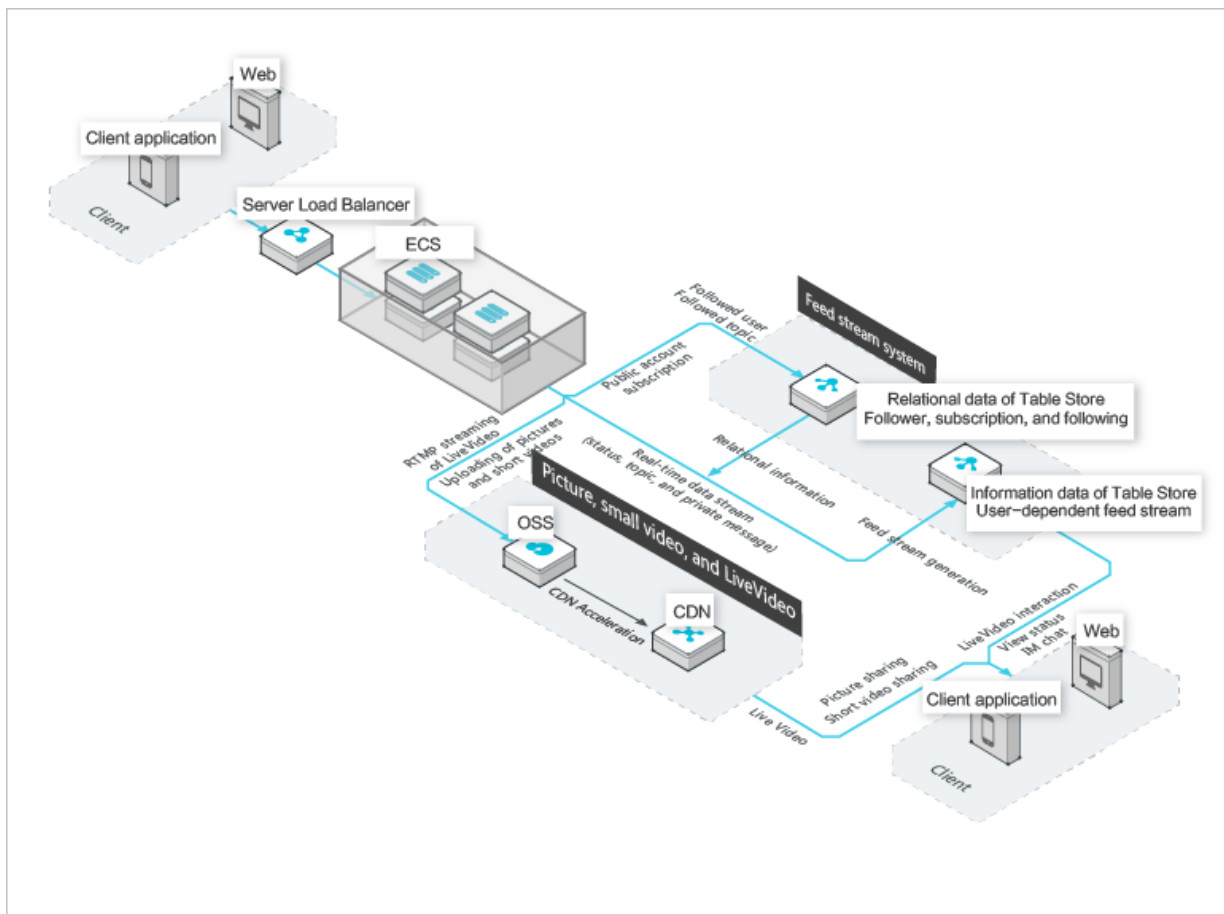


Scenario 2: Social media feeds on the Internet

You can use Table Store to store instant messaging (IM) chats, and social media feed information such as comments, follow-up posts, and likes. The elastic resources stored in Table Store are billed based on the Pay-As-You-Go model. At relatively low costs, Table Store can meet the application requirements such as significant traffic fluctuation, high concurrency, and low latency.

Specific features are as follows:

- Built-in auto-increment primary key columns simplify external system dependencies.
- Average read and write performance of high-performance instances is not affected by the data volume.
- High-availability storage of large amounts of messages and multi-terminal message synchronization are supported.

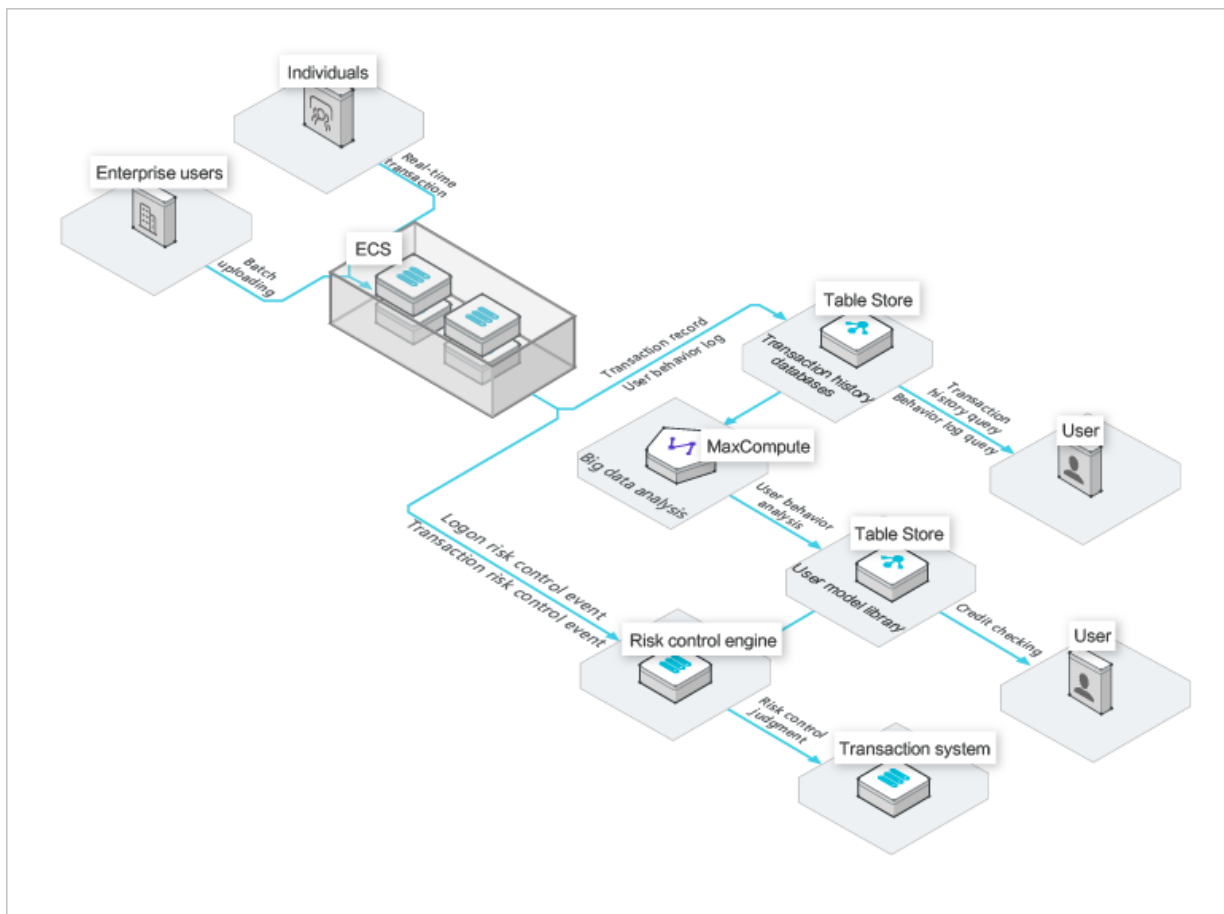


Scenario 3: Storage and real-time queries of large amounts of transaction records and user models

The Pay-As-You-Go billing method for elastic resources, low latency, and high concurrency allow your risk control system to always operate in optimal conditions. You can strictly control transaction risks. Furthermore, the flexible data structure allows your business model to rapidly evolve to meet market demands.

Specific features are as follows:

- A table contains 1 trillion records and easily stores full history transaction records.
- Three replicas are used to ensure high consistency and data security.
- Rapid service development is supported by a schema-free model and attribute columns that can be added as required.

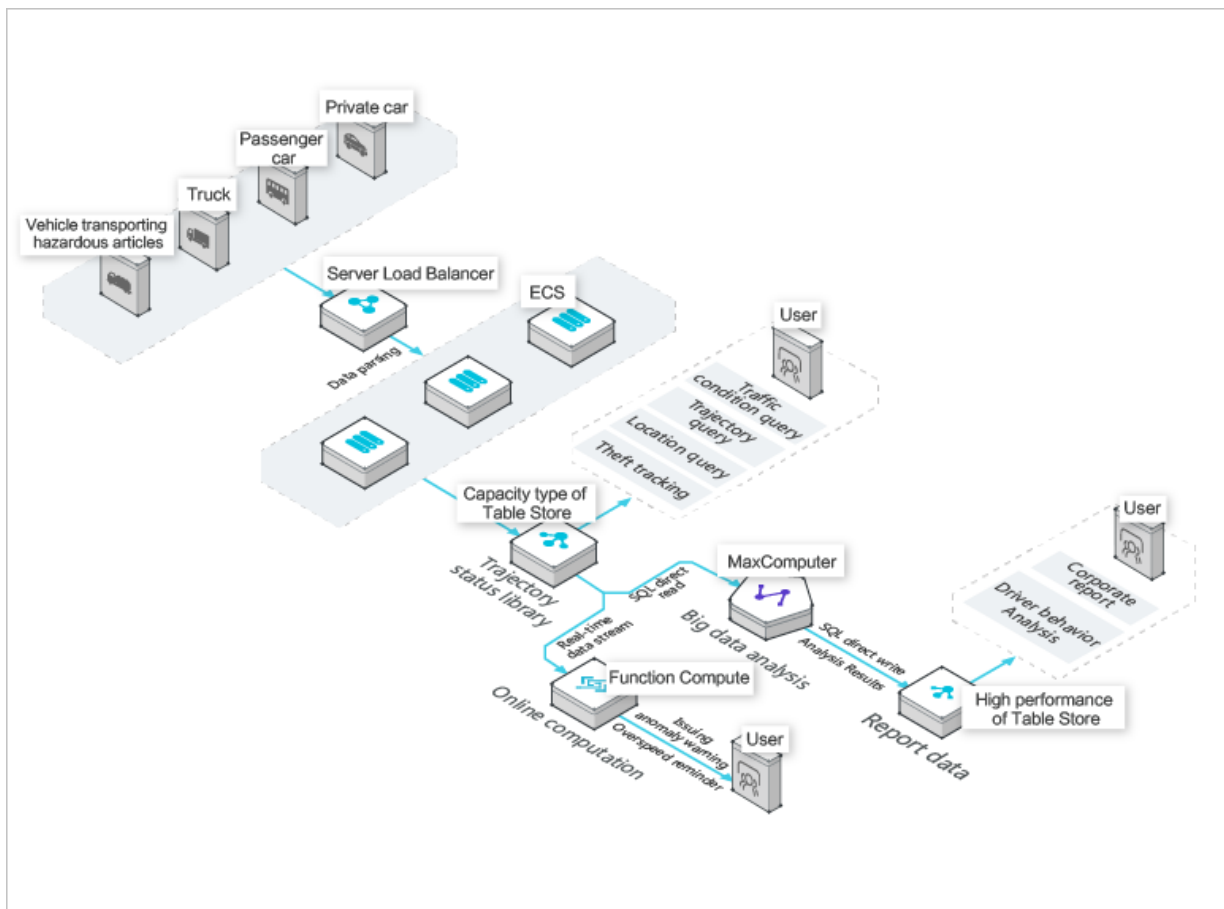


Scenario 4: Efficient and flexible IoV data storage

A table can store PBs of data without needing to partition data in databases and tables. This capacity simplifies the service logic. The schema-free data model enables easy access to the data collected from different vehicle-mounted devices. Table Store can be seamlessly integrated with multiple big data analysis platforms and real-time computing services for ease of real-time online queries and business report analysis.

Specific features are as follows:

- A table stores 10 PB of data without needing to partition data in databases and tables. This capacity simplifies the service logic.
- The query performance for vehicle conditions and routes is stable and predictable.
- The schema-free data model enables easy access to the data collected from different vehicle-mounted devices.

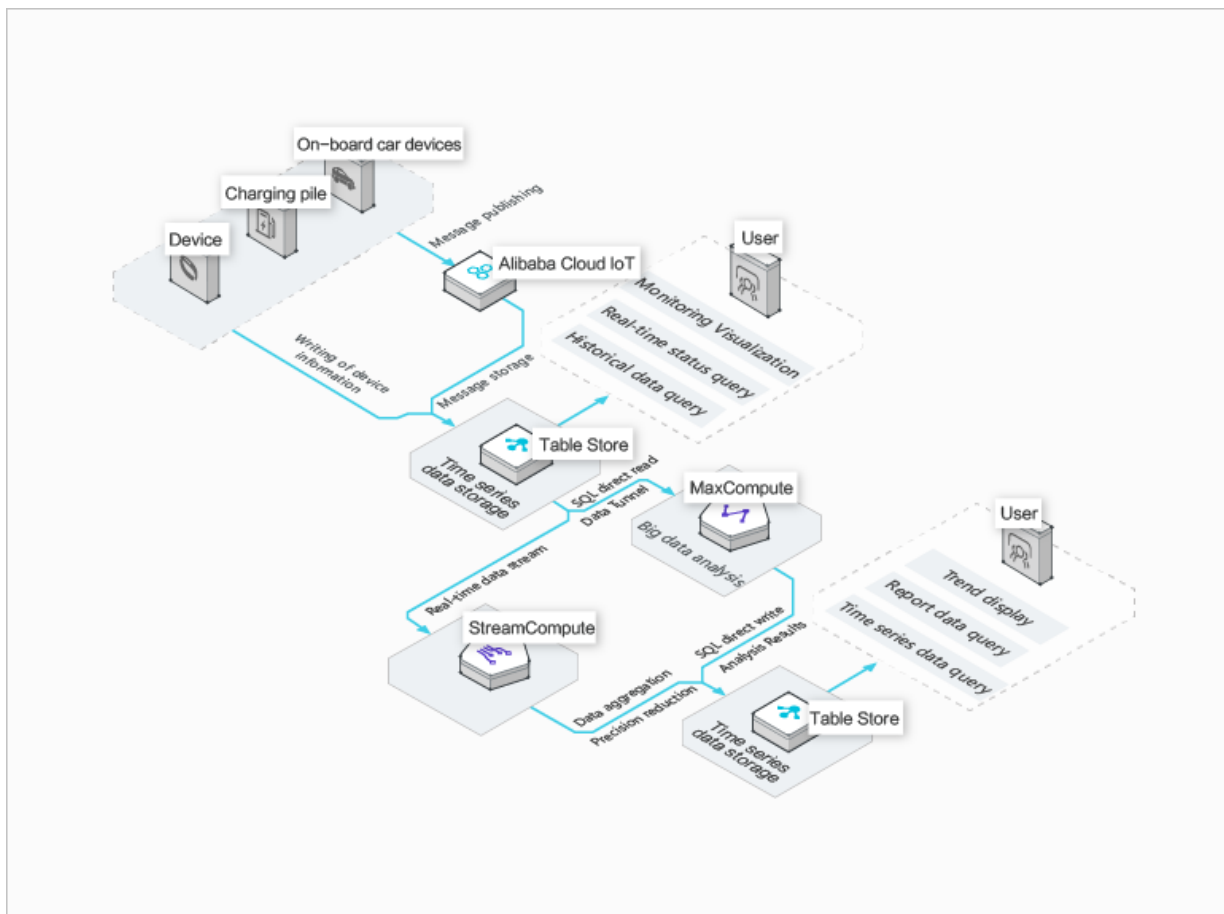


Scenario 5: Storage of a large amount of IoT data and efficient queries and analysis

A table stores PBs of data, allowing Table Store to easily store the time series data from IoT devices and monitoring systems. The direct SQL read for big data analysis and the efficient incremental streaming read API allow easy offline data analysis and real-time stream computing.

Specific features are as follows:

- A table stores 10 PB of data. This capacity enables Table Store to meet the data write and storage requirements of ultra large-scale IoT devices and monitoring systems.
- A single piece of data can be used in analysis and computing scenarios for different services through the interconnection with multiple offline and stream data analytics platforms.
- Pay-As-You-Go, low costs, and TTL management are supported.

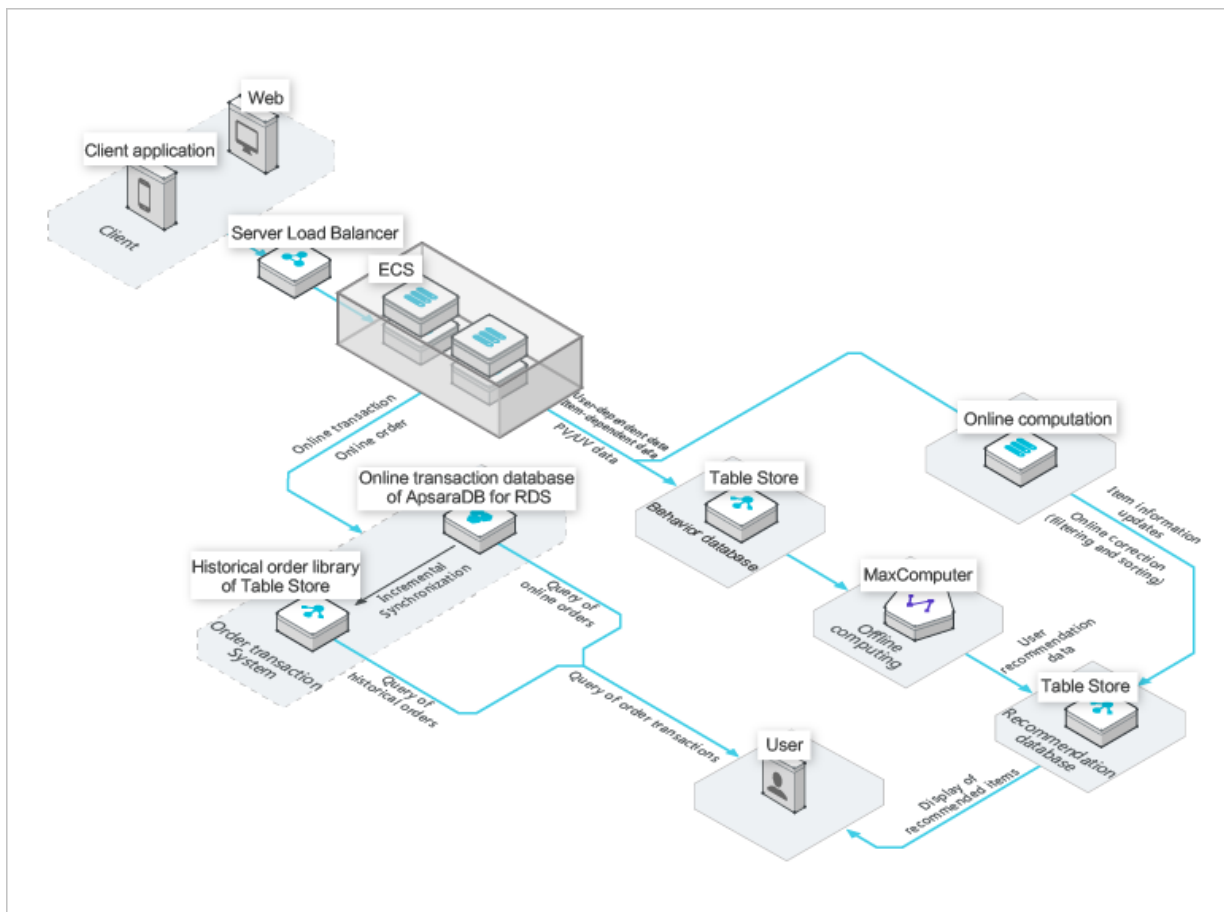


Scenario 6: Large-scale e-commerce transaction orders and user database recommendations

Table Store can easily manage large volumes of historical transaction data and improve access performance. Combined with MaxCompute, Table Store enables precision marketing, elastic resource storage, and Pay-As-You-Go billing. You can easily handle services during peak hours when all users go online.

Specific features are as follows:

- Auto scaling of data volumes and access concurrency meets the requirements of scenarios with access fluctuations during various periods.
- Various big data analytics platforms are supported for direct analytics of user behavior.
- Single-digit millisecond query latency of large amounts of transaction data.



9.6 Limits

[Table 9-1: Limits](#) shows the limits for Table Store. Some limit ranges indicate the maximum value that can be used rather than the suggested value. Table structure and row data size can be tailored to enhance performance.

Table 9-1: Limits

Limit	Limit range	Description
Number of instances created under an Alibaba Cloud account	1,024	If you need to increase the maximum number of instances, contact an administrator.
Number of tables in an instance	1,024	If you need to increase the maximum number of tables, contact an administrator.
Number of primary key columns	1–4	A primary key can contain one to four columns.
Size of string type primary key column values	1 KB	The size of a string type primary key column value cannot exceed 1 KB.

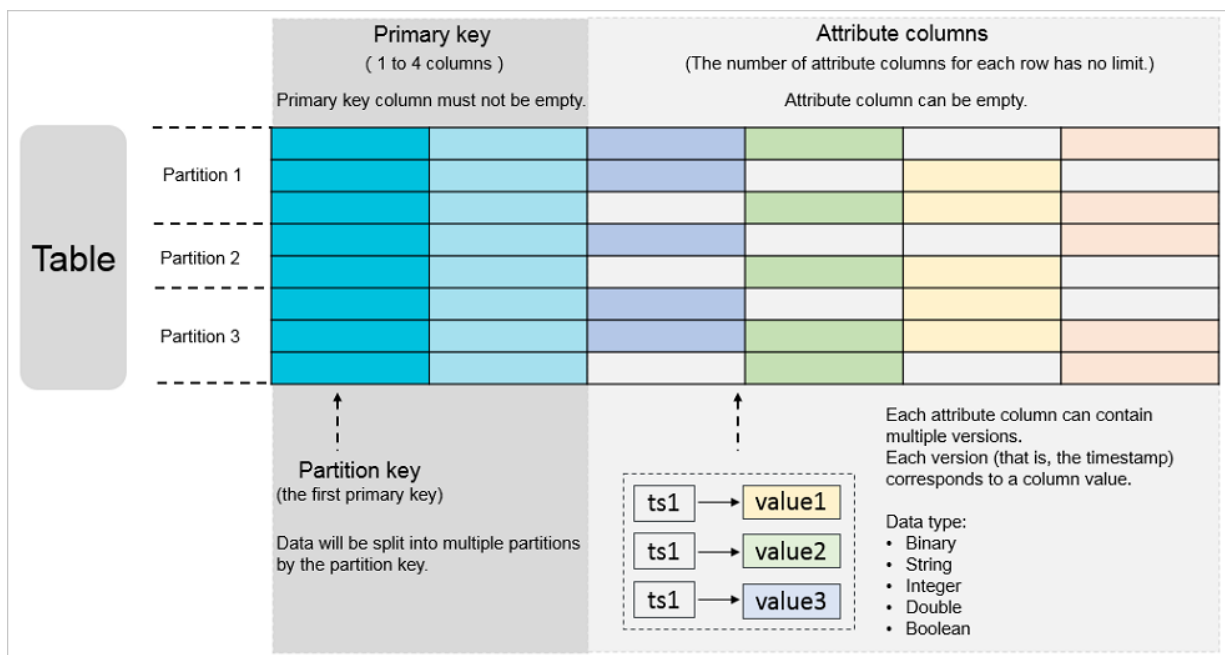
Limit	Limit range	Description
Size of string type attribute column values	2 MB	The size of a string type attribute column value cannot exceed 2 MB.
Size of binary type primary key column values	1 KB	The size of a binary type primary key column value cannot exceed 1 KB.
Size of binary type attribute column values	2 MB	The size of a binary type attribute column value cannot exceed 2 MB.
Number of attribute columns in a single row	Unlimited	A single row can contain an unlimited amount of attribute columns.
The number of attribute columns written by one request	1,024 columns	The number of attribute columns written by a PutRow, UpdateRow, or BatchWriteRow request in a row cannot exceed 1,024.
Data size of a single row	Unlimited	The total size of all column names and column value data for a row is unlimited.
Number of columns involved in columns\to\get in the read request	0–128	The maximum number of columns obtained in a row of data in the read request is up to 128.
Number of UpdateTable operations for a table	<ul style="list-style-type: none"> Upper limit : Unlimited Lower limit : Unlimited 	You need to follow the adjustment frequency limit for a table.
UpdateTable frequency for a table	Maximum of one update every 2 minutes	The reserved read/write throughput for a table cannot be adjusted beyond the frequency of once every two minutes.
The number of rows read by one BatchGetRow request	100	N/A
The number of rows written by one BatchWriteRow request	200	N/A
Data size of one BatchWriteRow request	4 MB	N/A
Data returned by one GetRange operation	5,000 rows or 4 MB	The data returned by a request cannot exceed 5,000 rows or 4 MB. When any of the preceding conditions is satisfied, data that exceeds the limits is truncated at the row-level. The data primary key information in the next row is returned.

Limit	Limit range	Description
The data size of an HTTP request body	5 MB	N/A

9.7 Terms

data model

A model that involves tables, rows, primary keys, and attributes, as shown in the following figure.



max versions

A data table attribute that indicates the maximum number of data versions in each attribute column of a data table. If the number of versions in an attribute column exceeds the value of Max Versions, the earliest version is deleted asynchronously.

Time To Live (TTL)

A data table attribute measured in seconds. It indicates the validity period of data. To save data storage space and reduce storage costs, the Table Store backend automatically clears any data that exceeds TTL.

max version offset

A table attribute measured in seconds. To prevent the writing of unexpected data, the server checks the attribute column versions when processing writing requests. Writing data to a specified row fails if the row has an attribute column in which:

- Its version is earlier than the current writing time minus the value of Max Version Offset.
- Its version is later than or equal to the current writing time plus the value of Max Version Offset.

The valid version range for attribute columns is calculated based on the formula: Valid version range = [Data written time - Value of max version offset, Data written time + Value of max version offset). The data written time is the sum of the seconds counted from 1970-01-01 00:00:00 UTC to the time data is written. Versions of the attribute columns (in milliseconds) must, after being divided by 1,000 and converted to seconds, fall into the valid version range.

primary key and attribute

The unique identifier of each row in a table. It consists of one to four primary key columns. When you create a table, you must define a primary key. Specifically, you must specify the name, data type, and sequence of each primary key column. The data type of primary key columns can only be string, integer, or binary. For a primary key column of the string or binary type, the size of the column must be smaller than 1 KB.

An attribute stores data in a row. The number of attribute columns for each row is unlimited.

read/write throughput

A Table Store attribute that is measured by read/write capacity units (CUs). The CU is the smallest billing unit for the data read and write operations.

region

An Apsara Stack physical data center. Table Store is deployed across many Apsara Stack regions . You can select a region that suits your business requirements.

instance

A logical entity in Table Store. It is used to manage tables that correspond to a database in a relational database management system (RDBMS). An instance is the basic unit of the Table Store resource management system. Table Store allows you to control access and measure resources at the instance level.

endpoint

A connection URL (also known as an endpoint) for each instance. You need to specify the endpoint before you perform any operations on Table Store tables and data.

stream

A data table attribute used for real-time analysis of incremental data streams and incremental data synchronization.

Serial ATA (SATA)

A disk that is based on serial connections and provides stronger fault-tolerance capabilities. It aims to improve the reliability of data transmission.

10 Network Attached Storage (NAS)

10.1 What is NAS?

Network Attached Storage (NAS) is a file storage service that can be mounted to compute nodes such as ECS instances. NAS allows you to use standard file access protocols to access distributed file systems without making any changes to your existing applications. NAS features unlimited capacity and performance expansion, single namespace, data sharing, high reliability, and high availability.

After you create a NAS instance and a mount point, you can use the standard NFS protocol to mount the instance to multiple compute nodes (such as ECS instances), and use the standard POSIX interface to access the instance. You can mount a NAS instance to multiple compute nodes for file and directory sharing.

10.2 Benefits

Data sharing

A NAS instance can be mounted to multiple compute nodes for data sharing. This significantly reduces data replication and synchronization costs.

High reliability

NAS provides high data reliability. Compared with traditional user-created storage, NAS greatly reduces maintenance costs and data security risks.

Elastic scalability

A NAS instance has a maximum storage capacity of 10 PB. It can scale to adapt to growing business data.

High performance

The throughput of a NAS instance scales linearly with the storage capacity. The cost is substantially lower than a high-end NAS device.

Ease of use

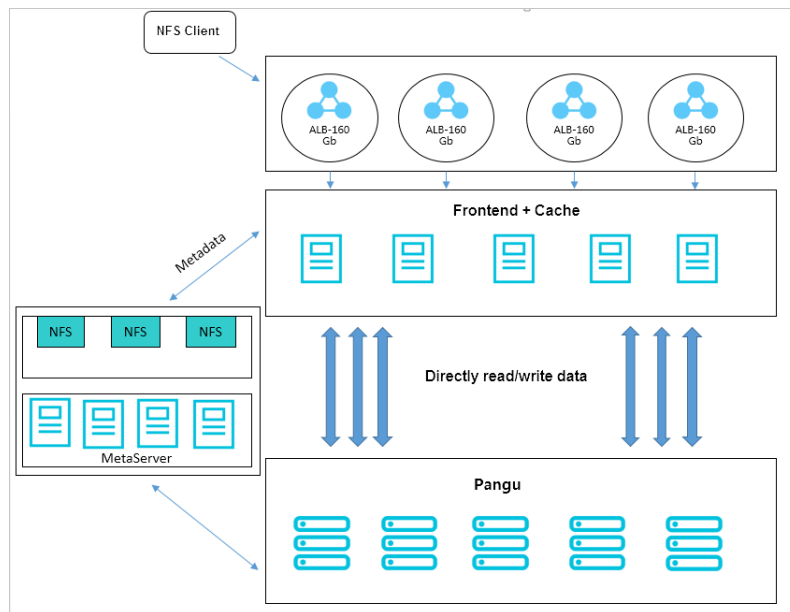
NAS supports the NFSv3 and NFSv4 protocols. Compute nodes such as ECS instances can use the standard POSIX interface to access NAS instances.

10.3 Architecture

NAS is based on Apsara Distributed File System. It maintains three copies of each data across multiple storage nodes. The frontend nodes receive connection requests from NFS clients and provide the cache function. These nodes are stateless and distributed to ensure high availability of the frontend. The metadata of NAS instances is stored in MetaServer. When the frontend nodes receive I/O requests, they access the MetaServer to obtain the metadata. Then, the frontend nodes access the backend data nodes for user data.

Both the frontend and backend can expand elastically, ensuring high availability, high throughput, and low latency.

Figure 10-1: System architecture



10.4 Features

Seamless integration

NAS supports the NFSv3 and NFSv4 protocols, and uses standard file system semantics to access data. No changes are required for mainstream applications and workloads to work with NAS.

Shared access

A NAS instance can be accessed by multiple compute nodes simultaneously. This makes it suitable for scenarios where applications deployed across multiple ECS instances access the same data source.

Security control

NAS provides multiple security mechanisms to guarantee system data security. These mechanisms include network isolation for VPCs, user isolation for classic networks, standard permission control for file systems, permission groups, and RAM accounts.

Linear scalability

NAS provides high throughput, high IOPS, and low latency for application workloads. The performance of NAS scales linearly with its capacity, making it suitable to meet growing business demands.

10.5 Scenarios

Scenario 1: shared storage and high availability for SLB

Your SLB instance is connected to multiple ECS instances. You can store the data of the applications on these ECS instances on a shared NAS instance. This implements data sharing and ensures high availability of the SLB servers.

Scenario 2: file sharing within an enterprise

The employees of an enterprise need to access the same datasets for work purposes. The administrator can create a NAS instance and configure different file or directory permissions for users or user groups.

Scenario 3: data backup

You want to back up the data stored in the data center to the cloud and use a standard interface to access the cloud storage service. You can back up the data in the data center to a NAS instance.

Scenario 4: server log sharing

You want to store the application server logs of multiple compute nodes on the shared file storage . You can store these server logs on a NAS instance for centralized log processing and analysis.

10.6 Limits

- NAS supports the NFSv3 and NFSv4 protocols.
- The following table lists the attributes that are not supported by NFSv4.0 and NFSv4.1, and their client errors.

Protocol	Unsupported attribute	Client error
NFSv4.0	FATTR4_MIMETYPE, FATTR4_QUOTA_AVAIL_H ARD, FATTR4_QUO TA_AVAIL_SOFT, FATTR4_QUOTA_USED, FATTR4_TIME_BACKUP, and FATTR4_TIME_CREATE	NFS4ERR_ATTRNOTSUPP
NFSv4.1	FATTR4_DIR_NOTIF_DEL AY, FATTR4_DIR ENT_NOTIF_DELAY , FATTR4_DACL , FATTR4_SACL, FATTR4_CHANGE_POLICY , FATTR4_FS_STATUS, FATTR4_LAYOUT_HINT, FATTR4_LAYOUT_TYPES, FATTR4_LAYOUT_ALIGNM ENT, FATTR4_FS_ LOCATIONS_INFO, FATTR4_MDSTHRESHOLD, FATTR4_RETENTION_GET, FATTR4_RETENTION_SET, FATTR4_RETEN_EVT_GET , FATTR4_RETEN_EVT_SET , FATTR4_RETENTION_HOL D, FATTR4_MOD E_SET_MASKED, and FATTR4_FS_CHARSET_CA P	NFS4ERR_ATTRNOTSUPP

- NFSv4 does not support the following OPs: OP_DELEGPURGE, OP_DELEGRETURN, and NFS4_OP_OPENATTR. The client displays an NFS4ERR_NOTSUPP error.
- NFSv4 does not support Delegation.
- About UID and GID:
 - For NFSv3, if the file UID or GID exists in a Linux local account, the corresponding username and group name is displayed based on the mapping between the local UID and GID. If the file UID or GID does not exist in the local account, the UID and GID is displayed.

- For NFSv4, if the version of the local Linux kernel is earlier than 3.0, the UIDs and GIDs of all files are displayed as "nobody." If the kernel version is later than 3.0, the display rule is the same as that of NFSv3.

**Note:**

If you use NFSv4 to mount a NAS instance and the Linux kernel version is earlier than 3.0, we recommend that you do not change the owner or group of local files or directories. Such changes can cause the UIDs and GIDs of the files or directories to become "nobody."

- You can mount a NAS instance to up to 10,000 compute nodes.

10.7 Terms

mount point

A mount point is the access address of a NAS instance in a VPC or classic network. Each mount point corresponds to a domain name. To mount a NAS instance to a local directory, you must specify the domain name of the mount point.

permission group

The permission group is a whitelist mechanism provided by NAS. You can add rules to a permission group of a NAS instance to allow users from specified IP addresses or address segments to access the NAS instance with different permissions.

**Note:**

Each mount point must be bound with a permission group.

authorized object

An authorized object is an attribute of the permission group rule. It specifies the IP address or address segment to which the permission group rule is applied. In a VPC, an authorized object can be a single IP address or an address segment. In a classic network, an authorized object must be an IP address, generally the intranet IP address of an ECS instance.

11 Distributed File System (DFS)

11.1 What is DFS

Apsara Stack Distributed File System (DFS) is a file storage service for computing resources such as Apsara Stack ECS and Container Service instances. It supports standard Hadoop FileSystem interfaces. You can use DFS without the need to modify the existing applications of big data analytics. DFS features unlimited capacity, performance scale-out, single namespace, multi-tenancy, high reliability, and high availability.

After you create a DFS instance, computing resources such as ECS and Container Service instances can access the DFS instance through standard Hadoop FileSystem interfaces. Multiple computing nodes can access the same DFS and share files and directories.

11.2 Benefits

High reliability

Data is stored in three copies to improve reliability. Compared with user-created HDFS, DFS minimizes O&M costs and data security risks.

Elastic scalability

The capacity of a single DFS is unlimited and can be scaled up or down to meet business requirements at any time.

High performance

Throughput performance is specially optimized for small files. Compared with user-created HDFS, DFS significantly improves the throughput performance of small files.

Multi-tenancy

Multiple DFS instances can be created in the storage system, with unified permission and space management.

Ease of use

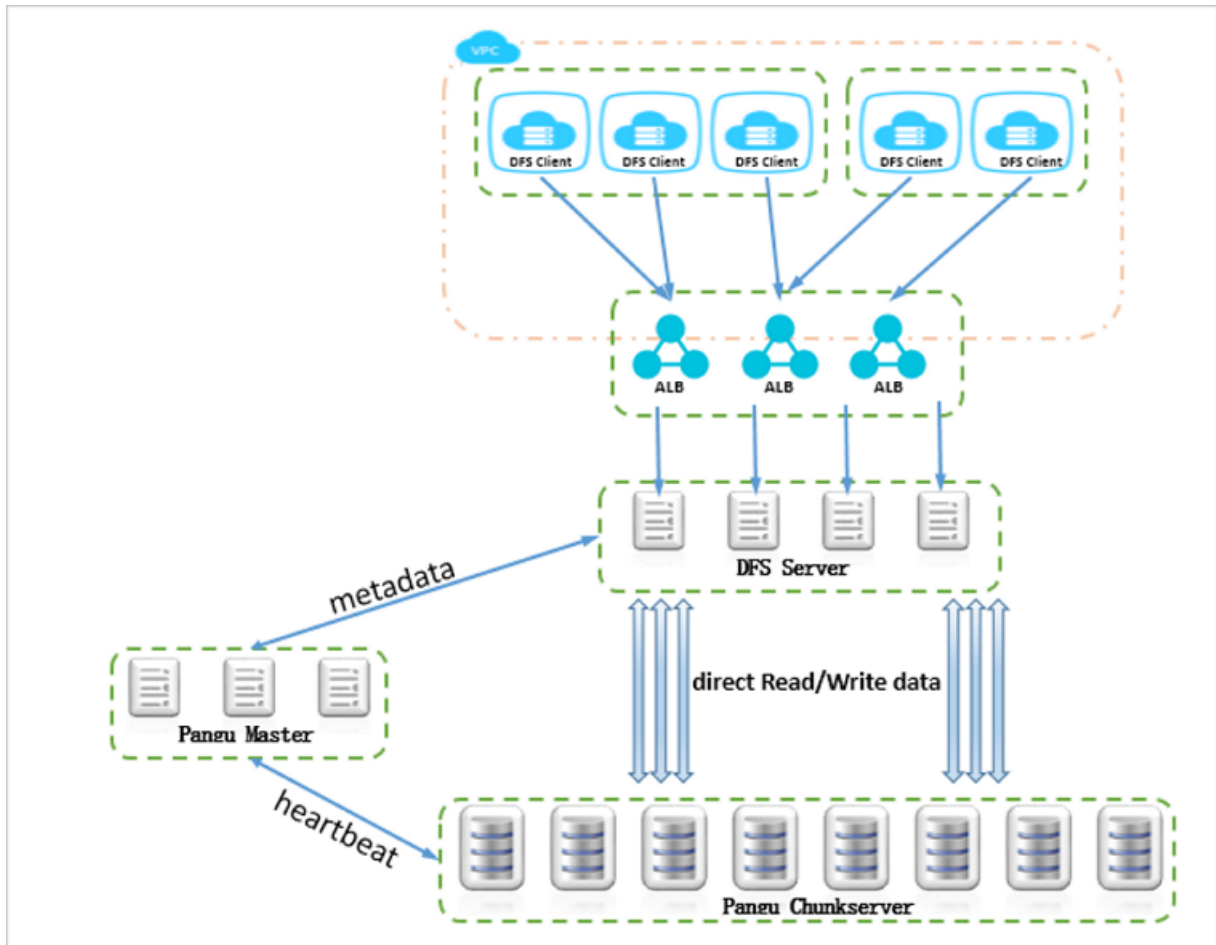
Automatic O&M is supported, which reduces O&M workloads and human errors, and secures O&M.

11.3 Architecture

The architecture of DFS consists of the frontend and backend.

The backend is based on Apsara Distributed File System. Data is replicated into multiple copies and stored in Apsara Distributed File System. Frontend access nodes of DFS receive and cache connection requests that are sent from computing resources such as ECS instances, Hadoop computing applications (such as MapReduce and Spark), or Container Service instances. Apsara Distributed File System also manages metadata and data of DFS.

The following figure shows the overall DFS architecture.



11.4 Features

Seamless integration

DFS supports the Hadoop 2.7.X protocol and accesses data through standard Hadoop FileSystem interfaces. Mainstream Hadoop applications and workloads can be integrated seamlessly without modifications.

Shared access

A single DFS instance can be accessed by multiple computing nodes simultaneously. It applies to scenarios where applications deployed across multiple ECS or Container Service instances access the same data source.

Security control

Multiple security mechanisms are implemented to secure system data. These security mechanisms include network isolation (such as the use of VPCs), standard permission control of file systems, permission groups, and RAM.

Linear scalability

DFS stores application workloads with high throughput, high IOPS, and short latency. Additionally, the linear relationship between the performance and capacity meets the rising requirements for capacity and storage performance when the business volume increases.

11.5 Scenarios

Scenario 1: Shared storage and high availability

DFS meets the following business requirements:

- Shared access to files
- High availability of files

Application method: DFS supports standard Hadoop FileSystem interfaces that allow you to store files in your DFS instance in real time.

Scenario 2: Big data analytics and machine learning

In big data analytic and machine learning scenarios, applications require high throughput performance and short latency for data access. DFS can provide access of high throughput and short latency. You do not need to migrate data to on-premises computing resources.

Application method: Data is stored in your DFS instance so that computing resources such as ECS instances can access the data directly. Hadoop and machine learning applications are deployed on multiple computing resources. In this way, computing resources can use Hadoop FileSystem interfaces to access data, perform offline or online computing, and export the computing result to DFS for permanent storage.

11.6 Limits

DFS does not support the following operations, features, or commands:

Hadoop FileSystem or AbstractFileSystem

- Configurations of the directory modification time (mtime) and access time (atime) through `setTimes`
- Symbolic links
- File truncation (`truncate`)
- File concatenation (`concat`)
- Extended attributes (XAttrs)-relevant operations
- Snapshot-relevant operations
- Delegation token-relevant operations
- Checksum-relevant operations (`setWriteChecksum` and `setVerifyChecksum`)
- Access control list (ACL)-relevant operations
- File block locations

Hadoop fs command line tools

- Snapshot-relevant commands (`createSnapshot`, `deleteSnapshot`, and `renameSnapshot`)
- ACL-relevant commands (`setfacl` and `getfacl`)
- XAttr-relevant commands (`setfattr` and `getfattr`)
- File truncation-relevant commands (`truncate`)

11.7 Terms

mount point

The access URL of a DFS instance in a VPC or classic network. Each mount point is mapped to a domain name. You need to modify the `core-site.xml` configuration to access files in a DFS instance.

Permission group

A whitelist in DFS. You can add rules to provide IP addresses or network segments with different permissions to access DFS.

**Note:**

Each mount point must belong to a specified permission group.

Authorized object

An attribute of a permission group rule. It specifies the IP address or network segment to which the permission group rule is applied. In a VPC, an authorized object can be a single IP address or a CIDR block. In a classic network, an authorized object can only be a single IP address (the private IP address of an ECS instance).

12 ApsaraDB for RDS

12.1 What is ApsaraDB for RDS?

ApsaraDB for Relational Database Service (RDS) is a stable, reliable, and automatically scalable online database service.

Based on the distributed file system and high-performance storage, ApsaraDB for RDS allows you to easily perform database operations and maintenance with its complete set of solutions for disaster recovery, backup, recovery, monitoring, and migration.

ApsaraDB for RDS supports three storage engines including MySQL, PostgreSQL, and PPAS. They help you conveniently and rapidly create database instances suitable for your scenarios.

ApsaraDB RDS for MySQL

Originally based on a branch of MySQL, ApsaraDB RDS for MySQL has proven its performance and throughput during the high-volume traffic from concurrent users during Double 11. ApsaraDB RDS for MySQL provides management for instances, accounts, and databases, whitelist configuration for instances, backup and recovery, transparent data encryption, and data migration. It also provides the following advanced functions:

- **Read-only instance:** In scenarios where there are a few write requests but a large number of read requests, you can enable read/write splitting to distribute read requests away from the primary instance. Read-only instances allow ApsaraDB RDS for MySQL 5.6 to auto-scale reading capabilities and increase the application throughput when large amounts of data is being read.
- **Read/write splitting:** The read/write splitting function provides an extra read/write splitting address. This address links the primary instance with all its read-only instances to enable an automatic link for read/write splitting. An application can use this method to read and write data by connecting to the same read/write splitting address. Write requests are automatically routed to the primary instance while read requests are routed to the read-only instances based on their weights. To scale up the reading capacity of the system, you can add more read-only instances.
- **Data compression:** ApsaraDB RDS for MySQL 5.6 allows you to compress data by using the TokuDB storage engine. Extensive tests show that the data volume is reduced by 80% to 90% after data tables are transferred from the InnoDB storage engine to the TokuDB storage engine. 2 TB data can be compressed to 400 GB or less by using TokuDB. In addition to data

compression, TokuDB supports transaction and online DDL operations. It is compatible with MyISAM and InnoDB applications.

ApsaraDB RDS for PostgreSQL

ApsaraDB RDS for PostgreSQL is an advanced open source database system with full SQL compliance and support for a diverse range of data formats such as JSON, IP, and geometric data. In addition to excellent support for features such as transactions, subqueries, multi-version concurrency control (MVCC), and data integrity checks, ApsaraDB RDS for PostgreSQL integrates a series of important functions including high availability, backup, and recovery that help ease your operations and maintenance burden.

ApsaraDB RDS for PostgreSQL provides basic functions such as whitelist configuration for instances, backup and recovery, data migration, and management for instances, accounts, and databases.

ApsaraDB RDS for PPAS

ApsaraDB RDS for Postgres Plus Advanced Server (PPAS) is a stable, secure, and scalable enterprise-class relational database. Based on ApsaraDB RDS for PostgreSQL, ApsaraDB RDS for PPAS features enhanced performance, application solutions, and compatibility. It is able to directly run Oracle applications. You can run enterprise-class applications on PPAS stably and obtain cost-effective services.

ApsaraDB RDS for PPAS provides basic functions such as whitelist configuration for instances, backup and recovery, data migration, and management for instances, accounts, and databases.

12.2 Benefits

12.2.1 Ease of use

ApsaraDB for RDS is a ready-to-use service featuring on-demand upgrades, convenient management, high transparency, and high compatibility.

Ready-to-use

ApsaraDB for RDS instances can be specified and created through APIs right away.

On-demand upgrade

When the database load and data storage capacity change, you can upgrade the RDS instance by changing its type. The upgrades do not interrupt the data link service.

Transparency and compatibility

ApsaraDB for RDS is used in the same way as the native database engine, allowing it to be adopted easily without the need to learn entirely new database engines. ApsaraDB for RDS is compatible with your existing programs and tools. Data can be easily migrated to ApsaraDB for RDS by using ordinary import and export tools.

Easy management

Alibaba Cloud is responsible for the routine maintenance and management tasks for ApsaraDB for RDS, such as hardware or software troubleshooting and database patch updates. You can also manually add, delete, restart, back up, and restore databases through Apsara Stack Management Console.

12.2.2 High performance

ApsaraDB for RDS provides parameter optimization, SQL optimization, and high-end back-end hardware to achieve high performance.

Parameter optimization

All RDS instance parameters have been optimized in years of production. The professional database administrators continually optimize RDS instances over the instance life cycles to ensure that ApsaraDB for RDS runs at optimal performance.

SQL optimization

Based on your scenario, ApsaraDB for RDS will lock inefficient SQL statements and provide recommendations to optimize code based on your business scenario.

High-end back-end hardware

All servers used by ApsaraDB for RDS have been evaluated by multiple parties to ensure stability.

12.2.3 High security

ApsaraDB for RDS provides high security by implementing anti-DDoS protection, access control, system security, and transparent data encryption (TDE).

DDoS attack prevention

**Note:**

You must activate Alibaba Cloud security services to use this feature.

When you access an ApsaraDB for RDS instance from the Internet, the instance is vulnerable to DDoS attacks. If a DDoS attack is detected, the RDS security system first performs traffic scrubbing. If traffic scrubbing is ineffective or the attack traffic reaches the black hole threshold, black hole filtering is triggered.

Access control

You can define IP addresses that are allowed to access ApsaraDB for RDS. The system will deny access from other IP addresses that have not been defined.

Each account can only view and operate their own respective database.

System security

ApsaraDB for RDS is protected by multiple layers of firewalls. These firewalls can effectively block a variety of attacks and secure data.

ApsaraDB for RDS servers cannot be logged onto directly. RDS services can only be accessed through specific ports.

The ApsaraDB for RDS servers cannot initiate an external connection. They can only receive access requests.

TDE encryption

Transparent Data Encryption (TDE) can be used to perform real-time I/O encryption and decryption on instance data files. Data is encrypted before being written to disks, and decrypted before being read from disks to the memory. TDE will not increase the size of data files.

Developers do not need to modify any applications before using the TDE feature.

12.2.4 High reliability

ApsaraDB for RDS provides hot standby, multi-copy redundancy, data backup, and data recovery to achieve high reliability.

Hot standby

ApsaraDB for RDS adopts a hot standby architecture. If the primary server fails, services are failed over to the secondary server within seconds. Applications running on the servers will be unaware of the failover process and continue to function normally.

Multi-copy redundancy

The data on the ApsaraDB for RDS server is built on RAID storage, and data backups are stored on OSS.

Data backup

ApsaraDB for RDS provides an automatic backup mechanism. You can set up a schedule for periodic backups or manually initiate temporary backups at any time to meet the needs of your business.

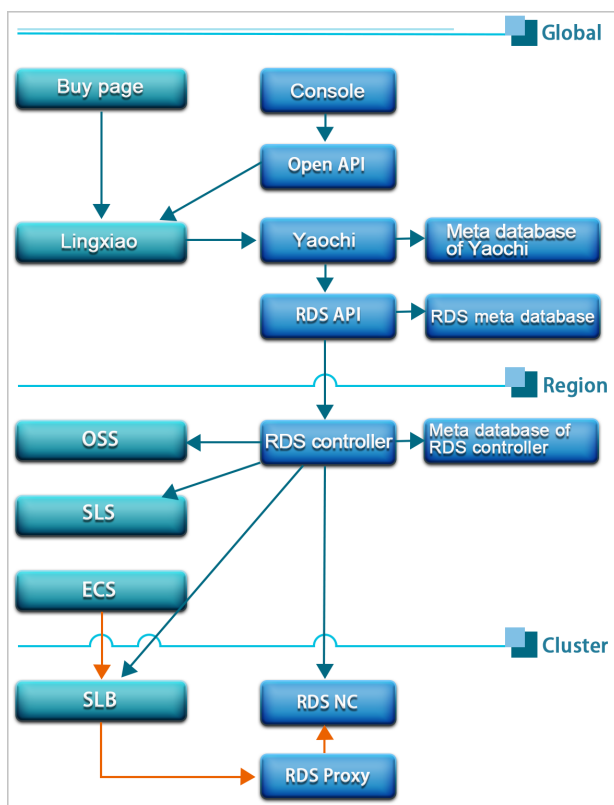
Data recovery

Data can be restored from backup sets or cloned instances created at previous points in time. After data is verified, the data can be migrated back to the primary RDS instance.

12.3 Architecture

The following figure shows the system architecture of ApsaraDB for RDS.

Figure 12-1: RDS system architecture

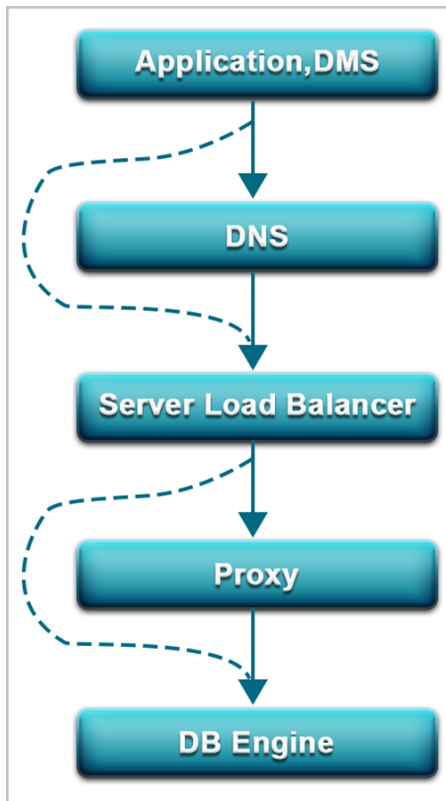


12.4 Features

12.4.1 Data link service

ApsaraDB for RDS provides all data link services, including DNS, Server Load Balancer (SLB), and Proxy.

ApsaraDB for RDS uses native DB engines with similar database operations across engines, minimizing the learning cost. Additionally, DMS greatly facilitates access to databases.



DNS

The DNS module can dynamically resolve domain names to IP addresses. Therefore, IP address changes do not affect the performance of RDS instances. After the domain name of an RDS instance is configured in the connection pool, the RDS instance can be accessed even if its corresponding IP address changes.

For example, the domain name of an ApsaraDB for RDS instance is `test.rds.aliyun.com`, and its corresponding IP address is `10.10.10.1`. The instance can be accessed when either `test.rds.aliyun.com` or `10.10.10.1` is configured in the connection pool of a program.

After a zone migration or version upgrade is performed for this ApsaraDB for RDS instance, the IP address may change to `10.10.10.2`. If the domain name `test.rds.aliyun.com` is configured in the connection pool, the instance can still be accessed. However, if the IP address `10.10.10.1` is configured in the connection pool, the instance will no longer be accessible.

SLB

The SLB module provides both the private IP address and public IP address of an ApsaraDB for RDS instance. Therefore, server changes do not affect the performance of the instance.

For example, the private IP address of an RDS instance is `10.1.1.1`, and the corresponding Proxy or DB Engine runs on `192.168.0.1`. The SLB module typically redirects all traffic destined for `10.1.1.1` to `192.168.0.1`. If `192.168.0.1` fails, another server in hot standby status with the IP address `192.168.0.2` takes over for the server with the IP address `192.168.0.1`. In this case, the SLB module will redirect all traffic destined for `10.1.1.1` to `192.168.0.2`, and the RDS instance will continue to provide services normally.

Proxy

The Proxy module provides a number of functions including data routing, traffic detection, and session persistence.

- Data routing: aggregates the distributed complex queries found in big data scenarios and provides the corresponding capacity management capabilities.
- Traffic detection: reduces SQL injection risks and supports SQL log backtracking when necessary.
- Session persistence: prevents database connection interruptions when faults occur.

DB Engine

The following table describes the mainstream database protocols supported by ApsaraDB for RDS.

Table 12-1: RDS database protocols

RDBMS	Version
MySQL	5.6 (including read-only instances)
PostgreSQL	9.4
PPAS	9.3/9.6

12.4.2 High-availability service

The high-availability (HA) service consists of modules such as the Detection, Repair, and Notice.

The HA service guarantees the availability of data link services and processes internal database exceptions.

Detection

The Detection module checks whether the primary and secondary nodes of the DB Engine are providing services normally. The HA node uses heartbeat information taken at 8 to 10 second intervals to determine the health status of the primary node. This information, along with the health status of the secondary node and heartbeat information from other HA nodes, provides a reference for the Detection module. All this information helps the module avoid misjudgment caused by exceptions such as network jitter. Failover can be completed within 30 seconds.

Repair

The Repair module maintains the replication relationship between the primary and secondary nodes of the DB Engine. It can also correct errors that occur on either node during normal operations.

For example:

- It can automatically restore primary/secondary replication after a disconnection.
- It can automatically repair table-level damage to the primary or secondary node.
- It can save and automatically repair the primary or secondary node in case of crashes.

Notice

The Notice module informs the SLB or Proxy module of status changes to the primary and secondary nodes to ensure that you always access the correct node.

For example, imagine that the Detection module discovers problems with the primary node and instructs the Repair module to resolve these problems. If the Repair module fails to resolve a problem, it instructs the Notice module to perform traffic switchover. The Notice module forwards the switching request to the SLB or Proxy module, and then all traffic is redirected to the secondary node. Meanwhile, the Repair module creates a new secondary node on a different physical server and synchronizes this change back to the Detection module. The Detection module rechecks the health status of the instance.

HA policy

Each HA policy defines a combination of service priorities and data replication modes defined to meet the needs of your business.

There are two service priorities:

- Recovery time objective (RTO): The database preferentially restores services to maximize the availability time. Use the RTO policy if you require longer database uptime.

- Recovery point objective (RPO): The database preferentially ensures data reliability to minimize data loss. Use the RPO policy if you require high data consistency

There are three data replication modes:

- Asynchronous replication (Async): When an application initiates an update request such as add, delete, or modify operations, the primary node responds to the application immediately after the primary node completes the operation. The primary node then replicates data to the secondary node asynchronously. This means that the operation of the primary database is not affected if the secondary node is unavailable. Data inconsistencies may occur if the primary node is unavailable.
- Forced synchronous replication (Sync): When an application initiates an update request such as add, delete, or modify operations, the primary node replicates data to the secondary node immediately after the primary node completes the operation. The primary node then waits for the secondary node to return a success message before the primary node responds to the application. The primary node replicates data to the secondary node synchronously. Unavailability of the secondary node will affect the operation on the primary node. Data will remain consistent even when the primary node is unavailable.
- Semi-synchronous replication (Semi-Sync): Data is typically replicated in Sync mode. When trying to replicate data to the secondary node, if an exception occurs causing the primary and secondary nodes to be unable to communicate with each other, the primary node will suspend response to the application. If the connection cannot be restored, the primary node will degrade to Async mode and restore response to the application after the Sync replication times out. In a situation such as this, the primary node becoming unavailable will lead to data inconsistency. After the secondary node or network connection is recovered, data replication between the two nodes is resumed, and the data replication mode will change from Async to Sync.

You can select different combinations of service priorities and data replication modes to improve availability based on the business features.

12.4.3 Backup and recovery service

This service supports data backup, dump, and recovery functions.

ApsaraDB for RDS can initiate database backup at any time. It can also restore databases to the status of any point in time based on backup policy, improving the traceability of data.

Backup

The Backup module compresses and uploads data and logs on both the primary and secondary nodes. ApsaraDB for RDS uploads backup files to OSS by default and dumps the backup files to a more cost-effective and persistent Archive Storage system. When the secondary node is operating properly, backup is always initiated on the secondary node. This will not affect the services on the primary node. When the secondary node is unavailable or damaged, the Backup module initiates backup on the primary node.

Recovery

The Recovery module restores backup files stored on OSS to a destination node.

- Primary node rollback: when an operation error occurs, rolls back the primary node to a specified point in time.
- Secondary node repair: when an irreparable fault occurs on the secondary node, creates a new secondary node to reduce risk.
- Read-only instance creation: creates a read-only instance from backup files.

Dump

The Dump module uploads, dumps, and downloads backup files. Currently, all backup data is uploaded to OSS for storage. You can obtain temporary links to download data as needed. In certain scenarios, the Dump module allows you to dump backup files from OSS to Archive Storage for more cost-effective and longer-term offline storage.

**Note:**

PPAS cannot support the download of backup files. It must back up data through pg_dump.

12.4.4 Monitoring service

ApsaraDB for RDS provides multilevel monitoring services across the physical, network, and application layers to ensure service availability.

Service

The Service module tracks the status of services that RDS depends on, such as SLB, OSS, log service, and Archive Storage, to ensure they are operating properly. Monitored metrics include functionality and response time. The Service module also uses logs to determine whether the internal RDS services are operating properly.

Network

The Network module tracks statuses at the network layer. It monitors the connectivity between ECS and RDS and between physical RDS servers. It also monitors the rates of packet loss on the VRouter and VSwitch.

OS

The OS module tracks the status of hardware and the OS kernel. The monitored items include:

- Hardware maintenance: The OS module constantly checks the operating status of the CPU, memory, motherboard, and storage device. It can predict faults in advance and automatically submit repair reports when it determines a fault is likely to occur.
- OS kernel monitoring: The OS module tracks all database calls and analyzes the causes of slow calls or call errors based on the kernel status.

Instance

The Instance module collects the following information on RDS instances:

- Instance availability information
- Instance capacity and performance metrics
- Instance SQL execution records

12.4.5 Scheduling service

The Resource module implements the scheduling of resources and services.

Resource

The Resource module allocates and integrates underlying RDS resources when you activate and migrate instances. When you use the RDS console or API to create an instance, the Resource module calculates the most suitable host to carry the traffic to and from the instance. This module also allocates and integrates the underlying resources required to migrate RDS instances. After repeated instance creation, deletion, and migration operations, the Resource module calculates the degree of resource fragmentation. It also regularly integrates resources to improve the service carrying capacity.

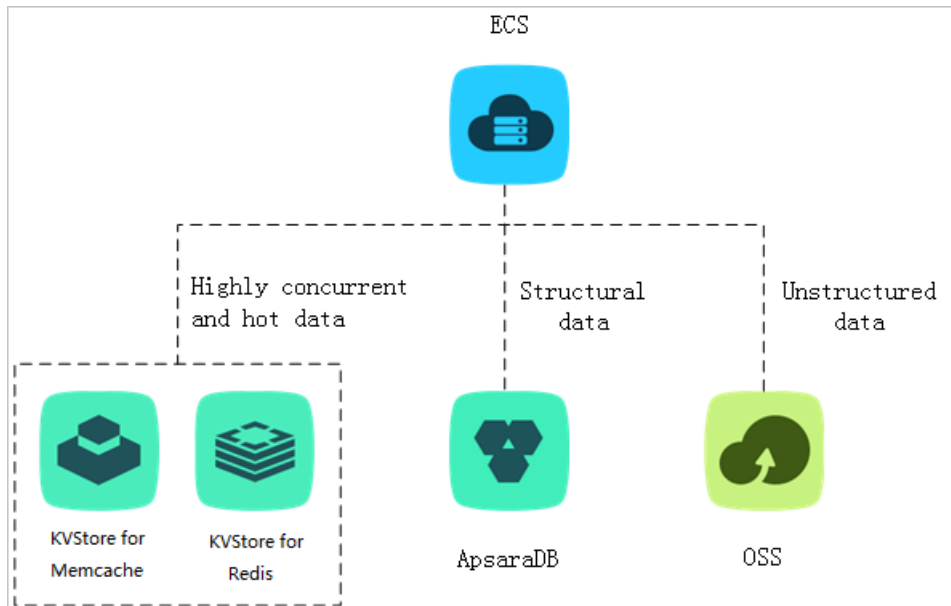
12.5 Scenarios

12.5.1 Diversified data storage

ApsaraDB for RDS provides cache data persistence and multi-structure data storage.

You can diversify the storage capabilities of ApsaraDB for RDS through services such as KVStore for Memcache, KVStore for Redis, and OSS, as shown in [Figure 12-2: Diversified data storage](#).

Figure 12-2: Diversified data storage



Cache data persistence

ApsaraDB for RDS can be used with KVStore for Memcache and KVStore for Redis to form a high-throughput and low-latency storage solution. ApsaraDB cache services have the following benefits over ApsaraDB for RDS:

- High response speed: The request latency of KVStore for Memcache and KVStore for Redis is usually within just a few milliseconds.
- The cache area supports a higher number of queries per second (QPS) than ApsaraDB for RDS.

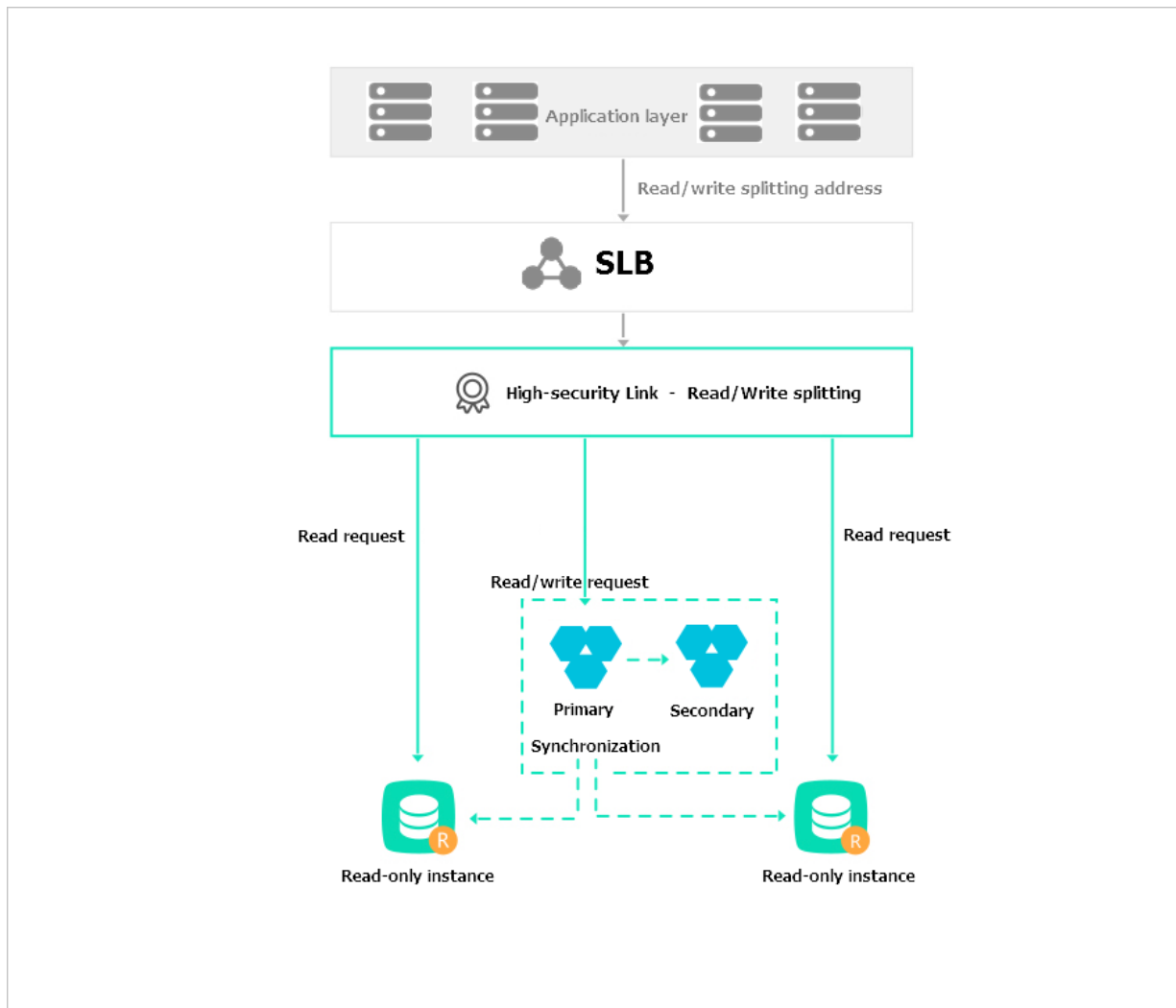
Multi-structure data storage

OSS is a secure and reliable high-capacity storage service from Alibaba Cloud with low costs. ApsaraDB for RDS can be used with OSS to form a multi-type data storage solution. For example, imagine ApsaraDB for RDS and OSS are used together to implement an online forum. Resources such as the images of registered users and posts on the forum can be stored in OSS to reduce storage needs on ApsaraDB for RDS.

12.5.2 Read/write splitting

This feature allows you to split read requests and write requests across different instances to expand the processing capability of the system.

ApsaraDB RDS for MySQL allows you to directly attach read-only instances to ApsaraDB for RDS to reduce read pressure on the primary instance. The primary instance and read-only instances of ApsaraDB RDS for MySQL each have their own connection addresses. The system also offers an extra read/write splitting address after read/write splitting is enabled. This address associates the primary instance with all of its read-only instances for automatic read/write splitting, allowing applications to send all read and write requests to a single address. Write requests are automatically routed to the primary instance, and read requests are routed to each read-only instance based on their weights. You can scale out the processing capability of the system by adding more read-only instances. There is no need to modify applications, as shown in [Read/write splitting](#).

Figure 12-3: Read/write splitting

12.6 Usage limits

12.6.1 Usage limits of ApsaraDB RDS for MySQL

Before you use ApsaraDB RDS for MySQL, you need to understand its limits and take precautions against them.

To guarantee instance stability and security, ApsaraDB RDS for MySQL has some service limits, as listed in [Table 12-2: Limits on ApsaraDB RDS for MySQL](#).

Table 12-2: Limits on ApsaraDB RDS for MySQL

Operation	Description
Database parameter modification	Database parameters can only be modified from the RDS console or through APIs. Due to security and stability considerations, only specific parameters can be modified.
Root permission of databases	The root and SA permissions are not provided.
Database backup	<ul style="list-style-type: none"> Logical backup can be performed from the command line interface (CLI) or graphical user interface (GUI). Physical backup can only be performed from the RDS console or through APIs.
Database restoration	<ul style="list-style-type: none"> Logical restoration can be performed from the CLI or GUI. Physical restoration can only be performed from the RDS console or through APIs.
Data import	<ul style="list-style-type: none"> Logical import can be performed from the CLI or GUI. Data can only be immigrated by using the MySQL command-line client.
ApsaraDB RDS for MySQL storage engine	<ul style="list-style-type: none"> Only InnoDB and TokuDB are supported. Due to the inherent defects of the MyISAM engine, data may be lost. Only some stock instances are using MyISAM engine. MyISAM engine tables in newly created instances will be automatically converted to InnoDB engine tables. For safety performance and security considerations, we recommend that you use the InnoDB storage engine. The Memory engine is not supported. Newly created Memory tables will be automatically converted into InnoDB tables.
Database replication	ApsaraDB RDS for MySQL provides dual-node clusters based on a primary/secondary replication architecture. The secondary instances in this replication architecture are hidden and cannot be accessed directly.
RDS instance restart	Instances must be restarted through the RDS console or APIs.
Account and database management	ApsaraDB RDS for MySQL uses the RDS console to manage accounts and databases by default. ApsaraDB RDS for MySQL also allows you to create a superuser account to manage users, passwords, and databases.
Standard account	<ul style="list-style-type: none"> Custom authorization is not supported. The account management and database management interfaces are provided in the RDS console.

Operation	Description
	<ul style="list-style-type: none"> Instances that support standard accounts also support superuser accounts.
Superuser account	<ul style="list-style-type: none"> Custom authorization is supported. The account management and database management interfaces are not provided in the RDS console. The relevant operations can only be performed through code or DMS. The superuser account cannot be reverted back into a standard account.

12.6.2 Usage limits of ApsaraDB RDS for PostgreSQL

Before you use ApsaraDB RDS for PostgreSQL, you need to understand its limits and take precautions against them.

To guarantee instance stability and security, ApsaraDB RDS for PostgreSQL has some service limits, as listed in [Table 12-3: Limits on ApsaraDB RDS for PostgreSQL](#).

Table 12-3: Limits on ApsaraDB RDS for PostgreSQL

Operation	Description
Database parameter modification	Not supported.
Root permission of databases	Superuser permissions are not provided.
Database backup	Data can only be backed up by using pg_dump.
Data migration	Only PostgreSQL can be used to restore data that was backed up by using pg_dump.
Database replication	<ul style="list-style-type: none"> The system automatically builds HA databases based on PostgreSQL streaming replication without user input. PostgreSQL standby nodes are hidden and cannot be accessed directly.
RDS instance restart	RDS instances must be restarted from the RDS console or through APIs.
Network settings	For instances that are operating in safe mode, net.ipv4.tcp_timestamps cannot be enabled in SNAT mode.

12.6.3 Usage limits of ApsaraDB RDS for PPAS

Before you use ApsaraDB RDS for PPAS, you must understand its limits and take precautions against them.

To guarantee instance stability and security, ApsaraDB RDS for PPAS has some service limits, as listed in [Table 12-4: Limits on ApsaraDB RDS for PPAS](#).

Table 12-4: Limits on ApsaraDB RDS for PPAS

Operation	Description
Database parameter modification	Not supported.
Root permission of databases	Superuser permissions are not provided.
Database backup	Data can only be backed up by using <code>pg_dump</code> .
Data migration	Only PostgreSQL can be used to restore data that was backed up by using <code>pg_dump</code> .
Database replication	<ul style="list-style-type: none">The system automatically builds HA databases based on PPAS streaming replication without user input.PPAS standby nodes are hidden and cannot be accessed directly.
RDS instance restart	RDS instances must be restarted from the RDS console or through APIs.
Network settings	For instances that are operating in safe mode, <code>net.ipv4.tcp_timestamps</code> cannot be enabled in SNAT mode.

12.7 Terms

Term	Description
Region	The geographical location where the server of your RDS instance resides. You must specify a region when you create an RDS instance. The region of an instance cannot be changed after instance creation. RDS must be used together with ECS and only supports intranet access. Because of this, RDS instances must be located in the same region as their corresponding ECS instances.
Zone	The physical area with an independent power supply and network in a region. Zones in a region can communicate through the intranet. Network latency for resources within the same zone is lower than for those across zones.

Term	Description
	Faults are isolated between zones. Single zone refers to the case where the three nodes in the RDS instance replica set are all located in the same zone . Network latency is reduced if an ECS instance and its corresponding RDS instance are both deployed in the same zone.
Instance	The most basic unit of RDS. An instance is the operating environment of ApsaraDB for RDS and works as an independent process on a host. You can create, modify, or delete an RDS instance from the RDS console. Instances are mutually independent and their resources are isolated. They do not compete for resources such as CPU, memory, or I/O. Each instance has its own features, such as database type and version. RDS controls instance behavior by using corresponding parameters.
Memory	The maximum amount of memory that can be used by an ApsaraDB for RDS instance.
Disk capacity	The amount of disk space selected when creating an ApsaraDB for RDS instance. Instance data that occupies disk space includes aggregated data as well as data required for normal instance operations such as system databases , database rollback logs, redo logs, and indexing. Ensure that the disk capacity is sufficient for the RDS instance to store data. Otherwise, the RDS instance may be locked. If the instance is locked due to insufficient disk capacity, you can unlock the instance by expanding the disk capacity.
IOPS	The maximum number of read/write operations performed per second on block devices at a granularity of 4 KB.
CPU core	The maximum computing capability of the instance. A single Intel Xeon series CPU core has at least 2.3 GHz of computational power with hyper-threading capabilities.
Number of connections	The number of TCP connections between a client and an RDS instance. If the client uses a connection pool, the connection between the client and RDS instance is a persistent connection. Otherwise, it is a transient connection.

12.8 Instance types

Instances of different series, versions, and types each perform differently from one another.

Table 12-5: ApsaraDB RDS for MySQL-type instance parameters

Series	Version	Type	Code	CPU and memory	Maximum connections	Maximum IOPS	Disk capacity	Zone-disaster recovery deployment
High-availability (HA) edition	5.6	Shared instance	rds.mysql.t1.small	1 core, 1 GB	300	600	5 GB to 2 TB	Single-IDC deployment Dual-IDC deployment
			rds.mysql.s1.small	1 core, 2 GB	600	1,000		
			rds.mysql.s2.large	2 cores, 4 GB	1,200	2,000		
			rds.mysql.s2.xlarge	2 cores, 8 GB	2,000	4,000		
			rds.mysql.s3.large	4 cores, 8 GB	2,000	5,000		
			rds.mysql.m1.medium	4 cores, 16 GB	4,000	7,000		
			rds.mysql.c1.large	8 cores, 16 GB	4,000	8,000		
			rds.mysql.c1.xlarge	8 cores, 32 GB	8,000	12,000		
			rds.mysql.c2.xlarge	16 cores, 64 GB	16,000	14,000		

Series	Version	Type	Code	CPU and memory	Maximum connections	Maximum IOPS	Disk capacity	Zone-disaster recovery deployment
			rds.mysql.c2.xlp2	16 cores, 96 GB	24,000	16,000		
		Dedicated instance (X8)	mysql.x8.medium.2	2 cores, 16 GB	2,500	4,500	250 GB	
			mysql.x8.large.2	4 cores, 32 GB	5,000	9,000	500 GB	
			mysql.x8.xlarge.2	8 cores, 64 GB	10,000	18,000	1 TB	
			mysql.x8.2xlarge.2	16 cores, 128 GB	20,000	36,000	2 TB	
		Dedicated instance (X4)	mysql.x4.large.2	4 cores, 16 GB	2,500	4,500	250 GB	
			mysql.x4.xlarge.2	8 cores, 32 GB	5,000	9,000	500 GB	
			mysql.x4.2xlarge.2	16 cores, 64 GB	10,000	18,000	1 TB	
			mysql.x4.4xlarge.2	32 cores, 128 GB	20,000	36,000	2 TB	
		Dedicated host	rds.mysql.st.d13	30 cores, 220 GB	64,000	20,000	3 TB	
ApsaraDB RDS for MySQL finance edition (3 nodes)	5.6	Dedicated instance (computing)	mysql.x4.large.3	4 cores, 16 GB	2,500	4,500	250 GB to 500 GB	Single-IDC deployment
			mysql.x4.xlarge.3	8 cores, 32 GB	5,000	9,000	500 GB to 1 TB	

Series	Version	Type	Code	CPU and memory	Maximum connections	Maximum IOPS	Disk capacity	Zone-disaster recovery deployment
			mysql.x4.2xlarge.3	16 cores, 64 GB	10,000	18,000	1 TB to 2 TB	
			mysql.x4.4xlarge.3	32 cores, 128 GB	20,000	36,000	2 TB to 3 TB	
		Dedicated instance (high memory)	mysql.x8.medium.3	2 cores, 16 GB	2,500	4,500	250 GB	
			mysql.x8.large.3	4 cores, 32 GB	5,000	9,000	500 GB	
			mysql.x8.xlarge.3	8 cores, 64 GB	10,000	18,000	1 TB	
			mysql.x8.2xlarge.3	16 cores, 128 GB	20,000	36,000	2 TB	
			mysql.x8.4xlarge.3	32 cores, 256 GB	40,000	72,000	3 TB	
ApsaraDB RDS for MySQL finance edition (4 nodes)	5.6	Dedicated instance (high memory)	mysql.x8.medium.4	2 cores, 16 GB	2,500	4,500	250 GB	Dual-IDC deployment
			mysql.x8.large.4	4 cores, 32 GB	5,000	9,000	500 GB	
			mysql.x8.xlarge.4	8 cores, 64 GB	10,000	18,000	1 TB	
			mysql.x8.2xlarge.4	16 cores, 128 GB	20,000	36,000	2 TB	
			mysql.x8.4xlarge.4	32 cores, 256 GB	40,000	72,000	3 TB	

Series	Version	Type	Code	CPU and memory	Maximum connections	Maximum IOPS	Disk capacity	Zone-disaster recovery deployment
Read-only instance	5.6	Common instance	rds.mysql.t1.small	1 core, 1 GB	300	600	5 GB to 2 TB	Single-IDC deployment, Dual-IDC deployment
			rds.mysql.s1.small	1 core, 2 GB	600	1,000		
			rds.mysql.s2.large	2 cores, 4 GB	1,200	2,000		
			rds.mysql.s2.xlarge	2 cores, 8 GB	2,000	4,000		
			rds.mysql.s3.large	4 cores, 8 GB	2,000	5,000		
			rds.mysql.m1.medium	4 cores, 16 GB	4,000	7,000		
			rds.mysql.c1.large	8 cores, 16 GB	4,000	8,000		
			rds.mysql.c1.xlarge	8 cores, 32 GB	8,000	12,000		
			rds.mysql.c2.xlarge	16 cores, 64 GB	16,000	14,000		
			rds.mysql.c2.xlp2	16 cores, 96 GB	24,000	16,000		
		Dedicated instance (X8)	mysqlro.x8.	2 cores, 16 GB	2,500	4,500	250 GB	

Series	Version	Type	Code	CPU and memory	Maximum connections	Maximum IOPS	Disk capacity	Zone-disaster recovery deployment
			medium .1					
			mysqlro.x8.large .1	4 cores , 32 GB	5,000	9,000	500 GB	
			mysqlro.x8.xlarge .1	8 cores , 64 GB	10,000	18,000	1 TB	
			mysqlro.x8.2xlarge.1	16 cores, 128 GB	20,000	36,000	2 TB	
		Dedicated instance (X4)	mysqlro.x4.large .1	4 cores , 16 GB	2,500	4,500	250 GB	
			mysqlro.x4.xlarge .1	8 cores , 32 GB	5,000	9,000	500 GB	
			mysqlro.x4.2xlarge.1	16 cores, 64 GB	10,000	18,000	1 TB	
			mysqlro.x4.4xlarge.1	32 cores, 128 GB	20,000	36,000	2 TB	
		Dedicated host	rds.mysql.st.d13	30 cores, 220 GB	64,000	20,000	3 TB	

Table 12-6: ApsaraDB RDS for PostgreSQL-type instance parameters

Series	Version	Type	Code	CPU and memory	Maximum connections	Maximum IOPS	Disk capacity
HA edition	9.4	Common instance	rds.pg.t1.small	1 core, 1 GB	2,000	400	5 GB to 2 TB
			rds.pg.s1.small	1 core, 2 GB	2,000	400	
			rds.pg.s2.large	2 cores, 4 GB	2,000	1,000	
			rds.pg.s3.large	4 cores, 8 GB	2,000	2,000	
			rds.pg.c1.large	8 cores, 16 GB	2,000	4,000	
			rds.pg.c1.xlarge	8 cores, 32 GB	2,000	4,000	
			rds.pg.c2.xlarge	16 cores, 64 GB	2,000	14,000	
		Dedicated instance (X8)	pg.x8.medium.2	2 cores, 16 GB	2,500	4,500	250 GB
			pg.x8.large.2	4 cores, 32G	5,000	9,000	500 GB
			pg.x8.xlarge.2	8 cores, 64 GB	10,000	18,000	1 TB
			pg.x8.2xlarge.2	16 cores, 128 GB	12,000	36,000	2 TB
		Dedicated instance (X4)	pg.x4.large.2	4 cores, 16 GB	2,500	4,500	250 GB
			pg.x4.xlarge.2	8 cores, 32 GB	5,000	9,000	500 GB
			pg.x4.2xlarge.2	16 cores, 64 GB	10,000	18,000	1 TB
			pg.x4.4xlarge.2	32 cores, 128 GB	12,000	36,000	2 TB

Series	Version	Type	Code	CPU and memory	Maximum connections	Maximum IOPS	Disk capacity
		Dedicated host	rds.pg.st.d13	30 cores, 220 GB	4,000	20,000	3 TB
			rds.pg.st.h43	60 cores, 470 GB	4,000	50,000	3 TB

Table 12-7: ApsaraDB RDS for PPAS-type instance parameters

Series	Version	Type	Code	CPU and memory	Maximum connections	Maximum IOPS	Disk capacity
HA edition	9.6	Common instance	rds.ppas.t1.small	1 core, 1 GB	100	600	5 GB to 2 TB
			rds.ppas.s1.small	1 core, 2 GB	200	1,000	
			rds.ppas.s2.large	2 cores, 4 GB	400	2,000	
			rds.ppas.s3.large	4 cores, 8 GB	800	5,000	
			rds.ppas.m1.medium	4 cores, 16 GB	1,500	8,000	
			rds.ppas.c1.xlarge	8 cores, 32 GB	2,000	12,000	
			rds.ppas.c2.xlarge	16 cores, 64 GB	2,000	14,000	
		Dedicated instance	ppas.x8.medium.2	2 cores, 16 GB	2,500	4,500	250 GB
			ppas.x8.large.2	4 cores, 32 GB	5,000	9,000	500 GB
			ppas.x8.xlarge.2	8 cores, 64 GB	10,000	18,000	1 TB
			ppas.x8.2xlarge.2	16 cores, 128 GB	12,000	36,000	2 TB

Series	Version	Type	Code	CPU and memory	Maximum connections	Maximum IOPS	Disk capacity
		Dedicated host	rds.ppas.st.d13	30 cores, 220 GB	4,000	20,000	3 TB
			rds.ppas.st.h43	60 cores, 470 GB	4,000	50,000	3 TB

13 KVStore for Redis

13.1 What is KVStore for Redis

KVStore for Redis is an online storage service compatible with the Redis protocol. It supports multiple data types, such as the string, list, set, sorted set, and hash. It also supports advanced features such as transactions and subscribe-publish (Sub/Pub). KVStore for Redis meets persistent storage requirements and provides fast read/write capabilities by using a combined flash memory and hard disk storage architecture.

KVStore for Redis is used as a cloud computing service, with hardware and data deployed on the cloud, providing comprehensive infrastructure planning, network security protection, and system maintenance services.

13.2 Benefits

Cluster functions

- The cluster function supports ultra-high capacity and performance. Cluster instances provide 128 GB or larger capacity, meeting requirements for large capacity and high performance.
- Master-slave dual-node instances are of 64 GB or smaller capacity, meeting average users' requirements for capacity and performance.

Elastic resizing

- One-click storage resizing: You can use the console to adjust the storage capacity of your instances as needed.
- Online resizing with no service interruption: You can adjust the instance capacity online without suspending your services or affecting your business.

Resource isolation

Instance-level resource isolation provides enhanced stability for individual services.

Data security

- Persistent data storage: With memory plus hard disk storage, KVStore for Redis provides high-speed data read/write capability and meets the data persistence requirements.
- Master-slave dual-backup for data: All data on the master node has a backup copy on the slave node.
- Access control: Password authentication is required for secure and reliable access.

- Data transmission encryption: Secure Sockets Layer (SSL) and Transport Layer Security (TLS) are supported for data transmission security.

High availability

- Each instance has a master node and a slave node: This prevents service interruption caused by SPOF.
- Automatic detection and recovery of hardware failure: This feature can automatically detect hardware failures and fail over to the slave node, restoring service in a matter of seconds.

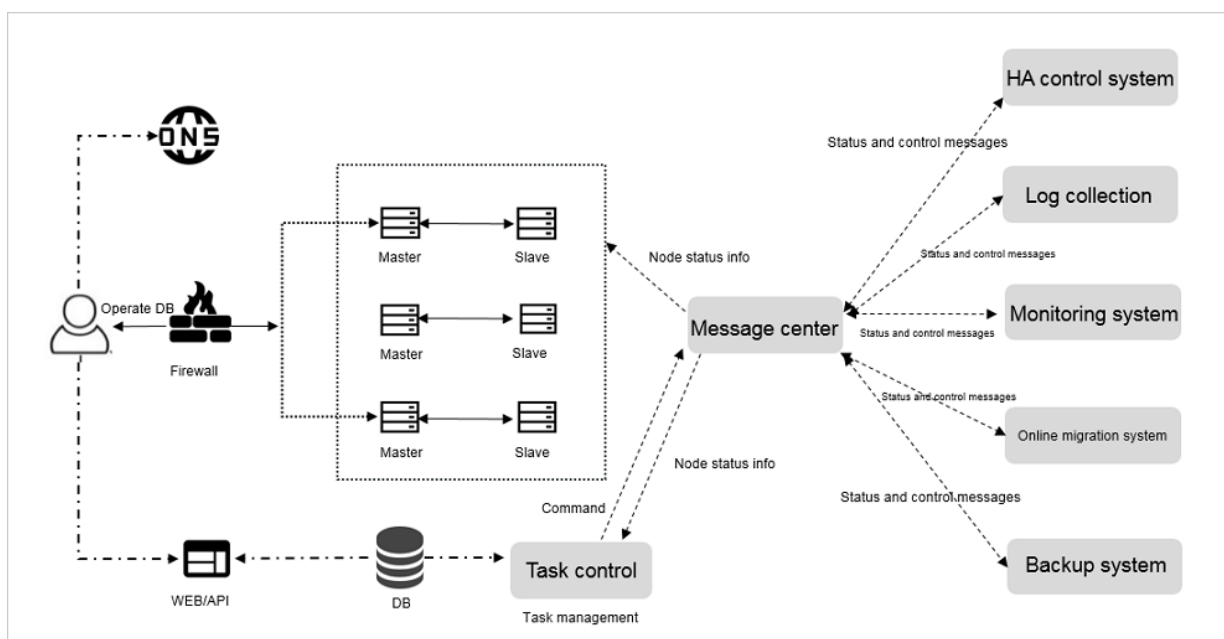
Easy to use

- Out-of-the-box service: This product requires no setup or installation and can be used right after purchase for quick and convenient business deployment.
- Compatible with open-source Redis: This product is compatible with Redis commands, and any Redis client can easily establish a connection with KVStore for Redis to perform data operations.

13.3 Architecture

Figure 13-1: Architecture of KVStore for Redis shows the architecture of KVStore for Redis.

Figure 13-1: Architecture of KVStore for Redis



KVStore for Redis automatically constructs a primary/secondary instance architecture.

- **HA control system**

The HA control system is a high-availability detection module. It is used to detect and monitor the operating status of KVStore for Redis instances. If this module determines that a primary node is unavailable, it fails over to the secondary node to ensure high availability of the KVStore for Redis instances.

- **Log collection**

This module collects instance operation logs, including slow query logs and RAM logs.

- **Monitoring system**

This module collects performance monitoring information of KVStore for Redis instances, including monitoring information for basic groups, key groups, and string groups.

- **Online migration system**

When a physical server that hosts an instance fails, the online migration system recreates an instance based on the backup files in the backup system. This ensures business continuity.

- **Backup system**

This module generates and stores the backup files of KVStore for Redis instances on OSS. At present, the backup system allows you to customize the backup settings and temporary backup configurations. The backup files are retained for seven days.

- **Task control**

KVStore for Redis instances support various management and control tasks, including instance creation, configuration changes, and instance backup. The task control module flexibly controls tasks and executes task tracking and error management based on the commands you provide.

13.4 Features

- **HA technology ensures business continuity**

During system operations, data is synchronized between the primary and secondary nodes. If the primary node becomes faulty, the system automatically fails over to the secondary node within seconds. The entire process is automatic without affecting your business. The primary/secondary architecture ensures high availability of your business.

Cluster instances adopt a distributed architecture, with each node working in a primary/secondary mode. This provides automatic failover to ensure high availability of your business.

- **One-click backup and recovery and custom backup policies are supported**

You can use the console to perform backup operations and create automatic backup policies. Data is saved automatically and can be recovered with a single click within seven days. This helps minimize the impact of incorrect data operations.

- **Multiple network protection measures are used to protect data security**

VPC provides network isolation protection at the TCP layer. The anti-DDoS feature can monitor and clean large-volume attacks in real time. You can add up to 1,000 IP addresses to the whitelist.

- **Optimized kernels are used to prevent vulnerability attacks**

Alibaba Cloud expert team has performed deep kernel optimization on the Redis source code, effectively preventing out-of-memory errors and fixing security vulnerabilities to safeguard your business.

- **Auto scaling helps you overcome capacity and performance bottlenecks**

KVStore for Redis supports multiple memory configurations. Your memory can be automatically scaled depending on your business.

Cluster architecture allows automatic scaling of storage space and throughput of databases to eliminate performance bottlenecks.

- **Multiple instance types deliver flexible configurations**

The service supports the single-node cache architecture and dual-node storage architecture to suit different business scenarios. Instance configurations can be changed flexibly based on business scenarios.

- **The monitoring and alarming function allows you to view the instance status in real time**

The service provides monitoring information about the CPU utilization, connections, and disk usage, as well as the alarming function, so that you are fully aware of the real-time status of all instances.

- **The visual O&M platform simplifies O&M operations**

The visual O&M platform allows you to perform frequent and risky operations with a single click, such as instance cloning, backup, and restoration.

- **Automatic upgrade of database kernels prevents software bugs**

Automatic instance upgrade and fast bug fixing free you from routine version management.

- **Custom parameter configurations are supported**

You can customize Redis parameters to make full use of system resources.

13.5 Scenarios

Gaming industry applications

Game companies can use KVStore for Redis as an important component of their deployment architecture.

- **Scenario 1: Using KVStore for Redis for data storage**

Game deployment architecture is usually relatively simple. With the main program deployed on ECS, all business data are stored in Redis as a persistent database. KVStore for Redis supports persistence function, with master-slave dual-node redundant data storage.

- **Scenario 2: Using KVStore for Redis as a cache to accelerate application access**

Using Redis as a cache layer will accelerate application access. Data are stored in a backend database (RDS).

Reliability is critical to KVStore for Redis services. Once a KVStore for Redis service becomes unavailable, business access may overload the backend database. KVStore for Redis uses a hot standby high-availability architecture to ensure extremely high service reliability. The master node provides external services. If this node fails, the system will automatically set up the standby node to take over the services. The entire failover process is completely transparent to users.

Live video applications

Live video services are often highly reliant on KVStore for Redis to store user data and friends interaction information.

- **Dual-node hot standby ensures high availability**

KVStore for Redis provides the hot-standby mode to maximize service availability.

- **Cluster version solves the performance bottleneck**

KVStore for Redis provides cluster version instances to break through the performance bottleneck of Redis single-thread mechanism. This approach can effectively cope with spikes in live video broadcast traffic and meet high performance requirements.

- **Easy resizing helps cope with business peaks**

KVStore for Redis can support one-click resizing. The entire upgrade process is fully transparent to you and helps you easily cope with traffic bursts.

E-commerce industry applications

In the e-commerce industry, Redis is extensively used, mostly for item display, shopping recommendations, and other modules.

- **Scenario 1: Seckill-type shopping systems**

During large-scale seckill promotions, a shopping system will be overwhelmed by traffic, which far exceeds the Read/Write capability of common databases.

The persistence function supported by KVStore for Redis allows you to directly use Redis as a database system.

- **Scenario 2: Inventory system with a counter**

In such a system, the underlying architecture usually keeps actual data in RDS and count information in database fields. KVStore for Redis reads the counts information while ApsaraDB for RDS stores the count information. In this scenario, KVStore for Redis is deployed on a physical machine, with an underlying architecture based on SSD high-performance storage that can provide high-level data reading capabilities.

13.6 Limits

Item	Description
List data type	You can create an unlimited number of lists. The size of a single list element cannot exceed 512 MB. We recommend that each list contain a maximum of 8.192 elements with a maximum value size of 1 MB.
Set data type	You can create an unlimited number of sets. The size of a single set element cannot exceed 512 MB. We recommend that each set contain a maximum of 8.192 elements with a maximum value size of 1 MB.
Sorted set data type	You can create an unlimited number of sorted sets. The size of a single sorted set element cannot exceed 512 MB. We recommend that each sorted set contain a maximum of 8.192 elements with a maximum value size of 1 MB.
Hash data type	You can create an unlimited number of fields. The size of a single field element cannot exceed 512 MB. We recommend that each field contain a maximum of 8.192 elements with a maximum value size of 1 MB.
Number of databases	Each instance supports up to 256 databases.

Item	Description
Supported commands	For more information, see Supported commands in KVStore for Redis in <i>KVStore for Redis User Guide</i> .
Monitoring and alarming	KVStore for Redis does not provide the capacity alarming feature. You can configure it in CloudMonitor. We recommend that you set alarms for the following metrics: instance fault, primary/secondary instance failover, connection usage, failed operation count, capacity usage, write bandwidth usage, and read bandwidth usage.
Expired data deletion policies	<ul style="list-style-type: none"> Active expiration: The system periodically detects and deletes expired keys in the background. Passive expiration: The system deletes expired keys upon key access by users.
Idle connection reclaim mechanism	Idle Redis connections are not automatically recovered by the server, but managed by users.
Data persistence policies	AOF_FSYNC_EVERYSEC is enabled, and fsync is performed every second.

13.7 Concepts

KVStore for Redis

A high-performance key-value storage system (cache and store) based on the BSD protocol.

Instance ID

An instance corresponds to a user space, and serves as the basic unit of using KVStore for Redis.

KVStore for Redis sets limits on parameters such as connections, bandwidth, and CPU capacity based on the capacity specifications of individual instances. On the console, you can view the list of IDs for the instances you have purchased. There are two types of KVStore for Redis instances: primary/secondary instances and high-performance cluster instances.

Primary/secondary instance

A Redis instance that adopts a primary/secondary architecture. Primary/secondary instances are limited in terms of capacity and performance.

High-performance cluster instance

A Redis instance that adopts a scalable cluster architecture. Cluster instances have higher scalability and performance, but they have limited functionality.

Connection address

The host address used to connect to KVStore for Redis. It is displayed as a domain name and can be found in **Connection Information** on the **Instance Information** tab page.

Eviction policy

The policy to evict old data from the KVStore for Redis memory when the threshold configured with the `maxmemory` parameter is exceeded. The eviction policy of KVStore for Redis is consistent with that of the Redis protocol. For more information, see [Redis eviction policies](#).

Database

KVStore for Redis supports up to 256 databases. Data is stored in database 0 by default.

13.8 Instance types



Note:

The maximum bandwidth applies to the maximum upstream bandwidth and the maximum downstream bandwidth.

Standard edition

Table 13-1: Standard package

Instance type	Service code	Maximum connections	Maximum bandwidth (MB)	Processing capacity	Description	Zone-disaster recovery deployment
1 GB primary/secondary standard edition with zone-disaster recovery	redis.logic.sharding.drredisdb1g.1db.0rodb.4proxy.default	10,000	10	Single-core	Primary/secondary zone-disaster recovery	2-IDC zone-disaster recovery mode
2 GB primary/secondary standard edition	redis.logic.sharding.drredisdb2g.1db.0rodb.4proxy.default	10,000	16	Single-core	Primary/secondary zone-disaster recovery	2-IDC zone-disaster recovery mode

Instance type	Service code	Maximum connections	Maximum bandwidth (MB)	Processing capacity	Description	Zone-disaster recovery deployment
with zone-disaster recovery						
4 GB primary/secondary standard edition with zone-disaster recovery	redis.logic.sharding.drredisdb4g.1db.0rodb.4proxy.default	10,000	24	Single-core	Primary/secondary zone-disaster recovery	2-IDC zone-disaster recovery mode
8 GB primary/secondary standard edition with zone-disaster recovery	redis.logic.sharding.drredisdb8g.1db.0rodb.4proxy.default	10,000	24	Single-core	Primary/secondary zone-disaster recovery	2-IDC zone-disaster recovery mode
16 GB primary/secondary standard edition with zone-disaster recovery	redis.logic.sharding.drredisdb16g.1db.0rodb.4proxy.default	10,000	32	Single-core	Primary/secondary zone-disaster recovery	2-IDC zone-disaster recovery mode
32 GB primary/secondary standard edition with zone-disaster recovery	redis.logic.sharding.drredisdb32g.1db.0rodb.4proxy.default	10,000	32	Single-core	Primary/secondary zone-disaster recovery	2-IDC zone-disaster recovery mode

Table 13-2: Premium package

Instance type	Service code	Maximum connections	Maximum bandwidth (MB)	Processing capacity	Description	Zone-disaster recovery deployment
1 GB primary/secondary advanced edition	redis.master.small.special2x	20,000	48	Single-core	Primary/secondary instance	Single-IDC mode
2 GB primary/secondary advanced edition	redis.master.mid.special2x	20,000	48	Single-core	Primary/secondary instance	Single-IDC mode
4 GB primary/secondary advanced edition	redis.master.stand.special2x	20,000	48	Single-core	Primary/secondary instance	Single-IDC mode
8 GB primary/secondary advanced edition	redis.master.large.special1x	20,000	48	Single-core	Primary/secondary instance	Single-IDC mode
16 GB primary/secondary advanced edition	redis.master.2xlarge.special1x	20,000	48	Single-core	Primary/secondary instance	Single-IDC mode
32 GB primary/secondary advanced edition	redis.master.4xlarge.special1x	20,000	48	Single-core	Primary/secondary instance	Single-IDC mode

Table 13-3: Cluster edition

Instance type	Service code	Maximum connections	Maximum bandwidth (MB)	Processing capacity	Description
16 GB cluster edition	redis.sharding.small.default	80,000	384	4-core	High-performance cluster instance
32 GB cluster edition	redis.sharding.small.default	80,000	384	8-core	High-performance cluster instance
64 GB cluster edition	redis.sharding.large.default	80,000	384	8-core	High-performance cluster instance
128 GB cluster edition	redis.sharding.2xlarge.default	160,000	768	16-core	High-performance cluster instance
256 GB cluster edition	redis.sharding.4xlarge.default	160,000	768	16-core	High-performance cluster instance

Table 13-4: Cluster edition with zone-disaster recovery

Instance type	Service code	Maximum connections	Maximum bandwidth (MB)	Processing capacity	Description
16 GB cluster edition with zone-disaster recovery	redis.logic.sharding.drredisdb16g.8db.0rodb.8proxy.default	80,000	384	8-core	Cluster edition with zone-disaster recovery

Instance type	Service code	Maximum connections	Maximum bandwidth (MB)	Processing capacity	Description
32 GB cluster edition with zone-disaster recovery	redis.logic.sharding.drredisdb32g.8db.0rodb.8proxy.default	80,000	384	8-core	Cluster edition with zone-disaster recovery
64 GB cluster edition with zone-disaster recovery	redis.logic.sharding.drredisdb64g.8db.0rodb.8proxy.default	80,000	384	8-core	Cluster edition with zone-disaster recovery
128 GB cluster edition with zone-disaster recovery	redis.logic.sharding.drredisdb128g.16db.0rodb.16proxy.default	160,000	768	16-core	Cluster edition with zone-disaster recovery
256 GB cluster edition with zone-disaster recovery	redis.logic.sharding.drredisdb256g.16db.0rodb.16proxy.default	160,000	768	16-core	Cluster edition with zone-disaster recovery

QPS reference

Table 13-5: QPS reference

Size (GB)	Maximum connections	Maximum bandwidth (MB)	Processing capacity	QPS reference value
8	10,000	24	Single-core	80,000



Note:

The QPS reference value ranges from 80,000 to 100,000 for non-cluster instances, while for cluster instances, it is 80,000–100,000 times the number of nodes.

14 ApsaraDB for MongoDB

14.1 What is ApsaraDB for MongoDB

ApsaraDB for MongoDB is a dedicated high-performance distributed data storage service. It is fully compatible with the MongoDB protocol and can provide stable, reliable, and automatically scalable database services. It offers a full range of database solutions, such as disaster recovery, backup and restoration, monitoring, and alarms.

ApsaraDB for MongoDB offers the following basic features:

- Automatically creates a three-node MongoDB replica set, which encapsulates advanced functions such as disaster tolerance switchover and failover.
- Supports one-click database backup and restoration. You can perform conventional database backup and database rollback with a single click on the ApsaraDB for MongoDB console.
- Provides more than 20 performance metrics for monitoring and alarm functions, giving you a full overview of database performance.
- Provides visual data management tools for convenient operations and maintenance.

14.2 Benefits

- **High availability**

- The three-node replica set HA architecture that delivers extremely high service availability.

The ApsaraDB for MongoDB service uses a three-node replica set HA architecture. The three data nodes are located on different hosts and automatically synchronize data. The primary and secondary nodes provide services. When the primary node fails, the system automatically elects a new primary node. When the secondary node is unavailable, the standby node takes over the services.

- Automatic backup and one-click data restoration

Data is automatically backed up and uploaded to Object Storage Service (OSS) every day. This improves data disaster recovery capabilities while effectively reducing disk space consumption. The backup files can restore the instance data to the original instance. This effectively prevents irreversible effects on service data caused by incorrect operations or other reasons.

- **High security**

- Anti-DDoS protection: provides real-time monitoring at the network entry point. When high-traffic attacks are identified, the source IP addresses are scrubbed. If scrubbing is ineffective, the black hole mechanism is triggered.
- IP whitelist configuration: Up to 1,000 IP addresses are allowed to connect to an ApsaraDB for MongoDB instance, directly controlling risks at the source.

- **Ease of use**

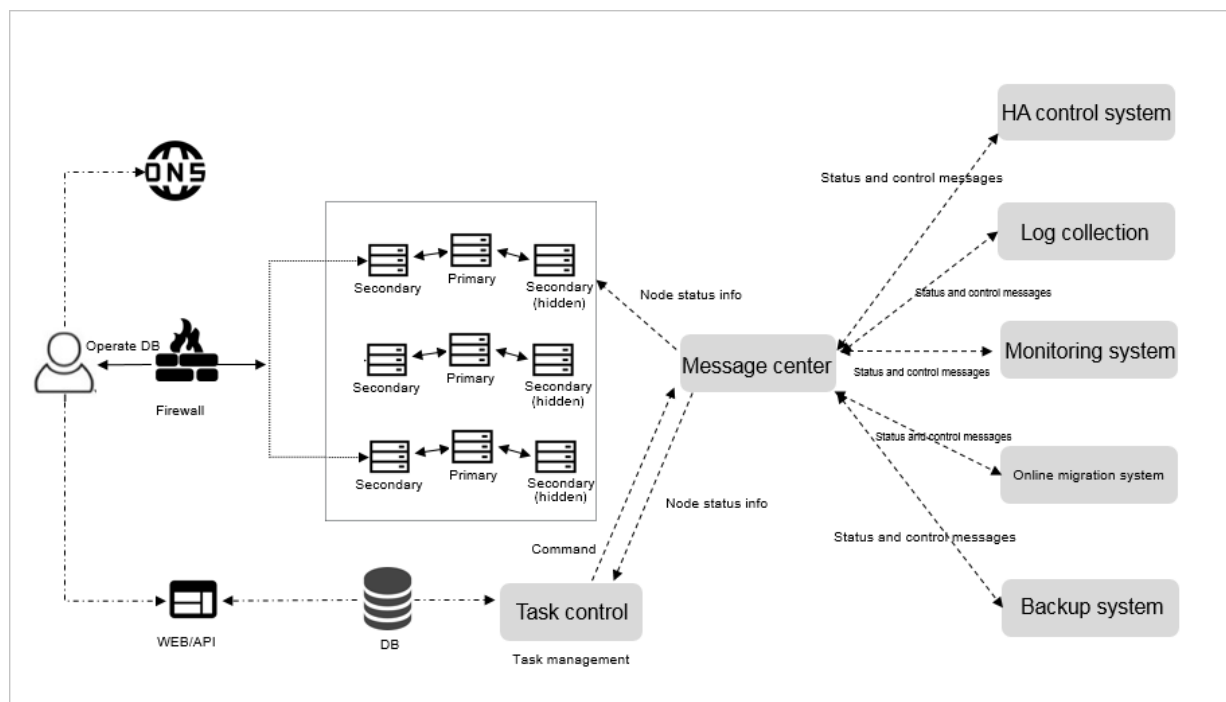
Sound performance monitoring: The product monitors instance information in real time, such as CPU utilization, IOPS, connections, and disk space and report alarms. This helps to keep you updated on instance statuses at all times.

- **Scalability**

The replica sets can be elastically scaled. ApsaraDB for MongoDB supports three-node replica sets to allow elastic scaling. You can change the configuration of your instance if the current configuration is too high or is unable meet the performance requirements of your application. The configuration change process is completely transparent and will not affect your business.

14.3 Architecture

ApsaraDB for MongoDB provides a three-node replica set for your use. You can directly use a primary or secondary node. The system architecture is shown in the following figure:



- **HA control system:** This is the instance HA detection module. The module is used to detect and listen to the running status of ApsaraDB for MongoDB instances. If the system detects that the primary node instance is unavailable, it will switch over to the secondary node to ensure the high availability of MongoDB instances.
- **Log collection:** This module collects MongoDB running logs, including the slow query logs and RAM logs of an instance.
- **Monitoring system:** This module collects performance monitoring information about MongoDB instances, including basic metrics, disk capacity, network requests, operation counts, and other core information.
- **Online migration system:** This module is used when the host on which the instance is deployed fails. The module will re-build an instance based on the backup files in the backup system and help to avoid service interruptions.
- **Backup system:** This module generates ApsaraDB for MongoDB instance backups and stores the backup files on the OSS system. Currently, this module allows you to customize the backup settings and create temporary backups. Files are retained for 7 days.
- **Task control:** This module is used to support various management and control tasks, including instance creation, configuration modification, and instance backup. The module flexibly controls tasks and performs troubleshooting based on your operations.

14.4 Features

Flexible architecture

ApsaraDB for MongoDB provides a three-node replica set for use. You can directly use a primary and secondary node. If the system detects that the primary node is unavailable, it will switch over to the secondary node to ensure the high availability of MongoDB instances.

Elastic scaling

- One-click scaling of storage capacity: You can scale the storage capacity of an instance on the ApsaraDB for MongoDB console based on business requirements.
- Online scaling without interrupting services: Instance storage capacity can be scaled online without having to stop services or affecting your business.

Data security

- Automatic backup: ApsaraDB for MongoDB allows you to set backup cycles. You can flexibly configure backup start times based on your service off-peak times. The backup files are retained for free for up to seven days.

- **Temporary backup:** You can initiate temporary backup as needed. The backup files are retained for free for up to seven days.
- **Data restore:** You can use backup files to directly overwrite existing data and restore an instance to a previous status.
- **Backup file download:** ApsaraDB for MongoDB retains your backup files for free for up to seven days. During this period, you can log on to the ApsaraDB for MongoDB console and download the backup files to a local device.
- **Creating instances from backup sets:** You can create an instance from backup files on the ApsaraDB for MongoDB console with a single click for fast deployment.
- **IP-based whitelist:** You can use IP address-based filtering to manage access to an instance. You can log on to the ApsaraDB for MongoDB console and configure a whitelist of up to 1000 IP addresses. This provides the most advanced access security protection.
- **Multi-layer network security protection:** VPCs provide network isolation at the TCP layer. Anti-DDoS can monitor and clean heavy attack traffic in real time. Up to 1,000 IP addresses can be added to the whitelist.

Intelligent operations and maintenance

- **Monitoring platform**

The platform provides monitoring information about the CPU utilization, connections, and disk utilization, as well as the alarm reporting function. The platform enables you to be fully aware of all instance statuses.

- **Graphical O&M platform**

The platform allows you to perform frequent and risky operations, such as instance cloning, backup, and one-click data restoration.

- **Database kernel version management**

This function proactively performs upgrades and quickly repairs defects, freeing you from routine version management. It also optimizes ApsaraDB for MongoDB parameter configurations and maximizes utilization of system resources.

14.5 Scenarios

- **Businesses that require read/write splitting**

The ApsaraDB for MongoDB service uses a high-availability architecture that features a three-node replica set. The three data nodes are located on different hosts. The secondary

and standby nodes automatically synchronize data from the primary node. The primary and secondary nodes provide services. The two nodes have independent domain names. Along with ApsaraDB for MongoDB Driver, the nodes can independently distribute read requests.

- **Flexible businesses**

As a schema-free database, ApsaraDB for MongoDB is particularly suitable for businesses in initial stages because it does not require you to change table structures. You can store data with fixed structures in ApsaraDB for RDS databases, data for flexible businesses in ApsaraDB for MongoDB databases, and frequently accessed data in ApsaraDB for Memcache databases or ApsaraDB for Redis databases. This helps you achieve efficient data storage and reduce costs.

- **Mobile applications**

ApsaraDB for MongoDB supports two-dimensional space indexes, making it ideal for location-based mobile app services. ApsaraDB for MongoDB adopts a dynamic storage method, which is suitable for storing heterogeneous data from multiple systems and meets the needs from mobile apps.

- **IoT applications**

ApsaraDB for MongoDB provides excellent performance and an asynchronous data writing function. In some scenarios, the service can provide the performance of an in-memory database, which makes it suitable for IoT high concurrency writing scenarios. The MapReduce function of the service supports aggregated analysis of large amounts of data.

- **Core log systems**

In asynchronous disk scenarios, ApsaraDB for MongoDB can provide excellent plugin performance and processing capabilities of an in-memory database. The service supports secondary indexes to meet the need for dynamic queries. It can use the MapReduce aggregation framework to perform multidimensional data analysis.

14.6 Limits

Operation	Limit
Create a database copy	The system automatically creates a three-node replica set. The primary and secondary nodes are provided to you. The standby node is hidden. Secondary nodes cannot be manually created by users.

Operation	Limit
Restart a database	Instances must be restarted on the ApsaraDB for MongoDB console.

14.7 Glossary

Concept	Description
Region	The geographical location of the server for the ApsaraDB for MongoDB instance that you created. You can specify a region when you create the instance. The region cannot be changed once the instance is created. When you create an ApsaraDB for MongoDB instance, you must use it with an Alibaba Cloud ECS instance. ApsaraDB for MongoDB only supports internal network access, so the specified region must be the same as that of the ECS instance.
Instance	An ApsaraDB for MongoDB instance or instance for short. An instance is the basic unit of the ApsaraDB for MongoDB service that you create. An instance is the operating environment for ApsaraDB for MongoDB and exists as a separate process on a host. You can create, modify, and delete an instance on the ApsaraDB for MongoDB console. Instances are mutually independent and their resources are isolated. They do not compete for CPU, memory, IO, and other resources. Each instance has its own features, such as database type and version. The system has corresponding parameters to control instance behavior.
Memory	The maximum memory that an instance can use.
Disk capacity	The disk size that you select when you create an instance. The disk capacity occupied by the instance includes set data and the space required for normal instance operations, such as the system database, database rollback log, redo log, and indexing. Ensure that the disk capacity is sufficient for the ApsaraDB for MongoDB instance to store data, otherwise, the instance may be locked. If an instance is locked because the disk capacity is insufficient, you can purchase a larger disk to unlock the instance.
IOPS	The maximum number of block device reads/writes per second. It is measured in units of 4 KB.
CPU core	The maximum computing power of an instance. A single core CPU has a minimum of 2.3 GHz hyperthreading (Intel Xeon series Hyper-Threading) computing power.
Connections	The TCP connections between clients and ApsaraDB for MongoDB instances. If a client uses a connection pool, the connections between the client and instance will be persistent connections. Otherwise, they will be short connections.

Concept	Description
Mongos	The request portals of ApsaraDB for MongoDB clusters. All requests must be coordinated through mongos that act as request distribution centers. Mongos forward data requests to the corresponding shard servers. You can use multiple mongos as request portals. If one goes offline, other mongos can process MongoDB requests.
ConfigServer	The server that stores all database metadata (routes and shards) configuration. Mongos do not have storage so they cache shard server information and data routing information in memory. The information is actually stored on the ConfigServer. When mongos are started for the first time or shut down and then restarted, they load configuration information from the ConfigServer. If the ConfigServer information changes, all mongos are notified to update their status. In this way, mongos always have correct routing information. The ConfigServer stores shard route metadata and has high requirements for service availability and data reliability. Therefore, ApsaraDB for MongoDB uses a three-node replica set to comprehensively ensure the reliability of the ConfigServer.

14.8 Instance specifications

Table 14-1: ApsaraDB for MongoDB replica set specifications

Type	Specification	Code	Maximum number of connections	Maximum IOPS
General specifications	1 core, 2 GB	dds.mongo.mid	500	1,000
	2 cores, 4 GB	dds.mongo.standard	1,000	2,000
	4 cores, 8 GB	dds.mongo.large	2,000	4,000
	8 cores, 16 GB	dds.mongo.xlarge	4,000	8,000
	8 cores, 32 GB	dds.mongo.2xlarge	8,000	14,000
	16 cores, 64 GB	dds.mongo.4xlarge	16,000	16,000
Dedicated specifications	2 cores, 16 GB	mongo.x8.medium	2,500	4,500
	4 cores, 32 GB	mongo.x8.large	5,000	9,000
	8 cores, 64 GB	mongo.x8.xlarge	10,000	18,000

Type	Specification	Code	Maximum number of connections	Maximum IOPS
	16 cores, 128 GB	mongo.x8.2xlarge	20,000	36,000
	32 cores, 256 GB	mongo.x8.4xlarge	40,000	72,000
Dedicated host	60 cores, 440 GB	dds.mongo.2xmonopolize	100,000	100,000

15 KVStore for Memcache

15.1 What is KVStore for Memcache

KVStore for Memcache is a memory-based cache service that supports high-speed access to large volumes of small data. KVStore for Memcache can greatly cut down the load on back-end storage, and speed up the response of websites and applications.

KVStore for Memcache supports the key-value data structure. KVStore for Memcache can communicate with clients compatible with the Memcached protocol.

KVStore for Memcache supports out-of-the-box deployment. It also relieves the loads of dynamic Web applications on databases through the cache service, thus improving the overall response speed of the website.

Similar to local self-built Memcached databases, KVStore for Memcache is also compatible with the Memcached protocol and user environments. You can use ApsaraDB for Memcache directly. The difference is that the hardware and data of ApsaraDB for Memcache are deployed on the cloud, which provides complete infrastructure, network security, and system maintenance services.

15.2 Benefits

Ease of use

- Immediate availability: An instance is immediately available after it is created, facilitating fast business deployment.
- Compatibility with open-source Memcached: KVStore for Memcache is compatible with Memcached Binary Protocol. All clients that support this protocol and SASL can connect to KVStore for Memcache.
- Visualized management and monitoring panel: The console provides multiple monitoring metrics for your convenience to manage KVStore for Memcache instances.

Cluster features

KVStore for Memcache supports super large capacity and provides super high performance. The default cluster output utilizes super large cluster instances to meet demands for large capacity and high performance.

Elastic scalability

- Scale-out of storage capacity with a single click: You can adjust the storage capacity of an instance in the console based on business requirements.
- Online scale-out without service interruption: You can adjust the instance capacity online without suspending your services or affecting your business.

Resource isolation

Instance-level resource isolation provides enhanced stability for individual services.

High security and reliability

- Password authentication is supported to ensure secure and reliable access.
- Persistent data storage: The use of memory and hard disks meets data persistence demands while high-speed data reading and writing are provided.

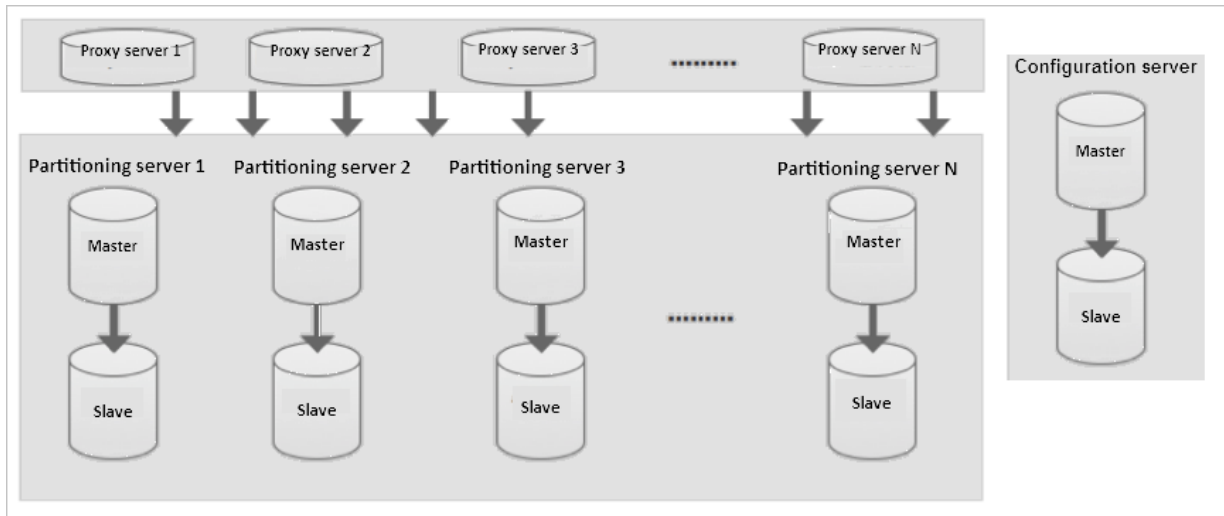
High availability

- Each instance has a primary node and a secondary node. This prevents service interruption caused by single point of failures (SPOFs).
- Automatic detection and recovery of hardware faults: KVStore for Memcache automatically detects hardware faults and fails services over within seconds to recover services.

15.3 Architecture

KVStore for Memcache uses a cluster-based architecture. It is embedded with data partitioning and reading algorithms. The whole process is transparent to you, saving development and O&M troubles. Each partition node uses master-slave architecture to ensure high availability of services.

KVStore for Memcache is comprised of three components, namely, the proxy server (service proxy), the partitioning server, and the configuration server.

Figure 15-1: Memcache architecture**Proxy server**

Single-noded. A cluster structure may contain multiple proxies. The system automatically implements load balancing and fail-over for proxies.

Partitioning server

Each partitioning server is in a dual-copy high-availability architecture. The system automatically implements the master-slave switchover in case of a fault in the master node to ensure high availability of services.

Configuration server

The server is used to store cluster configuration information and partitioning policies. It adopts dual-copy architecture to ensure high availability.

**Note:**

- The number and configuration of the three components are specified by the system at purchase and are not customizable. Specification details are as follows:

Specification	Number of proxies	Number of partitioning server	Memory size of single partitioning server
1 GB	1	1	1 GB
2 GB	1	1	2 GB
4 GB	1	1	4 GB

Specification	Number of proxies	Number of partitioning server	Memory size of single partitioning server
8 GB	1	1	8 GB
16 GB	2	2	8 GB
32 GB	4	4	8 GB
64 GB	8	8	8 GB
128 GB	16	16	8 GB
256 GB	16	16	16 GB
512 GB	32	32	16 GB

- A Memcache cluster exposes a uniform domain for access. You can visit this domain for normal access to and data operations on Memcache. The proxy server, the partitioning server, and the configuration server do not provide domain access and you cannot directly access them for operations.

15.4 Features

Distributed architecture, freeing businesses from the impact of SPOFs

- KVStore for Memcache uses a distributed cluster architecture. Each node is composed of two servers for hot backups and is capable of automatic disaster recovery and failovers.
- Multiple specifications are provided to handle different business stresses with unlimited database performance expansion.
- KVStore for Memcache supports data persistence and backup and recovery policies to effectively secure data reliability and avoid the impact of huge stress to the backend databases when the cache becomes invalid because of physical node faults.

A multi-level security defense system which helps you resist more than 90% of network attacks

- DDoS defense: Real-time monitoring of inbound network traffic is enabled. The source IP address will be cleaned in the event of flood attacks. If the cleaning turns out to be ineffective, the IP address will be silently discarded.
- IP address whitelist mechanism: A maximum of 1,000 server IP addresses can be configured in the whitelist for accessing an instance, directly putting risks under control at the source.
- VPC: KVStore for Memcache is fully compatible with VPCs and you can build an isolated network environment based on Alibaba Cloud.

- SASL authentication: SASL-enabled user identify authentication secures data access.

15.5 Scenarios

Frequently-accessed businesses

Businesses such as social networks, e-businesses, games and advertisements, can store frequently-accessed data in KVStore for Memcache and the underlying data in RDS.

Large promotion businesses

Large promotion or flash sales systems are usually under high access pressure. The average database simply cannot handle this amount of read stress, but KVStore for Memcache can be a viable alternative.

Inventory system with a counter

ApsaraDB for RDS and KVStore for Memcache can be used in combination. RDS stores the specific data information, while the database fields store the specific counter statistics. KVStore for Memcache reads the statistics, while RDS stores the statistics.

Data analysis businesses

KVStore for Memcache can be used in combination with open data processing service MaxCompute. It implements distributed analysis and processing of big data, which is suitable for big data processing scenarios such as business analysis and data mining. Data Integration service allows you to synchronize data between KVStore for Memcache and MaxCompute on your own, simplifying data operations.

15.6 Limits

Before using KVStore for Memcache, you need to understand the limits listed in the following table

Limit	Description
Data type	KVStore for Memcache supports only data formatted as key-value pairs and does not support complex data types such as array, map, and list.
Data reliability	KVStore for Memcache stores data in the memory, and does not guarantee that the cached data is never lost. Therefore, KVStore for Memcache is unsuitable for storing data that requires high consistency.
Data size	KVStore for Memcache supports a maximum of 1 KB in key size and 1 MB in value size for a single piece of cached data. KVStore for Memcache is unsuitable for storing sizable data.

Limit	Description
Transaction support	KVStore for Memcache does not support transactions. Therefore, KVStore for Memcache is unsuitable for storing transaction data. Such data must be written directly to the database.
Scenarios	When data access traffic is evenly distributed, and there is no obvious hotspot or less popular data, a large number of access requests cannot hit the cached data in KVStore for Memcache. Therefore, KVStore for Memcache does not effectively function as the database cache. You must give full consideration to the data access requirements of the business model when selecting the database cache.

15.7 Glossary

Memcached

Memcached is a high-performance distributed caching system for memory objects. For the official introduction of Memcached, see [here](#). KVStore for Memcache is compatible with Memcached binary protocol and text protocol.

Instance ID

An instance corresponds to a user space. It is the basic unit for using KVStore for Memcache. KVStore for Memcache imposes different QPS and traffic limits on single instances of different capacity specifications. You can view the instance ID list on the console.

Connection address

The host address used for connecting to KVStore for Memcache is displayed in the form of domain names. You can query the connection address on the **Basic Information** page of the KVStore for Memcache console.

Connection password

The password used for connecting to KVStore for Memcache. You can set the password at purchase, or reset the password after purchase.

Hit rate

Hit rate=Number of successful reads by the user/Number of total reads.

15.8 Instance specifications

KVStore for Memcache uses a cluster-based architecture. The following table describes its specifications.

Specification	Code	CPU processing capability	Number of nodes	Max connections	Maximum internal network bandwidth	Description
1 GB	memcache.master.small.default	Single-core	1	10,000	10	Primary/secondary dual-node architecture
2 GB	memcache.master.mid.default	Single-core	1	10,000	16	Primary/secondary dual-node architecture
4 GB	memcache.master.stand.default	Single-core	1	10,000	24	Primary/secondary dual-node architecture
8 GB	memcache.master.large.default	Single-core	1	10,000	24	Primary/secondary dual-node architecture
16 GB	memcache.sharding.small.default	Dual-core	2	10,000	96	High-performance cluster computing architecture
32 GB	memcache.sharding.mid.default	4-core	4	40,000	192	High-performance cluster computing architecture
64 GB	memcache.sharding.large.default	8-core	8	80,000	384	High-performance cluster computing architecture
128 GB	memcache.sharding.2xlarge.default	16-core	16	160,000	768	High-performance cluster

Specification	Code	CPU processing capability	Number of nodes	Max connections	Maximum internal network bandwidth	Description
						computing architecture
256 GB	memcache. sharding. 4xlarge. default	16-core	16	160,000	768	High-performance cluster computing architecture

16 Data Management Service (DMS)

16.1 What is DMS?

Data Management Service (DMS) is a Web-based data management service used to manage relational databases (such as MySQL, SQL Server, and PostgreSQL) and OLAP databases.

DMS integrates data management and structure management services.

16.2 Benefits

Support for multiple data sources

- Relational databases: MySQL, SQL Server, and PostgreSQL
- OLAP

Data analysis

- Visual analysis of table operations, including row read, add, delete, and update operations

Efficient development

- Table structure comparing
- Automatic completion of SQL statements
- Reuse of customized SQL statements and templates
- Automatic recovery of operation environments
- Dictionary and document export

16.3 Architecture

DMS is made up of three layers: the business layer, the scheduling layer, and the connection layer. DMS processes real-time data access and schedules data-related background tasks for relational databases.

Business layer

- The DMS business layer provides online GUI-based database operations. The business layer can be extended linearly to improve the general service capabilities of DMS.
- DMS supports stateless failovers, ensuring 24/7 availability.

Scheduling layer

- The scheduling layer allows you to import and export tables and compare table structures. This layer schedules tasks by using the thread pool in two modes: real-time scheduling and background periodic scheduling.
- Real-time scheduling allows you to quickly schedule and execute tasks in the frontend. You do not need to wait for the execution result after you have submitted a task. The DMS backend automatically executes the task. After the task has been executed, you can download or view the execution result.
- Background periodic scheduling allows you to periodically obtain specified data, such as data trends. DMS collects business data in the background based on scheduled tasks, allowing you to query and analyze the collected data.

Connection layer

The connection layer is the core component for data access in DMS. It has the following features:

- Process requests from MySQL, SQL Server, and PostgreSQL databases.
- Isolate sessions and provide session persistence. You can open multiple SQL windows using DMS. The SQL window sessions are isolated from each other. Additionally, the session in each SQL window is persistent to simulate the client experience.
- Control the number of instance connections to avoid a large number of connections to a single instance.
- Have different connection release policies to improve user experience and reduce the number of connections to individual databases.

16.4 Features

Relational database management

- Data management: includes functions such as SQL windows, SQL command lines, table data management, automatic completion of SQL statement, SQL formatting, custom SQL statements, SQL templates, SQL execution plans, import, and export.
- Structure management: includes functions such as table structure comparison, and management of objects (databases, tables, views, functions, storage procedures, triggers, events, series, and synonyms).

16.5 Scenarios

16.5.1 Convenient data operations

Scenario

You need an easy-to-use and multi-functional product to create SQL statements, save frequently used SQL statements, and use these statements in your business.

Solution

- You can use DMS to open tables and edit their data similar to how you do in Excel. Even if you are not familiar with SQL, you can add, delete, modify, query, and analyze data.
- You can customize SQL statements, save frequently used SQL statements, and apply these SQL statements to databases or instances.

16.5.2 Prohibiting data export

Scenario

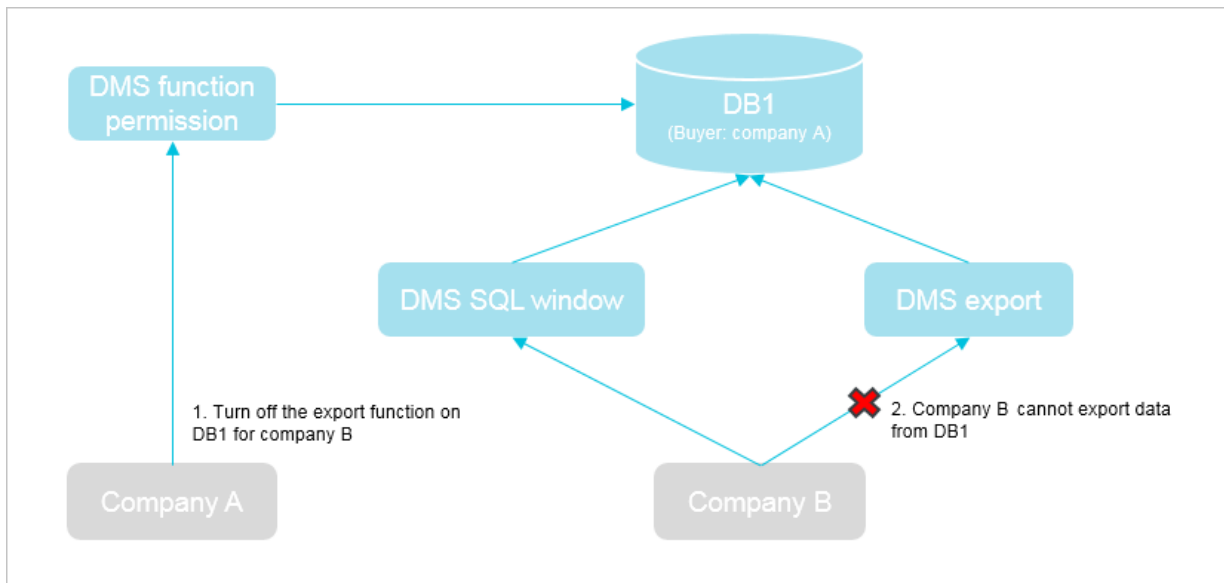
When cooperating with a partner, an enterprise manages data and its partner develops functions . The partner needs to have access to view the enterprise's data but cannot have the ability to export data to maintain data security.

Solution

Enterprise users can log on to the DMS console to grant their partners access permissions on the corresponding database instances, disabling data exporting to protect their data.

Partners who have been granted permission are only able to query and view data, eliminating the risks of data leaks.

Figure 16-1: Function-based authorization shows how to use the function-based authorization feature to prohibit partners from exporting data.

Figure 16-1: Function-based authorization

16.5.3 SQL statement reuse

Scenario

SQL statements are used when you access a database. While simple SQL queries are easy to use, rewriting SQL queries for complex data analysis or SQL queries that contain service logic is time-consuming. Even if you save these SQL queries to files, you have to maintain the files and you cannot use them without access to the files.

Solution

You can use the **My SQL** function provided by DMS to save frequently used SQL statements. As the SQL statements are not saved locally, they can be reused in any databases or instances.

16.6 Limits

Relational databases

Table 16-1: Support for relational databases

Module	Function	MySQL	SQL Server	PostgreSQL
Data management	Table data management	√	√	√
	SQL windows	√	√	√
	SQL command lines	√		√
	SQL templates	√		

Module	Function	MySQL	SQL Server	PostgreSQL
	SQL formatting	√	√	√
	Custom SQL statements	√	√	
	Automatic completion of SQL statements	√	√	
	SQL execution plans	√	√	√
Structure management	Library management	√	√	√
	Table management	√	√	√
	Management of objects such as indexes, views, storage processes, functions, triggers, and events	√	√	√
	Entity relationship diagram display	√	√	
	Data dictionaries	√		
Import and export	Basic import and export functions	√	√	√
	Export of large volumes of data	√	√	√

17 Server Load Balancer (SLB)

17.1 What is Server Load Balancer?

Server Load Balancer (SLB) is a traffic distribution control service that distributes the incoming traffic among multiple ECS instances according to the configured forwarding rules. SLB expands the service capabilities of the application and enhances application availability.

By setting a virtual service address, SLB virtualizes the added ECS instances into an application service pool with high-performance and high availability, and distributes client requests to ECS instances in the server pool based on forwarding rules.

SLB also checks the health status of added backend servers, and automatically isolates abnormal ECS instances to eliminate single point of failure (SPOF), thus improving the overall service capability of your application. Additionally, working with Alibaba Anti-DDoS, SLB is able to defend DDoS attacks.

Components

Server Load Balancer consists of the following components:

- SLB instances

An SLB instance is a running load balancing service that distributes incoming traffic to backend servers. To use the load balancing service, you must create an SLB instance, and then add at least one listener and two backend servers to the instance.

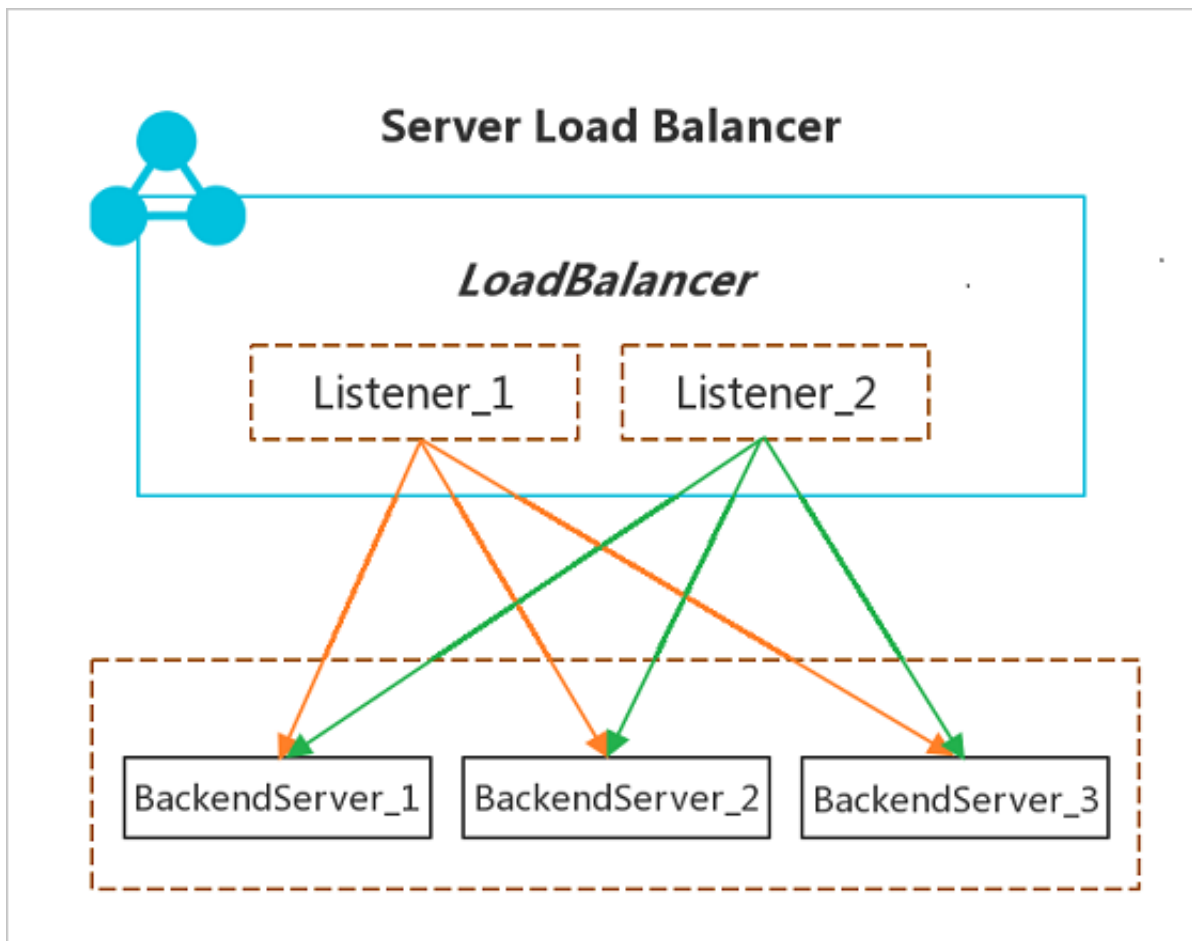
- Listeners

A listener checks client requests and forwards the requests to backend servers according to the configured rules. It also performs health check on backend servers.

- Backend servers

Backend servers are the ECS instances added to an SLB instance to receive and process distributed requests. You can classify ECS instances running different applications or playing different roles by creating server groups.

As shown in the following figure, once a request arrives at an SLB instance, Server Load Balancer will distribute the request to the corresponding backend server according to the listener configurations.



17.2 Benefits

High availability

Server Load Balancer is designed to work in the full-redundancy mode without SPOF. Server Load Balancer supports local disaster tolerance and can flexibly scale its service based on the application load without interrupting external services during traffic fluctuation.

Low cost

Server Load Balancer is more cost-efficient than traditional hardware load-balancing systems without generating any O&M cost.

Security

Combined with Alibaba Cloud Security, Server Load Balancer can defend against up to 5 Gbps DDoS attacks, such as HTTP flood and SYN flood attacks.

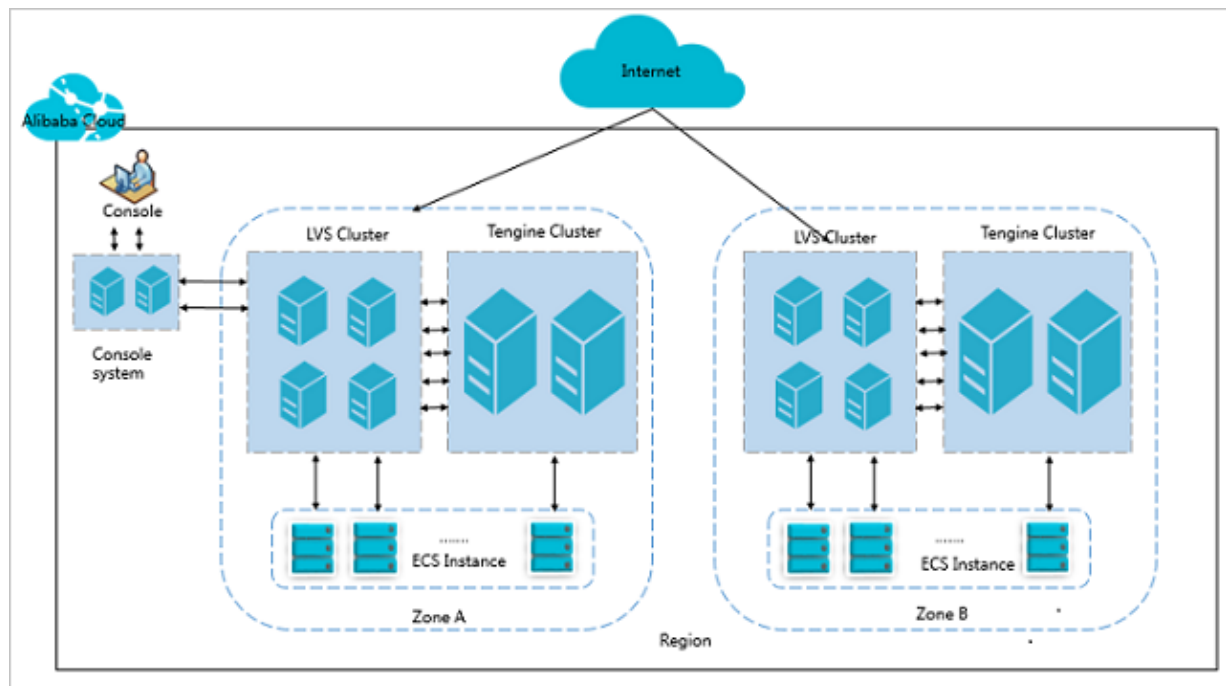
17.3 Architecture

Server Load Balancer is deployed in clusters. The cluster deployment model achieves session synchronization, eliminates SPOF, improves redundancy and increases service stability.

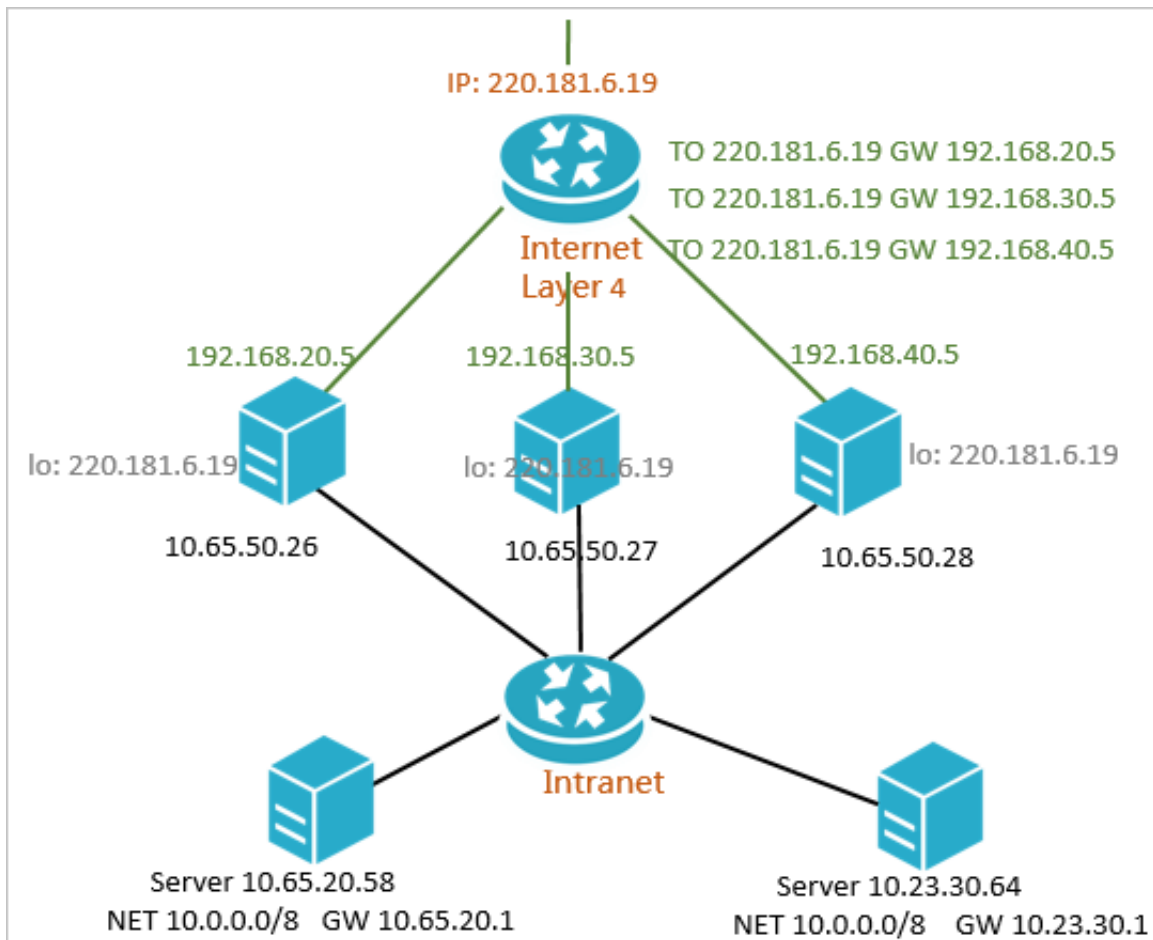
Apsara Stack provides the Layer-4 (TCP protocol and UDP protocol) and Layer-7 (HTTP protocol and HTTPS protocol) load balancing services.

- Layer 4 uses the open source software Linux Virtual Server (LVS) with keepalived to achieve load balancing, and also makes some customization to it according to cloud computing requirements.
- Layer 7 uses Tengine to achieve load balancing. Tengine is a Web server project launched by Taobao. Based on Nginx, it adds a wide range of advanced features dedicated for high-traffic websites and also adds the [CreateLoadBalancer](#) feature.

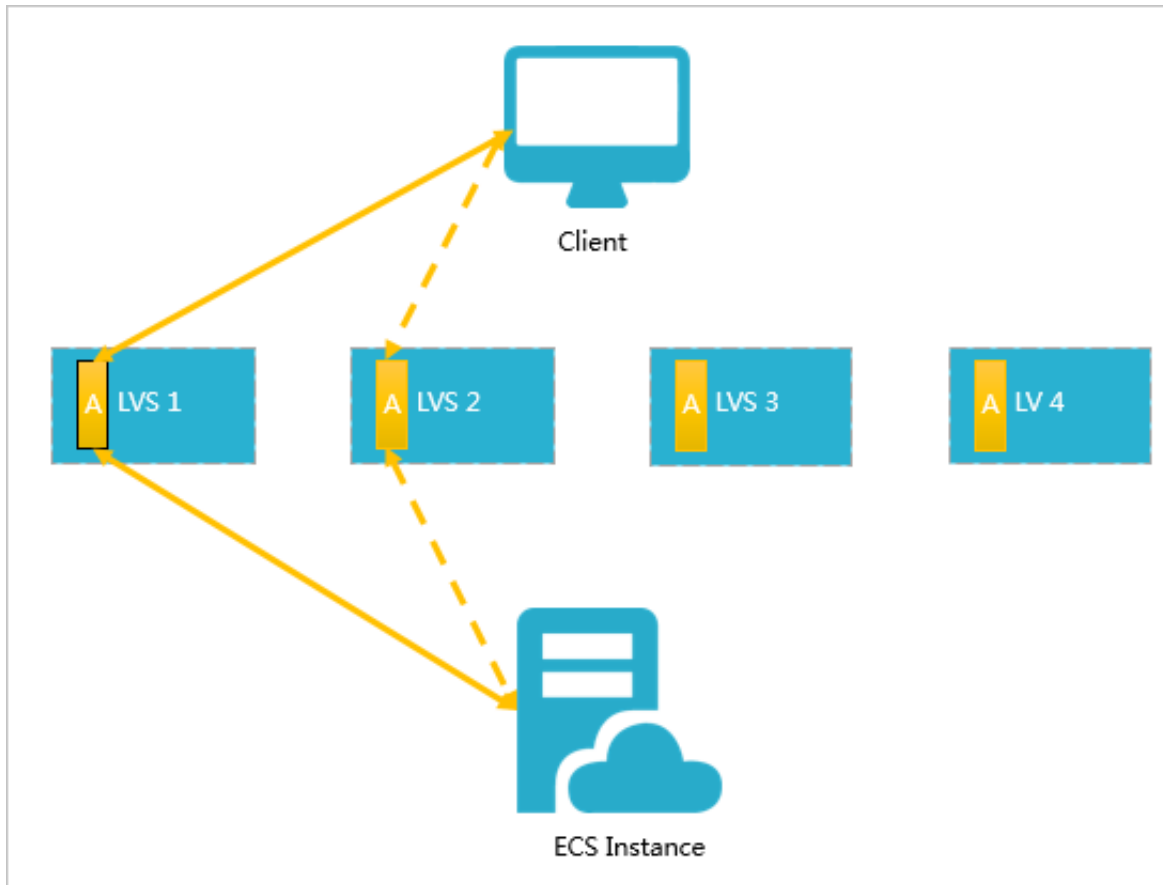
Figure 17-1: SLB architecture



As shown in the following figure, Layer-4 Server Load Balancer actually runs in a cluster of LVS machines. The cluster deployment model strengthens the availability, stability, and scalability of the load balancing services in abnormal circumstances.

Figure 17-2: Cluster deployment

Additionally, each LVS machine in the LVS cluster uses multicast packets to synchronize sessions to other LVS machines. As shown in the following figure, session A established on LVS1 is synchronized to other LVS machines after three packets are transferred. In normal situations, a session request is sent to LVS1 as the solid line shows. If LVS1 is abnormal or being maintained, the session request will be sent to LVS2 that is working normally, as the dotted line shows. Therefore, SLB clusters support hot update, and machine failure or cluster maintenance will not affect your service.

Figure 17-3: Session persistence

17.4 Features

Layer-4 and layer-7 load balancing

Alibaba Cloud provides layer-4 (TCP and UDP) and layer-7 (HTTP and HTTPS) load balancing services. You can create different listeners to balance loads of different applications. For example , create an HTTP listener to balance loads of HTTP applications.

Health check

Server Load Balancer monitors the health of added backend servers. When a backend server is declared as unhealthy, Server Load Balancer will stop forwarding requests to it and forward the requests to other healthy backend servers.

Session persistence

Server Load Balancer supports session persistence. With session persistence enabled, Server Load Balancer can forward requests from the same client to the same backend server.

Scheduling algorithms

Server Load Balancer supports the following scheduling algorithms:

- Round robin: Requests are distributed across backend servers sequentially.
- Least connections: A backend server with less connections receives more requests.

Access control

You can set a whitelist to control which IP addresses can access the load balancing service.

Certificate management

Server Load Balancer supports HTTPS load balancing service and provides the certificate management function. You do not need to upload certificates to backend servers. Deciphering is performed on Server Load Balancer to reduce the CPU usage of backend servers.

Virtual server group

A virtual server group consists of a group of ECS instances. You can add ECS instances that run different applications or provide different functions to different virtual server groups, and then you can create different listeners for different applications to forward requests to specified server groups.

17.5 Scenarios

Server Load Balancer is applicable to the following scenarios:

Load balance your applications

Server Load Balancer can automatically distribute incoming traffic across multiple backend servers (ECS instances). Additionally, the requests from the same client can be distributed to the same backend server by configuring session persistence.

Scale your applications

To meet the demand of your customers, you can increase the number of backend servers at any time to scale your applications. SLB applies to web servers and app servers.

Protect your applications from single point of failures

You can add multiple backend servers to an SLB instance. If some of the backend servers are unhealthy, Server Load Balancer will stop distributing incoming traffic to them and distribute the traffic to the healthy ones. Once the backend servers become healthy, Server Load Balancer will automatically resume distributing traffic to them.

17.6 Limits

- For the Layer-4 SLB service, a backend ECS instance cannot act both as a real server and a client that sends requests to an SLB instance. The returned packets can be transmitted only inside the ECS instance and cannot pass through SLB, so the SLB instance cannot be accessed from backend ECS instances.
- Before using Server Load Balancer to provide service, make sure you have configured the application on the backend ECS instances of the SLB instance and the service can be accessed through the service addresses of the ECS instances.
- SLB cannot synchronize data among ECS instances. If the application deployed on the backend ECS instances of the SLB instance is stateless, you can store data through independent ECS instances or RDS service; if the application deployed on the backend ECS instances of the SLB instance is stateful, make sure data on these ECS instances is synchronized.
- When the service address of an SLB instance is subjected to domain name resolution, do not randomly delete the SLB instance while providing external service. The service address will be released and the service will be interrupted if you delete the SLB instance.

17.7 Terms

Server Load Balancer

Server Load Balancer instance

A Server Load Balancer instance is a running entity of the Server Load Balancer service. To use Server Load Balancer, you must first create a Server Load Balancer instance.

Server Load Balancer IP

An IP address allocated to an SLB instance. According to the instance type, the IP address is either a public IP or a private IP.

Listener

A listener defines how to forward incoming requests to backend servers. A listener includes the listening port, the forwarding policy, health check configurations and more. Each listener corresponds to a backend application.

Backend server

The ECS instances that are added to an SLB instance to receive the distributed requests. The SLB service forwards requests to added backend ECS instances according to configured rules.

18 Virtual Private Cloud (VPC)

18.1 What is VPC?

A Virtual Private Cloud (VPC) is a private network established in Apsara Stack. VPCs are logically isolated from each other.

You have full control over your VPC. For example, you can select its IP address range and configure routing tables and gateways. You can also use Alibaba Cloud resources such as ECS, RDS, and SLB in your own VPCs. You can connect a VPC to other VPCs or a local network to form an on-demand customizable network environment. This allows you to smoothly migrate applications to the cloud.

Components

Each VPC consists of a private Classless Inter-Domain Routing (CIDR) block, a VRouter, and at least a VSwitch.

- CIDR block

A CIDR block is a private IP address range in a VPC. The IP addresses of all cloud resources deployed in the VPC are within the specified CIDR block. When creating a VPC or a VSwitch, you must specify the private IP address range in the form of a CIDR block.

You can use any of the following standard CIDR blocks and their subnets as the IP address range of the VPC.

CIDR block	Number of available private IP addresses (system reserved ones excluded)
192.168.0.0/16	65,532
172.16.0.0/12	1,048,572
10.0.0.0/8	16,777,212

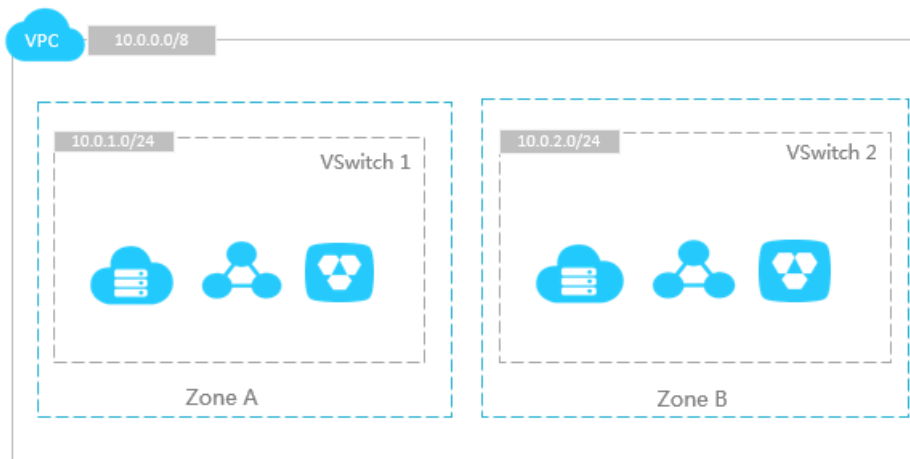
- VRouter

The VRouter is the hub of a VPC. As an important component of a VPC, the VRouter connects the VSwitches in a VPC and serves as the gateway connecting the VPC with other networks. After you create a VPC, the system automatically creates a VRouter, which is associated with a routing table.

- VSwitch

A VSwitch is a basic network device in a VPC and is used to connect different cloud product instances. After creating a VPC, you can further divide the VPC to one or more subnets by creating VSwitches. The VSwitches within a VPC are interconnected. You can deploy applications in VSwitches of different zones to improve the service availability.

Figure 18-1: VPC



18.2 Benefits

VPC features high security and flexible configurations, and supports multiple connection methods.

Secure

Each VPC is identified by a unique tunnel ID. VPCs are completely isolated from one another. You can use security groups or whitelists to control access to cloud resources in the VPC.

Easy to use

You can create and manage a VPC in the VPC console. After a VPC is created, the system automatically creates a VRouter and a routing table for it.

Scalable

You can create multiple subnets in a VPC to deploy different services. Additionally, you can connect a VPC to a local data center or other VPCs to extend the network architecture.

18.3 Architecture

VPCs are isolated virtual networks achieved by using tunneling technology. Each VPC is identified by a unique tunnel ID.

Background information

The continuous development of cloud computing technologies leads to increasing virtual network requirements such as scalability, security, reliability, privacy, and performance. This scenario has hastened the birth of a variety of network virtualization technologies.

Earlier solutions combined virtual and physical networks to form a flat network architecture, such as large layer-2 networks. As the scale of virtual networks grew, earlier solutions faced more serious problems. A few notable problems include ARP spoofing, broadcast storms, and host scanning. Various network isolation technologies emerged to resolve these problems by completely isolating the physical networks from the virtual networks. One of the technologies utilized VLAN to isolate users, but due to VLAN limitations, it could only support up to 4096 nodes. It is insufficient to support the huge amount of users in the cloud.

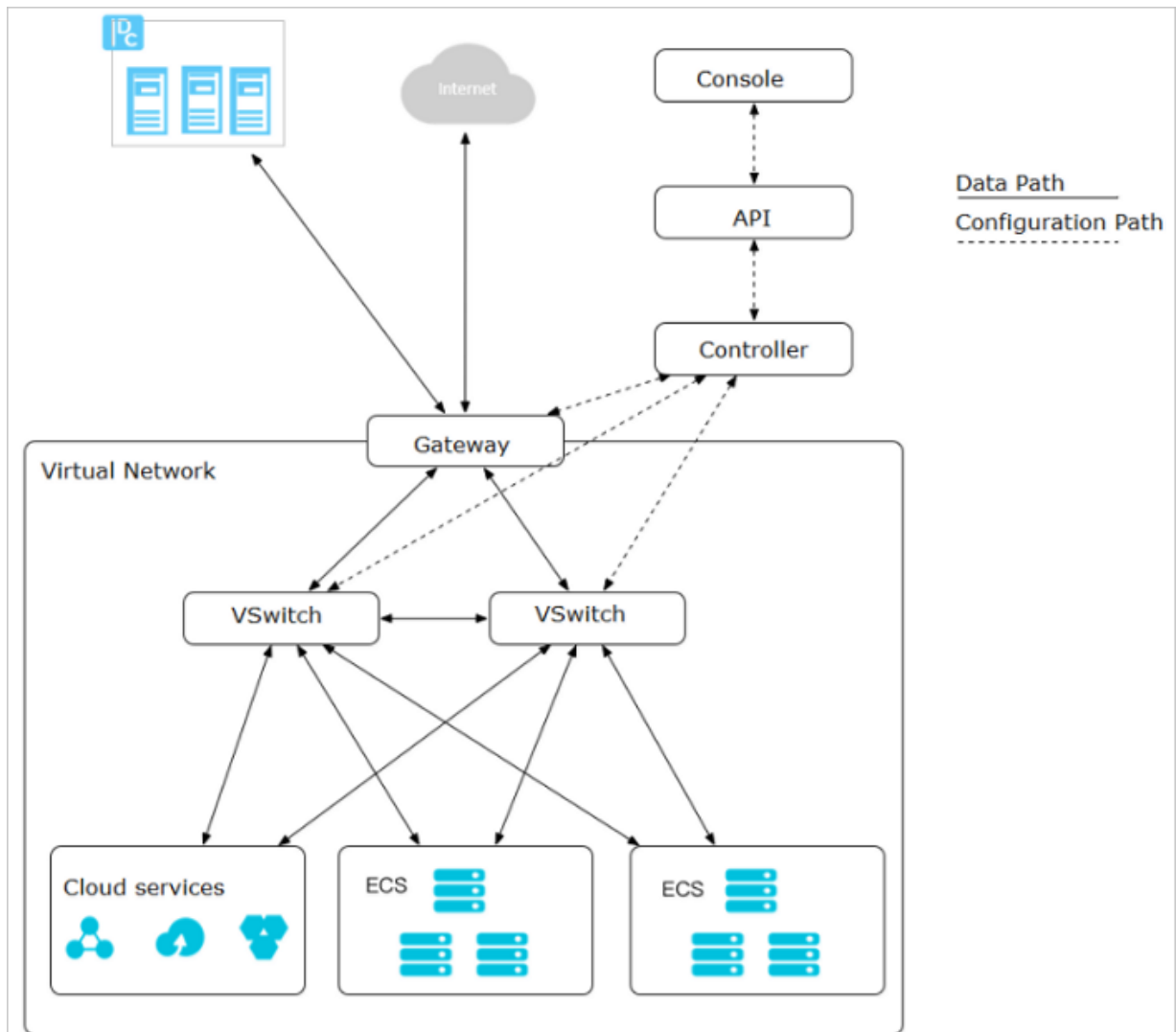
VPC basis

VPCs are isolated virtual networks achieved by using tunneling technology. Each VPC is identified by a unique tunnel ID. A unique tunnel ID is generated when tunnel encapsulation is performed on each data packet transmitted between the ECS instances within a VPC. Then, the data packet is transmitted over the physical network. ECS instances in different VPCs cannot communicate with each other. They have different tunnel IDs and therefore are on different routing planes.

Alibaba Cloud developed technologies such as VSwitch, Software Defined Network (SDN), and hardware gateway based on tunneling technology. These technologies serve as the basis for VPC.

Logical architecture

As shown in the following figure, the VPC architecture contains three main components: VSwitches, gateway, and controller. VSwitches and gateways form the key data path. Controllers use the protocol developed by Alibaba Cloud to forward the forwarding table to the gateway and VSwitches, completing the key configuration path. In the overall architecture, the configuration path and data path are separated from each other. VSwitches are distributed nodes, the gateway and controller are deployed in clusters, backup for disaster recovery is supported by multiple data centers, and all links have redundant disaster recovery. This improves the overall availability of the VPC.

Figure 18-2: VPC architecture

18.4 Features

Custom VPCs

You can customize VPCs. When you create VPCs or VSwitches, you can specify private IP addresses for them. You can divide a VPC into multiple subnets and deploy services in different subnets to improve the service availability.

Custom routes

You can add custom routes to the VPC routing table to forward traffic to the specified next hop. The routing table uses the longest prefix matching rule for traffic routing. The routing entry with the longest subnet mask will be used because it is the most specific route.

18.5 Scenarios

VPCs are applicable to scenarios that require high communication security and service availability.

Host applications

Applications that provide external services can be hosted within a VPC. Security group rules and a whitelist can be created to control Internet access. You can also isolate application servers from databases to implement access control. For example, you can deploy Web servers in a subnet that can access the Internet, and deploy its application databases in a subnet that cannot access the Internet.

Host applications that require Internet access

An application that requires Internet access can be hosted in a subnet of a VPC. The traffic is then routed through NAT. You can configure SNAT rules so that instances in a subnet can access the Internet without exposing their private IP addresses. SNAT can replace the private IP addresses with public IP addresses to protect the instances against external attacks.

Zone-disaster recovery

You can divide a VPC into one or multiple subnets by creating VSwitches. Different VSwitches within the same VPC can communicate with each other. Resources can be deployed to VSwitches of different zones to achieve zone-disaster recovery.

Isolate business systems

VPCs are logically isolated from each other. To isolate multiple business systems, such as the production and test environments, you can create a VPC for each environment. When the VPCs need to communicate with each other, you can create a peer connection between them.

Extend the local network architecture

To extend the local network architecture, you can connect the on-premises data center to a VPC. You can also seamlessly migrate local applications to the cloud without changing the application access method.

18.6 Limits

VPC

Resource	Default limit
Maximum number of VRouters in a VPC	1

Resource	Default limit
Maximum number of routing tables in a VPC	1
Maximum number of VSwitches in a VPC	24
Maximum number of routing entries in a routing table	48

VRouter and VSwitch

Resource	Default limit
VRouter	<ul style="list-style-type: none"> Each VPC can have only one VRouter. Each VRouter can have only one routing table. Dynamic routing protocols such as BGP and OSPF are not supported.
VSwitch	<ul style="list-style-type: none"> Layer-2 broadcasting and multicasting are not supported.

18.7 Terms

Virtual Private Cloud (VPC)

A private network established in Apsara Stack. VPCs are logically isolated from each other. You can create and manage cloud service instances in your VPC, such as ECS instances, SLB instances, and RDS instances.

VRouter

A hub in a VPC. It connects all VSwitches in the VPC and serves as a gateway that connects the VPC to other networks. A VRouter routes the network traffic to their destinations based on the configured routing entries.

VSwitch

A basic network device of a VPC. It is used to connect different cloud service instances. When creating a cloud service instance in a VPC, you must specify the VSwitch that is used by the instance.

Routing table

A list of routing entries in a VRouter.

Routing entry

An entry in a routing table. A routing entry specifies the next hop address for the network traffic destined to a CIDR block. There are two types of entries, system routing entry and custom routing entry.

19 Log Service

19.1 What is Log Service?

As a one-stop service for log data, Log Service (Log for short) provides you with multiple functions such as log data collection, query, analysis, and consumption.

Log Service has been honed by countless big data scenarios at Alibaba Group. Without any development, you can quickly collect, consume, query, and analyze log data by using Log Service. It helps increase the O&M efficiency and build capabilities to process large-volume logs in this data technology (DT) era.

Log Service provides you with the following functions:

- **Log collection:** Supports collecting multiple formats of log data such as Event, Binlog, and TextLog in real time through the Logtail client, JS, and other methods.
- **Query and analysis:** Provides real-time query and analysis for collected log data and supports generating visual charts and dashboards based on analysis results.
- **Status alarm:** Supports executing query and analysis statements regularly based on the query and analysis function. When query results meet alarm conditions, real-time alarms are reported based on configured alarm tasks.
- **Real-time consumption:** Provides real-time consumption interfaces for log data collected to the server.

19.2 Benefits

Fully managed service

- The service is easy to use. You can access the service for usage in five minutes.
- LogHub has all the functions of Kafka, and provides complete functional data, such as monitoring and alarms. It also supports automatic scaling (by PB/day), saving costs of more than 50% compared to self-built systems.
- LogSearch/Analytics provides query saving, dashboard, and alarm functions, saving costs of more than 80% compared to self-built systems.
- It has more than 30 access methods, can seamlessly interwork with open-source software (such as Storm and Spark).

Rich ecosystem

- LogHub supports over 30 log data sources and can be easily connected using embedded devices, webpages, servers, and programs. It can also interconnect with consumption systems such as Spark Streaming and Storm.
- LogSearch/Analytics have complete syntaxes and are compatible with SQL-92, supporting interconnection with Grafana by using JDBC protocol.

Strong real-timelines

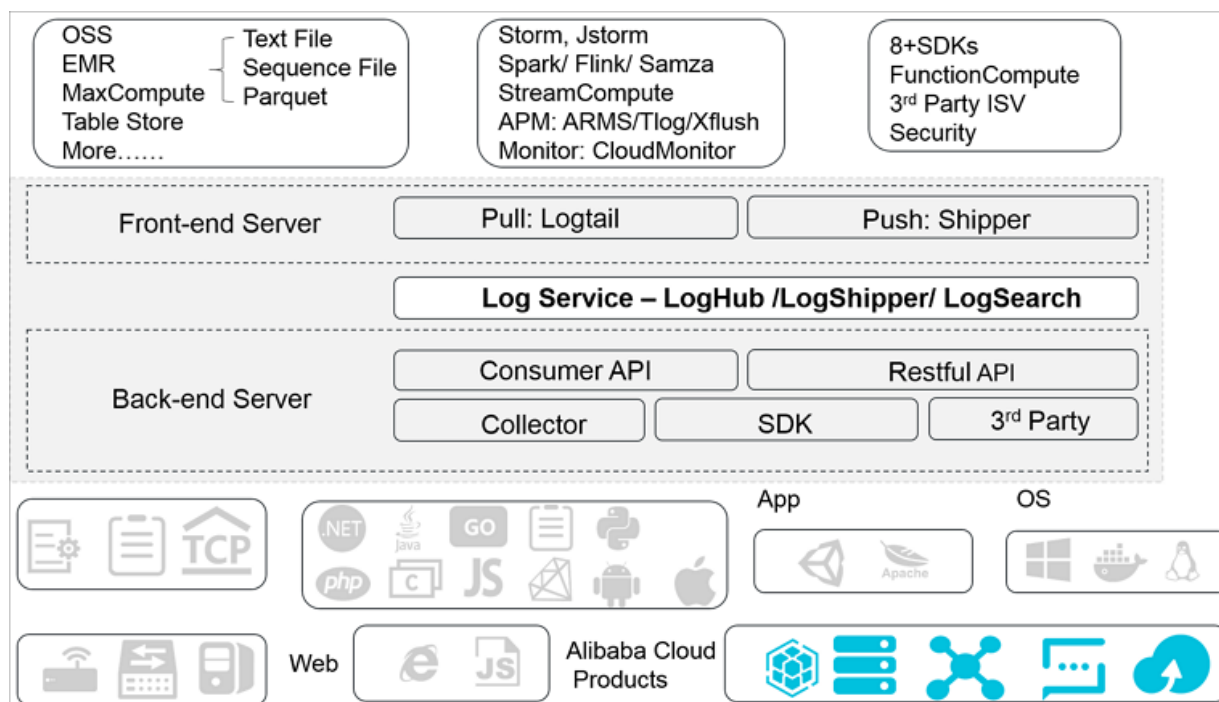
- LogHub: Data can be used immediately after it is written. Logtail (the data collection agent) collects and transfers data in real time.
- LogSearch/Analytics: Data can be searched and analyzed immediately after it is written. If multiple search conditions are used, billions of data pieces can be searched within one second. When multiple aggregation conditions are used, hundreds of millions of data pieces can be analyzed within one second.

Complete APIs and SDKs

- Log Service supports user-defined management and secondary development.
- All functions can be implemented using APIs and SDKs. SDKs for multiple languages are provided. Services and millions of devices can be managed in an easy way.
- The query and analysis syntax is simple (compatible with SQL-92). The interfaces can be used to interconnect with ecological software.

19.3 Architecture

The Log Service architecture is shown in the following figure.

Figure 19-1: Log Service architecture

Logtail

Logtail is an agent that helps you quickly collect logs and has the following features:

- Non-invasive log collection based on log files
 - Only reading of files.
 - Non-invasive during the reading process.
- Security and reliability
 - Supports file rotation to prevent loss of data.
 - Supports local caching.
 - Provides a network exception retry mechanism.
- Convenient management
 - Management on Web.
 - Visualization configuration.
- Comprehensive self-protection
 - Monitors the CPU and memory consumed by the process in real time.
 - Restricts the upper limit of CPU/memory usage.

Frontend servers

Frontend servers are the frontend machines built with LVS and Nginx and have the following features:

- HTTP and REST protocols
- Horizontal scaling
 - Frontend servers allow horizontal scaling when traffic increases.
 - Frontend servers can be quickly added to improve processing capabilities.
- High throughput and low latency
 - Pure asynchronous processing. A single request exception does not affect other requests.
 - Adopts the Lz4 compression, which is specially for logs, to increase the processing capabilities of individual machines and reduce network bandwidth.

Backend servers

The backend service is a distributed process deployed on multiple machines. It provides real-time Logstore data persistence, indexing, and query. Features of the backend service are as follows:

- High data security
 - Each log you write is saved in triplicate.
 - Data are automatically recovered if damage to disks or machine downtime occurs.
- Stable service
 - Logstores automatically migrate in case of a process crash or machine downtime.
 - Automatic load balancing makes sure that traffic is distributed evenly among different machines.
 - Strict quota restrictions help prevent abnormal behavior of a single user from affecting other users.
- Horizontal scaling
 - Horizontal scaling is performed by using shards as the unit.
 - You can dynamically add shards as needed to increase throughput.

19.4 Key features

19.4.1 Core features

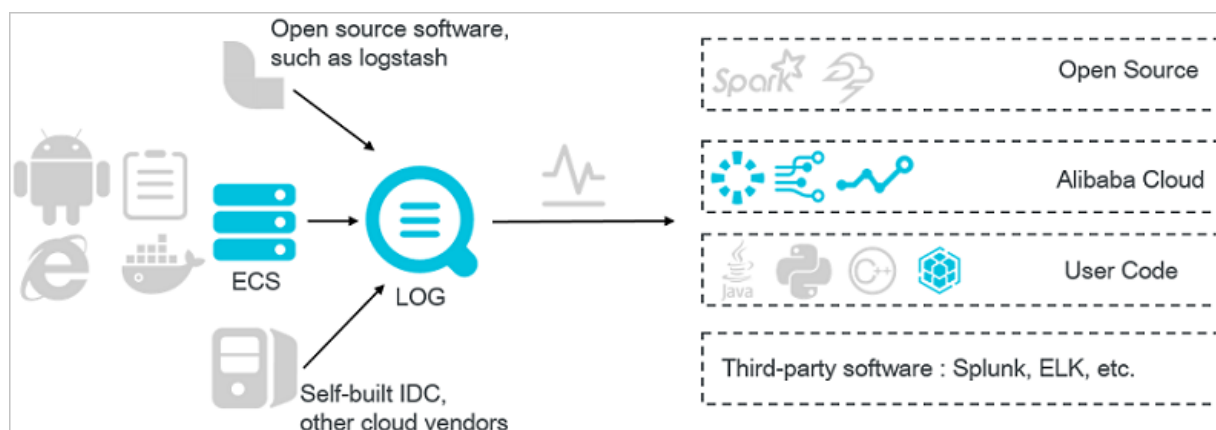
Real-time log collection and consumption (LogHub)

LogHub supports multiple log lossless collection methods such as clients, webpages, protocols, SDKs and APIs (for mobile terminals and gaming), and consumption ways of SDK, Storm Spout, Spark Client, and more. By supporting multiple formats of real-time log collection and consumption, LogHub helps you streamline the processing of multi-device and multi-source log collection and consumption.

Log Service has the following core functions.

- Use Elastic Compute Service (ECS), containers, mobile terminals, open-source software, and JS to access real-time log data (such as Metric, Event, BinLog, TextLog, and Click data).
- A real-time consumption interface is provided to interconnect with real-time computing and service.

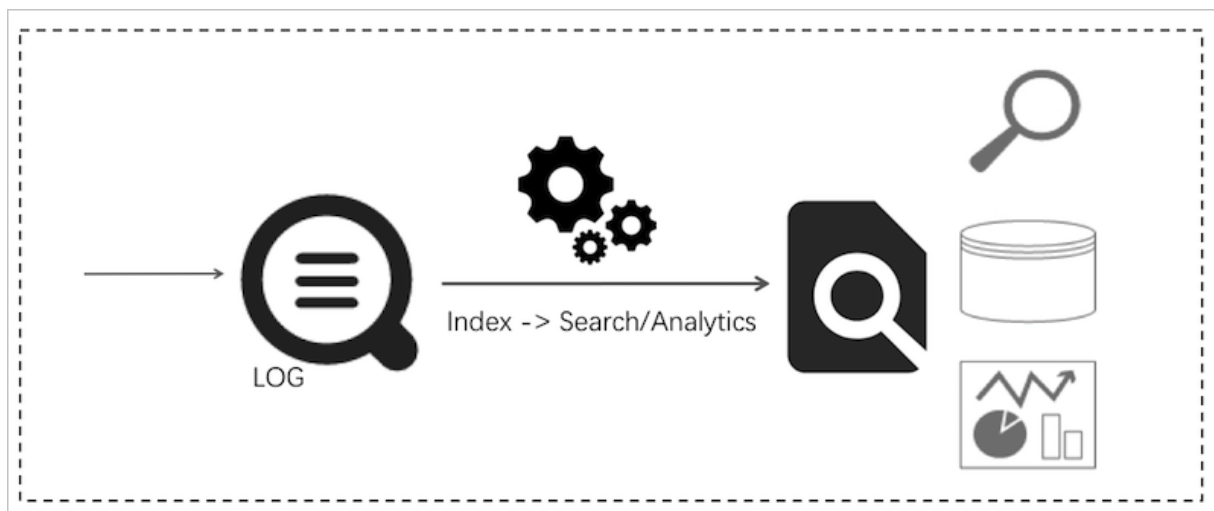
Figure 19-2: Real-time log collection and consumption



Query and real-time analysis (Search/Analytics)

It can index, query, and analyze log data collected to the server in real time and generate dynamic data reports based on query and analysis results. Visualization analysis of log data in multi-scenarios is supported.

- Query: keyword, fuzzy match, context, and range.
- Statistics: Rich query methods such as SQL aggregation.
- Visualization: Dashboard and report functions.
- Interconnection: Grafana, JDBC, and SQL92.

Figure 19-3: Query and real-time analysis

19.4.2 Other features

19.4.2.1 Log

Logs in Log Service

A log is an abstraction of system changes during the running process. The content is the time-ordered collection of some operations and operation results of specified objects. LogFile, Event, BinLog and Metric data are different carriers of logs. In LogFile, every log file is composed of one or more logs, and every log describes a single system event. A log is the minimum data unit processed in Log Service.

Log Service uses a semi-structured data model to define a log. This model is composed of four data fields: Topic, Time, Content, and Source.

Furthermore, Log Service has different format requirements for different fields, as described in the following table:

Data field	Description	Format
Topic	This is a custom field to mark a batch of logs. For example, access logs can be marked according to sites.	Any string up to 128 bytes in length, including null strings. This field is a null string by default.
Time	This is a reserved field in the log and is used to indicate the generation time of the log. It is	It must be an integer in standard UNIX time format. The unit is in seconds. This field indicates the

Data field	Description	Format
	typically generated directly based on the time in the log.	number of seconds from 00:00:00 Thursday, 1 January 1970 UTC.
Content	This field is used to record the specific content of the log. Content is composed of one or more content items, and each content item is a Key-Value pair.	A key is a UTF-8 encoded string up to 128 bytes in length, and can contain letters, numbers, and underscores (_). It cannot start with a number. The following keywords cannot be used in the key: __time__, __source__, __topic__, __partition_time__, __extract_others__, and __extract_others__. The value can be any string up to 1024 x 1024 bytes.
Source	This field indicates the source of the log. For example, the IP address of the machine where the log is generated.	Any string up to 128 bytes in length. This field is null by default.

Various log formats are used in actual application scenarios. For better understanding, the following example describes how to map an original Nginx access log to the Log Service log data model. Assume that the IP address of your Nginx server is 10.249.201.117. The following is an original log of this server.

```
10.1.168.193 - - [01/Mar/2012:16:12:07 +0800] "GET /Send?AccessKeyId=8225105404 HTTP/1.1" 200 5 "-" "Mozilla/5.0 (X11; Linux i686 on x86_64; rv:10.0.2) Gecko/20100101 Firefox/10.0.2"
```

Map the original log to the Log Service log data model as follows.

Data Field	Content	Description
Topic	""	Use the default value (null string).
Time	1330589527	Precise generation time of the log, indicating the number of seconds from 00:00:00 Thursday, 1 January 1970 UTC. This time is converted from the timestamp in the original log.

Data Field	Content	Description
Content	Key-Value pair	The content of the log.
Source	"10.249.201.117"	Use the IP address of the server as the log source.

You can then decide how to extract the original content of the log and combine them into Key-Value pairs. The following table is an example.

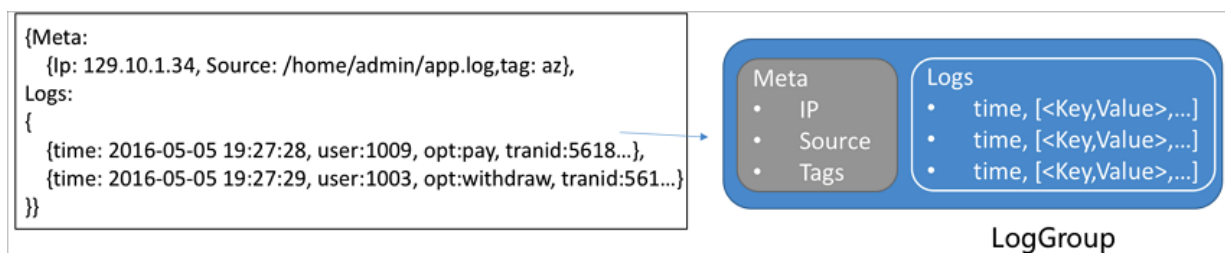
Key	Value
ip	"10.1.168.193"
method	"GET"
status	"200"
length	"5"
ref_url	"_"
browser	"Mozilla/5.0 (X11; Linux i686 on x86_64; rv:10.0.2) Gecko/20100101 Firefox/10.0.2"

Log group

A log group is a collection of logs and is the basic unit for writing and reading.

The maximum capacity of a log group is 4096 logs or 10 MB.

Figure 19-4: Log group



19.4.2.2 Project

A project is the resource management unit of Log Service. Projects are used to isolate and control resources. You can use a project to manage logs and related log sources of one application. A project is used to manage Logstores of a user and the machine configurations for log collection. It also serves as the portal for users to access the Log Service resources.

Projects provide following functions:

- Help you organize and manage different Logstores. You can use Log Service to centrally collect and store the logs of different projects, products, or environments. You can classify different logs for management in different projects to facilitate subsequent log consumption, exporting, or indexing. In addition, projects are the carriers of the log access permission management.
- Provide you with a portal to access Log Service resources. Log Service allocates a unique access portal to each created project. This access portal supports writing, reading, and managing logs through the network.

19.4.2.3 Logstore

Logstores are the units used in Log Service for log data collection, storage, and query. Each Logstore belongs to one project, and multiple Logstores can be created for a single project. You can create multiple Logstores for one project as needed. Typically, an independent Logstore is created for each type of log in one application. For example, assume that you have the game application "big-game", and three types of logs are on the server: operation_log, application_log, and access_log. You can first create a project named "big-game", and then create three Logstores for the three types of logs under this project.

Whether writing or querying logs, you must specify a Logstore for the operation. If you want to ship the log data to MaxCompute for offline analysis, the data will be shipped in Logstore units for data synchronization (that is, the data in a single Logstore is shipped to a single MaxCompute table).

Logstores provide the following functions:

- Log collection, supporting real-time logging
- Log storage, supporting real-time consumption
- Index creation, supporting real-time log query

19.4.2.4 Shard

Logstore read/write logs must be saved in a shard. Each Logstore is divided into several shards and each shard is composed of MD5 left-closed, right-open intervals. Each interval range does not overlap with others and the total range of all the intervals is the entire MD5 value range.

Range

When a Logstore is being created, the entire MD5 range is automatically divided evenly based on the specified number of shards. Each shard has a certain range within the following value range: [00000000000000000000000000000000,ffffffffffffffffffffffffffffffff).

Each shard is composed of the following two keys:

- **BeginKey:** Indicates the start of the shard. This key is included in the shard range.
- **EndKey:** Indicates the end of the shard. This key is excluded from the shard range.

With the shard range, you can write logs by specifying the Hash Key, as well as splitting or merging shards. To read data from a shard, you must specify the corresponding shard. To write data to a shard, you can use Server Load Balancer or specify the Hash Key. By using Server Load Balancer, each data packet is written to an available shard randomly. By specifying the Hash Key, data is written to the shard whose range includes the specified key.

For example, a Logstore has four shards and the MD5 value range of this Logstore is [00,FF).

Each shard range is as follows:

Shard No.	Range
Shard0	[00,40)
Shard1	[40,80)
Shard2	[80,C0)
Shard3	[C0,FF)

If you specify the MD5 key as 5F by specifying the Hash Key when writing logs, the log data is written to Shard1 that contains the MD5 key 5F. If you specify the MD5 key as 8C, the log data is written to Shard2 that contains the MD5 key 8C.

Shard read/write capacities

The service capacities of a shard are:

- Write: 5 MBit/s, 2000 times/s
- Read: 10 MBit/s, 100 times/s

We recommend that you plan the number of shards based on the actual data traffic. If the traffic exceeds the read/write capacities, split the shard in a timely manner to increase the number of shards to achieve greater read/write capacities. If the traffic is far less than the maximum read/write capacities of shards, we recommend that you merge the shards to reduce the number of shards to save the rental costs of shards.

For example, assume that you have two shards in read/write status and can write data at 10 MBit/s at maximum. If you write data at 14 MBit/s in real time, we recommend that you split a shard to make the number of shards in read/write status reach three. If you write data at only 3 MBit/s in

real time, we recommend that you merge these two shards because one shard can meet the need

**Note:**

- During log writing, if the API consistently reports a 403 or 500 error, refer to Logstore CloudMonitor metrics to view the traffic and status code and determine whether you need to increase the number of shards.
- For read/write operations that exceed the service capacities of shards, the system attempts to provide the needed services, but the service quality cannot be guaranteed.

Status

The shard status includes:

- read/write: Supports reading and writing data.
- readonly: Only supports reading data.

When a shard is created, all the shards are in read/write status. Split or merge operations change the shard status to readonly and generate a new shard in read/write status. The shard status does not affect the performance of reading data. Shards in read/write status maintain normal data writing performance, while shards in readonly status do not support writing data.

When splitting a shard, you must specify a ShardId in read/write status and an MD5. The MD5 must be greater than the shard BeginKey and smaller than the shard EndKey. Split operations can split two other shards from one, that is, the number of shards is increased by 2 after the split. After the split, the status of the original shard specified to be split is changed from read/write to readonly. Data can still be consumed, while new data cannot be written. The two newly generated shards are in read/write status and arranged behind the original shard. The MD5 range of these two shards covers the range of the original shard.

When merging shards, you must specify a shard in read/write status. Make sure the specified shard is not the last shard in read/write status. The server automatically finds the adjacent shard at the right of the specified shard and merges these two shards. After the merge, the specified shard and the adjacent shard on the right are in readonly status. Data can still be consumed, while new data cannot be written. A new shard in read/write status is generated and its MD5 range covers the total range of the original two shards.

19.4.2.5 Log topic

Logs in a Logstore can be classified by log topics. You can specify the topic when writing and querying logs. For example, as a platform user, you can use your user ID as the log topic when writing logs. In this way, you can choose to only view your own logs based on the log topic when querying logs. If you do not need to classify logs in a Logstore, use the same topic for all of the logs.

**Note:**

A null string is a valid log topic and is the default log topic when writing and querying logs. If you do not need to use the log topic, the easiest way is to use the default log topic, the null string, for writing and querying logs.

19.5 Scenarios

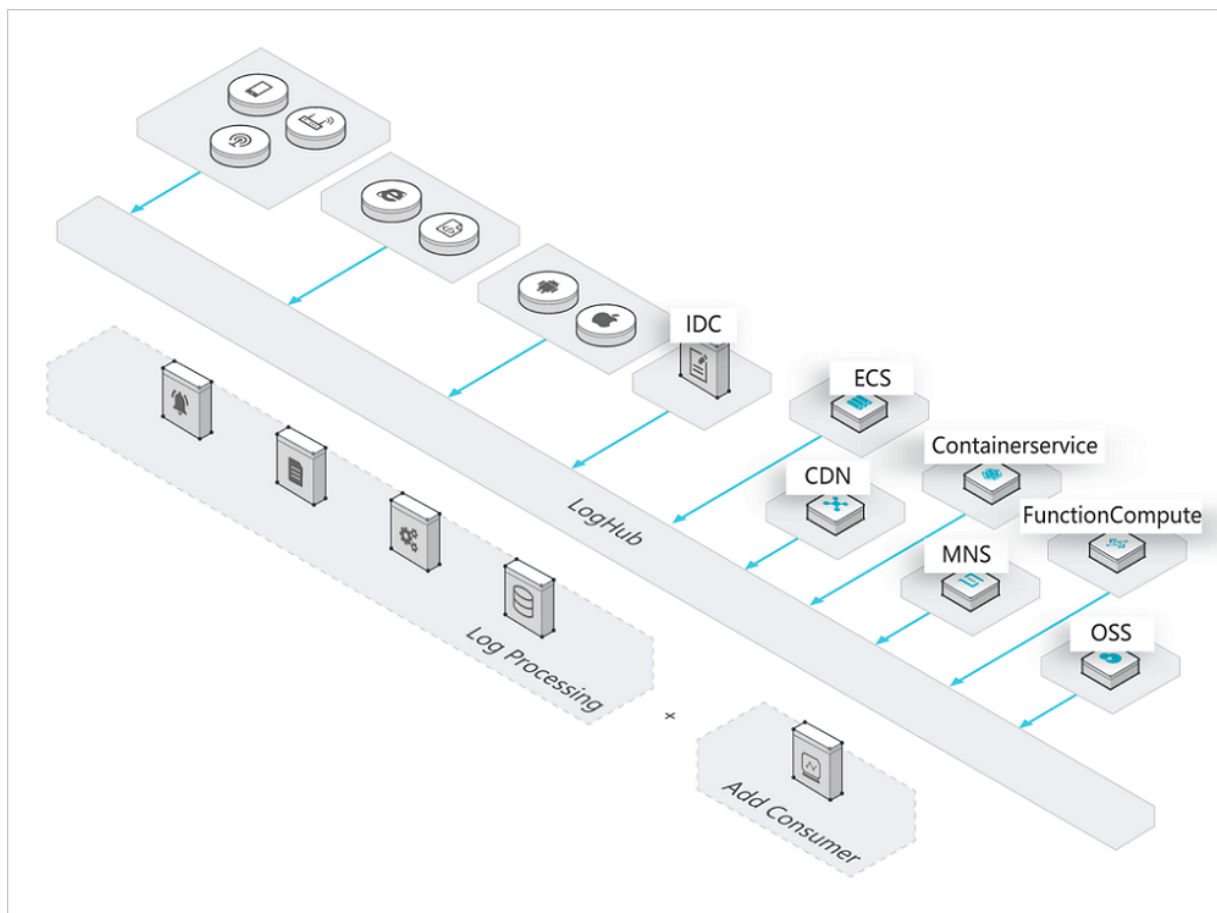
Typical Log Service application scenarios include data collection, real-time computing, data warehousing and offline analysis, product operation and analysis, and O&M and management.

Data collection and consumption

The LogHub function of Log Service enables access to large amounts of real-time log data (including Metric, Event, BinLog, TextLog, and Click data) at low costs.

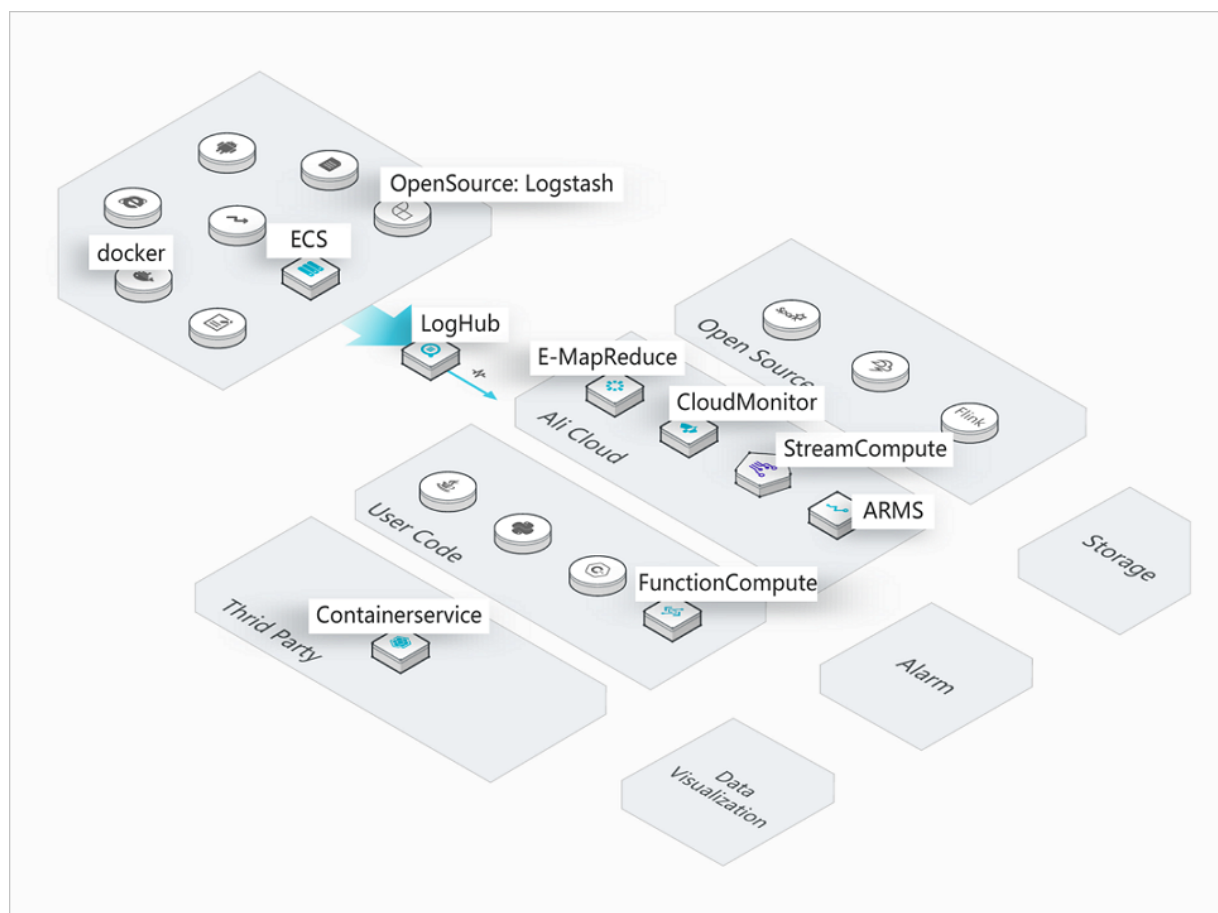
Advantages:

- Easy to use: Over 30 real-time data collection methods are provided for you to quickly set up your platform and reduce O&M workload.
- Automatic scaling: It helps easily cope with traffic peaks and business growth

Figure 19-5: Data collection and consumption**ETL and stream processing**

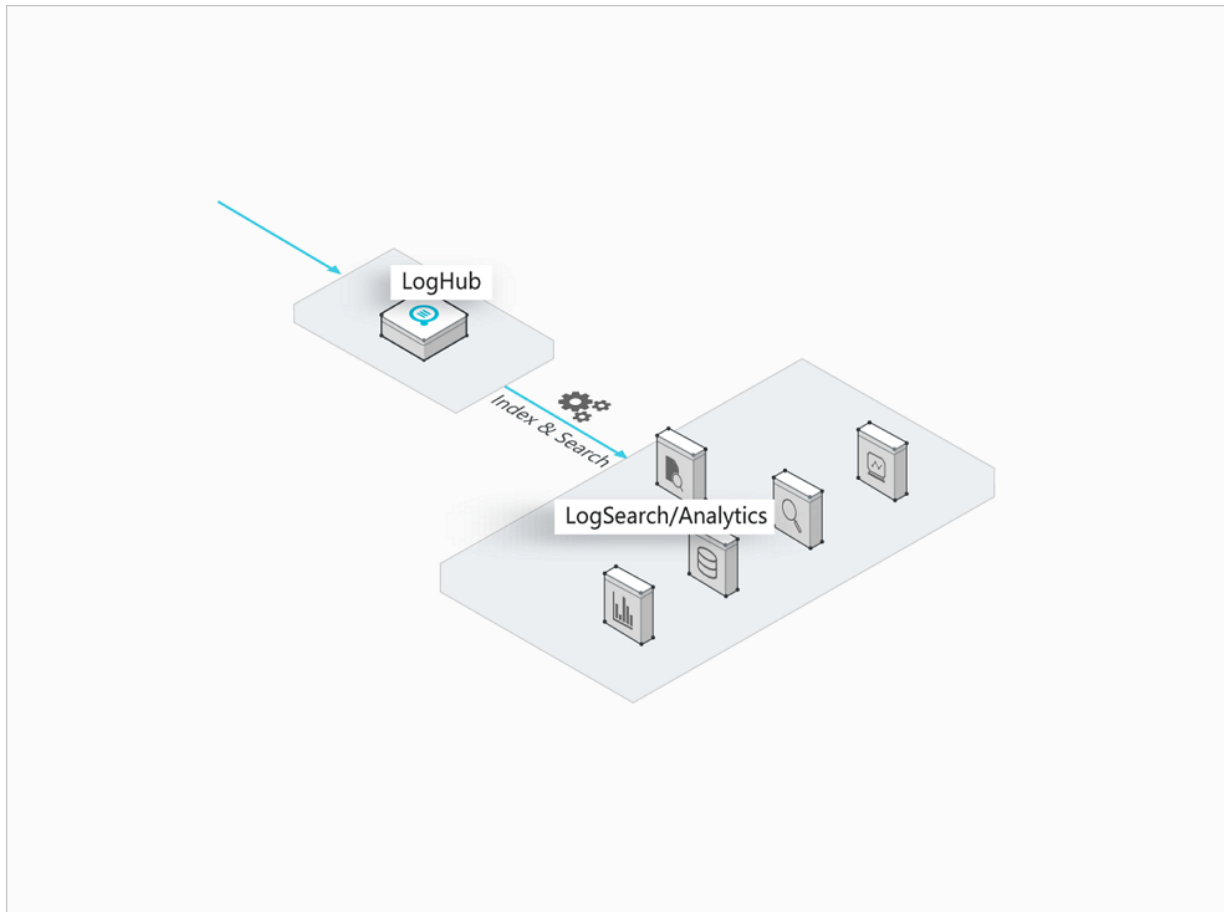
LogHub can interwork with many real-time computing and services to provide complete progress monitoring and alarm notification functions, and support SDK- and API-based custom consumption.

- Easy to operate: It provides various SDKs and programming frameworks and can interconnect with various stream computing engines seamlessly.
- Rich functions: Various monitoring data and alarm reporting are provided.
- Automatic scaling: PB-level elasticity and zero latency.

Figure 19-6: ETL and stream processing**Real-time query and analysis of logs**

The LogAnalytics function supports indexing LogHub data in real time and provides rich query methods such as keywords, fuzzy match, context, range, and SQL aggregation.

- Strong real-timeliness: Supports data query immediately after the data is written.
- Ultra low cost: Supports PB/day indexing capacity, saving costs of up to 85% compared to self-built systems.
- Strong analytical capability: Supports multiple query methods and SQL for aggregation analysis , and provides visualization and alarm notification function.

Figure 19-7: Real-time query and analysis of logs

19.6 Limits

Resource limits

Item	Description	Remarks
Project	Up to 100 projects can be created in each department.	For more, open a ticket.
Logstore	Up to 100 Logstores can be created in each project.	For more, open a ticket.
Shard	<ul style="list-style-type: none"> Up to 10 shards can be created in each Logstore. You can also split a shard to increase the number of shards. Up to 100 shards can be created in each project. 	For more, open a ticket.

Item	Description	Remarks
Dashboard	<ul style="list-style-type: none"> Up to 5 dashboards can be created in each project. Up to 10 charts can be added in each dashboard. 	For more, open a ticket.
Saved search	Up to 10 saved search can be created in each project.	For more, open a ticket.
Logtail configuration	Up to 100 Logtail configurations can be created in each project.	For more, open a ticket.
Consumer group	Up to 10 consumer groups can be created in each project.	For more, open a ticket.
Machine group	Up to 100 machine groups can be created in each project.	For more, open a ticket.
Log retention time	Logs collected to the server can be kept for up to 365 days.	For more, open a ticket.

19.7 Glossary

Log

Log is an abstraction of system changes during the running process. The log content is a time-ordered collection of some operations and the corresponding operation results of specified objects. LogFile, Event, BinLog, and Metric data are different carriers of logs. In LogFile, every log file is composed of one or more logs, and every log describes a single system event. A log is the minimum data unit processed in Log Service.

Log group

A log group is a collection of logs and is the basic unit for writing and reading.

Log topic

Logs in a Logstore can be classified by log topics. You can specify the topic when querying logs.

Project

Project is the resource management unit in Log Service and is used to isolate and control resources. You can manage all the logs and related log sources of an application by using projects. Projects manage the information of all your Logstores and the log collection machine configuration, and serve as the portals where you can access Log Service resources.

Logstore

Logstore is a unit in Log Service to collect, store, and query log data. Each Logstore belongs to a project and multiple Logstores can be created in each project.

Shard

Each Logstore is divided into several shards and each shard is composed of MD5 left-closed and right-open intervals. Each interval range does not overlap with others and the total range of all the intervals is the entire MD5 value range.

20 Apsara Stack Security

20.1 What is Apsara Stack Security

Apsara Stack Security is a solution that provides Apsara Stack with a full suite of security features, such as network security, server security, application security, data security, security management, and security operations services.

In today's cloud computing environment, new technologies are developed every day. Border security protection methods that use traditional detection technologies are insufficient to secure cloud businesses. Apsara Stack Security combines the powerful data analysis capabilities of Alibaba Cloud with the expertise of the Alibaba Cloud security operations team. It provides integrated security protection services at the network layer, application layer, and server layer.

Apsara Stack Security protects core business applications that provide services for the Internet. It provides real-time protection capabilities, including DDoS detection and prevention, Web attack detection and prevention, Web vulnerability detection and fix, server vulnerability detection and fix, and server intrusion prevention. Using a large amount of local security data and the intelligence collected from the cloud, this service performs centralized security big data analysis in the security data analysis engine cluster. It then presents security administrators with the overall security situation and intrusion tracing results, including targeted attack detection, user information leak alerts, and intrusion cause analysis. Based on this core security information, security administrators can understand the security status and use the custom analysis interface provided by the security data analysis engine to perform scenario-based analysis on security data for flexible customization of security analysis capabilities.

20.2 Benefits

This topic describes the benefits of Apsara Stack Security.

Pioneer of cloud security in China

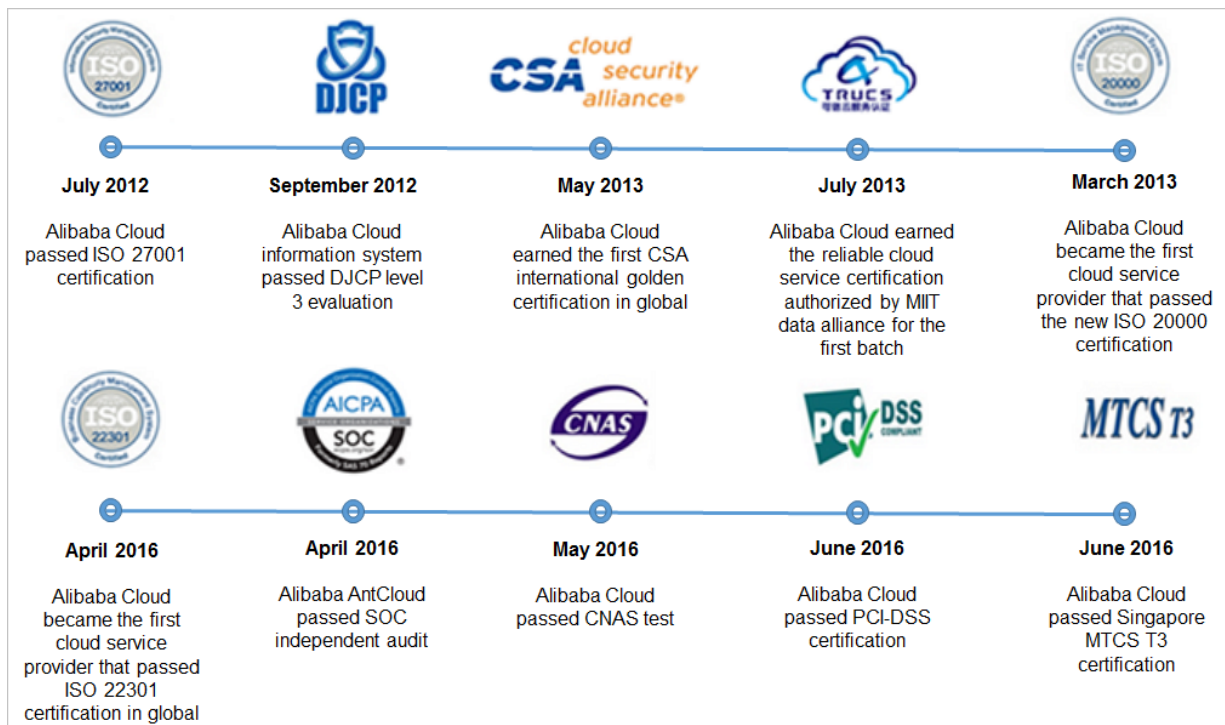
The Apsara Stack Security team has accumulated a wealth of security experience by protecting all internal business systems of Alibaba Group since 2005. Since its release in 2011, Apsara Stack Security has become a pioneer in providing comprehensive protection for cloud security.

Apsara Stack Security protects more than 40% of all websites in China. It prevents more than 50 % of all DDoS attacks and blocks up to 3.6 billion attacks every day. It has fixed over 6.13 million vulnerabilities over the last year.

Security and reliability proved by authoritative certifications

Alibaba Cloud has received multiple national and international cloud security certifications. Our certifications encompass the security features of the Alibaba Cloud platform and the attack prevention features of Apsara Stack Security.

Figure 20-1: Certifications



- Alibaba Cloud is the first cloud service provider in the world to receive CSA STAR Certification.
- Alibaba Cloud is the first cloud security service provider in China to be certified by the ISO 27001 international information security management system.
- Alibaba Cloud is the first cloud computing system in China to pass the Ministry of Public Security's classified security protection test (DJCP).
- Alibaba GovCloud is one of the first platforms to pass the cloud service cyber security review (enhanced level) by government departments.
- Alibaba Cloud is a leading platform in the pilot project of cloud classified security protection in China.
- Alibaba Finance Cloud has passed the Level-4 classified security protection test and is the first Level-4 cloud platform in China.

Mature systems and advanced technologies

Apsara Stack Security is the product of ten years of experience in providing security solutions. After a decade of experience protecting the internal businesses of Alibaba Group, Apsara Stack Security has accumulated a wealth of security research achievements, security data, and security operations and management approaches. A team of cloud security experts has been built. Apsara Stack Security brings together the rich experience of these experts to develop sophisticated systems that provide enhanced security for cloud computing platforms. This product can protect the cloud platform, cloud network environments, and cloud business systems of Apsara Stack users.

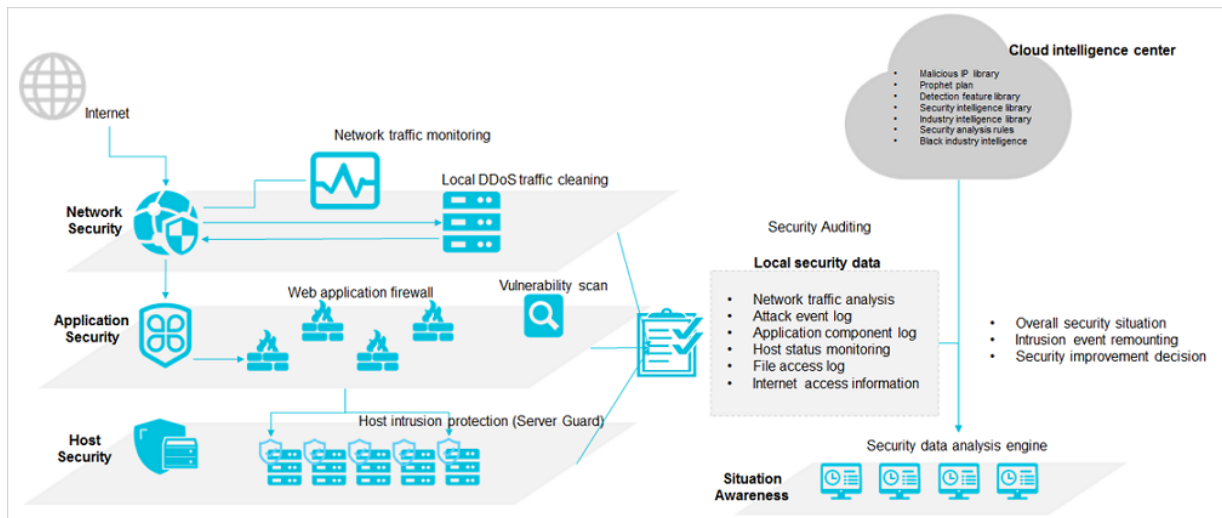
Comparison with traditional security products

Feature	Traditional security products	Apsara Stack Security
Comprehensive industry-leading security capabilities among Internet enterprises	A traditional security service provider only has limited products and features and cannot provide a comprehensive security protection system.	Alibaba has accumulated a large amount of intelligence sources through years of attack prevention experience. This has allowed it to detect common Internet attacks including zero day exploits, and provide comprehensive security capabilities.
Early risk detection	Traditional security service providers cannot detect risks due to a lack of complete monitoring systems.	Apsara Stack Security can detect and respond to critical vulnerabilities and security events quickly to prevent security issues.
Security big data modeling analysis	Traditional security service providers cannot detect threats using signature scanning. The traditional log analysis feature only provides data collection and reporting. It does not provide data modeling analysis.	Big data modeling analysis enables Apsara Stack Security to detect threats in the entire network and display the security data. More than 30 algorithmic models are used to analyze the historical data, network data, and server data. This enables security situation awareness.
Scalability and decoupling with hardware	Traditional security products are developed based on existing hardware devices.	<ul style="list-style-type: none"> Hardware and software decoupling: All modules are developed based on the

Feature	Traditional security products	Apsara Stack Security
	Security product software also relies on virtual machines on virtual platforms.	<p>cloud computing architecture and the common x86 hardware platform, and therefore do not rely on specific hardware.</p> <ul style="list-style-type: none"> Scalability: You can simply increase the number of software for higher performance without the need to change the network architecture.
Collaboration between the network and servers	Traditional service providers increase security features by adding devices. The devices can only collect device logs and status data and display the data on the management platform. They cannot collaborate to provide more features.	Apsara Stack Security provides complete Internet protection to ensure the security of networks, applications, and servers. The security modules interact with each other to form a comprehensive protection system that blocks attacks effectively.
Compatibility with all data center environments and not requiring specific cloud platforms	Most traditional security products are provided in hardware appliances. This makes the product incompatible with cloud platforms based on SDN technology.	Based on interactions between servers and the operating system, Apsara Stack Security detects threats at the network perimeter through data analysis. This enables the service compatibility with all data center environments by avoiding the complex network topology inside data centers.

20.3 Architecture

The architecture of Apsara Stack Security Standard Edition is shown in [Figure 20-2: Apsara Stack Security Standard Edition architecture](#).

Figure 20-2: Apsara Stack Security Standard Edition architecture

- **Traffic Security Monitoring:** This module is deployed on the network perimeter of Apsara Stack. It allows you to inspect and analyze each inbound or outbound packet of an Apsara Stack network by traffic mirroring. The analysis results are used by other Apsara Stack Security modules.
- **Server Intrusion Detection:** This module collects information and performs detection through the client deployed on physical servers. It detects file tampering, suspicious processes, suspicious network connections, suspicious port listening, and other suspicious activities on all servers in the Apsara Stack environment. This helps you detect server data security risks in time.
- **Server Guard:** This module provides security protection features such as vulnerability management, baseline check, intrusion detection, and asset management for ECS instances using log monitoring, file analysis, and signature scanning.
- **Security Audit:** This module collects database logs, server logs, user console operation logs, IT administrator console operation logs, and network device logs in Apsara Stack. This module can store and analyze these logs and trigger alerts on unusual events.
- **Web Application Firewall (WAF):** This module protects Web applications against common Web attacks reported by OWASP, such as SQL injections, XSS, exploitation of Web server plugin vulnerabilities, trojan uploads, and unauthorized access. It blocks a large number of malicious visits to avoid website data leaks. This ensures the security and availability of your websites.
- **Threat Detection Service:** This service collects traffic data and server information and detects potential intrusions or attacks through machine learning and data modeling. It detects vulnerability exploitation and new virus attacks launched by advanced attackers, and

shows you the information of ongoing attacks, enabling business security visualization and awareness.

Apsara Stack Security Basic Edition also provides on-premises security operations services. On-premises security operations services help users make good use of the features of Apsara Stack products and Apsara Stack Security products to ensure the user application security. Security operations services include pre-release security assessment, access control policy management, Apsara Stack Security product configuration, periodic security check, routine security inspection, and urgent event handling. These services cover the entire lifecycle of the user businesses in Apsara Stack. On-premises security operations services help users create a security operations system for cloud businesses. This system enhances the security of application systems and ensures the security and stability of user businesses.

You can also choose the following optional services based on your own business needs to enhance your system security.

- **DDoS Traffic Scrubbing:** This service detects and filters out DDoS attack traffic to block DDoS attacks.

20.4 Features

Apsara Stack Security is developed based on the Apsara Stack environment and adopts a cloud security architecture that enables in-depth defense and multi-module collaboration. Our product is unlike traditional software and hardware security products. Apsara Stack Security provides comprehensive and integrated cloud security protection capabilities that protect the network layer, application layer, server layer, and other layers.

Features

The features provided by Apsara Stack Security Standard Edition are shown in [Table 20-1: Apsara Stack Security Standard Edition features](#).

Table 20-1: Apsara Stack Security Standard Edition features

Module	Feature	Description
Traffic Security Monitoring	Traffic data collection and analysis	Collects inbound and outbound traffic through the interconnection switch (ISW) using a bypass in the traffic mirroring mode and generates a traffic diagram.

Module	Feature	Description
	Malicious server detection	Detects attacks launched by malicious servers in the Apsara Stack network and locates the malicious servers.
	Unusual traffic detection	Uses a bypass in the traffic mirroring mode to detect unusual traffic that has exceeded the threshold.
	Web application protection	Uses a bypass to block common Web attacks at the network layer based on default Web attack detection rules.
Server Intrusion Detection	Key directory integrity check	Checks the integrity of files in a specific system directory (<i>/etc/init.d</i>) to detect file tampering and generate alerts.
	Suspicious process alerts	Detects suspicious processes, and generates alerts.
	Suspicious port alerts	Detects new port listening tasks, and generates alerts.
	Suspicious network connection alerts	Detects active connections with the public network, and generates alerts.
Server Guard	Baseline check	Checks the security baselines of the ECS instances, including the account security, weak passwords, and at-risk configuration items. This ensures that the ECS instances comply with security standards for enterprise servers.
	Vulnerability management	<ul style="list-style-type: none"> Scans for vulnerabilities in the software on ECS instances, and provides suggestions on vulnerability fixes. Provides quick fixes for critical vulnerabilities in applications and the operating system on your ECS instance, such as Web application vulnerability fixes and system file repair.
	Webshell detection and removal	Accurately detects and removes webshells by rule matching, and allows you to manually quarantine webshells.
	Brute-force attack blocking	Detects and blocks brute-force attacks in real time.

Module	Feature	Description
	Unusual logon alerts	Detects unusual logons based on the approved logon settings, and generates alerts.
	Suspicious server detection	Detects suspicious processes, such as reverse shells, Java process running CMD commands, and unusual file downloads using bash.
	Asset fingerprints	Collects information on the servers, including ports, accounts, processes, and software. Uses the data to learn the server running status and perform event tracing.
	Log retrieval	Centrally manages server logs on processes, networks, and system logons. This allows you to quickly locate the cause of a problem by log retrieval.
Security Audit	Raw log collection	Collects the following logs: <ul style="list-style-type: none"> Database logs and server logs Operation logs of both the user console and the IT administrator console Network device logs
	Audit query	Allows you to query audit logs by audit type , audit target , operation type , operation risk level , alert , or creation time . Full-text search is supported.
	Policy setup	Allows you to configure audit rules using the following parameters: initiator , target , command , result , and cause . Identifies at-risk operations in raw logs and generates alerts.
Web Application Firewall	Web application protection	Blocks Web attacks including SQL injections, XSS, file upload vulnerabilities, file inclusion vulnerabilities, sensitive information leakage, common CMS vulnerabilities, code injections, webshells, hacking tools, and scanner attacks.
	HTTP flood mitigation	Detects and mitigates HTTP flood attacks by analyzing the HTTP response status codes, the distribution of URL requests, unusual HTTP referers, and the user agent features. Controls the frequency of access from a specific IP address based on custom rules. Supports

Module	Feature	Description
		URL redirection for authentication and can identify bot-initiated requests.
	Precise access control at the application layer	You can combine different HTTP fields, such as IP address, URL, HTTP referer, and user agent , to implement precise access control.
	Enhanced Web attack prevention	<ul style="list-style-type: none"> Automatic blocking of malicious IP addresses: WAF automatically blocks the access of an IP address that initiates consecutive Web attacks on a domain name. Region blocking: WAF allows you to block IP addresses in a specified province or region outside China. Business analysis: WAF collects business data of the traffic, including the Top N statistics of the visitor IP addresses, visitor regions, URL access latency, and browser types.
Threat Detection Service	Security situation overview	Provides the overall security information, including the number of emergencies, the current day's attacks, the current day's flaws , attack trend, latest threat analysis, latest intelligence, and protected assets information.
	Access analysis	Analyzes all information about the access to the protected Web services, including the top 10 accessed services, number of normal source IP addresses, number of malicious source IP addresses, number of crawler source IP addresses, and detailed access samples.
	Screens	Provides map-based traffic data screens and server security screens.
	Security event analysis	Uses big data algorithms and models to detect zombies, brute-force attacks, backdoors, DDoS attacks, hacking tools, suspicious network connections, unusual traffic, and other security events.
	Traffic data collection and analysis	Collects the traffic data in the monitored IP range, including the current day's traffic, traffic in the last 30 days, traffic in the last 90 days

Module	Feature	Description
		, and the QPS. Displays the traffic data of a specific IP address.
	Malicious server identification	Detects attacks launched by internal malicious servers, such as HTTP flood and DDoS attacks , and identifies the controlled malicious servers.
	Web attack detection	Detects Web vulnerability exploitation, malicious scanning tools, webshell uploads and connections, SQL injections, XSS, local and remote file inclusion attacks, code or command execution, and other attacks.
	Server vulnerability exploitation detection	Converts packet feature characters into binary strings and matches these strings with signatures to detect security events such as the exploitation of Redis server vulnerabilities.
	Application vulnerability analysis	Detects Web application vulnerabilities and provides advice on vulnerability fixes and vulnerability fix verification. Periodically and automatically scans NAT assets and servers for application vulnerabilities, verifies the detected vulnerabilities, and updates the vulnerability status.
	Server vulnerability analysis	Detects server vulnerabilities, and provides the scanning results and advice on vulnerability fixes.
	Weak password analysis	Detects weak passwords of accounts in common systems such as Web, SSH, FTP , MySQL, and SQL Server and allows users to customize weak password policies. Automatically scans NAT assets and servers at a scheduled time each day, verifies the detected weak passwords, and updates the weak password detection time.
	At-risk configuration detection	Scans the access to external service pages , generates alerts on leaks of Web page configuration items, verifies the detected leaks of configuration items at a scheduled time every day, and updates the detection time.

Security operations services

Apsara Stack Security Standard Edition provides on-premises security operations services that ensure the security of user business systems. The included services are shown in [Table 20-2: On-premises security operations services](#).

Table 20-2: On-premises security operations services

Category	Service	Description
User business security operations	User asset research	With user authorization, this service periodically analyzes the cloud businesses of the user and develops a business list containing information such as the business system name, ECS, RDS, IP address, domain name, and owner.
	New business security assessment	<ul style="list-style-type: none"> Before a user migrates a new business system to the cloud, this service detects system vulnerabilities and application vulnerabilities in the new business system using both automation tools and manual operations. Provides advice and verification on vulnerability fixes.
	Periodic business security assessment	<ul style="list-style-type: none"> Periodically uses automated tools to detect system vulnerabilities, application vulnerabilities, and security risks in running businesses. Provides advice on handling detected risks, including but not limited to security policy settings, patch updates, and application vulnerability handling.
	Access control management	Provides inspection and guidance on applying access control policies when new business is migrated to the cloud.
	Access control routine inspection	Periodically checks for access control risks of user businesses.
	Security risk routine inspection	Monitors and inspects security events in Apsara Stack Security. Informs the user of verified events and provides advice on event handling.

Category	Service	Description
Apsara Stack Security operations	Rule update	Periodically updates the rule libraries of Apsara Stack Security products.
	Product integration	<ul style="list-style-type: none"> Provides technical support for integrating Apsara Stack Security products with the application systems of users. Helps users customize and optimize security policies.
Security event response	Event alerts	Synchronizes recent security events information from Alibaba Cloud, and helps users remove the risks.
	Event handling	Handles urgent events such as attacker intrusions.

Optional services

Apsara Stack Security provides the following optional services:

Table 20-3: Optional services

Service	Feature	Description
DDoS Traffic Scrubbing	Traffic scrubbing	Detects and prevents attacks such as SYN flood, ACK flood, ICMP flood, UDP flood, NTP flood, DNS flood, and HTTP flood.
	DDoS attack display	Allows you to view DDoS attacks in the console and search for DDoS attacks by IP address, status, and event information.
	DDoS traffic analysis	Allows you to monitor and analyze the traffic of a DDoS attack, and view the attack traffic protocol and the 10 IP addresses that have launched most attacks.

20.5 Restrictions

None

20.6 Concepts

DDoS attacks

An attacker attempts to cause network failures by initiating a large number of valid requests to consume network resources.

SQL injections

An attacker makes the server run malicious SQL commands by inserting these commands into Web tables or inserting malicious strings in URL requests.

Traffic scrubbing

The traffic scrubbing module monitors the inbound traffic of a data center in real time and detects unusual traffic that may be from DDoS attacks and other attacks. The module scrubs the unusual traffic without affecting businesses.

Password cracking based on brute-force attacks

Brute-force attacks work by iterating through all possible combinations that can make up a password.

Webshells

A webshell is a script written in languages such as ASP and PHP. Attackers can run a webshell on a Web server to perform risky operations. This enables attackers to obtain sensitive information or control the server through server penetration or privilege escalation.

Server intrusion detection

By analyzing server logs, Apsara Stack Security can detect attacks, such as system password cracking and logons from unusual IP addresses, and generate real-time alerts.

21 Key Management Service (KMS)

21.1 What is KMS

Key Management Service (KMS) is a secure and easy-to-use key management service provided by Apsara Stack. KMS allows you to create and manage CMKs with ease and use a DEKs to encrypt your data.

KMS integrates many Alibaba Cloud products and services to help protect your data in the cloud.

[Table 21-1: KMS solutions](#) describes how KMS provides solutions for a variety of concerns and issues.

Table 21-1: KMS solutions

Role	Requirement	Solution
Application or website developer	<ul style="list-style-type: none"> My program needs keys or certificates for encryption or signature, and I want secure and independent key management services. I want to securely access keys regardless of where my application is deployed , and cannot take the risk of deploying plaintext keys elsewhere. 	KMS provides envelope encryption, allowing you to store the Customer Master Key (CMK) in KMS and deploy only the DEKs. You can simply call a KMS API to decrypt DEKs only when necessary.
Service developer	<ul style="list-style-type: none"> I do not want to be responsible for securing users keys and data. I want users to manage their own keys. I want to use specified keys to encrypt their data after obtaining their authorization . In this way, I can focus on developing service features. 	Envelope encryption and KMS APIs allow service developers to use specified CMKs to encrypt and decrypt DEKs. Plaintexts are not directly stored in a storage device. This method helps service developers manage CMKs.
Chief security officer (CSO)	<ul style="list-style-type: none"> There are compliance requirements that I expect 	KMS can connect to RAM for unified authorization management.

Role	Requirement	Solution
	<p>our key management activities to meet.</p> <ul style="list-style-type: none">• I need to ensure that keys are reasonably authorized and that the use of any keys is audited.	

21.2 Benefits

Cost-friendly

Traditional key management solutions require the purchase of secure key management equipment to construct a secure physical environment, as well as the design and implementation of key management solutions and specifications. This mode leads to high costs in hardware and software.

KMS enables you to manage your keys on the cloud platform in a unified manner while minimizing hardware and software investment.

Ease of use

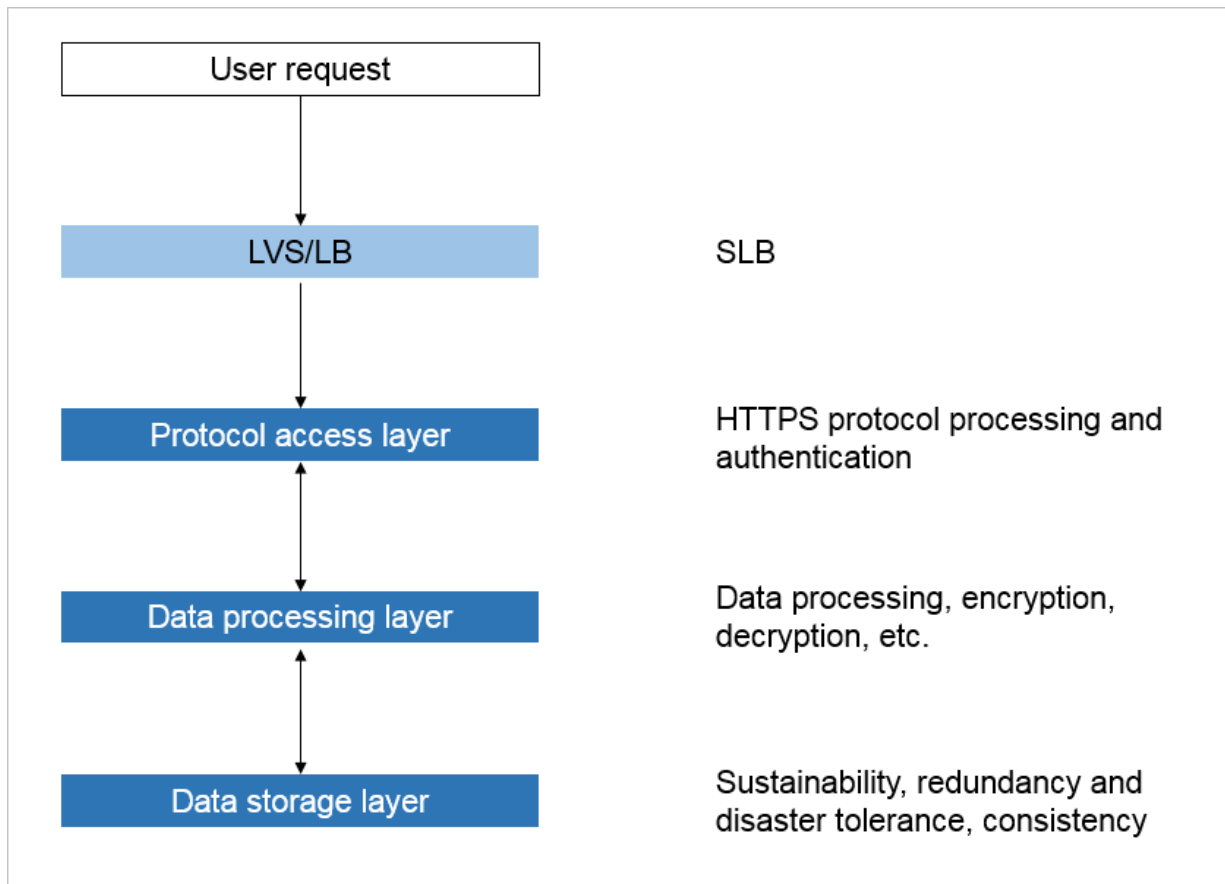
KMS uses the unified APIs and standard HTTPS for ease-of-use.

Reliability

KMS uses a distributed system to enhance reliability.

21.3 Architecture

The following figure shows the KMS architecture.

Figure 21-1: Architecture

- The protocol access layer of KMS receives HTTPS requests sent from a user to KMS, verifies the user identity, and authenticates the permission. After the verification and authentication succeed, the user request is forwarded to the data processing layer. The data processing layer receives the processing result and sends it to the user. If the verification and authentication fail, no data is processed and error information is returned.
- The data processing layer of KMS processes requests. Data processing in KMS involves cryptography-relevant operations such as encryption and decryption. The protocol access layer and data processing layer communicate with each other based on RPC of TLS. The data processing layer adopts distributed deployment. The nodes are independent of each other. Requests sent from the protocol access layer can be properly processed on any node at the data processing layer.
- The storage layer of KMS stores core root keys, uses Raft to ensure data consistency, and uses TPM to implement persistent encrypted storage.

21.4 Features

The following table describes the features of KMS.








Feature	Description
Create a CMK	You need to create at least one CMK before KMS can be used. CMKs can be used to encrypt a small amount of data (less than 4 KB). However, in most cases, CMKs are used to call the GenerateDataKey API to generate DEKs.
Create a DEK	You must use a specified CMK to create a DEK when using KMS for envelop encryption. You can use the DEK to encrypt local data.
Encrypt data	You can use a specified CMK to encrypt a small amount of data (less than 4 KB) such as RSA keys, database passwords, or other sensitive user data.
Decrypt data	You can decrypt data encrypted through KMS.
View CMKs	You can obtain IDs of all CMKs that belong to the current region in your account.
View CMK details	You can view detailed information about a specified CMK, such as its creation date and time, description, globally unique identifier, CMK status, purpose, scheduled deletion time, creator, source of the CMK material, and expiration time of the CMK material.
Enable a CMK	You can change the status of a CMK from disabling to enabling.
Disable a CMK	If the status of a CMK is changed from enabling to disabling, the disabled CMK cannot be used to encrypt or decrypt data.
Schedule a CMK to be deleted	You can schedule a CMK to be automatically deleted after a specified deletion period.
Cancel scheduled key deletion	You can cancel the scheduled deletion of a CMK to change the CMK status back to enabling.
API	You can call HTTPS APIs to use KMS.
SDK	You can use SDKs in mainstream languages.

21.5 Scenarios

This topic describes the following typical scenarios of KMS:

- Use KMS to encrypt and decrypt data.
- Use envelope encryption to encrypt and decrypt data locally.

Table 21-2: Example description

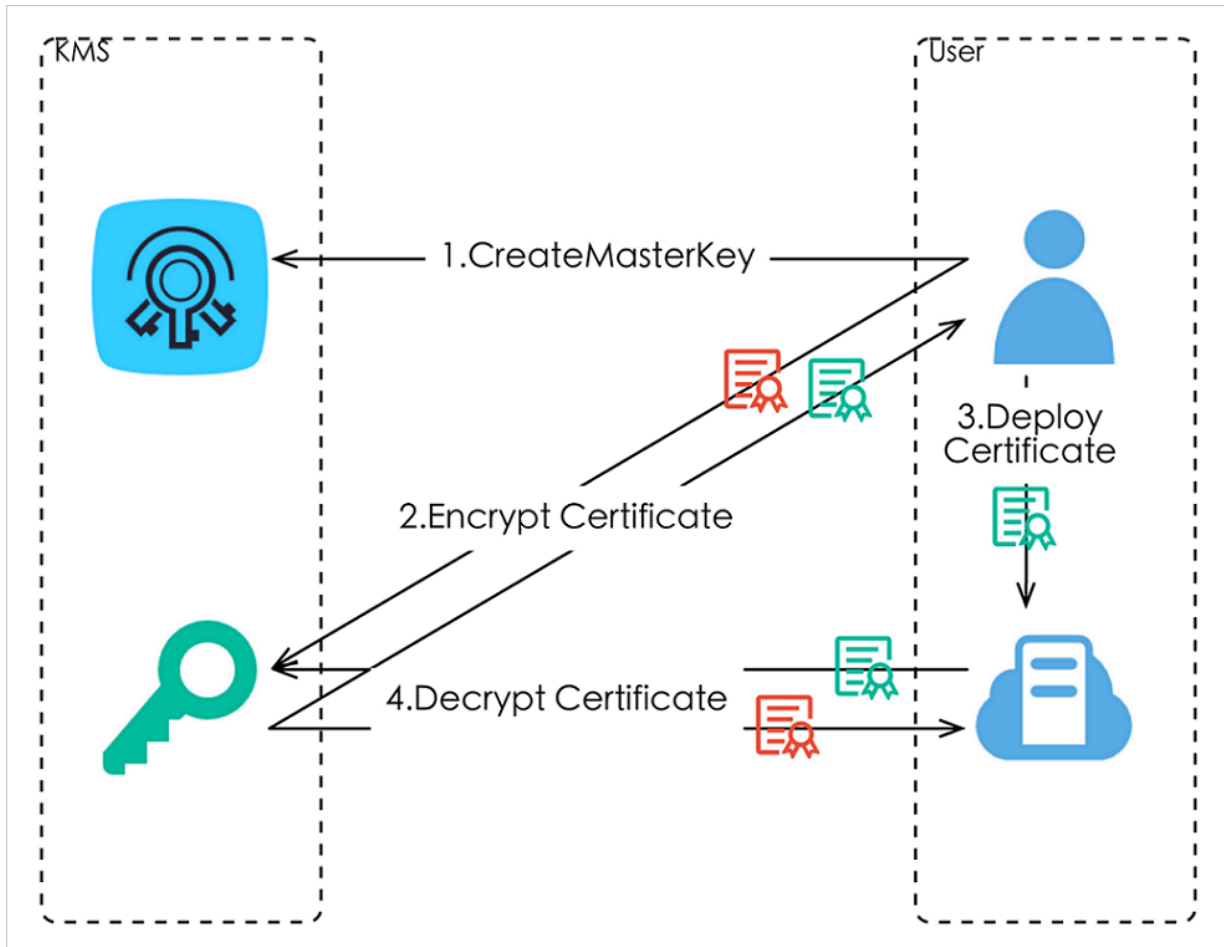
Example	Description	Example	Description
	CMK		The ciphertext key.
	The plaintext certificate.		The plaintext file.
	The ciphertext certificate.		The ciphertext file.
	The plaintext key.	-	-

Directly use KMS to encrypt and decrypt data

You can directly call KMS APIs to encrypt and decrypt data with a specified CMK.

This scenario applies to the encryption and decryption of a small amount of data (less than 4 KB). Data is transmitted to and from, and encrypted or decrypted on the KMS server over secure channels.

Example: Encrypt the HTTPS certificate on the server, as shown in [Encrypt the HTTPS certificate on the server](#).

Figure 21-2: Encrypt the HTTPS certificate on the server

The procedure is as follows:

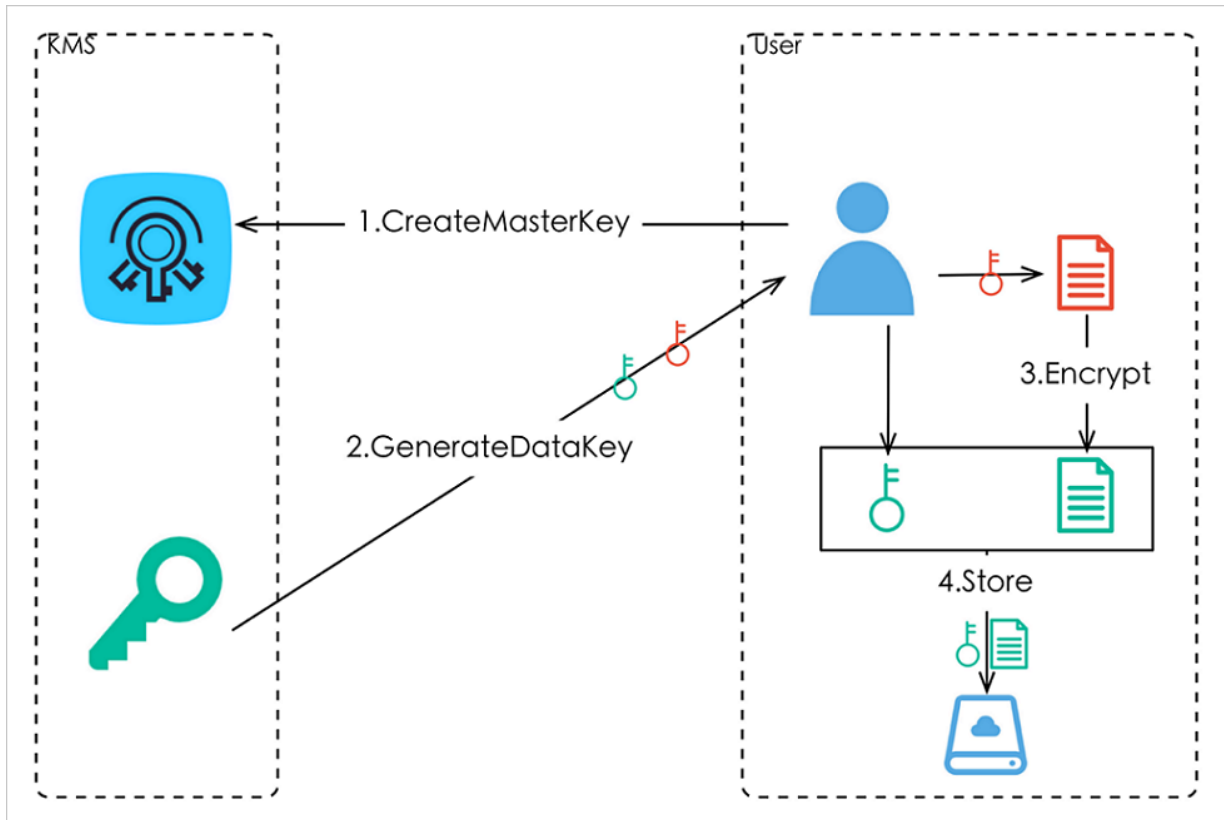
1. Create a CMK.
2. Call the Encrypt API to encrypt the plaintext certificate.
3. Deploy the ciphertext certificate on the server.
4. When the server has been started and needs the plaintext certificate, call the Decrypt API to decrypt the ciphertext certificate.

Use envelope encryption to encrypt and decrypt data locally

You can directly call a KMS API to use a specified CMK to generate and decrypt a DEK, and use the DEK to encrypt and decrypt data locally.

This scenario applies to encryption and decryption of large amounts of data that does not need to be transmitted over the network, which minimizes costs.

Example: Encrypt a local file, as shown in [Encrypt a local file](#).

Figure 21-3: Encrypt a local file

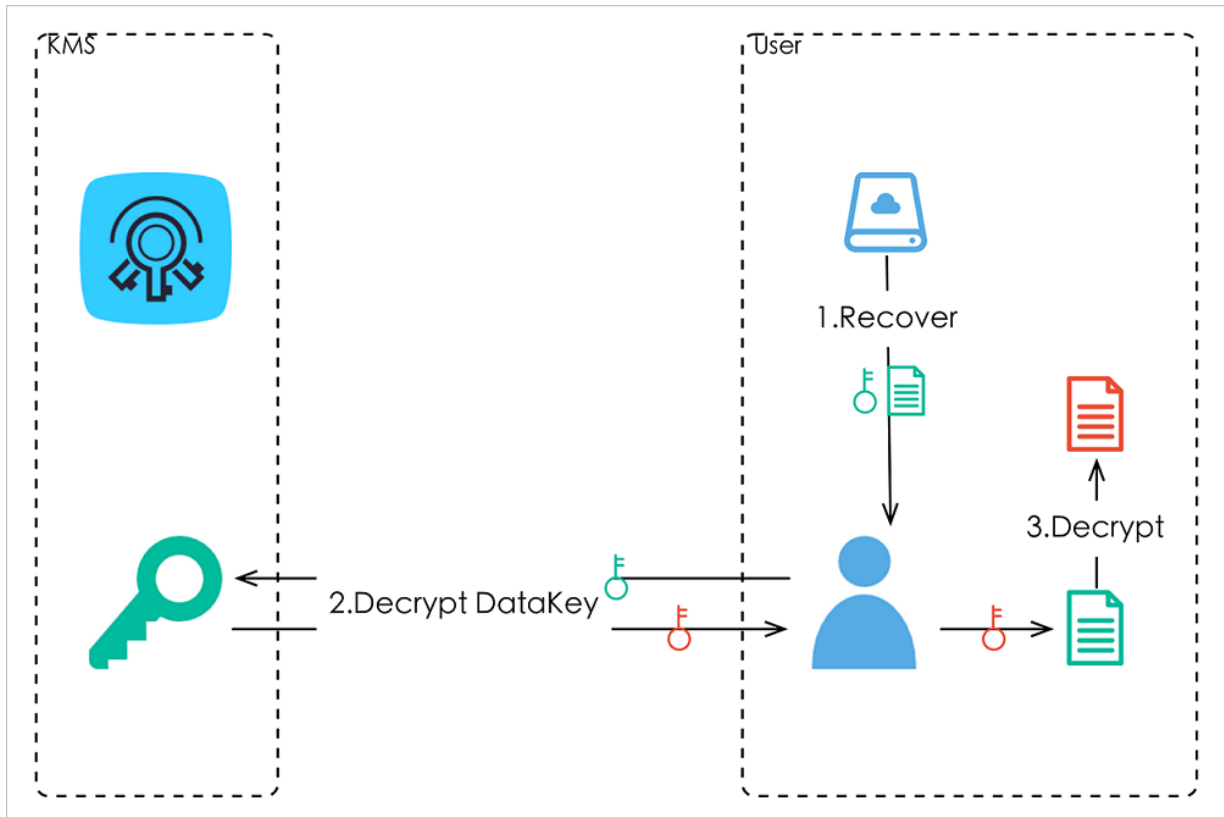
The encryption process is as follows:

1. Create a CMK.
2. Call the GenerateDataKey API to generate a DEK.

You can obtain a DEK and an EDK.

3. Use the DEK to encrypt the file and generate a ciphertext file.
4. Save the EDK and the ciphertext file to a persistent storage device or service.

[Decryption process](#) shows how to decrypt the encrypted file.

Figure 21-4: Decryption process

The decryption process is as follows:

1. Read the EDK and the ciphertext file from the persistent storage device or service.
2. Call the Decrypt API to decrypt the EDK and obtain the DEK.
3. Use the DEK to decrypt the file.

Notes

1. You must authenticate the Alibaba Cloud server HTTPS certificate to prevent phishers from stealing your information.
2. We recommend that you assign different permissions to users based on their CMKs.

21.6 Limits

A maximum of 200 CMKs can be created for a department.

21.7 Terms

envelope encryption

The practice of encrypting plaintexts by using a unique DEK, which is then encrypted with CMK.

The ~~EDK~~ is stored and transferred directly over unsecured communication processes. You need to retrieve the EDK only when you need it.

customer master key (CMK)

A master key created by a user in Apsara Stack KMS, which is used to encrypt DEKs and generate EDKs. It can also be used to encrypt a small amount of data.

enveloped data key (EDK)/data encryption key (DEK)

EDK: the ciphertext key generated by using envelop encryption. DEK: the plaintext key used to encrypt data.

22 Domain Name System (DNS)

22.1 What is Apsara Stack DNS

Apsara Stack DNS is a service that runs on Apsara Stack and translates domain names. Based on the rules you have set, Apsara Stack DNS translates domain names that you have requested and direct requests from the client to the corresponding cloud services, business systems in enterprise internal networks, and services provided by Internet service providers.

Apsara Stack DNS provides basic domain name translation and scheduling services for VPC environments. You can perform the following operations through Apsara Stack DNS in your VPC:

- Access other ECS servers deployed in VPCs.
- Access cloud service instances provided by Apsara Stack.
- Access custom enterprise business systems.
- Access Internet services and business.
- Establish network connections between Apsara Stack DNS and user-created DNS through a leased line.

22.2 Benefits

Domain name management for enterprise domains

Apsara Stack DNS provides domain name management and translation services for enterprise domains.

- Apsara Stack DNS supports DNS resolution and reverse DNS resolution for domain names of cloud service instances, including ECS instance domain names.
- It also supports DNS resolution and reverse DNS resolution for your internal domain names.
- You can add, modify, and delete DNS records, including A, AAAA, CNAME, NS, MX, TXT, SRV, and PTR.
- You can add multiple DNS records, including A, AAAA, and PTR, for one host. By default, the resolution finds all matching records. The records can be randomly rotated to balance the load.

Flexible networking

Apsara Stack DNS provides the domain name forwarding service for enterprise domains, which allows you to flexibly create or combine networks.

- Supports forwarding all domain names.

- Supports forwarding specific domain names.

Access the Internet from your server

When the public network is accessible, Apsara Stack DNS supports recursive queries for public domain names and Internet domain names. This service allows your servers to access the Internet.

A unified management platform

The management system of Apsara Stack DNS is built on the unified management platform of Apsara Stack. You can use one account to manage all services. Apsara Stack DNS has the following benefits:

- Data management and service management support Web actions, which are easy to learn and operate.
- Apsara Stack DNS is deployed on clusters. You can add more clusters based on your needs.
- You can deploy Apsara Stack DNS in multiple zones. Apsara Stack DNS supports active-active deployment in the same city and disaster recovery deployment in the same city.
- Apsara Stack DNS is deployed based on anycast. High availability and disaster recovery can be automatically enabled.

22.3 Architecture

The architecture of Apsara Stack DNS

- Deploys two physical servers for network connections and you can add more servers based on your needs.
- Uses two control interfaces for bond, which is uplinked to the ASW. The gateway is the default gateway of the internal network.
- Two service interfaces are uplinked to the LSW (ECMP is supported). These interfaces support OSPF to advertise anycast VIP routes, and are connected to the Internet.
- The control system is deployed in a container in the control area.

22.4 Features

Internal domain name management

Apsara Stack DNS provides data management for internal domain names. You can register, search, and delete internal domain names and add remarks. You can also add, delete, and modify

DNS records. Supported DNS record types include A, AAAA, CNAME, NS, MX, TXT, SRV, and PTR.

Internal domain name management can translate internal domain names for servers deployed in a VPC. The DNS server addresses are deployed based on anycast, which ensures the continuity of services if errors occur.

Domain name forwarding management

Apsara Stack DNS can forward a specific domain name to other DNS servers for translation.

The domain name forwarding feature includes two forwarding modes: forward all requests (with recursion) and forward all requests (without recursion).

- Forward all requests (without recursion): Uses the target DNS server to translate domain names. If the domain names cannot be translated, or the request is timed out, a message is returned to the DNS client indicating that the query fails.
- Forward all requests (with recursion): Uses the target DNS server to translate domain names. If the domain names cannot be translated, then uses the local DNS server to translate them.

Recursive query management

Apsara Stack DNS supports recursive queries, which enables your servers to access the Internet.

Option configuration

You can enable, modify, or disable global default forwarding for Apsara Stack DNS.

22.5 Scenarios

Scenario A: Access cloud resources from a VPC environment

Apsara Stack DNS allows VPC-connected ECS or Docker instances to access Alibaba Cloud instances such as RDS, SLB, and OSS instances.

Scenario B: Access ECS hostnames from a VPC environment

If you need to define hostnames for your VPC-connected ECS and Docker instances according to your own rules, then use Apsara Stack DNS to remotely access and control the ECS instances and Docker instances using their hostnames.

Scenario C: Access the service domain name in the internal network from a VPC environment

If you need to develop your own SaaS service on Apsara Stack and assign a domain name that only allows internal access, Apsara Stack DNS helps you access the SaaS service through the domain name in a VPC environment.

Scenario D: Perform round-robin traffic redistribution for the internal network services provided by Apsara Stack

If you need to develop your own SaaS service on Apsara Stack and assign a domain name that only allows internal access, and this service is deployed in multiple zones or regions, Apsara Stack DNS helps you access your SaaS service in a VPC environment and redistribute traffic to different nodes.

Scenario E: Access the Internet from a VPC environment

You can use Apsara Stack DNS to access the Internet from a VPC environment.

Scenario F: Establish network connections among multiple networks on Apsara Stack

You can use Apsara Stack DNS to establish network connections between your internal network and Apsara Stack networks.

22.6 Limits

Apsara Stack DNS clusters have the following restrictions:

Table 22-1: Restriction

Cluster	Module	Server type	Configuration requirement	Quantity requirement
Service cluster	Basic edition-resolution module	Q46S1.2B	Minimum configuration : 16-core CPU + 96 GB of memory + two GE ports + two 10-GE ports + 600 GB of hard disk	2
Control cluster	Basic edition-resolution module	Base container	Minimum configuration : 4-core CPU + 8 GB of memory + 60 GB of hard disk + network connection support	2

22.7 Basic concepts

DNS

Domain Name System (DNS) is a distributed database used for TCP/IP applications. It translates domain names into IP addresses, and selects paths for emails.

Domain name resolution

This is a process that translates domain names into IP addresses based on the DNS system.

Domain name resolution includes authoritative DNS and recursive DNS.

Recursive DNS

Recursive DNS queries domain names cached on the local DNS server or sends a request to the authoritative resolver to obtain the corresponding IP addresses. You can use recursive DNS to translate Internet domain names.

Authoritative DNS

Authoritative DNS translates root domains, top-level domains, and other levels of domains.

Authoritative domain names

Authoritative domain names are domain names translated by the local DNS server. You can configure and manage DNS records on the local DNS server.

DNS forwarding

DNS forwarding uses two local DNS servers to provide DNS resolution services. One DNS server is used to configure and manage the domain name resolution data. The other DNS server is used to translate domain names.

Default forwarding

DNS queries for authoritative domain names are forwarded to another DNS server for resolution if they are not translated by the local DNS server.

23 API Gateway

23.1 Product overview

API gateway is a complete API hosting service. It helps you use APIs to provide capabilities, services, and data to your partners. You can also publish APIs to the API marketplace for other developers to purchase and use.

- API Gateway provides a range of mechanisms to enhance security and reduce risks arising from APIs. These mechanisms include attack prevention, replay prevention, request encryption, identity authentication, permission management, and request throttling.
- API Gateway provides a full range of API lifecycle management functions, including creating, testing, publishing, and unpublishing APIs. It also generates SDKs and API documentation to improve API management and iteration efficiency.
- API Gateway provides convenient O&M functions to reduce API O&M costs, including monitoring, alarms, and log analysis.

API Gateway maximizes capability multiplexing. It allows enterprises to share capabilities and focus more on their core businesses, which benefits all parties involved.

23.2 Features

API lifecycle management

- API lifecycle management enables you to manage APIs throughout their full lifecycle, including publishing, testing, and unpublishing APIs.
- API lifecycle management supports maintenance features such as routine management, version management, and quick rollback.

Comprehensive security protection

- API Gateway supports multiple authentication methods and HMAC (SHA-1 and SHA-256) algorithms.
- API Gateway supports HTTPS and SSL encryption.
- API Gateway provides multiple security mechanisms to prevent injections, replay attacks, and tempering.

Flexible access control

- Applications are used to make API requests. API Gateway implements access control for applications.
- If an application attempts to call an API, the application must first be authorized.
- API providers can authorize applications to call APIs.

Precise throttling

- You can use throttling to control API access frequency, application request frequency, and user request frequency.
- The unit of time for throttling can be set to minute, hour, or day.

Request validation

API Gateway validates parameter types and values by using range, enumeration, and regular expression. When an API request fails to be validated, API Gateway immediately rejects the request. This helps reduce the amount of back-end resources wasted on invalid requests and significantly lowers the processing costs of back-end services.

Data conversion

API Gateway enables you to configure mapping rules to translate front-end and back-end data.

- API Gateway supports data conversion for front-end requests.

23.3 Benefits

Easy maintenance

After you create APIs in API Gateway, API Gateway performs all the other API management functions. This significantly reduces routine maintenance costs.

Large scale and high performance

API Gateway uses a distributed deployment and automatic scaling model to respond to a large number of API access requests at very low latencies. It provides highly secure and efficient gateway functions for your backend services.

Security and stability

You can securely open your services to API Gateway on the intranet. API Gateway also provides enhanced permission management functions, and precise request throttling functions. It makes your services secure, stable, and controllable.

23.4 Concepts

It is important to familiarize yourself with the following basic concepts when you use API Gateway.

Application

Application

An application defines the identity of an API caller. To call an API, you must first create an application.

AppKey and AppSecret

Each application has an AppKey and AppSecret pair. This pair is encrypted and attached to a request as the signature.

Encrypted signature

An encrypted signature is attached to each API request and is authenticated by API Gateway.

Authorize

The API service provider can open an API to an application by granting authorization to the application. Only authorized applications can call the specified API.

API lifecycle

The API service provider manages an API by stages, including creating an API, testing the API, publishing the API, unpublishing the API, and changing the version.

API definition

An API definition is a set of rules defined by the API service provider when creating an API. The API definition specifies the backend service, request format, received format, and returned format.

Parameter mapping

Parameter mapping is configured by the API service provider. It is used when the parameters in a request are inconsistent from those of the API backend service.

Parameter verification

Parameter verification is performed based on a set of rules defined by the API service provider. API Gateway filters out invalid requests based on these rules.

Constant parameter

API users do not have to input the constant parameters. The constant parameters are always received by the backend service.

System parameter

You can configure API Gateway to add certain system parameters such as CaClientIP (request IP address) to the requests sent to you by the backend service.

API group

An API group is a group of APIs that are managed by the API service provider as a whole. Before you create an API, you must first create an API group.

Second-level domain name

A second-level domain name is a domain name that you bind to an API group when creating the group. The second-level domain name is used to test API calling.

Independent domain name

An independent domain name is a domain name that you bind to an API group when opening an API in the group. Users must access the independent domain name to call the API.

Signature key

A signature key is created by the API service provider and bound to an API. The signature is added to each request sent from API Gateway to the backend service. The backend service checks the signature for security purposes.

Throttling policy

The API service provider can configure a throttling policy to limit the maximum number of requests for an API, and the maximum number of API requests that can be initiated by a user or an application. The throttling granularity can be day, hour, or minute.