

Alibaba Cloud Elastic Compute Service

Product Introduction

Issue: 20190918

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 content of the Alibaba Cloud website, including but not limited to works, products, images, archives, information, materials, website architecture, website graphic layout, and webpage design, 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 the Alibaba Cloud website, product programs, or content 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.

Generic conventions

Table -1: Style conventions

Style	Description	Example
	This warning information 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.
	This warning information 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 restore business.
	This indicates warning information, supplementary instructions, and other content that the user must understand.	 Notice: Take the necessary precautions to save exported data containing sensitive information.
	This indicates supplemental instructions, best practices, tips, and other content that is good to know for the user.	 Note: You can use Ctrl + A to select all files.
>	Multi-level menu cascade.	Settings > Network > Set network type
Bold	It is used for buttons, menus, page names, and other UI elements.	Click OK .
Courier font	It is used for commands.	Run the <code>cd / d C :/ windows</code> command to enter the Windows system folder.
<i>Italics</i>	It is used for parameters and variables.	<code>bae log list --instanceid <i>Instance_ID</i></code>
[] or [a b]	It indicates that it is an optional value, and only one item can be selected.	<code>ipconfig [-all -t]</code>

Style	Description	Example
<code>{}</code> or <code>{a b}</code>	It indicates that it is a required value, and only one item can be selected.	<code>swich {stand slave}</code>

Contents

Legal disclaimer.....	I
Generic conventions.....	I
1 What is ECS?.....	1
2 Benefits of ECS.....	6
3 Scenarios.....	11
4 Regions and zones.....	13
5 Limits.....	17
6 Usage notes.....	26

1 What is ECS?

Elastic Compute Service (ECS) is a high-performance, stable, reliable, and scalable IaaS-level service provided by Alibaba Cloud. It is more convenient and efficient compared with physical servers. You can immediately acquire ECS instances and scale computing resources on-demand. ECS provides a variety of instance types that suit various business needs and help boost business growth.

Why ECS?

ECS provides the following benefits:

- You do not have to purchase hardware or construct data centers up front.
- Instances are delivered within minutes, enabling rapid deployment and reducing time to market.
- You can make use of resources in data centers and BGP grids around the world.
- You can scale and release resources based on actual business needs.
- Provides heterogenous computing servers such as GPU, FPGA, bare metal servers, and general x86 servers.
- You can use ECS to access other Alibaba Cloud services over the internal network, reducing traffic costs.
- Provides a host of security solutions, such as virtual firewalls, permission control, internal network isolation, virus protection, and traffic throttling.
- Provides a performance monitoring framework and proactive O&M systems.
- Provides standardized APIs to improve ease of use and applicability.

For more benefits, see [Benefits of ECS](#) and [Scenarios](#).

Architecture

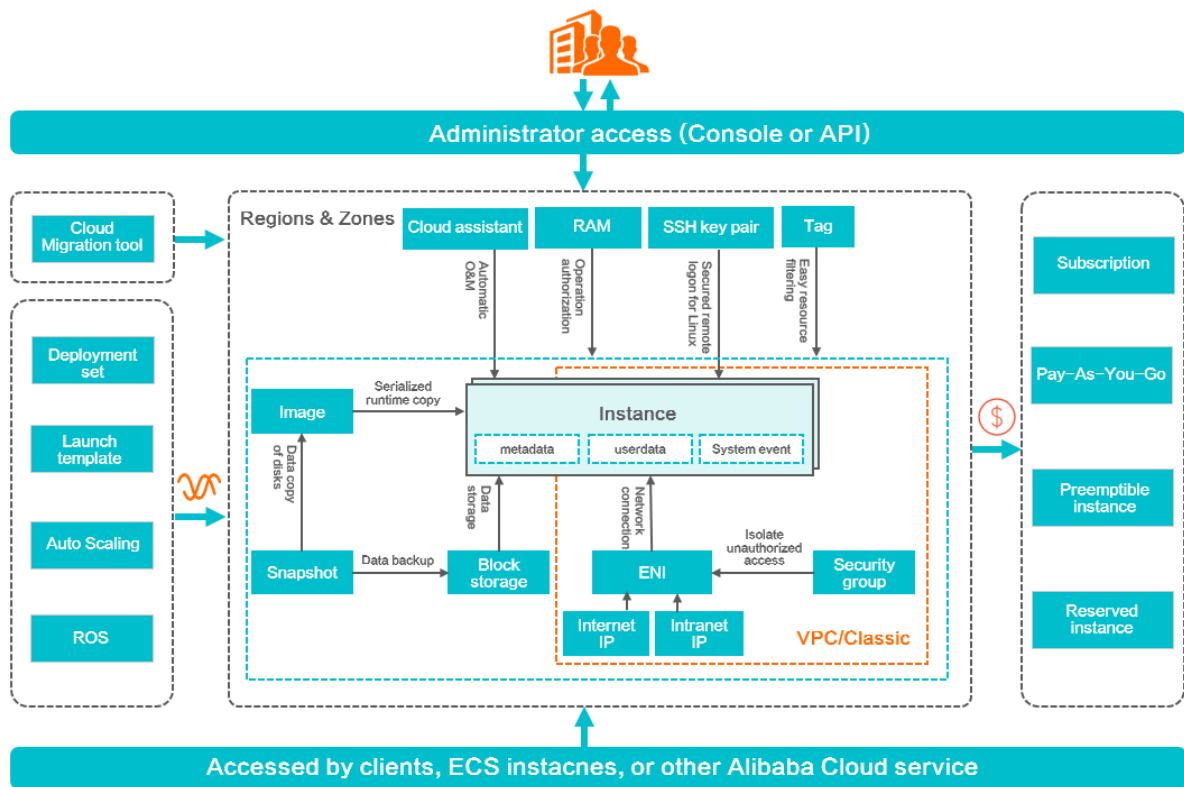
ECS comprises the following major components:

- **Instance**: a virtual computing environment that includes basic computing components such as CPU, memory, operating system, bandwidth, and disks. The computing performance, memory specifications, and applicable scenarios of an instance are determined by its **instance type**. The performance of an instance is determined by its vCPUs, memory capacity, and network performance.

- **Image**: provides the operating system, initial application data, and pre-installed software for instances. Multiple Linux operating systems and Windows Server operating systems are supported.
- **Block storage**: a block storage device which features high performance and low latency. Distributed storage architecture-based **cloud disks**, and physical storage-based **local disks** can be used.
- **Snapshot**: a stateful data file of a cloud disk or shared block storage at a certain point in time. It is often used to back up and restore data, and to create custom images.
- **Security group**: a logical group of instances located in the same region that have the same security requirements and require access to each other. A security group works as a virtual firewall for the ECS instances inside it.
- **Network**:
 - **VPC**: a logically isolated private cloud network. You can configure a private IP address range, a route table, and a gateway for a VPC.
 - **Classic network**: all classic instances are primarily built on a shared infrastructure network of Alibaba Cloud, which is planned and managed in a centralized manner.

For more information, see [Elastic Compute Service](#).

The following figure shows the architecture of ECS components. For more information about the functional components in the figure, see the related documentation.



Pricing

ECS supports multiple billing methods, such as subscription, pay-as-you-go, reserved instances (beta phase), and preemptible instances. For more information, see [Pricing overview](#) and [Product pricing](#).

Management tools

After registering an Alibaba Cloud account, you can create, use, or release an ECS instance in any region by using the following methods provided by Alibaba Cloud:

- ECS console: a Web service page used to manage ECS instances. For more information, see [Quick reference](#).
- ECS APIs: RPC APIs that support GET and POST requests. For more information, see [API Reference](#). The following developer tools can be used to call ECS APIs:
 - **CLI**: a flexible and scalable management tool based on Alibaba Cloud APIs. You can use CLI to encapsulate Alibaba Cloud native APIs to develop custom features.
 - **OpenAPI Explorer**: allows you to retrieve APIs, call APIs, and dynamically generate SDK example code.
 - **Alibaba Cloud SDK**: provides SDKs for multiple programming languages such as Java, Python, and PHP.

- **Resource Orchestration Service:** automatically creates and configures Alibaba Cloud resources based user-defined templates.
- **Terraform:** an open-source tool that uses configuration files to call computing resources on Alibaba Cloud and other platforms that support Terraform. Terraform also implements version control.

Deployment tips

Before you purchase an ECS instance, consider the following factors:

- **Regions and zones**

A region represents an Alibaba Cloud data center. The region and zone determine the physical location of an ECS instance. After an instance is created, its **metadata** is established and its region cannot be changed. Select a region and zone based on the target geographical location, availability of Alibaba Cloud services, application availability requirements, and whether internal network communication is required. For example, if you want to access both ECS and **ApsaraDB for RDS**, the RDS instance and ECS instance must be in the same region. For more information, see **Regions and zones**.

- **High availability**

To ensure business consistency and continuity, we recommend that you use **snapshots** to back up data, and use multi-zone deployment, **deployment sets**, and **Server Load Balancer (SLB)** for disaster recovery.

- **Network planning**

We recommend that you use **VPC** to plan your own private IP addresses. VPC supports all new features and new instance types. It also supports business system isolation and multi-region system deployment.

- **Security solutions**

You can use ECS security groups to control inbound and outbound access policies and the port listening status of ECS instances. For applications deployed on ECS, you can use free **Anti-DDoS Basic**. Alibaba Cloud Security is also available.

- For example, you can use Anti-DDoS Pro to ensure the stability and reliability of source sites. For more information, see **Anti-DDoS Pro documentation**.
- Server Guard can be used to safeguard the security of ECS instances. For more information, see **Server Guard documentation**.

Learning path

[ECS Learning Path](#) provides a sequence of courses to help you master the basics of ECS and gain advanced knowledge.

Related services

Together with ECS, you can select the following Alibaba Cloud services:

- [Auto Scaling](#) automatically adjusts the number of ECS instances based on business and policy changes.
- [Container Service for Kubernetes \(ACK\)](#) manages application lifecycles on groups of ECS instances.
- [Server Load Balancer \(SLB\)](#) distributes traffic among multiple ECS instances.
- [CloudMonitor](#) develops monitoring solutions for instances, system disks, and public network bandwidth.
- [ApsaraDB for RDS](#) provides database services accessible over internal networks to ECS instances, reduces network latency, access fees, and delivers top-notch performance. ApsaraDB for RDS supports multiple database engines, including MySQL, SQL Server, PostgreSQL, PPAS, and MariaDB.
- [Alibaba Cloud Marketplace](#): a platform to purchase software infrastructure, business software, website construction, hosted O&M, security, data & APIs, and solutions provided by third-party partners. You can also provide software programs and services as a service provider. For more information, see [Alibaba Cloud Marketplace documentation](#).

For more solutions, see [Alibaba Cloud Solutions](#).

2 Benefits of ECS

ECS provides better availability, security, and elasticity than integrated data centers (IDCs) and cloud server providers.

High Availability

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

Each Alibaba Cloud region has multiple zones. You can create active/standby or active/active ECS instances in multiple zones to achieve higher availability. You can build fault tolerant systems across multiple regions and zones to implement a financial-grade solution that spans three data centers across two regions. Alibaba Cloud provides mature solutions for fault tolerant services such as disaster recovery.

The Alibaba Cloud framework allows you to seamlessly switch between services. For more information about industry solutions, see [Solutions by industry](#). Alibaba Cloud industry solutions support a variety of services, such as finance, e-commerce, and video services.

Alibaba Cloud provides you with the following support services:

- Products and services for availability improvement, such as Elastic Compute Service (ECS), Server Load Balancer (SLB), Relational Database Service (RDS), and Data Transmission Service (DTS).
- Industry partners and ecosystem partners that help you build a more advanced and stable architecture and guarantee service continuity.
- Diverse training services that help you achieve high availability from businesses to underlying services.

Security

Alibaba Cloud has passed a host of international information security certifications, such as ISO 27001 and MTCS, which demands strict confidentiality of user data and user information, as well as user privacy protection.

We recommend that you use ECS in a [Virtual Private Cloud](#) (VPC). VPCs provide a stable, secure, controllable network environment that can be delivered in a short

period of time. The capability and architecture of VPC hybrid cloud bring the technical advantages of cloud computing to enterprises in traditional industries that have not implemented cloud computing.

- Breadth of network products

You only need to perform a simple configuration to connect your business environment to global IDCs, making your business more flexible, stable, and extensible.

- Interconnection with your IDC

You can use Express Connect to connect Alibaba Cloud VPC to your IDC to build a hybrid cloud architecture. You can use a variety of hybrid cloud architectures to provide network services and robust networking.

- Stability

After constructing your VPC, you can update your network architecture and obtain new network functions daily to constantly evolve your network infrastructure and ensure your business is always running steadily.

- Security

VPC features traffic isolation and attack isolation to protect your services from cyber attacks. After you build your business in a VPC, the first line of defense is established.

Elasticity

Elasticity is a key benefit of cloud computing. Alibaba Cloud is capable of providing IT resources required by a medium-sized Internet enterprise within a few minutes. In this way, most enterprises that build business on the cloud can process huge business volumes.

Alibaba Cloud provides elastic computing, storage, networking, and business architecture planning and allows you to combine your businesses as needed.

- Elastic computing

- Vertical scaling

Vertical scaling is the process where the configurations of an ECS instance are modified. After you purchase an ECS instance or storage capacity from Alibaba Cloud, you can configure the instance based on your transaction volume, whereas it may be difficult to change the configurations of a server in

a traditional IDC. For more information about vertical scaling, see [Overview of configuration changes](#).

- Horizontal scaling

Horizontal scaling allows the re-division of resources between applications. A traditional IDC may not be able to immediately provide sufficient resources for online gaming or live video streaming applications during peak hours. The elasticity of cloud computing makes it possible to provide the resources required during peak hours. When the load returns to normal levels, you can release unnecessary resources to reduce operation costs. The combination of ECS vertical and horizontal elasticity enables you to scale resources up and down by specific quantities as scheduled or against business load. For more information about horizontal scaling, see [#unique_30](#).

- Elastic storage

In a traditional IDC, you must add servers to increase the storage space. However, the number of servers that you can add is limited. Alibaba Cloud provides unlimited storage capacity and allows you to order as much storage space based on the business requirements. For more information about elastic storage, see [Resize a disk](#).

- Elastic network

Alibaba Cloud provides network elasticity. When purchasing Alibaba Cloud VPCs, you can configure the VPCs in the same way as the IDCs. In addition, VPCs have the following benefits: interconnection between data centers, separate secure domains in data centers, and flexible network configurations and planning within a VPC. For more information about elastic networks, see [Virtual Private Cloud](#).

Comparison between ECS and traditional IDCs

The following table lists the benefits of ECS compared with traditional IDCs.

Item	ECS	Traditional IDC
Equipment room deployment	Provides independently developed DC-powered servers with low PUE.	Provides traditional AC-powered servers with high PUE.
	Provides backbone equipment rooms with high outbound bandwidth and dedicated bandwidth.	Provides equipment rooms with various quality levels and shared bandwidth primarily, difficult for users to choose from.

Item	ECS	Traditional IDC
	Provides multi-line BGP equipment rooms, enabling smooth and balanced access among regions.	Provides equipment rooms with a single or dual line primarily.
Ease of operation	Provides mainstream OSs, including activated Windows OS.	Requires users to purchase and install OSs manually.
	Easily switches between OSs online.	OSs have to be manually reinstalled.
	Provides a Web-based console for online management.	Users must manually perform management and maintenance operations.
	Provides mobile phone verification for password setting, increasing data security.	Brings difficulty in resetting passwords, and exposes high risk of password cracking.
Disaster recovery and backup	Stores three copies of each piece of data. When one copy is corrupted, the data can be quickly restored.	Users must build a disaster recovery environment by themselves, and use traditional storage devices.
	Users can customize automatic snapshot policies to create automatic snapshots for data recovery.	Does not support automatic recovery because the snapshot function is not provided.
	Hardware failures can be recovered quickly and automatically.	Users must restore corrupted data manually.
Security and reliability	Effectively prevents MAC spoofing and ARP attacks.	Fails to prevent MAC spoofing and ARP attacks.
	Effectively defends against DDoS attacks by using black holes and traffic scrubbing.	Needs additional costs for devices for traffic scrubbing and black hole shielding systems.
	Provides additional services, such as port scanning, trojan scanning, and vulnerability scanning.	Typically encounters problems such as port scanning, trojan scanning, and vulnerability scanning.
Flexible scalability	Activates cloud servers on demand and upgrades configurations online.	Needs a long time for server delivery.

Item	ECS	Traditional IDC
	Adjusts outbound bandwidth as required.	Requires one-off purchase of outbound bandwidth that cannot be adjusted.
	Combines with Server Load Balancer online, enabling scaling up applications quickly and easily .	Uses hardware-based server load balancing, which is expensive and difficult to set up.
Cost effectiveness	Low cost.	High cost.
	Small up-front investment.	Large up-front investment , causing serious waste of resources.
	Provides pay-as-you-go and flexible payment options to allow you to flexibly respond to business changes.	Requires users to purchase up front to meet configuration requirements during peak hours.

3 Scenarios

ECS is a highly flexible solution that can be used independently as a simple web server or an application server, or used with other Alibaba Cloud products to deliver advanced solutions.



Note:

Typical scenarios of ECS are included but not limited to the applications described in this topic.

Official websites and simple web applications

A new website has low traffic and requires only low-configuration ECS instances to run applications, databases, storage files, and other resources. As your business expands, you can upgrade the ECS configuration or add ECS instances at any time without worrying about resource insufficiency during peak traffic.

Multimedia and high-traffic apps or websites

ECS can be used with [OSS](#) to store static images, videos, and downloaded packages, reducing the storage fees. In addition, it can work with [CDN](#) and [Load Balancer](#) to significantly reduce your waiting time, cut down public bandwidth costs, and improve service availability.

Databases with high I/O requirements

ECS supports databases with high I/O requirements, such as OLTP and NoSQL databases. A high-configuration I/O-optimized ECS instance can be used with the [#unique_37](#) to achieve high level of I/O concurrency with higher data reliability. Alternatively, multiple lower-configuration I/O-optimized ECS instances can be used with Server Load Balancer (SLB) to deliver a high availability architecture.

Apps and websites with sharp traffic fluctuations

Some applications may experience sharp 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 way, you can meet the changing resource requirements at a low cost. ECS can also work with SLB to deliver a high availability architecture.

Big data and real-time online and offline analysis

ECS allows you to use [big data instance type families](#) to support Hadoop distributed computing, log processing, and large data warehouses. By adopting the local storage architecture, the big data instance type families can deliver better network performance for Hadoop and Spark clusters while providing abundant storage space with higher storage performance.

AI applications

ECS allows you to use [compute optimized and GPU-equipped instance type families](#) to build AI applications with TensorFlow. Such instance type families have lower requirements on the computing capacity of clients and are suitable for image processing, and real-time online rendering for cloud gaming and AR/VR applications.

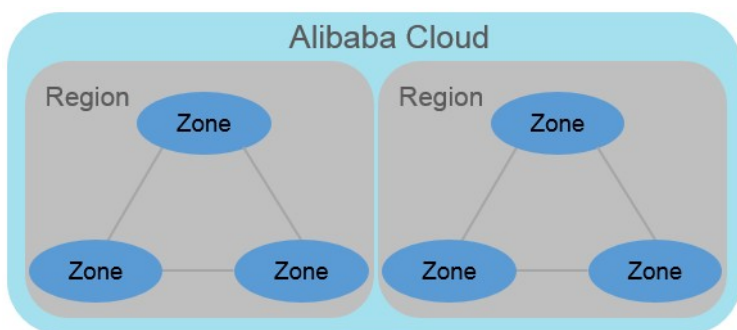
Customer case

For more scenarios of ECS, see [Alibaba Cloud solutions](#).

4 Regions and zones

This topic provides a complete list of Alibaba Cloud regions and zones.

Regions in Alibaba Cloud are independent. Zones in Alibaba Cloud are completely isolated. However, zones in the same region are connected through links with low latency. The following figure shows the relationship between regions and zones in Alibaba Cloud.



Regions

Alibaba Cloud regions are physical locations (data centers) that spread all over the world to reduce the network latency. The region cannot be changed once a resource is created in it. The following table lists all Alibaba Cloud regions, corresponding cities, and Region IDs.



Note:

The available regions vary according to the product. To view the list of regions available for each product, see [Alibaba Cloud Global Infrastructure](#).

• Regions in Mainland China

Region name	City	Region ID	Number of zones
China North 1	Qingdao	cn-qingdao	2
China North 2	Beijing	cn-beijing	8
China North 3	Zhangjiakou	cn-zhangjiakou	2
China North 5	Hohhot	cn-huhehaote	2
China East 1	Hangzhou	cn-hangzhou	8
China East 2	Shanghai	cn-shanghai	6
China South 1	Shenzhen	cn-shenzhen	5

Region name	City	Region ID	Number of zones
China SW 1	Chengdu	cn-chengdu	1

- Regions outside Mainland China

Region name	City	Region ID	Number of zones
China (Hong Kong)	Hong Kong	cn-hongkong	2
Asia Pacific SE 1	Singapore	ap-southeast-1	3
Asia Pacific SE 2	Sydney	ap-southeast-2	2
Asia Pacific SE 3	Kuala Lumpur	ap-southeast-3	2
Asia Pacific SE 5	Jakarta	ap-southeast-5	2
Asia Pacific SOU 1	Mumbai	ap-south-1	2
Asia Pacific NE 1	Tokyo	ap-northeast-1	2
US West 1	Silicon Valley	us-west-1	2
US East 1	Virginia	us-east-1	2
EU Central 1	Frankfurt	eu-central-1	2
UK (London)	London	eu-west-1	2
Middle East 1	Dubai	me-east-1	1

When choosing a region, consider the following factors:

- Geographic location

Choose a region based on the geographic locations of you and your target users.

- Regions in Mainland China

In general, we recommend that you choose the data center closest to your target users so that they have fast access. Alibaba Cloud's data centers of the regions in Mainland China are similar to each other in terms of infrastructure, BGP network quality, service quality, ECS operation and configuration. The regions offer BGP backbone network lines covering all provinces, autonomous regions, and municipalities to guarantee fast, stable access within Mainland China.

- Regions outside Mainland China

Regions outside Mainland China offer access at international bandwidth, targeting areas outside Mainland China. Users in Mainland China that access services deployed in these regions may experience high latency. Therefore, we

do not recommend deploying services in regions outside Mainland China if users from Mainland China require access.

- If you have business operations in China (Hong Kong) or Southeast Asia, choose the region China (Hong Kong), Asia Pacific SE 1, Asia Pacific SE 3, or Asia Pacific SE 5.
- If you have business operations in Japan and Korea, choose the Asia Pacific NE 1.
- If you have business operations in India, choose the region Asia Pacific SOU 1.
- If you have business operations in Australia, choose the region Asia Pacific SE 2.
- If you have business operations in America, choose the region US East 1 or US West 1.
- If you have business operations in Continental Europe, choose the region EU Central 1.
- If you have business operations in the Middle East, choose the Eastern Middle East 1 region.

- **Intranet communication between Alibaba Cloud products across regions**

Intranet communication between Alibaba Cloud products that are not in the same region is not supported, which means:

- ECS instances and other products, such as ApsaraDB for RDS and OSS instances, cannot communicate with one another over the intranet if they are in different regions.
- Server Load Balancer cannot be deployed for ECS instances in different regions.

- **Resource pricing**

The price of resources may vary according to the region. For details, see the [product pricing page](#).

- **Business license filing**

When you choose a region, consider the special requirements of some areas. If you have purchased an ECS instance in Mainland China and used it for a web server, complete business license filing first.

To complete business license filing, pay attention to the following points:

- If your company is located in Beijing, choose the China North 2 region for the ECS instance you have purchased.
- If your company is located in Guangdong, choose the China South 1 region for the ECS instance you have purchased.



Note:

The communications administrations in different provinces, autonomous regions, and municipalities have different approval requirements for business license filing. In case of any change, the information published on the business license filing website of your local administration prevails.

Zones

Zones are physical areas with independent power grids and networks within one region. The network latency for instances within the same zone is lower.

Intranet communication can take place between instances in different zones of the same region, and fault isolation can be achieved between zones. Whether or not instances can be deployed in the same zone depends on the requirements for disaster recovery and network latency.

- If your applications require high disaster recovery capabilities, we recommend that you deploy your instances in different zones of the same region.
- If your applications require low network latency between instances, we recommend that you create your instances in the same zone.

For more information about regions and zones, see the [Alibaba Cloud Global Infrastructure](#).

5 Limits

This topic describes service limits of ECS and how to request exceptions to the limits.

Overview

ECS has the following service limits:

- ECS does not allow you to install virtualization software, such as VMware Workstation, or use it for re-virtualization. Only [ECS Bare Metal Instances](#) and [Super Computing Clusters](#) support re-virtualization.
- ECS does not support sound card applications.
- External hardware devices, such as hardware dongles, USB flash drives, external hard disks, and hardware tokens, cannot be attached. Software verification methods such as two-factor authentication and dynamic passwords can be used.
- ECS does not support IP address translation services such as SNAT. Instead, it achieves this function by using VPNs and proxies.
- ECS does not support multicast protocols. We recommend that you use unicast instead.
- Log Service does not support 32-bit Linux ECS instances. For information about regions that support Log Service, see [Service endpoint](#). For information about operating systems that support Log Service, see the [Overview](#) section about Logtail collection.

View quotas

On the Overview page of the ECS console, click Privileges & Quotas and select a region to view the resource quotas in that region. If the quota of a resource cannot meet your needs, submit a ticket to request an increase for the quota. For more information, see [#unique_44](#). You can also view resource quotas by calling an API operation. For more information, see [#unique_45](#).

Instance limits

Item	Limit	Request an exception to a limit
Permissions to create instances	To create ECS instances in any Mainland China regions, you must first complete real-name authentication.	No exceptions
Instance types for which you can create pay-as-you-go instances	Instance types with less than 16 vCPUs.	Submit a ticket
Total number of vCPUs of pay-as-you-go instances in each region for an account	50 vCPUs.	Submit a ticket
Total number of vCPUs of preemptible instances in each region for an account	0 vCPUs. You can submit a ticket to request the quota of 50 vCPUs.	Submit a ticket
Total number of launch templates in each region for an account	30.	No exceptions
Total number of versions of a launch template	30.	No exceptions
Switch from pay-as-you-go to subscription	Not supported by the following instance types (families): t1, s1, s2, s3, c1, c2, m1, m2, n1, n2, and e3.	No exceptions
Switch from subscription to pay-as-you-go	<ul style="list-style-type: none"> Whether this feature is supported depends on your ECS instance resource usage. 5,000 vCPU hours each month. A maximum monthly refund limit is set. For more information about the refund limit, see the Switch to Pay-As-You-Go page. 	No exceptions

Block Storage limits

Item	Limit	Request an exception to a limit
Permissions to create pay-as-you-go disks	Complete real-name authentication before creating disks in any Mainland China regions.	No exceptions
Total number of all pay-as-you-go disks in all regions for an account	Number of pay-as-you-go instances in all regions for an account \times 5. Each account can create at least 10 pay-as-you-go disks.	Submit a ticket
Total number of system disks for each instance	1.	No exceptions
Maximum number of data disks for an instance	16 (including disks and Shared Block Storage devices).	No exceptions
Number of instances to which a Shared Block Storage device can be attached	8.	No exceptions
Total number of Shared Block Storage devices in all regions for an account	10.	Submit a ticket
Capacity of a basic disk	5 GiB to 2,000 GiB.	No exceptions
Capacity of an SSD	20 GiB to 32,768 GiB.	No exceptions
Capacity of an ultra disk	20 GiB to 32,768 GiB.	No exceptions
Capacity of a local SSD	5 GiB to 800 GiB.	No exceptions
Total capacity of all local SSDs for an instance	1,024 GiB.	No exceptions
Capacity of a local NVMe SSD	1,456 GiB.	No exceptions
Total capacity of all local NVMe SSDs for an instance	2,912 GiB.	No exceptions
Capacity of a local SATA HDD	5,500 GiB.	No exceptions
Total capacity of all local SATA HDDs for an instance	154,000 GiB.	No exceptions
Capacity of an SSD Shared Block Storage device	32,768 GiB.	No exceptions

Item	Limit	Request an exception to a limit
Total capacity of all SSD Shared Block Storage devices for an instance	128 TiB.	No exceptions
Capacity of an ultra Shared Block Storage device	32,768 GiB.	No exceptions
Total capacity of all ultra Shared Block Storage devices for an instance	128 TiB.	No exceptions
Capacity of an enhanced SSD	32,768 GiB.	No exceptions
Capacity of a system disk	<ul style="list-style-type: none"> · Windows: 40 GiB to 500 GiB. · Linux (excluding CoreOS) and FreeBSD: 20 GiB to 500 GiB. · CoreOS: 30 GiB to 500 GiB. 	No exceptions
Capacity of a data disk	<ul style="list-style-type: none"> · Basic disk: 5 GiB to 2,000 GiB. · SSD, ultra disk, SSD Shared Block Storage device, or ultra Shared Block Storage device: 20 GiB to 32,768 GiB. · Local disk: dependent on specific disks. 	No exceptions
Attach new local disks to instances that have local disks	Not supported.	No exceptions
Change the specifications of instances that have local disks	Only bandwidth can be changed.	No exceptions
System disk mount points	/dev/xvda.	No exceptions
Data disk mount points	/dev/xvd[b-z].	No exceptions

**Note:**

Block storage capacity is measured in binary units. 1 KiB equals 1,024 Bytes. 1 GiB equals 1,024 MiB.

Snapshot limits

Item	Limit	Request an exception to a limit
Number of snapshots that can be created for each disk and Shared Block Storage device	64	No exceptions
Number of automatic snapshot policies that can be created in a region for an account	100	No exceptions

Image limits

Item	Limit	Request an exception to a limit
Total number of custom images in each region for an account	100.	Submit a ticket
Maximum number of users with whom an image can be shared	50.	Submit a ticket
Images supported by instance types	Instance types that have 4 GiB or more memory do not support 32-bit images.	No exceptions

SSH key pair limits

Item	Limit	Request an exception to a limit
Total number of SSH key pairs in each region for an account	500	No exceptions
Instance types supporting SSH key pairs	All instance types except non-I/O optimized instance types in Generation I	No exceptions
Images supporting SSH key pairs	Linux images only	No exceptions

Public bandwidth limits

Item	Limit	Request an exception to a limit
Maximum inbound public bandwidth	200 Mbit/s.	No exceptions
Maximum outbound public bandwidth	<ul style="list-style-type: none"> · Subscription: up to 200 Mbit/s. · Pay-as-you-go: up to 100 Mbit/s. 	No exceptions
Change the assigned public IP address for an instance	Public IP addresses can be changed within six hours after the instance is created, and can be changed a maximum of three times.	No exceptions

Security group limits

Item	Limit on basic security groups	Limit on advanced security groups
Total number of security groups in each region for an account	100.	Same as that on basic security groups.
Number of classic network-type ECS instances that can be added to a classic network-type security group	1,000.*	Classic network is not supported.
Number of VPC-type ECS instances that can be added to a VPC-type security group	Depends on the number of private IP addresses in a VPC-type security group .	No limit.
Number of security groups that can contain the same ECS instance	5.	Same as that on basic security groups.
Number of security groups that can contain the same ENI of an ECS instance	You can submit a ticket to raise the limit to 10 or 16 security groups.	

Item	Limit on basic security groups	Limit on advanced security groups
Maximum number of rules (both inbound and outbound) in a security group	100. ^{***}	Same as that on basic security groups.
Maximum number of rules (both inbound and outbound) in all security groups that contain the same ENI	500.	Same as that on basic security groups.
Number of private IP addresses in a VPC-type security group	2,000. ^{**}	No limit.
Number of ENIs in a VPC-type security group	Depends on the maximum number of rules in a security group .	50,000.
Public network access port	The default STMP port for outbound traffic is port 25, which is disabled by default. It cannot be enabled by security group rules.	Same as that on basic security groups.

* If more than 1,000 classic network-type instances need access to each other, you can distribute them across multiple security groups and authorize mutual access.

** If you require mutual access among more than 2,000 private IP addresses, you can distribute them across multiple security groups and authorize mutual access.

*** If you increase the number of security groups that can contain the same ECS instance, the maximum number of rules in a security group will decrease. The product of the number of security groups that can contain the same ECS instance and the maximum number of rules (both inbound and outbound) in a security group must be less than or equal to 500. For example, $5 \times 100 = 500$, $10 \times 50 = 500$, and $16 \times 30 < 500$.

Deployment set limits

Item	Limit	Request an exception to a limit
Total number of deployment sets in each region for an account	2.	No exceptions
Total number of instances that can be included in a deployment set	Seven instances are allowed in a zone. The maximum number of instances allowed in a region is $7 \times$ number of zones.	No exceptions
Instance types that can be created in a deployment set	c5, g5, hfc5, hfg5, r5, se1ne, sn1ne, and sn2ne.	No exceptions

Cloud Assistant limits

Item	Limit	Request an exception to a limit
Total number of Cloud Assistant commands in each region for an account	100	Submit a ticket
Total number of Cloud Assistant commands that can be used by an account each day in a region	5,000	Submit a ticket

ENI limits

Item	Limit	Request an exception to a limit
Total number of ENIs in each region for an account	100	Submit a ticket

Tag limits

Item	Limit	Request an exception to a limit
Total number of tags that can be bound to an instance	20	No exceptions

API limits

Item	Limit	Request an exception to a limit
Total number of CreateInstance calls	200 calls per minute	Submit a ticket

**Note:**

For more information about the VPC limits, see [#unique_51](#).

6 Usage notes

This topic describes the general limits for the usage of ECS instances and limits that are specific to the installed OS.

General limits

- After you create an ECS instance, only you have administrator permissions on this instance.
- You are not allowed to resell or sublicense any bandwidth assigned to your ECS instances without authorization. Failure to comply will result in a suspension or termination of your right to use any portions or all Alibaba Cloud services.
- Do not use your ECS instances for malicious, fraudulent, or unlawful activities, such as click farming or fraudulent transactions, on e-commerce websites such as Taobao. Failure to comply will result in suspension or termination of your account.
- Do not uninstall relevant hardware drivers.
- Do not modify the MAC address of ENIs unless necessary.
- Do not enable SELinux.
- An ECS instance with more than 4 GiB RAM must use a 64-bit operating system (OS). The 64-bit OSs currently supported are as follows:
 - Aliyun Linux 64-bit
 - CoreOS 64-bit
 - CentOS 64-bit
 - Debian 64-bit
 - FreeBSD 64-bit
 - OpenSUSE 64-bit
 - SUSE Linux 64-bit
 - Ubuntu 64-bit
 - Red Hat 64-bit
 - Windows 64-bit

For the latest available OSs, go to the instance purchase page in the ECS console.

- To guarantee service continuity, we recommend that you enable automatic startup upon instance boot for relevant software. If service applications are connected

to databases, we recommend that you enable automatic re-connection for these service applications.

- We recommend that you do not upgrade the kernel and the operating system. If you need to upgrade the kernel, see [How to avoid Linux instance startup failure after kernel upgrade](#).

Windows limits

- Do not stop the built-in AliyunService or shutdownmon.exe process. Otherwise, ECS instances may not be stopped or restarted properly.
- Do not modify the hostname of the domain controller.
- We recommend that you do not create custom images by using a virtual machine that acts as a domain controller.
- Do not rename, delete, or disable the administrator account.
- We recommend that you do not use the virtual memory if Basic Cloud Disks are used. For Ultra Cloud Disks, SSD Cloud Disks, or ESSD Cloud Disks, you can use the virtual memory as needed.
- Exercise caution when you use the administrator account to resize cloud disks, operate spanned volumes or the registry, and update the system. Failure to comply can result in data loss. For more information, see [Resize a cloud disk](#).
- Windows 32-bit OSs support up to four CPU cores.
- Ensure that a minimum of 2 GiB RAM is available when you build a website or deploy a Web environment on a Windows instance. Instances with 1-core CPU and 1 GiB RAM cannot start MySQL.
- For more information see [#unique_53](#).

Linux limits

- Do not modify the content of the default `/etc/issue` file on Linux instances. Otherwise, if you create a custom image of such an instance and then use the image to create a new instance, the new instance cannot start properly because the operating system edition cannot be recognized.
- Do not modify permissions of the directories in the root partition, especially `/etc`, `/sbin`, `/bin`, `/boot`, `/dev`, `/usr`, and `/lib`. Improper modifications of permissions may cause system errors.
- Do not rename, delete, or disable the root account.
- Do not compile the Linux kernel or perform any other operations on it.

- We recommend that you do not use the swap partition if Basic Cloud Disks are used . However, for Ultra Cloud Disks, SSD Cloud Disks, or ESSD Cloud Disks, you can use the swap partition as needed.
- Exercise caution when you use the root account to run `fiio`, `mkfs`, or `fsck` commands or resize cloud disks. Failure to comply can result in data loss.
- For more information see [#unique_53](#).