

Alibaba Cloud

Elastic Compute Service Product Introduction

Document Version: 20201030

Legal disclaimer

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

1. You shall download and obtain this document from the Alibaba Cloud website or other Alibaba Cloud-authorized channels, and use this document for your own legal business activities only. The content of this document is considered confidential information of Alibaba Cloud. You shall strictly abide by the confidentiality obligations. No part of this document shall be disclosed or provided to any third party for use without the prior written consent of Alibaba Cloud.
2. No part of this document shall be excerpted, translated, reproduced, transmitted, or disseminated by any organization, company or individual in any form or by any means without the prior written consent of Alibaba Cloud.
3. The content of this document may be changed because of product version upgrade, adjustment, or other reasons. Alibaba Cloud reserves the right to modify the content of this document without notice and an updated version of this document will be released through Alibaba Cloud-authorized channels from time to time. You should pay attention to the version changes of this document as they occur and download and obtain the most up-to-date version of this document from Alibaba Cloud-authorized channels.
4. This document serves only as a reference guide for your use of Alibaba Cloud products and services. Alibaba Cloud provides this document based on the "status quo", "being defective", and "existing functions" of its products and services. Alibaba Cloud makes every effort to provide relevant operational guidance based on existing technologies. However, Alibaba Cloud hereby makes a clear statement that it in no way guarantees the accuracy, integrity, applicability, and reliability of the content of this document, either explicitly or implicitly. Alibaba Cloud shall not take legal responsibility for any errors or lost profits incurred by any organization, company, or individual arising from download, use, or trust in this document. Alibaba Cloud shall not, under any circumstances, take responsibility for any indirect, consequential, punitive, contingent, special, or punitive damages, including lost profits arising from the use or trust in this document (even if Alibaba Cloud has been notified of the possibility of such a loss).
5. By law, all the contents in Alibaba Cloud documents, including but not limited to pictures, architecture design, page layout, and text description, are intellectual property of Alibaba Cloud and/or its affiliates. This intellectual property includes, but is not limited to, trademark rights, patent rights, copyrights, and trade secrets. No part of this document shall be used, modified, reproduced, publicly transmitted, changed, disseminated, distributed, or published without the prior written consent of Alibaba Cloud and/or its affiliates. The names owned by Alibaba Cloud shall not be used, published, or reproduced for marketing, advertising, promotion, or other purposes without the prior written consent of Alibaba Cloud. The names owned by Alibaba Cloud include, but are not limited to, "Alibaba Cloud", "Aliyun", "HiChina", and other brands of Alibaba Cloud and/or its affiliates, which appear separately or in combination, as well as the auxiliary signs and patterns of the preceding brands, or anything similar to the company names, trade names, trademarks, product or service names, domain names, patterns, logos, marks, signs, or special descriptions that third parties identify as Alibaba Cloud and/or its affiliates.
6. Please directly contact Alibaba Cloud for any errors of this document.

Document conventions









Style	Description	Example
 Danger	A danger notice indicates a situation that will cause major system changes, faults, physical injuries, and other adverse results.	 Danger: Resetting will result in the loss of user configuration data.
 Warning	A warning notice indicates a situation that may cause major system changes, faults, physical injuries, and other adverse results.	 Warning: Restarting will cause business interruption. About 10 minutes are required to restart an instance.
 Notice	A caution notice indicates warning information, supplementary instructions, and other content that the user must understand.	 Notice: If the weight is set to 0, the server no longer receives new requests.
 Note	A note indicates supplemental instructions, best practices, tips, and other content.	 Note: You can use Ctrl + A to select all files.
>	Closing angle brackets are used to indicate a multi-level menu cascade.	Click Settings > Network > Set network type .
Bold	Bold formatting is used for buttons, menus, page names, and other UI elements.	Click OK .
Courier font	Courier font is used for commands	Run the <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</code> <i>Instance_ID</i>
[] 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>

Table of Contents

1.What is ECS? -----	05
2.Benefits of ECS -----	08
3.Scenarios -----	12
4.Limits -----	13
5.Usage notes -----	22

1. What is ECS?

Elastic Compute Service (ECS) is a high-performance, stable, reliable, and scalable IaaS-level service provided by Alibaba Cloud. ECS eliminates the need to invest in IT hardware up front and allows you to quickly scale computing resources on demand, making ECS more convenient and efficient than physical servers. ECS provides a variety of instance types that suit different 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.
- Instances can be delivered within minutes, enabling rapid deployment and reducing time to market.
- You can make use of resources in data centers and BGP servers around the world.
- You can scale resources up or down based on your actual business needs at a transparent, easy-to-understand cost.
- x86 architecture-based ECS instances, ECS Bare Metal instances, and heterogeneous computing ECS instances such as GPU-accelerated and FPGA-accelerated instances are provided.
- You can use ECS to access other Alibaba Cloud services over the internal network, reducing Internet traffic costs.
- A host of security solutions such as virtual firewalls, role-based permission control, internal network isolation, virus protection, and traffic throttling are provided.
- ECS comes with a performance monitoring framework and active O&M system.
- A standardized API is provided to improve ease-of-use and applicability.

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

Architecture

ECS comprises the following major components:

- **Instance:** An ECS instance is a virtual server 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. Each instance type has particular specifications, including the number of vCPUs, memory capacity, and network performance.
- **Image:** Images provide operating systems, initial application data, and pre-installed software for instances. Multiple Linux distributions and Windows Server operating systems are supported.
- **Block Storage:** Block Storage devices offer high performance and low latency. ECS comes with distributed storage architecture-based cloud disks and physical storage-based local disks.
- **Snapshot:** A snapshot is a stateful data file of a cloud disk at a certain point in time. Snapshots are often used to back up and restore data or to create custom images.
- **Security group:** A security group is a logical grouping of instances located within 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:**
 - **Virtual private cloud (VPC):** A VPC is 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 network-type instances are built on a shared infrastructure network that is centrally planned and managed by Alibaba Cloud.

For more information, see the [Product page of Elastic Compute Service](#).

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



Pricing

ECS supports multiple billing methods such as subscription and pay-as-you-go and purchasing options such as reserved instances and preemptible instances. For more information, see [Billing overview](#) and the Pricing tab of the [Product page of Elastic Compute Service](#).

Management tools

After you register an Alibaba Cloud account, you can create, use, or release ECS instances in any region by using one of 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 API:** an RPC API that supports GET and POST requests. For more information, see [API Reference](#). The following developer tools can be used to call ECS API operations:
 - **Alibaba Cloud 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 API operations, call API operations, and dynamically generate SDK sample code.
 - **Alibaba Cloud SDK:** SDKs for a variety of programming languages such as Java, Python, and PHP are provided.
- **Resource Orchestration Service (ROS):** automatically creates and configures Alibaba Cloud resources based on user-defined templates.
- **Terraform:** an open-source tool that uses configuration files to call computing resources of Alibaba Cloud and other platforms that support Terraform. Terraform also implements version control.

Deployment suggestions

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

- **Region and zone**

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. You can obtain metadata only of the ECS instances located within VPCs. 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 ApsaraDB for RDS instance and ECS instance must be within 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.

- Set up network connections

We recommend that you use VPC to plan your own private IP addresses. VPC supports all new features and new instance types. VPC also supports business system isolation and multi-region system deployment. For more information, see [What is a VPC?](#)

- 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 instances, Anti-DDoS Basic are provided for free. 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 [What is Anti-DDoS Pro](#).
- Security Center can be used to safeguard the security of ECS instances. For more information, see [What is Security Center?](#)

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. For more information, see [What is Auto Scaling?](#)
- Dedicated Host (DDH): allows you to deploy ECS instances on a dedicated host to have dedicated use of its physical resources. DDH also allows you to migrate your businesses to the cloud at minimal costs while meeting compliance requirements. For more information, see [What is DDH?](#)
- Container Service for Kubernetes: manages application lifecycles on groups of ECS instances. For more information, see [What is Container Service for Kubernetes?](#)
- Server Load Balancer (SLB): distributes traffic among multiple ECS instances. For more information, see [What is Server Load Balancer?](#)
- CloudMonitor: develops monitoring solutions for instances, system disks, and public bandwidth. For more information, see [Overview](#).
- ApsaraDB for RDS: provides database services accessible over internal networks to ECS instances, reduces network latency and access fees, and delivers top-notch performance. ApsaraDB for RDS supports multiple database engines, including MySQL, SQL Server, PostgreSQL, PPAS, and MariaDB. For more information, see [What is ApsaraDB for RDS?](#)
- [Alibaba Cloud Marketplace](#): a platform where third-party partners provide software infrastructure, business software, and various software and services related to website construction, hosted O&M, security, data, APIs, and solutions. You can also provide software and services as a service provider in Alibaba Cloud Marketplace.

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 [What is Auto Scaling?](#)

- 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
	Provides independently developed DC-powered servers with low PUE.	Provides traditional AC-powered servers with high PUE.

Item	ECS	Traditional IDC
Equipment room deployment	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.
	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.
	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.

Item	ECS	Traditional IDC
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 [ESSD](#) 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.Limits

This topic describes the limits on ECS and how to apply for extensions on some limits.

Overview

ECS has the following limits:

- You cannot install virtualization software such as VMware Workstation or use ECS instances for secondary virtualization. Only ECS bare metal instances and Super Computing Clusters (SCCs) support secondary 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 directly attached to ECS instances. Software verification methods such as two-factor authentication and dynamic passwords can be used.
- ECS does not support multicast protocols. We recommend that you use unicast protocols instead.
- Log Service does not support 32-bit Linux ECS instances.

For information about the ECS instances that support Log Service, see [Logtail overview](#).

- To apply for ICP filings for websites that are deployed on your ECS instance, ensure that the instance meets ICP filing requirements. You can apply for a limited number of ICP filing service numbers for each ECS instance. For more information, see [Prepare and check the instance and access information](#). For information about how to apply for ICP filings, see [ICP备案流程概述](#).

View quotas

You can click **Privileges** on the **Overview** page of the ECS console, and select a region to view the resource usage and quotas in that region. If the quota for a resource type is insufficient, submit a ticket to request a quota increase. For more information about how to view privileges and quotas, see [View and increase quotas \(new version\)](#) or [DescribeAccountAttributes](#).


Instance limits


Item	Limit	How to apply for an extension on the limit
Permissions to create an ECS instance	To create an ECS instance within a region in mainland China, you must first complete real-name verification.	None.
Instance types that can be used to create pay-as-you-go instances	Instance types that have less than 16 vCPUs.	Submit a ticket.
Quota for vCPUs available for pay-as-you-go instances per region in an account	50 vCPUs	Submit a ticket.
Quota for vCPUs available for preemptible instances per region in an account	Submit a ticket to apply for a quota of up to 50 vCPUs.	Submit a ticket.

Item	Limit	How to apply for an extension on the limit
Quota for instance launch templates per region in an account	30	None.
Quota for versions for an instance launch template	30	None.
Permissions to change the billing method from pay-as-you-go to subscription	The following instance types or families do not support billing method changes: t1, s1, s2, s3, c1, c2, m1, m2, n1, n2, and e3.	None.
Permissions to change the billing method from subscription to pay-as-you-go	<ul style="list-style-type: none"> Whether you have these permissions depends on your ECS usage. 5,000 vCPUs × Number of hours per month. The billing method change from subscription to pay-as-you-go may result in a refund. Each account is limited to a maximum monthly refund amount. The actual refund details are displayed on the Switch to Pay-As-You-Go page in the ECS console. 	None.

Reserved instance limits

Item	Limit	How to apply for an extension on the limit
Quota for regional reserved instances in an account	20	Submit a ticket.
Quota for zonal reserved instances per zone in an account	20	Submit a ticket.

Item	Limit	How to apply for an extension on the limit
<p>Instance families that support reserved instances</p>	<p>Instance families that support reserved instances include:</p> <ul style="list-style-type: none"> • General purpose instance families: g6e, g6, g5, and sn2ne • Compute optimized instance families: c6e, c6, c5, ic5, and sn1ne • Memory optimized instance families: r6e, r6, r5, and se1ne • Big data instance family: d2s • Instance families with local SSDs: i2 and i2g • Instance families with high clock speed: hfc6, hfc5, hfg6, hfg5, and hfr6 • GPU-accelerated compute optimized instance families: gn6i and gn6e • ECS Bare Metal Instance families: ebmc6, ebmg6, ebmr6, ebmhfc6, ebmhfg6, and ebmhfr6 • Burstable instance families: t6 and t5 <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin-top: 10px;"> <p> Note The t5 instance family supports only zonal reserved instances.</p> </div>	<p>None.</p>


 **Note** For more information, see the "Limits" section in [Reserved instance overview](#).

Elastic Block Storage limits

Item	Limit	How to apply for an extension on the limit
<p>Permissions to create a pay-as-you-go disk</p>	<p>You must complete real-name verification before you can create a disk within a region in mainland China.</p>	<p>None.</p>

Item	Limit	How to apply for an extension on the limit
Quota for pay-as-you-go disks in all regions for an account	This quota is calculated by using the following formula: Number of ECS instances across all regions × 5. A minimum of ten pay-as-you-go disks can be created in each account. If only one instance exists in your account, you can create a maximum of ten pay-as-you-go disks in all regions, instead of five pay-as-you-go disks.	Submit a ticket.
Quota for the capacity of pay-as-you-go disks that are used as data disks in an account	This limit is subject to ECS resource usage, region, and disk category. You can go to the Privileges & Quotas page in the ECS console to view this information. For more information, see View and increase quotas (new version) .	Submit a ticket.
Quota for system disks on an instance	1	None.
Quota for data disks on an instance	16.	None.
Capacity of a basic disk	5 GiB to 2,000 GiB.	None.
Capacity of a standard SSD	20 GiB to 32,768 GiB.	None.
Capacity of an ultra disk	20 GiB to 32,768 GiB.	None.
Capacity of an enhanced SSD (ESSD)	20 GiB to 32,768 GiB.	None.
Capacity of a local SSD	5 GiB to 800 GiB.	None.
Total capacity of all local SSDs on an instance	1,024 GiB.	None.
Capacity of a system disk	<ul style="list-style-type: none"> Windows Server: 40 GiB to 500 GiB CoreOS and FreeBSD: 30 GiB to 500 GiB Linux systems excluding CoreOS: 20 GiB to 500 GiB 	None.
Permissions to attach new local disks to instances that are equipped with local disks	Not allowed.	None.
Permissions to change configurations of instances that are equipped with local disks	Only bandwidth configurations of instances that are equipped with local disks can be changed.	None.
Mount points of system disks	/dev/vda	None.

Item	Limit	How to apply for an extension on the limit
Mount points of data disks	/dev/vd[b-z]	None.

 **Note** Block storage capacity is measured in binary units. In binary units, 1 KiB equals 1,024 bytes. Example: 1 GiB = 1,024 MiB.

Snapshot limits

Item	Limit	How to apply for an extension on the limit
Quota for user-created snapshots (also called manual snapshots) that can be created for each disk	256	None.
Quota for automatic snapshots that can be created for each disk	1,000	None.
Quota for automatic snapshot policies that can be created per region in an account	100	None.

Image limits

Item	Limit	How to apply for an extension on the limit
Quota for custom images that can be created per region in an account	100	Submit a ticket.
Quota for users to whom a single image can be shared	50	Submit a ticket.
Support of instance types for images	Instance types that have 4 GiB or larger memory do not support 32-bit images.	None.


SSH key pair limits

Item	Limit	How to apply for an extension on the limit
Quota for SSH key pairs per region in an account	500	None.
Instance types that support SSH key pairs	Non-I/O optimized instances of Generation I instance families do not support SSH key pairs	None.

Item	Limit	How to apply for an extension on the limit
Images that support SSH key pairs	Only Linux images	None.

Public bandwidth limits

Item	Limit	How to apply for an extension on the limit
Peak inbound bandwidth	<ul style="list-style-type: none"> If the purchased peak outbound bandwidth is less than or equal to 10 Mbit/s, Alibaba Cloud allocates an inbound bandwidth of 10 Mbit/s. If the purchased peak outbound bandwidth is greater than 10 Mbit/s, Alibaba Cloud allocates an inbound bandwidth that is equal to the purchased peak outbound bandwidth. 	None.
Peak outbound bandwidth	<ul style="list-style-type: none"> Subscription instance: 200 Mbit/s Pay-as-you-go instance: 100 Mbit/s 	None.
Change to the public IP address of an instance	The public IP address of an instance can be changed within six hours after the instance is created, and can be changed a maximum of three times.	None.

 **Notice** When the pay-by-traffic billing method is used, the peak inbound and outbound bandwidths indicate the upper limits for bandwidths and are only for reference. In the event of resource contention, the peak bandwidths cannot be guaranteed. If you want guaranteed bandwidths for your instance, use the pay-by-bandwidth billing method.

Security group limits

Item	Basic security group	Advanced security group
Quota for security groups that can be created per region in an account	100 ⁽¹⁾	Same as the limit on basic security groups.
Quota for classic network-type ECS instances that can be contained in a classic network-type security group	1,000 ⁽²⁾	The classic network is not supported.

Item	Basic security group	Advanced security group
Quota for VPC-type ECS instances that can be contained in a VPC-type security group	Depends on the number of private IP addresses that can be contained in the VPC-type security group.	No limits.
Quota for security groups to which an ECS instance can belong	5	Same as the limit on basic security groups.
Quota for security groups to which an ENI of an ECS instance can belong	You can submit a ticket to raise the limit to 10 or 16 security groups.	
Quota for rules (including inbound and outbound rules) in a security group	200 ⁽³⁾	Same as the limit on basic security groups.
Quota for rules (including inbound and outbound rules) in all security groups to which an ENI belongs	1,000	Same as the limit on basic security groups.
Quota for private IP addresses that can be contained in a VPC-type security group	2,000 ⁽⁴⁾	65,536.
Internet access port	The default SMTP port for outbound traffic is port 25, which is disabled by default. It cannot be enabled by security group rules.	Same as the limit on basic security groups.

- ⁽¹⁾ The following regions share the quota for security groups that can be created in each account: China (Hangzhou), China (Shanghai), China (Qingdao), China (Beijing), China (Shenzhen), China (Hong Kong), US (Silicon Valley), and Singapore (Singapore). A maximum of 100 security groups can be created in all of these regions in an account.
- ⁽²⁾ If more than 1,000 classic network-type instances need mutual access over the internal network, you can assign them to multiple security groups and authorize mutual access among these security groups.
- ⁽³⁾ If you increase the quota for security groups to which an ECS instance can belong, the quota for rules in each security group decreases. The product of the quota for security groups to which an ECS instance belongs and the quota for rules in each security group cannot exceed 1,000. For example, if the quota for security groups to which the ECS instance can belong is 5, 10, or 16, the corresponding quota for rules in each security group is 200, 100, or 60, which are verified by using the following formulas: $5 \times 200 = 1000$, $10 \times 100 = 1000$, and $16 \times 60 \leq 1000$.
- ⁽⁴⁾ If more than 2,000 private IP addresses need mutual access over the internal network, you can distribute the ECS instances of these private IP addresses to multiple security groups and authorize mutual access among these security groups.

Deployment set limits

Item	Limit	How to apply for an extension on the limit
Quota for deployment sets per region in an account	2	Submit a ticket.
Quota for instances that can be contained in a deployment set	A maximum of seven instances are allowed in each zone. The number of instances allowed in a deployment set within a region is calculated by using the following formula: 7 × Number of zones in the region.	None.
Instance families that support deployment sets	Instances in the following instance families can be created in deployment sets: c6, g6, r6, hfc6, hfg6, hfr6, d2, d2s, d2c, c5, d1, d1ne, g5, hfc5, hfg5, i2, i2g, i1, ic5, r5, se1ne, sn1ne, and sn2ne. For more information about instance types and their performance, see Instance families .	None.

Cloud Assistant limits

Item	Limit	How to apply for an extension on the limit
Quota for Cloud Assistant commands that can be created per region in an account	100	Submit a ticket.
Quota for Cloud Assistant commands that can be run each day by an account within a region	5000	Submit a ticket.

ENI limits


Item	Limit	How to apply for an extension on the limit
Quota for elastic network interfaces (ENIs) per region in an account	100	Submit a ticket.

Tag limits

Item	Limit	How to apply for an extension on the limit
Quota for tags that can be bound to an instance	20	None.

API

Item	Limit	How to apply for an extension on the limit
Quota for calls to the CreateInstance operation	200 calls per minute	Submit a ticket.

 **Note** For more information about the limits on VPCs, see [Limits](#).

5.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:
 - Alibaba Cloud 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 disks,

standard SSDs, or ESSDs, 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 [Select an image](#).

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 disks, standard SSDs, or ESSDs, 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 [Select an image](#).