

Alibaba Cloud Elastic Compute Service

User Guide

Issue: 20181219

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.	 Note: 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 Instance_ID</code>
[] or [a b]	It indicates that it is an optional value, and only one item can be selected.	<code>ipconfig [-all -t]</code>
{ } or {a b}	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 Quick reference	1
2 Instructions on using ECS	6
3 Limits	8
4 Instances	15
4.1 Create an instance.....	15
4.1.1 Create an instance by using the wizard.....	15
4.1.2 Create an instance of the same configuration.....	19
4.1.3 Create an instance from a custom image.....	20
4.1.4 Create a ga1 instance.....	21
4.1.5 Create a compute optimized instance with GPUs.....	23
4.1.6 Create a preemptible instance.....	28
4.1.7 Create an f1 instance.....	29
4.1.8 Create an f2 instance.....	30
4.1.9 Create an f3 instance.....	30
4.1.10 Create an EBM instance.....	31
4.1.11 Create an SCC server instance.....	32
4.2 Launch template.....	33
4.2.1 Create a template.....	33
4.2.2 Create a template version.....	35
4.2.3 Use a launch template.....	37
4.2.4 Delete a template or version.....	38
4.3 Check instance information.....	39
4.4 Change configurations.....	41
4.4.1 Overview of configuration changes.....	41
4.4.2 Upgrade configurations of Subscription instances.....	43
4.4.3 Change configurations of Pay-As-You-Go instances.....	45
4.4.4 Instance type families that support instance type upgrades.....	47
4.4.5 Change EIP Internet bandwidth.....	49
4.5 Change the operating system.....	50
4.6 Reset an instance password.....	51
4.7 Start or stop an instance.....	52
4.8 Restart an instance.....	55
4.9 Reactivate an instance.....	55
4.10 Release an instance.....	56
4.11 Change IP addresses.....	59
4.11.1 Change public IP address.....	59
4.11.2 Convert public IP address to EIP address.....	61
4.11.3 Change the private IP of an ECS instance.....	62

4.12 User-defined data and metadata.....	63
4.12.1 Metadata.....	63
4.12.2 User data.....	67
4.12.3 Instance identity.....	75
4.13 Instance RAM roles.....	79
4.13.1 What is the RAM role of an instance.....	79
4.13.2 Use the instance RAM role in the console.....	81
4.13.3 Use the instance RAM role by calling APIs.....	87
5 Connect to instances.....	91
5.1 Overview.....	91
5.2 Connect to an instance by using the Management Terminal.....	93
5.3 Connect to a Linux instance by using an SSH key pair.....	97
5.4 Connect to a Linux instance by using a password.....	102
5.5 Connect to a Windows instance.....	105
5.6 Connect to an instance on a mobile device.....	113
6 Cloud disks.....	127
6.1 Create a cloud disk.....	127
6.2 Create a cloud disk from a snapshot.....	128
6.3 Attach a cloud disk.....	131
6.4 Partition and format data disk more than 2 TiB.....	134
6.5 Detach a cloud disk.....	141
6.6 Resize cloud disks.....	144
6.6.1 Overview.....	144
6.6.2 Increase system disk size.....	145
6.6.3 Windows _ Resize a data disk.....	150
6.6.4 Linux _ Resize a data disk.....	153
6.7 Reinitialize a cloud disk.....	159
6.8 Roll back a cloud disk.....	166
6.9 Convert billing methods of cloud disks.....	168
6.10 Replace the system disk (public image).....	169
6.11 Replace the system disk (non-public image).....	175
6.12 Monitor a cloud disk.....	180
6.13 Release a cloud disk.....	181
7 Snapshots.....	183
7.1 Create a snapshot.....	183
7.2 Create and delete an automatic snapshot policy.....	186
7.3 Apply automatic snapshot policies to disks.....	187
7.4 Delete automatic snapshots when releasing disks.....	189
7.5 Delete snapshots or automatic snapshot policies.....	190
7.6 View a snapshot chain.....	191
8 Images.....	193
8.1 Create custom image.....	193
8.1.1 Create a custom image by using a snapshot.....	193

8.1.2 Create a custom image by using an instance.....	197
8.2 Manage custom images.....	199
8.3 Copy images.....	201
8.4 Share images.....	203
8.5 Import images.....	205
8.5.1 Image compliance tool.....	205
8.5.2 Notes for importing images.....	210
8.5.3 Install cloud-init for Linux images.....	213
8.5.4 Install virtio driver.....	216
8.5.5 Customize Linux images.....	222
8.5.6 Convert image file format.....	228
8.5.7 Import custom images.....	229
8.6 Export custom images.....	232
8.7 Marketplace images.....	234
8.8 Open source tools.....	236
8.8.1 Create and import on-premises images by using Packer.....	236
8.8.2 Use Packer to create a custom image.....	242
9 Security groups.....	246
9.1 Typical applications of security group rules.....	246
9.2 Scenarios.....	248
9.3 Default security group rules.....	258
9.4 Create a security group.....	260
9.5 Add security group rules.....	263
9.6 View the security group list.....	268
9.7 Modify security group attributes.....	269
9.8 View the security group rules.....	269
9.9 Delete a security group rule.....	269
9.10 Delete a security group.....	270
9.11 Clone a security group.....	271
9.12 Introduction to common ECS instance ports.....	273
9.13 Restore security group rules.....	275
9.14 Add to or remove from a security group.....	277
10 Deployment sets.....	278
10.1 Create an instance in the deployment set.....	278
10.2 Manage deployment sets.....	280
10.3 Create deployment sets.....	281
11 Key pairs.....	283
11.1 Create an SSH key pair.....	283
11.2 Import an SSH key pair.....	284
11.3 Attach or remove an SSH key pair.....	285
11.4 Delete an SSH key pair.....	286
12 Cloud assistant.....	288
12.1 Cloud assistant client.....	288

12.2 Manage commands.....	290
12.3 Create commands.....	292
12.4 Run commands.....	294
12.5 Query execution results and statuses.....	297
13 Elastic Network Interfaces.....	301
13.1 Attach an ENI when creating an instance.....	301
13.2 Create an ENI.....	302
13.3 Attach an ENI to an instance.....	303
13.4 Detach an ENI from an instance.....	304
13.5 Modify attributes of an ENI.....	305
13.6 Delete an ENI.....	306
13.7 Configure an ENI.....	307
14 Tags.....	311
14.1 Limits.....	311
14.2 Add a tag to resources.....	311
14.3 Delete a tag.....	312
14.4 Filter resources by tags.....	314
15 Monitoring.....	315
15.1 Monitoring.....	315
15.2 System events.....	319
15.3 Console output and screenshot.....	324

1 Quick reference

This article is a quick reference guide for common features of Alibaba Cloud ECS instances and resources. It offers solutions for such scenarios as connecting to an instance, resizing a disk, upgrading or downgrading configurations, and using snapshots or images.

Operation instructions and limits

To guarantee proper operation of your ECS instance, please read the [ECS operation instructions](#) and [limits](#) sections carefully before using your instance.

Create and manage ECS instances

Basic operations

1. [Create an ECS instance](#).
2. Connect to the ECS instance. Depending on the operating system running on your ECS instance and your actual scenario, use one of the following methods:
 - For any type of operating system, use the [Management Terminal](#) for scenarios involving troubleshooting and maintenance.
 - For Linux or Unix-like OSs, you can [connect to a Linux instance by using a password](#), or [connect to a Linux instance by using an SSH key pair](#).
 - For Windows OSs, you can [connect to a Windows instance](#).
3. [Stop the ECS instance](#).
4. [Release the instance](#).

To use an ECS instance, follow these steps:

Change configurations

You can change the instance type, IP addresses, and network bandwidth of your instance.

- Subscription instances: [Upgrade configurations of Subscription instances](#) or [renew for configuration downgrade](#)
- [Change configurations of Pay-As-You-Go instances](#)
- [Change public IP address](#)
- [Convert public IP address to EIP address](#)

If the current operating system no longer meets your business needs, you can [change the operating system](#).

Billing

You can [switch from Pay-As-You-Go to Subscription](#).

Select either of the following methods to renew your Subscription instances:

- [Manual renewal](#)
- [Auto-renewal](#)

Refined management of and control over ECS instances

You can use the following features to refine your management of and control over ECS instances:

- [User data](#)
- [Metadata](#), including [instance identity](#)
- [Instance RAM roles](#)

Create and manage cloud disks

Basic operations

To use a cloud disk as a data disk, follow these steps:

1. [Create a cloud disk](#).
2. [Attach a cloud disk](#).
3. [\(Linux\) Format and mount a data disk](#) or [\(Windows\) Format a data disk](#).
4. [Create snapshots](#) to back up data.
5. [Detach a cloud disk](#).
6. [Release a cloud disk](#).

Change configurations

To adjust the capacity of your system disks or data disks, you can [increase the system disk size](#) or resize the data disks. For more information about resizing a data disk, see [Linux _ Resize a data disk](#) and [Windows _ Resize a data disk](#).

Manage data on a cloud disk

If data errors occur on a cloud disk, you can use a snapshot to [roll back a cloud disk](#) and restore data.

If you want to restore a cloud disk to its initial status after it is created, you can [reinitialize a cloud disk](#).

If you want to copy data on an existing cloud disk to a new, empty cloud disk, you can [create a cloud disk from a snapshot](#).

Create and manage snapshots

Basic operations

To use a snapshot, follow these steps:

1. Create a snapshot by using either of the following methods:
 - [Create snapshots](#).
 - [Create and delete an automatic snapshot policy](#), and [apply automatic snapshot policies to disks](#), to enable automatic snapshot creation.
2. [View a snapshot chain](#).
3. [Delete unnecessary snapshots](#) to reduce charges and free disk space

Using snapshots

To copy or back up data, you can use a snapshot to [create a cloud disk from a snapshot](#), or [roll back a cloud disk](#).

To simplify deployment, you can use a system disk snapshot to [create a custom image using a snapshot](#), and [create an instance from a custom image](#).

Create and manage custom images

Only custom images can be operated in the ECS console.

You can run a custom image by using the following methods:

- [Create a custom image using a snapshot](#)
- [Create a custom image by using an instance](#)
- [Use Packer to create a custom image](#)
- [Copy custom images](#) across different regions.
- [Share custom images](#) across different accounts.
- [Import custom images](#)
- [Create and import custom images stored on an on-premises server by using Packer](#)

You can also [export custom images](#) to back up your environment and delete custom images when they are no longer required.

Create and manage security groups

Basic operations

To use a security group, follow these steps:

1. [Create a Security Group](#).
2. [Add security group rules](#).
3. [Add to or remove from a security group](#)
4. [Delete a security group rule](#).
5. [Delete a security group](#).

Manage security groups and their rules

To simplify business deployment, you can [clone a security group](#) across regions or network types.

If new security group rules disrupt your online business application, you can [restore security group rules](#) fully or partially.

Create and manage SSH key pairs

To use an SSH key pair, follow these steps:

1. [Create an SSH key pair](#), or [import an SSH key pair](#).
2. [Bind a SSH key pair](#), or bind the SSH key pair after a Linux instance is created or when you [create an instance](#).
3. [Connect to a Linux instance by using an SSH key pair](#).
4. [Unbind an SSH key pair](#).
5. [Delete a SSH key pair](#).

Create and manage ENIs

To use an ENI, follow these steps:

1. [Create an ENI](#).
2. [Attach an ENI to an instance](#), or [attach an ENI when creating an instance](#).
3. Optional. [Configure an ENI](#).
4. [Detach an ENI from an instance](#).
5. [Delete an ENI](#).

Use tags

You can apply tags to group resources for easier resource organization. To use tags, follow these steps:

1. [Add a tag to resources](#).
2. [Filter resources by tags](#).
3. [Delete a tag](#).

2 Instructions on using ECS

This article describes usage restrictions and recommendations of an ECS instance.

General instructions

Restrictions

- You are prohibited from using your instances for flow-through services. Any violations will lead to punishments including shutdown and lockout of instances, and termination of services.
- You are prohibited from using instances for click farming, advertising, or fraudulent transactions.
- Do not enable SELinux.
- Do not uninstall relevant hardware drivers.
- Do not arbitrarily modify the MAC address of the network adapter.

Recommendations

- For an instance with more than 4 GiB RAM, we recommend that you use a 64-bit operating system as a 32-bit operating system only supports up to 4 GiB RAM. Currently, the following 64-bit operating systems are supported (please refer to the instance purchase page for the latest details):
 - 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
 - Windows 64-bit
- Windows 32-bit supports vCPUs with up to 4 cores.
- A minimum of 2 GiB RAM is required for building a website or deploying a Web environment on a Windows instance.
- An instance type with 1 vCPU core and 1 GiB RAM cannot be used for MySQL service.

- To guarantee service continuity and avoid service downtime, we recommend that you enable auto-start upon instance boot for relevant software. In the case of databases that are connected to service applications, auto-reconnect should be enabled for them.
- For I/O-optimized instances, do not disable the aliyun-service process.
- For Windows users, exercise caution when using the administrator or other accounts to perform actions involving capacity expansion, spanned volume, registry, system update, and other related actions, in order to avoid data corruption due to misoperations.
- For Linux users, exercise caution when using the root or other accounts to perform actions involving fio, mkfs, fsck, capacity expansion, and other related actions, in order to avoid data corruption due to misoperations.
- We do not recommend that you 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 instructions

- Do not kill the built-in shutdownmon.exe process. Otherwise, the server may take a longer time to restart.
- Do not rename, delete, or disable the administrator account.
- We do not recommend that you use the virtual memory if Basic Cloud Disks are used. For Ultra Cloud Disks or SSD Cloud Disks, you can use the virtual memory as needed.

Linux instructions

- Do not modify the contents of the default /etc/issue file on Linux instances . Otherwise, if you create a custom image of the instance and then use it to create a new instance, the new instance cannot start properly because the operating system edition cannot be recognized.
- Do not arbitrarily modify permissions of the directories in the root partition, especially /etc, /sbin, /bin, /boot, /dev, /usr, and /lib. Improper modification of permissions may cause errors.
- Do not rename, delete, or disable the Linux root account.
- Do not compile or perform any arbitrary operations on the Linux kernel.
- We recommend you do not use the swap partition if Basic Cloud Disks are used. For Ultra Cloud Disks or SSD Cloud Disks, you can use the swap partition as needed.
- Do not enable the NetWorkManager service. This service conflicts with the internal network service of the system which can result in network errors.

For more information, see [Limits](#).

3 Limits

When using ECS, note the following limits:

- ECS does not support virtual application installation or revirtualization (such as installation of VMware). Currently, only [ECS Bare Metal Instance and Super Computing Clusters](#) support revirtualization.
- ECS does not support sound card applications.
- ECS does not support external hardware devices directly (such as hardware dongles, USB drives, external hard drives, and USB security keys issued by banks). Instead, it supports a software protection dongle or two-step verification with dynamic passwords.
- ECS does not support IP address translation services such as SNAT. Instead, it supports a VPN or proxy.
- ECS does not support multicast protocols. If multicasting services are required, we recommend that you use point-to-point unicast instead.
- Currently, 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 [Overview](#).

In addition to the preceding limits, the following table details further limits of ECS and states whether you can open a ticket to request changing the limit.

ECS instances

Item	Limit	Can I open a ticket to raise the limit?
Permission to create instances	Complete real-name registration before creating ECS instances in any mainland China regions	No
Instance types for which you can create Pay-As-You-Go instances	Instance types with less than 16 vCPUs	Yes
Default quota of Pay-As-You-Go instances in each region for one account	50 vCPUs (increases with membership levels)	Yes

Item	Limit	Can I open a ticket to raise the limit?
Default quota of preemptible instances in each region for one account	Up to 50 vCPUs are allowed once the permission is granted . The quota increases with membership levels.	Yes
Default quota of launch templates in each region for one account	30	No
Default quota of versions of one launch template	30	No
Switch from Pay-As-You-Go to Subscription	The following instance types (families) are not supported: t1, s1, s2, s3, c1, c2, m1, m2, n1, n2, and e3	No

Block storage

Item	Limit	Can I open a ticket to raise the limit?
Create Pay-As-You-Go cloud disks	Complete real-name registration before creating cloud disks in any mainland China regions	No
Default quota of Pay-As-You-Go cloud disks in all regions for one account	Number of Pay-As-You-Go instances in all regions under the user account x 5	Yes
Default quota of system disks for one instance	1	No
Default quota of data disks for one instance	16 (including cloud disks and Shared Block Storage)	No
Default quota of instances to which one shared block storage can be attached	8	No
Default quota of shared block storage in all regions for one account	10	Yes
Capacity of one Basic Cloud Disk	5 GiB–2,000 GiB	No

Item	Limit	Can I open a ticket to raise the limit?
Capacity of one SSD Cloud Disk	20 GiB–32,768 GiB	No
Capacity of one Ultra Cloud disk	20 GiB–32,768 GiB	No
Capacity of one local SSD disk	5 GiB–800 GiB	No
Capacity of local SSD disks for one instance	1,024 GiB	No
Capacity of one local NVMe SSD disk	1,456 GiB	No
Capacity of local NVMe SSD disks for one instance	2,912 GiB	No
Capacity of one local SATA HDD disk	5,500 GiB	No
Capacity of local SATA HDD disks for one instance	154,000 GiB	No
Capacity of one SSD Shared Block Storage	32,768 GiB	No
Capacity of SSD Shared Block Storage for one instance	128 TiB	No
Capacity of one Ultra Shared Block Storage	32,768 GiB	No
Capacity of Ultra Shared Block Storage for one instance	128 TiB	No
Capacity of one ESSD disk	32,768 GiB	No
Capacity of one system disk	<ul style="list-style-type: none"> Windows: 40 GiB–500 GiB Linux (excluding CoreOS) and FreeBSD: 20 GiB–500 GiB CoreOS: 30 GiB–500 GiB 	No
Capacity of one data disk	<ul style="list-style-type: none"> Basic Cloud Disk: 5 GiB–2,000 GiB SSD Cloud Disk/Ultra Cloud Disk/SSD Shared Block Storage/Ultra Shared Block 	No

Item	Limit	Can I open a ticket to raise the limit?
	Storage: 20 GiB–32,768 GiB <ul style="list-style-type: none"> Local disk: dependent on specific disks 	
Attach a new local disk to an instance with local disks	This feature is not supported.	No
Change configuration of an instance with local disks	Only bandwidth changes are allowed.	No
System disk mount points	/dev/xvda	No
Data disk mount points	/dev/xvd[b-z]	No

**Note:**

Block storage capacity is measured in binary units. 1 KiB is 1,024 bytes. 1 MiB is 1,024 KiB. 1 GiB is 1,024 MiB. 1 TiB is 1,024 GiB.

Snapshots

Item	Limit	Can I open a ticket to raise the limit?
Quota of snapshots	Number of elastic block storage devices x 64	No

Images

Item	Limit	Can I open a ticket to raise the limit?
Quota of custom images in one region for one account	100 (increases with membership levels)	Yes
Maximum number of users with whom a single image can be shared	50	Yes
Usage of images on instance types	32-bit images are not supported on an instance with 4 GiB or more RAM.	No

Key pairs

Item	Limit	Can I open a ticket to raise the limit?
Quota of key pairs in one region for one account	500	No
Instance types supporting key pairs	All instance types except non-I/O optimized instance types in Generation I	No
Images supporting key pairs	Linux images only	No

Internet bandwidth

Item	Limit	Can I open a ticket to raise the limit?
Maximum inbound Internet bandwidth	200 Mbit/s	No
Maximum outbound Internet bandwidth	100 Mbit/s	Open a ticket to increase the limit to 200 Mbit/s
Change the assigned public IP address for one instance	The instance has existed for less than six hours. You can change the public IP address of an instance three times.	No

Security groups

Item	Limit	Can I open a ticket to raise the limit?
Quota of instances/IP addresses for one security group	<ul style="list-style-type: none"> Security groups for classic network instances: 1,000 classic network instances Security groups for VPC instances: 2,000 private IP addresses (shared by primary and secondary network cards) 	No
Quota of authorization rules for one security group	100	No

Item	Limit	Can I open a ticket to raise the limit?
Quota of security groups in one region for an account	100 (increases with membership levels)	Yes
Quota of security groups to which each Elastic Network Interface (ENI) belongs for one instance	5	Yes
Port	For the outbound Internet traffic, the default STMP port is 25, which is disabled by default and cannot be enabled through security group rules.	Open a ticket to enable it. For more information, see Request for enabling TCP port 25 .

Deployment sets

Item	Limit	Can I open a ticket to raise the limit?
Quota of deployment sets in one region for an account	2	No
The number of instances that can be included in a deployment set	Seven instances are allowed in one zone. As for the number of instances allowed in one region, it equals 7 x number of zones.	No
Instance types that can be created in a deployment set	c5, g5, hfc5, hfg5, r5, se1ne, sn1ne, and sn2ne	No

ENIs

Item	Limit	Can I open a ticket to raise the limit?
Quota of ENIs in one region for one account	100 (increases with membership levels)	Yes

Tags

Item	Limit	Can I open a ticket to raise the limit?
Quota of tags that can be bound to one instance	20	No

APIs

Item	Limit	Can I open a ticket to raise the limit?
Quota of CreateInstance calls	200 times per minute	Yes

**Note:**

For the limits of VPC products, see [Limits](#).

4 Instances

4.1 Create an instance

4.1.1 Create an instance by using the wizard

This article describes how to create an instance using the console wizard.

Prerequisites

- Before creating an ECS instance, you must complete the [preparation work](#).
- To bind an SSH key pair when creating a Linux instance, you must [create an SSH key pair](#) in the target region.
- To set the user-defined data, you must prepare the [user data](#).
- To authorize an instance to play a role, you must [create an instance RAM role and grant it permissions](#).

Procedure

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Instances**.
3. Click **Create Instance**.
4. Complete the **Basic Configurations** as follows:
 - a) Select a **Billing Method: Subscription** or **Pay-As-You-Go**.
 - b) Select a region and zone. By default, a zone is assigned randomly. You can also select an applicable one. For more information about regions and zones, see [regions and zones](#).



Note:

- After an instance is created, you cannot change its region and zone.
 - Some instance type families are not supported in all regions. For more information, see [Create a compute optimized instance with GPUs](#), [Create an f1 instance](#), [Create an f2 instance](#), [Create an SCC server instance](#), and [Create an EBM instance](#).
- c) Select an instance type and specify the quantity of instances. The availability of an instance type family is determined by the selected region. For the application scenarios of each instance type, see [instance type families](#).



Note:

- The quota of Pay-As-You-Go or preemptible instances for your account is shown on the page.
- To use Elastic Network Interfaces (ENIs), select an enterprise-level instance type with no less than two vCPU cores or an entry-level instance type with no less than four vCPU cores. For more information about the maximum number of ENIs that can be attached to one instance, see [instance type families](#).
- To use an SSD Cloud Disk, select an I/O-optimized instance.

d) Select an image. You can select a system image, custom image, shared image, or marketplace image.



Note:

- To use an SSH key pair, select a Linux image.
- To use user-defined data, select an image as instructed in [user data](#).

e) Select storage devices:

- **System Disk:** Required. A system disk is required for the image. Specify the cloud disk category and size for the system disk:
 - Cloud disk category: The available categories are determined by the selected region.
 - Size: The default size range is [40, 500] GiB. If the selected image file is greater than 40 GiB, the size is defaulted to the image file size. The available size range varies with the selected image, as shown in the following table.

Image	Available size range
Linux (excluding CoreOS) FreeBSD	[max{20, ImageSize}, 500] GiB. Where, the public image size is 40 GiB for Ubuntu 14.04 32-bit, Ubuntu 16.04 32-bit, and CentOS 6.8 32-bit.
CoreOS	[max{30, ImageSize}, 500] GiB
Windows	[max{40, ImageSize}, 500] GiB

- **Data Disk:** Optional. If you create a cloud disk as a data disk at this time, you must select the disk type, capacity, and quantity, and set whether to [encrypt](#). You can create an empty data disk or create a data disk from a snapshot. Up to 16 data disks can be added.



Note:

The data disks added here have the following features:

- The billing method is the same as that of the instance.
 - A Subscription data disk has to be released along with the instance, but a Pay-As-You-Go data disk can be set to being released along with the instance.
- If you have selected an instance type family that has local disks (such as i1, d1, or d1ne), the local disk information is displayed. You cannot specify the quantity or category of local disks, which is determined by the selected instance type. For information about the local disks corresponding to various instance types with local disk, see [instance type families](#).

5. Click **Next: Networking** to finish the network and security group configuration:

a) Select a network:

- **VPC:** Must select a VPC and a VSwitch. If you do not have a VPC and a VSwitch, you can use the default ones.
- **Classic network:** If you purchased the ECS instance for the first time after June 16, 2016, 12:00 (UTC + 8), you can no longer select a classic network.

b) Configure the **Network Billing Method**:

- To assign a public IP address to the instance, select **Assign public IP**. Then, select **PayByTraffic** as the network billing method and specify the bandwidth. For public IP addresses assigned in this way, you cannot unbind them from the instance. For more information about network billing, see [billing of network bandwidth](#).
- If your instances do not need to access the Internet or your VPC instances [use an Elastic IP \(EIP\) address to access the Internet](#), you do not need to assign a public IP. You can unbind an EIP address from an instance.

c) Select a security group. You can use the default security group if you do not create one. For the rules of the default security group, see [default security group rules](#).

d) Add an Elastic Network Interface (ENI). If your selected instance type supports ENI, you can add one and specify a VSwitch for it.



Note:

By default, the ENI is released along with the instance. You can detach it from the instance in the [ECS console](#) or by using the [DetachNetworkInterface](#) interface.

6. (Optional.) Click **Next: System Configurations** to finish the following configuration:

- Select and set logon credentials. You can choose to [set the credentials after creating an instance](#) or do it now. Select a credential based on the image:
 - Linux: You can select a password or SSH key pair as a logon credential.
 - Windows: You can only select a password as a logon credential.
 - Specify the instance name, which is displayed in the ECS console, and the host name, which is displayed inside the guest operating system.
 - Set the advanced options:
 - Instance RAM role: Assign a RAM role to the instance.
 - UserData: Customize the startup behaviors of an instance or pass data into an instance.
7. (Optional) Click **Next: Grouping** to manage instances by group. You can add tags to instances to simplify future management.
8. Confirm the order:
- In the **Configurations Selected** area, confirm all the configurations. You can click the edit icon to re-edit the configuration.
 - (Optional) Click **Save as launch template** to record your configuration as a launch template for further reference and a quick launch. For more information, see [instance launch template](#).
 - (Optional) Click **View Open API** to acquire the best API practice about how to create instances. At the left side, **API Workflow** explains the related API and request parameter values for the current operation. At the right side, the programming language-specific sample are given for you to use, currently, **Java** and **Python** samples are provided. For more information about ECS API, see *Cite Left* [ECS API Reference](#) *Cite Right* [Overview](#).
 - (Optional) If the billing method is **Pay-As-You-Go**, you can set **Automatic Release**.
 - (Optional) If the billing method is **Subscription**, you can set the duration and select whether to enable **Auto Renewal**.
 - Confirm the configuration costs. The billing methods for an instance and Internet bandwidth determine the displayed cost information, as shown in the following table.

Instance billing method	Fees estimated
Pay-As-You-Go or preemptible instance	Internet traffic fee + configuration fee. Configuration fees include: the instance type (vCPU and memory), the system disk, data disks (if any), and local disks (if any).

Instance billing method	Fees estimated
Subscription	Internet traffic fee + configuration fee. Configuration fees include: the instance type (vCPU and memory), the system disk, data disks (if any), and local disks (if any).

- Read and confirm **Terms of Service**.

9. Click **Create Instance**.

Result

When the instance is activated, click **ECS console** to view the instance details on the console. In the **Instances** list of the relevant region, you can view the information of the new instance, including the instance name, the Internet IP address, and the private IP address.

What's next

- You can create an FTP site on the instance for transferring files. For more information, see [build an FTP site on an ECS instance](#).
- To secure your instance after creation, we recommend that you perform security compliance inspection and configuration:
 - Linux instances: See [harden operating system security for Linux](#) in *Cite LeftSecurity AdvisoriesCite Right*.
 - Windows instances: See [harden operating system security for Windows](#) in *Cite LeftSecurity AdvisoriesCite Right*.
- If a data disk is created along with the instance, you must partition and format the disk before use. For more information, see [format a data disk for Windows instances](#) or [format and mount data disks for Linux instances](#).

4.1.2 Create an instance of the same configuration

To duplicate ECS instances of the same configurations, use the **Buy Same Type** feature.

Procedure

1. Log on to the [ECS console](#).
2. Select the target region.
3. In the left-side navigation pane, click **Instances**.
4. Find your ECS instance and, in the **Actions** column, select **More > Buy Same Type**.

5. On the **Buy the Same Configuration** page, confirm the selected configurations in the **Overview** section. If you want to modify any configurations, select **View More** to change the billing method, security group, network billing method, bandwidth, logon credential, or instance name.
6. To purchase a Subscription ECS instance, you can change the **Purchase Time**.
7. Set the **Amount**.
8. Read and confirm you agree to the **ECS Service Terms and Product Terms of Service**.
9. Confirm the order.

4.1.3 Create an instance from a custom image

If you want to create an ECS instance that has the same operating system, software applications, and data as an existing instance, you can create a custom image and use it to create the new ECS instance. This method improves the deployment efficiency.

Context

- If the image and the instance are in the same region, create a custom image by using one of the following methods:
 - [Import an image](#)
 - [Create a custom image by using an instance](#)
 - [Create a custom image by using a snapshot](#)
- If the custom image and the instance are in different regions, copy the custom image to the target region. For more information, see [Copy images](#).
- If the image to be used is owned by another account, it must be shared with you. For more information, see [share images](#).

Procedure

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Instances**.
Alternatively, you can click **Images** to find the target image, and then click **Create Instance** in the **Actions** column.
3. In the upper-right corner of the **Instances** page, click **Create Instance**.
4. Follow the steps when you [create an instance by using the wizard](#). When creating an ECS instance, note the following:
 - **Region**: Select the region where the image is located.

- **Image:** Select **Custom Image** or **Shared Image**, and then select an image from the drop-down list.

**Note:**

If the selected custom image contains more than one data disk snapshot, an equal number of cloud disks are automatically created to function as data disks. By default, the size of each data disk is equal to that of the source snapshot. You can only increase the size of a data disk. You cannot decrease it.

5. Confirm the order.

4.1.4 Create a ga1 instance

Image description

GPU visualization computing ga1 instances use the AMD S7150 series GPU. Alibaba Cloud and AMD work together to optimize GPU drivers, so you can select the following images with preinstalled drivers from **Image Marketplace**:

- Ubuntu16.04 with AMD GPU driver preinstalled
- Windows Server 2016 English version with AMD GPU driver preinstalled
- Ubuntu16.04 with AMD GPU driver and KDE preinstalled

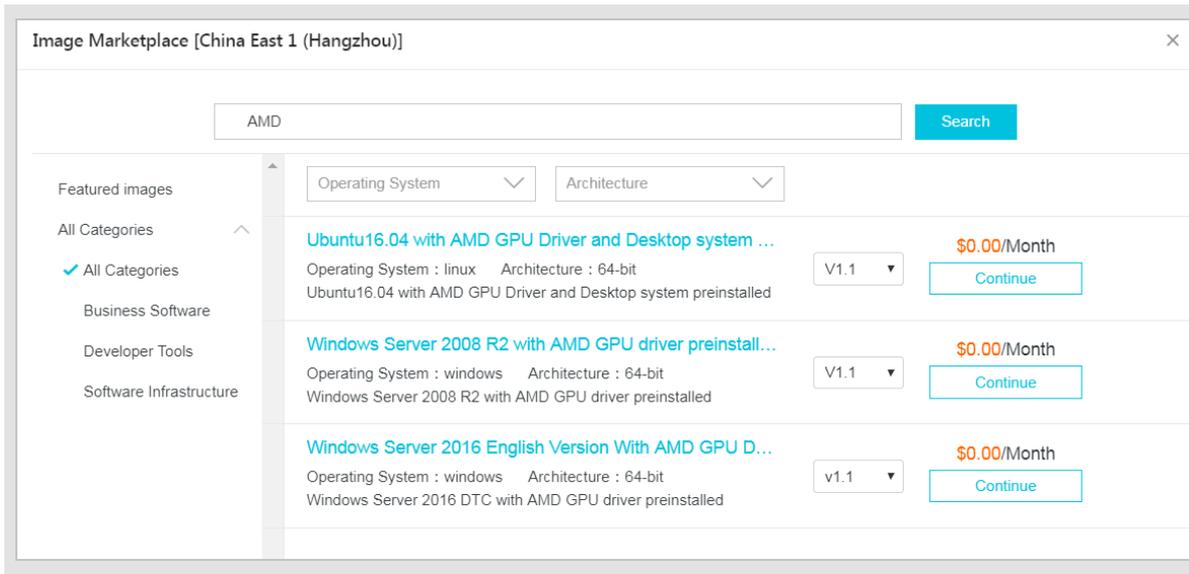
Procedure

You can create a ga1 instance by following the instructions provided in [creating an instance](#). Use the following configurations when configuring an instance:

- **Network:** Select **VPC**. The current GPU-rendered ga1 instances only support VPCs.
- **Instance:** Select **Heterogeneous Computing > GPU Visualization Compute > GPU Rendering Type ga1**.
- **Image:** Click **Marketplace Image**, and then click **Select from image market (including operating system)**. Enter GPU or AMD in the search box to search images.

**Note:**

It is recommended that you purchase or subscribe to these images because you can select images later for another instance directly from **Purchased Images** or **Subscribed Images**.



Precautions

- The driver used for GPU visualization computing ga1 instances is optimized jointly by Alibaba Cloud and AMD. It is currently available only in images provided by Alibaba Cloud. No driver download link is provided, so you cannot install the driver on your own.
- If the driver does not work due to uninstallation or deletion of GPU driver-related components, [change a system disk](#) to restore the GPU functions.



Warning:

Changing a system disk may cause data loss.

- If you use an unspecified image when creating a GPU visualization computing ga1 instance, the instance driver will not work. In this case, [change a system disk](#) to use the image with an AMD GPU driver preinstalled.
- For Windows, after the GPU driver takes effect, the **Connect** function cannot be used, and the **Management Terminal** page displays a black screen or the startup page. In this case, use another protocol to enter the OS, for example, Windows embedded Remote Desktop Protocol (RDP).
- RDP does not support DirectX, OpenGL, or other related applications. You need to install the VNC service and client or use other protocols, such as XenDesktop HDX 3D and PCOIP.

4.1.5 Create a compute optimized instance with GPUs

Compute optimized instances with GPUs (or GPU instances for short) are available in the gn4, gn5, gn5i, and gn6v instance families.

Create an instance

To create a GPU instance, you must configure the following settings. For detailed steps, see [create an ECS instance](#).

- **Region:** Each instance type family is available only in certain regions:
 - gn4: North China 2 (Zone A), East China 2 (Zone B), South China 1 (Zone C)
 - gn5: North China 2 (Zone C and E), North China 5 (Zone A), East China 1 (Zone G and F), East China 2 (Zone D, B, and E), South China 1 (Zone D), Hong Kong (Zone C and B), Asia Pacific SE 1 (Zone A and B), Asia Pacific SE 2 (Zone A), Asia Pacific SE 3 (Zone A), Asia Pacific SE 5 (Zone A), US West 1 (Zone A and B), US East 1 (Zone A and B), and EU Central 1 (Zone A)

**Note:**

If you want to deploy NVIDIA GPU CLOUD (NGC) on a gn5 instance, see [deploy an NGC on a gn5 instance when selecting a region](#).

- gn5i: North China 2 (Zone C, E and A), East China 1 (Zone B), East China 2 (Zone D and B), and South China 1 (Zone A)
- gn6v: East China 2 (Zone F)

The preceding regions and zones are for reference purpose only. Please check the purchase page for the latest information about the availability of instance types.

- **Image:**
 - To install a GPU driver and a CUDA library, select one of the following options:
 - Select **Public Image**. From the drop-down list, select a CentOS 64-bit image (any version), the Ubuntu 16.04 64-bit image, or the SUSE Linux Enterprise Server 12 SP2 64-bit image. Then, select **Auto-install GPU Driver**. Finally, select a CUDA library and GPU driver.

**Note:**

- You can select a GPU driver based on your business needs. For a new business system, we recommend that you select the latest GPU driver.

- If **Auto-install GPU Driver** is selected, **UserData** will be generated automatically. It is a shell script to install the CUDA library and the GPU driver automatically. You can view it when setting up the **Advanced** options in **System Configurations**. After the instance is created and started successfully for the first time, the cloud-init runs the script to install the GPU driver. For more information, see the note for the GPU installation script.
- Select **Marketplace Image**, search for `NVIDIA`, and select an image from the results. Currently, only CentOS 7.3 and Ubuntu 16.04 are supported.
- To use a GPU instance for deep learning, you can select an image pre-installed with the deep learning framework. Select **Marketplace Image**, search for `Deep Learning`, then select the desired image from the results. Currently, only CentOS 7.3 and Ubuntu 16.04 are supported.
- For other images, you must [download and install a GPU driver](#) after the instance is created.
- **Instance Type**: Select **Heterogeneous Computing > GPU Compute**, and select an instance type.
- **Network**: Select **VPC**.
- **Network Billing Method**: Select a bandwidth based on your needs.

**Note:**

If Windows 2008 R2 or earlier is running on the instance, you cannot use the [Management Terminal](#) in the ECS console to connect to the GPU instance after the GPU driver is effective. Therefore, you must select **Assign Public IP**, or [bind EIP](#) after creating the instance.

- **Log on Credentials**: Set a credential to log on to the instance.

**Note:**

We recommend that you do not select **Set Later**. This is because, to log on to the instance after it is created and before the GPU driver is installed, you have to reset the password or bind an SSH key pair, and then restart the instance to make the change take effect. The restart operation, however, causes the failure of driver installation.

- **User Data**: If **Auto-install GPU Driver** is enabled, a script to install a CUDA library and a GPU driver is displayed here. Read the script and note carefully.

Check the installation progress

If **Auto-install GPU Driver** is enabled, after the instance is created, you can [connect to the instance](#), and read the log file `/root/nvidia_install.log` to check the installation progress.



Note:

To avoid installation failure, do not operate GPU or install any other GPU-related software before the GPU driver is installed successfully.

Download and install a GPU driver

If the image is not pre-installed with a GPU driver, you must install one for the instance.

1. Obtain the GPU driver package:
 - a. Go to the [NVIDIA official website](#).
 - b. Manually locate the appropriate driver for the instance, and click **Search**. Complete the following parameters and search for the driver.

Item	gn4	gn5	gn5i	gn6v
Product Type	Tesla	Tesla	Tesla	Tesla
Product Series	M-Class	P-Series	P-Series	V-Series
Product	M40	Tesla P100	Tesla P4	Tesla V100
Operating System	Select an OS version according to the image of your instance. If the server operating system is not displayed in the drop-down list, click Show all Operating Systems at the bottom of the drop-down list.			

Item	gn4	gn5	gn5i	gn6v

c. Confirm the driver and click **DOWNLOAD**.

2. Install the GPU driver:

- Windows: Double-click the installation file to install the driver.
- Linux:
 1. Download and install the appropriate version of kernel-devel and kernel-header packages.
 2. Run the following commands to verify that the packages are downloaded and installed successfully.

```
sudo rpm -qa | grep $(uname -r)
```

Take CentOS 7.3 as an example. If the following message returns, the packages are installed successfully.

```
kernel-3.10.0-514.26.2.el7.x86_64
kernel-headers-3.10.0-514.26.2.el7.x86_64
kernel-tools-libs-3.10.0-514.26.2.el7.x86_64
python-perf-3.10.0-514.26.2.el7.x86_64
```

```
kernel-tools-3.10.0-514.26.2.el7.x86_64
```

3. Go to the download page of the driver on the NVIDIA website, click the **ADDITIONAL INFORMATION** tab, and follow the tips to install the driver.

Take Linux 64-bit Ubuntu 14.04 as an example.

TESLA DRIVER FOR LINUX OPENSUSE 13.2

Version: 375.66
 Release Date: 2017.5.9
 Operating System: Linux 64-bit Opensuse 13.2
 Language: English (US)
 File Size: 133.05 MB

DOWNLOAD

RELEASE HIGHLIGHTS **SUPPORTED PRODUCTS**

ADDITIONAL INFORMATION

Once you accept the download please follow the steps listed below

- ``rpm -i nvidia-diag-driver-local-repo-opensuse132-375.66-1.x86_64.rpm``
- ``zypper refresh``
- ``zypper install cuda-drivers``
- ``reboot``

Install a GRID driver

If your gn5, gn5i or gn6v instance requires OpenGL, you must install a GRID driver. For more information, see [install a GRID driver on a gn5/gn5i/gn6v instance](#).

Note

For remote access

For Windows 2008 R2 or earlier, if you want to connect to your instance by using the [Management Terminal](#) on the ECS console after the GPU driver is effective, the **Management Terminal** is non-operational. It shows either a black screen or the startup interface. If the instance can access the Internet, connect to the instance remotely by using other protocols, such as the Remote Desktop Protocol (RDP) developed by Microsoft.

RDP does not support DirectX, OpenGL, and other related applications. Therefore, you must install the VNC server and client, or other protocols that support these applications, such as PCOIP or XenDesktop HDX 3D.

For the GPU driver installation script

For the GPU auto installation script, note the following:

- The script automatically downloads and installs a NVIDIA GPU driver and a CUDA library.

- The installation takes 4.5 to 10 minutes, depending on the intranet bandwidth and the quantity of vCPU cores of the instance. To avoid installation failure, do not operate the GPU or install any GPU-related software until the GPU driver is installed successfully.
- After the GPU is installed successfully, the instance restarts automatically to bring the driver into effect.
- The script turns on the **Persistence Mode** of the driver and add this setup to the automatic startup script of the instance to make sure that this mode is turned on automatically when the instance is restarted. The GPU driver works in a more stabilized manner in this mode.
- When [changing the operating system](#), note the following:
 - Replacing Ubuntu 16.04 64-bit or SUSE Linux Enterprise Server 12 SP2 64-bit with other distributions or versions causes failure of automatic installation of the GPU driver.
 - Replacing one version of CentOS with another version does not cause installation failure of the GPU driver.
 - Replacing the image with another image that does not support the automatic installation script causes failure of automatic installation of the GPU driver.
- A log is generated during installation, and is saved to the `/root/nvidia_install.log`. You can read the log to check the installation result of the driver. If the installation fails, you can view the causes.

4.1.6 Create a preemptible instance

You can create a preemptible instance in the ECS console. This document describes the steps and relevant operations.

Procedure

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Instances**.
3. Click **Create Instance**.
4. Select **Preemptible Instance** for **Billing Method**.
5. Set the **Single instance price limit**.

You need to enter the maximum price you are willing to pay. When your bid is higher than the current market transaction price, the instance starts to run. Eventually, the market transaction price is charged for the instance.

6. Select or enter the number of instances to purchase.
7. Complete other settings. For more information, see [create an instance by using the wizard](#).

8. After the order is confirmed, click **Create Instance**.

After a preemptible instance is created, you can view its information in the instance list. A preemptible instance is marked as a **Pay-As-You-Go-Preemptible Instance**. After opening the instance details page, you can view the bidding policy set during instance creation in the **Payment Information** area.

4.1.7 Create an f1 instance

This article describes how to create an f1 instance.

Prerequisites

You must use an image that is pre-installed with the Intel development environment to create an f1 instance. To obtain the image, [open a ticket](#).

Procedure

Follow the steps described in [create an ECS instance](#). The following configurations must be selected:

- **Region:** Select **China East 1 (Hangzhou) > China East 1 Zone F**.
- **Instance Type:** Select **Heterogeneous Computing > FPGA > Compute**, and then select the appropriate f1 instance type.
- **Image:** Select **Shared Image**, and then select the shared image.



Note:

You must use an image that is pre-installed with the Intel development environment to create an f1 instance. This image is not available in the Alibaba Cloud Marketplace directly. To obtain the image, please find quartus17.0, vcs2017.3, dcp sdk in the `opt` directory.

- **Network:** Select **VPC**, and select a created VPC and VSwitch.

After an f1 instance is created, [connect to the instance](#) and run the following command to check whether the licence is configured.

```
echo $LM_LICENSE_FILE #Check whether the variable is set.
```

Best practices

See best practices of f1 instances:

- [Use OpenCL on an f1 instance](#)
- [Use f1 RTL \(Register Transfer Level\)](#)

4.1.8 Create an f2 instance

This article describes how to create an f2 instance.

Prerequisites

You must use an image that is pre-installed with the Xilinx development environment to create an f2 instance. To obtain the image, [open a ticket](#).

Procedure

Follow the steps described in [creating an ECS instance](#). However, the following configurations must be selected:

- **Region:** Select **China East 1 (Hangzhou) > China East 1 Zone B**.
- **Instance Type:** Select **Heterogeneous Computing > FPGA Compute**.
- **Image:** Click **Shared Image**, and then select the shared image.
- **Network:** Select **VPC**, and select a created VPC and VSwitch.

Best practice

[Use OpenCL on an f2 instance](#)

4.1.9 Create an f3 instance

This article describes how to create an f3 instance.



Note:

Due to limited computing resources, we recommend that you use instances with at least four cores, such as instance type family g5-ecs.g5.2xlarge (8 vCPU core, 32 GiB). Create an f3 instance when you need to download the image to the FPGA chipset.

Prerequisite

[Open a ticket](#) to request a free f3 instance for testing. This instance type is currently in public beta release.

Procedure

For more information about how to create an f3 instance, see [create an instance by using the wizard](#). However, the following configurations are recommended:

- **Billing Method:** Select **Pay-As-You-Go** or **Subscription**.



Note:

- During the testing phase, f3 instances are available for free. Other ECS resources including cloud disks, public network bandwidth, and snapshots will incur usage fees.
- f3 instances are not available as preemptible instances.
- **Region:** Select **China East 2 (Shanghai)**.
- **Instance Type:** Select **Heterogeneous Computing > FPGA Compute**, and then select your required instance type.
- **Image:** Click **Shared Image**, and then select the specified image.

**Note:**

A Xilinx image is available for use (recommended). The image is only available as a Shared image. To obtain the image, open a ticket.

- **System Disk:** Allocate a 200 GiB Ultra Disk for the system image.
- **Network:** Select **VPC**.

Best practices

[Use OpenCL on an f3 instance](#)

[Use RTL compiler on an f3 instance](#)

4.1.10 Create an EBM instance

Follow the steps in [creating an instance by using the wizard](#) to create an EBM instance. However, the following configurations are recommended:

- **Region:** Currently, EBM instances are available in the following regions and zones: **China East 2 (Shanghai), Zone D, China North 2 (Beijing), Zone C, China East 1 (Hangzhou), Zone G, and China South 1 (Shenzhen), Zone D**.
- **Instance Type:** In ebmhfg5, ebmc4, and ebmg5 type families are available. For more information about instance types, see [instance type families](#).
- **Image:** The following public images are supported.

Operating system	Image
Linux	<ul style="list-style-type: none"> • CentOS 7.2/7.3/7.4/6.9/6.8 64-bit • Ubuntu 14.04/16.04 64-bit • Debian 8.9/9.2 64-bit • OpenSUE 42.3 64-bit • SUSE Linux Enterprise Server 12 SP2 64-bit

Operating system	Image
	<ul style="list-style-type: none"> Aliyun Linux 17.1 64-bit
Windows	<ul style="list-style-type: none"> 2016 Data Center Edition 64-bit Chinese Edition 2016 Data Center Edition 64-bit English Edition 2012 R2 Data Center Edition 64-bit Chinese Edition 2012 R2 Data Center Edition 64-bit English Edition

- **Storage:** EBM instances support up to 16 data disks. You can add a data disk during or after instance creation, and then [mount the data disk](#).
- **Network:** Only VPC is supported.

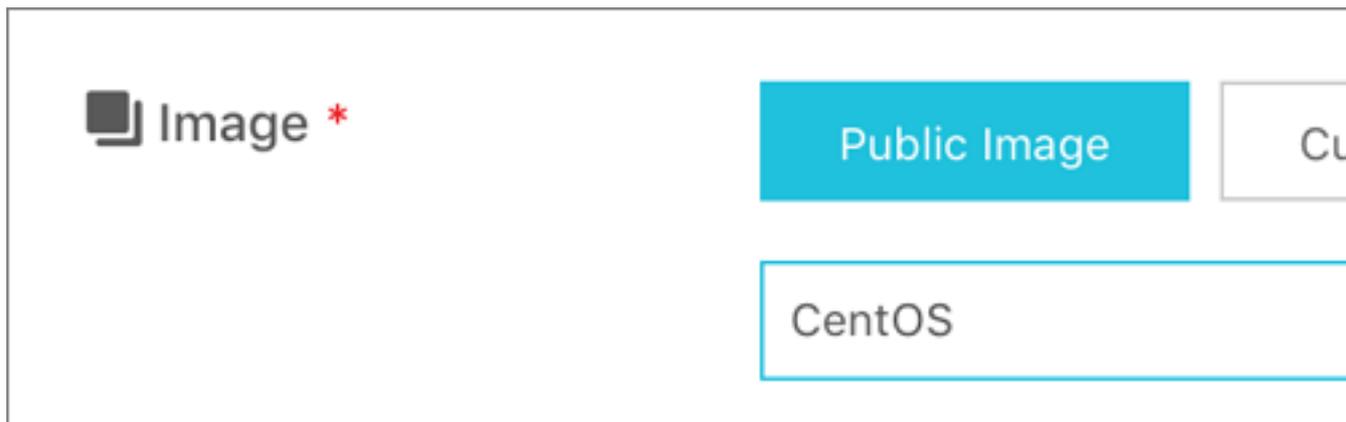
4.1.11 Create an SCC server instance

Super Computing Cluster (SCC) is based on the ECS Bare Metal (EBM) instance product. Utilizing the high-speed interconnectivity of RDMA (Remote Direct Memory Access) technology, SCC greatly improves network performance and increases the acceleration ratio of large-scale clusters. SCC has all the advantages of EBM instances, and provides high-quality network performance featuring high bandwidth and low latency. For more information, see [ECS Bare Metal instance and Super Computing Clusters](#).

This article describes how to create an SCC instance. For more information about instance creation, see [create an image by using the wizard](#).

The following configurations are recommended for SCC instances:

- **Region:** Currently, only **China East 2 (Shanghai) Zone D** and **Zone B** support SCC instances.
- **Instance Type:** Instance type families scch5 and sccg5 are available. For more information about instance types, see [instance type families](#).
- **Image:** Select **Public Image**. Currently, only a custom Linux CentOS 7.5 image for SCC is supported.



- **Storage:** SCC support up to 16 data disks. You can add a data disk during or after instance creation, and then [mount the data disk](#).
- **Network:** Only VPC is supported.

4.2 Launch template

4.2.1 Create a template

You can create a launch template using the following methods:

- [Create a launch template in the ECS console](#) if you want to create launch templates first, and then create instances using a specific launch template in one click.
- [Create a launch template on the ECS buy page](#) to create an instance and save its configuration information as a launch template.

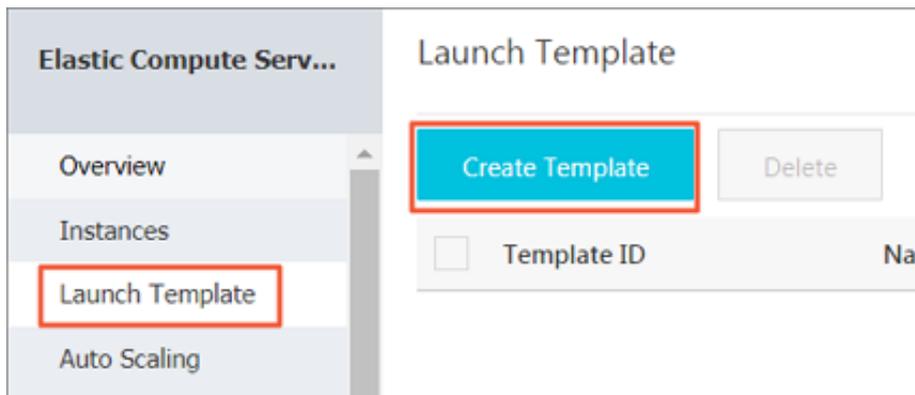


Note:

- Each account can create a maximum of 30 launch templates per region.
- All parameters are optional when you create a template using the ECS console. However, if the template that you want to use to create an instance does not have all required parameters (such as an image), then you must specify the required parameters at instance creation.
- A template cannot be modified after it is created.

Create a template in the ECS console

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Launch Templates**, and then click **Create Template**.



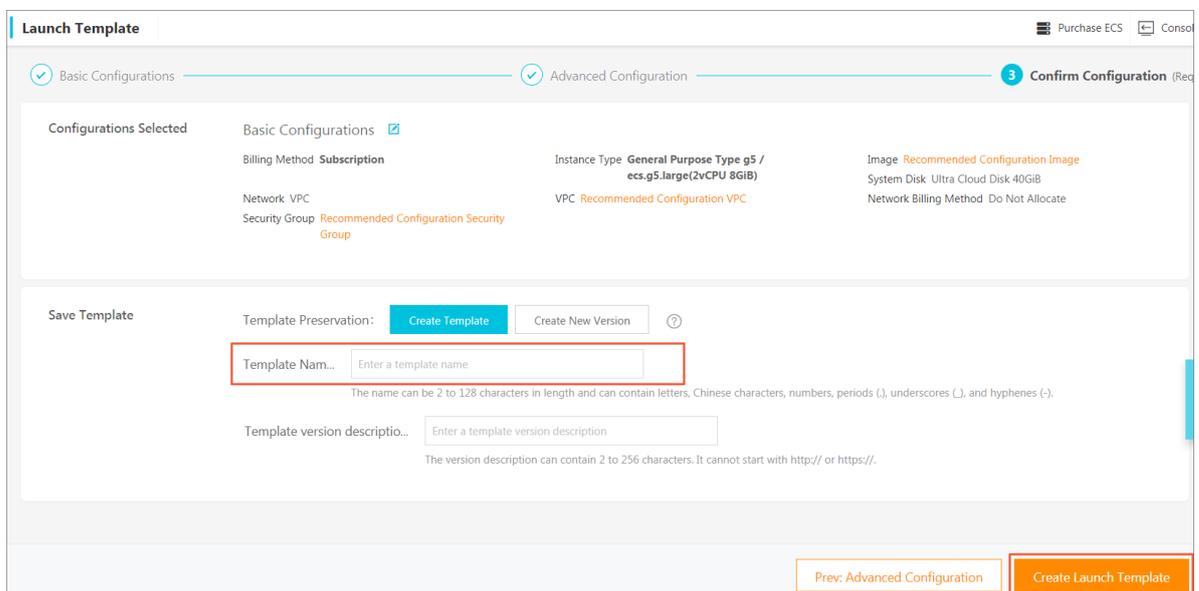
3. Go to the **Launch Template** page and complete the basic configurations and advanced configurations.



Note:

During your first template creation, the **Clone Template** area is unavailable. If you have already created templates, you can select an existing template, and version, and then modify its configurations.

4. On the **Confirm Configuration** page, enter a template name and description, and then click **Create Launch Template**.



Note:

All parameters are optional when you create a template. However, on the **Confirm Configuration** page, we recommend that you configure the required parameters so that you can create instances in one click as needed.

Click **View Template** in the **Activated** dialog box to view the template you have created.

Create a template on the ECS buy page

1. Go to the [ECS product details page](#), and then click **Buy Now**.
2. Configure the required parameters, and then click **Save as launch template**.
3. In the dialog box that appears, select **Create Template**, enter a template name and description, and then click **Save**.

Click **View Template** in the **Activated** dialog box to view the template you have created.

4.2.2 Create a template version

One template can have multiple versions. The default version number of a newly created template is 1, and you can create additional versions based on this template. The version number increments automatically as you create a new version. You cannot customize the version number, but you can set any of the template versions as the default version.



Note:

- Each template can have a maximum of 30 versions.
- All parameters are optional when you create a template version.
- A template version cannot be modified once you have created it.

You can create a template version using the following methods:

- [Create an instance using the ECS console](#) to create versions of a template for future use.
- [Create an instance on the ECS buy page](#) to create an instance, save its configurations, and create versions of a template.

Prerequisite

You have already [created a template](#).

Create an instance using the ECS console

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Launch Template**.
3. Select a template ID to view its configurations, and then click **New Version**. You can also click **New Version** in the **Actions** column.

The screenshot shows the 'Launch Template' management interface. At the top, there are buttons for 'Create Template' and 'Delete', along with a search bar. Below is a table listing templates:

Template ID	Name	Created At	Default Version	Latest Version	Actions
It-...	testcl0603	2018-06-03 14:05	1	1	Create Instance Delete New Version

Below the table, there is a 'Version Information' section with a 'New Version' button highlighted in red, and a 'Configuration Information' section showing details like Pricing Model, Region, Instance Type, Image, and Storage.

4. On the **Launch Template** page, set the parameters.



Note:

You can also go to the **Clone Template** area, select an existing template and version, and then set the parameters.

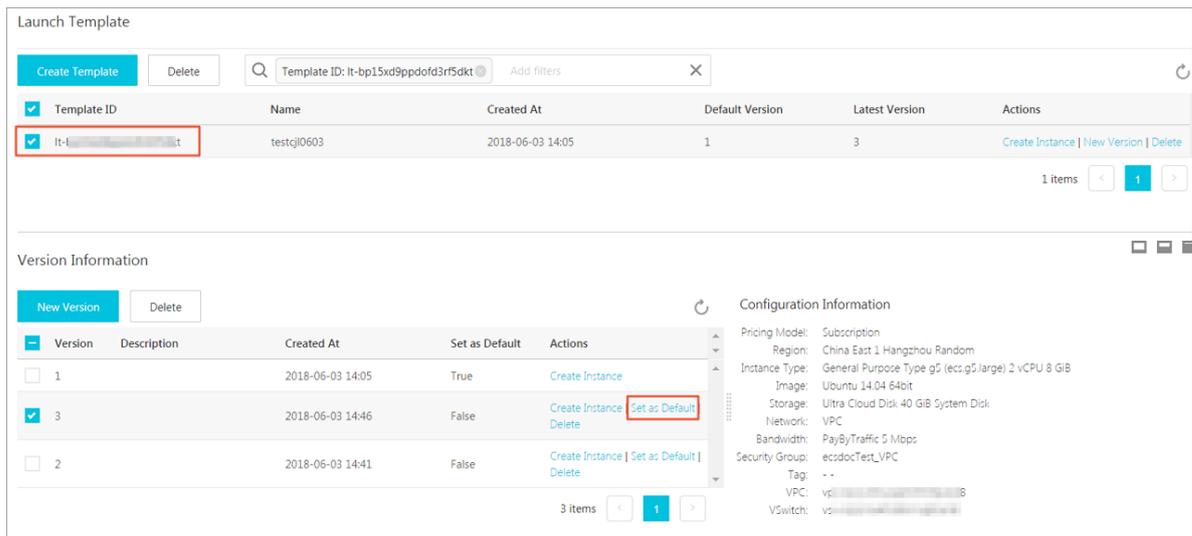
5. On the **Confirm Configuration** page, select **Create New Version**, and then select a template to save the version.
6. Click **Create Launch Template**.
7. In the dialog box that appears, click **View New Version** to view the version you have created.

Create an instance on the ECS buy page

1. Go to the [ECS product details page](#), and then click **Buy Now**.
2. On the ECS buy page, configure the parameters.
3. On the **Preview** page, click **Save as launch template**.
4. In the dialog box that appears, click **Create New Version**, and then select a template to save the version.
5. In the **Activated** dialog box, click **View New Version** to view the version you have created.

Change the default version

1. In the ECS console, select a template ID that has multiple versions.
2. Locate the version you want to set as default, and then click **Set as Default** in the **Actions** column.



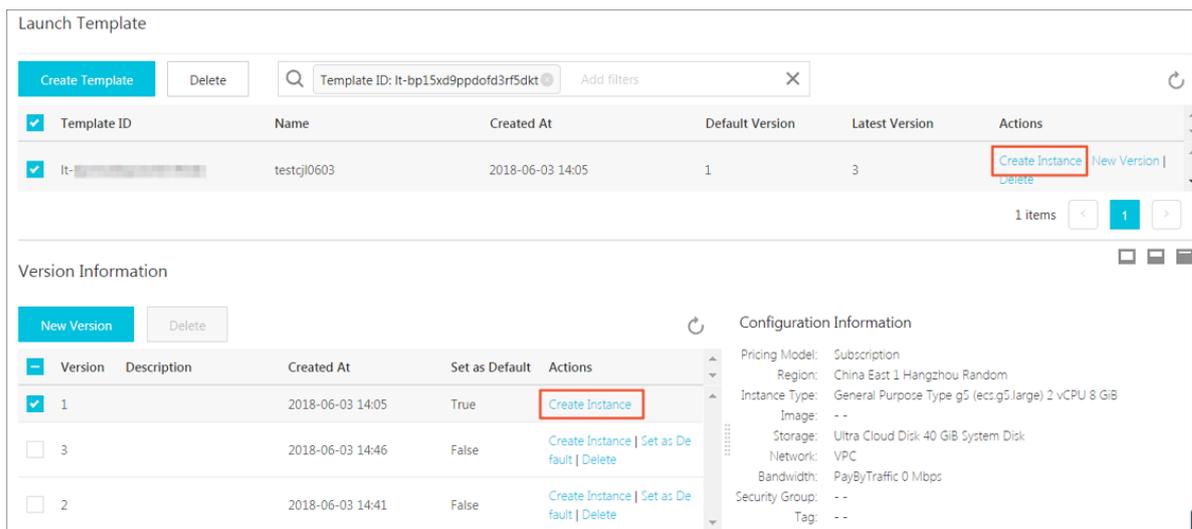
4.2.3 Use a launch template

Prerequisites

You have *created a template* or *created a version*.

Procedure

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Launch Templates**.
3. Locate the template or version that you want to use, and then click **Create Instance** in the **Actions** column.



4. On the ECS buy page, select the required template and version.



Note:

You can click the edit icon next to the target launch template to modify its configurations.

- If you want to create an instance using the Subscription billing method, select a subscription duration, read and confirm you agree to the Terms of Service, and then click **Create Order**. After you complete the payment, you can view the newly created instance in the ECS console.

If you want to create an instance using the Pay-As-You-Go billing method, read and confirm you agree to the Terms of Service, and then click **Create Instance**. After the instance is created successfully, you can view its details in the ECS console.

4.2.4 Delete a template or version

You can delete templates and versions through the ECS console. Once you delete a template, all associated versions of that template are also deleted.

Delete a version

- Log on to the [ECS console](#).
- In the left-side navigation pane, click **Launch Templates**.
- Select the target template ID.
- In the **Version Information** area, locate the version you want to delete and, in **Actions** column, click **Delete**.



Note:

You cannot delete the default template version. If the version you want to delete is the default version, change it to a non-default version, and then delete it. If you no longer need any versions of a single template, delete the template.

Version Information

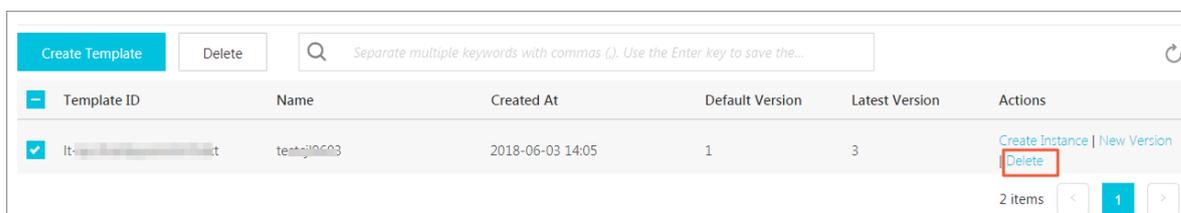
New Version
Delete
↻

	Version	Description	Created At	Set as Default	Actions
<input checked="" type="checkbox"/>	1		2018-06-03 14:05	True	Create Instance
<input type="checkbox"/>	3		2018-06-03 14:46	False	Create Instance Set as Default Delete

- Click **OK**.

Delete a template

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Launch Templates**.
3. Locate the version you want to delete, and click **Delete** in the **Actions** column.



4. Click **OK**.



Note:

When you delete a template, all versions of the template are also deleted.

4.3 Check instance information

Through the console, you can:

- [View all ECS instances under your account on the Overview page.](#)
- [View details of an ECS instance on the Instance Details page.](#)
- [View details of an ECS instance on Instance Details page](#)

View all ECS instances under your account on the Overview page

You can view information of all the ECS instances created by your account on the ECS **Overview** page, including:

- Total number of ECS instance, and numbers of instances under each status.
- Number of resources in different regions and numbers of ECS instances under each status.

The homepage of the ECS console is the **Overview** page by default.

View the information of ECS instances on the Instance List page

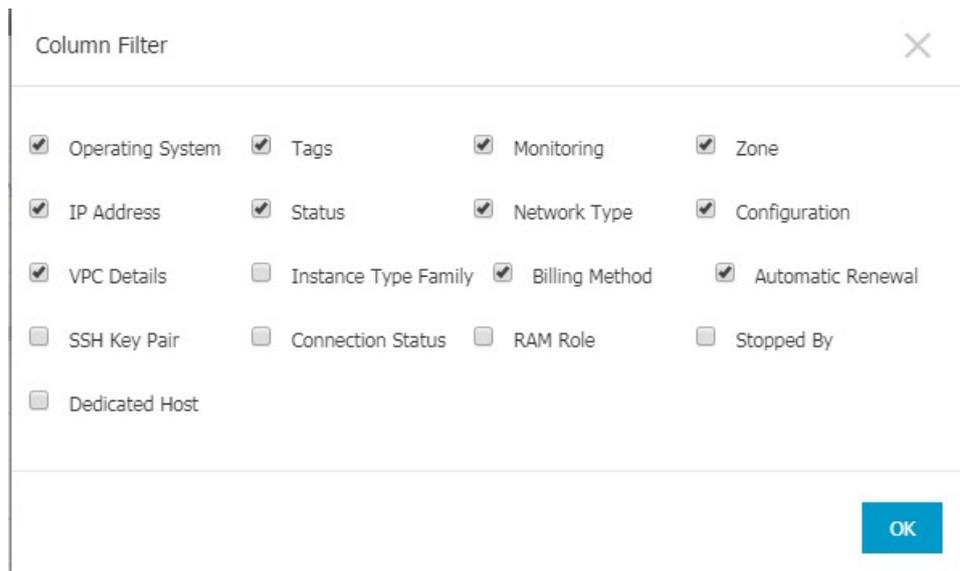
To navigate to the Instance List page, follow these steps:

1. Log on to the [ECS console](#).
2. On the left-side navigation pane, click **Instances**.
3. Select a region.

Here, you can see information of all the existing ECS instances in the selected region, including ECS instance ID/name, zone, IP addresses, status, network type, billing method, and actions. You can show or hide the displayed information of an instance by using the **Column Filter** feature.

1. In the upper-right corner of the Instance List, click the  icon.

2. Select the instance information to be displayed and then click **OK**.



View details of an ECS instance on Instance Details page

The **Instance Details** displays detailed information of a selected ECS instance.

To navigate to the **Instance Details** page, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Instances**.
3. Select a region.
4. Find the ECS instance you want to view the details of, and then click its instance ID.

On the **Instance Details** page, you can view the following information:

- **Basic Information**, including the ECS instance ID, instance name, region, zone, instance type, instance type family, image ID, key pair name (applies to Linux instances only), instance RAM role, and tags.
- **Configuration Information**, including CPU, memory, I/O optimization, operating system, IP addresses, billing method for bandwidth, current bandwidth, and VPC network information (applies to VPC instances only).

- **Payment Information**, including billing method, the mode to stop an instance, creation time, and automatic release schedule (applies to Pay-As-You-Go instances only).
- **Monitoring Information**, including CPU and network usage.

You can also switch from the **Instance Details** page to the **Disks**, **Instance Snapshots**, or **Security Groups** page to view resources related to this instance.

4.4 Change configurations

4.4.1 Overview of configuration changes

You can change the configurations of an instance and its Internet bandwidth after it is created.

Upgrade or downgrade instance configurations

You can only upgrade or downgrade the configurations of vCPU and memory (that is, *instance type family*) simultaneously by changing the instance type. Depending on the method of billing applied to your instance, you can change an instance type as follows:

- Subscription:
 - Upgrade: See [upgrade configurations](#). The new configurations take effect after you *restart the instance* in the console or by using the *RebootInstance* interface.
 - Downgrade: See [renewal for configuration downgrade](#). You can downgrade the configuration of an instance when you renew the instance. The new configuration takes effect after you *restart the instance* in the ECS console within the first seven days of the new billing cycle.
- Pay-As-You-Go: See [change configurations of Pay-As-You-Go instances](#). You must stop the instance to use this feature.

**Note:**

Stopping an instance disrupts services. Exercise caution when performing this action.

Adjust Internet bandwidth

You can adjust the Internet bandwidth of an instance. The methods vary according to your business needs and the billing method of the instance. The following table lists the methods.

Billing method	Supports permanent upgrade?	Is it effective immediately?	Available feature	Description
Subscription	Yes	Yes	Upgrade configurations of Subscription instances	Only applicable to VPC-Connected ECS instances to which no EIP addresses are attached, or classic network-connected ECS instances. The Internet and intranet IP addresses remain unchanged after you upgrade your configurations.
Subscription	Yes	Effective from next billing cycle	Renew for configuration downgrade	Adjust bandwidth in the new billing cycle. When the Internet bandwidth is set to 0 Mbit/s, the Internet IP address of a VPC-Connected instance is released in the new billing cycle, but that of a classic network-connected ECS instance is retained.
Pay-As-You-Go or Subscription	Yes	Yes	Change EIP Internet bandwidth	Only applicable to VPC-Connected instances to which EIP addresses are bound. You can adjust the Internet

Billing method	Supports permanent upgrade?	Is it effective immediately?	Available feature	Description
				bandwidth on an EIP address at any time.

Assign a public IP address

Assign a public IP address to an ECS instance while *creating it*. If you skip it, you can even assign after an ECS instance is created. However, the feature is only available for Subscription instances. For more information, see the following table.

Feature	Is it effective immediately?	Description
Upgrade configurations of Subscription instances	Yes	Only applicable to VPC-Connected ECS instances to which no EIP addresses are attached, or classic network-connected ECS instances. Set the Internet bandwidth to a non-zero value to assign a public IP address.
Renew for configuration downgrade	Effective from next billing cycle	

4.4.2 Upgrade configurations of Subscription instances

You can upgrade a Subscription-billed instance type.

you can also:

- **Convert the billing method of data disks** from Pay-As-You-Go to Subscription. The billing method of system disks cannot be changed.
- **Change the Internet bandwidth.** This applies to the instances in a classic network and instances in a VPC that are not bound with EIPs. If you do not purchase Internet bandwidth when creating an instance, no public IP address is assigned. In this case, you can use this feature to assign a public IP address to the instance when needed.

Fees

After upgrading the configuration, you must make up the difference for the rest of the current billing cycle.

Limits

This feature has the following limits:

- Only applicable to Subscription instances.
- You can upgrade an instance multiple times, but the time period between each upgrade must be at least five minutes.
- You must upgrade both the vCPU cores and memory size of an instance type. That is, you cannot upgrade one item separately.
- Not supported within or between such instance type families: d1, d1ne, i1, i2, ga1, gn5, f1, f2, f3, ebmc4, ebmg5, sccg5, and scch5. For the instance type families that support this feature and the rules for upgrading instance types, see [instance type families that support upgrading instance types](#).
- This feature can be used to change the Internet bandwidth only for VPC instances bound with no EIPs and classic network instances.
- You can change the billing method from Pay-As-You-Go to Subscription only for data disks, not for system disks.
- In the current billing cycle, if you have already performed the [renewal for configuration downgrade](#) operation, you cannot upgrade the configuration until a new billing cycle begins.
- After upgrading an instance type or changing the Internet bandwidth of a classic network instance from 0 Mbps to a non-zero value for the first time, you must restart the instance on the console or through the [RebootInstance](#) API to activate the new configuration.

Procedure

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Instances**.
3. Select a region.
4. Select the Subscription instance to upgrade and, in the **Actions** column, click **Change Configuration**.
5. Select **Upgrade Configuration** and click **Continue**.
6. On the **Upgrade Configuration** page, perform any of the following operations:
 - Select a new **Instance Type**.

**Note:**

The page displays all the new instance types that are available for your instance.

- If a [Pay-As-You-Go-billed data disk is attached](#) to your instance, you can convert its billing method to Subscription.
- If the instance is a classic network instance, or is VPC-Connected and not bound with an EIP, you can modify its Internet bandwidth.

**Note:**

If you do not purchase Internet bandwidth when creating an instance, no public IP address is assigned. In this case, you can use this feature to assign a public IP address to the instance when needed.

7. Confirm your order details, and then click **Create Order**. Follow additional instructions as required.
8. After upgrading an instance type or changing the Internet bandwidth of a classic network instance from 0 Mbps to a non-zero value for the first time, you must restart the instance through the console or through the [RebootInstance](#) API to activate the new configuration.

**Note:**

You do not have to restart a VPC instance if this upgrade configuration is the first time its Internet bandwidth is increased from 0 Mbps to a non-zero value.

You can also use the [DescribeResourcesModification](#) API to query the instance types that can be upgraded.

4.4.3 Change configurations of Pay-As-You-Go instances

This article describes how to change configurations of Pay-As-You-Go instances. For information about how to change configurations of (Subscription) instances, see [overview of configuration changes](#).

**Note:**

Changing instance configurations requires stopping your instance, which disrupts services. Exercise caution when performing this action.

Limits

- You can upgrade an instance multiple times, but the time period between each upgrade must be at least five minutes.

- Not supported within or between such instance type families: d1, d1ne, i1, i2, ga1, gn5, f1, f2, f3, ebmc4, ebmg5, sccg5, and scch5. For more information, see [instance type families that support upgrading instance types](#).

Prerequisite

The instance has been stopped.

Procedure

To change instance type configurations of the instance, follow these steps:

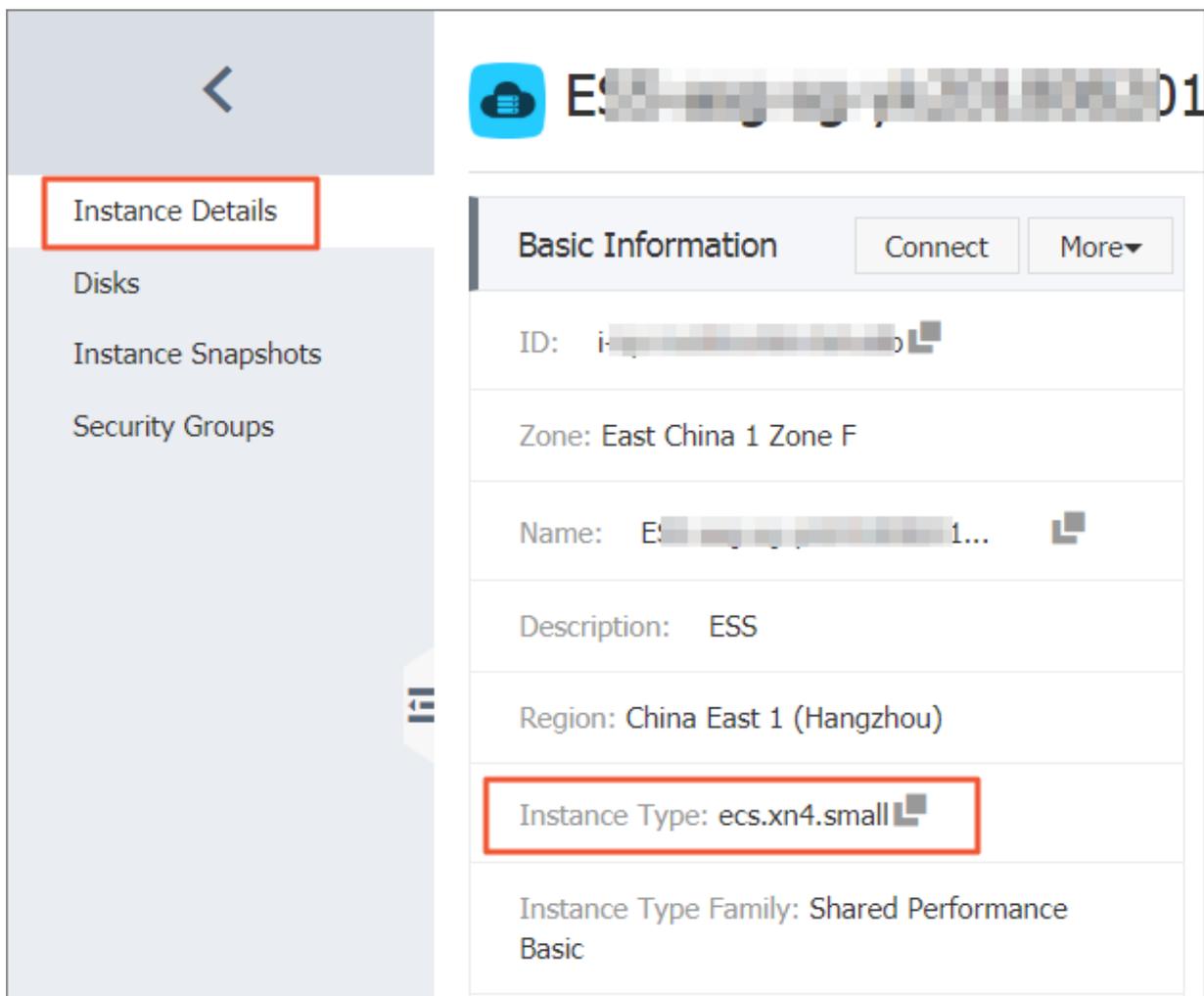
1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Instances**.
3. Select the target region.
4. In the **Actions** column, click **Change Instance Type**.
5. On the **Instance Type** page, select the desired instance type and click **Confirm**.



Note:

You can also enter the instance type information in the search box to filter instance types.

Once the change is complete, it takes effect immediately. You can view the instance type information in the **Basic Information** area of the **Instance Details** page, as shown in the following figure.



Then, restart the instance to restore your services.

You can also use the [DescribeResourcesModification](#) API to query the instance types that can be changed.

4.4.4 Instance type families that support instance type upgrades

This article describes the instance type families that support instance type upgrades.

Restrictions

Upgrading instance types has the following impacts:

- Classic network instances:
 - For *phased-out instance types*, when a non-I/O optimized instance is upgraded to an I/O optimized instance, changes are made to the private IP address, the driver name, and the software authorization code. For Linux instances, Basic Cloud Disks (`cloud`) are

recognized as `xvda` or `xvdb`, while Ultra Cloud Disks (`cloud_efficiency`) and SSD Cloud Disks (`cloud_ssd`) are recognized as `vda` or `vdb`.

— For *available instance types*, changes are made to the private IP address of the instance.

- VPC instances:

For *phased-out instance types*, when a non-I/O optimized instance is upgraded to an I/O optimized instance, changes are made to the driver name and the software authorization code.

For Linux instances, Basic Cloud Disks (`cloud`) are recognized as `xvda` or `xvdb`, while Ultra Cloud Disks (`cloud_efficiency`) and SSD Cloud Disks (`cloud_ssd`) are recognized as `vda` or `vdb`.

Instance type families that support upgrading instance types



Note:

Each instance type is available only in specific zones. Before upgrading an instance type, check if the target instance type (family) is available in the current zone.

In the following table, the target instance type families apply to both Subscription and Pay-As-You-Go instances.

Source instance type family	Target instance type family
g5, r5, c5, ic5	<ul style="list-style-type: none"> • g5, r5, c5, ic5 • sn1ne, sn2ne, se1ne, c4, cm4, ce4, hfc5, hfg5, re4, t5, n4, mn4, xn4, e4
sn1ne, sn2ne, se1ne	<ul style="list-style-type: none"> • sn1ne, sn2ne, se1ne • c4, cm4, ce4, hfc5, hfg5, g5, r5, c5, ic5, re4, t5, n4, mn4, xn4, e4
se1	<ul style="list-style-type: none"> • se1 • sn1, sn2, n1, n2, e3, sn1ne, sn2ne, se1ne, c4, cm4, ce4, hfc5, hfg5, g5, r5, c5, ic5, re4, t5, n4, mn4, xn4, e4
n4, mn4, xn4, e4	<ul style="list-style-type: none"> • n4, mn4, xn4, e4 • sn1, sn2, se1, n1, n2, e3, sn1ne, sn2ne, se1ne, c4, cm4, ce4, hfc5, hfg5, g5, r5, c5, ic5, re4, t5
re4	<ul style="list-style-type: none"> • re4

Source instance type family	Target instance type family
	<ul style="list-style-type: none"> sn1ne, sn2ne, se1ne, c4, cm4, ce4, hfc5, hfg5, g5, r5, c5, ic5, t5, n4, mn4, xn4, e4, ecs.se1.14xlarge
hfc5, hfg5	<ul style="list-style-type: none"> hfc5, hfg5 sn1ne, sn2ne, se1ne, c4, cm4, ce4, g5, r5, c5, ic5, re4, t5, n4, mn4, xn4, e4
gn4	gn4
gn5i	gn5i
gn6v	gn6v
t5	<ul style="list-style-type: none"> t5 sn1ne, sn2ne, se1ne, c4, cm4, ce4, hfc5, hfg5, g5, r5, c5, ic5, re4, n4, mn4, xn4, e4
t1, s1, s2, s3, m1, m2, c1, c2	<ul style="list-style-type: none"> t1, s1, s2, s3, m1, m2, c1, c2 sn1, sn2, se1, n1, n2, e3, sn1ne, sn2ne, se1ne, c4, cm4, ce4, hfc5, hfg5, g5, r5, c5, ic5, re4, t5, n4, mn4, xn4, e4
n1, n2, e3	<ul style="list-style-type: none"> n1, n2, e3 sn1, sn2, se1, sn1ne, sn2ne, se1ne, c4, cm4, ce4, hfc5, hfg5, g5, r5, c5, ic5, re4, t5, n4, mn4, xn4, e4
sn1, sn2	<ul style="list-style-type: none"> sn1, sn2 se1, n1, n2, e3, sn1ne, sn2ne, se1ne, c4, cm4, ce4, hfc5, hfg5, g5, r5, c5, ic5, re4, t5, n4, mn4, xn4, e4
c4, ce4, cm4	<ul style="list-style-type: none"> c4, ce4, cm4 sn1ne, sn2ne, se1ne, hfc5, hfg5, g5, r5, c5, ic5, re4, t5, n4, mn4, xn4, e4

4.4.5 Change EIP Internet bandwidth

If you are using a VPC-Connected ECS instance, and [an Elastic IP \(EIP\) address is bound to it](#), you can use the **Change Bandwidth** feature to change the Internet bandwidth as needed, regardless of your billing method.

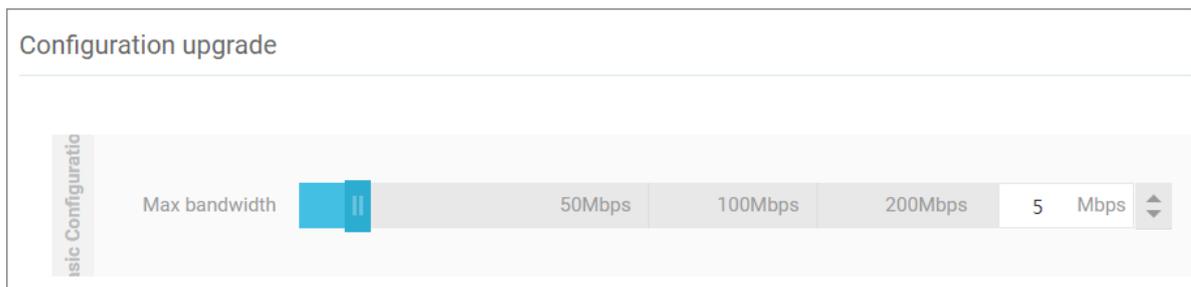
Restrictions

The **Change Bandwidth** feature only applies to VPC-Connected instances, that are bound with an Elastic IP (EIP) address.

Change bandwidth

To change the Internet bandwidth of an EIP address, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Instances**.
3. Select a region.
4. Find an instance to which an EIP address is bound, and in the **Actions** column, click **Change Configuration**.
5. In the **Change Configurations** dialog box, select change bandwidth, and click **Continue**.
6. On the **Confirm Order** page, set the new peak bandwidth.



7. Click **Activate** and follow the instructions to complete the configuration.

Related operations

For Subscription instances of the classic network type or VPC-Connected ECS instances to which no EIP addresses are bound, you can use the following features to change the Internet bandwidth:

- Use the [Upgrade Configuration](#) feature to change the Internet bandwidth immediately.
- Use the [Renew for Configuration Downgrade](#) feature to change the Internet bandwidth for the next billing cycle.

4.5 Change the operating system

You can convert the OS running on your ECS instance to another supported OS through the ECS console.

To change the operating system, you must change the system disk of an instance:

- If you want to use a custom image, see [change the system disk \(custom image\)](#).

- If you want to use a public image, see [change a system disk \(public image\)](#).

**Note:**

Currently, instances that are hosted in regions outside of mainland China do not support swapping between Linux and Windows OSs. If your instance is hosted in one of these regions, you can only change its version of Windows OS to another version of Windows, or replace its current Linux OS with another Linux OS.

4.6 Reset an instance password

This article describes how to use the Reset Password feature to specify a new logon password for an instance.

**Note:**

You must restart an instance after its password is reset, which may disrupt services. Exercise caution when performing this action.

Prerequisite

The instance must be in a stable status, such as **Stopped** and **Running**. For more information, see [ECS instance life cycle](#).

Procedure

To reset a password for one or multiple ECS instances, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Instances**.
3. Select the target region.
4. According to the number of instances to be operated, do the following:
 - To reset the password for one instance, find the target instance and, in the **Actions** column, select **More** > **Password/Key Pair** > **Reset Password**.
 - To reset the passwords for multiple instances, select the target instances and, under the instance list, click **Reset Password**.
5. Specify a new valid password, and click **Submit**.
6. To make the password change take effect, change the instance status to one of the following:
 - **Running**: [Restart an instance](#) in the console.
 - **Stopped**: Start the instance in the console.

Related operations

To modify the password: follow [Step 3: Connect to an instance](#), then [change the password in the ECS instance](#).

4.7 Start or stop an instance

This article describes how to start or stop an ECS instance.

Start an instance

You can start an instance in the ECS console. When an instance starts successfully, it is in the **Running** status.

Prerequisite

The instance must be in the **Stopped** status.

Procedure

To start an instance, follow these steps:

1. Log on to the [ECS Management console](#).
2. In the left-side navigation pane, click **Instances**.
3. Select the target region.
4. Find the instance to be started and, in the **Actions** column, select **More > Instance Status > Start**. If you want to start multiple **Stopped** instances, select the required instances and then, under the instance list, click **Start**.
5. Read and confirm you agree to the note displayed in the dialog box by clicking **OK**.

The instance is in the **Running** status after it is started.

Stop an instance

To stop an instance is to shut it down. You can stop an ECS instance in the ECS console. When an instance stops successfully, it is in the **Stopped** status.



Note:

Stopping an instance disrupts services. Exercise caution when performing this action.

If you stop a (Subscription) instance before its billing cycle is completed, the bill for that cycle is not affected. If the auto-renewal service is activated, you are still billed for the stopped instance at the start of each new billing period.

For a Pay-As-You-Go instance, its network type and the No Fees for Stopped Instances (VPC-Connected) feature determine billing:

- VPC: If the **No Fees for Stopped Instances (VPC-Connected)** feature is enabled, you can decide whether to continue being billed for the instance. However, you are still billed for other ECS-related resources. For more information, see [no fees for stopped instances \(VPC-Connected\)](#). If this feature is not enabled, billing continues after the instance is stopped.
- Classic network: A stopped instance still incurs fees. Billing stops only after you [release the instance](#).

Prerequisite

The instance is in the **Running** status.

Procedure

To stop an instance, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Instances**.
3. Select the target region.
4. Find the instance to be stopped and, in the **Actions** column, select **More > Instance Status > Stop**. If you want to stop multiple **Running** instances, select the required instances and then, under the instance list, click **Stop**.
5. According to the billing method and network type of the instance, complete the required actions:
 - Subscription instance or classic network pay per volume instance: In the **Stop Instance** dialog box, select **Stop** or **Force Stop**, and then click **OK**.
 - A VPC-Connected Subscription instance:
 - If the **No Fees for Stopped Instances (VPC-Connected)** feature is enabled, read the **Notice**, select **Stop** or **Force Stop** in the **Stop Instance** dialog box, select a mode (whether to keep the instance after stopping and continue charging), and then click **OK**.

Stop
✕



Operation will be executed on the selected **1 instance(s)** ▼. Are you sure you want to proceed?

I want to :

Stop
 Force Stop

Stop Instance

Keep Instance with Fees

Operation will be executed on the selected **1 instance(s)** ▼ after instance is stopped, it will not be billed.

The expiration date will not change after the Subscription instance is stopped.

If you stop the instance to replace the system disk, reinitialize the disk, change the instance specifications, modify the IP address of the private network, etc., you are advised to select the "Always keep the instance after the stop and continue to charge" option to avoid startup failure.

OK

Cancel

- If the **No Fees for Stopped Instances (VPC-Connected)** feature is disabled, in the **Stop Instance** dialog box, select **Stop** or **Force Stop**.



Note:

To disable the **No Fees for Stopped Instances (VPC-Connected)** feature, see [no fees for stopped instances](#).

Once the instance is successfully stopped, the instance enters the **Stopped** status. For a VPC-Connected Pay-As-You-Go instance, if you select not to keep the instance, **Stop Instance, No Fees** is shown in the instance list. Otherwise, **Keep Instance, Fees Apply** is shown. For other ECS instances, no information is shown.

Related APIs

Start instance: [StartInstance](#)

Stop instance: [StopInstance](#)

4.8 Restart an instance

You can restart your instances through the ECS console.

**Note:**

- Only instances in the **Running** status can be restarted.
- Restarting an instance may disrupt services. Exercise caution when performing this action.

Procedure

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Instances**.
3. Select the target region.
4. Select one or multiple instances as needed. All selected instances must be in the **Running** status.
5. Click **Restart**, and then click **OK**.

4.9 Reactivate an instance

For a Pay-As-You-Go instance, if the due date of an overdue payment (T) is not settled within 15 days after the due date (T+15), the instance is stopped due to overdue payment and its status changes to **Expired**. You must submit a ticket to settle the payment and reactivate your instance within 30 days after the due date (T+30). Otherwise, the instance is released and the data cannot be recovered.

**Note:**

If you fail to reactivate the ECS instance within 30 days after the due date (T+30), the instance is automatically released 30 days after the due date and the data cannot be recovered.

Prerequisites

The Pay-As-You-Go instance is in the **Expired** status.

Submit a job statement to clear the bill.

Procedure

To reactivate an instance in the ECS console, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Instances**.

3. Select the target region.
4. Select the instance to be reactivated, and then select **More > Reactivate** at the bottom of the instance list.
5. Choose whether to reactivate the instance immediately or later at a specified time.

If you choose to reactivate immediately, the selected instance returns to the **Running** status after about 10 minutes.

Alternatively, you can also complete the task by calling the ECS API [ReactivateInstances](#).

4.10 Release an instance

You can release a Pay-As-You-Go instance when you no longer need it to avoid excess charges.

For a Pay-As-You-Go instance, if the [No fees for stopped VPC instances](#) feature is not enabled, charges continue to incur until the instance is released.

For a Subscription instance, the instance is automatically released after the billing cycle expires.

To release a Pay-As-You-Go instance, you can choose either of the following options:

- Release immediately, which releases the pay-per-order instance at once.
- Scheduled Release, which customizes the releasetime of your Pay-As-You-Go instance. The time to release the instance must be at least 30 minutes from the current time. Applying new schedules overwrites the previous ones.



Note:

After an instance is released, its data cannot be recovered. We recommend that you [create a snapshot](#) to back up data before releasing an instance.

Release an instance immediately

To release an instance immediately, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Instances**.
3. Select the target region.
4. Set release:
 - If you want to release only one instance, find the instance that you want to release and then, in the **Actions** column, select **Manage > Release**.

- If you want to release multiple instances, find the Pay-As-You-Go instances according to the **Billing Method**, select multiple instances to release, and click **Release Setting** at the bottom of the list.

5. In the dialog box, select **Release Now**.

6. Click **Next**, and then click **OK**.

Enable automatic release

To enable automatic release, follow these steps:

1. Log on to the [ECS console](#).

2. In the left-side navigation pane, click **Instances**.

3. Select the target region.

4. Set the release:

- If you want to release only one instance, find the instance that you want to release, and in the **Actions** column, select **Manage > Release**.
- If you want to release multiple instances, find the Pay-As-You-Go instances according to the **Billing Method**, select multiple instances to release, and click **Release Setting** at the bottom of the list.

5. In the dialog box, select **Scheduled Release**.

6. Turn on the Automatic Release switch, and specify the release date and time. The earliest setting can only be set to automatically release an instance after 30 minutes.

Release Setting ⓘ Release instance ✕

*Release Mode: Release Now Scheduled Release

Automatic Release:

*Released At: 

*Released At:  : 

Note:

- The scheduled release interval is five minutes. The system will stop charging the instance at the scheduled release time.

7. Click **Next**, and then click **OK**.

Disable automatic release

If you want to cancel the automatic release schedule of a Pay-As-You-Go instance, you can disable the feature.

To disable the automatic release feature, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Instances**.
3. Select the target region.
4. Set the release:
 - If you want to disable automatic release for an instance, find the instance and then, in the **Actions** column, select **Manage > Release**.
 - If you want to disable automatic release for multiple instances, find the Pay-As-You-Go instances according to the **Billing Method**, select the target instances, and click **Release Setting** at the bottom of the list.
5. In the dialog box, select **Scheduled Release**.
6. Turn off the Automatic Release switch.

7. Click **Next**, and then click **OK**.

Related API

[DeleteInstance](#)

4.11 Change IP addresses

4.11.1 Change public IP address

If your instance is assigned a public IP address, you can change the address within six hours after the instance is created regardless of whether the instance is in a classic network or in a VPC network.

Limits

- The instance must be assigned a public IP address. To verify the public IP address, view the public IP address in the **IP Address** column from the **Instance List** in the ECS console, as displayed in the following figure.

<input type="checkbox"/> Instance ID/Name	Zone	IP Address
<input type="checkbox"/> i- launch-advisor-2018022...gj	China East 1 Zone G	4 1(Internet IP Address) 1 3(Private IP Address)



Note:

- If the public network IP address is not assigned at the time of creation of the instance, after the instance is created successfully, you can assign the public IP address by upgrading or downgrading the network bandwidth configuration. For more information, see [overview of configuration changes](#).
 - If the public network IP address is not assigned during the creation of a Pay-As-You-Go instance, after the instance is created successfully, public IP address cannot be assigned. You can only [bind an elastic IP \(EIP\) address](#).
- The instance must be in the **Stopped** status.
 - The instance has existed for less than six hours.



Note:

After six hours, for a VPC instance in a VPC network, you can [convert public IP address to EIP address](#). Instances in the classic network cannot have their public IP address converted.

- You can change the public IP address of an instance a maximum of three times.

Prerequisite

The instance must be in the **Stopped** status.

Procedure

To change the public IP address, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Instances**.
3. Select the target region.
4. Find the target instance to change the public IP address and then, in the **Actions** column, select **More > Network and Security Group > Change Public IP**.

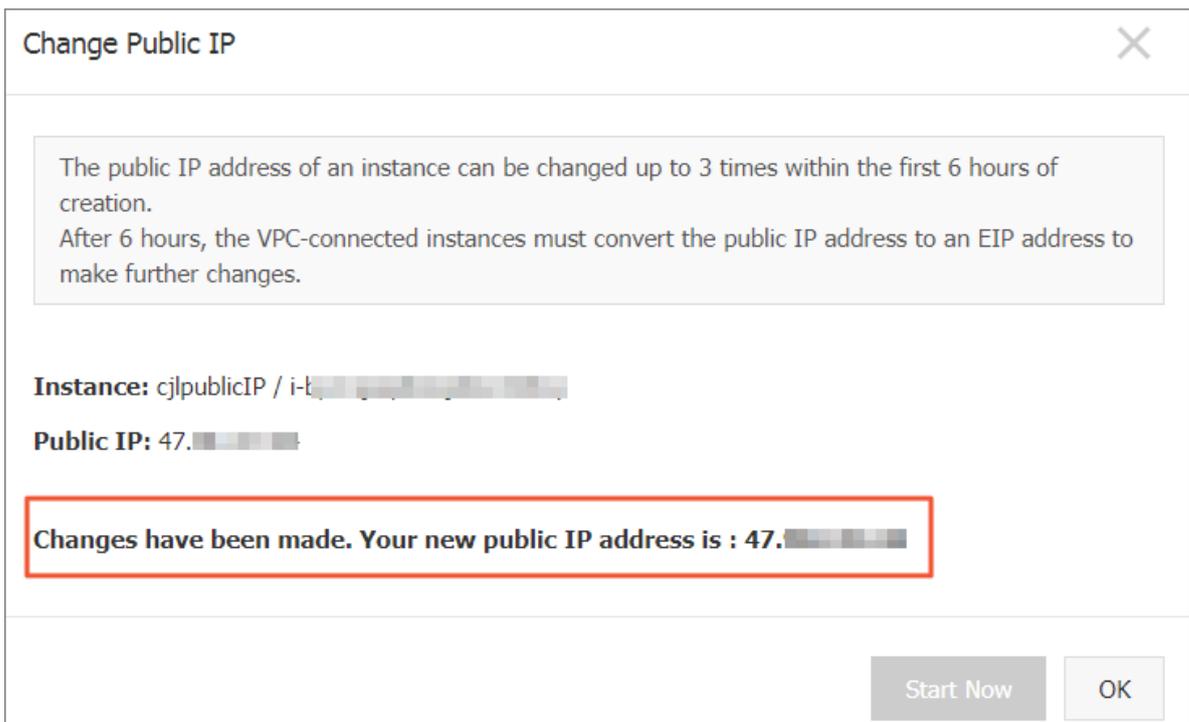


Note:

If the instance has existed for more than six hours, the **Change Public IP** option in the **More** drop-down menu is not available.

5. Click **Start Now**.

A new public IP address is displayed as shown in the following figure.



6. Click **OK**.

Related operation

You can [change the private IP of an ECS instance](#).

4.11.2 Convert public IP address to EIP address

This article describes how to convert the assigned public IP address of an ECS instance in a VPC network, (referenced as VPC instance for short in this article), to an elastic public IP (EIP) address. After conversion, you can retain the public IP address and bind it to another ECS instance.

Limits

To convert a public IP address to an EIP address, consider the following limits:

- You cannot undo this action. Exercise caution when converting an assigned public IP address to an EIP address.
- Only a VPC instance assigned a public IP address is supported.
- Only a VPC instance in the **Stopped** or **Running** status is supported.
- Only a VPC instance that does not have any inactivated specification changes is supported.
- Only a VPC instance that is not within the last 24 hours of its life cycle is supported.



Note:

- The conversion has no effect on the Internet access of the VPC instance. It does not cause transient traffic interruption.
- The billing method of the public traffic remains unchanged.
- After conversion, the EIP address is charged separately. For more information about billing of EIP addresses, see [EIP billing](#). You can go to the [Usage Records](#) page in the **Billing Management** to download the **Elastic Public IP** usage record.

Procedure

To convert a public IP address to an elastic public IP (EIP) address, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Instances**.
3. Select the region.

4. Find the target **VPC instance** to convert the public IP address, in the **Actions** column, select **More > Convert to EIP**.
5. In the **Convert to EIP** dialog box, read the note and click **OK**.
6. Refresh the instance list.

After the public IP address is converted to an EIP address, the IP address is followed by **(Elastic IP Address)**.

<input type="checkbox"/> Instance ID/Name	Zone	IP Address	Status ▾	Network Type ▾
<input type="checkbox"/> i-b[redacted] ConvertIpTest	China East 1 Zone B	4[redacted]3(Elastic IP Address) 172.16.21.212(Private IP Address)	● Running	VPC

Click the IP address to go to the EIP console to manage the EIP address.

Follow-up operations

After the public IP address is converted to an EIP address, you can unbind the EIP address from the instance and bind it to another instance. You can also release the EIP address. For more information, see [unbind and release an EIP](#).

Related API

You can use the [ConvertNatPublicIpToEip](#) interface to convert a public IP address to an EIP address. Currently, only SDK 4.3.0 or a later version supports this interface.

[Download](#) the latest SDK.

4.11.3 Change the private IP of an ECS instance

After creating an ECS instance in a VPC network, you can change the private IP address and can change the VSwitch of the ECS instance.

Procedure

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Instances**.
3. Select the target region.
4. In the **Actions** column, click **More > Instance Status > Stop**.
5. When the instance is stopped, click the instance ID to go to its **Instance Details** page.
6. In the **Configuration Information** panel, click **More > Modify Private IP Address**.

7. In **Modify Private IP Address** dialog, select a VSwitch, and then click **Modify**.

Make sure the current VSwitch and the selected VSwitch are in the same zone.

**Note:**

Enter a new IP address if you do not want to change the VSwitch of the ECS instance.

Modify Private IP Address ✕

Instance: i-b [redacted]

Zone: China East 1 Zone G

VSwitch: vs-[redacted] 4090 private IP addresses available

The VSwitch must be in the same zone as the instance.

Private IP Address: 17[redacted]7

The specified private IP address must be unoccupied in the VSwitch network segment. If no private IP address is specified, an idle private IP address will be automatically assigned to the ECS instance.

Modify **Cancel**

8. Go back to the instance page and, in the **Actions** column, click **More > Instance Status > Restart** to make the new private IP address take effect.

4.12 User-defined data and metadata

4.12.1 Metadata

The metadata of an instance is the basic information of the ECS instance, including the Instance ID, IP Address, OS, and other related information. You can use an instance's metadata to better manage and configure an instance.

**Note:**

If you manually change some instance information, this change will not be reflected in the instance metadata.

Limits

The metadata is only applicable for VPC-Connected instances.

Get the metadata

Linux instance

1. [Connect to a Linux instance by using a password.](#)
2. Run `curl http://100.100.100.200/latest/meta-data/` to access the root directory of the metadata.
3. Add the specific metadata name to the preceding command to access the specified metadata.

For example:

- Run `curl http://100.100.100.200/latest/meta-data/instance-id` to get the ID of an ECS instance.
- Run `curl http://100.100.100.200/latest/meta-data/image-id` to get the image ID of an ECS instance.

Windows instance

1. [Connect to a Windows instance.](#)
2. Use PowerShell to run `Invoke-RestMethod http://100.100.100.200/latest/meta-data/` to get the metadata.
3. Add the specific metadata name to the preceding command to access the specified metadata.

For example:

- Run `Invoke-RestMethod http://100.100.100.200/latest/meta-data/instance-id` to get the ID of an ECS instance.
- Run `Invoke-RestMethod http://100.100.100.200/latest/meta-data/image-id` to get the image ID of an ECS instance.

List of instance metadata

Metadata name	Description	Version
<code>dns-conf/nameservers</code>	DNS configurations for an instance.	2016-01-01
<code>eipv4</code>	EIP address	2016-01-01
<code>hostname</code>	The OS name of an instance.	2016-01-01

Metadata name	Description	Version
dns-conf/nameservers	DNS configurations for an instance.	2016-01-01
image-id	ID of the image that is selected at the time of instance creation.	2016-01-01
image/market-place/product-code	Product code of the image in the Alibaba Cloud Marketplace.	2016-01-01
image/market-place/charge-type	Billing method of the image in the Alibaba Cloud Marketplace.	2016-01-01
instance-id	The instance ID	2016-01-01
mac	The MAC address of the instance. If multiple network interface cards exist in an instance, this metadata indicates the MAC address of eth0.	2016-01-01
network-type	Network type (only applicable for VPC network).	2016-01-01
ntp-conf/ntp-servers	The address of a NTP server.	2016-01-01
owner-account-id	The aliuid of the instance owner.	2016-01-01
private-ipv4	Private IP address.	2016-01-01
public-ipv4	Public network IP address.	2016-01-01
public-keys	A list of all public keys of the current instance.	2016-01-01
region-id	The region where the instance is located.	2016-01-01
zone-id	Zone ID of the zone where the ECS instance is located.	2016-01-01
serial-number	The serial number of an instance.	2016-01-01
source-address	The source of YUM/APT (only applicable for a Linux instance).	2016-01-01
kms-server	Activate the server (only applicable for a Windows instance).	2016-01-01
wsus-server/wu-server	Update the server (only applicable for a Windows instance).	2016-01-01
wsus-server/wu-status-server	The server that monitors the update status of an instance (only applicable for a Windows instance).	2016-01-01
vpc-id	ID of the VPC that an instance is in.	2016-01-01

Metadata name	Description	Version
dns-conf/nameservers	DNS configurations for an instance.	2016-01-01
vpc-cidr-block	The CIDR block of the VPC that an instance is in.	2016-01-01
vswitch-cidr-block	The CIDR block of the VSwitch that an instance is in.	2016-01-01
vswitch-id	ID of the VSwitch that an instance is in.	2016-01-01
ram/security-credentials/[role-name]	<p>The temporary STS credential is generated according to the policy of a RAM role. Only available when you specify a RAM role to an ECS instance. When you use this metadata to get the STS credential, <code>[role-name]</code> must be replaced with the actual RAM role name you create or you have created.</p> <div style="border: 1px solid #ccc; background-color: #f9f9f9; padding: 5px; margin-top: 10px;">  Note: A new STS credential is available 30 minutes prior to the expiration of the old one. </div>	2016-01-01
instance/spot/termination-time	The spot instance release time is based on the instance operating system time zone. It is specified in UTC format. For example, YYYY-MM-DDThh:mm:ssZ For example, 2018-04-07T17:03:00Z.	2016-01-01
network/interfaces/macs	The MAC address list of Network Interface Controllers (NICs).	2016-01-01
network/interfaces/macs/[mac]/network-interface-id	The unique ID of the NIC, <code>[mac]</code> must be replaced with the actual MAC address.	2016-01-01
instance/virtualization-solution	The virtualization solution, which is ECS Virt 1.0 / 2.0	2016-01-01
instance/virtualization-solution-version	The internal build version.	2016-01-01
instance/last-host-landing-time	The latest update time of the physical server to which your instance is hosted on.	2016-01-01
instance-identity/document	<i>Instance identity</i> Instance identity document.	2016-01-01
instance-identity/pkcs7	Instance identity signature.	2016-01-01

Other data exposing to instance metadata

- **Maintenance:** For on-going [system events](#), run the following command to get the latest data:

```
curl http://100.100.100.200/latest/maintenance/active-system-events
```

- **Signature:** For the [instance identity](#) document, run the following command to get the latest data:

```
curl http://100.100.100.200/latest/dynamic/instance-identity/document
```

- **Boot behavior configuration:** For an [instance's user data](#), run the following command to get the latest data:

```
curl http://100.100.100.200/latest/user-data
```

4.12.2 User data

You can use user data of an ECS instance to customize its startup behavior and to pass data into the instance. You can specify user data when creating an instance ([RunInstances](#)) and customize startup behavior such as automatically update software packages, enable services, print logs, install dependencies, initialize web services, and more. User data of an ECS instance is implemented primarily through different types of scripts. User data can also be used as common data to be referenced in the instances.

Instructions for use

To configure instance user data, note that:

- Only VPC-Connected instances are supported.
- For [phased-out instance types](#), they must be I/O optimized. Other [instance type families](#) are not limited for I/O optimized.
- Instance user data requires Base64 encoding before being passed in, and the user data before encoding cannot exceed 16 KB.
- The instance must use an official image or a user image that is created from an official image.

The operating system must be one of the following:

Windows instances	Linux instances
Windows Server 2016 64-bit Windows Server 2012 64-bit Windows Server 2008 64-bit	CentOS Ubuntu SUSE Linux Enterprise OpenSUSE Debian Aliyun Linux

Module frequency

After the instance enters the **Running** state, use your Alibaba Cloud primary account to run the user data of the instance, followed by the initialization or `/etc/init` information.

After you modify the instance user data, depending on the type of scripts and modules that are used, the modified user data is or is not run. For example:

- If you configure user data by using a shell script, such as a [user-data script](#), the modified user data is not run.
- If the user data configures modules such as Byobu, Set Hostname, and Set Passwords, the modified user data is not run.
- If the user data configures modules such as `bootcmd`, `update_etc_hosts`, and `yum_add_repo`, the modified user data is run.

For more information, see [modules](#).

Set user data

For this example, assume that you write user data development in a Windows environment, and you use [Upstart Job](#) to configure the user data.

1. Use an editor to create a text file, such as Notepad++.
2. Edit the script related to the user data in the text file.



Note:

The first line must meet the format requirements of the instance user data script, such as `#!/bin/sh`, `#cloud-config`, `#upstart-job`, `[bat]` and `[powershell]`. For more information, see [Linux instance user data](#) and [Windows instance user data](#).

3. Debug the script file to confirm that the content is valid.
4. (Optional) If you make a [Gzip compression content](#), compress the script file in .gz format.
5. (Optional) If you are creating an [Include file](#) or a [Gzip compression script](#), upload script file to available storage services, obtain the link, and set the valid period of the link.

We recommend that you use Alibaba Cloud OSS to create links. For more information, see [upload an object](#) or [set lifecycle](#).

6. Log on to the [ECS console](#).
7. Follow the instructions in [creating an instance](#) to create a Linux instance.

**Note:**

The instance must be VPC-Connected, and you must select a [image](#) that meets the requirement. For [phased-out instance types](#), I/O optimized instances are required. Other [instance type families](#) are not limited in terms of I/O optimized.

After creating the instance, select **Advanced (based on instance RAM roles or cloud-init) use text form** and enter your **user data**. If your user data has been encrypted by Base64 encoding, click **The text is Base64-encoded**.

8. Wait for the instance to be created.
9. [Connect](#) to your instance.
10. View the results of the user data. If a failure occurs, check the relevant log files. The following is an output example of user data on a CentOS instance by using the upstart job script:

```
[root@ ~]# cd /etc/init
init/  init.d/  inittab
[root@ ~]# cd /etc/init/
[root@ init]# ls
part-001.conf
[root@ init]# cat part-001.conf
#upstart-job
#!/bin/sh
echo "Hello World. The time is now $(date -R)!" | tee /root/output.txt[root@ ~]#
```

In the preceding figure, the startup job file `part-001.conf` is generated in the `/etc/init` folder.

Related API: [RunInstances](#) + Parameters `UserData`

View user data

You can view user data of an instance from the server `100.100.100.200`. To do so, follow these steps:

1. Connect to the target instance.
2. In the instance, depending on your OS, run one of the following:

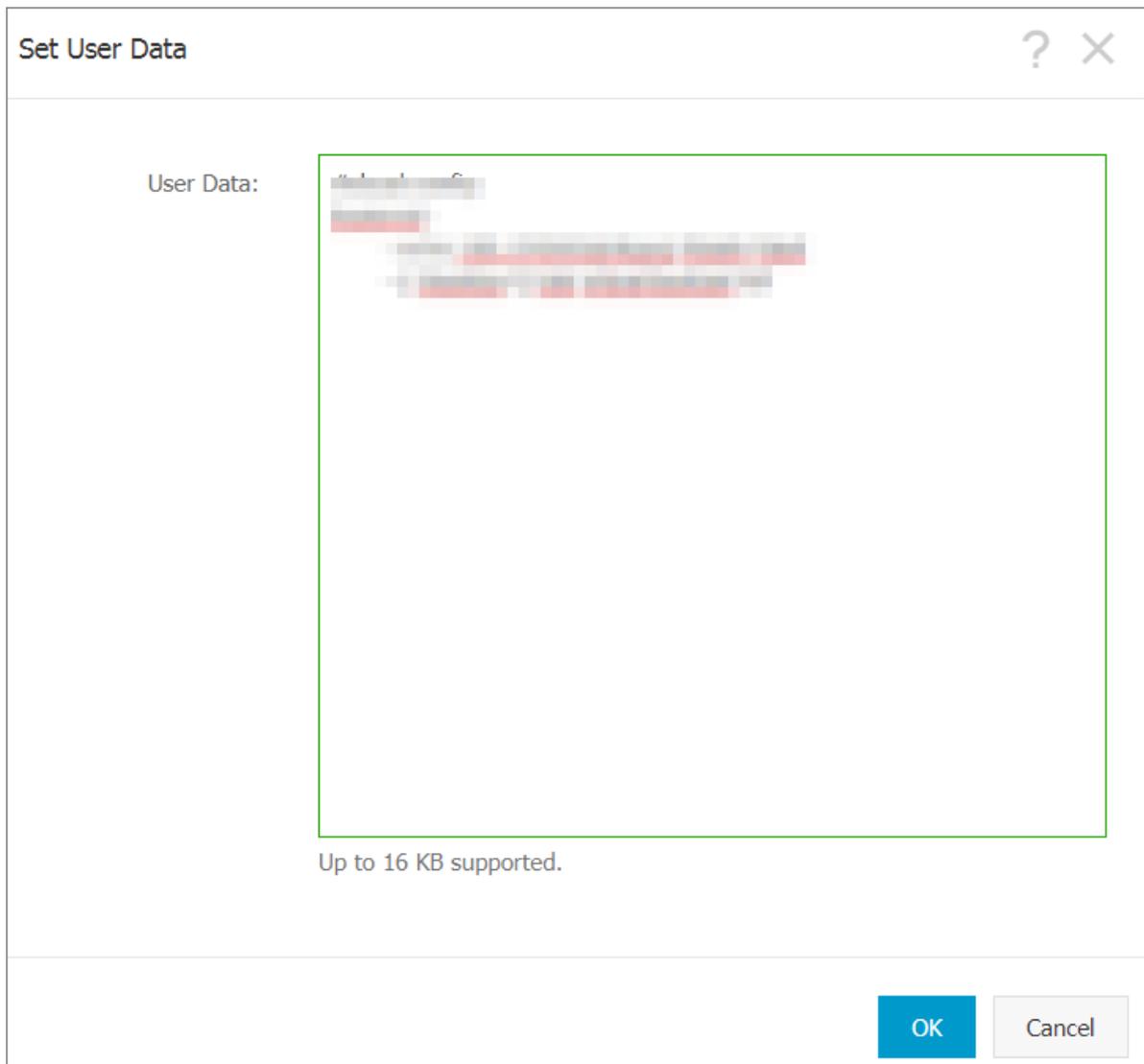
- For Linux, run `curl http://100.100.100.200/latest/user-data` to view the user data.
- For Windows, run `Invoke-RestMethod http://100.100.100.200/latest/user-data/` to view the user data.

Related API: [DescribeUserData](#)

Modify user data

You must stop the instance before modifying its current user data. If you need to restart a Pay-As-You-Go VPC-Connected instance immediately after you modify the user data, we recommend that you disable the No fees for stopped instances option. To modify user data of an instance, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Instances**.
3. Select the target region.
4. Select the target instance and then, in the **Actions** column, click **Sets User Data**.
5. Enter the user data and then click **OK**.



 **Note:**
After you modify the user data, depending on the script type and the module type, the modified user data is or is not run.

Related API: [ModifyInstanceAttribute](#) + Parameters `Userdata`

Linux instance user data

Linux instance user data can be configured by several types of script, such as [User-data Script](#), [Cloud Config](#), [Include Files](#), [Gzip compression scripts](#), and [Upstart Job](#). The scripts follow the format of open source cloud-init, and reference the [Metadata](#) for data sources. The configuration of Linux instances are automated at boot. For more information, see [formats](#).

User-data script

User-data can be a shell script. It runs once at the instance first boot. The first line is fixed as `#!`, for example `#! /bin/sh`. The content of user-data script before Base64 encoding cannot exceed 16 KB. The following is a User-Data script example:

```
#! /bin/sh
echo "Hello World. The time is now $(date -R)!" | tee /root/output10.txt
service httpd start
chkconfig httpd on
```

After the instance has been created, connect to the instance and run `cat [file]` to view the results of the user-data script.

```
[root@XXXXXX2z ~]# cat output.txt
Hello World. The time is now Mon, 24 Jul 2017 13:03:19 +0800!
```

Cloud-Config

You can use Cloud-Config to configure services such as updating yum sources, importing SSH keys, installing dependency packages, and more. The first line of Cloud-Config is fixed as `#cloud-config`, and the header cannot have spaces. The file must be valid yaml syntax. Depending on the service you configured, the instance user data runs differently.

Cloud Instance user data requires Base64 encoding before being passed in, and the pre-encoding cloud config data cannot exceed 16 KB. The following is a Cloud-Config script example:

```
#cloud-config
apt:
  primary:
    - arches: [default]
    uri: http://us.archive.ubuntu.com/ubuntu/
  bootcmd:
    - echo 192.168.1.130 us.archive.ubuntu.com >> /etc/hosts
```

After the instance has been created, connect to the instance to view the results.

```
localhost localhost.localdomain localhost4 localhost4.localdomain4
:1 localhost localhost.localdomain localhost6 localhost6.localdomain6
192.168.1.130 us.archive.ubuntu.com
```

Include files

The contents of an Include File consist of a script link, with one link on one line. When the instance starts, cloud-init reads the contents of the script link in the Include File. If there is an error reading script content in a row, the instance stops performing user data. The first line of Include File is

fixed as `#include` and the header cannot have spaces. The update frequency of the instance user data follows the script type configured in the include file.

Instance user data requires Base64 encoding before being passed in. The file before Base64 encoding cannot exceed 16 KB. The following is an Include File example:

```
#include
http://ecs-image-test.oss-cn-hangzhou.aliyuncs.com/UserData/myscript.
sh
```

After the instance has been created, connect to the instance to view the results.

Gzip compressed content

The content of a [User-Data Script](#), [Cloud-Config](#), and [Include File](#) cannot exceed 16 KB. If your script content is larger than 16 KB, you can use Gzip to compress the content, then upload the compressed script to an available storage service (we recommend OSS), obtain the link, and use the Include File format to render the link. The first line of a Gzip compressed script is fixed as `#include` and the header cannot have spaces. The update frequency of the instance user data follows the script type configured in the Gzip file. The following is a Gzip compressed file example:

```
#include
http://ecs-image-test.oss-cn-hangzhou.aliyuncs.com/userdata/config.gz
```

Upstart Job

Upstart service is required for an init system if you use Upstart Job to configure user data. For example, CentOS 6, Ubuntu 10/12/14, and Debian 6/7 use upstart as the init system. Upstart Job script places your instance user data into a file in `/etc/init` directory. The first line of Upstart Job script is fixed as `#upstart-job` and the header cannot have spaces. We perform the instance user data for every instance boot. The following is a Upstart Job script example:

```
#upstart-job
description "upstart test"
start on runlevel [2345]
stop on runlevel [! 2345]
exec echo "Hello World. The time is now $(date -R)!" | tee /root/
output.txt
```

Windows instance user data

Windows instance user data is supported by Alibaba Cloud ECS, and offers Windows-based instances the ability to run initialization scripts. Instance user data requires Base64 encoding before being passed in, and the pre-encoding user data cannot exceed 16 KB. Only SBC case

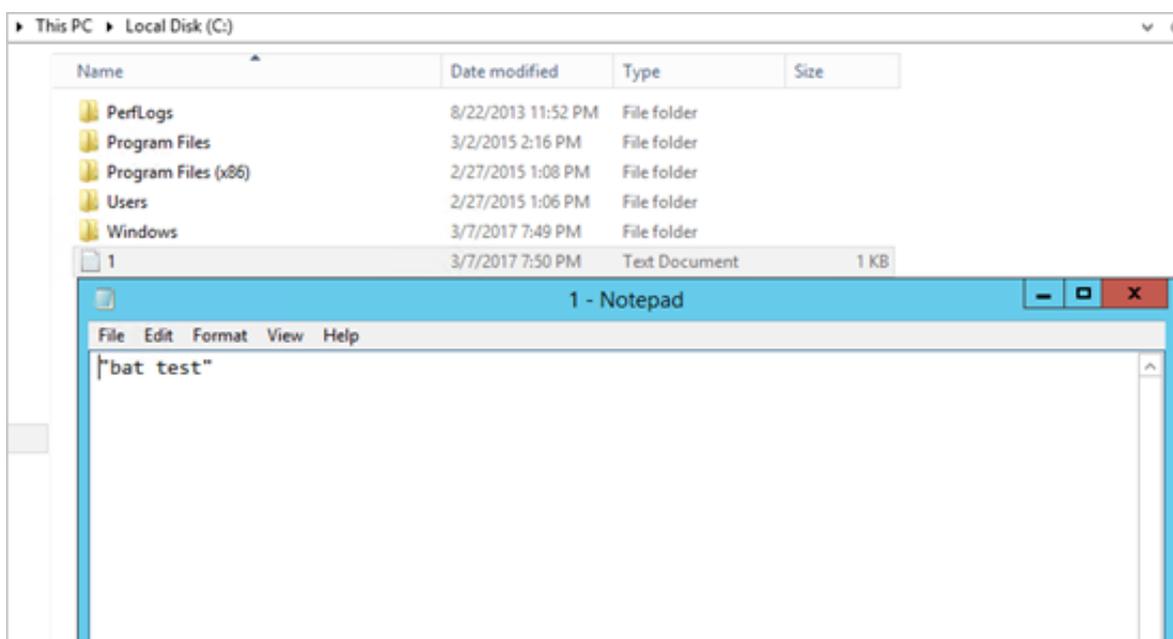
characters are allowed. You can write Bat script or PowerShell script to configure the instance user data.

Bat scripts

The first line is fixed as `[bat]` and the header cannot have spaces. For example:

```
[bat]
echo "bat test" > c:\1.txt
```

After the instance has been created, connect to the instance to view the results. In the following example, a `1.txt` text file is shown under the `C:\` drive.



The first line of PowerShell scripts

is fixed as `[powershell]` and the header cannot have spaces. For example:

```
[powershell]
write-output "Powershell Test" | Out-File C:\2.txt
```

Reference

For more information about Linux instance user data, see cloud-init [formats](#).

For more information about the update frequency of Linux instance user data, see cloud-init [modules](#).

4.12.3 Instance identity

Instance identity is a part of [Metadata](#) that describes and validates an instance. Instance identity enables you quickly locate a target instance,

and provides authentication for such actions as software updates, access control, or application activation. The signature of instance identity is encrypted by the [PKCS#7](#) standard.

Use cases

You can use **instance identity** (`instance-identity`) in scenarios such as authentication, granting access, or instance identification, as follows.

- A typical software activation (with one serial number for one device) does not work in the environment of cloud computing where the sales model of the [Alibaba Cloud Marketplace](#) is flexible. In this case, you can use instance identity to complete the software activation. For more information, see [Sample 1. No audience in the signature](#).
- When you write sensitive data in the instance, you can use instance identity to verify that the server is your instance.
- Scenarios whereby you want to confirm the source of the target server.

Feature details

Instance identity consists of a dynamically generated **instance identity document** (`document`) and **instance identity signature** (`signature`).

- **Instance identity document:** Describes the attributes of an instance. The following table lists instance identity document items.

Properties	Description	Can it be changed?
account-id	ID of the Alibaba Cloud account to which the instance belongs	No
create-time	Instance creation time	No
instance-id	Instance ID.	No
mac	MAC address of the instance primary network interface	No
region-id	ID of the region to which the instance belongs	No
serial-number	Serial number of the instance	No

Properties	Description	Can it be changed?
zone-id	ID of the zone to which the instance belongs	No
instance-type	Instance types	Yes. It changes after you change the instance type.
image-id	Image ID of the instance	Yes. It changes after you replace the system disk of the instance.
private-ip	Private IP of the instance	Yes. It changes after you change the private IP of a VPC-Connected instance.

- **Instance identity signature:** Verifies the instance identity in the cryptographic method of the PKCS#7 standard.
 - To enhance the security of the signature, you can protect it by specifying the **audience** parameter in it. However, even if you specify **audience**, another user may get information about the identity document and the identity signature. Therefore, we recommend the value of the **audience** parameter is a random string, timestamp, regularly changed data, or some output generated by a specific algorithm.
 - If you specify the **audience** parameter, you must modify the instance identity document and signature simultaneously. For example, if you have specified the **audience** parameter while obtaining the signature, before you verify the signature by using the OpenSSL commands, you must add the value of the **audience** parameter at the end of the dynamically obtained instance identity document in the format of `"audience": "Value of the audience"`, and separate the parameters with a comma (,).

Usage

The instance identity is verified by using the OpenSSL commands. Make sure that you have the OpenSSL configured in your instance. Visit <https://www.openssl.org/source> to download and update OpenSSL service.

Take CentOS 7.4 as an example to use the instance identity.

1. Connect to your Linux instance.
2. Run `curl http://100.100.100.200/latest/dynamic/instance-identity/document` to query the file of instance identity document.

- Run `curl http://100.100.100.200/latest/dynamic/instance-identity/pkcs7` or `curl http://100.100.100.200/latest/dynamic/instance-identity/pkcs7?audience=XXXX` to get the instance identity signature.

```
[root@LocalHost ~]# curl http://100.100.100.200/latest/dynamic/instance-identity/document
{"region-id":"cn-hangzhou","instance-id":"i-bp1b39...","serial-number":"52aelaf6-64aa-407f-88fd-...","private-ip4":"172.17.0.174","mac":"00:16:3f:00:00:00","image-id":"centos_7_04_64_20G_alibase_201701015.vhd","zone-id":"cn-hangzhou-g","owner-account-id":"...","instance-type":"ecs.g5.large"}[root@LocalHost ~]# curl http://100.100.100.200/latest/dynamic/instance-identity/pkcs7?audience=test
MIIDMwYJKoZIhvcNAQcCoIIDJCCAYACAQExCzAJBgUrDgMCGGUAMIIBbQYJKoZIhvcNAQcBoIIBXGSCBSsOAwIaBQAwDQYJKoZIhvcNAQEBBQAEggEAUOghxG3i3hKgjPggPX6NInYNPOZJusp//fy15Pr+gZoqLgvxonLzOXxBG1yy1aEzjb2b2zUFZGfjuNDbk1kVcsGq3kKEbpBCXF2GNm9FaS54r0szTcw9HkNPsklaCqR9Z6LvBB/sPMTz8i3dY8pu/zhiZoulHdpvKCYYP8Q89sN+QVlsS2eQDXnqNSFbi/QS/hZOz83gIuacH6+rWxojBf3Hs7bft4YdmNBhaTpxq8R3w16rWvtq3W58ULraHgmZq/Kn9e1SCnSAiqETLj7i60As7h/hd/DebOvoF0QANiXAIvdhLci4KK1rKJOHOW4BzOy44s/jc1flASYIsAQrug== [root@LocalHost ~]#
```

- Verify the instance identity by using OpenSSL.

```
openssl smime -verify -in $signature -inform PEM -content $DOCUMENT
-certfile AliyunPubkey -noverify > /dev/null
```



Note:

- Specify the variable `$signature` with the responded *instance identity signature*.
- Specify the variable `$DOCUMENT` with the responded *instance identity document*.

(Optional) In [step 3](#), if you have specified the `audience` parameter, add the value of the audience parameter at the end of the dynamically obtained instance identity document in the format of `"audience": "Value of the audience"`, and separate the parameters with a comma (,).

- Specify the variable `AliyunPubkey` with the *Alibaba Cloud public certificate*.

The public certificate of Alibaba Cloud in all regions is as follows.

```
-----BEGIN CERTIFICATE-----
MIIDdzCCA1+gAwIBAgIEZmbRhZANBgkqhkiG9w0BAQsFAADBsMRAwDgYDVQQGEwdV
bmtub3duMRAwDgYDVQQIEwdVbmtub3duMRAwDgYDVQQHEwdVbmtub3duMRAwDgYD
VQQKEwdVbmtub3duMRAwDgYDVQQLEwdVbmtub3duMRAwDgYDVQQDEwdVbmtub3du
MB4XDTE4MDIyMzAxMjEzOFoXDTE4MDIyMzAxMjEzOFowbDEQMA4GA1UEBHMHVW5r
bm93b3JlEQMA4GA1UECBMHVW5rbm93b3JlEQMA4GA1UEBxMHVW5rbm93b3JlEQMA4GA1UE
ChMHVW5rbm93b3JlEQMA4GA1UECXMHVW5rbm93b3JlEQMA4GA1UEAxMHVW5rbm93b3JlCC
ASIdQYJKoZIhvcNAQEBBQADggEPADCCAQoCggEBAlJwy5sbZDiNyX4mvdP32pqM
YMK4k7+5lRnVR2Fky/5uwyGSPbddNXaXzWem+u4wIsJiaAN3OZgJpYIoCGik+9lG
5gVAIr0+/3rZ6lIbeVE+vDenDd8g/m/YIdYBfC2IbZgS9EVGaf/gJdtDODXrDfQj
Fk2rQsvpftVOUs3Vp190+jeCQLoRbZYM0c5v7jP/L2lK0MjhiyWPF2kpDeismtNd
/ArkSPilglqVYm3F19v3pa6ZioM2hnxG5DibYlgVvsIBGhvYqdQ1KosNvcVGGQa
HCUuVGdS7vHjYp3byH0vQYYygzxUJT2TqvK7pD57eYMN5drc7e19oyRQvcbPQ3kkC
AwEAAAMhMB8wHQYDVR0OBBYEFawwrnH1RgFvPGo+UD5zS1xAkC91MA0GCSqGSIb3
DQEBCwUAA4IBAQBBLhDRgezD/OOppuYEVNB9+XiJ9dNmcuHUhJnTnjKQWVk/YDA
v+T2V3t9y18L8o61tRIVKQ++lDhj1Vmur/mbBN25/UNRpJllfpUH6oOaqvQAze4a
nRgyTnBwVBZkdJ0d1sivL9NZ4pKelJF3Y1w6rp0YMqV+cwkt/vRtZRJ31ZEeBhs7
```

```
vKh7F6BiGCHL5ZAwEUYe8O3akQwjgrMUcfuiFs4/sAeDMnmgN6Uq8DFEBXDpAxVN
sV/6Hockdfinx85RV2AUwJGfClcVcu4hMhOvKROpcH27xu9bBIeMuY0vvzP2VyOm
DoJeqU7qZjyCaUBkPimsz/1eRod6d4P5qxTj
-----END CERTIFICATE-----
```

Sample 1. No audience in the signature

Assume that you have published an image in the image market. The following example shows you how to grant access to the instances of your customers.

1. Connect to the target instance.
2. Verify whether the image used by the instance is from the [Alibaba Cloud Marketplace](#), or from another source, by calling the metadata items of `product-code` and `charge-type`. For more information, see [Metadata](#).

```
curl http://100.100.100.200/latest/meta-data/image/market-place/
product-code
curl http://100.100.100.200/latest/meta-data/image/market-place/
charge-type
```

3. Create a temporary file `cert.cer` in the working directory and save the [public certificate](#) to the file.
4. Determine the identity of the instance by running the following script.

```
#!/usr/bin/bash
function verify_signature_without_audience(){
curl 100.100.100.200/latest/dynamic/instance-identity/document >
document
echo "-----BEGIN CERTIFICATE-----" > signature
curl 100.100.100.200/latest/dynamic/instance-identity/pkcs7 >>
signature
echo "" >> signature
echo "-----END CERTIFICATE-----" >> signature
openssl smime -verify -in signature -inform PEM -content document -
certfile cert.cer -noverify > /dev/null
}
verify_signature_without_audience
```

5. Once the response result shows `Verification successful`, remove the restriction and run the image in the instance.

Sample 2. Audience in the signature

Assume that you published an image in the image market. The following example shows you how to grant access to the instances of your customers by specifying an `audience` parameter during the process of validation. To make sure that the instance identity is not maliciously acquired and used, you can implement the access control at the application server by combining your

audience parameter. We recommend the value of the `audience` parameter is a random string, timestamp, regularly changed data, or some output generated by a specific algorithm.

1. Connect to the target instance.
2. Verify whether the image used by the instance is from the [Alibaba Cloud Marketplace](#), or another source, by calling the metadata items of `product-code` and `charge-type`.

```
curl http://100.100.100.200/latest/meta-data/image/market-place/  
product-code  
curl http://100.100.100.200/latest/meta-data/image/market-place/  
charge-type
```

3. Create a temporary file `cert.cer` in the working directory and save the [public certificate](#) to the file.
4. Determine the identity of the instance by running the following script.

```
#!/usr/bin/bash  
function verify_signature_with_specified_audience(){  
  audience='your audience' #Here is your audience parameter.  
  document=$(curl 100.100.100.200/latest/dynamic/instance-identity/  
  document)  
  audience_json=', "audience": "'${audience}'"''  
  echo -n "${document%?}" ${audience_json} > document  
  echo "-----BEGIN CERTIFICATE-----" > signature  
  curl 100.100.100.200/latest/dynamic/instance-identity/pkcs7?  
  audience=${audience} >> signature  
  echo "" >> signature  
  echo "-----END CERTIFICATE-----" >> signature  
  openssl smime -verify -in signature -inform PEM -content document -  
  certfile cert.cer -noverify > /dev/null  
}  
verify_signature_with_specified_audience
```

5. Once the response result shows `Verification successful`, remove the restriction and run the image in the instance.

4.13 Instance RAM roles

4.13.1 What is the RAM role of an instance

Instance RAM (Resource Access Management) roles allow you to authorize role-based permissions to ECS instances.

You can assign a [role](#) to an ECS instance to allow applications hosted on that instance to access other cloud services by using a temporary STS (Security Token Service) credential. This helps guarantee the security of your `AccessKey` and allows you to apply fine-grained access control of your instances.

Background

Generally, applications within an ECS instance need to use the AccessKey of the **primary account** or *RAM user account*, which includes an AccessKeyId and AccessKeySecret, to access various cloud services on the Alibaba Cloud platform.

This means that, to make a call, you must apply the AccessKey directly in the instance, such as in the configuration file. However, if Alibaba Cloud writes the AccessKey into the instance for calling purposes, the AccessKey may be mistakenly exposed. To ensure the security of your account and resources, Alibaba Cloud provides instance RAM roles to support .

Benefits

Instance RAM roles enable you to:

- Associate a *role* to an ECS instance.
- Access other cloud services securely (such as OSS, SLB, and ApsaraDB for RDS) by using the STS credential from the applications within the ECS instance.
- Assign roles that have different policies for different ECS instances, and allow those instances have restrictive access level to other cloud services to obtain fine-grained access control.
- Maintain the access permission of ECS instances by modifying only the policy of the RAM role, meaning no changes to the AccessKey are required.

Pricing

Instance RAM roles are free to use.

Limits

Instance RAM roles have the following limits:

- Instance RAM roles are only applicable to VPC instances.
- An ECS instance can only be authorized to one instance RAM role.

How to use an instance RAM role

The instance RAM role can be used by any of the following methods:

- *Use the instance RAM role in the console.*
- *Use the instance RAM role by calling APIs.*

References

- For a list of cloud services that support STS, see *cloud services supporting RAM*.

- See [access other Cloud Product APIs by the Instance RAM Role](#) for instruction on how to access other cloud services.

4.13.2 Use the instance RAM role in the console

Limits

Instance RAM roles have the following limits:

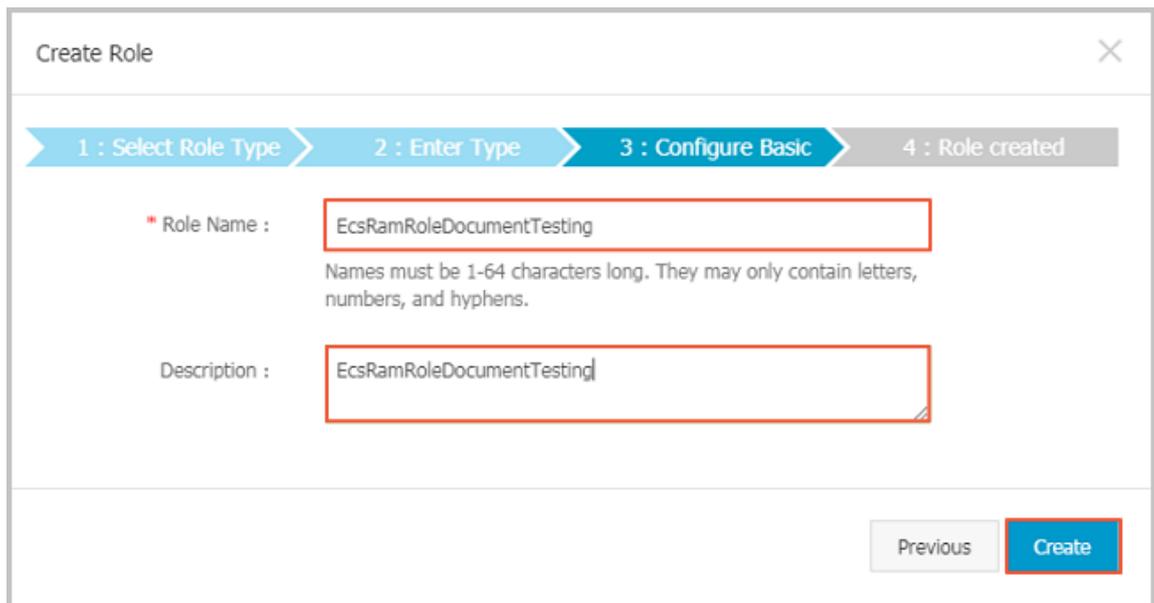
- Instance RAM roles are only applicable to VPC-Connected instances.
- An ECS instance can only be authorized to one instance RAM role at a time.
- After an instance RAM role is bound to an ECS instance, if you want to access other cloud services (such as OSS, SLB, or ApsaraDB for RDS) from applications within the ECS instance, you must obtain the authorization credential of the instance RAM role by using [Metadata](#). For more information, see [obtain authorization credentials](#).
- If you are using an instance RAM role through a RAM user account, you must use a primary account to [authorize a RAM user to use the instance RAM role](#).

Prerequisites

You must have activated the RAM service. See [activation method](#) to activate the RAM service.

1. Create an instance RAM role

1. Log on to the [RAM console](#).
2. In the left-side navigation pane, click **Roles**.
3. Click **Create Role**.
4. In the dialog box:
 - a. Select **Service Role** for **Role Type**.
 - b. Select **ECS Elastic Compute Service** for **Type**.
 - c. Enter a role name and description, for example, EcsRamRoleDocumentTesting.



Create Role

1 : Select Role Type > 2 : Enter Type > 3 : Configure Basic > 4 : Role created

* Role Name : EcsRamRoleDocumentTesting
Names must be 1-64 characters long. They may only contain letters, numbers, and hyphens.

Description : EcsRamRoleDocumentTesting

Previous Create

d. Click **Create**.

2. Authorize the instance RAM role

1. Log on to the [RAM console](#).
2. In the left-side navigation pane, click **Policies**.
3. Click **Create Authorization Policy**.
4. In the dialog box:
 - a. Select **Blank Template** for **authorization policy template**.
 - b. Enter a **Authorization Policy Name** and **Policy Content**. In this example, they are EcsRamRoleDocumentTestingPolicy.



Note:

For information about how to write the authorization policy in JSON format, see [policy syntax structure](#).

Create Authorization Policy

Step 1: Select an authorization policy | Step 2: Edit permissions and submit. | Policy creation complete.

Authorization Policy Name :

Names must be 1-128 characters long. They may only contain the letters A-Z, numbers 0-9, and hyphens.

Description :

Policy Content :

```
1 {
2   "Version": "1",
3   "Statement": [
4     {
5       "Effect": "Allow",
6       "Action": [
7         "oss:Get*",
8         "oss:List*"
9       ],
10      "Resource": ""
11    }
12  ]
13 }
```

[Authorization Policy Format](#)
[Authorization Policy FAQ](#)

Previous Cancel

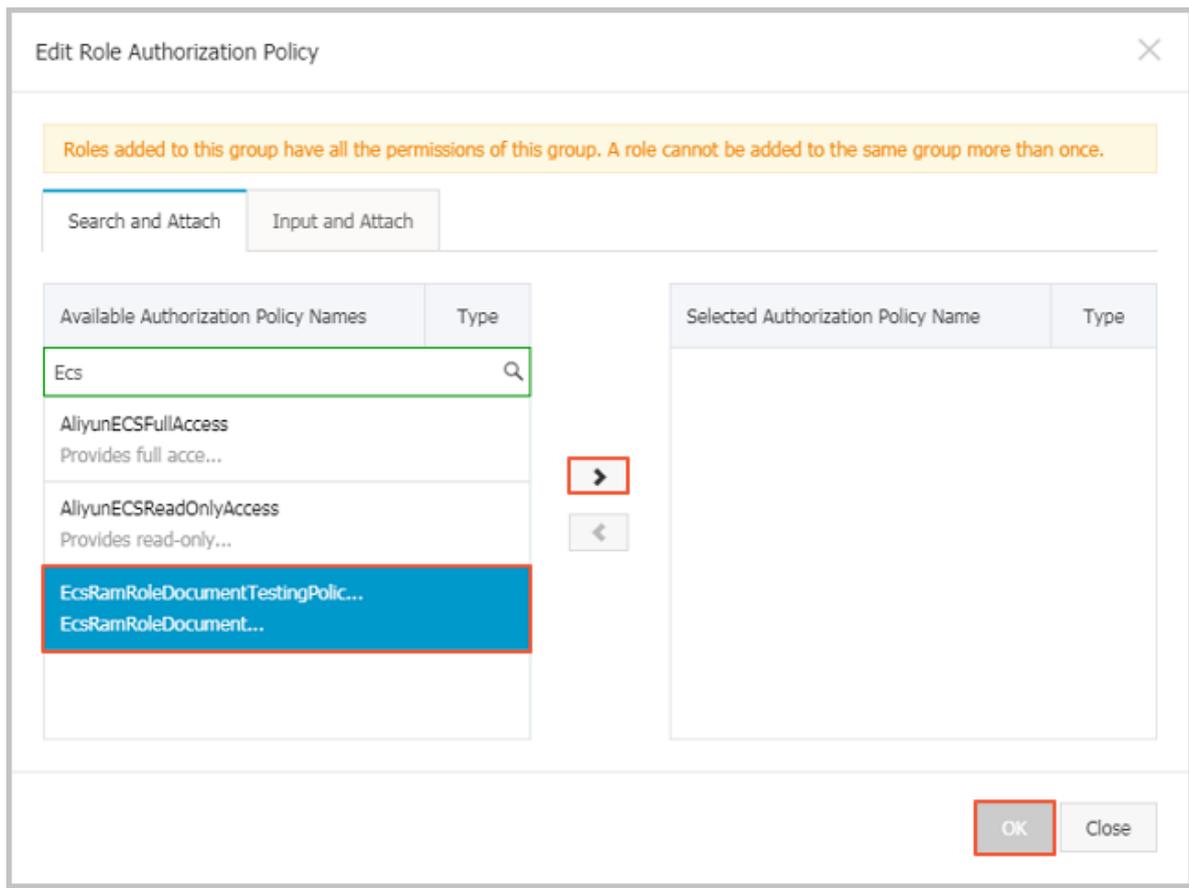
c. Click **Create Authorization Policy**.

5. In the left-side navigation pane, click **Roles**.

6. Select a role, for example, `EcsRamRoleDocumentTesting`, and click **Authorize**.

7. Enter the **Authorization Policy Name** and select it from the drop-down menu. In this example, `EcsRamRoleDocumentTestingPolicy` is selected.

8. Click the icon **>** to select the policy name, and then click **OK**.

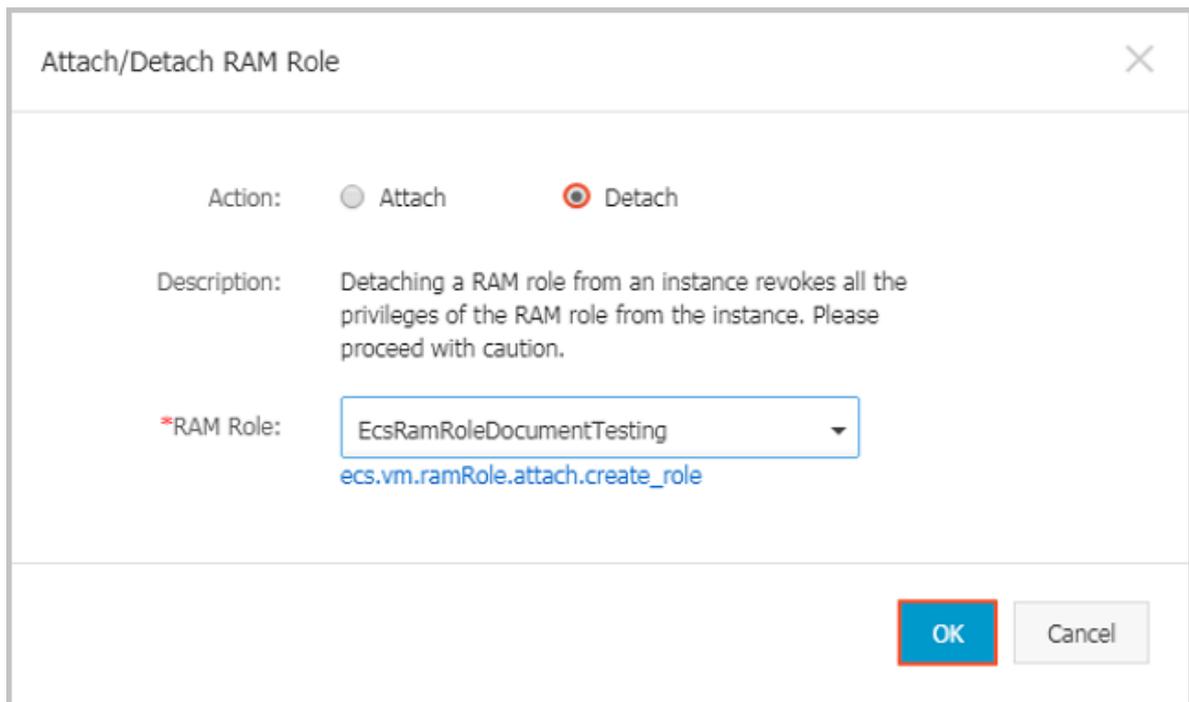


3. Bind an instance RAM role

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Instances**.
3. Select the target region.
4. Find the target ECS instance and select **More > Instance Settings > Bind/Unbind RAM Role**.
5. Select **Bind** for Action, select a role (for example, EcsRamRoleDocumentTesting), and then click **OK**.

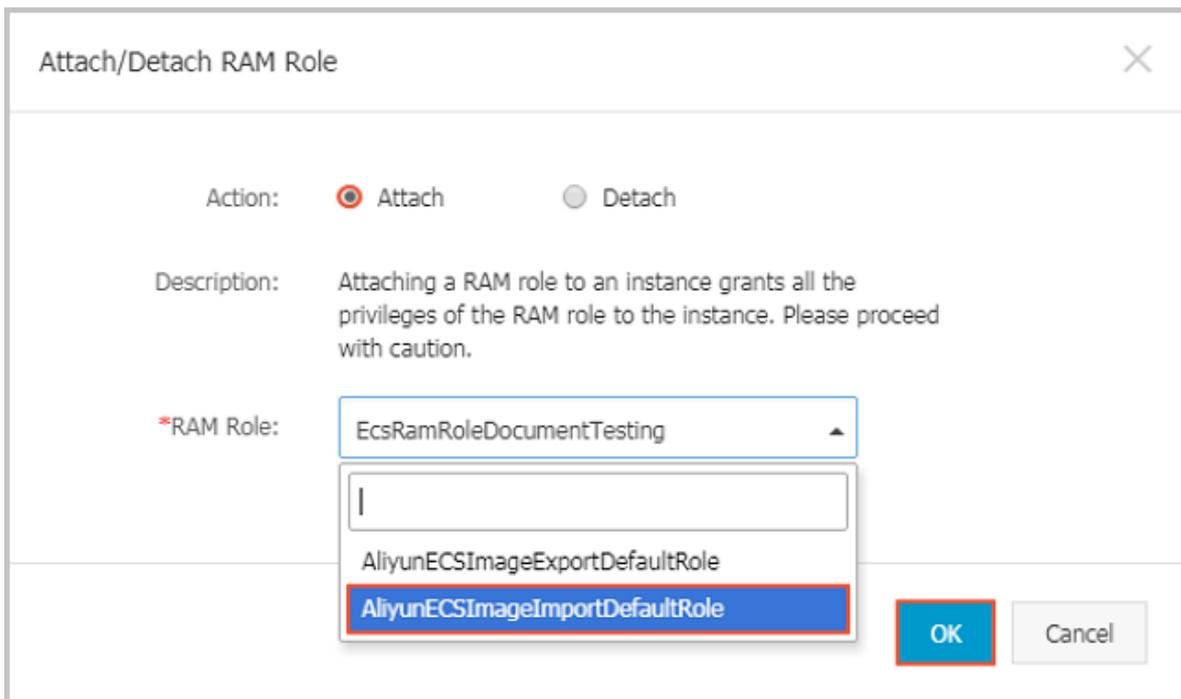
4. (Optional). Unbind an instance RAM role

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Instances**.
3. Select the target region.
4. Find the target ECS instance and select **More > Instance Settings > Bind/Unbind RAM Role**.
5. Select **Unbind** for **Action**, and click **OK**.



5. (Optional). Replace an instance RAM role

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Instances**.
3. Select the target region.
4. Find the target ECS instance and select **More > Instance Settings > Bind/Unbind RAM Role**.
5. Select **Bind** for **Action**, select another instance RAM role in the list of **RAM Role**, and then click **OK**.



6. (Optional). Obtain authorization credentials

To access an internal application of an ECS instance, you can obtain STS credentials of the instance RAM role (which is part of the metadata of an instance) to access the role-authorized permissions and resources. The credential is updated periodically. To access an instance by STS , follow these steps:

1. Connect to the target ECS instance.
2. Obtain the STS credential of the instance RAM role. In this example, it is EcsRamRoleDocumentTesting:

- For a Linux instance: run `curl http://100.100.100.200/latest/meta-data/Ram/security-credentials/EcsRamRoleDocumentTesting`.
- For a Windows instance: see [Metadata](#).

3. Get the credential. An example return is as follows:

```
{
  "AccessKeyId" : "XXXXXXXXXX",
  "AccessKeySecret" : "XXXXXXXXXX",
  "Expiration" : "2017-11-01T05:20:01Z",
  "SecurityToken" : "XXXXXXXXXX",
  "LastUpdated" : "2017-10-31T23:20:01Z",
  "Code" : "Success"
```

```
}

```

7. (Optional). Authorize a RAM user to use the instance RAM role



Note:

You must grant the RAM user with the **PassRole** permission to use the instance RAM role feature. Without the **PassRole** permission, a RAM user cannot carry out the permission of the authorization policy that is attached to the RAM user.

Log on to the RAM console and authorize the target RAM user by means of [authorizing RAM users](#) to complete the authorization. The following is an authorization policy example:

```
{
  "Version": "2016-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "ecs: [ECS RAM Action]",
        "ecs: CreateInstance",
        "ecs: AttachInstanceRamRole",
        "ecs: DetachInstanceRAMRole"
      ],
      "Resource": "*"
    },
    {
      "Effect": "Allow",
      "Action": "ram:PassRole",
      "Resource": "*"
    }
  ]
}
```

The parameter `[ECS RAM Action]` indicates that a RAM user can be authorized for certain actions. For more information, see [authorization rules](#).

References

- Click the following link to learn how to [use the instance RAM role by calling APIs](#).
- Click the following link to see how to [access other cloud products by using the instance RAM role](#).

4.13.3 Use the instance RAM role by calling APIs

Limits

Instance RAM roles have the following limits:

- Instance RAM roles are only applicable to VPC-Connected instances.

- An ECS instance can only be authorized to one RAM role at a time.
- After an instance RAM role is attached to an ECS instance, if you want to access other cloud services (such as OSS, SLB, or ApsaraDB for RDS) from applications within the ECS instance, you must obtain the authorization credential of the instance RAM role by using [Metadata](#). For more information, see [obtain the on-demand authorization credential](#).
- If you are using an instance RAM role through a RAM user account, you must use a primary account to [authorize a RAM user to use the instance RAM role](#).

Prerequisite

If you are using a RAM user account, it must be authorized to use the instance RAM role. See [activation method](#) to activate the RAM service.

1. Create an instance RAM role

1. Call the `CreateRole` [CreateRole](#) to create an instance RAM role.
2. Set a parameter `RoleName`, for example, `EcsRamRoleDocumentTesting`.
3. Set the `AssumeRolePolicyDocument` as follows:

```
"Statement": [  
  {  
    "Action": "sts:AssumeRole",  
    "Effect": "Allow",  
    "Principal": {  
      "Service": [  
        "ecs.aliyuncs.com"  
      ]  
    }  
  }  
]  
  
"Version": "1"
```

2. Authorize the instance RAM role

1. Call the `CreatePolicy` [CreatePolicy](#) to create an authorization policy.
2. Set a parameter `RoleName`, for example, set it to `EcsRamRoleDocumentTestingPolicy`.
3. Set the `PolicyDocument` as follows.

```
"Statement": [  
  {  
    "Action": [  
      "oss:Get*",  
      "oss:List*"  
    ],  
    "Effect": "Allow",  
    "Resource": "*"  
  }  
]
```

```
"Version": "1"
```

4. Call the [AttachPolicyToRole](#) to authorize the role policy.
5. Set `PolicyType` to Custom.
6. Set a parameter `PolicyName`, for example, `EcsRamRoleDocumentTestingPolicy`.
7. Set a parameter `RoleName`, for example, `EcsRamRoleDocumentTesting`.

Attach the instance RAM role

1. Call the [AttachInstanceRamRole](#) to attach an instance RAM role to an ECS instance.
2. Set the parameters `RegionId` and `InstanceIds` to specify an ECS instance.
3. Set a parameter `RamRoleName`, for example, `EcsRamRoleDocumentTesting`.

4. (Optional). Detach an instance RAM role

1. Call the [DetachInstanceRamRole](#) to detach an instance RAM role.
2. Set the parameters `RegionId` and `InstanceIds` to specify an ECS instance.
3. Set a parameter `RamRoleName`, for example, `EcsRamRoleDocumentTesting`.

5. (Optional). Obtain the on-demand authorization credential

For the internal application of an ECS instance, you can obtain the STS credential of the instance RAM role, which is a metadata of an instance, to access the role-authorized permissions and resources. The credential is updated periodically. Example:

1. Obtain the STS credential of the instance RAM role, for example,

`EcsRamRoleDocumentTesting`:

- Linux instance: run `curl http://100.100.100.200/latest/meta-data/Ram/security-credentials/EcsRamRoleDocumentTesting`.
- Windows instance: see [Metadata](#).

2. Get the credential Token. Return example:

```
"AccessKeyId" : "XXXXXXXXXX",
"AccessKeySecret" : "XXXXXXXXXX",
"Expiration" : "2017-11-01T05:20:01Z",
"SecurityToken" : "XXXXXXXXXX",
"LastUpdated" : "2017-10-31T23:20:01Z",
"Code" : "Success"
```

6. (Optional). Authorize a RAM user to use the instance RAM role



Note:

You must grant the RAM user with the `PassRole` permission to use the instance RAM role feature.

Log on to the RAM console and follow the steps to [authorize RAM users](#). Then, authorize the RAM user to complete the authorization, see the following code snippet as an authorization policy example:

```
"Version": "2016-10-17",
"Statement": [

  "Effect": "Allow",
  "Action": [
    "ecs: [ECS RAM Action]",
    "ecs: CreateInstance",
    "ecs: AttachInstanceRamRole",
    "ecs: DetachInstanceRAMRole"

  "Resource": "*"

  "Effect": "Allow",
  "Action": "ram:PassRole",
  "Resource": "*"

```

The parameter `[ECS RAM Action]` indicates that a RAM user is authorized for certain actions. See [authorization rules](#).

References

- Click the following link to see how to [use the instance RAM role in the console](#).
- For instruction on how to access other cloud services, see [access other Cloud Product APIs by the Instance RAM Role](#).
- APIs related to the instance RAM role include:
 - [CreateRole](#): Create an instance RAM role
 - [ListRoles](#): Query the list of instance RAM roles
 - [CreatePolicy](#): Create an instance RAM role policy
 - [AttachPolicyToRole](#): Authorize an instance RAM role policy
 - [AttachInstanceRamRole](#): Attach an instance RAM role
 - [DetachInstanceRamRole](#): Detach an instance RAM role
 - [DescribeInstanceRamRole](#): Query an instance RAM role

5 Connect to instances

5.1 Overview

Based on the network type and operating system of your ECS instance, and the operating system of your local machine, use one of the following methods to connect to an ECS instance.

Connect to a Linux instance

The following table details different methods by which to remotely connect to a Linux instance.

Is Internet access required?	Operating system of the local machine	Connection method
Yes/No	Windows or Unix-like OS	Connect to an instance by using the Management Terminal.
Yes	Windows	Use a remote connection tool to create remote connection: <ul style="list-style-type: none"> • Use an SSH key pair as the credential. For details, see connect to a Linux instance by using an SSH key pair. • Use a password as the credential. For details, see connect to a Linux instance by using a password.
Yes	Linux, Mac OS, or other Unix-like OS	Use commands to create remote connection: <ul style="list-style-type: none"> • Use an SSH key pair as the credential. For details, see connect to a Linux instance by using an SSH key pair. • Use a password as the credential. For details, see connect to a Linux instance by using a password.
Yes	iOS or Android	User apps, such as SSH Control Lite or JuiceSSH, to create remote connection.

Is Internet access required?	Operating system of the local machine	Connection method
		For details, see connect to an instance on a mobile device .

Connect to a Windows instance

The following table details different methods by which to remotely connect to a Windows instance.

Is Internet access required?	Operating system of the local machine	Connection method
Yes/No	Windows or Unix-like OS	Connect to an instance by using the Management Terminal .
Yes	Windows	Use mstsc to create remote connection. For details, see connect to a Windows instance .
Yes	Linux	Use a remote connection tool, such as rdesktop, to create remote connection. For details, see connect to a Windows instance .
Yes	Mac OS	Use Microsoft Remote Desktop Connection for Mac to create remote connection. For details, see connect to a Windows instance .
Yes	iOS or Android	Use Microsoft Remote Desktop to create a remote connection. For details, see connect to an instance on a mobile device .

5.2 Connect to an instance by using the Management Terminal

You can use the Management Terminal, also known as VNC, to connect to an ECS instance. This method is suitable for when other remote access software programs such as PuTTY, Xshell, or SecureCRT, do not work.

Scenarios

The Management Terminal can be used to:

- Check the status of an ECS instance.
- Reconfigure the firewall if a remote connection fails due to software error within the ECS instance.
- End abnormal processes that consume excessive CPU usage or bandwidth.



Note:

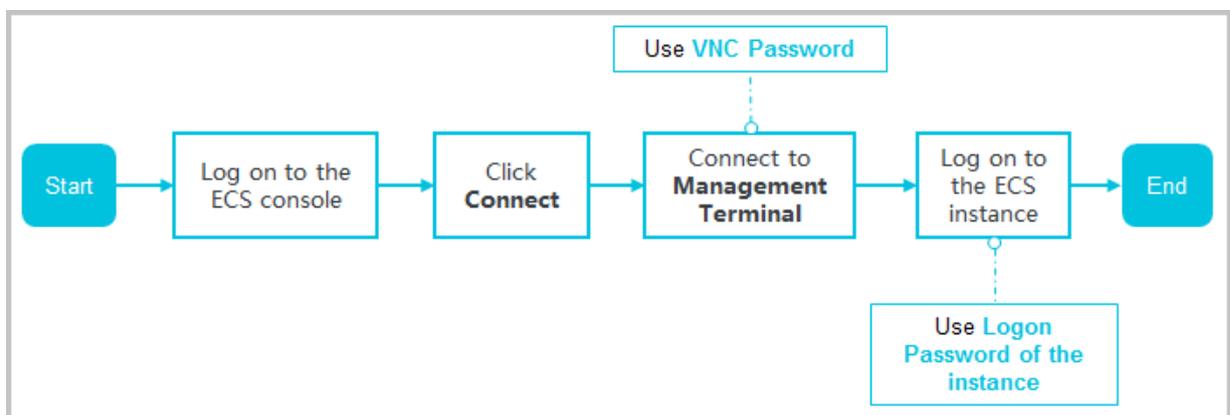
The Management Terminal can be used to connect to an instance even if no public IP address is assigned to your instance.

Prerequisites

- You have an ECS instance. For more information, see [create an ECS instance](#).
- You have set the logon password of the ECS instance. If not, you can use the [reset password](#) function.

Procedure

The following figure illustrates how to use the Management Terminal to connect to an ECS instance.



To connect to the ECS instance by using the Management Terminal, follow these steps:

1. Log on to the [ECS console](#).

2. In the left-side navigation pane, click **Instances**.
3. Select the target region.
4. In the instance list, find your instance and then, in the **Actions** column, click **Connect**.
5. In the **Management Terminal** page, follow the instructions to connect to the Management Terminal:

- If you log on as an Alibaba Cloud account to connect to the Management Terminal for the first time, follow these steps:

1. In the **VNC Connection Password** dialog box, copy the password and click **Close**.

**Note:**

- The VNC password appears only once. You must save the password immediately and store it securely for future use. If you need to change the VNC password, see [change the VNC connection password](#).
- If you log on as a RAM user to connect to the Management Terminal for the first time, you will not see this dialog box.

2. In the **Enter VNC Password** dialog box, paste the VNC connection password that you have copied, and click **OK**.

- If you log on as a RAM user to connect to the Management Terminal for the first time, or if you have forgotten your VNC connection password, follow these steps:

— [Change the VNC connection password](#).

— In the upper-left corner of the **Management Terminal** page, select **Send Remote Command > Connect to Management Terminal**.

— In the **Enter VNC Password** dialog box, enter the new password and click **OK**.

- If this is not your first connection to the Management Terminal, enter the VNC connection password in the **Enter VNC Password** dialog box and click **OK**.

6. To log on to the ECS instance, follow these steps according to the operating system:

- For a Linux instance: Enter the user name (`root`) and the logon password.

**Note:**

- If you do not know the logon password of your instance, [reset the password](#).
- The logon password input is invisible.

- If you want to perform additional operations within the instance, in the upper-left corner of the **Management Terminal** page, select **Send Remote Command > CTRL + ALT + Fx**, of which **Fx** can be any key from **F1** to **F10**, to switch the interfaces for different operations.
- If see a black screen, the Linux instance may be in sleep mode. To exit sleep mode, click your mouse or press any key.
- For a Windows instance: In the upper-left corner of the **Management Terminal** page, select **Send Remote Command > CTRL+ALT+DELETE**. The Windows logon interface is displayed. Enter the user name (**Administrator**) and the logon password.

**Note:**

If you do not know the logon password of your instance, [reset the password](#).

Other Operations

Change the VNC connection password

If you do not know your VNC connection password, follow these steps to change the password.

**Note:**

If the instance that you are connecting to is not I/O optimized, you must restart your instance in the ECS console to apply the new VNC password. The restart operation stops your instance and interrupts your business operations. Therefore, proceed with caution.

1. Open the **Management Terminal** page.
2. Close any dialog box that displays.
3. In the upper-right corner of the **Management Terminal** page, click **Modify Management Terminal Password**.
4. Enter a new password. The password must be at least six characters in length and can contain letters and numbers only.
5. Depending on the instance, the new password takes effect as follows:
 - For an I/O-optimized instance, the new password takes effect immediately.
 - For a non-I/O-optimized instance, [restart the instance](#) in the ECS console.

**Note:**

Restarting the operating system does not apply the new password.

Input commands

If you are connecting to a Linux instance, use the **Input Commands** feature to type long text, such as a complex command or a URL.

To use Input Commands, follow these steps:

1. Open the **Management Terminal** page.
2. In the upper-right corner of the **Management Terminal** page, click **Input Commands**.
3. Enter the commands and click **OK**.
4. Press the **Enter** key to run the commands.

FAQ

- Can multiple users simultaneously connect to the Management Terminal?

No. Only one user can connect to the Management Terminal at a time.

- Why am I unable to connect to an instance by using the Management Terminal even after changing the password?

Make sure that you enter the correct VNC password. If the instance that you are connecting to is not I/O optimized, you must restart the instance in the ECS console. This action helps the new VNC password to take effect.

- Why do I see a black screen after logging on to my instance?

A black screen indicates that the instance is in sleep mode.

For a Linux instance, click your mouse or press any key to activate the screen.

For a Windows instance, click **Send remote command** > **CTRL+ALT+DELETE** to view the logon interface.

- Why am I unable to access the Management Terminal?

Open your browser, connect to the Management Terminal, and press **F12** to open the developer tool. You can then go to the Console tab to analyze the Management Terminal information and locate errors under.

- Can I use IE or Firefox to access the Management Terminal?

IE version 10 and later support the Management Terminal. Only certain versions of Firefox are supported.



Note:

We recommend that you use Google Chrome as it offers the best support for Management Terminal.

5.3 Connect to a Linux instance by using an SSH key pair

This document describes how to use an SSH key pair to log on to a Linux instance in the following OSs:

- [Local Windows OS](#)
- [Local Linux OS or other OSs supporting SSH commands](#)



Note:

You can also use your account and password to connect to a Linux instance. For detailed operations, see [connect to a Linux instance by using a password](#) and [connect to an instance by using the Management Terminal](#).

Local Windows OS

The following uses PuTTY and PuTTYgen as an example to describe how to use a key pair generated by Alibaba Cloud to log on to a Linux instance through the SSH remote access tool on a Windows OS.

Prerequisites

- You have downloaded and installed PuTTY and PuTTYgen. The download links are as follows:
 - PuTTY: <https://the.earth.li/~sgtatham/putty/latest/w64/putty.exe>
 - PuTTYgen: <https://the.earth.li/~sgtatham/putty/latest/w64/puttygen.exe>
- You have a Linux instance allocated with a key pair. You can allocate a key pair when creating an instance or [bind a key pair](#) for the instance.
- The following security group rules must be added to the security group where the instance resides. For detailed operations, see [add security group rules](#).

Network type	Network card type	Rule direction	Authorization policy	Protocol type	Port range	Authorization type	Authorization object	Priority
VPC	Not required	Inbound	Allow	SSH (22)	22/22	IP address segment-based access	0.0.0.0/0	1
Classic network	Internet							

Procedure

1. (Optional) If you are using a .pem private key file generated by Alibaba Cloud, you must do the following to convert it to a .ppk key file: If you are using a .ppk private key file, skip this step.

**Note:**

You can download the .pem private key file when you [create an SSH key pair](#).

- a. Start PuTTYgen. PuTTYgen 0.68 is used in this example.
- b. In the **Parameters** area, select `RSA` for **Type of key to generate**.

**Note:**

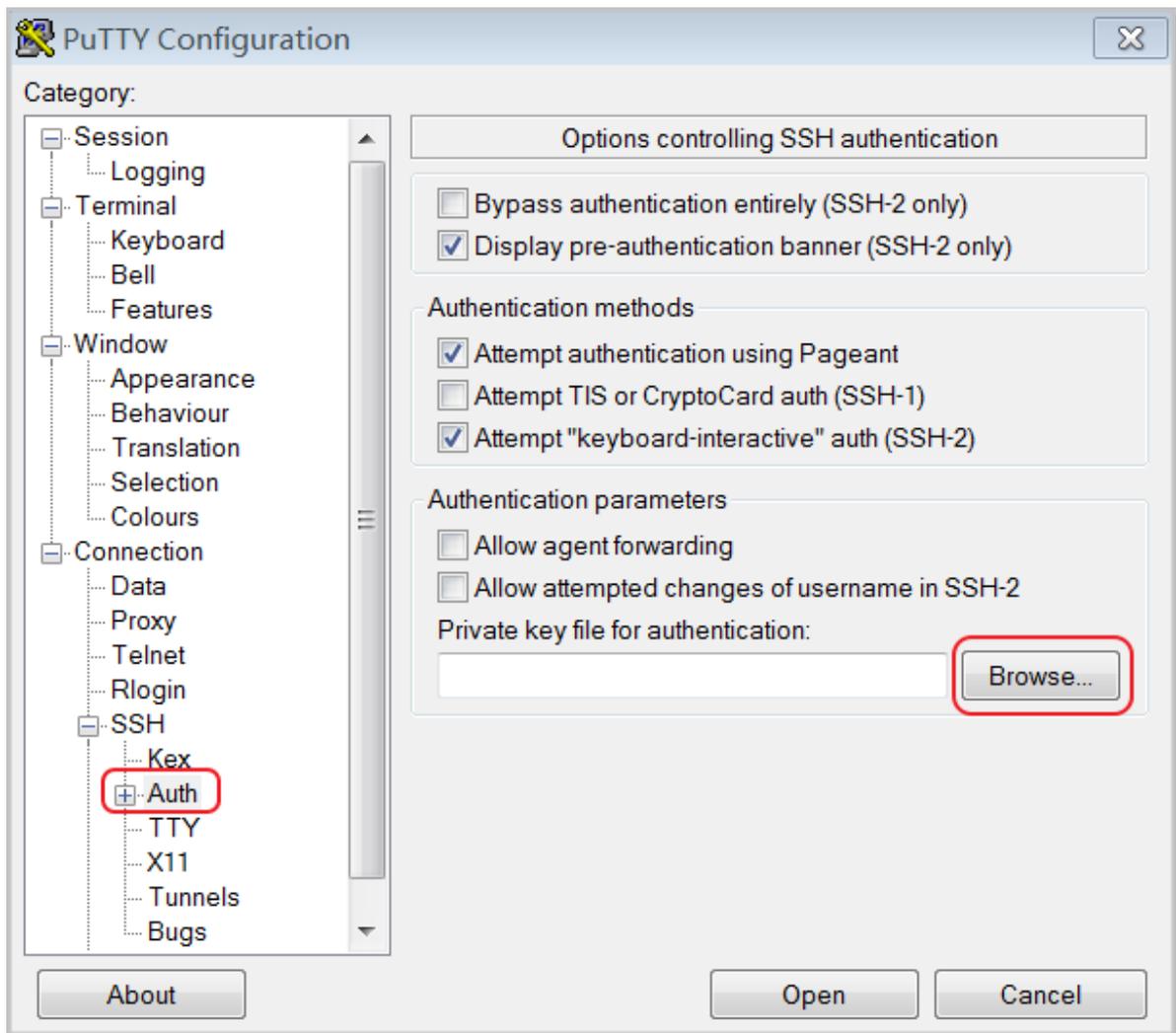
You do not need to set **Number of bits in a generated key**. PuTTYgen will automatically updates the parameter value according to the information about the imported private key.

- c. Click **Load**, select All Files (*.*) from the drop-down list after the file name field, and then locate your .pem file.

**Note:**

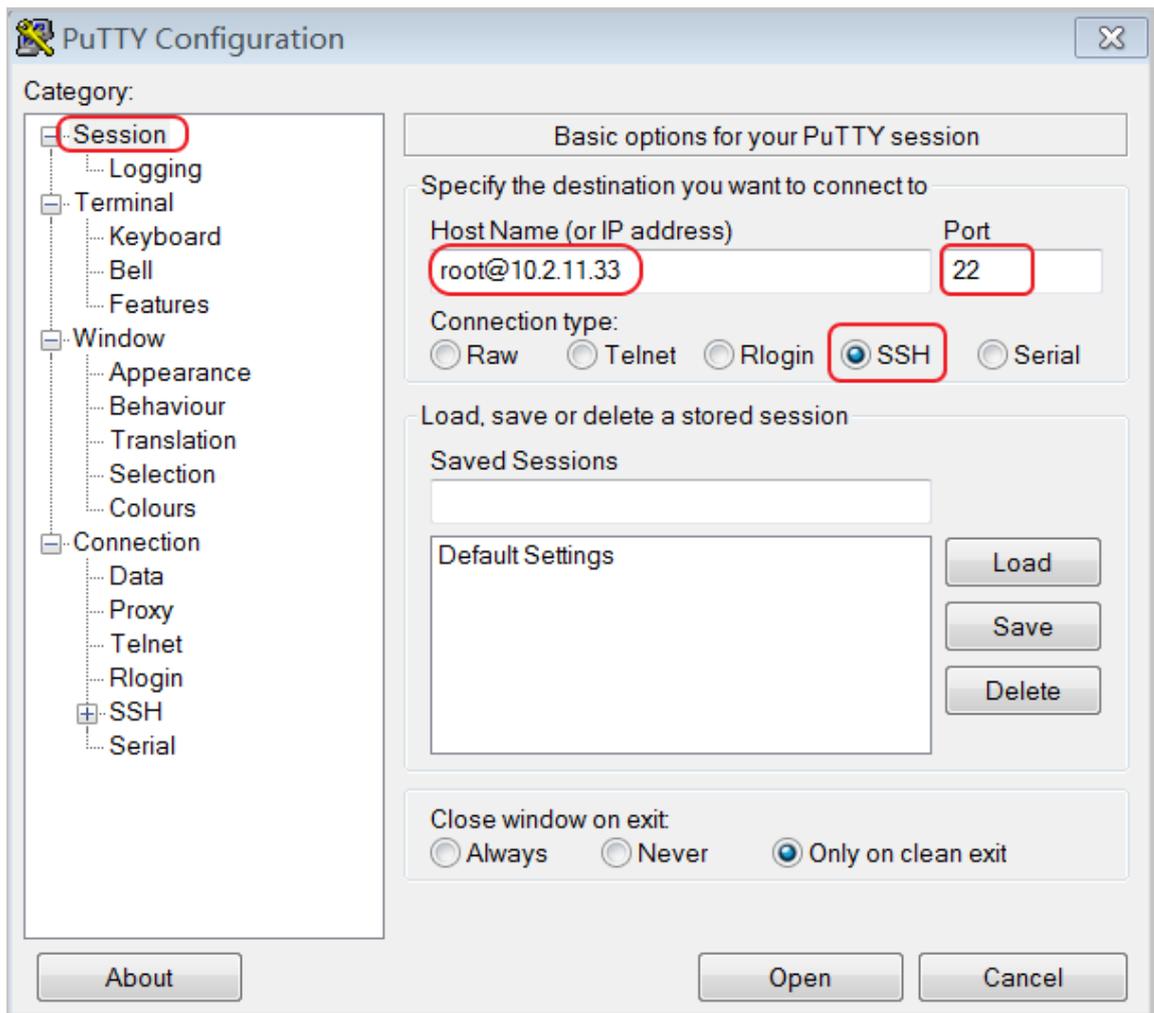
By default, only .ppk files are displayed.

- d. Select the .pem private key file you have downloaded from Alibaba Cloud and click **Open**.
 - e. Click **OK** to close the confirmation dialog box.
 - f. Click **Save private key**. In the PuTTYgen Warning dialog box indicating saving the key without a passphrase to protect it, click **Yes**.
 - g. Set the private key name to the key pair name and save the name. PuTTY will automatically add the .ppk extension to the file.
2. Start PuTTY.
 3. Choose **Connection > SSH > Auth** from the left navigation pane, click **Browse...** in the right pane, and then select the generated .ppk file.



4. In the left navigation pane, click **Session**.

- In the right pane, enter your account and the Internet IP address of the instance to be connected in the **Host Name (or IP address)** text box in `root@ IP address` format.
- In the **Port** text box, enter the port number 22.
- Select SSH for **Connection type**.



5. Click **Open** to start connecting to your Linux instance.

When `Connection established.` is displayed, you have successfully logged on to the instance by using the key pair.

Local Linux OS or other OSs supporting SSH commands

This section describes how to use an SSH key pair to log on to a Linux instance on a Linux OS or an OS supporting SSH commands, for example, Windows MobaXterm.

Prerequisites

You have a Linux instance with a key pair allocated. You can [allocate a key pair when creating an instance](#) or [bind a key pair](#) for the instance.

The following security group rules must be added to the security group where the instance resides. For detailed operations, see [add security group rules](#).

Network type	Network card type	Rule direction	Authorization policy	Protocol type	Port range	Authorization type	Authorization object	Priority
VPC	Not required	Inbound	Allow	SSH (22)	22/22	IP address segment-based access	0.0.0.0/0	1
Classic network	Internet							

Procedure

• Method 1

1. Locate the directory for saving the .pem private key file on your local PC, for example, `/root/xxx.pem`.



Note:

You can download the .pem private key file when you [create an SSH key pair](#). `xxx.pem` is your private key file.

2. Run `chmod 400 [Directory for saving the .pem private key file on your local PC]` to modify the attributes of the private key file, for example, `chmod 400 /root/xxx.pem`.
3. Run `ssh -i [Directory for saving the .pem private key file on your local PC] root@[Internet IP address]` to connect to the instance, for example, `ssh -i /root/xxx.pem root@10.10.10.100`.

• Method 2

You can simplify the connection commands through SSH configurations.

1. Enter the `ssh` directory in the root directory and do the following to modify the `config` file:

```
Host ecs // Set the name of your ECS instance.
HostName 192. *. *. * // Enter the Internet IP address of your ECS instance.
Port 22 / Enter the port number, which is 22 by default.
User Root // Enter your logon account.
IdentityFile ~/.ssh/ecs.pem // Enter the directory for saving the .pem private key file on your local PC.
```

2. Save the `config` file.
3. Restart SSH.
4. Run `ssh [ECS name]` to connect to your ECS instance, for example, `ssh ecs`.

5.4 Connect to a Linux instance by using a password

You can connect to a Linux instance by using different authentication methods:

- If you are using an SSH key pair, see [connect to a Linux instance by using an SSH key pair](#).
- If you are using a password, you can [connect to an instance by using the Management Terminal](#) or by using software applications or command lines.

Prerequisites

- The instance must be in the **Running** status. If not, [start it](#).
- You have set a logon password for the instance. If the password is lost, you can [reset the password](#).
- The instance can access the Internet:
 - In a VPC, a public IP address is assigned to the instance or [an EIP address is bound to the instance](#).
 - In the classic network, a public IP address is assigned to the instance by using either of the following methods:
 - For a Subscription or a Pay-As-You-Go instance, you can select **Assign public IP** when creating the instance.
 - For a Subscription instance without a public IP address, you can assign one by [upgrading the bandwidth](#).
- The following security group rules must be added to the security group that the instance joins. For more information, see [add security group rules](#).

Network type	NIC	Rule direction	Authorization policy	Protocol type	Port range	Authorization type	Authorization object	Priority
VPC	N/A	Inbound	Allow	SSH (22)	22/22	Address Field Access	0.0.0.0/0	1
Classic	Internet							

Procedure

Based on the operating system of your local machine, use one of the following methods to connect to a Linux instance by using the SSH protocol:

- [Windows OS](#)
- [Linux or Mac OS X](#)

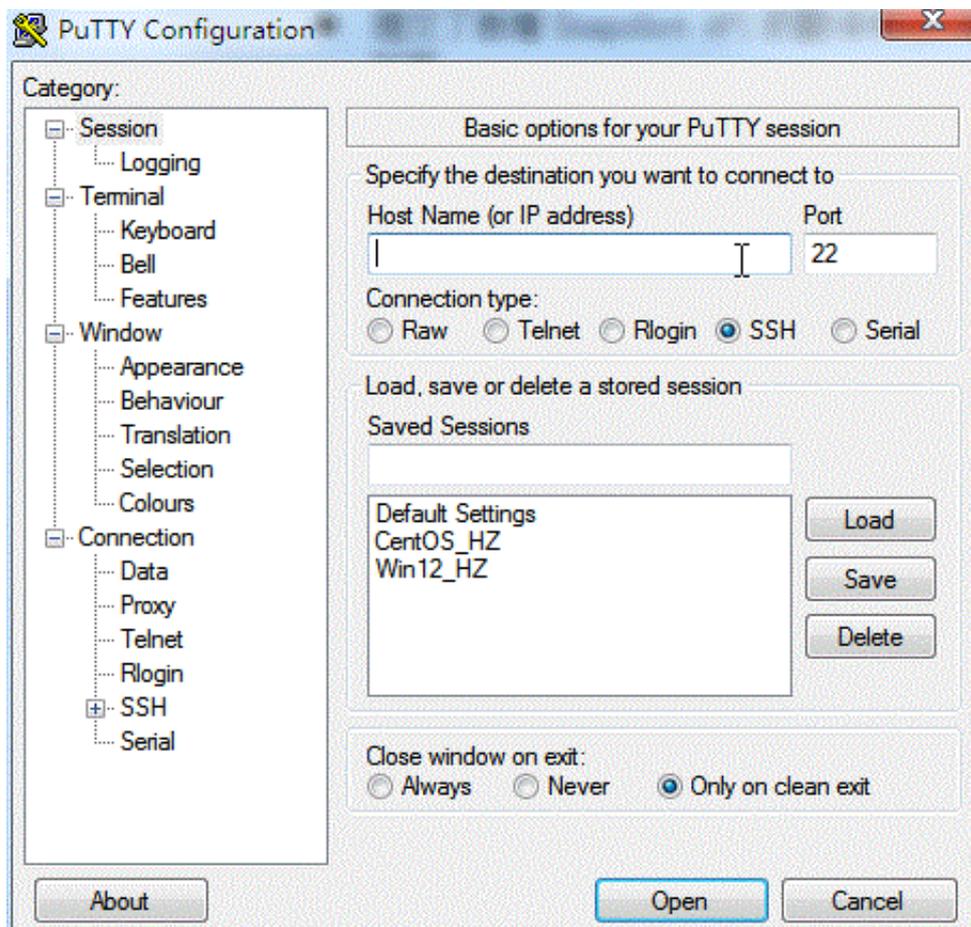
- [Android or iOS](#)

Windows OS

If your local machine is running Windows OS, you can use a remote connection tool, such as PuTTY, to connect to a Linux instance. In this article, we use PuTTY as an example to describe how to connect to a Linux instance by using the password authentication method. Before you start, download [PuTTY](#).

Follow these steps to connect to a Linux instance:

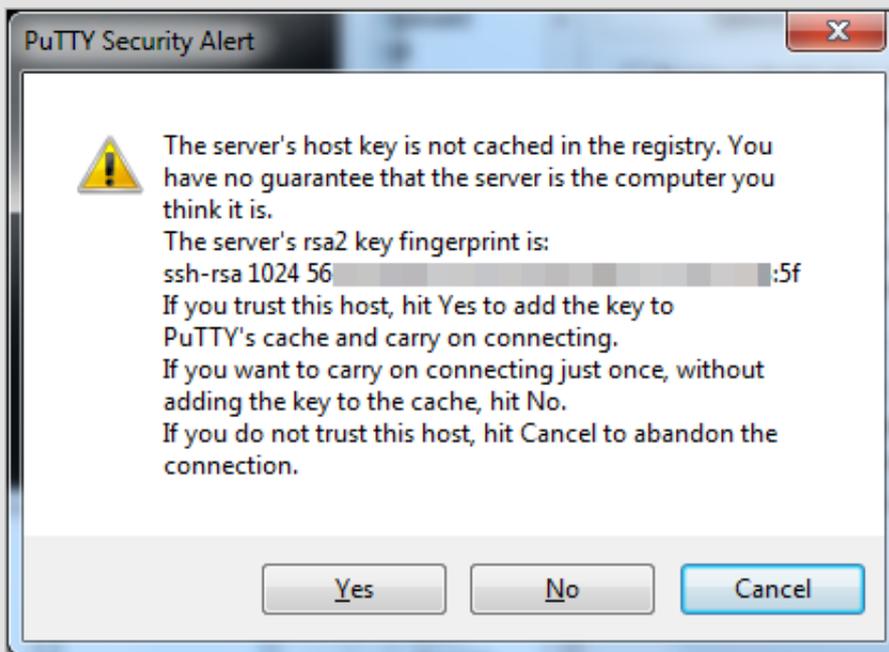
1. Start putty.exe.
2. In the left-side navigation pane, click **session**, and configure the following parameters:
 - **Host Name:** Type the public IP address or EIP address of the instance.
 - **Port:** Type 22.
 - **Connection Type:** Select SSH.
 - (Optional) **saved session:** If you do not want to repeat the configurations during the next logon, add a name for the session, and click **Save**.



3. Click **Open** to connect, and in the PuTTY Security Alert dialog box, click **Yes**.

**Note:**

For the first connection to an ECS instance, you have the PuTTY Security Alert as follows, which means PuTTY cannot guarantee the instance is the one that you think it is, so it can only provide the public key fingerprint of the instance for you to decide to trust the instance or not. If you select **Yes**, the public key will be added to the PuTTY's cache and you will not be alerted again during your next connection. If you select Yes but are alerted again, a *man-in-the-middle attack (MITM)* may occur. For more information, see [PuTTY User Manual](#).



4. Enter the user name and password for the Linux instance, and then press **Enter**.

**Note:**

The password is not displayed on screen.

If you are successfully connected to the instance, the following message is displayed.

```
Welcome to Alibaba Cloud Elastic Compute Service !
```

Linux or Mac OS X

If your local machine is running Linux OS or Mac OS X, follow these steps:

1. Run the command `ssh root@[Public IP address or EIP address of the instance]`.
2. Type the password and then press **Enter**.

If you are successfully connected to the instance, the following message is displayed.

```
Welcome to Alibaba Cloud Elastic Compute Service !
```

Android or iOS

If your local machine is running Android OS or iOS, see [connect to an instance on a mobile device](#).

Reference

You can run a script to install a graphical desktop on an instance running CentOS. For more information, see [automatic installation tool for Linux instance](#).

5.5 Connect to a Windows instance

If your Windows instance can access the Internet, you can use remote connection tools to connect to it. Otherwise, you can use the [Management Terminal](#).

Prerequisites

- The instance is in the **Running** status. If not, [start it](#).
- You have set a logon password for the instance. If the password is lost, you can [reset the password](#).
- The instance can access the Internet:
 - In a VPC, a public IP address is assigned to the instance or [an EIP address is bound to the instance](#).
 - In the classic network, a public IP address is assigned to the instance by using either of the following methods:
 - For a Subscription or a Pay-As-You-Go instance, you can select Assign public IP when creating the instance.
 - For a Subscription instance without a public IP address, you can assign one by [upgrading bandwidth](#).
- The following security group rules must be added to the security group that the instance joins. For more information, see [add security group rules](#).

Network Type	NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
VPC	N/A	Inbound	Allow	RDP(3389)	3389/3389	Address Field Access	0.0.0.0/0	1
Classic	Internet							

Procedure

Based on the operating system of your local machine, use one of the following methods to connect to a Windows instance:

- [Windows OS](#)
- [Linux](#)
- [Mac OS](#)
- [Android or iOS](#)

Windows OS

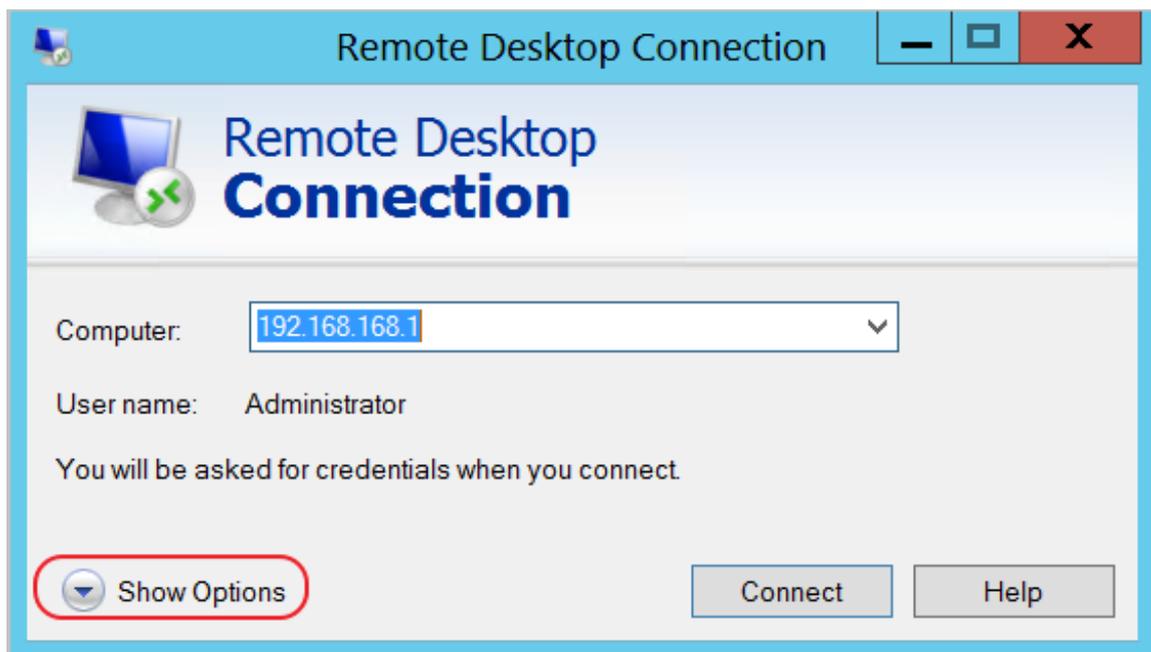
If the local machine is running Windows OS, you can use the `mstsc` to create a remote connection to a Windows instance.

1. Use any one of the following methods to start **mstsc**:

- Select **Start** > **icon** > **Remote Desktop Connection**.
- Click the **Start** icon and search for `mstsc`.
- Press the **Windows key** + **R** to open the **Run** window, type `mstsc`, and then press **Enter**.

2. In the **Remote Desktop Connection** dialog box, follow these steps:

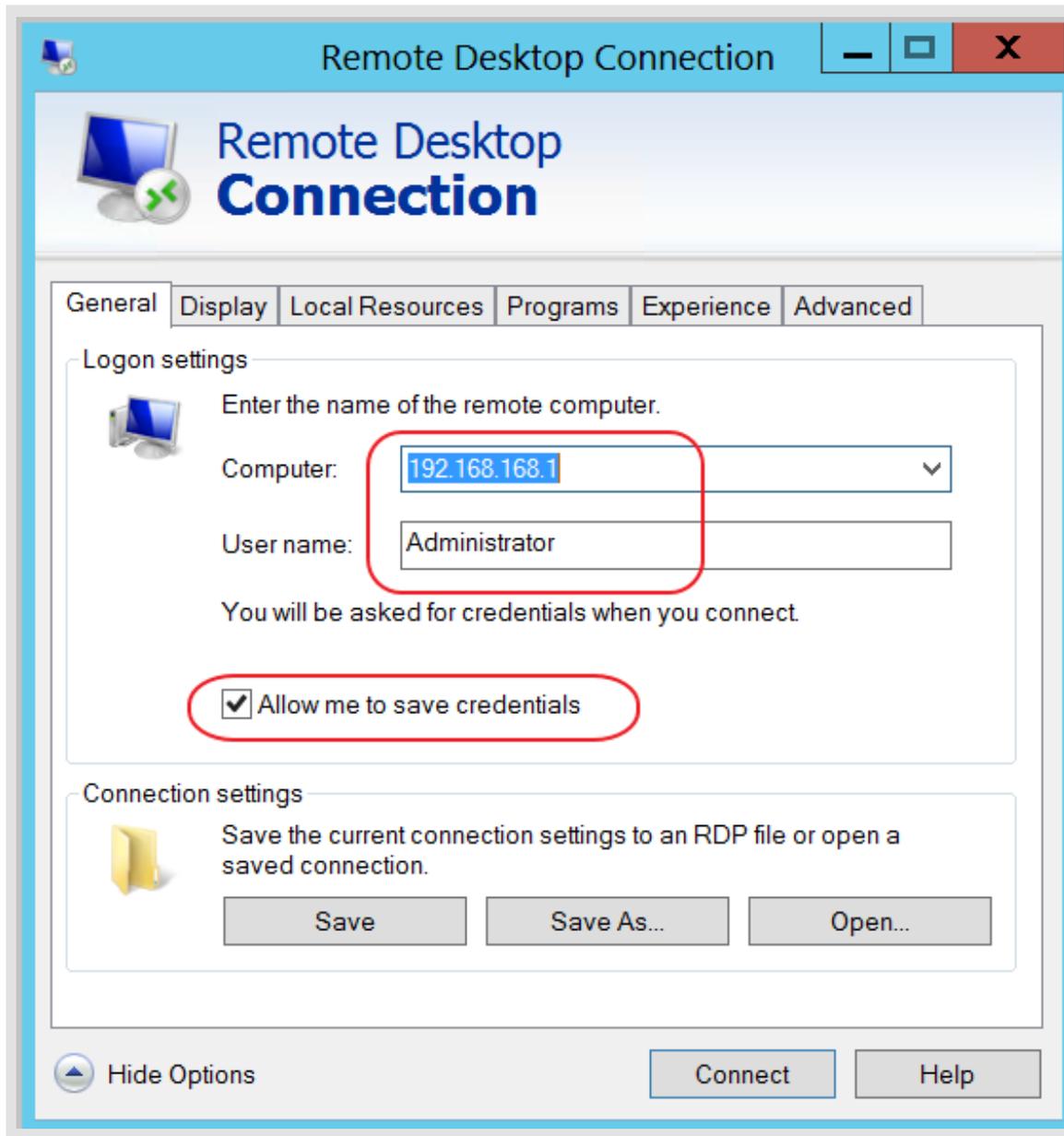
- a. Click the **Show Options** drop-down box.



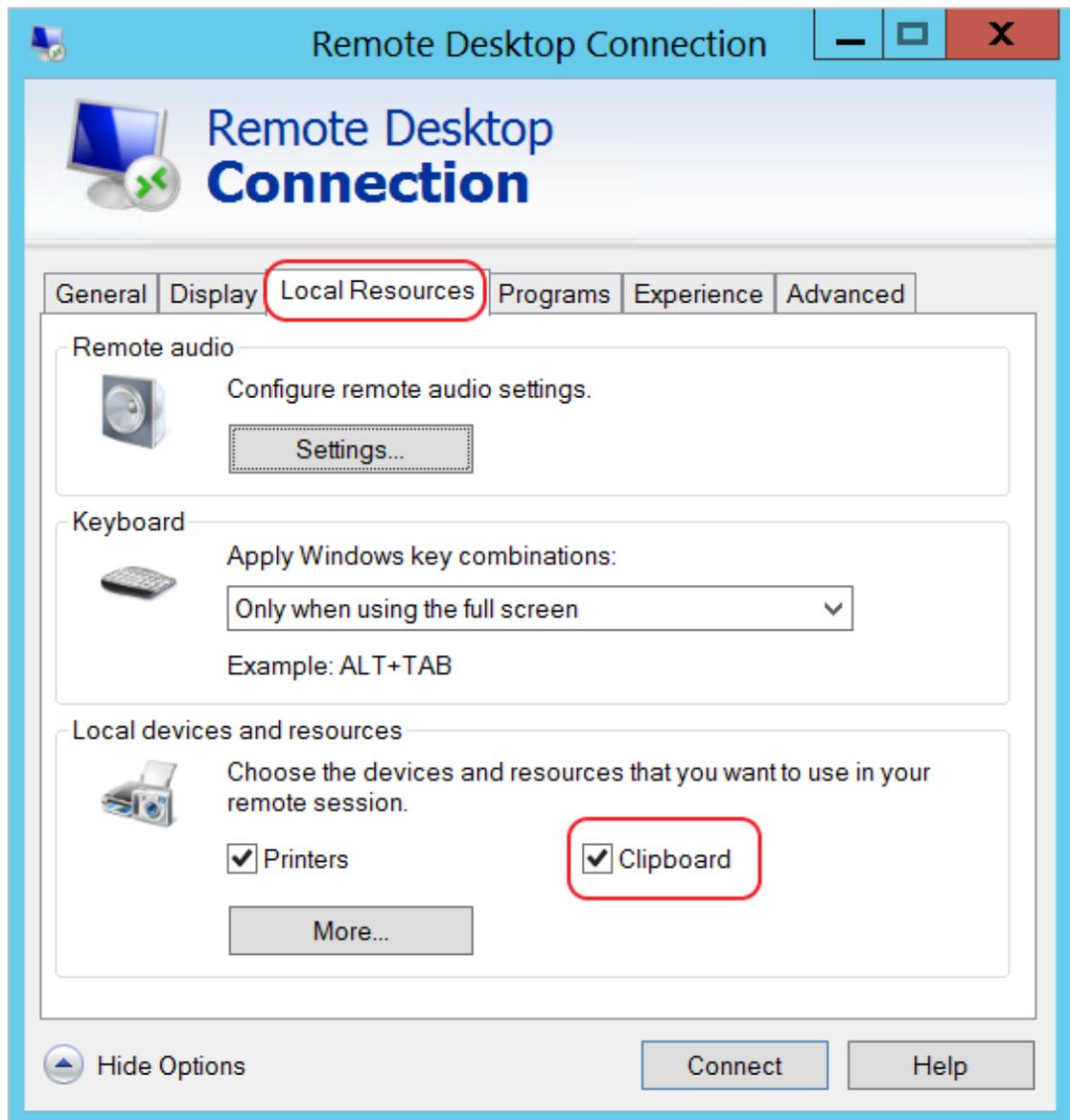
- b. Type the public IP address or EIP address of the instance.
- c. Type the user name. The default user name is **Administrator**

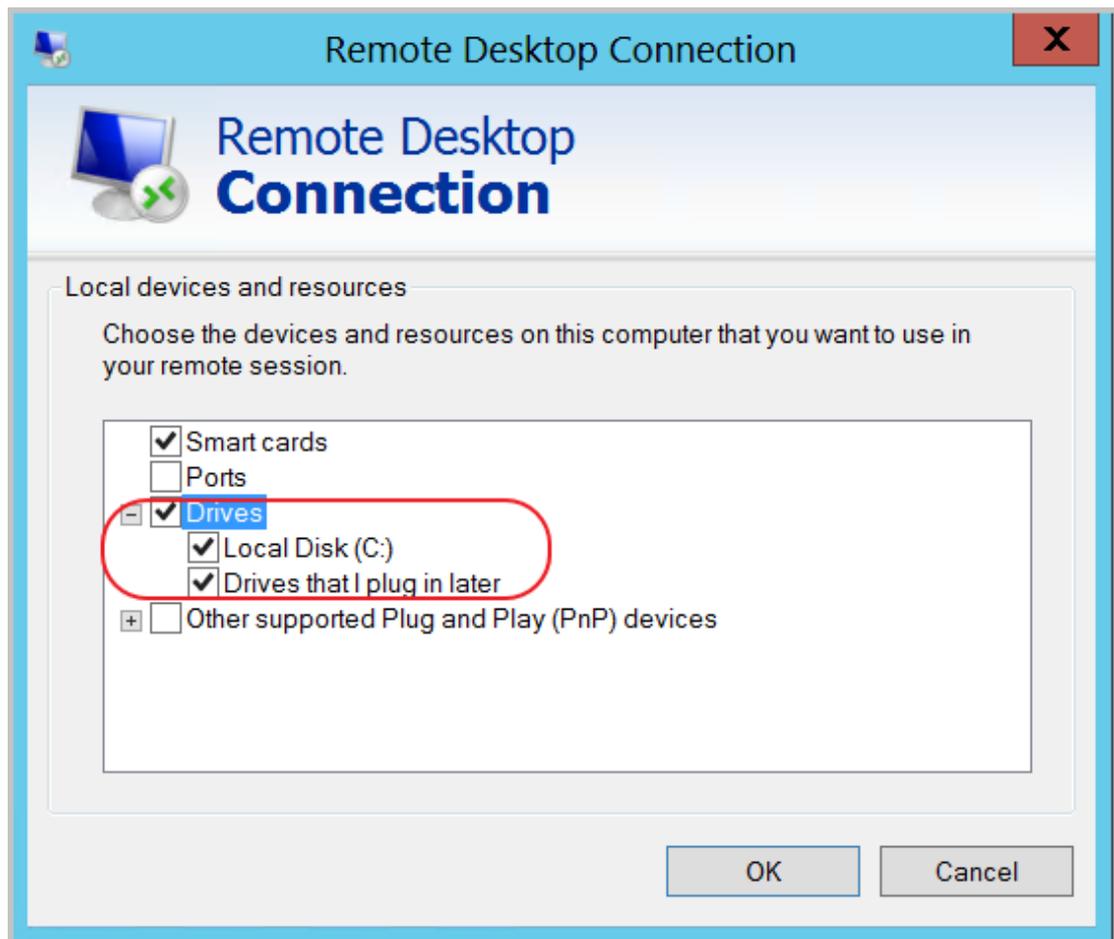
**Note:**

If you want to log on to the instance next time using the same credentials, select **Allow me to save credentials**.

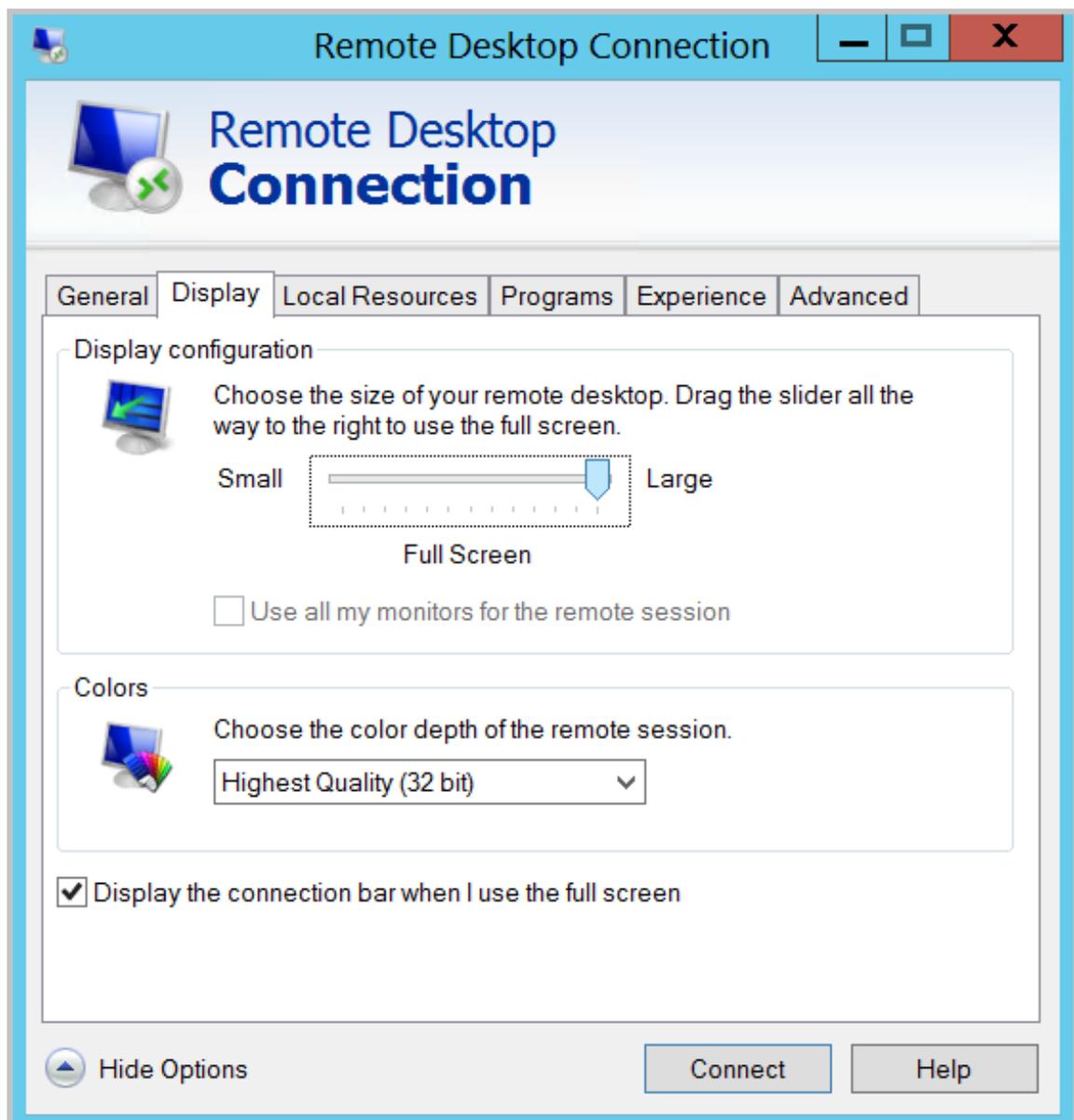


- d. Optional. If you want to copy text or files from the local machine to the instance, click the **Local Resources** tab to see options for sharing local computer resources.
- If you want to copy text only, select **Clipboard**.
 - If you also want to copy files, select **More** and select the drive letters from which you want to copy files to your instance and click **OK**.





- e. Optional. Click the **Display** tab to resize the remote desktop window. Full Screen is recommended.



- f. Click **Connect**.

Linux

If the local machine is running Linux OS, you can use a remote connection tool to create a remote connection to a Windows instance. This article takes rdesktop as an example to describe how to connect a Windows instance from a local machine running Linux.

1. Download and start rdesktop.

2. Run the command to connect to a Windows instance. Replace the parameter values with your own configurations.

```
rdesktop -u administrator -p password -f -g 1024*720 192.168.1.1 -r
clipboard:PRIMARYCLIPBOARD -r disk:sunray=/home/yz16184
```

The following table describes the parameters involved.

Parameters	Description
-u	The user name. The default user name for a Windows instance is Administrator.
-p	The password used to log on to the windows instance.
-f	Full screen by default. Use Ctrl+Alt+Enter to switched the mode.
-g	Resolution. Asterisks (*) are used for separation. If omitted, full-screen display is used by default.
192.168.1.1	The IP address of the server that requires remote connection. Replace it with the public IP or EIP address of your Windows instance.
-d	Domain name. For example, if the domain name is INC, then the parameter is <code>-d inc</code> .
-r	Multimedia reorientation. For example: <ul style="list-style-type: none"> Turn on the sound: <code>--r sound</code>. Use a local sound card: <code>-r sound : local</code>. Open the U Disk: <code>-r disk:usb=/mnt/usbdevice</code>.
-r clipboard:PRIMARYCLIPBOARD	Realizes direct word copying and pasting between Linux and Windows instances of local devices. Supports Chinese words copying and pasteing.
-r disk:sunray=/home/yz16184	Specifies that a directory on the Linux system of a local device maps to a hard disk on a Windows instance. If this is configured, Samba and FTP are not recommended for file transfers.

For more information about parameters of the `rdesktop` command, see [rdesktop documentation](#).

Mac OS

To connect to a Windows instance from a local machine running Mac OS, see [get started with Remote Desktop on Mac](#).

Android or iOS

If your local machine is running Android OS or iOS, see [connect to an instance on a mobile device](#).

5.6 Connect to an instance on a mobile device

This article describes how to connect to an ECS instance on a mobile device. Depending on the operating system of your instance, select the required method as follows.

- [Connect to a Linux instance](#): This example uses SSH Control Lite to describe how to connect to a Linux instance on an iOS device, and JuiceSSH to describe how to connect to a Linux instance on an Android device.
- [Connect to Windows instances](#): This example uses Microsoft Remote Desktop to describe how to connect to a Windows instance on an iOS or Android device.

Connect to a Linux instance

Prerequisites

- The instance is **Running** .
- The instance has a public IP address and is accessible from the Internet.
- You have set the logon password for the instance. If the password is lost, you can [reset the instance password](#).
- The security group of the instance has the [the following security group rules](#):

Network type	NIC	Rule direction	Authorization policy	Protocol type	Port range	Authorization type	Authorization object	Priority
VPC	No configuration required	Inbound	Allow	SSH(22)	22/22	Address Field Access	0.0.0.0/0	1
Classic	Internet							

- You have downloaded and installed the appropriate app:
 - For an iOS device, install SSH Control Lite.

- For an Android device, install JuiceSSH.

Procedure

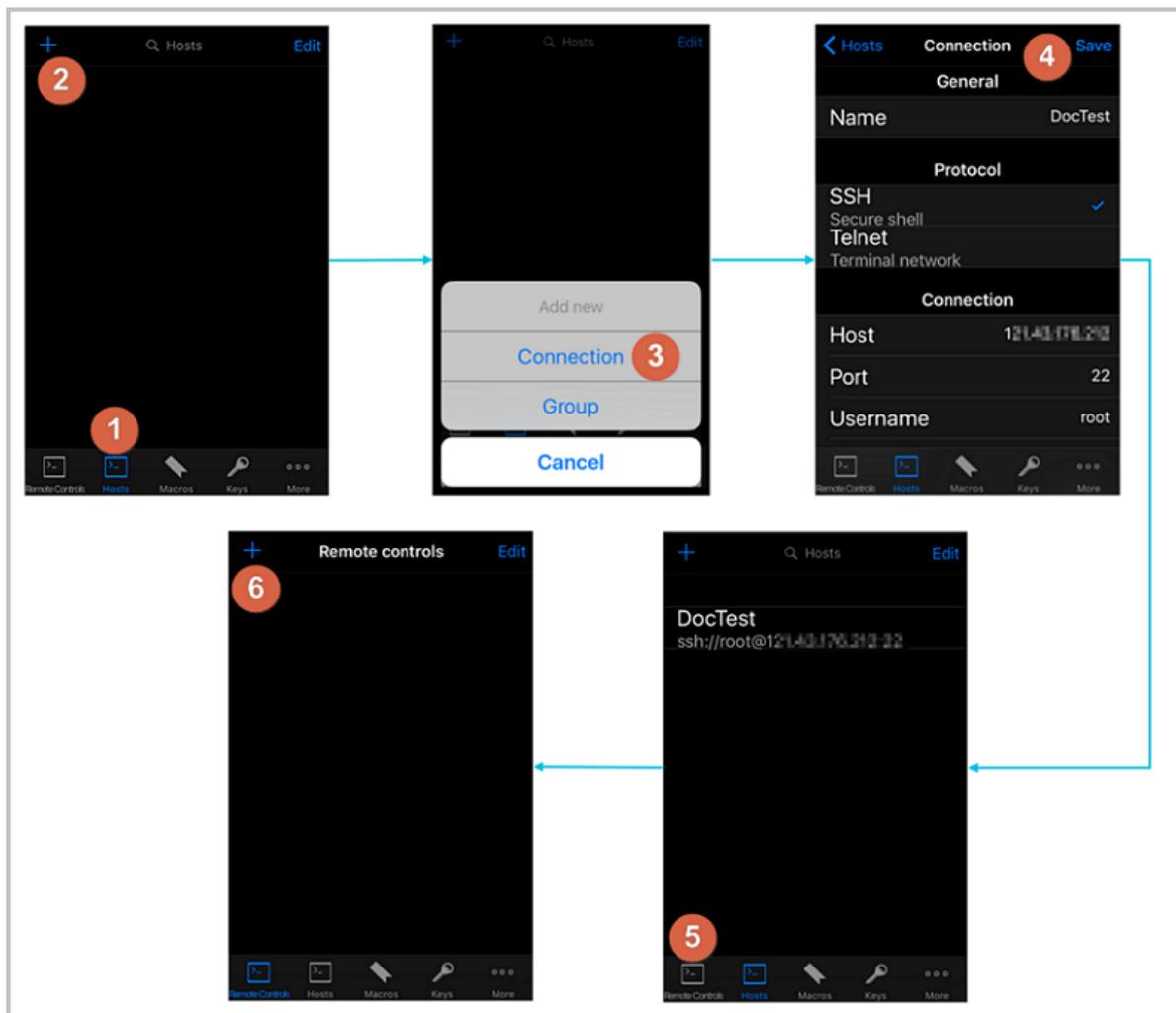
For iOS devices, see [use SSH Control Lite to connect to a Linux instance](#). In this example, a user name and password are used for authentication.

For Android devices, see [use JuiceSSH to connect to a Linux instance](#). In this example, a user name and password are used for authentication.

Use SSH Control Lite to connect to a Linux instance

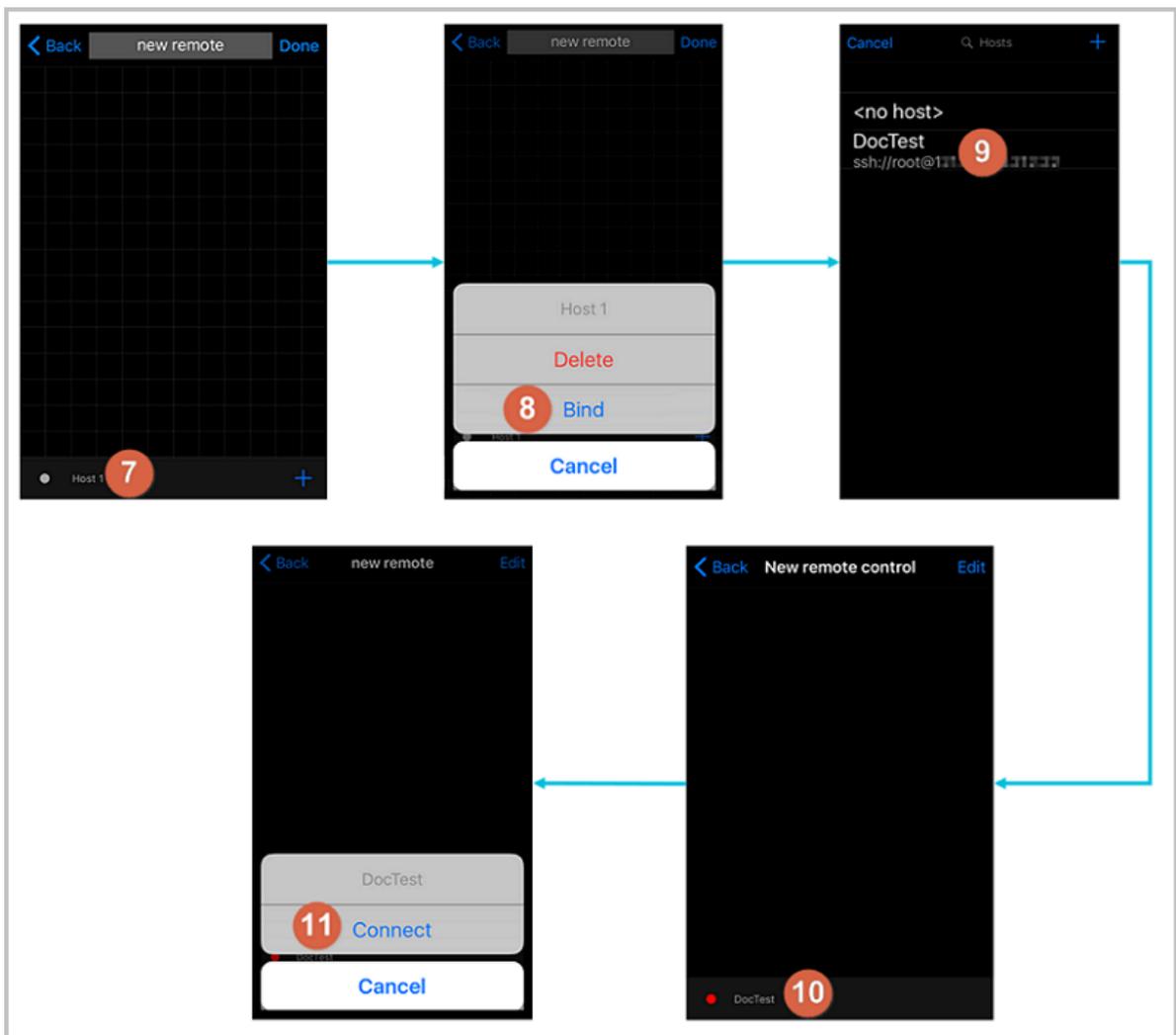
1. Start SSH Control Lite, and tap **Hosts**.
2. Tap the **+** icon in the upper left corner of the **Hosts** page.
3. In the action sheet, tap **Connection**.
4. On the **Connection** page, set the connection information and tap . The following connection information is required:
 - **Name:** Specify the Host name. `DocTest` is used in this example. .
 - **Protocol:** Use the default value SSH.
 - **Host:** Type the public IP address of the Linux instance to connect to.
 - **Port:** Type the port number for SSH protocol. 22 is used in this example.
 - **Username:** Type `root` for the user name.
 - **Password:** Type the logon password of the instance.
5. In the tool bar, tap **Remote Controls**.
6. On the **Remote Controls** page, tap the **+** icon in the upper left corner to create a remote connection session. `New remote` is used in this example.

The following figure shows Steps 1 through 6.



7. On the **New remote** page, tap **Host1**.
8. In the action sheet, tap **Bind**.
9. Select the new Linux instance. In this example, select `DocTest`.
10. On the **New remote** page, tap **Done** to switch it to the **Edit** mode, and then tap **DocTest**.
11. In the action sheet, tap **Connect**.

The following figure shows Steps 7 through 11.

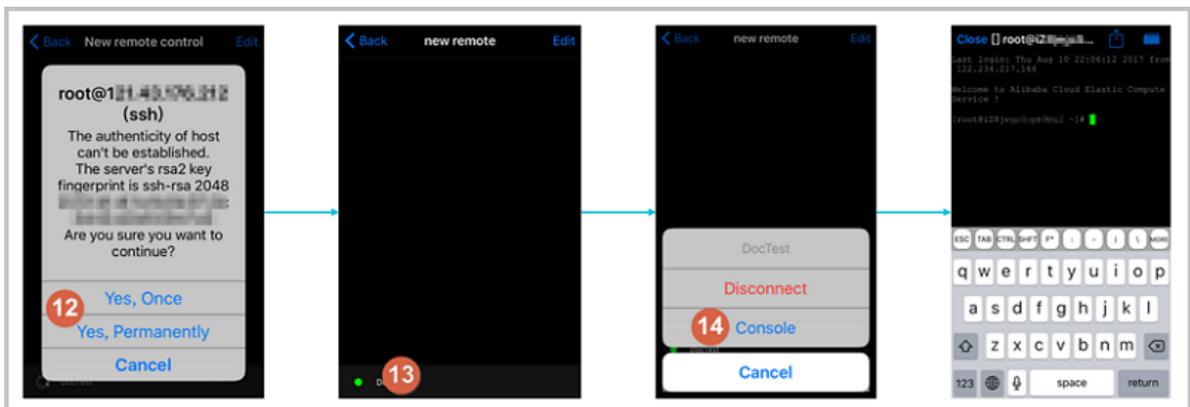


12.In the action sheet, select **Yes, Once** or **Yes, Permanently**. Once the connection is successful, the indicator in front of **DocTest** turns green.

13.On the **New remote** page, tap **DocTest**.

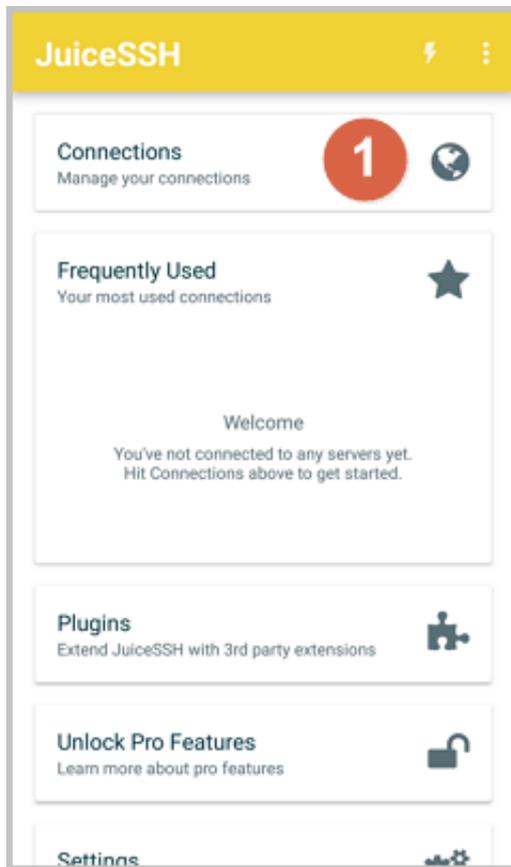
14.In the action sheet, tap **Console** to open Linux instance console.

The following figure shows Steps 12 through 14:

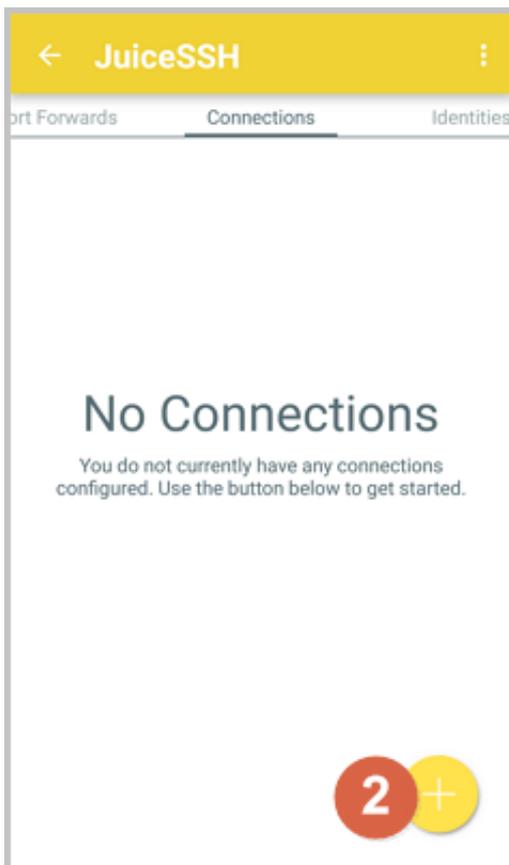


Use JuiceSSH to connect to a Linux instance

1. Start JuiceSSH, and tap **Connections**.



2. Under the **Connections** tab, tap the + icon.



3. On the **New Connection** page, add the connection information and then tap the  icon. The

following connection information is required:

- **Nickname:** Specify the name of the connection session. `DocTest` is used in this example.
- **Type:** Use the default value SSH.
- **Address:** Type the public IP address of the Linux instance to connect to.
- To set **Identity**, follow these steps:

1. Tap **Identity**, and tap **New** in the drop-down list.

2. On the **New Identity** page, add the connection information and then tap the  icon.

The following connection information is required:

- **Nickname:** Optional. `DocTest` is used in this example.
- **Username:** Type `root` for the user name.
- **Password:** Tap **SET(OPTIONAL)**, and type the logon password of the instance.

← New Identity ✓

IDENTITY

Nickname: DocTest

Username: root

Password: UPDATE / CLEAR

Private Key: SET (OPTIONAL)

SNIPPET

JuiceSSH Pro users can take advantage of an automatically generated snippet to add a public key to a servers ~/.ssh/authorized_keys file and set the correct permissions.

GENERATE SNIPPET

- **Port:** Type the port number for SSH protocol. In this example, 22 is used.

← New Connection 3 ✓

BASIC SETTINGS

Nickname: DocTest

Type: SSH

Address: 121.43.176.212

Identity: DocTest

ADVANCED SETTINGS

Port: 22

Connect Via: (Optional)

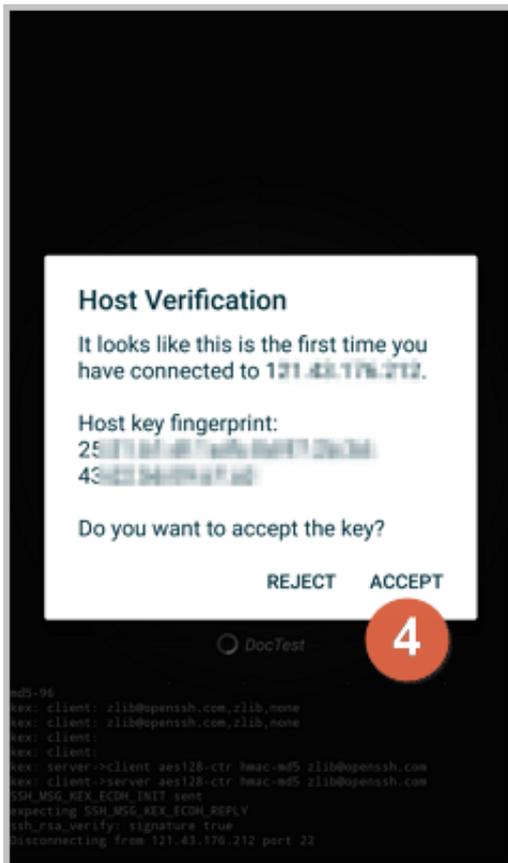
Run Snippet: (Optional)

Backspace: Default (sends DEL)

GROUPS

ADD TO GROUP

4. Confirm the message, and tap **ACCEPT**.



5. (Optional) For a first-time connection, the app will show a prompt of helpful tips. Tap **OK - I'VE GOT IT!**.



If you are successfully connected to the Linux instance, the following screen is displayed.

```

Last failed login: Tue Aug 29 10:42:15 CST 2017 from 180.180.180.89 on ssh:nd
tty
There were 8 failed login attempts since the last successful login.
Last login: Tue Aug 29 09:56:28 2017 from 36.23.155.166
Welcome to Alibaba Cloud Elastic Compute Service !
[root@i-instance-qa1000 ~]#

```

Connect to Windows instances

In this section, Microsoft Remote Desktop is used as an example to describe how to use an app to connect to a Windows instance on a mobile device.

Prerequisites

- The instance is **Running**.
- The instance has a public IP address and is accessible from the Internet.
- You have set the logon password for the instance. If the password is lost, you must [reset the instance password](#).
- The security group of the instance has [the following security group rules](#):

Network type	NIC	Rule direction	Authorization policy	Protocol type	Port range	Authorization type	Authorization object	Priority
VPC	No configuration required	Inbound	Allow	RDP(3389)	3389/3389	Address field access	0.0.0.0/0	1

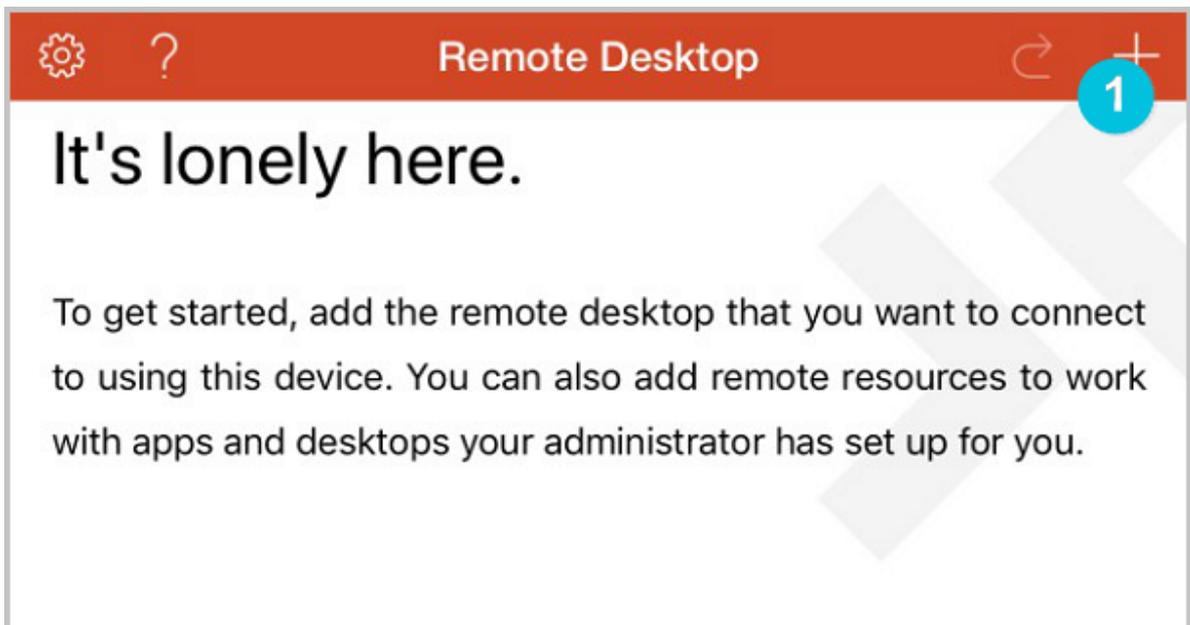
Network type	NIC	Rule direction	Authorization policy	Protocol type	Port range	Authorization type	Authorization object	Priority
Classic	Internet							

- You have downloaded and installed Microsoft Remote Desktop.
 - For iOS devices, download the app from iTunes.
 - For Android devices, download the app from Google Play.

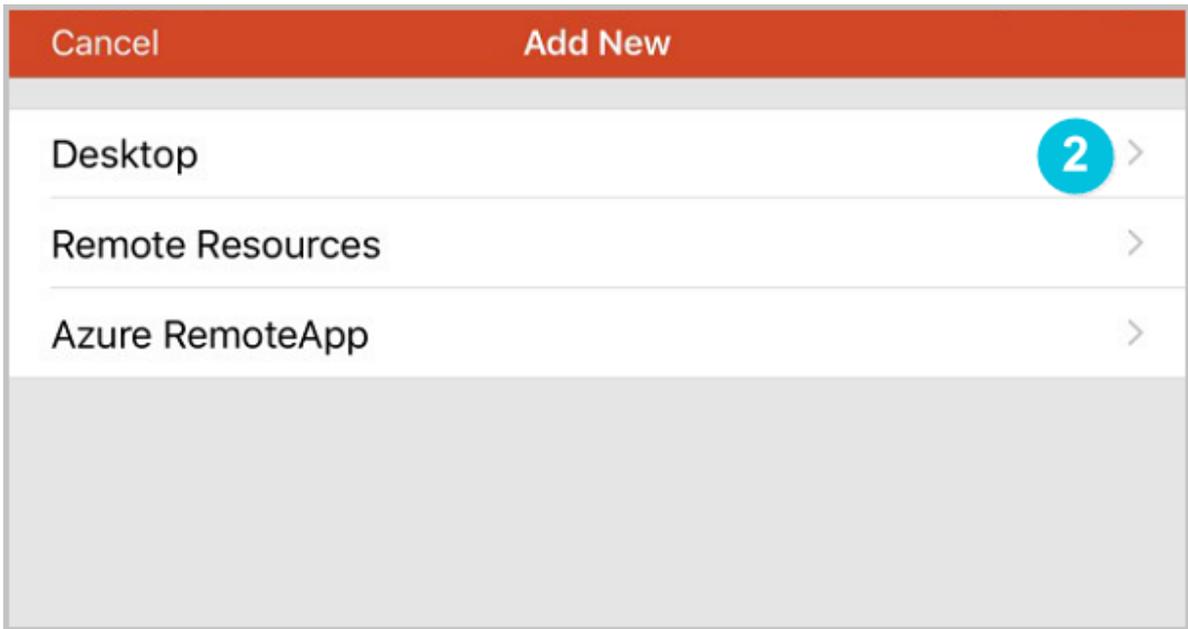
Procedure

To connect to a Windows instance by using Microsoft Remote Desktop, follow these steps:

1. Start RD Client. In the navigation bar, tap the + icon.

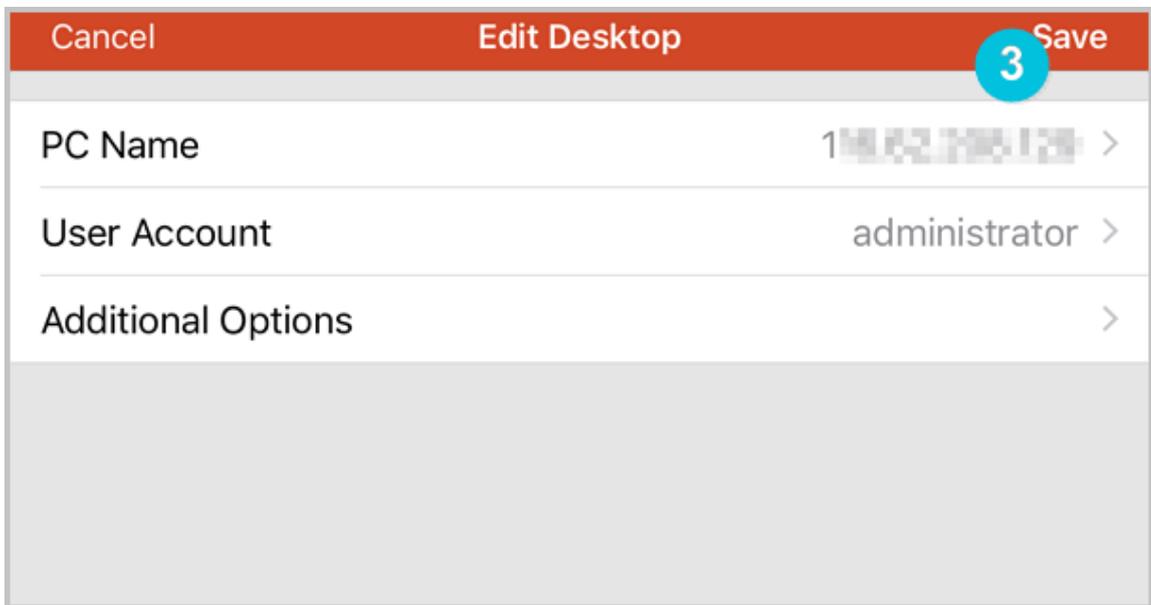


2. On the **Add New** page, select **Desktop**.

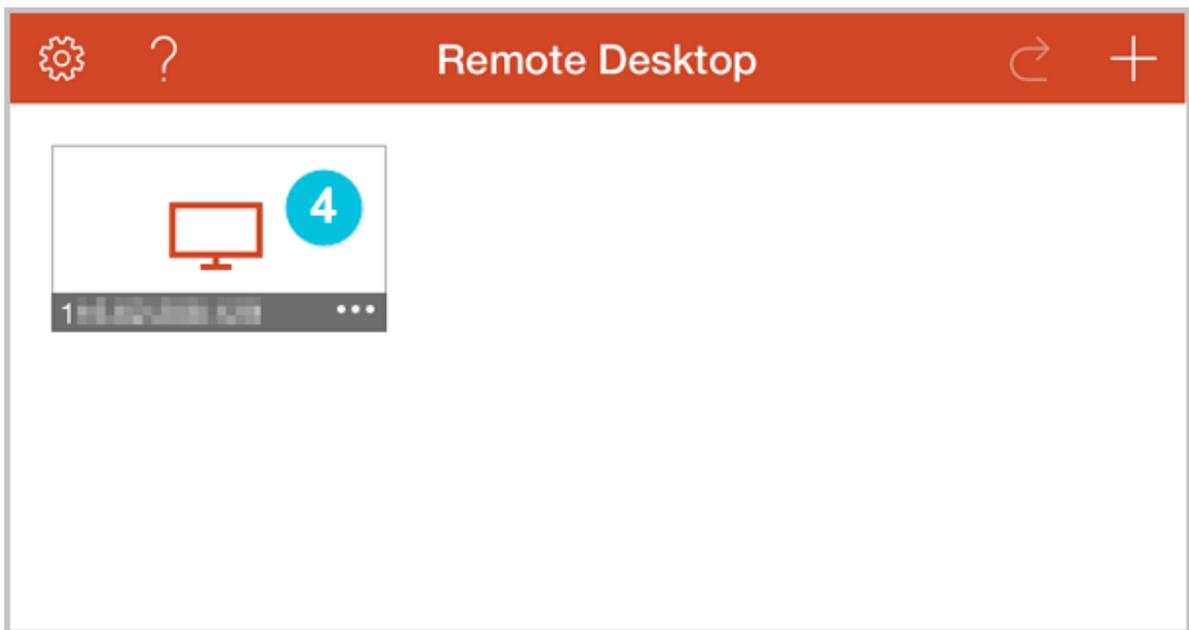


3. On the **Edit Desktop** page, type the connection information and tap **Save**. The following connection information is required:

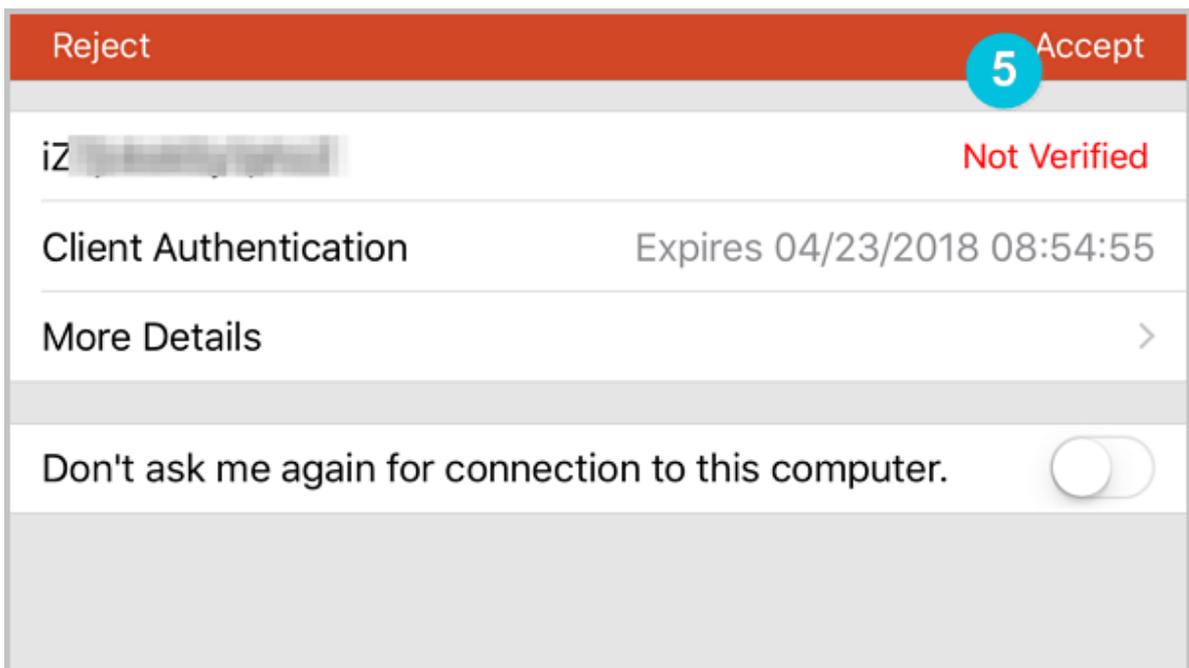
- **PC Name:** Type the public IP address of the Windows instance to connect to.
- **User Account:** Type the account name `administrator` and the logon password of the Windows instance.



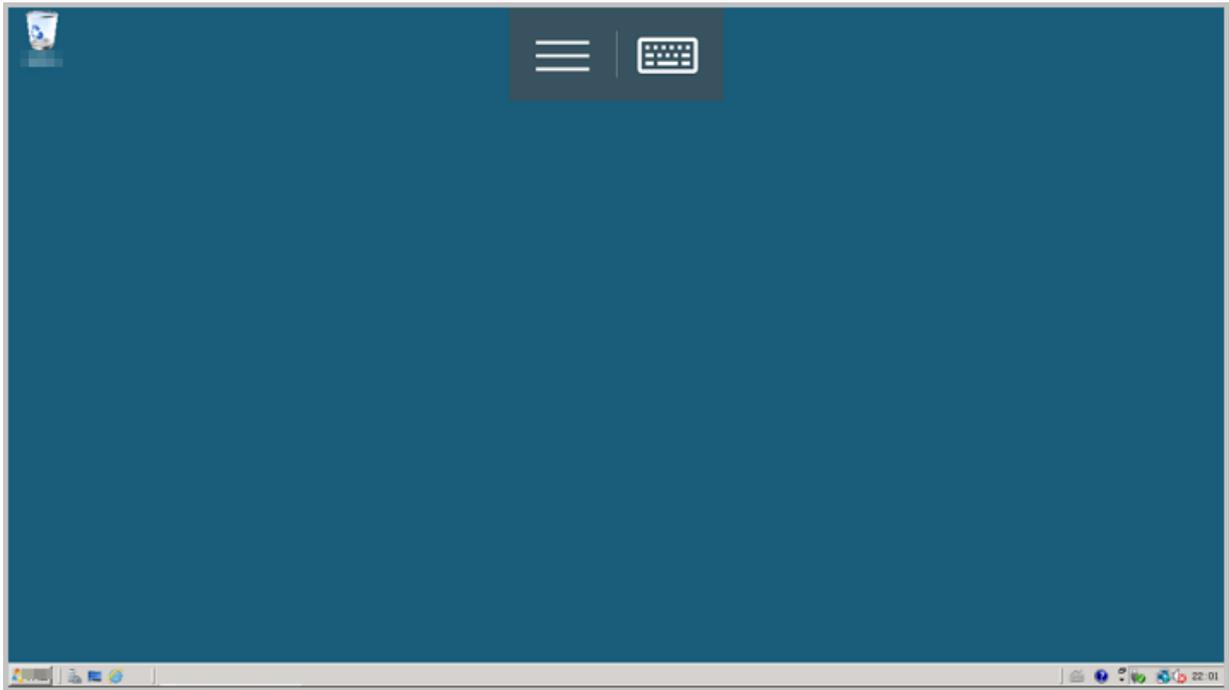
4. On the **Remote Desktop** page, tap the icon of a Windows instance.



5. On the confirmation page, confirm the message and tap **Accept**.



If you are successfully connected to the Windows instance, the following screen is displayed.



6 Cloud disks

6.1 Create a cloud disk

You can create a cloud disk to work as a data disk in the ECS console or by using the API. This article introduces how to create a new empty cloud disk in the ECS console.

Notes

Before you create a cloud disk, consider the following:

- Only [Pay-As-You-Go](#) cloud disks can be created in this way, and they can be used as data disks only.

**Note:**

You can create cloud disks as data disks when creating an ECS instance. Those disks have the same billing method of the instance.

- You can create a new empty cloud disk or [create a cloud disk from a snapshot](#).
- The quota of the Pay-As-You-Go cloud disks that are used as data disks of each account in all regions is five times than that of the Pay-As-You-Go instances. For more information, see [limits](#).
- Currently, you cannot merge multiple cloud disks. After cloud disks are created, they are independent from each other, and you cannot merge their space by formatting. We recommend that you determine the number of disks and disk sizes required for your business before you create cloud disks.
- Because you can create a snapshot for a single cloud disk, we do not recommend that you create LVM (Logical Volume Manager) volumes as the volumes may result in data loss if you use the snapshot to roll back the cloud disk.
- You can convert a Pay-As-You-Go billed cloud disk to Subscription as follows:
 - [Upgrade configurations of Subscription instances](#).
 - [Switch from Pay-As-You-Go to subscription](#).
- If a cloud disk is created in this way, and its billing method is retained as Pay-As-You-Go, you can [detach a cloud disk](#) and [release a cloud disk](#) at any time.

Prerequisites

If you want to [attach a cloud disk](#) to an instance, make sure they are in the same region and zone.

Procedure

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Block Storage > Disks**.
3. In the upper-right corner of the **Disks** list page, click **Create Disk** to go to the **Create** page.
4. Select the target region and zone.



Note:

If you want to attach the cloud disk to an ECS instance, they must be in the same zone and the same region.

5. Select a cloud disk category and specify the disk size and the quantity. You can also select [create a cloud disk from a snapshot](#).
6. Confirm the configuration and the **Total** cost.
7. Click **Preview**, confirm your order, and click **Create**.

After you complete the payment, return to the **Disks** page and refresh it. The new disk is displayed and its status is **Available**.

Additional operations

[Attach a cloud disk](#).

Related APIs

To create a disk after creating an instance, see [CreateDisk](#).

To create a cloud disk when creating an instance, see [RunInstances](#) or [CreateInstance](#).

6.2 Create a cloud disk from a snapshot

This article describes how to create a cloud disk from a snapshot in the ECS console. You can take a snapshot of an existing system disk, or data disk, and create a cloud disk from the snapshot. The new disk can be attached to any instance in the same zone of the same region.

Scenario

If you need to access data from a snapshot, but do not want to [roll back a cloud disk](#), you can create a cloud disk from the snapshot to access data that you need. For example, if your instance encounters a system disk failure, you can use an existing snapshot to create a cloud disk, and attach the disk to a healthy instance. By doing so, you can restore the data of the affected instance.

Disk Performance

If a cloud disk is created from a snapshot, the initial disk performance decreases because data needs to be accessed from OSS before being written into the disk. We recommend that you write and read every data block at least once before production use. For more information about OSS, see [what is OSS](#).

Considerations

- Only [Pay-As-You-Go](#) cloud disks can be created in this way, and they can only be used as data disks.



Note:

You can set cloud disks as data disks when creating an ECS instance. The disks then have the same billing method as that of the instance.

- You can create a new empty cloud disk. For more information, see [create a cloud disk](#).
- The quota of Pay-As-You-Go cloud disks that are used as data disks of each account in all regions is five times the quota of Pay-As-You-Go instances. For more information, see [limits](#).
- Currently, you cannot merge multiple cloud disks. After cloud disks are created, they are independent from each other, and you cannot merge their space by formatting. We recommend that you confirm the amount and size required before you create cloud disks.
- You can create a snapshot for a single cloud disk, so we do not recommend that you create LVM (Logical Volume Manager) volumes, which may cause data loss when you use the snapshot to rollback the cloud disk.
- After a Pay-As-You-Go cloud disk is created, you can convert its billing method to Subscription:
 - If it is attached to a Subscription instance, use the [upgrade configurations of Subscription instances](#) feature.
 - If it is attached to a Subscription instance, use the [switch from Pay-As-You-Go to Subscription](#) feature.
- If a cloud disk is created in this way, and its billing method is not converted, you can [detach a cloud disk](#) and [release a cloud disk](#) at any time.

Prerequisites

- You have created a snapshot for your instance, and you make sure the region and zone. For specific actions, see [create snapshots](#).
- [Attach a cloud disk](#). The instance and the cloud disk must be in the same region and zone.

Procedure

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Block Storage > Disks**.
3. In the upper-right corner of the **Disks** list page, click **Create Disk**.
4. Select a region and zone.



Note:

If you want to attach the cloud disk to an ECS instance, they must be in the same zone of the same region.

5. Configure the cloud disk:
 - a. Select a cloud disk category. The category of the source disk of the snapshot does not modify the configuration.
 - b. Click **Create from snapshot** and select a snapshot.
 - c. Specify the size of the cloud disk. The size range is 20 GiB to 32768 GiB. If the selected snapshot is smaller than 20 GiB, you can adjust the size manually. For a snapshot larger than 20 GiB, the size is adjusted automatically according to the snapshot size. However, if you replace the snapshot, you must manually set the size.
 - d. For Purchase Plan, set the quantity.
6. Check the cost.
7. Click **Preview**, confirm you order, and click **Create**.

After you complete the payment, return to the **Disks** page and refresh it. The new disk is displayed and its status is **Available**.

Additional operation

[Attach a cloud disk](#).

Related API

Create a cloud disk: [CreateDisk](#)

6.3 Attach a cloud disk

You can create a cloud disk and attach it to an ECS instance to function as a data disk by going to the Instance Disks page or on the Disk List page.

Note

Before you attach a cloud disk to an ECS instance, consider the following:

- If a cloud disk is created at the same time as an ECS instance, you do not have to attach the disk.
- You can attach a cloud disk to work as a data disk only, but not as a system disk.
- To attach a cloud disk to an ECS instance, the instance must meet the following requirements:
 - The instance must be in the **Running** or **Stopped** status. It cannot be in **Locked** status.
 - The instance must not have any overdue payments.
- The disk to be attached must be in the **Available** status.
- The cloud disk and the ECS instance must be in the same region and the same zone.
- Up to 16 cloud disks can be attached to an ECS instance to work as data disks. However, a cloud disk cannot be attached to multiple instances simultaneously.
- A cloud disk can be attached to an ECS instance, regardless of the billing method of the instance.

Prerequisites

You must create an ECS instance and a cloud disk in the same region and zone. For more information, see [create a cloud disk](#) and [create an instance](#) in *Quick Start*.

Attach a cloud disk on the Instance Disks page

To attach one or multiple cloud disks to a specified ECS instance, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Instances**.
3. Select the target region.
4. Find the target ECS instance and click its ID to go to its **Instance Details** page.
5. In the left-side navigation pane, click **Disks** and then, on the **Disks** page, click **Mount**.
6. In the dialog box, complete the following configurations:
 - **Target Disk:** Select a cloud disk in the **Unmounted** status in the same region and zone.

- **Release Disk with Instance:** If you select this option, the disk is released when you release its corresponding instance.
- **Delete Automatic Snapshots While Releasing Disk:** If you select this option, all the automatic snapshots of the target disk are deleted when you release it. However, all the manual snapshots are retained. To keep a complete data backup, we recommend that you do not select this option.

Click **OK** and then click **Mount**.

7. Refresh the Disk List.

When the status of the cloud disk is **In Use**, the attachment is successful.

8. According to the content of the cloud disk and the operating system of the ECS instance, perform follow-up operations as required to make the disk ready for use. The following table details the different follow-up operations available.

Disk content	Operating system of the ECS instance	Follow-up operations
A new empty cloud disk	Linux	Format a data disk for Linux instance . If the cloud disk is larger than 2 TiB, see partition and format data disk more than 2 TiB .
	Windows	Format a data disk for Windows instances . If the cloud disk is larger than 2 TiB, see partition and format data disk more than 2 TiB .
A cloud disk from a snapshot	Linux	Connect to the Linux instance and run the <code>mount</code> command to mount the partitions to make the disk ready for use.
	Windows	No follow-up operations are required. The cloud disk is ready for use.

Attach a cloud disk on the Disk List page

To attach a cloud disks to an ECS instances, follow these steps:

1. Log on to the [ECS console](#).

2. In the left-side navigation pane, select **Block Storage > Disks**.
3. Select the target region.
4. Find a cloud disk in the **Unmounted** status and then, in the **Actions** column, select **More > Mount**.
5. In the dialog box, complete the following configurations:
 - **Target Instance:** Select an ECS instance in the same zone.
 - **Release Disk with Instance:** If you select this option, the disk is released when you release its instance.
 - **Delete Automatic Snapshots While Releasing Disk:** If you select this option, all the automatic snapshots of the selected disk are deleted when you release the disk. However, all the manual snapshots are retained. To keep complete data backup, we recommend that you do not select this option.

Click **Mount**.

6. Refresh the disk list.

When the status of the cloud disk is **In Use**, the attachment is successful.

7. According to the content of the cloud disk and the operating system of the ECS instance, perform follow-up operations as required to make the disk ready for use. The following table details the different follow-up operations available.

Disk content	Operating system of the ECS instance	Follow-up operations
A new empty cloud disk	Linux	Format a data disk for Linux instance . If the cloud disk is larger than 2 TiB, see partition and format data disk more than 2 TiB .
	Windows	Format a data disk for Windows instances . If the cloud disk is larger than 2 TiB, see partition and format data disk more than 2 TiB .
A cloud disk from a snapshot	Linux	Connect to the Linux instance and run the <code>mount</code> command to mount the partitions to make the disk ready for use.

Disk content	Operating system of the ECS instance	Follow-up operations
	Windows	No follow-up operations are required. The cloud disk is ready for use.

Additional operations

After a cloud disk is attached to an ECS instance, you can perform any of the following operations according to your business needs:

- You can [reinitialize a cloud disk](#) to restore it to the initial status after it is created.
- You can increase the size of the cloud disk by resizing it. For more information, see [Linux _ Resize a data disk](#) or [Windows _ Resize a data disk](#).
- You can [create snapshots](#) of the cloud disk to back up data. Alternatively, you can [apply automatic snapshot policies to disks](#).
- You can use a snapshot to [roll back a cloud disk](#) to restore the cloud disk to a previous state.
- You can [detach a cloud disk](#) and [release a cloud disk](#) when you no longer require a cloud disk to reduce costs.

Related API

[AttachDisk](#)

6.4 Partition and format data disk more than 2 TiB

This article describes how to partition and format a large data disk in different operating systems using the GPT format. Note that a large data disk indicates a disk greater than 2 TiB, while a disk size smaller than 2 TiB indicates a small data disk.



Note:

If you want to partition and format a data disk less than 2 TiB, please see [Format a data disk for Linux instances](#) and [Format a data disk for Windows instances](#).

Note

Before partition and formatting a large data disk, note the following:

- Large data disks support the partition tools and file systems shown in the following table.

Operating system	Partition tool	File system
Linux	parted	ext4 or xfs
Windows	Disk management	NTFS

- **We recommend that you do not create a large data disk by using a snapshot of a small data disk.**

If you create a large data disk by using the snapshot of a small data disk, the following risks may occur:

- The system expands the block-level of the device's disk, but does not automatically convert between the partition format and the file system.
- If the MBR format is used in the snapshot of the small data disk, none of the supported partition tools (`parted` on Linux and **Disk Management** on Windows) can convert the MBR to GPT and retain the data. Therefore, even if you create a large data disk by using a snapshot of a small data disk, while partitioning and initializing, you must delete the original data and partition with the GPT format. If you have created large data disk by using a snapshot of a small data disk, see [use windows to partition and format a large data disk created by a snapshot of a small data disk](#).



Note:

This is not the case if the snapshot of the small data disk is in GPT format, or if you have another powerful partitioning tool. You can select based on your own situation.

Instead, create an empty large data disk, or create large data disk by using snapshots of large data disks, because of the following reasons

- **Effect of data disk snapshots**

No matter if you are using a large data disk or a small data disk, the process for creating a snapshot from a data disk is proportional to the total data volume of the disk. If the disk contains a large amount of compromised or damaged data, or residual data fragments, then the snapshot creation time is extended

Use Windows to partition and format an empty large data disk

The following examples uses a Windows Server 2008 R2 64-bit operating system to describe how to partition and format a large data disk in Windows instance. Assume the data disk to be processed is a 4 TiB empty disk.

Prerequisites

The data disk has been attached to an instance. For more information, see [attach a cloud disk](#).

Procedure

To partition and format a large data disk, follow these steps:

1. [Connect to a Windows instance](#).

2. Click the  icon in the task bar.

3. In the left-side navigation pane select **Storage > Disk Management**.

4. Find the disk that is to be partitioned and formatted (in this example, **Disk 4**). The disk status should be shown as **Offline**.

5. Right click the blank area around Disk 4, and then click **Online**.

After going online, Disk 4 enters the **Not Initialized** status.

6. Right click the blank area around Disk 4, and then select **Initialize Disk** in the context menu.

7. In the **Initialize Disk** dialog box, select **Disk 4** and select **GPT** as the disk partitioning method.

8. In the **Disk Management** window, right click the **Unallocated** area of Disk 4, and then select **New Simple Volume** to create a 4 TiB volume in the NTFS format.

9. In the **New Simple Volume Wizard**, follow these steps:

a. Click **Next**.

b. Choose a volume size: designate size of simple volume. If you want to create a master area only, use the default value. Click **Next**. You can also partition **Disk 4** into several partitions.



Note:

The maximum NTFS volume is, theoretically, the maximum volume of NTFS containing $2^{64}-1$ clusters. However, in Windows XP Pro, the maximum volume of NTFS is $2^{32}-1$ clusters. For example, for a 64 KiB cluster, the maximum NTFS volume is approximately 256 TiB. If you select a 4 KiB cluster, the maximum NTFS volume is 16 TiB. NTFS selects the size of a cluster automatically based on the disk capacity.

c. Assign drive letter and path: select a drive letter, then select G in this instance. Click **Next**.

d. Format Partition: Select the formatting settings, including file system, distributed unit size, and volume label, and then confirm whether to **Perform a quick format** and **Enable file and folder compression**. Select **Perform a quick format** here only. Click **Next**.

- e. Start creating a new simple volume by following the prompts displayed by the wizard. Then, click **Finish**.

After the formatted partition is completed, in **Disk Management**, the status of **Disk 4** is shown in the following screenshot.

Use Windows to partition and format a large data disk created by a snapshot of a small data disk

If you create a large data disk by using snapshots of a small data disk, you first need to convert the partition format of data disk from MBR to GPT, and then format the data disk. However, because data of the original snapshots is saved, we recommend you do not create large data disk by using a snapshot of a small data disk.

If you have already created large data disks in this method, perform the following actions to partition and format this data disk. The following example operating system uses a Windows Server 2012 R2 64-bit operating system. Assume the data disk to be processed is a 3 TiB disk.

Prerequisites

The data disk has been *attached* to an instance.

Procedure

To partition and format a large data disk, follow these steps:

1. *Connect to a Windows instance*.
2. On the Windows Server desktop, right click the **Start** icon, and select **Disk Management**.

The data disk (Disk 2, in this example) that has not been formatted or partitioned is in the **Offline** status.
3. Right click the blank area around Disk 2, and then select **Offline** in the context menu.
4. Right click a simple volume, and then select **Delete Volume** in the context menu.
5. Right click the blank area around Disk 2, and then select **Convert to GPT Disk** in the context menu.
6. In the **Disk Management** window, right click **Unallocated** area of Disk 2, and then select **New Simple Volume** to create a 3 TiB volume in the NTFS format.
7. In the **New Simple Volume Wizard**, follow these steps:
 - a. Click **Next**.
 - b. Specify the size of the simple volume. If you need only one primary partition, use the default value, and then click **Next**. You can also partition **Disk 2** into several partitions.

**Note:**

The maximum NTFS volume is, theoretically, the maximum volume of NTFS containing $2^{64}-1$ clusters. However, in Windows XP Pro, the maximum volume of NTFS is $2^{32}-1$ clusters. For example, for a 64 KiB cluster, the maximum NTFS volume is approximately 256 TiB. If you select a 4 KiB cluster, the maximum NTFS volume is 16 TiB. NTFS selects the size of a cluster automatically based on the disk capacity.

- c. Assign Drive Letter or Path: Select a drive letter, and then click **Next**.
- d. Format Partition: Select the formatting settings, including file system, distributed unit size and volume label, and then confirm whether to **Perform a quick format** and **Enable file and folder compression**. Select **Perform a quick format** here only. Click **Next**.
- e. Start creating a new simple volume by following the prompts displayed by the wizard. Then, click **Finish**.

After the formatted partition is completed, in **Disk Management**, the status of **Disk 4** is shown in the following screenshot.

Use Linux to partition and format a large data disk

To partition and format a large data disk that is attached to a Linux instance, use the GPT format. In Linux system, large data disks normally uses xfs or ext4 file system.

The following example uses a CentOS 7.4 64-bit operating system. This section describes how to use **parted** and **e2fsprogs** tools to partition and format a large data disk on a Linux instance. Assume the data disk to be processed is an empty 3 TiB new disk, and the device name is `/dev/vdd`.

Prerequisites

Your Linux instance has installed **parted**. If not, run `yum install -y parted`.

Your Linux instance has installed **e2fsprogs**. If not, run `yum install -y e2fsprogs`.

The data disk has been attached to the instance. For more information, see [attach a cloud disk](#).

Procedure

To partition and format a large data disk and mount the file system, follow these steps:

1. Run `fdisk -l` to check whether the data disk exists. If the data disk is successfully mounted, the following result is returned

```
Disk /dev/vdd: 3221.2 GB, 3221225472000 bytes, 6291456000 sectors
```

```
Units = sectors of 1 * 512 = 512 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes
```

2. Run `parted /dev/vdd` to start partitioning:

- a. Run `mklabel gpt`, to convert partitioning format from MBR to GPT.
- b. Run `mkpart primary ext4 <StartSector> <EndSector>` to partition a primary partition by using the ext4 file system, and specify a start sector and end sector for the partition. If a data disk is partitioned into one partition only, run `mkpart primary ext4 0 -1`.



Note:

You can also use xfs file system.

- c. Run `print` to check partition table.

```
(parted) mkpart primary ext4 0 -1
Warning: The resulting partition is not properly aligned for best
performance.
Ignore/Cancel? ignore
(parted) print
Model: Virtio Block Device (virtblk)
Disk /dev/vdd: 3221 GB
Sector size (logical/physical): 512B/512B
Partition Table: gpt
Disk Flags:
Number Start End Size File system Name Flags
1 17.4kB 3221GB 3221GB primary
```

- d. Run `quit` to exit **parted** .

3. Run `partprobe` to make system re-read the partition table.

4. Run the following commands to create an ext4 file system, and make `/dev/vdd1` partition use ext4.

```
mke2fs -O 64bit,has_journal,extents,huge_file,flex_bg,uninit_bg,
dir_nlink,extra_isize /dev/vdd1
```



Note:

- If you want to disable the lazy init function of ext4 file system to avoid its effect on data disk I/O performance, see [disable lazy init function](#).
- If capacity of the data disk is 16 TiB, you must format it by using e2fsprogs in the designated version. See [update e2fsprogs](#).
- If you want to create an xfs file system, run `mkfs -t xfs /dev/vdd1`.

5. Run `mkdir /test` to create a mount point with the name `/test`.

6. Run `mount /dev/vdd1 /test` to mount `/dev/vdd1` to `/test`.
7. Run `df -h` to check current disk space and usage.

If the returned result shows the newly created file system information, the mount operation was successful, and you can use the new file system directly (that is, you do not need to restart the instance).

```
[root@izXXXXz ~]# df -h
Filesystem Size Used Avail Use% Mounted on
/dev/vda1 40G 6.4G 31G 18% /
devtmpfs 487M 0 487M 0% /dev
tmpfs 497M 0 497M 0% /dev/shm
tmpfs 497M 364K 496M 1% /run
tmpfs 497M 0 497M 0% /sys/fs/cgroup
tmpfs 100M 0 100M 0% /run/user/0
/dev/vdd1 2.9T 89M 2.8T 1% /test
```

8. (Optional) Write new partition information to `/etc/fstab` to enable automatic mount partition while the instance is started.
 - a. (Optional) Run `cp /etc/fstab /etc/fstab.bak` to back up `etc/fstab`.
 - b. Run `echo /dev/vdd1 /test ext4 defaults 0 0 >> /etc/fstab` to write new partition information to `/etc/fstab`.
 - c. Run `cat /etc/fstab` to check `/etc/fstab` information.

If the new partition information is in the returned result, the write operation was successful.

Appendix 1: Update e2fsprogs

If the disk capacity is 16 TiB, you must use e2fsprogs of version 1.42 or later to format its partitions to ext4 file system. If e2fsprogs version is too low (for example, e2fsprogs 1.41.11), the following error occurs.

```
mkfs.ext4: Size of device /dev/vdd too big to be expressed in 32 bits
using a blocksize of 4096.
```

To install e2fsprogs of later version, such as 1.42.8 in this example, follow these steps:

1. Run `rpm -qa | grep e2fsprogs` to check the current version of e2fsprogs.

```
$sudo rpm -qa | grep e2fsprogs
e2fsprogs-libs-1.41.12-3
e2fsprogs-1.41.12-3
e2fsprogs-libs-1.39-33.1.ali5
e2fsprogs-devel-1.39-33.1.ali5
```

If the current version is earlier than 1.42, update the software by following these steps.

2. Run the following command to download e2fsprogs in version 1.42.8. Go to [e2fsprogs](#) to find the latest software package.

```
wget https://www.kernel.org/pub/linux/kernel/people/tytso/e2fsprogs/v1.42.8/e2fsprogs-1.42.8.tar.gz
```

3. Run the following commands in turn to compile tools in later versions.

```
tar xvzf e2fsprogs-1.42.8.tar.gz
cd e2fsprogs-1.42.8
./configure
make
make install
```

4. Run `rpm -qa | grep e2fsprogs` to check whether the software of the later version has been installed successfully.

Appendix 2: Disable lazy init function

The lazy init function of ext4 file system is enabled by default. While the function is enabled, it will run in the system background and initiate a thread to initialize metadata of ext4 file system continuously to delay metadata initialization. Therefore, immediately after formatting a data disk, IOPS can be affected.

If you need to test performance of data disk immediately after formatting, run the following commands to disable lazy init function while formatting the file system.

```
mke2fs -O 64bit,has_journal,extents,huge_file,flex_bg,uninit_bg,
dir_nlink,extra_isize -E lazy_itable_init=0,lazy_journal_init=0 /dev
/vdd1
```

If the lazy init is disabled, it may take longer time to format a partition. For example, it may take 10–30 minutes to format a 32 TiB data disk.

6.5 Detach a cloud disk

If a Pay-As-You-Go cloud disk is attached to an ECS instance as a data disk, you can detach it from the instance and release it. However, if the disk is used as a system disk, you cannot detach it.

When detaching a cloud disk, consider the following:

- Only the Pay-As-You-Go cloud disks in the **In Use** status and used as a **Data Disk** can be detached.
- You cannot detach a local disk.
- For a Windows instance:

- To guarantee data integrity, we recommend that you stop writing or reading the files on the cloud disk. Otherwise, data may be lost.
- Before detaching a cloud disk in the ECS console, you must [connect to the instance](#) and set its status as **Offline** in **Disk Management**.
- For a Linux instance:
 - Before detaching a cloud disk in the ECS console, you must [connect to the instance](#) and run `umount` to unmount the partitions.
 - If you have configured the `/etc/fstab` file to automatically mount the partitions at the startup of the instance, before detaching it, you must delete the configurations from the `/etc/fstab` file. Otherwise, you cannot connect to the instance after the instance is restarted.

The following table describes the actions available for you to detach a cloud disk in the ECS console.

Scenario	Action
You want to detach one or more cloud disks from one instance.	Detach cloud disks on the Instance Disk page.
You want to detach one specific cloud disk.	Detach a cloud disks on the Disk List page.

Detach cloud disks on the Instance Disk page

On the **Instance Disks** page, you can delete one or more cloud disks that are attached to the instance.

Prerequisites

The cloud disks have been [attached to the instance](#) and its status is **In Use**.

If you are detaching a cloud disk from a Linux instance, and you have configured the `/etc/fstab` file to mount the partitions at the startup of the instance, you must first delete the configurations.

Procedure

To detach a cloud disk from the **Instance Disks** page, follow these steps:

1. Connect to the instance and unmount its partitions. According to the operating system, follow the recommended steps detailed in the following table.

Operating system	Steps
Linux	Run <code>umount [partition]</code> . For example, <code>umount /dev/vdb1</code> .

Operating system	Steps
Windows	Start Disk Management , right-click the disk name (for example, Disk 2) and then click Offline .

2. Log on to the [ECS console](#).
3. In the left-side navigation pane, click **Instances**.
4. Select the target region.
5. Find the target instance and click its ID to go to its Instance Details page.
6. In the left-side navigation pane, click **Disks**.
7. Find the target cloud disk and then, in the **Actions** column, select **More > Unmount**.

Only cloud disks that have the following attributes can be detached:

- **Status** must be **In Use**.
 - **Unmountable** must be **Yes**.
 - **Type** must be **Data Disk**.
8. In the dialog box, click **Confirm**.
 9. Optional. If you want to detach multiple cloud disks, repeat steps 7 and 8 as required.

When the status of the cloud disk becomes **Unmounted**, the disk is detached.

Detach a cloud disks on the Diskspage

On the Disk List page, you can detach a specific cloud disk from an ECS instance.

Prerequisites

The cloud disk has been [attached to the instance](#) and are in the **In Use** status.

If you are detaching a cloud disk from a Linux instance, and you have configured the `/etc/fstab` file to mount the partitions at the startup of the instance, delete the configurations.

Procedure

To detach a cloud disk on the Disk List page, follow these steps:

1. Connect to the instance and unmount the partitions. According to the operating system, follow the recommended steps detailed in the following table.

Operating system	Steps
Linux	Run <code>umount [partition]</code> . For example, <code>umount /dev/vdb1</code> .

Operating system	Steps
Windows	Start Disk Management , right-click the disk name (for example, Disk 2) and then click Offline .

2. Log on to the [ECS console](#).
3. In the left-side navigation pane, select **Block Storage > Disks**.
4. Select the target region.
5. Find the target cloud disk and then, in the **Actions** column, select **More > Unmount**.

Only cloud disks that have the following attributes can be detached:

- **Status** must be **In Use**.
- **Unmountable** must be **Yes**.
- **Type** must be **Data Disk**.

6. In the dialog box, click **Confirm**.

When the status of the cloud disk becomes **Unmounted**, the disk is detached.

Related API

[DetachDisk](#)

Additional operation

If you no longer need the disk, you can [release it](#).

6.6 Resize cloud disks

6.6.1 Overview

Depending on the disk type, you can resize a disk as follows:

- For a system disk: **Change System Disk**
- For a data disk: **Resize Disk**

Limitations

Limitations of resizing disks vary between system disks and data disks.

System disks

The **Change system disk** feature allows you to increase the disk size only. The size limit for disk resizing is determined by the image and the current size of the system disk, as displayed in the following table.

Image	Size limit (GiB)
Linux (excluding CoreOS) and FreeBSD	20-500
CoreOS	30-500
Windows	40-500

Data disk

The **Resize Disk** feature allows you to increase the disk size only. The following table lists the capacity limits of different data disk typics after resizing, which is determined by the cloud disk types.

Cloud disk type	Current capacity	Capacity after resizing
Basic Cloud Disk	Any	2,000 GiB
SSD Cloud Disk or Ultra Cloud Disk	equal or less than 2,048 GiB	2,048 GiB
SSD Cloud Disk or Ultra Cloud Disk	> 2,048 GiB	Cannot be resized
ESSD Cloud Disk	Any	32,768 GiB

Additional operations

- To increase the size of the system disk of an ECS instance, see [increase system disk size](#).
- To resize a data disk attached to a Windows instance, see [Windows _ Resize a data disk](#).
- To resize a data disk attached to a Linux instance, see [Linux _ Resize a data disk](#).

6.6.2 Increase system disk size

You can increase the size of the system disk of your ECS instance by using **Change System Disk** feature. This article describes how to increase the size of a system disk while keeping the operating system and environment intact.



Note:

You can change the operating system while increasing the size of a system disk. For more information, see [change the operating system](#).

Notes

Before you begin, consider the following.

Risks

Risks that may occur when you replace a system disk are as follows:

- If you attempt to replace the system disk while the instance is running, your business services may be disrupted. We recommend you stop your instance before replacing the system disk.
- After disk replacement, you must redeploy the business runtime environment on the new system disk. This may result in a long period of disruption to your business services.
- After the system disk is changed, a new cloud disk with a new disk ID is assigned, and the old disk ID is released. Therefore, you cannot roll back the system disk by using any snapshot of the released cloud disk.

**Note:**

After the system disk is changed, you can still use manually created snapshots of the released disk to create custom images. If you have applied an automatic snapshot policy to the old system disk and set the automatic snapshots to release when the disk is released, you must apply the policy to the new disk. Furthermore, all the automatic snapshots of the old disk are released.

Limits and recommendations

When changing the system disk, you must consider the following:

- After the system disk is changed, your instance is assigned a new cloud disk as the system disk, with a new disk ID, and the old one is released.
- You cannot replace the Cloud Type of the system disk.
- You cannot reduce the capacity of a system disk. You can only increase it. The maximum capacity of a system disk is 500 GiB.
- You cannot increase the size of the system disk that runs Windows 2003.
- If your Subscription instance has been [renewed for configuration downgrade](#), you cannot modify the system disk capacity until you enter the next billing cycle.
- The IP address and the MAC address remain unchanged after the system disk is changed.
- We recommend that you create a snapshot for the system disk before you change the disk.

When creating a snapshot, consider the following:

- We recommend that you create snapshots during off-peak business hours as it may take an extended amount of time to complete. For example, it takes about 40 minutes to create a snapshot of 40 GiB. Creation of a snapshot may also reduce the I/O performance of a block storage device.

- Make sure the system disk has enough available storage space when creating a snapshot (at least 1 GiB). Otherwise, the system may fail to start after the system disk is changed.
- To make sure you have enough quota for automatic snapshots of the new system disk, delete any unnecessary snapshots of the old system disk. For more information, see [delete snapshots or automatic snapshot policies](#).

Procedure

If you want to increase the size of the system disk while keeping the operating system and environment intact, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Instances**.
3. Select the target region.
4. Find the instance for which to change the system disk and click its instance ID to go to the **Instance Details** page.
5. Follow these steps to create a snapshot of the system disk:
 - a. In the left-side navigation pane, click **Disks**.
 - b. Locate the required system disk and then, in the **Actions** column, click **Create Snapshot**.



Note:

For more information about the limits or note for creating a snapshot, see [create snapshots](#).

6. Follow these steps to create a custom image by using the snapshot:
 - a. In the left-side navigation pane, click **Instance Snapshots** to check the creation status and progress. When the progress is 100% and the status is **Success**, go to the **Actions** column and click **Create Custom Image**.



Note:

- For more information about the limitations of creating a custom image, see [create a custom mirror using a snapshot](#).
- The custom image is displayed in the dropdown list of the **Custom Image** on the Replace System Disk page.

- b. Go back to the **Instances** page and then, in the left-side navigation pane, select **Snapshots and Images > Image** to check the creation status and progress of the custom image.

- When the progress is 100% and the status is **Available**, in the left-side navigation pane, click **Instances**.
- In the Instances list, find the instance, and in the **Actions** column, select **More > Instance Status > Stop**.

**Note:**

For a Pay-As-You-Go VPC-Connected ECS instance, if the [No fees for stopped instances \(VPC-Connected\)](#) feature is enabled, in the **Notes** dialog box, click **OK**. Then, in the **Stop** dialog box, select **Keep Stopped Instances and Continue Billing**, and click **OK**. If you use the **No Fees for Stopped Instances (VPC-Connected)**, you may not be able to start the instance successfully after changing the system disk.

Stop Stop instance

Operation will be executed on the selected 1 Instances . Are you sure you want to proceed?

Stopped By: Stop
 Force Stop

Stopped By: Keep Stopped Instances and Continue Billing

Stopping Subscription based instances does not change their expiration date.
If you need to stop an instance for system disk replacement, disk reinitialization, instance upgrade, or private IP address modification, we recommend that you select Keep Stopped Instances and Continue Billing to avoid startup failure.

OK Cancel

- When the instance is in the **Stopped** status, go to the **Actions** column and select **More > Disk and Image > Replace System Disk**.
- In the pop-up dialog box, read and confirm you agree to the notice by clicking **OK**.
- On the **Replace System Disk** page, complete the configurations as follows:
 - Image Type**: Click the **Custom Image** tab and select the created custom image in the drop-down list.

b. Security enhancement:

- **System Disk:** Specify a new size for the system disk according to your business needs. The maximum size is 500 GiB. The size limit for changing is determined by the image and the current size of the system disk, as displayed in the following table.

Image	Limit for capacity expansion (GiB)
Linux (excluding CoreOS) and FreeBSD	20-500
CoreOS	30-500
Windows	40-500

**Note:**

You cannot modify the Cloud Type of the system tray.

- If a Windows image is used, set a logon password.
 - If a Linux image is used and the instance is I/O optimized, you can choose to set a password or bind an SSH key pair for logon.
- c.** Confirm the **Instance Cost**, which includes the price of the mirror and the price of the system disk. For more information, see [cloud product price](#).
- d.** Read and confirm you agree to the ECS Service Terms and Product Terms of Service, check the box, and then click **Confirm to change**.

Go back to the ECS console to check the status of the process. It may take a few minutes to process the change. After the system disk is changed, the instance starts automatically.

Follow-up operations

After the system disk is changed, you may have to perform the following:

- If your instance is running a Linux image, and a data disk was attached to the instance and set to automatically mount the file systems at the beginning, the mount information is now lost while changing the system disk. Therefore, you must write the new partition and mounting information to the `/etc/fstab` file on the new system disk and mount the file systems. You must not partition or format the data disk again. For more information about the commands, see [Linux _ Format and mount a data disk](#). Follow these steps to mount the file systems:
 1. (Optional) Create a backup of `/etc/fstab` file.
 2. Write the new partition and mounting information to the `/etc/fstab` file.
 3. Check the new partition information in the `/etc/fstab` file.

4. Mount the file systems.
5. To view disk space and usage: run the command `df -h`.

After mounting, you do not need to restart the instance to use the new file system directly.

- [Apply automatic snapshot policies to disks](#). Note that the link between an automatic snapshot policy ID and a disk ID is broken after the system disk is changed. You must set up an automatic snapshot policy for the new system disk.

6.6.3 Windows _ Resize a data disk

As your business grows, the current capacity of your data disks may not be able to meet your data storage needs. You can use the **Resize Disk** function to resize your data disks as necessary.



Note:

- We recommend that you manually create a snapshot to back up your data before resizing a data disk.
- You can resize a data disk when the data disk is either in the **Available** status or in the **In Use** status.
- If a snapshot is being created for a data disk, you cannot resize the data disk.
- If you have renewed a Subscription ECS instance for configuration downgrade ([renew for configuration downgrade](#)) during its current billing cycle, you cannot resize the attached Subscription cloud disks, including its data or system disks.
- You can resize data disks, but not file system.
- You can resize data disks, but not system disks or local disks.
- Resize the data disks that are attached to the instance only when the instance is in the **Running** (Running) or **Stopped** (Stopped) status. The changes are applied when you restart the instance in the ECS console. This action stops your instance from working and interrupts your business. Hence, proceed with caution.

This example uses a data disk of the ultra cloud disk type and an ECS instance running 64-bit Windows Server 2008 R2 Enterprise Edition to show how to resize a data disk and extend the available capacity. In this example, the current disk capacity is 20 GiB, and we resize it to 24 GiB.

To resize a data disk, follow these steps:

[Step 1. Resize a data disk in the ECS console](#)

[Step 2. Log on to the instance to enable the extended storage space](#)

Step 1. Resize a data disk in the ECS console

To resize a data disk in the ECS console, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Block Storage > Disks**.



Note:

If the data disk you want to resize is attached to an instance, click **Instances** in the left-side navigation pane, find the instance, go to the Instance Details page, and then click **Disks**.

3. Select a region.
4. Find the disk to be resized, and in the **Actions** column, select **More > Resize Disk**.
5. On the **Resize Disk** page, set **Capacity after resizing**. In this example, 24 GiB. The capacity after resizing must be larger than the current capacity.
6. When the cost is calculated, click **Confirm to resize**.



Note:

If your data disk is attached to an instance, [restart the instance](#) in the ECS console to make the disk resize take effect.

Once the data disk resizing completes, you can do the following:

- If the data disk is attached to an instance, [log on to the instance to enable the extended storage space](#).
- If the data disk is not attached to an instance, attach the disk to an instance in the console first, and then proceed depending on the data disk:
 - If it is not formatted or partitioned, format and mount the data disk. For more information, see [format a data disk for Windows instances](#).
 - If it is formatted and partitioned, [log on to the instance to enable the extended storage space](#).

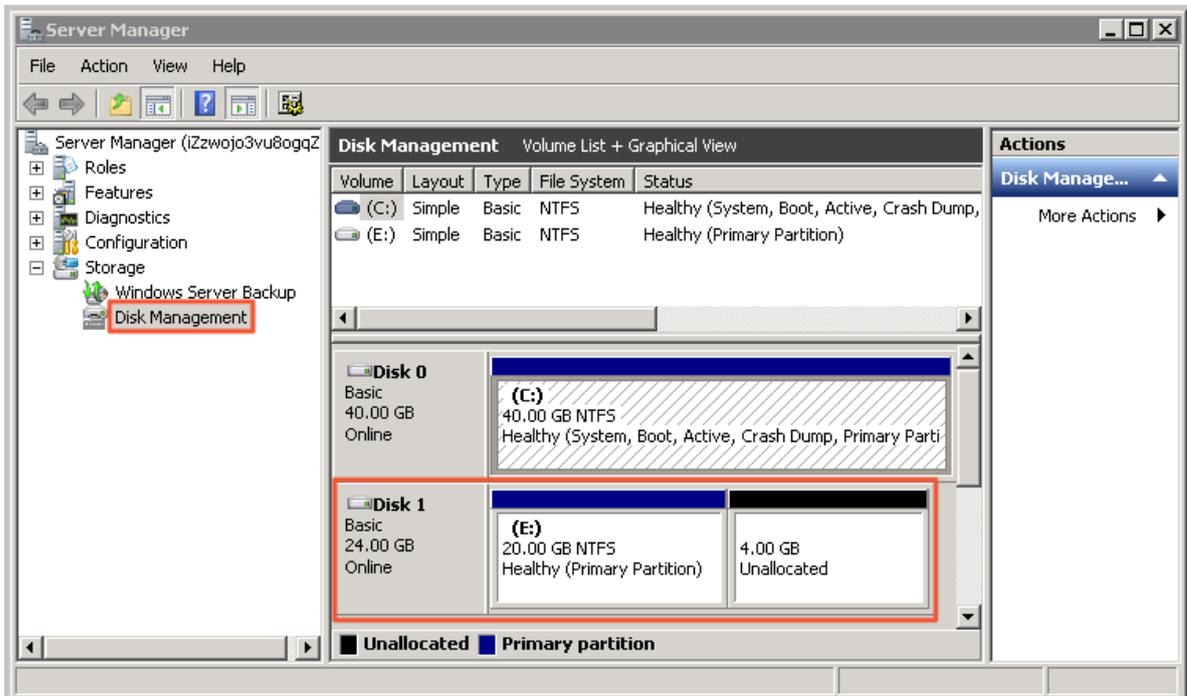
Step 2. Log on to the instance to enable the extended storage space

To resize a data disk within the instance, follow these steps:

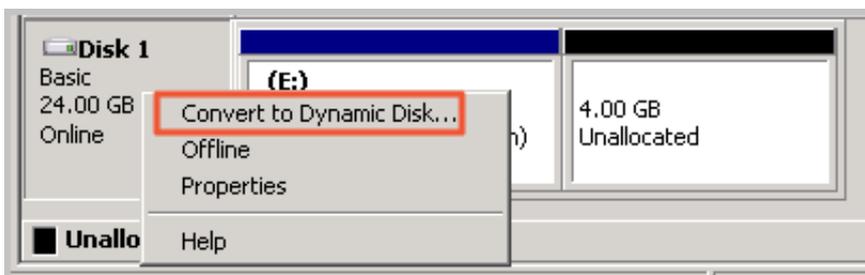
1. [Connect to a Windows instance](#).
2. On the Windows Server desktop, click the Server Manager icon



3. In the left-side navigation pane of **Server Manager**, select **Storage > Disk Management**. In the disk management area, you can see the relationship between the new and the original data disk spaces. In this example, **Disk 1** is the resized data disk.



4. Right click **Disk 1**, select **Convert to Dynamic Disk**, and follow the wizard to convert a basic disk to a dynamic disk.



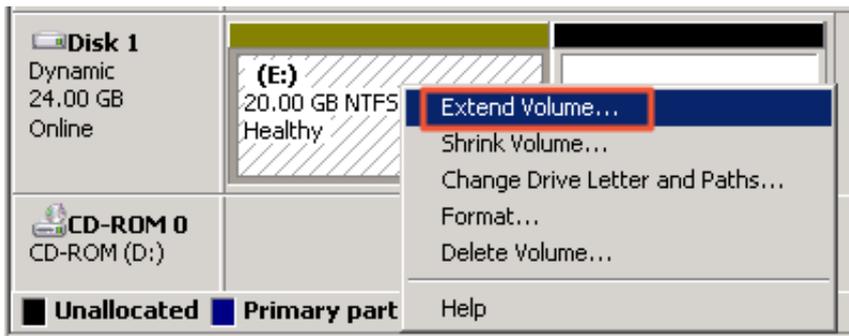
Note:

Converting a basic disk into a dynamic disk unmounts the disk from the system. Applications installed on the data disk, if any, are temporarily unavailable during the conversion process. The conversion process does not cause any data loss.

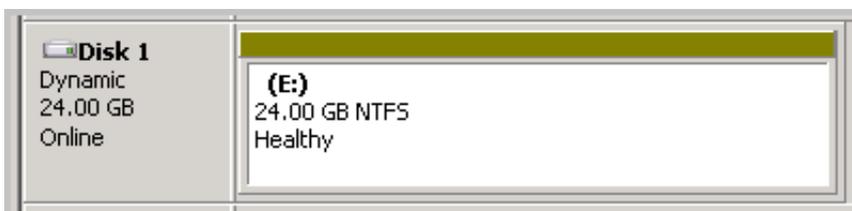
After the conversion, the **Disk 1** information shows in the Disk Manager as follows.



- Right click an empty area of the simple volume of Disk 1, and select **Extend Volume**.



- Follow the **Extend Volume Wizard** to extend the volume. When the wizard is complete, the new data disk space is automatically merged into the original volume and the **Disk 1** information showed in the Disk Manager as follows.



Note:

On Windows Server 2003, the extended storage space is added to the data disk but it is displayed as a separate volume in Disk Manager. On Windows Server 2008, one separate volume is created for each expansion and is not merged into the original volume, which does not affect the availability of the extended storage space.

You have resized a data disk successfully and the extended storage space is ready for use.

6.6.4 Linux _ Resize a data disk

As your business grows, the current capacity of your data disks may not be able to meet your data storage needs. You can use the **Resize Disk** feature to resize your data disks as necessary.



Note:

- Resize the data disks that are attached to an instance only when the instance is in the **Running** or **Stopped** status. You must restart the instance in the ECS console to apply the changes. This action causes your instance to stop working and may cause your business to be interrupted, so please proceed with caution.
- We recommend that you manually create a snapshot to back up your data before resizing your data disk.
- You can resize a data disk when the data disk is either in the **Available** status or in the **In Use** status.
- If you have renewed a Subscription ECS instance for configuration downgrade ([renew for configuration downgrade](#)), during its current billing cycle, you cannot resize the attached Subscription cloud disks, including its data or system disks.
- If a snapshot is being created for a data disk, you cannot resize the data disk.
- You can resize data disks, but not system disks or local disks.

This example uses a data disk of the ultra cloud disk type and an ECS instance running 64-bit CentOS 7.3 to describe how to resize data disk and extend the available capacity.

To resize a data disk, follow these steps:

[Step 1. Increase the size of a data disk in the ECS console](#)

[Step 2. Log on to the instance to resize the file system](#)

Step 1. Increase the size of a data disk in the ECS console

To increase the size of a data disk in the ECS console, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Block Storage > Disks**.



Note:

If the data disk you want to resize has been attached to an instance, in the left-side navigation pane, click **Instances**, find the corresponding instance, go to the instance details page, and click **Disks**.

3. Select a region.
4. Find the disk to be resized, and in the **Actions** column, select **More > Resize Disk..**
5. On the **Resize Disk** page, set **Capacity after resizing** (in this example, 30 GiB). The capacity after resizing must be larger than the current capacity.
6. When the cost is calculated, click **Confirm to resize**.

**Note:**

After the resizing, you can view the new disk size in the console. However, if the data disk is attached to an ECS instance, you must *restart the instance* in the ECS console to view the new disk size when you log on to the instance.

After the disk size is increased,

- If the data disk is attached to an instance, *log on to the instance to resize the file system*.
- If the data disk is not attached to an instance, attach the disk to an instance in the console (*attach a cloud disk*) first, and then proceed depending on the data disk:
 - If it is a new data disk, which has not been formatted, format it. For more information, see *format a data disk for Linux instances*.
 - If it has been formatted and partitioned, *log on to the instance to resize the file system*.

Step 2. Log on to the instance to resize the file system

After the disk size is increased, you must log on to the instance to resize the file system.

In this example, the data disk is attached to a Linux instance running the 64-bit CentOS 7.3. The data disk before resizing has only one primary partition (/dev/vdb1, ext4 file system), the mount point of the file system is */resizetest*, and after resizing is completed, the data disk still has only one primary partition.

1. *Connect to a Linux instance by using a password*.
2. Run the `umount [file system name]` command to unmount the primary partition.

```
umount /dev/vdb1
```

**Note:**

Run the `df -h` command to check whether the unmounting is successful. If you do not see the `/dev/vdb1` information, unmounting is successful. The following is the sample output.

```
[root@iXXXXXXX ~]# df -h
Filesystem Size Used Avail Use% Mounted on
/dev/vda1 40G 1.5G 36G 4% /
devtmpfs 487M 0 487M 0% /dev
tmpfs 497M 0 497M 0% /dev/shm
tmpfs 497M 312K 496M 1% /run
tmpfs 497M 0 497M 0% /sys/fs/cgroup
tmpfs 100M 0 100M 0% /run/user/0
```

3. Run the `fdisk` command to delete the original partition and create a new partition:

**Note:**

If you use the `parted` tool to manipulate partitions, you cannot use it in conjunction with `fdisk`. Otherwise, this results in an inconsistent first sector of the partition. Instructions on how to use the `parted` tool can be found [here](#).

- a. Run the `fdisk -l` command to list the partition details and record the final size of the partition and its first sector before resizing.
- b. Run the `fdisk [device name of data disk]` command to go to `fdisk`. In this example, the device name is `/dev/vdb`.
- c. Type `d` and press the Enter key to delete the original partition.

**Note:**

Deleting a partition does not cause loss of data in the data disk.

- d. Type `d` and press the Enter key to start creating a new partition.
- e. Type `p` and press the Enter key to create a primary partition. In this example, you are creating a single-partition data disk, so it is sufficient to create one primary partition.

**Note:**

If you want to create more than four partitions, create at least one extended partition, that is, type `e`.

- f. Type the partition number and press the Enter key. In this example, only one partition is to be created, so type `1`.
- g. Type a number for the First sector: For data consistency, the number for the First sector must be identical with that of the original partition. In this example, press the Enter key to use the default value of `1`.

**Note:**

If you find that the First sector is not identical with the recorded one, you may have used the `parted` tool for partitioning. In that case, stop the current `fdisk` operation and use [parted](#) to start over again.

- h. Type a number for the last sector: Because only one partition is to be created in this example, press the Enter key to use the default value.
- i. Type `wq` and press the Enter key to start partitioning.

```
[root@iXXXXXX ~]# fdisk /dev/vdb
```

```

Welcome to fdisk (util-linux 2.23.2).
Changes will remain in memory only, until you decide to write them
.
Be careful before using the write command.
Command (m for help): d
Selected partition 1
Partition 1 is deleted
Command (m for help): n
Partition type:
p primary (0 primary, 0 extended, 4 free)
e extended
Select (default p):
Using default response p
Partition number (1-4, default 1):
First sector (2048-62914559, default 2048):
Using default value 2048
Last sector, +sectors or +size{K,M,G} (2048-62914559, default
62914559):
Using default value 62914559
Partition 1 of type Linux and of size 30 GiB is set
Command (m for help): wq
The partition table has been altered!
Calling ioctl() to re-read partition table.
Syncing disks.

```

**Note:**

If you are using the `parted` tool, type `p` in the `parted` window to list the current partition details. If any partition is displayed, use `rm + serial number` to delete the original partition table, then run the `unit s` command to specify the start unit, calculated by the number of sectors, and finally run the `mkpart` command to create it, as shown in the following figure.

```

[root@ ~]# parted /dev/xvdb
GNU Parted 3.1
Using /dev/xvdb
Welcome to GNU Parted! Type 'help' to view a list of commands.
(parted) p
Model: Xen Virtual Block Device (xvd)
Disk /dev/xvdb: 5369MB
Sector size (logical/physical): 512B/512B
Partition Table: gpt
Disk Flags:

Number  Start  End  Size  File system  Name  Flags

(parted) unit s
(parted) mkpart primary ext3 56 5369MB
Warning: The resulting partition is not properly aligned for best performance.
Ignore/Cancel? i
(parted) p
Model: Xen Virtual Block Device (xvd)
Disk /dev/xvdb: 10485760s
Sector size (logical/physical): 512B/512B
Partition Table: gpt
Disk Flags:

Number  Start  End  Size  File system  Name  Flags
1       56s    10485726s  10485671s  ext3        primary

```

4. For some operating systems, the file system may be automatically mounted to the mount point after partitioning. We recommend that you run the `df -h` command to check the file system space and usage. Run the `umount [file system name]` to unmount the file system again.
5. Check the file system and resize the file system.

```
e2fsck -f /dev/vdb1 # check the file system
resize2fs /dev/vdb1 # resize the file system
```

**Note:**

- Running the `e2fsck` command is time-consuming because the system needs to check and revise the file system metadata during that process, so be patient.
- Properly running the `e2fsck` command and the `resize2fs` command does not cause data loss.

The following is the sample output.

```
[root@iXXXXXX ~]# e2fsck -f /dev/vdb1
e2fsck 1.42.9 (28-Dec-2013)
Pass 1: Checking inodes, blocks, and sizes
Pass 2: Checking directory structure
Pass 3: Checking directory connectivity
Pass 4: Checking reference counts
Pass 5: Checking group summary information
/dev/vdb1: 11/1835008 files (0.0% non-contiguous), 159218/7339776
blocks
[root@iXXXXXX ~]# resize2fs /dev/vdb1
resize2fs 1.42.9 (28-Dec-2013)
Resizing the filesystem on /dev/vdb1 to 7864064 (4k) blocks.
The filesystem on /dev/vdb1 is now 7864064 blocks long.
```

6. Mount the resized file system to the original mount point (in this example, `/resizetest`).

```
mount /dev/vdb1 /resizetest
```

7. Run the `df -h` command to check file system space and usage. If the correct information about the resized file system is displayed, the mounting is successful and the resized file system is ready for use.

**Note:**

After the mounting is completed, you can use the resized file system without restarting the instance.

The following is the sample output.

```
[root@iXXXXXX ~]# df -h
```

```
Filesystem Size Used Avail Use% Mounted on
/dev/vda1 40G 1.5G 36G 4% /
devtmpfs 487M 0 487M 0% /dev
tmpfs 497M 0 497M 0% /dev/shm
tmpfs 497M 312K 496M 1% /run
tmpfs 497M 0 497M 0% /sys/fs/cgroup
tmpfs 100M 0 100M 0% /run/user/0
/dev/vdb1 30G 44M 28G 1% /resizetest
```

6.7 Reinitialize a cloud disk

When a cloud disk is attached to an ECS instance, you can **reinitialize the disk** to restore the system disk or the data disks to the status when they were created. After a cloud disk is reinitialized:

- The system disk is restored to the initial status when it was created. For example, if you select a public image to create an ECS instance, after the system disk is reinitialized, the operating system is retained, but all other applications that were installed after the instance creation are deleted.



Note:

After you change the operating system or resize the system disk, the instance is not fully restored to the status at which it was created, but only to the status of the new system disk when it was created.

- Depending on how the data disk was created, it is restored to the following initial status:
 - Restored to an empty disk if it was an empty disk
 - Restored to a disk with only the data of the source snapshot if it was *created from a snapshot*
- If an automatic snapshot policy is applied to a cloud disk, the policy is retained and does not need to be applied again after reinitialization.
- If an automatic snapshot policy is applied to a cloud disk, the policy is retained and does not need to be applied again after reinitialization.
- After a cloud disk is reinitialized, all the snapshots, both automatically and manually created, are retained. You can use them to *roll back a cloud disk*.



Warning:

- Because you must stop your ECS instance to reinitialize a cloud disk, your business services may be disrupted. Exercise caution when performing this action

- After a cloud disk is reinitialized, its data is lost. Therefore, we recommend you back up the data. To do so, you can [create snapshots](#).

Reinitialize a system disk

Prerequisites

If an SSH key pair is used as the authentication method, check that you have [created an SSH key pair](#) or [import ed an SSH key pair](#).

Procedure

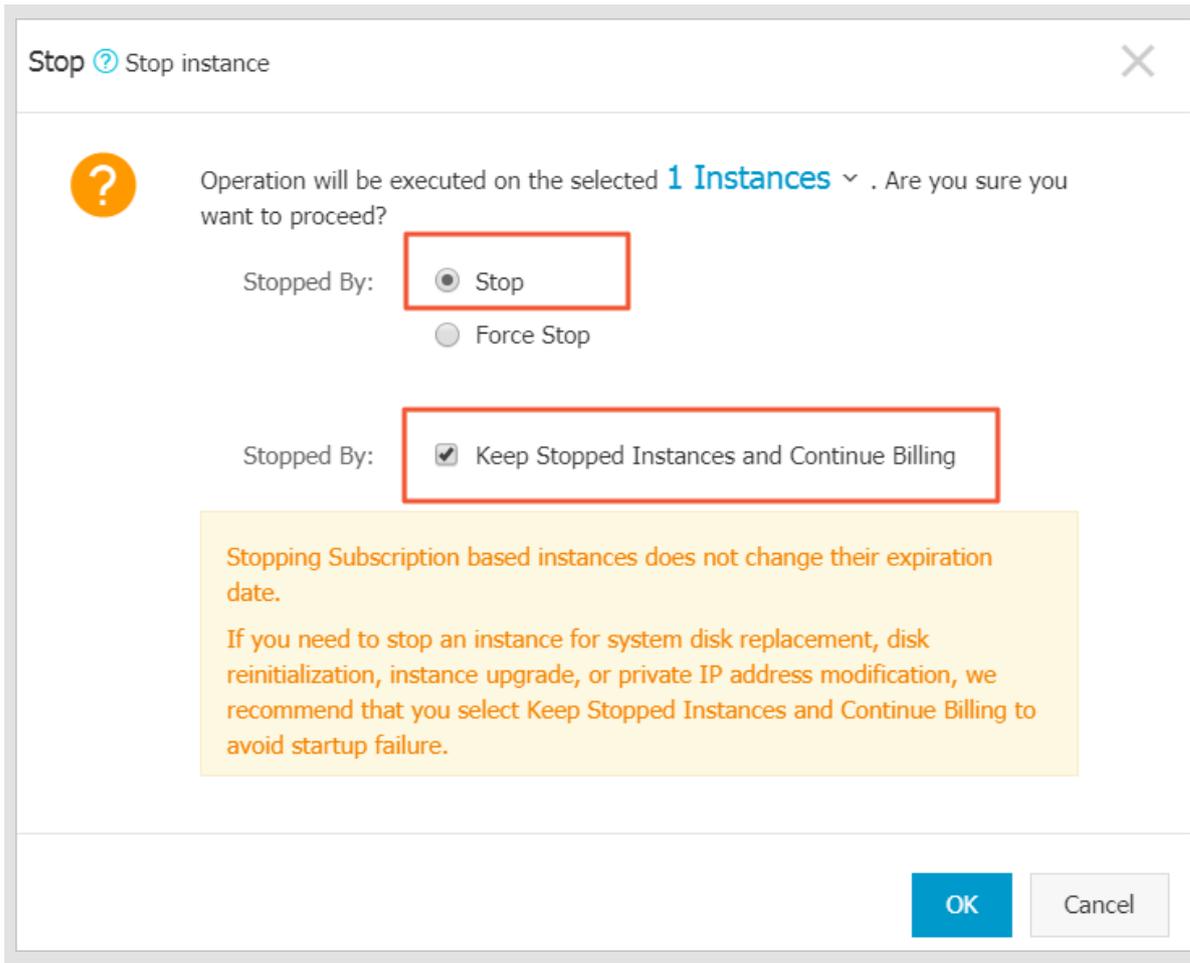
To reinitialize a system disk, follow these steps:

1. Log on to the [ECS console](#).
2. Select the target region.
3. In the left-side navigation pane, click **Instances**.
4. Find the target ECS instance and click its ID to go to its **Instance Details** page.
5. Click **Stop**.



Note:

For a Pay-As-You-Go VPC-Connected ECS instance, if the [No fees for stopped VPC instances](#) feature is enabled, in the **Notice** dialog box, click **OK**, and then in the **Stop** dialog box, select **Keep Instance with Fees**. If you select the **No Fees for Stopped Instances (VPC-Connected)** mode, you may not be able to start the instance successfully after you reinitialize the system disk.



6. After the instance is **stopped**, in the left-side navigation pane, click **Disks**.
7. Find the system disk and then, in the **Actions** column, click **Reinitialize Disk**.
8. In the **Reinitialize Disk** dialog box, complete the following configuration:
 - a. Authentication method:
 - For a Windows instance, you must specify a logon password. You can either use a previous password or specify a new one.

Reinitialize Disk ✕

Are you sure you want to reinitialize the following disks:

System Disk: ██████████

Security: Set SSH Key Set Password

*Logon Password:

The password can be 8 to 30 characters in length and must contain three types of the following characters: Uppercase letters, lowercase letters, numbers, and special characters. Special characters include parentheses (()), graves (`), tildes (~), exclamation points (!), at signs (@), number signs (#), dollar signs (\$), percent signs (%), carets (^), ampersands (&), asterisks (*), hyphens (-), plus signs (+), equal signs (=), vertical bars (|), curly braces ({ }), braces ([]), colons (:), semicolons (;), apostrophes ('), angle brackets (< >), commas (,), periods (.), question marks (?), and forward slashes (/).

*Confirm Password:

Security Enhancement: Activate i

Instance Startup Policy: Start Instance after Resetting Disk

Note: 1. After reinitialization, the selected system disk will be restored to its image.
2. You must reset the password for logging on to the instance when reinitializing the disk.

Note: Your automatic snapshot policy will become invalid soon and must be reconfigured. Alibaba Cloud is not responsible for any data loss caused by failure to reconfigure the policy.

Confirm Cancel

- For a Linux instance, select **Set SSH Key** or **Set Password** as the security setting. If Key Pair is selected, bind a key pair. If Password is selected, specify a logon password.

Reinitialize Disk
✕

Are you sure you want to reinitialize the following disks:

System Disk: XXXXXXXXXX

Security: Set SSH Key Set Password

SSH Key Pair:

Security Enhancement: Activate i

Instance Startup Policy: Start Instance after Resetting Disk

Note: 1. After reinitialization, the selected system disk will be restored to its image.
2. You must reset the password for logging on to the instance when reinitializing the disk.

Note: Your automatic snapshot policy will become invalid soon and must be reconfigured. Alibaba Cloud is not responsible for any data loss caused by failure to reconfigure the policy.

Confirm
Cancel

- b. (Optional) **Security Enhancement:** Select **Activate**. After the security enhancement feature is enabled, ECS security components are loaded. These components provide security features such as backdoor detection, remote logon reminders, brute-force cracking prevention mechanisms, and more.
 - c. (Optional) **Instance Startup:** Select **Start Instance Resetting Disk**.
 - d. Click **Confirm**.
9. For Linux instances: If you have attached a data disk to the instance, connect to the instance and [create a mounting point for the partitions of data disks](#), because the mounting points are lost after the system disk is reinitialized.



Note:

For a Windows instance, both the system disk and the data disks are ready for use. No additional operations are needed.

After the system disk is reinitialized, you must deploy all applications to restore your business operations.

Reinitialize a data disk

Once reinitialized, a data disk is in a different status according to its original status and the operating system of the instance:

- For a Windows instance, the data disk is ready to use without any additional operations required.
- For a Linux instance:
 - If the data disk was an empty disk after it was created, then all the data and partitions on the disk are lost. You must partition and format the disk, and mount the partitions again.



Note:

If you configured the `/etc/fstab` file to automatically mount the disk partitions at startup of the instance, you must comment out the lines from the `/etc/fstab` file before reinitializing a data disk. Otherwise, your instance will fail to start.

- If the data disk was created from a snapshot, then the data disk is recovered to the point in time at which the snapshot was generated. You do not have to mount the partitions again, but all the data generated after the disk creation is lost.

In this section, `/dev/vdb1` is the example partition and `/lnitTest` is the example mounting point.

Replace these details with your actual information.

Prerequisites

The data disk to be reinitialized must be attached to an ECS instance. For more information, see [attach a cloud disk](#).

Procedure

To reinitialize a data disk, follow these steps:

1. For Linux instances: If the data disk was an empty disk after it was created, and the mounting configuration was added to the `/etc/fstab` file, you must comment out the mounting configuration from the `/etc/fstab` file. To do so, follow these steps:
 - a. [Connect to the Linux instance](#).
 - b. Run `vim /etc/fstab`.
 - c. Press the `i` key to enter the Insert mode.

- d. Locate the mounting configuration lines and type # before the lines. For example:

```
# /dev/vdb1 /InitTest ext3 defaults 0 0
```

- e. Press the `ESC` key to exit the Insert mode, and then run `:wq` to save and exit.

2. Log on to the [ECS console](#).
3. In the left-side navigation pane, click **Instances**.
4. Select the target region.
5. Find the target ECS instance and click its ID to go to its **Instance Details** page.
6. Click **Stop**.



Note:

For a Pay-As-You-Go VPC-Connected ECS instance, if the [No fees for stopped VPC instances](#) feature is enabled, in the **Notice** dialog box, click **OK**, and then in the **Stop** dialog box, select **Keep Instance with Fees**. If you select the **No Fees for Stopped Instances (VPC-Connected)** mode, you may not be able to start the instance successfully after you reinitialize the system disk.

Stop ⓘ Stop instance
✕

?

Operation will be executed on the selected **1 Instances** . Are you sure you want to proceed?

Stopped By: Stop

Force Stop

Stopped By: Keep Stopped Instances and Continue Billing

Stopping Subscription based instances does not change their expiration date.

If you need to stop an instance for system disk replacement, disk reinitialization, instance upgrade, or private IP address modification, we recommend that you select **Keep Stopped Instances and Continue Billing** to avoid startup failure.

OK

Cancel

7. After the instance is **stopped**, in the left-side navigation pane, click **Disks**.
8. Find the target data disk and in the **Actions** column, click **Reinitialize Disk**.
9. In the **Reinitialize Disk** dialog box, read the notes and click **Confirm**.
10. In the left-side navigation pane, click **Instance Details**.
11. Click **Start**.
12. For Linux instances: If the data disk was an empty disk after it was created, [format and mount data disks for Linux instances](#).

After the data disk is reinitialized, you may need to deploy applications to restore your business operations.

API

[ReInitDisk](#)

6.8 Roll back a cloud disk

If you have [created snapshots](#) for a cloud disk, you can use the **Disk Rollback** feature to restore a cloud disk to a specific snapshot status at a given time point.

Note

Before you roll back a cloud disk, note the following:

- Rolling back a cloud disk is an irreversible action. Once rollback is complete, data cannot be restored. Exercise caution when performing this action.
- After the disk is rolled back, data from the creation date of the snapshot to the rollback date is lost.
- After a system disk is restored, the logon password or the SSH key pair of the ECS instance is retained.

Prerequisites

Before rolling back a cloud disk, check that:

- You have [created a snapshot](#) for the cloud disk, and no snapshot creation is in progress.
- You have not released the cloud disk.
- The cloud disk has been [attached to an ECS instance](#) and the instance is in the [Stopped](#) status.



Note:

For a Pay-As-You-Go VPC-Connected ECS instance, if the [No fees for stopped instances \(VPC-Connected\)](#) feature is enabled, to stop an instance, in the **Notice** dialog box, click **OK**. Then in the **Stop** dialog box, select **Keep Instance with Fees**, and click **OK**. If you use the **No fees for stopped instances (VPC-Connected)** feature, you may not be able to start the instance successfully after changing the system disk.

Stop ? Stop instance

Operation will be executed on the selected **1 Instances** . Are you sure you want to proceed?

Stopped By: Stop
 Force Stop

Stopped By: Keep Stopped Instances and Continue Billing

Stopping Subscription based instances does not change their expiration date.

If you need to stop an instance for system disk replacement, disk reinitialization, instance upgrade, or private IP address modification, we recommend that you select Keep Stopped Instances and Continue Billing to avoid startup failure.

OK Cancel

Procedure

To roll back a cloud disk , follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Instances**.
3. Select the target region.
4. Find the target instance and click its ID to go to its **Instance Details** page.
5. In the left-side navigation pane, click **Instance Snapshots**.
6. Find the target snapshot and then in the **Actions** column, click **Disk Rollback**.
7. In the dialog box, click **OK**.

**Note:**

If you select **Start Instance after Rollback**, the instance starts automatically after the disk is restored.

Related API

[ResetDisk](#)

Additional operations

If you resize a cloud disk after creating a snapshot, you can connect to the instance to resize its file system. For more information, see:

- [Linux _ Resize a data disk](#)
- [Windows _ Resize a data disk](#)

6.9 Convert billing methods of cloud disks

The billing method of a cloud disk depends on how it is created:

- For cloud disks created with Subscription instances, prepayment of the service fee is required for it to be available for use. For more information, see [Subscription](#).
- For cloud disks created jointly with Pay-As-You-Go instances, or created separately the billing is on a Pay-As-You-Go basis. For more information, see [Pay-As-You-Go](#).

You can change the billing method of a cloud disk, as shown in the following table.

Conversion of billing methods	Conversion method	Suitable for	Effective date
Subscription → Pay-As-You-Go	Renew for configuration downgrade	Subscription cloud disks attached to Subscription instances . The billing method of the system disk cannot be changed.	Effective from the next billing cycle
Pay-As-You-Go → Subscription	Upgrade configurations	Pay-As-You-Go data disks attached to Subscription instances . The billing method of the system disk cannot be changed.	Effective immediately

Conversion of billing methods	Conversion method	Suitable for	Effective date
	Switch from Pay-As-You-Go to subscription	The system disks and data disks attached to the Pay-As-You-Go instances.	

6.10 Replace the system disk (public image)

You can replace the system disk if you select an incorrect OS when creating an ECS instance or you need to replace the current OS. The new system disk will be allocated a new ID, and the previous system disk ID will be released.

You can replace the image of the system disk with a public image, shared image, customized image, or any other image from the Alibaba Cloud Marketplace.



Note:

Microsoft has ended extended technical support for Windows Server 2003. For the purpose of data security, we recommend that you do not continue running Windows Server 2003 on your ECS instance. Its image is no longer provided. For more information, see [offline announcement of Windows Server 2003 system image](#).

After you replace the system disk, note that:

- A new system disk with a new disk ID is allocated to your instance, and the original ID is released.
- The cloud type of the cloud disk cannot be replaced.
- The IP address and the MAC address remain unchanged.
- We recommend you [delete snapshots or automatic snapshot policies](#) to ensure sufficient snapshot quota for executing automatic snapshot policies of the new system disk.

This article describes how to replace an existing image with a public image. If you need to use a non-public image, see [replace the system disk \(non-public image\)](#).

Precautions

Replacing the system disk exposes the system to multiple risks. Read the following precautions carefully before you begin:

Risks

The risks of replacing the system disk are as follows:

- Replacing the system disk will stop your instances, which means your business services will be disrupted.
- After replacing the system disk, you must redeploy the service running environment on the new system disk, which may result in a long interruption to services.
- After you replace the system disk, a new system disk with a new disk ID will be assigned to your instance. This means that you cannot use snapshots of the original system disk to roll back the new system disk.

**Note:**

After you replace the system disk, the snapshots you have manually created are not affected. You can still use them to create customized images. If you have configured automatic snapshot policies for the original system disk to allow automatic snapshots to be released along with the disk, the snapshot policies no longer apply and all automatic snapshots of the original system disk will be automatically deleted.

Precautions for cross-OS disk replacement

Cross-OS disk replacement refers to replacing the system disk between Linux and Windows.

**Note:**

Regions outside mainland China do not support disk replacement between Linux and Windows. Disk replacement between Linux editions or Windows editions are supported.

During cross-OS disk replacement, the file format of the data disk may be unidentifiable.

- If no important data exists on the data disk, we recommend that you [reinitialize the disk](#) and format it to the default file system of your OS.
- If important data exists in your data disk, perform the following actions as required:
 - From Windows to Linux, you must install a software application, for example, NTFS-3G, because NTFS cannot be identified by Linux by default.
 - From Linux to Windows, you must install a software application, for example, Ext2Read or Ext2Fsd, because ext3, ext4, and XFS cannot be recognized by Windows by default.

If you replace Windows with Linux, use a password or an SSH key pair for authentication.

Preparations

- Make sure that there is sufficient system disk space. We recommend that you reserve 1 GiB in case the OS cannot properly start after system disk replacement.

- If you want to replace the OS to Linux and use an SSH key pair for authentication, [create an SSH key pair](#) first.
- Replacing the system disk may cause data loss or service interruption. To minimize impact to your business services, we recommend that you [create snapshots](#) for the original system disk before replacement.

**Note:**

We recommend that you create snapshots during off-peak hours and plan for sufficient time. For example, to create a snapshot of 40 GiB for the first time, the process takes about 40 minutes. Additionally, creating a snapshot may reduce I/O performance of a block storage device by up to 10%.

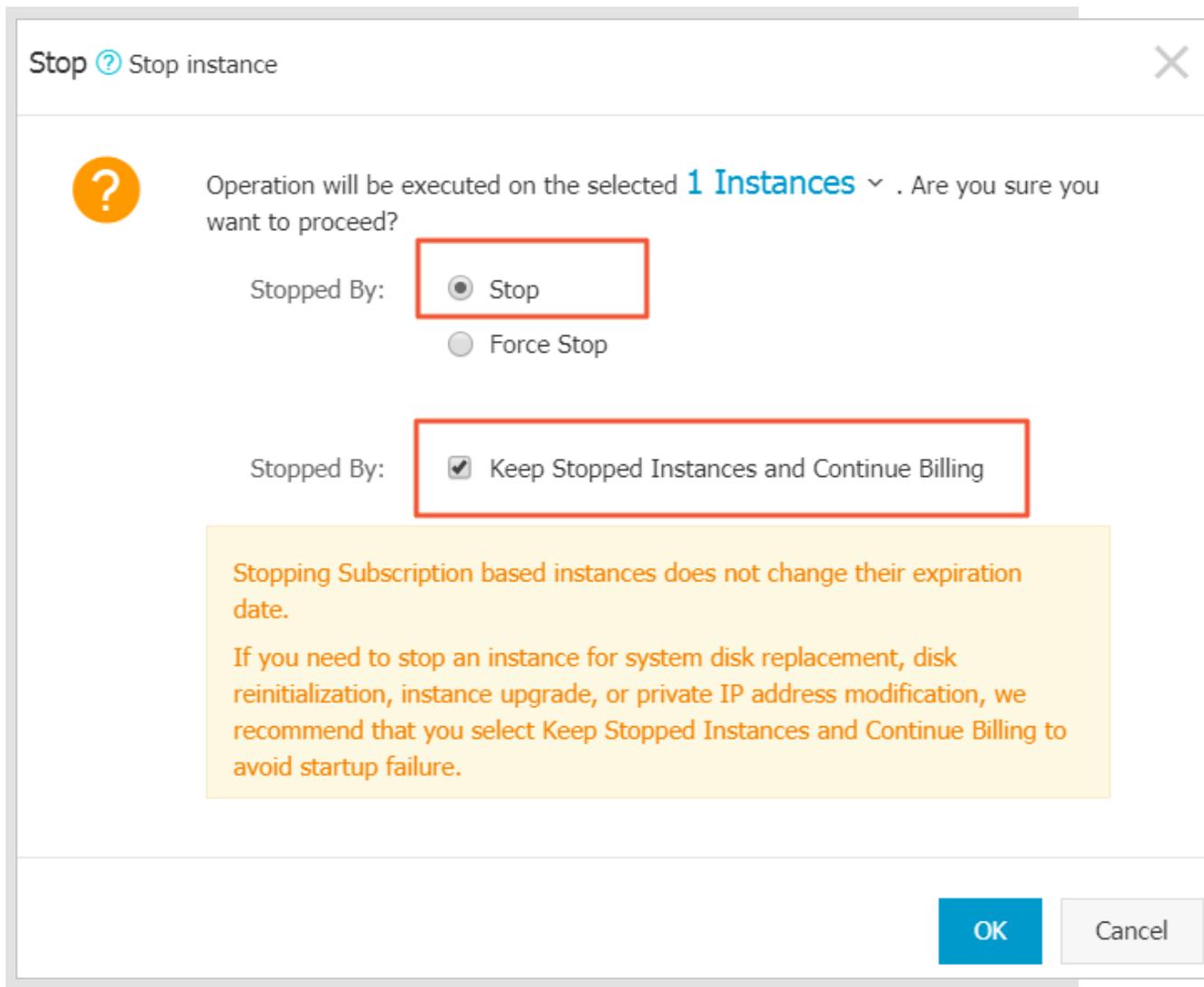
Procedure

To replace the system disk, follow these steps:

1. Log on to the [ECS console](#).
2. In the left navigation pane, click **Instances**.
3. Select the target region.
4. In the **Actions** column of the target instance, select **More > Instance Status > Stop** and follow the instructions in the prompt to stop the instance.

**Note:**

If the instance is a Pay-As-You-Go instance using a VPC with the No Fees for Stopped Instances function enabled, in the displayed **Notes** dialog box, click **OK**. In the displayed **Stop** dialog box, select **Keep Stopped Instances and Continue Billing**. If you select **No Fees for Stopped Instances (VPC-Connected)**, the instance may not be properly started after system disk replacement.



5. After the instance is stopped, in the **Actions** column, select **More > Disk and Image > Replace System Disk**.
6. In the displayed dialog box, read the precautionary statement about system disk replacement and then click **OK**.
7. On the **Replace System Disk** page, configure the following:
 - a. **Image Type**: Select **Public Image** and then select the image version.

**Note:**

If you need to use a non-public image, see [replace the system disk \(non-public image\)](#).

- b. **System Disk**: Unchangeable. However, you can expand the disk space to meet the requirements of your system disk and services. The maximum disk space is 500 GiB. The minimum space of the system disk you can configured is determined by the current disk space and image type.

Image	Allowed range (GiB)
Linux (excluding CoreOS) + FreeBSD	20-500
CoreOS	30-500
Windows	40-500

**Note:**

If your instance has been configured with renewal for configuration downgrade, you cannot change the system disk space until the next billing cycle.

c. Security enhancement:

- If the new OS is Windows, you can only use a password for authentication.

Image Type:

Public Image

Custom Image

Shared Image

Marketplace Image

Public Image:

Windows Server

Version 1709 DataCenter Edition 64bit Chinese ...

Selection advice >

Security enhancement ?

System Disk:

Ultra Cloud Disk

40 GB

2120 IOPS The default system disk device name : /dev/xvda

To learn how to select SSD cloud disks, ultra cloud disks, and basic cloud disks, [Learn More >](#)

Login name:
administrator

Login password:

It must be 8 - 30 characters long and contain three types of

Confirm password:

- If the instance is an I/O optimized instance, and the new OS is Linux, you can use either a password or an SSH key pair for authentication. In this case, we recommend you set a login password or bind an SSH key pair.

Image Type:

Public Image Custom Image Shared Image Marketplace Image

Public Image:

CentOS 7.4 64bit Selection advice >

Security enhancement ?

System Disk:

Ultra Cloud Disk 40 GB 2120 IOPS The default system disk device name : /dev/xvda
To learn how to select SSD cloud disks, ultra cloud disks, and basic cloud disks, [Learn More](#) >

Security:

Key Pair Password Set Now

A key pair includes a public key and a private key. Currently only I/O-optimized instances support the use of key pair. Using a key pair, you cannot log on with a user name and password.

Key Pair:

Select the Key Pair ↕

Also, you can go to the console to [create an access key](#) >

- d. Confirm **Instance Cost**, which includes the image fee and system disk fee. For more information about system disk price, see the [pricing](#) page of ECS.
- e. Check the configuration and click **Confirm to change**.

Log on to the ECS console to monitor the system status. It may take about 10 minutes to replace the OS and update the system status. After the OS is replaced, the instance automatically starts.

Additional operations

After replacing the system disk, you can perform the following operations:

- (Optional) [Apply automatic snapshot policies to disks](#). Automatic snapshot policies are bound to the disk ID. After the system disk is replaced, automatic snapshot policies applied on the original disk automatically fail. You need to configure automatic snapshot policies for the new system disk.
- If the OS before and after disk replacement is Linux, and if a data disk is mounted to the instance and the partition is set to be mounted automatically at instance startup, then all mounting information will be lost. In this case, you need to write the new partition information

into the `/etc/fstab` file of the new system disk and mount the partition, but do not need to partition or format the data disk for another time. The steps are described as follows. For more information about operation commands, see [format and mount data disks for Linux instances](#).

1. (Recommended) Back up the `/etc/fstab` file.
2. Write information about the new partition into the `/etc/fstab` file.
3. Check the information in the `/etc/fstab` file.
4. Run `mount` to mount the partition.
5. Run `df -h -h` to check the file system space and usage.

After the data partition is mounted, the data disk is ready for use without the need for instance restart.

Related API

[ReplaceSystemDisk](#)

6.11 Replace the system disk (non-public image)

You can replace the system disk if you select an incorrect OS when creating an ECS instance or you need to replace the current OS. The new system disk will be allocated a new ID, and the previous system disk ID will be released.

You can replace the image of the system disk with a public image, shared image, customized image, or any other image from the marketplace.



Note:

Microsoft has ended extended technical support for Windows Server 2003. For the purpose of data security, we recommend that you do not continue running Windows Server 2003 on your ECS instance. Its image is no longer provided. For more information, see [offline announcement of Windows Server 2003 system image](#).

After you replace the system disk, note that:

- A new system disk with a new disk ID is allocated to your instance, and the original ID is released.
- The cloud type of the cloud disk cannot be replaced.
- The IP address and the MAC address remain unchanged.
- We recommend you [delete snapshots or automatic snapshot policies](#) to ensure sufficient snapshot quota for executing automatic snapshot policies of the new system disk.

This article describes how to replace an existing image with a non-public image. If you need to use a public image, see [replace the system disk \(public image\)](#).

Precautions

Replacing the system disk exposes the system to multiple risks. Read the following sections carefully before you begin:

Risks

Risks of replacing the system disk are as follows:

- Replacing the system disk will stop your instances, which means your business services will be disrupted.
- After replacing the system disk, you must redeploy the service running environment on the new system disk, which may result in a long interruption to services.
- After you replace the system disk, a new system disk with a new disk ID will be assigned to your instance. This means that you cannot use snapshots of the original system disk to roll back the new system disk.



Note:

After you replace the system disk, the snapshots you have manually created are not affected. You can still use them to create customized images. If you have configured automatic snapshot policies for the original system disk to allow automatic snapshots to be released along with the disk, the snapshot policies no longer apply and all automatic snapshots of the original system disk will be automatically deleted.

Precautions for cross-OS disk replacement

Cross-OS disk replacement refers to replacing the system disk between Linux and Windows.



Note:

Regions outside mainland China do not support disk replacement between Linux and Windows. Disk replacement between Linux editions or Windows editions are supported.

During cross-OS disk replacement, the file format of the data disk may be unidentifiable.

- If no important data exists on the data disk, we recommend that you [reinitialize the disk](#) and format it to the default file system of your OS.
- If important data exists in your data disk, perform the following actions as required:

- From Windows to Linux, you must install a software application, for example, NTFS-3G, because NTFS cannot be identified by Linux by default.
- From Linux to Windows, you must install a software application, for example, Ext2Read or Ext2Fsd, because ext3, ext4, and XFS cannot be recognized by Windows by default.

If you replace Windows with Linux, use a password or an SSH key pair for authentication.

Prerequisites

Before replacing the existing image with a non-public image, note the following:

- If the target image is a custom image:
 - If you want to use an image of a specified ECS instance, you must [create a snapshot for the system disk of the specified instance](#) and [create a custom image using a snapshot](#). If the specified instance and the one whose system disk you want to change are located in different regions, you need to [copy the images](#).
 - To use a local physical image file, [import it on the ECS console](#) or [use Packer to create and import the local image](#). The region where the image is located must be the same as that of your instance.
 - To use an image in a region other than that of your instance, you must [copy the image](#) first.



Note:

Imported or duplicated images will be displayed in the **Custom Image** drop-down list.

- To use an image owned by another Alibaba Cloud account, the account must first [share the image](#).
- If you want to replace the OS to Linux and use an SSH key pair for authentication, you must first [create an SSH key pair](#).
- Replacing the system disk may cause data loss or service interruption. To minimize impact to your business services, we recommend that you [create snapshots](#) for the original system disk before replacement.
- If you want to replace the OS to Linux, make sure that there is sufficient system disk space. We recommend that you reserve 1 GiB in case the OS cannot properly start after system disk replacement.



Note:

We recommend that you create snapshots during off-peak hours and plan for sufficient time. For example, to create a snapshot of 40 GiB for the first time, the process takes about 40 minutes. Additionally, creating a snapshot may reduce I/O performance of a block storage device by up to 10%.

Procedure

1. Log on to the [ECS console](#).
2. In the left navigation pane, click **Instances**.
3. Select the target region.
4. In the **Actions** column of the target instance, select **More > Instance Status > Stop** and follow the instructions in the prompt to stop the instance.

The action is successful when the instance status is **Stopped**.

5. In the **Actions** column, select **More > Disk and Image > Replace System Disk**.
6. In the displayed dialog box, read the precautionary statement about system disk replacement, and then click **OK**.
7. On the **Replace System Disk** page, configure the following:
 - a. **Image Type:** Select **Custom Image**, **Shared Image**, or **Marketplace Image**, and then select the image version.
 - b. **System Disk:** Unchangeable. However, you can expand the disk space to meet the requirements of your system disk and services. The maximum disk space is 500 GiB. The minimum space of the system disk you can configured is determined by the current disk space and image type.

Image	Allowed range (GiB)
Linux (excluding CoreOS) + FreeBSD	20-500
CoreOS	30-500
Windows	40-500



Note:

If your instance has been configured with [renewal for configuration downgrade](#), you cannot change the system disk space until the next billing cycle.

c. Security enhancement:

- If the new OS is Windows, you can only use a password for authentication.

- If the instance is an I/O optimized instance, and the new OS is Linux, you can use either a password or an SSH key pair for authentication. In this case, set a login password or bind an SSH key pair.
- d. Confirm **Instance Cost** : includes the image fee and system disk fee. For more information, see the [pricing](#) page of ECS.
- e. Check the configuration and click **Confirm to change**.

Log on to the ECS console to monitor the system status. It may take about 10 minutes to replace the OS. After the OS is replaced, the instance automatically starts.

Additional operations

After replacing the system disk, you can perform the following operations:

- (Optional) [Apply automatic snapshot policies to disks](#). Automatic snapshot policies are bound to the disk ID. After the system disk is replaced, automatic snapshot policies applied on the original disk automatically fail. You need to configure automatic snapshot policies for the new system disk.
- If the OS before and after disk replacement is Linux before and after disk replacement, and if a data disk is mounted to the instance and the partition is set to be mounted automatically at instance startup, then all mounting information will be lost. In this case, you need to write the new partition information into the `/etc/fstab` file of the new system disk and mount the partition, but do not need to partition or format the data disk for another time. The steps are described as follows. For more information about operation commands, see [format and mount data disks for Linux instances](#).
 1. (Recommended) Back up the `/etc/fstab` file.
 2. Write information about the new partition into the `/etc/fstab` file.
 3. Check the information in the `/etc/fstab` file.
 4. Run `mount` to mount the partition.
 5. Run `df -h -h` to check the file system space and usage.

After the data partition is mounted, the data disk is ready for use without the need for instance restart.

Related API

[ReplaceSystemDisk](#)

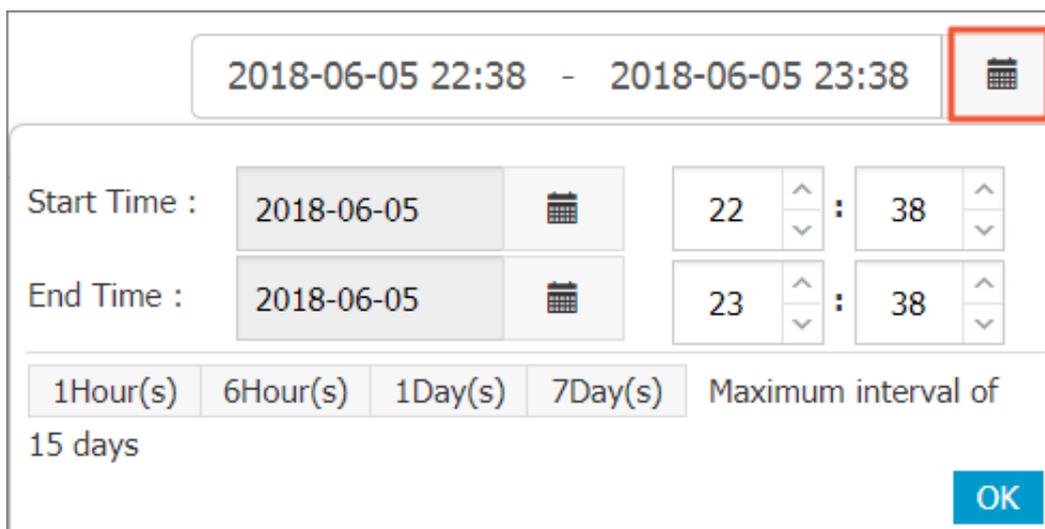
6.12 Monitor a cloud disk

You can monitor the IOPS and throughput of a cloud disk in the ECS console or, if you have installed the CloudMonitor agent, you can monitor the disk in the CloudMonitor console.

To monitor the IOPS and throughput of a cloud disk in the ECS console, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Block Storage** > **Disks**.
3. Select the target region.
4. Find a cloud disk and click its ID to go to the **Details** page.
5. In the left-side navigation pane, click **Disk Monitoring**.
6. On the **Monitoring Information** page, click the  icon and set Start Time and End Time

for monitoring information. You can check the monitoring information of a cloud disk for up to 15 days.



2018-06-05 22:38 - 2018-06-05 23:38

Start Time : 2018-06-05 22 : 38

End Time : 2018-06-05 23 : 38

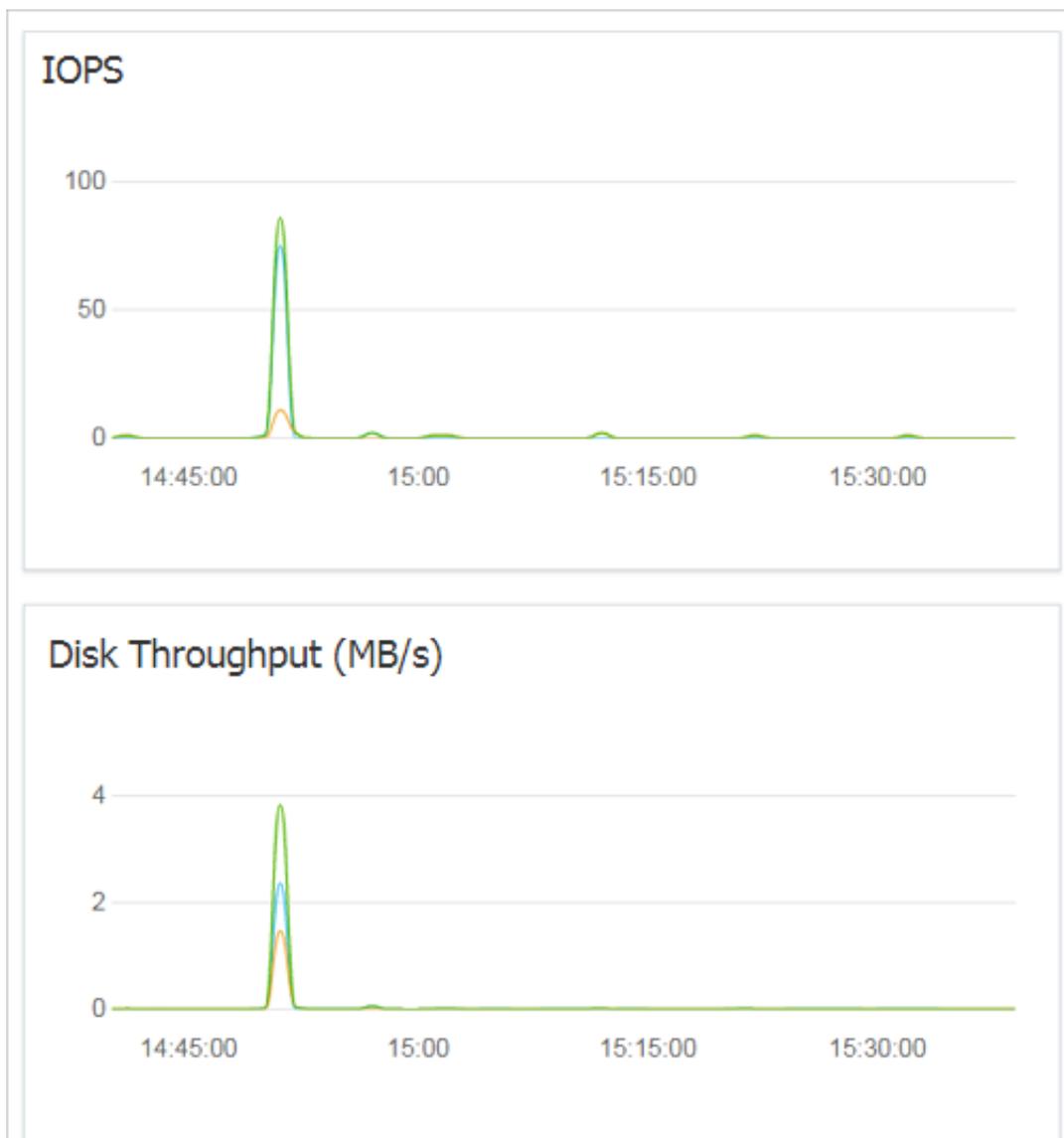
1Hour(s) 6Hour(s) 1Day(s) 7Day(s) Maximum interval of 15 days

OK

7. View the IOPS and throughput of the cloud disk.

**Note:**

Click a legend in the chart to view a single performance index of a cloud disk.



6.13 Release a cloud disk

We recommend you release a cloud disk when you no longer require it to avoid incurring excess fees. Releasing a cloud disk is a permanent, irreversible action and is irreversible. After its release, data on the cloud disk cannot be restored. You can only release a cloud disk in the Available status. Exercise caution when performing this action.

Note

When releasing a cloud disk, note that

- Only cloud disks that are in the **Available** status can be released independently. Other cloud disks, such as those used as system disks, or Subscription billed cloud disks used as data

disks, can only be released together with ECS instances. If a cloud disk is in the **In Use** status, you must first Detach it from the instance.

- By default, automatic snapshots are released together with their cloud disks. However, manually created snapshots are not. You can change the snapshot release configuration when attaching a cloud disk.

**Note:**

Each cloud disk can have up to 64 snapshots. To make sure you have sufficient storage space for the automatic snapshots, we recommend that you release automatically or manually created snapshots that you no longer require.

- You can have data backed up before releasing a cloud disk. For example, by creating a snapshot.

Procedure

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Block Storage > Disks**.
3. Select the target region.
4. Select the disk that you want to release and check it is in the **Unmounted** status. Then, in the **Actions** column, select **More > Release**.
5. In the Release dialog box, read the note and click **Confirm Release**.

Related API

[DeleteDisk](#)

7 Snapshots

7.1 Create a snapshot

A snapshot records the disk data state at a specific point in time that can be used for easier data backup and image customization.

Scenarios

You can create a snapshot for data backup scenarios. For example, you can create a snapshot if you need to modify critical system files, [migrate instances from a classic network to a VPC](#), back up data routinely, mitigate network attacks, [change the operating system](#), or provide data support for a production environment.

Additionally, you can use a snapshot to [create customized images](#) for quick application environment deployment in a large number of instances.

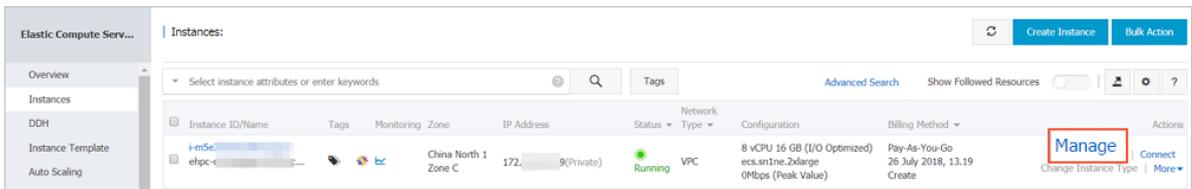
Precautions

- Creating a snapshot may impact disk performance and I/O performance. We recommend that you create a snapshot during off-peak business hours.
- A snapshot only records data at a specific point in time. Therefore, incremental data generated by disk operations during snapshot creation is not synchronized to a created snapshot.
- To guarantee successful snapshot creation, we recommend you do not modify the ECS instance status (that is, stop or restart the instance) during snapshot creation.
- If you want to create a snapshot based on an instance, the instance must be in **Running** or **Stopped** status.
- If you want to create a snapshot based on a disk, the disk must be in **Running** status.
- Manually created snapshots must be manually deleted. Therefore, you need to [delete unnecessary snapshot](#) to prevent continuous deduction caused by increasing snapshot capacity.
- If you create an extended volume by using a multi-partition single disk, the snapshots you created can be used for [rolling back a cloud disk](#).
- After you create a dynamic extended volume by using multiple disks and no I/O operation is performed on data in the extended volume, the snapshot you created can be used for disk rollback. If I/O operations are continuously performed in the extended volume, data consistency of the rolled-back disk is uncertain.

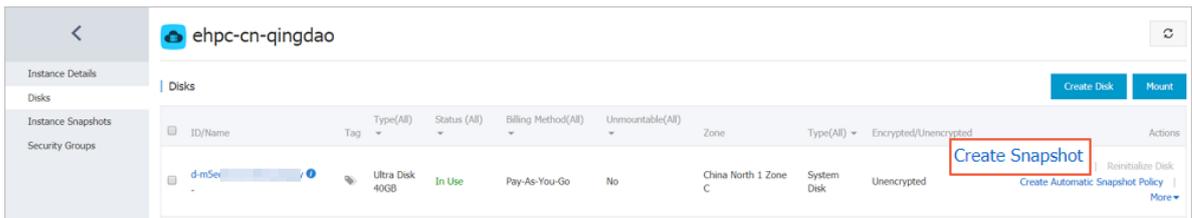
Procedure

To create a snapshot on the ECS console, follow these steps:

1. Log on to the [ECS console](#).
2. Select the target region.
3. In the left-side navigation pane, click **Instances**.
4. Locate the instance for which you want to create a snapshot and click **Manage** in the Actions column.



5. In the left-side navigation pane, click **Disks**, locate the target disk, and then click **Create Snapshot**. You can select only one disk at a time. **Type** can be either system disk or data disk.



6. Enter a name for the snapshot and click **OK**.

Create Snapshot ? X

Do not change the status of the instance during snapshot creation (for example, do not stop or restart the instance). Changes to the instance's status may cause the snapshot creation to fail.

ID: d-m5 [redacted]

Instance ID/Name: i-m5 [redacted]

Type: Ultra Disk

*Snapshot Name:
 The snapshot name can be 2 to 128 characters in length and cannot start with **auto**.

Tag:

1:11 X

OK
Cancel

7. In the left-side navigation pane, click **Instance Snapshots**. Here, you can see the creation progress, estimated remaining time, and status of the snapshot.

<p style="margin: 0;"><</p> <p style="margin: 0;">ehpc-cn-qingdao</p> <ul style="list-style-type: none"> <li style="padding: 2px;">Instance Details <li style="padding: 2px;">Disks <li style="padding: 2px; background-color: #f0f0f0;">Instance Snapshots <li style="padding: 2px;">Security Groups 	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th colspan="7" style="text-align: left; padding: 2px;">Snapshots</th> </tr> <tr> <th style="text-align: left; padding: 2px;">Snapshot ID/Name</th> <th style="text-align: left; padding: 2px;">Tags</th> <th style="text-align: left; padding: 2px;">Disk ID</th> <th style="text-align: left; padding: 2px;">Disk Size</th> <th style="text-align: left; padding: 2px;">Disk Type(All)</th> <th style="text-align: left; padding: 2px;">Encrypted/Unencrypted</th> <th style="text-align: left; padding: 2px;">Created At</th> </tr> </thead> <tbody> <tr> <td style="padding: 2px;">s-m5 [redacted]</td> <td style="padding: 2px;">[tag icon]</td> <td style="padding: 2px;">d-m5 [redacted]</td> <td style="padding: 2px;">40GB</td> <td style="padding: 2px;">System Disk</td> <td style="padding: 2px;">Unencrypted</td> <td style="padding: 2px;">15 August 2018, 14:41</td> </tr> <tr> <td style="padding: 2px;">CreateSnapshot</td> <td style="padding: 2px;"></td> </tr> </tbody> </table> <div style="border: 1px solid #e67e22; padding: 5px; margin-top: 5px; display: flex; justify-content: space-between;"> <div style="width: 60%;"> <p style="margin: 0;">Progress</p> <p style="margin: 0;">48%</p> <p style="margin: 0;">Time left 0Hours5minute</p> </div> <div style="width: 35%;"> <p style="margin: 0;">Status</p> <p style="margin: 0;">Progressing</p> </div> </div>	Snapshots							Snapshot ID/Name	Tags	Disk ID	Disk Size	Disk Type(All)	Encrypted/Unencrypted	Created At	s-m5 [redacted]	[tag icon]	d-m5 [redacted]	40GB	System Disk	Unencrypted	15 August 2018, 14:41	CreateSnapshot						
Snapshots																													
Snapshot ID/Name	Tags	Disk ID	Disk Size	Disk Type(All)	Encrypted/Unencrypted	Created At																							
s-m5 [redacted]	[tag icon]	d-m5 [redacted]	40GB	System Disk	Unencrypted	15 August 2018, 14:41																							
CreateSnapshot																													

Alternatively, you can use the ECS API [CreateSnapshot](#) to create a snapshot.

Time required

The time required for creating a snapshot depends on the disk capacity.

According to the *incremental snapshot mechanism*, the first disk snapshot records full disk data, which takes an extended amount of time. However, subsequent snapshots are generally created in a shorter time. However, the exact time required depends on the amount of data generated since the first snapshot creation. The larger the data amount is, the longer time the snapshot creation will take.

What to do next

After you create a snapshot, you can:

- [Roll back a cloud disk](#)
- [Create a cloud disk from a snapshot](#)
- [Create a custom image by using a snapshot](#)

7.2 Create and delete an automatic snapshot policy

An automatic snapshot policy is a set of defined parameters for automatically creating snapshots.



Note:

- We recommend you set the automatic snapshot creation time and repeat date during off-peak business hours to avoid interruptions to your business services.
- You can create a maximum of 100 automatic snapshot policies per region.

Prerequisite

You must have created an automatic snapshot policy.

Procedure

To create an automatic snapshot policy, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Snapshots and Images > Automatic Snapshot Policies**.
3. On the **Automatic Snapshot Policies** page, perform the following action as required:
 - If you want to create a policy, click **Create Policy** at the upper-right corner.
 - If you want to modify a policy, find the policy that you want to modify, and click **Modify Policy** in the **Actions** column.
4. In the **Create Policy** or **Modify Policy** dialog box, define the automatic snapshot policy as follows.

- Enter a policy name.
- Select a time after **Executed At** to specify the time of day for automatically creating snapshots.
- Specify the **Execution Frequency**.
- Set a number after **Keep Snapshots** to defines the number of days a snapshot can be retained. The value range is 1–65535 days, or permanently. By default, it is set to 30 days. You can also choose to permanently retain automatic snapshots.

**Note:**

When the number of snapshots reaches the limit, the system automatically removes the oldest automatic snapshots created. Manually created snapshots are not removed.

5. Click **OK**.

Additional operations

You can [apply automatic snapshot policies to disks](#).

Related APIs

- [CreateAutoSnapshotPolicy](#): Creates automatic snapshot policies.
- [DescribeAutoSnapshotPolicyEx](#): Queries automatic snapshot policies.
- [ModifyAutoSnapshotPolicyEx](#): Modifies automatic snapshot policies.

7.3 Apply automatic snapshot policies to disks

You can apply an automatic snapshot policy to disks according to your business needs.

Automatic snapshots are named in the format of auto_yyyymmdd_1, for example, auto_20140418_1.

**Note:**

- Creating snapshots may interrupt read/write operations on your disk. We recommend that you set the creation time of automatic snapshots during off-peak business hours to avoid interruptions to your business services.
- Automatic snapshot policies cannot be applied to unused basic cloud disks.
- Snapshots that are manually created do not conflict with automatic snapshots. However, if an automatic snapshot is being created on a disk, you must wait for it to finish before manually creating a snapshot.

You can apply an automatic snapshot policy to a disk using either of the following menus:

- From the Cloud Disks menu. This method is suitable for applying an automatic snapshot policy to a specific disk.
- From the Snapshots and Images menu. This method is suitable for applying a unified automatic snapshot policy to multiple disks.

From the Cloud Disks menu

To apply an automatic snapshot policy through the Cloud Disks menu, follow these steps:

1. Log on to the [ECS console](#).
2. Select the target region.
3. In the left-side navigation pane, click **Disks**.
4. Select the disk for which you want to execute the policy and click **Create Automatic Snapshot Policy**.
5. Enable the automatic snapshot function and select the desired snapshot policy.
6. Click **OK**.

From the Snapshots and Images menu

To apply or disable an automatic snapshot policy, follow these steps:

1. Log on to the [ECS console](#).
2. Select the target region.
3. In the left-side navigation pane, select **Snapshots and Images > Automatic Snapshot Policies**.
4. Select the automatic snapshot policy that you want to apply and click **Apply Policy**.
5. To enable an automatic snapshot policy, select **Disks without Policy Applied** to view the disks. Find the disk for which you want to enable the policy, and then click **Apply Policy** after it. Alternatively, if you select multiple disks, click **Apply Policy** at the lower-left corner.

Modify the automated Snapshot policy ✕

After you enable the automated snapshot policy, your Snapshot will be managed according to the automated Snapshot policy.

Disk without preset policy
 Disk with preset policy

Disk Name

<input checked="" type="checkbox"/>	Disk ID/Disk Name	Disk Category (All) ▾	Disk Property (All) ▾	Action
<input checked="" type="checkbox"/>	d-28eyf2ur4 ⓘ	General CloudDisk 40GB	System Disk	<input type="button" value="Enable autosnapshot"/>

Total: 1 item(s) , Per Page: 20 item(s)

6. To disable the automatic snapshot policy, select the **Disks with Policy Applied** tab to view the disks, select the disk for which you want to disable the policy, and then click **Disable Policy** at the right side. Note that if you select multiple disks, you need to click **Disable Policy** at the lower-left corner to disable the automatic snapshot policy for all selected disks.

Modify the automated Snapshot policy ✕

After you enable the automated snapshot policy, your Snapshot will be managed according to the automated Snapshot policy.

Disk without preset policy
 Disk with preset policy

Disk Name

<input checked="" type="checkbox"/>	Disk ID/Disk Name	Disk Category (All) ▾	Disk Property (All) ▾	Action
<input checked="" type="checkbox"/>	d-28eyf2ur4 ⓘ	General CloudDisk 40GB	System Disk	<input type="button" value="Disable autosnapshot"/>

Total: 1 item(s) , Per Page: 20 item(s)

7.4 Delete automatic snapshots when releasing disks

By default, automatic snapshots of cloud disks are not released at the same time as the cloud disk. However, you can modify the disk property so that automatic snapshots are released when you:

- [Replace the system disk \(non-public image\)](#): The previous system disks are released. If an automatic snapshot has been set up to release with the cloud disks, the automatic snapshots of the previous system disks are automatically deleted.
- [Detach a cloud disk](#).

Procedure

To delete automatic snapshot policies at the same time as a disk release, follow these steps:

1. Log on to the [ECS console](#).
2. Select the target region.
3. In the left-side navigation pane, click **Block Storage > Disks**.
4. Select the disk that you want to configure and then, in the **Actions** column, click **More > Modify Attributes**.
5. In the **Modify Disk Type** dialog box, select **Delete Automatic Snapshots while Releasing Disk**, and then click **OK**.

Modify Disk Attribute

Disk: d-t4nibr01p3bgwd2mjkrd

Instance Subordinated to: i-t4nactualy65ptudegzo / iZt4nactualy65Z

Device Name: /dev/xvdb

Disk Type: SSD Cloud Disk

Release Action: Release Disk with Instance
 Delete automatic snapshot when releasing disk ?

OK Cancel

Related API

[ModifyDiskAttribute](#)

7.5 Delete snapshots or automatic snapshot policies

When you no longer require a snapshot, or you have reached your snapshot quota, we recommend you delete unnecessary snapshots to free up space and avoid being charged for excessive fees.



Note:

- After a snapshot is deleted, it cannot be restored. Exercise caution when performing this action.

- If a snapshot has been used to create a custom image, you must delete the associated image before you can delete the snapshot.

Delete snapshots

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Snapshots and Images > Snapshots**.
3. Select the target region.
4. Select the snapshot or snapshots you want to delete.
5. Click **Delete** at the lower-left corner, and then click **OK**.

Delete snapshot policies

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Snapshots and Images > Automatic Snapshot Policies**.
3. Select the target region.
4. Find the snapshot policy you want to delete and then, in the **Actions** column, click **Delete Automatic Snapshot Policy**.
5. In the dialog box, confirm the information and click **OK**.

7.6 View a snapshot chain

The snapshot service fee is related to the snapshot size. This article describes how to view the snapshot size on a single disk or under a region.

View the snapshot size on a single disk

When you create snapshots of a cloud disk or a shared block storage device, you can view the snapshot size of the device by using the **Snapshot Chains** feature in the ECS console.

A snapshot chain is composed of all the snapshots of a cloud disk or shared block storage device. After you create a snapshot, the device has a snapshot chain. The snapshot chain uses an identical ID with that of the disk, and provides the following information:

- Snapshot nodes, whereby each snapshot node of the chain represents one snapshot of the device.
- Snapshot size, indicating the storage space occupied by all snapshots of the device.
- Snapshot quota, whereby each device can have up to 64 snapshots, including those created manually or automatically.

Prerequisite

You have [created snapshots](#).

Procedure

To view the total size of all the snapshots of a cloud disk or shared block storage device, follow these steps:

1. Log on to the [ECS console](#).
2. Select the target region.
3. In the left-side navigation pane, select **Snapshots and Images > Snapshots** .
4. Find the disk ID of the target snapshot.
5. In the left-side navigation pane, click **Snapshot Chains**.
6. View the size of all snapshots on the disk according to the disk ID found in step 4. You can view the total number and size of snapshots of the disk in the list.

In the **Actions** column, click **Details** to go to the **Snapshot Chain Details** page to see all snapshots of the disk. Here, you can choose to [roll back a cloud disk](#) or [create a custom image by using a snapshot](#).

View the snapshot size under a region

Follow these steps:

1. Log on to the [ECS console](#).
2. Select the target region.
3. In the left-side navigation pane, select **Snapshots and Images > Snapshots**.

Here, you can view the total number and size of snapshots of the disk in the list.

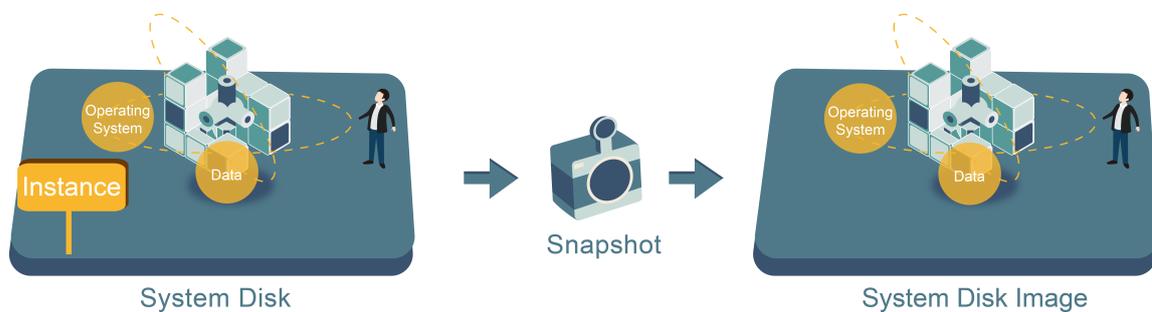
8 Images

8.1 Create custom image

8.1.1 Create a custom image by using a snapshot

Custom images allow you to create multiple ECS instances with identical OS and environment data.

Custom images are based on ECS disk snapshots. You can set up identical or different configurations for ECS instances that are created from images.



You can also use an instance to create an image. For more information, see [create a custom image by using an instance](#).

To enhance the security of custom images created from snapshots, see [security suggestions for Alibaba Cloud custom images](#).



Note:

- Custom images cannot be used across regions.
- You can change the operating system of an instance created from a custom image, and the custom image remains usable. For details, see [change the system disk \(custom image\)](#).
- You can upgrade the instance created from a custom image, including upgrading the CPU, memory, bandwidth, and disks.
- Custom images are created independently from the billing methods of the instances from which they were created. For example, custom images created from Subscription instances can be used for creating Pay-As-You-Go instances. The converse method also applies.
- If the ECS instance used for creating a custom image expires, or the data is erased (that is, the system disk used for the snapshot expires or is released), the custom image and the ECS

instances created from the custom image are not affected. However, automatic snapshots are cleared when an ECS instance is released.

Restrictions for Linux instances

- Do not load data disk information in the `/etc/fstab` file. Otherwise, instances created using this image cannot start.
- We recommend that you `umount` all data disks before creating a custom image, and then use a snapshot to create a custom image. Otherwise, ECS instances that are created based on this custom image may not start.
- Do not upgrade the kernel or operating system version.
- Do not change the system disk partitions. The system disk only supports single root partitions.
- We recommend you check the available space of the system disk to make sure that the system has available space.
- Do not modify critical system files such as `/sbin`, `/bin`, `/lib`, and so on.
- Do not modify the default logon user name root.

Procedure

1. Log on to the [ECS console](#).
2. Select the region.
3. In the left-side navigation pane, click **Instances**.
4. Find the target instance and click its instance ID, or click **Manage** in the **Actions** column.
5. In the left-side navigation pane, click **Instance Snapshots**. Find the target system disk and then click **Create Custom Image** in the **Actions** column.

Elastic Computing Se...		teslz	Disk	16:29:48	Create Custom Image
<input type="checkbox"/>	[blurred]	jgf	20G Data Disk	2016-12-27 16:29:34	100% Success Disk Rollback Create Custom Image
<input type="checkbox"/>	[blurred]	s0r5oi	40G System Disk	2016-12-21 11:12:08	100% Success Disk Rollback Create Custom Image
<input type="checkbox"/>	[blurred]	az680	40G System Disk	2016-12-13 11:07:47	100% Success Disk Rollback Create Custom Image
<input type="checkbox"/>	[blurred]	9gj3	40G System Disk	2016-11-25 08:57:49	100% Success Disk Rollback Create Custom Image
<input type="checkbox"/>	[blurred]	[blurred]	37G Data Disk	2016-08-05 13:38:07	100% Success Disk Rollback Create Custom Image
<input type="checkbox"/>	[blurred]	[blurred]	40G System Disk	2016-03-14 16:00:02	100% Success Disk Rollback Create Custom Image

The snapshot must be created from system disks. Data disks cannot be used to create custom images.

You can also click **Snapshots and Images > Snapshots**, and select a snapshot created from a system disk to **Create Custom Image**.

Elastic Computing Se...		testz		DISK	16:29:48	Create Custom Image	
<input type="checkbox"/>	[blurred]	[blurred]	jgf	20G	Data Disk	2016-12-27 16:29:34	100% Success Disk Rollback Create Custom Image
<input type="checkbox"/>	[blurred]	[blurred]	s0r5oi	40G	System Disk	2016-12-21 11:12:08	100% Success Disk Rollback Create Custom Image
<input type="checkbox"/>	[blurred]	[blurred]	az680	40G	System Disk	2016-12-13 11:07:47	100% Success Disk Rollback Create Custom Image
<input type="checkbox"/>	[blurred]	[blurred]	9gi3	40G	System Disk	2016-11-25 08:57:49	100% Success Disk Rollback Create Custom Image
<input type="checkbox"/>	[blurred]	[blurred]	[blurred]	37G	Data Disk	2016-08-05 13:38:07	100% Success Disk Rollback Create Custom Image
<input type="checkbox"/>	[blurred]	[blurred]	[blurred]	40G	System Disk	2016-03-14 16:00:02	100% Success Disk Rollback Create Custom Image

6. In the **Create Custom Image dialog box, complete the following:**

- Confirm the snapshot ID.
- Enter a name and description of the custom image.
- Optional. Check **Add Data Disk Snapshot**, select multiple snapshots of data disks for the image, and click **Add** to add a data disk.

 **Note:**

- We recommend that you remove sensitive data from the data disk before creating a custom image to guarantee data security.
- If the snapshot disk capacity is left blank, an empty disk is created with the default capacity of 5 GiB.
- If you select available snapshots, the disk size is the same as the size of the snapshots.

Create Custom Image
✕

When creating a custom image with Linux system, please do not load data disk information in the `/etc/fstab` file. Otherwise, you cannot launch the instance created through the image.

System Snapshot ID: `si-bp0dbycpjpp4hp7xvda29/`

* Image Name:

It must contain 2-128 characters and begin with English letters or Chinese characters. It can include numbers and the characters ".", "_", and "-".

* Image Description:

It must contain 2-256 characters and it cannot begin with `http://` or `https://`

Add Data Disk Snapshot

Snapshot Details:

Snapshot ID	Device Name:	Disk Capacity:	Action
<code>si-bp0dbycpjpp4hp7xvda29(System Disk)</code>	<code>/dev/xvda</code>	<input style="width: 30px;" type="text"/> GB	Delete

1. Leaving the snapshot ID blank will create an empty disk. Default capacity: 5 GB, with up to 2,000 GB supported.
 2. If a snapshot ID is selected, the default disk capacity will be the snapshot capacity.
 3. If the device name is blank, it will be randomly allocated.

7. Click **Create**. Then, in the left-side navigation pane, select **Snapshots and Images > Images** to view the images you have created.

Linux instance image FAQ

How to `umount` a disk and delete disk table data?

If `/dev/hda5` is attached to `/mnt/hda5`, run any of the following three commands to detach the file system.

```
umount /dev/hda5
umount /mnt/hda5
umount /dev/hda5/mnt/hda5
```

`/Etc/fstab` is an important configuration file in Linux. It contains the details of mounting the file system and storage devices upon startup. If you do not want to mount a specified partition when starting the instance, delete the corresponding lines from `/etc/fstab`. For example, you can delete the following statement to disconnect `xvdb1` upon startup: `/dev/xvdb1 /leejd ext4 defaults 0 0`.

How to determine whether a data disk is detached and a custom image can be created?

You must make sure that the statement line for automatically attaching mounting data disk has been deleted from the fstab file.

Use the `mount` command to view the information of all mounted devices. Make sure that the execution results do not contain the information of the data disk partition.

Relevant configuration files

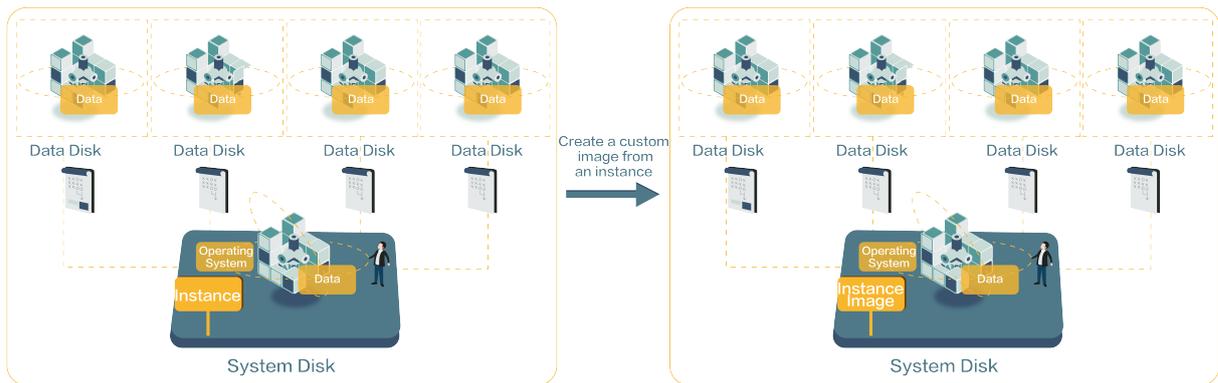
Before creating an image, make sure that the key configuration files listed in the following table have not been modified. Otherwise, the new instance cannot start.

Configuration file	Related to	Risks if modified
<code>/etc/issue*</code> , <code>/etc/*-release</code> , and <code>/etc/*_version</code>	System release version	Modifying <code>/etc/issue*</code> makes the system release version unidentifiable, which can cause instance creation failure.
<code>/boot/grub/menu.lst</code> and <code>/boot/grub/grub.conf</code>	System startup	Modifying <code>/boot/grub/menu.lst</code> results in kernel loading failure, which means the system cannot start.
<code>/etc/fstab</code>	Partitions upon startup	Modifying <code>/etc/fstab</code> causes partition mounting failure, which means the system cannot start.
<code>/etc/shadow</code>	System passwords	If this file is set to read-only, the password file cannot be edited, which means instance creation fails.
<code>/etc/selinux/config</code>	System security policies	Modifying <code>/etc/selinux/config</code> and enabling SELinux results in start failure.

8.1.2 Create a custom image by using an instance

You can create a custom image based on an ECS instance. That is,, you can fully copy all its disks and pack the data into an image.

During this process, snapshots are automatically created for all disks of the instance, including the system disk and data disks. All the created snapshots compose a new custom image. The following figure details this process.



For information about creating an image from a snapshot, see [create a custom image by using a snapshot](#).

Considerations

- Make sure you have deleted all confidential data in the ECS instance before creating a custom image to guarantee data security.
- During creation, do not change the status of the instance. Specifically, do not stop, start, or restart the instance.
- If your custom image contains data disks, new data disks along with the ECS instance are created together. The data on the data disk duplicates the data disk snapshot in your custom image according to the mount device.
- You can export custom images that contain data disks.
- You cannot use a custom image which contains data disks to replace the system disk.

Procedure

1. Log on to the [ECS console](#).
2. Select the target region.
3. In the left-side navigation pane, click **Instances**.
4. Find the target instance and click **More > Disk and Image > Create Custom Image**.
5. Enter a name and description for the image.
6. Click **Create**.

Create Custom Image
✕

You can create a complete image template for the current ECS instance, including all its disks. A new snapshot will be taken for each instance disk and can be viewed in the snapshot list. You must wait for the snapshots for each disk to be created before the image can be used. Please be patient.

* Image Name:
 2 - 128 characters long. Do not start with a special character or a digit. It can contain the following special characters: "-", "_", and ".".

* Image Description:
 It must contain 2-256 characters and it cannot begin with http:// or https://

The image is available after all snapshots of all disks have been created.

Additional operation

See [create a custom image by using a snapshot](#).

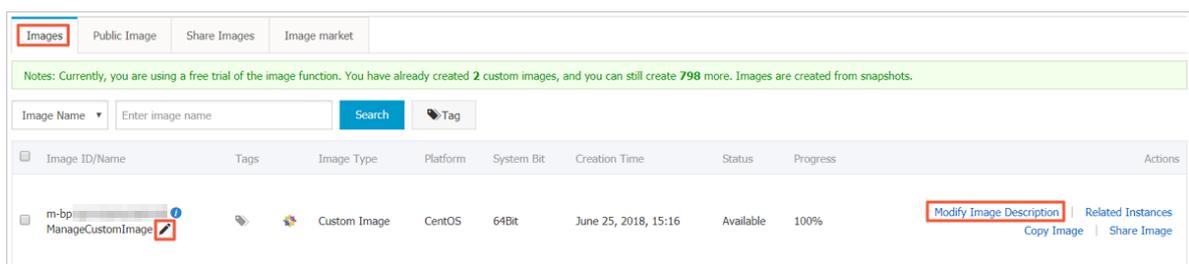
8.2 Manage custom images

You can modify the name and description of your custom images to help you organize and identify them, and you can delete custom images that you no longer require

Modify the name and description of a custom image

To modify the name and description of a custom image, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Snapshots and Images > Images**.
3. Select the target region.
4. Find the custom image to be edited and then click the  icon..
5. Enter a name for the custom image.



6. In the Actions column, click **Modify Description** and then, in the dialog box, enter a **Custom Image Description**.

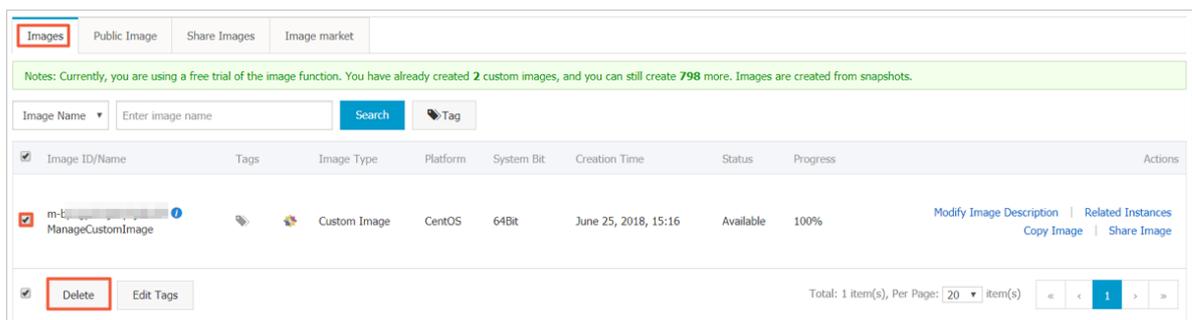
7. Click **Save**.

Alternatively, you can modify the name and description of a custom image by calling the ECS API [ModifyImageAttribute](#).

Delete custom images

To delete one or more custom images, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Snapshots and Images > Images**.
3. Select the target region.
4. Select one or more custom images that you want to delete, and then click **Delete**.



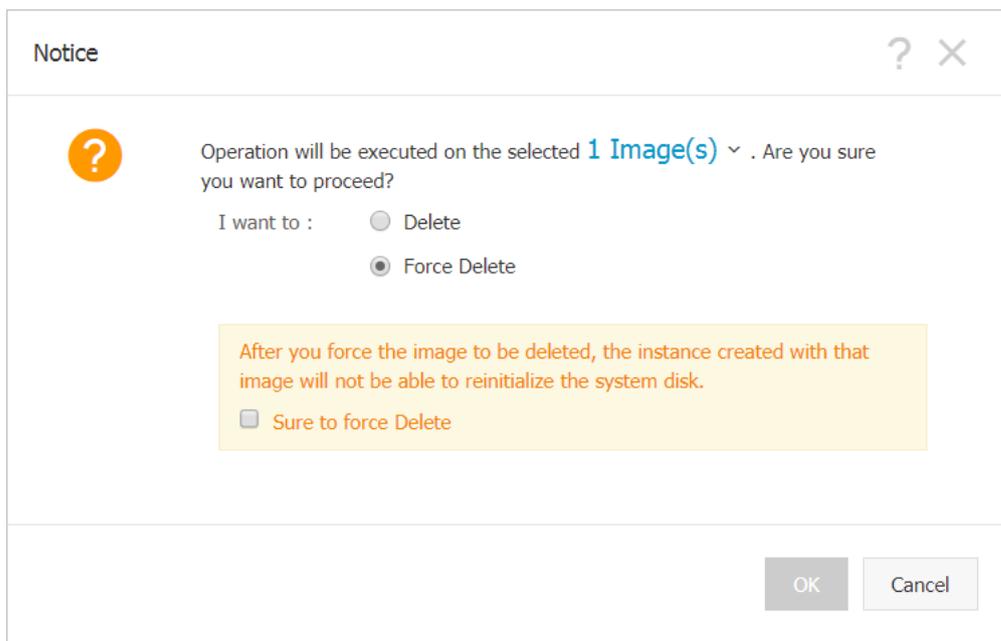
5. In the dialog box that appears, select the required method for deleting the custom images:

- **Delete:** The custom images are deleted normally.
- **Force Delete:** The custom images are deleted forcibly. Check **I confirm to forcibly Delete the selected instances**.



Note:

After you forcibly delete the custom images, [cloud disk reinitialization](#) of the instances that you have created from the images cannot be performed.



6. Click **OK**.

Alternatively, you can delete custom images by calling the ECS API [DeleteImage](#).

8.3 Copy images

Copying images allows you to deploy an application across regions that runs the same image environment. You can copy a custom image from one region to another. The task completion time depends on the network transfer speed and the number of concurrent tasks in the queue.

Attention

- Upon copying a custom image, a corresponding snapshot is created in the target region. The image is then created from that snapshot in the target region. As a result, fees are calculated due to data transfer between different regions. However, no fee is charged for such traffic. The billing policy is subject to change.
- After copying a custom image, an identical image is created in the target region. However, the relevant role and service authorization information is lost, which is also true for previously configured [user data](#).

Procedure

To copy images in the ECS console, follow these steps:

1. Log on to the [ECS console](#).
2. Select the target region.
3. In the left-side navigation pane, select **Snapshots and Images > Images**.

4. Select the custom image you want to copy. Note that **Type** must be **Custom Images**. Then, in the **Actions** column, click **Copy Image**.

**Note:**

If your custom image is larger than 200 GiB, when you click **Copy Image** you will be directed to open a ticket to complete the operation.

5. In the **Copy Image** dialog box, you can find the ID of the selected image. Complete the following configurations:
 - a. Select the **Target Region**.
 - b. Enter the **Custom Image Name** and **Custom Image Description** that are shown in the target region.
 - c. Click **OK**.
6. Switch to the target region and check the progress. When 100% is displayed, the image is copied successfully.

**Note:**

If **Progress** is not 100%, **Status** is **Creating**. In this case, you can click **Cancel Copy** to cancel the operation. After the operation is canceled, the image information is removed from the target region.

Image ID/Name	Tags	Image Type	Platform	System Bit	Creation Time	Status	Progress	Actions
copyImageTest		Custom Image	CentOS	64Bit	July 5, 2018, 15:20	Creating	43%	Modify Image Description Related Instances Cancel Copy Share Image

You can also use the ECS APIs [CopyImage](#) and [CancelCopyImage](#) to perform the operation.

Additional operations

When an image's status is **Available**, you can use it to [create an instance](#) or [change the system disk](#).

You can also view all copied snapshot in the target region.

FAQ

[FAQs about copying images](#)

8.4 Share images

After creating a custom image, you can share it with other Alibaba Cloud users. Shared images help new users adapt to ECS faster as they can quickly create ECS instances and set up business environments based on your custom images. Moreover, shared images do not consume the image quota of the account from which an image is shared.

Attention

You can only share custom images you have created, not custom images created and shared by other users. Each custom image can be shared with up to 50 users within the same Alibaba Cloud region. That is, an image cannot be shared across regions.

Before sharing an image, make sure that all sensitive data and files have been deleted from the image.



Note:

The integrity or security of shared images is not guaranteed. Make sure that you use only images shared by trusted accounts before using shared images. Besides, you shall bear the risk on your own. After you create an instance based on a shared image, be sure to [connect the instance](#) to check the integrity and security of the image.

Sharing image restrictions

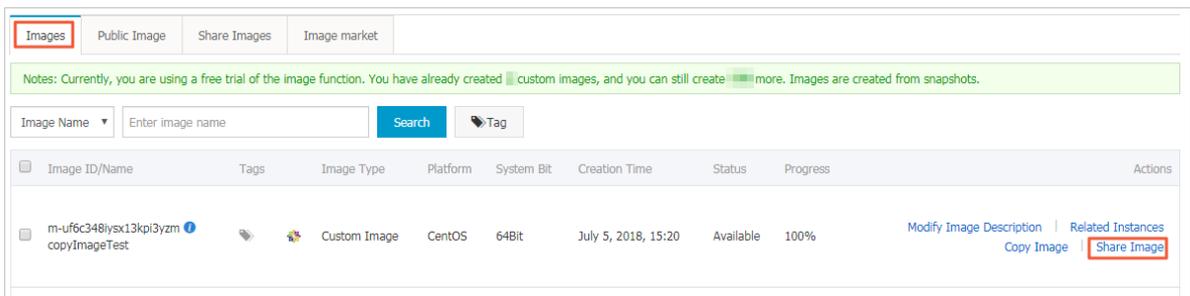
If your custom image has been shared with other accounts, you must remove all the sharing relationships for that image before you can delete the image. After deleting a shared custom image:

- Users who are using the shared image will no longer be able to find the image through the ECS console or ECS API, nor can they use the image to create ECS instances or replace system disks.
- ECS instances that are created from the shared image cannot re-initialize their system disks.

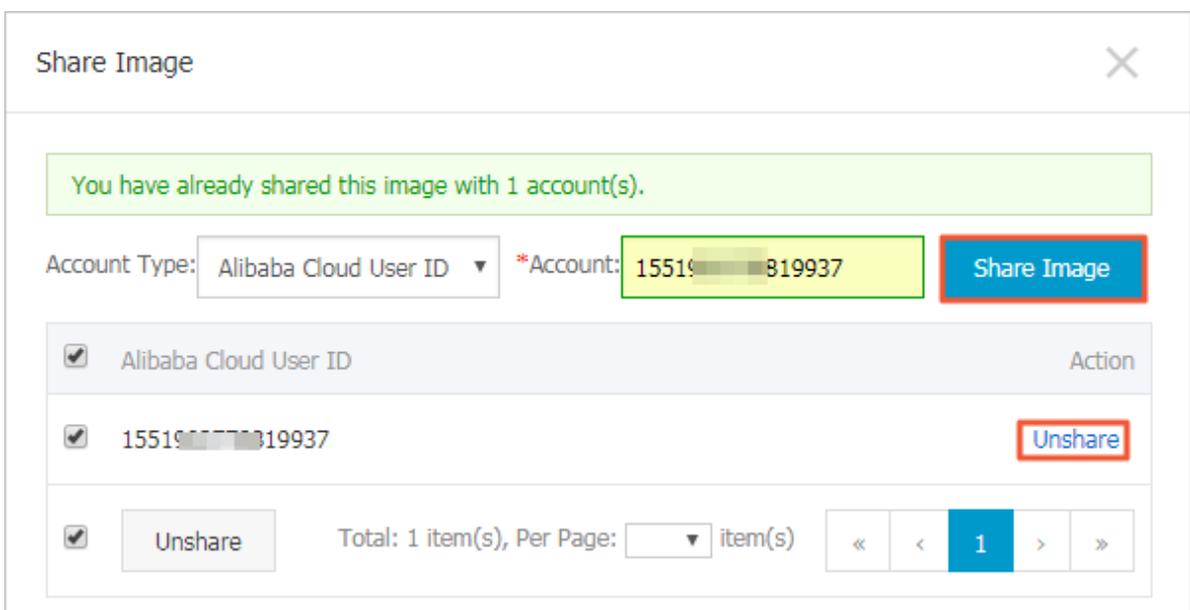
Share an image

To share an image in the ECS console, follow these steps:

1. Log on to the [ECS console](#).
2. Select the target region.
3. In the left-side navigation pane, choose **Snapshots and Images > Images**.
4. Select the target **Custom Image** the, in the **Actions** column, click **Share Image**.



- In the pop-up dialog box, select **Alibaba Cloud Account ID** in the **Account Type** drop-down list. Then, enter the account ID that you want to share the image with in the **Account** box. For more information, see [Appendix. How to get the account ID](#).



Note:

If you want to stop sharing the image with an account, click **Unshare** next to the account. After you cancel the sharing, that account will be unable to query and use the image. This means that if that account has already created an instance by using this shared image, the instance will be unable to *re-initialize the system disk*.

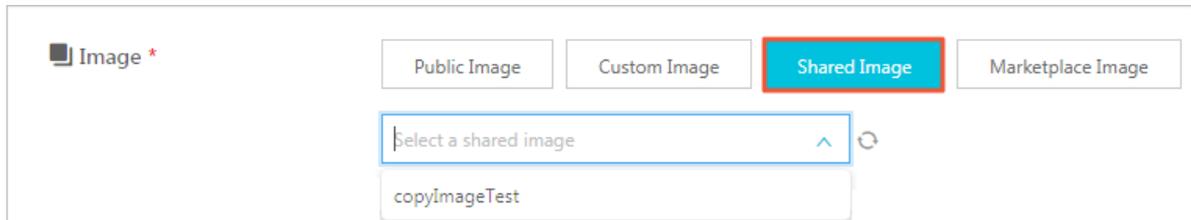
- (Optional) For the accounts with whom you share an image, these accounts can view the shared image in **Snapshots and Images > Images > Share Image** in the same region in the ECS console.

You can also use the ECS APIs [ModifyImageSharePermission](#) and [DescribeImageSharePermission](#) to share an image.

Next steps

After an image is shared with other users, they can use it to create one or more instances.

1. Log on to the [ECS console](#).
2. Create one or more instances by referring to [Step 2. Create an instance](#) Create an instance in [Quick Start](#). Note that you should select **Shared Image** during the procedure.

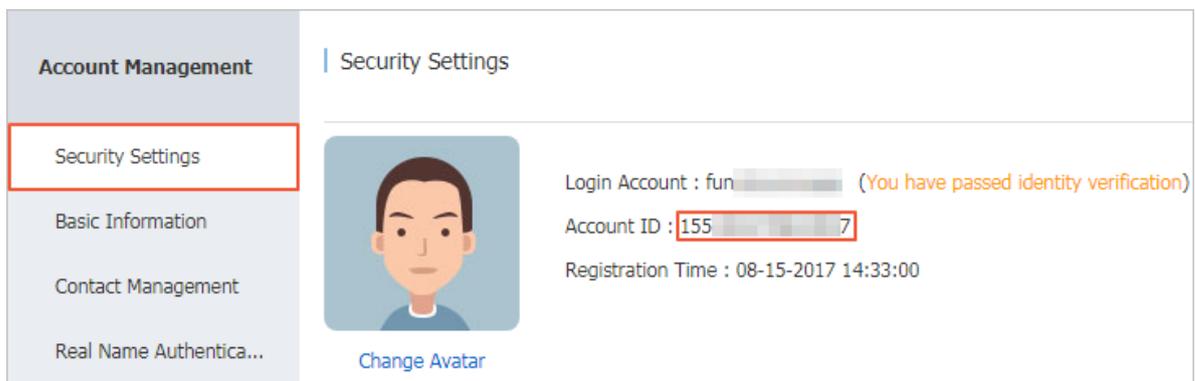


They can also use the shared image to [Replace the system disk \(non-public image\)](#) for instances.

Appendix. How to get the account ID?

To find your account ID, follow these steps:

1. Log on to the ECS console.
2. Hover your mouse over your avatar and then click **Security Settings** from the account menu.
3. On the page that appears, the account ID is displayed at the right as follows.



8.5 Import images

8.5.1 Image compliance tool

ECS allows you to create instances from imported custom images. Imported custom images can be created based on your offline server, virtual machine, or a cloud host on any cloud platform. The images you import must meet certain requirements. For more information, see [Notes for importing images](#). To reduce the time required for creating images and instances, we recommend that you use the **image compliance tool** of ECS (referenced in this document as **compliance tool**) to create images that comply with the relevant standards. The compliance tool can detect non-compliance of various configuration indicators and locations based on a given server environment, generate TXT and JSON detection reports, and offer possible solutions.

Limits

The compliance tool currently supports Linux images only, such as Ubuntu, CentOS, Debian, RedHat, SUSE Linux Enterprise Server (SLES), OpenSUSE, FreeBSD, CoreOS, and other Linux versions.

Sample

The following example use a CentOS 7.4 64-bit server.

1. Log on to your server, virtual machine, or cloud host.
2. [Download](#) the compliance tool.
3. Run `image_check` with root permissions to guarantee that the compliance tool can read configuration files under permission control.

```
chmod +x image_check
sudo image_check -p [destination path]
```



Note:

You can use `-p [destination path]` to specify the path where detection reports are generated. If you do not set this parameter, reports are generated in the compliance tool path by default.

4. Wait for the compliance tool to detect the system configuration.

```
Begin check your system...
The report is generating.
-----
The information you need to enter when you import your image to the
Alibaba Cloud website:
Current system: CentOS # System information 1: Server operating
system
Architecture: x86_64 # System information 2: System architecture
System disk size: 42 GB # System information 3: Server system disk
capacity
-----
# Detection item
Check driver [ OK ]
Check shadow file authority [ OK ]
Check security [ OK ]
Check qemu-ga [ OK ]
Check network [ OK ]
Check ssh [ OK ]
Check firewall [ OK ]
Check filesystem [ OK ]
Check device id [ OK ]
Check root account [ OK ]
Check password [ OK ]
Check partition table [ OK ]
Check lvm [ FAILED ]
Check lib [ OK ]
```

```

Check disk size [ OK ]
Check disk use rate [ WARNING ]
Check inode use rate [ OK ]
-----
15 items are OK
1 items are failed
1 items are warning
-----
The report is generated: /root/image_check_report_2018-05-14_18-18-10.txt
Please read the report to check the details

```

5. View the detection report. The report is generated in the format of `image_check_report_date_time.txt` or `image_check_report.json`.

Detection items

The compliance tool detects the following server configuration items to ensure that the ECS instances created from your custom image are fully functional.

Detection item	Non-compliance issue	Suggestion
driver	The ECS instance cannot start normally.	Install a virtualization driver. For example, install a virtio driver
/etc/shadow	You cannot modify the password file, so you cannot create an ECS instance from the custom image.	Do not use the <code>chattr</code> command to lock the <code>/etc/shadow</code> file.
SELinux	The ECS instance cannot start normally.	Do not modify <code>/etc/selinux/config</code> to start SELinux.
qemu-ga	Some of the services required by ECS are unavailable, and the instance is not fully functional.	Uninstall <code>qemu-ga</code> .
network	Network functions of the ECS instance are unstable.	Disable or delete the Network Manager and enable the network service. For the latest Linux versions, we recommend that you use the Network Manager and run the <code>nmcli</code> command to configure the network.
ssh	You cannot connect to the ECS instance from the console.	Enable the SSH service and do not set <code>PermitRootLogin</code> .
firewall	The system does not automatically configure your ECS instance environment.	Disable the firewall <code>iptables</code> , <code>firewalld</code> , <code>IPFilter (IPF)</code> , <code>IPFireWall (IPFW)</code> , or <code>PacketFilter (PF)</code> .
file system	You cannot resize the disk .	The XFS, Ext3, and Ext4 file systems are used, and the Ext2, UFS, and

Detection item	Non-compliance issue	Suggestion
		UDF file systems are allowed. The Ext4 file system does not support 64-bit features.
root	You cannot use your username and password to remotely connect to the ECS instance.	Reserve the root account.
passwd	You cannot add users to the ECS instance.	Retain or reinstall the passwd command.
Partition table	The ECS instance cannot start normally.	Use MBR partitioning.
Logical Volume Manager (LVM)	The ECS instance cannot start normally.	Switch to another partitioning service.
/lib	The ECS instance cannot be automatically configured.	The /lib and /lib64 files cannot be stored in absolute paths. Modify the storage paths of /lib and /lib64 to their relative paths.
system disk	N/A	Increase the system disk capacity. The optimal system disk capacity is 40 GiB to 500 GiB. When you import images, configure the system disk capacity based on the virtual file size of images, instead of the usage capacity of images.
disk_usage	You cannot install the necessary drivers or services for the ECS instance.	Make sure that sufficient disk space is available.
inode usage	You cannot install the necessary drivers or services for the ECS instance.	Make sure that sufficient inode resources are available.

The compliance tool provides a detection result `OK`, `FAILED`, or `WARNING` based on detection items.

- `OK`: The detection items all comply with requirements.
- `FAILED`: The detection items do not comply with requirements, which means a ECS instance created from the custom image cannot start normally. We recommend that you rectify the non-compliant items and recreate the image to improve instance startup efficiency.

- **WARNING:** The detection items do not comply with requirements, which means an ECS instance created from the custom image can start normally, but ECS cannot use valid methods to configure your instance. You can choose to immediately rectify the non-compliant items or temporarily retain the items and create an image.

Output items

The compliance tool provides detection reports in both TXT and JSON formats after it detects the system environment. You can use `-p [destination path]` to specify the path where detection reports are generated. If you do not specify this parameter, reports are generated in the compliance tool path by default.

- Reports in TXT format are named `image_check_report_date_time.txt`. The reports include server configuration information and detection results. The following example uses a CentOS 7.4 64-bit server.

```
The information you need to input when you import your image to
Alibaba Cloud Website:
Current system is: CentOS #Server operating system
Architecture: x86_64 #System architecture
System disk size: 42 GB #Server system disk capacity
-----
Check driver #Detection item name
Pass: kvm drive is exist #Detection result
Alibaba Cloud supports kvm virtualization technology
We strongly recommend installing kvm driver.
```

- Reports in JSON format are named `image_check_report.json`. The reports include server configuration information and detection results. The following example uses a CentOS 7.4 64-bit server.

```
"platform": "CentOS", \\Server operating system
"os_big_version": "7", \\Operating system version number (major)
"os_small_version": "4", \\Operating system version number (minor)
"architecture": "x86_64", \\System architecture
"system_disk_size": "42", \\Server system disk capacity
"version": "1.0.2", \\Compliance tool version
"time": "2018-05-14_19-18-10", \\Detection time
"check_items": [{
  "name": "driver", \\Detection item name
  "result": "OK", \\Detection result
  "error_code": "0", \\Error code
  "description": "Pass: kvm driver exists.", \\Description
  "comment": "Alibaba Cloud supports kvm virtualization
technology. We strongly recommend installing kvm driver."
}]
}
```

What to do next

1. View the [notes for importing images](#).

2. [Install the virtio driver.](#)
3. (Optional) [Convert the image file format.](#)
4. [Import custom images.](#)
5. [Create an instance from a custom image.](#)

8.5.2 Notes for importing images

To guarantee the usability of an imported image and improve the importing efficiency, the following considerations must be noted before importing an image:

Depending on the operating system, the notes vary for [Windows images](#) and [Linux images](#).

Windows images

Considerations

- Verify the integrity of the file system before importing images for Windows.
- Check that there is adequate space on the system disk for the image to be installed.
- Disable the firewall and allow access to RDP port 3389.
- The logon password for the administrator account must be 8-30 characters in length and can contain letters, numbers, and the following special characters () ` ~ ! @ # \$ % ^ & * - + = | { } [] ; : ' < > , . ? /
- Configure the system disk size for the importing based on the virtual disk size rather than the usage of the image. The size of the disk to be used for the image import must be a minimum of 40 GiB, and cannot exceed 500 GiB.
- Do not modify critical system files.

What are supported

- Multi-partition system disks.
- NTFS file systems and MBR partitions.
- Images in RAW, qcow2, or VHD format.



Note:

If you want to import an image in another format, you need to [convert image file format](#) before importing it. We recommended that you convert the format to VHD because it offers smaller transmission capacity.

- Images with the following operating system versions can be imported:

— Microsoft Windows Server 2016

- Microsoft Windows Server 2012 R2 (standard edition)
- Microsoft Windows Server 2012 (standard edition and data center edition)
- Microsoft Windows Server 2008 R2 (standard edition, data center edition, and enterprise edition)
- Microsoft Windows Server 2008 (standard edition, data center edition, and enterprise edition)
- Microsoft Windows Server 2003 with Service Pack 1 (SP1) (standard edition, data center edition, and enterprise edition) or higher

What are not supported

- The installation of qemu-ga in an image is not supported because some services needed by ECS will become unavailable.
- Windows XP, Windows 7 (professional and enterprise editions), Windows 8, and Windows 10.

Linux images

Considerations

- Verify the integrity of the file system before importing images for Linux.
- Check that there is adequate space on the system disk for the image to be installed.
- Disable the firewall and allow access to TCP port 22.
- Install the virtualization platform XEN or KVM drives.
- We recommended that you [install cloud-init](#), so as to guarantee that hostname, NTP, and yum sources can be configured successfully.
- Dynamic Host Configuration Protocol (DHCP) needs to be enabled.
- The logon password for the root account must be 8-30 characters long and must contain uppercase/lowercase letters, numbers, and special characters simultaneously. The special characters can be: () ` ~ ! @ # \$ % ^ & * - + = | { } [] ; ' < > , . ? /
- Do not modify critical system files, such as `/sbin`, `/bin`, and `/lib*`.

What are supported

- Images in RAW, qcow2, or VHD format.



Note:

If you want to import an image in another format, you need to [convert image file format](#) before importing it. We recommended that you convert the format to VHD because it offers smaller transmission capacity.

- The xfs, ext3, and ext4 file systems and MBR partitions.

**Note:**

The ext4 file system cannot include the `64bit` feature. Moreover, the `project` and `quota` features cannot appear in pairs. You can run the command `tune2fs -l <ext4 file system directory> | grep features` to view the features included in the ext4 file system.

What are not supported

- Multiple network interfaces.
- IPv6 addresses.
- System disk partitions cannot be adjusted. Currently, only a single root partition is supported.

Non-standard image usage notes

Depending on whether the Linux system image you are importing is a standard platform image, the following issues must be noted.

- Official operating system releases are defined as *standard platform images*. Currently, supported system releases include Aliyun Linux, CentOS 5/6/7, CoreOS 681.2.0+, Debian 6/7, FreeBSD, OpenSUSE 13.1, RedHat, Red Hat Enterprise Linux (RHEL), SUSE Linux 10/11/12, and Ubuntu 10/12/13/14.
- Operating system images that are not listed as public images provided by ECS are *non-standard platform images*. Such images, though based on the standard operating system, do not comply with the requirements for a standard operating system regarding critical system configuration files, basic system environments, and applications. If you want to use a non-standard platform image, you can only choose the following when importing an image:
 - Other Linux: Alibaba Cloud identifies all of these images as other Linux systems. Alibaba Cloud does not handle the instances created if you import an image of Other Linux type. If you enable DHCP before creating an image, Alibaba Cloud automatically configures your network. After creating the instance, you need to connect to the instance by using the [Management Terminal](#) feature in the console, and then manually configure the IP address, router, and password.
 - Customized Linux: Customized images. After importing a customized Linux image, configure the network and password of the instance according to the standard system configuration mode of Alibaba Cloud. For more information, see [customize Linux images](#).

Item	Standard platform image	Non-standard platform image
Requirements for critical system configuration files	<ul style="list-style-type: none"> Do not modify <code>/etc/issue*</code>. Otherwise, ECS cannot properly identify the system release, leading to system creation failure. Do not modify <code>/boot/grub/menu.lst</code>, or the ECS instance cannot be started. Do not modify <code>/etc/fstab</code>, or the exception partition cannot be loaded, leading to ECS instance start failure. Do not change <code>/etc/shadow</code> to read only, or you may be unable to modify the password file, leading to system creation failure. Do not enable SELinux by modifying <code>/etc/selinux/config</code>, or the system may fail to start. 	Does not meet the requirements of standard platform images
Requirements for applications	Do not install <code>qemu-ga</code> in an imported image, or some services required by Alibaba Cloud may become unavailable.	Does not meet the requirements of standard platform images

8.5.3 Install cloud-init for Linux images

To guarantee successful configuration of the hostname, NTP source, and yum source of an imported Linux image, we recommend that you install cloud-init in your on-premises server, virtual machine, or cloud host, before import.

Limitations

- Currently, cloud-init supports the following Linux OSs: CentOS, Debian, Fedora, FreeBSD, Gentoo, RHEL (Red Hat Enterprise Linux), SLES (SUSE Linux Enterprise Server), and Ubuntu.
- If cloud-init has been installed on your on-premises server, VM, or cloud host, make sure that the cloud-init version is 0.7.9 or later because images using an earlier-version cloud-init may cause instance configuration failure in NTP, hostname, and yum. To check your version of cloud-init, follow these steps:
 1. Log on to your on-premise server, VM, or cloud host.
 2. Run `cloud-init --version` to query the cloud-init version.

If your version is earlier than 0.7.9, see how to [install cloud-init](#).

Prerequisites

You have installed the following software on your on-premise server, VM, or cloud host:

- git: Source code package of cloud-init

Command for installing yum: `yum install git`.

- Python2.7: Basis of cloud-init installation and running.

Command for installing yum: `yum install python`

- pip: Python library on which cloud-init installation depends.

Command for installing yum: `yum install python-pip`

yum installation is used in the following example. If you manage packages using `zypper` or `apt-get`, the installation methods are similar to `yum`.

Install cloud-init

To install cloud-init, follow these steps:

1. Log on to your on-premise server, VM, or cloud host.
2. Run `git clone https://git.launchpad.net/cloud-init` to download the cloud-init source code package.
3. Run `cd cloud-init` to go to the cloud-init directory.
4. Run `python setup.py install` to install `setup.py`, which is the cloud-init installation file.
5. Run `vi /etc/cloud/cloud.cfg` to modify the `cloud.cfg` configuration file.

```
# The top level settings are used as module
# and system configuration.

# A set of users which may be applied and/or used by various modules
# when a 'default' entry is found it will reference the 'default_user'
# from the distro configuration specified below
users:
  - default

# If this is set, 'root' will not be able to ssh in and they
# will get a message to login instead as the default $user
disable_root: true

# This will cause the set+update hostname module to not operate (if true)
preserve_hostname: false

# Example datasource config
# datasource:
#   Ec2:
#     metadata_urls: [ 'blah.com' ]
#     timeout: 5 # (defaults to 50 seconds)
#     max_wait: 10 # (defaults to 120 seconds)

# The modules that run in the 'init' stage
cloud_init_modules:
```

Modify `cloud_init_modules` configurations to the following:

```
# Example datasource config
# The top level settings are used as module
# and system configuration.
# A set of users which may be applied and/or used by various
modules
# when a 'default' entry is found it will reference the 'default_us
er'
# from the distro configuration specified below
users:
  - default
user:
  name: root
  lock_passwd: False
# If this is set, 'root' will not be able to ssh in and they
# will get a message to logon instead as the above $user
disable_root: false
# This will cause the set+update hostname module to not operate (if
true)
preserve_hostname: false
syslog_fix_perms: root:root
datasource_list: [ AliYun ]
# Example datasource config
datasource:
  AliYun:
    support_xen: false
    timeout: 5 # (defaults to 50 seconds)
    max_wait: 60 # (defaults to 120 seconds)
#   metadata_urls: [ 'blah.com' ]
# The modules that run in the 'init' stage
cloud_init_modules:
```

Troubleshooting

The libraries missing from images may be different. You can install the missing libraries through `pip` and repeat [step 4](#).

When the `six` and `oauthlib` libraries are missing

- If the following message is displayed during installation, the `six` library is missing from Python.

You can run `pip install six` to install the `six` library through `pip`.

```
File "/root/cloud-init/cloudinit/log.py", line 19, in <module>
  import six
ImportError: No module named s )
```

- If the following message is displayed during installation, the `oauthlib` library is missing from Python. You can run `pip install oauthlib` to install the `six` library through `pip`.

```
File "/root/cloud-init/cloudinit/url_helper.py", line 20, in <module>
>
  import oauthlib.oauth1 as oauth1
ImportError: No module named oauthlib.oaut )
```

When an error message does not indicate missing libraries

You can run `pip install -r requirements.txt` to install all dependency libraries according to the library information in the requirements.txt file of cloud-init.

What to do next

You can [import the image to ECS](#).

Reference

cloud-init [Alibaba Cloud \(AliYun\)](#)

8.5.4 Install virtio driver

This topic details which images do and do not require the virtio driver to be installed on the source server before import.

Images requiring no manual installation

After you [import custom images](#), if the operating systems of your images are listed as follows, Alibaba Cloud automatically processes the virtio driver for you:

- Windows Server 2008
- Windows Server 2012
- Windows Server 2016
- CentOS 6/7
- Ubuntu 12/14/16
- Debian 7/8/9
- SUSE 11/12

You can skip to recover the temporary root file system of initramfs or initrd.

Images requiring manual installation

For Linux images that are not included in the preceding list, you must install the virtio driver on-premises before importing the images.

To check the availability of virtio driver on a server

1. Run `grep -i virtio /boot/config-$(uname -r)` to inspect whether the virtio driver is already built in the kernel of your server.

```
[root@izbp1lcnsefoj0kcvvddt1z ~]# grep -i virtio /boot/config-$(uname -r)
CONFIG_VIRTIO_VSOCKETS=m
CONFIG_VIRTIO_VSOCKETS_COMMON=m
CONFIG_VIRTIO_BLK=m
CONFIG_SCSI_VIRTIO=m
CONFIG_VIRTIO_NET=m
CONFIG_VIRTIO_CONSOLE=m
CONFIG_HW_RANDOM_VIRTIO=m
CONFIG_DRM_VIRTIO_GPU=m
CONFIG_VIRTIO=m
# Virtio drivers
CONFIG_VIRTIO_PCI=m
CONFIG_VIRTIO_PCI_LEGACY=y
CONFIG_VIRTIO_BALLOON=m
CONFIG_VIRTIO_INPUT=m
# CONFIG_VIRTIO_MMIO is not set
```

**Note:**

- If VIRTIO_BLK and VIRTIO_NET do not exist in the output, the virtio driver is not built in the kernel. You must install and configure the virtio driver on your server [to compile and install virtio driver](#).
- If the values of parameter CONFIG_VIRTIO_BLK and parameter CONFIG_VIRTIO_NET are y, the virtio driver is already built in the kernel. For more information, see [notes for importing images](#) and [import custom images](#).
- If the values of parameter CONFIG_VIRTIO_BLK and parameter CONFIG_VIRTIO_NET are m, continue to step 2.

2. Run `lsinitrd /boot/initramfs-$(uname -r).img | grep virtio` to make sure the virtio driver has been compiled in the temporary root file system of initramfs or initrd.

```
[root@izbp1lcnsefoj0kcvvddt1z ~]# lsinitrd /boot/initramfs-$(uname -r).img | grep virtio
Arguments: -f --add-drivers ' xen-blkfront xen-blkfront virtio_blk virtio_pci virtio_console virtio_console'
-rw-r--r-- 1 root root 7628 Sep 13 07:14 usr/lib/modules/3.10.0-693.2.2.el7.x86_64/kernel/drivers/block/virtio_blk.ko.xz
-rw-r--r-- 1 root root 12820 Sep 13 07:15 usr/lib/modules/3.10.0-693.2.2.el7.x86_64/kernel/drivers/char/virtio_console.ko.xz
-rw-r--r-- 1 root root 7980 Sep 13 07:16 usr/lib/modules/3.10.0-693.2.2.el7.x86_64/kernel/drivers/scsi/virtio_scsi.ko.xz
drwxr-xr-x 2 root root 0 Oct 24 14:09 usr/lib/modules/3.10.0-693.2.2.el7.x86_64/kernel/drivers/virtio
-rw-r--r-- 1 root root 4340 Sep 13 07:16 usr/lib/modules/3.10.0-693.2.2.el7.x86_64/kernel/drivers/virtio/virtio.ko.xz
-rw-r--r-- 1 root root 9480 Sep 13 07:16 usr/lib/modules/3.10.0-693.2.2.el7.x86_64/kernel/drivers/virtio/virtio_pci.ko.xz
-rw-r--r-- 1 root root 8136 Sep 13 07:16 usr/lib/modules/3.10.0-693.2.2.el7.x86_64/kernel/drivers/virtio/virtio_ring.ko.xz
[root@izbp1lcnsefoj0kcvvddt1z ~]#
```

**Note:**

- According to the preceding figure, the virtio_blk driver, including its dependency virtio.ko, virtio_pci.ko and virtio_ring.ko, has been compiled in the temporary root file system initramfs. For more information, see [notes for importing images](#) and [import custom images](#).
- If virtio driver is unavailable in the initramfs, you must recover the temporary root file system of initramfs or initrd before importing images or migration.

To recover the temporary root file system

If the virtio driver is supported by the kernel but not compiled in the temporary root file system, you must recover the temporary root file system. The following example uses CentOS:

- CentOS/RedHat 5

```
mkinitrd -f --allow-missing \
  --with=xen-vbd --preload=xen-vbd \
  --with=xen-platform-pci --preload=xen-platform-pci \
  --with=virtio_blk --preload=virtio_blk \
  --with=virtio_pci --preload=virtio_pci \
  --with=virtio_console --preload=virtio_console \
```

- CentOS/RedHat 6/7

```
mkinitrd -f --allow-missing \
  --with=xen-blkfront --preload=xen-blkfront \
  --with=virtio_blk --preload=virtio_blk \
  --with=virtio_pci --preload=virtio_pci \
  --with=virtio_console --preload=virtio_console \
  /boot/initramfs-$(uname -r).img $(uname -r)
```

- Debian/Ubuntu

```
echo -e 'xen-blkfront\nvirtio_blk\nvirtio_pci\nvirtio_console' >> \
  /etc/initramfs-tools/modules
mkinitramfs -o /boot/initrd.img-$(uname -r) "
```

To compile and install virtio driver

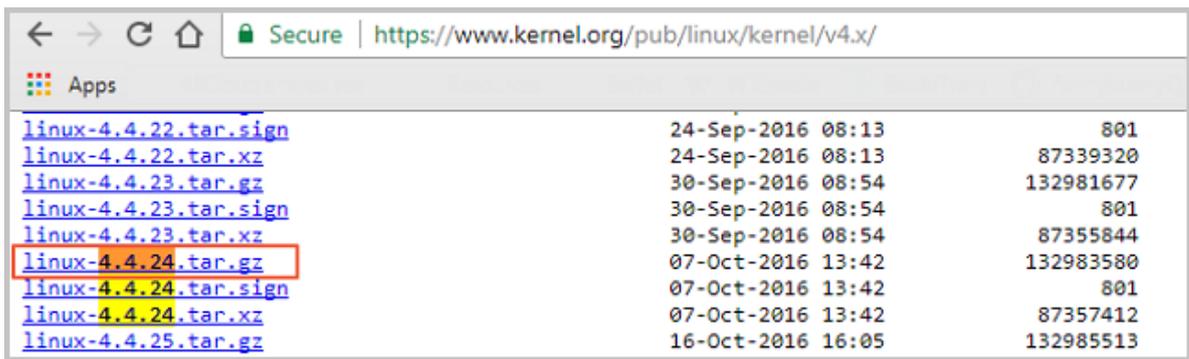
The following example uses a Red Hat server:

To download the kernel package

1. Run `yum install -y ncurses-devel gcc make wget` to install necessary components to compile the kernel.
2. Run `uname -r` to query the kernel version of your server, such as 4.4.24-2.a17.x86_64.

```
[root@i2bp1127hr3wi6p2cq9lnb2 ~]# uname -r
4.4.24-2.a17.x86_64
```

3. Visit [published Linux Kernel Archives](#) to download the source codes of kernel, for example, the download link of kernel version starting with 4.4.24 is <https://www.kernel.org/pub/linux/kernel/v4.x/linux-4.4.24.tar.gz>.



File Name	Date	Size
linux-4.4.22.tar.sign	24-Sep-2016 08:13	801
linux-4.4.22.tar.xz	24-Sep-2016 08:13	87339320
linux-4.4.23.tar.gz	30-Sep-2016 08:54	132981677
linux-4.4.23.tar.sign	30-Sep-2016 08:54	801
linux-4.4.23.tar.xz	30-Sep-2016 08:54	87355844
linux-4.4.24.tar.gz	07-Oct-2016 13:42	132983580
linux-4.4.24.tar.sign	07-Oct-2016 13:42	801
linux-4.4.24.tar.xz	07-Oct-2016 13:42	87357412
linux-4.4.25.tar.gz	16-Oct-2016 16:05	132985513

4. Run `cd /usr/src/` to change the directory.
5. Run `wget https://www.kernel.org/pub/linux/kernel/v4.x/linux-4.4.24.tar.gz` to download the installation package.
6. Run `tar -xzf linux-4.4.24.tar.gz` to decompress the package.
7. Run `ln -s linux-4.4.24 linux` to establish a link.
8. Run `cd /usr/src/linux` to change the directory.

To compile the kernel

1. Run the following commands to compile the driver into the kernel.

```
make mrproper
symvers_path=$(find /usr/src/ -name "Module.symvers")
test -f $symvers_path && cp $symvers_path .
cp /boot/config-$(uname -r) ./config
make menuconfig
```

2. Configure the corresponding settings of virtio driver in the following windows:



Note:

Select `*` to build the driver in the kernel, select `m` to compile it as a module.

- a. Press the space bar to select Virtualization.

```

Linux Kernel Configuration
Arrow keys navigate the menu. <Enter> selects submenus --->. Highlighted letters are hotkeys. Press
Press <Esc><Esc> to exit, <?> for Help, </> for Search. Legend: [*] built-in [ ] excluded <M> modul

General setup --->
[*] Enable loadable module support --->
-* Enable the block layer --->
Processor type and features --->
Power management and ACPI options --->
Bus options (PCI etc.) --->
Executable file formats / Emulations --->
-* Networking support --->
Device Drivers --->
Firmware Drivers --->
File systems --->
Kernel hacking --->
Security options --->
-* Cryptographic API --->
[ ] Virtualization --->
Library routines --->
---
Load an Alternate Configuration File
Save an Alternate Configuration File

```

Make sure that you have selected the option of KVM (Kernel-based Virtual Machine).

```

Virtualization
Arrow keys navigate the menu. <Enter> selects submenus --->. Highlighted letters are hotkeys. Pres
Press <Esc><Esc> to exit, <?> for Help, </> for Search. Legend: [*] built-in [ ] excluded <M> modu

--- Virtualization
< > Kernel-based Virtual Machine (KVM) support
< > KVM for Intel processors support
< > KVM for AMD processors support
< > PCI driver for virtio devices (EXPERIMENTAL)
< > Virtio balloon driver (EXPERIMENTAL)

```

```

Processor type and features --->
  [*] Paravirtualized guest support --->
    --- Paravirtualized guest support
    (128) Maximum allowed size of a domain in gigabytes
  [*] KVM paravirtualized clock

```

```
[*] KVM Guest support
```

```

Paravirtualized guest support
Arrow keys navigate the menu. <Enter> selects submenus --->. Highlighted letters are hotkeys. Pressing
Press <Esc><Esc> to exit, <?> for Help, </> for Search. Legend: [*] built-in [ ] excluded <M> module

--> Paravirtualized guest support
[*] Xen guest support
(128) Maximum allowed size of a domain in gigabytes
[*] Enable Xen debug and tuning parameters in debugfs
[*] KVM paravirtualized clock
[*] KVM Guest support
-- Enable paravirtualization code
[ ] Paravirtualization layer for spinlocks

```

```

Device Drivers --->
[*] Block devices --->
<M> Virtio block driver (EXPERIMENTAL)
-- Network device support --->
<M> Virtio network driver (EXPERIMENTAL)

```

- b. Press the Esc key to exit the kernel configuration windows, and save changes to file `.config` according to the dialog box.
- c. Inspect whether all the corresponding settings of virtio driver has been correctly configured or not.
- d. (Optional) If no configuration of virtio driver is settled after the inspect, run the following commands to edit the file `.config` manually.

```

make oldconfig
make prepare
make scripts
make
make install

```

- e. Run the following commands to check whether the virtio driver is installed.

```

find /lib/modules/"$(uname -r)"/ -name "virtio.*" | grep -E "
virtio.*"
grep -E "virtio.*" < /lib/modules/"$(uname -r)"/modules.builtin

```



Note:

If any of the output includes `virtio_blk` and `virtio_pci.virtio_console`, your server has correctly installed the virtio driver.

What to do next

After compiling the virtio driver, you can [migrate your server to Alibaba Cloud by using the Cloud Migration Tool](#).

8.5.5 Customize Linux images

If your selected OS is not supported by Alibaba Cloud, and cloud-init cannot be installed, you can select **Customized Linux** when importing a customized image. Alibaba Cloud will then regard the customized Linux image as an unrecognized OS type (that is, it lacks necessary standard configuration information for ECS instance start for the first time). In this case, you need to add a parsing script to the customized image before importing the image, so as to facilitate automatic configuration of the instance at the first start.

Limitations

- The first partition of the customized Linux image must be writable.
- The first partition type of the customized Linux image must be FAT32, EXT2, EXT3, EXT4, or UFS.
- The size of the virtual file of the customized Linux image must be larger than 5 GiB.
- Security requirements for customized Linux images are as follows:
 - There is no high-risk vulnerability that can be remotely exploited.
 - When you log on to an instance for the first time through the [Management Terminal](#) of the ECS console, you must change the initial default password (if there is any) before performing subsequent actions.
 - There is no default SSH private key pair. The initial SSH private key pair must be randomly generated by Alibaba Cloud.

Procedure

1. Create the `aliyun_custom_image` directory in the root directory of the first image partition.

When the instance that is created using the customized Linux image is started for the first time, Alibaba Cloud will write configuration information into the `os.conf` file in the `aliyun_custom_image` directory. Alibaba Cloud will automatically create an `os.conf` file if there is none.

2. Create a parsing script in the image to parse system configurations of the `os.conf` file. For details about how to write a script, see [attentions on script parsing](#) and [parsing script example](#).

Example of the `os.conf` file

For instances using classic networks

```
hostname=iz23r29djmjz
password=cXdlcjEyMzQK
eth0_ip_addr=10.171.254.123
eth0_mac_addr=00:8c:fa:5e:14:23
eth0_netmask=255.255.255.0
eth0_gateway=10.171.254.1
```

```
eth0_route="10.0.0.0/8 10.171.254.1;172.16.0.0/12 10.171.254.1"
eth1_ip_addr=42.120.74.105
eth1_mac_addr=00:8c:fa:5e:14:24
eth1_netmask=255.255.255.0
eth1_gateway=42.120.74.1
eth1_route="0.0.0.0/0 42.120.74.1"
dns_nameserver="7.7.7.7 8.8.8.8"
```

The following table describes the parameters.

Parameter	Parameter description
hostname	The host name
password	The password, which is Base64-encoded
eth0_ip_addr	The IP address of the eth0 NIC
eth0_mac_addr	The MAC address of the eth0 NIC
eth0_netmask	The network mask of the eth0 NIC
eth0_gateway	The default gateway of the eth0 NIC
eth0_route	The eth0 intranet route list, in which routes are separated by semicolons (;) by default
eth1_ip_addr	The IP address of the eth1 NIC
eth1_mac_addr	The MAC address of the eth1 NIC
eth1_netmask	The network mask of the eth1 NIC
eth1_gateway	The default gateway of the eth1 NIC
eth1_route	The eth1 internet route list, in which routes are separated by semicolons (;) by default
dns_nameserver	The DNS address list, in which addresses are separated by spaces by default

For instances using VPCs

```
hostname=iZ23r29djmjZ
password=cXd1cjEyMzQK
eth0_ip_addr=10.171.254.123
eth0_mac_addr=00:8c:fa:5e:14:23
eth0_netmask=255.255.255.0
eth0_gateway=10.171.254.1
eth0_route="0.0.0.0/0 10.171.254.1"
dns_nameserver="7.7.7.7 8.8.8.8"
```

The following table describes the parameters.

Parameter	Parameter description
hostname	The host name
password	The password, which is Base64-encoded
eth0_ip_addr	The IP address of the eth0 NIC
eth0_mac_addr	The MAC address of the eth0 NIC
eth0_netmask	The network mask of the eth0 NIC
eth0_gateway	The default gateway of the eth0 NIC
eth0_route	The eth0 intranet route list, in which routes are separated by semicolons (;) by default
dns_nameserver	The DNS address list, in which addresses are separated by spaces by default

Script parsing considerations

In normal cases, when an instance is started for the first time, Alibaba Cloud automatically writes information about configuration items into the `os.conf` file in the `aliyun_custom_image` directory in the root directory of the first partition. To configure a customized Linux image, you must create a pre-defined parsing script in the image. Then, Alibaba Cloud reads configuration information about the instance from the `os.conf` file to complete instance configuration. The following conditions must be met for script parsing:

- **Automatic start:** The parsing script should be automatically started. To do so, place the script in the `/etc/init.d/` directory.
- **Configuration item value rules:** As described in [example of the os.conf file](#), instances using classic networks and those using VPCs differ in rules of the number of configuration items and values of some configuration items.
- **Configuration file read path:** By default, names of the devices allocated for the first partition vary with types of the instances created for the customized Linux image, including I/O optimization instances and non-I/O optimization instances. Therefore, you are recommended to use `uuid` or `label` to indicate devices in the first partition. Because the user password is a Base64-encoded string, it therefore must be Base64-encoded in the script.
- **Network type:** When using the parsing script to determine the network type, you can check whether there is `eth1_route` or other eth1-related configuration item. To do so, parse and process the instance accordingly after determining whether it uses a classic network or VPC.

- Instances using VPCs are configured with Internet routes that are specified by the `eth0_route` parameter in the `os.conf` file.
- Instances using classic networks are configured with Internet routes that are specified by the `eth1_route` parameter in the `os.conf` file, and intranet routes are specified by the `eth0_route` parameter.
- **Configuration optimization:** Configurations in the `os.conf` file are executed only once during the instance life cycle. You are recommended to delete the `os.conf` file after the parsing script is successfully executed. The parsing script does not execute configurations in the `os.conf` file if it does not read any.
- **Customized image processing:** When you create a customized image based on the customized Linux image, the script requiring automatic start is also included in the new image. Alibaba Cloud will write `os.conf` file configurations when the instance is started for the first time. Then, the parsing script immediately executes the configurations upon detection.
- **Configuration change processing:** When instance configurations are changed through the Alibaba Cloud console or APIs, Alibaba Cloud writes related information into the `os.conf` file. Then, the parsing script executes the configurations again to issue the changes.

Parsing script example

The following uses a parsing script used for CentOS as an example. You can change the script content as needed. Make sure that the script has been successfully debugged in the image before you use the script.

```
#!/bin/bash

### BEGIN INIT INFO
# Provides:          os-conf
# Required-Start:    $local_fs $network $named $remote_fs
# Required-Stop:
# Should-Stop:
# Default-Start:     2 3 4 5
# Default-Stop:      0 1 6
# Short-Description: The initial os-conf job, config the system.
### END INIT INFO

first_partition_dir='/boot/'
os_conf_dir=${first_partition_dir}/aliyun_custom_image
os_conf_file=${os_conf_dir}/os.conf

load_os_conf() {
    if [[ -f $os_conf_file ]]; then
        . $os_conf_file
        return 0
    else
        return 1
    fi
}
```

```

}

cleanup() {
    # ensure $os_conf_file is deleted, to avoid repeating config system
    rm $os_conf_file >& /dev/null
    # ensure $os_conf_dir is existst
    mkdir -p $os_conf_dir
}

config_password() {
    if [[ -n $password ]]; then
        password=$(echo $password | base64 -d)
        if [[ $? == 0 && -n $password ]]; then
            echo "root:$password" | chpasswd
        fi
    fi
}

config_hostname() {
    if [[ -n $hostname ]]; then
        sed -i "s/^HOSTNAME=. */HOSTNAME=$hostname/" /etc/sysconfig/network
        hostname $hostname
    fi
}

config_dns() {
    if [[ -n $dns_nameserver ]]; then
        dns_conf=/etc/resolv.conf
        sed -i '/^nameserver.*\/d' $dns_conf
        for i in $dns_nameserver; do
            echo "nameserver $i" >> $dns_conf
        done
    fi
}

is_classic_network() {
    # vpc: eth0
    # classic: eth0 eth1
    grep -q 'eth1' $os_conf_file
}

config_network() {
    /etc/init.d/network stop
    config_interface eth0 ${eth0_ip_addr} ${eth0_netmask} ${eth0_mac_addr}
}

config_route eth0 ${eth0_route}
if is_classic_network ; then
    config_interface eth1 ${eth1_ip_addr} ${eth1_netmask} ${eth1_mac_addr}
    config_route eth1 ${eth1_route}
fi
/etc/init.d/network start
}

config_interface() {
    local interface=$1
    local ip=$2
    local netmask=$3
    local mac=$4
    interface_cfg="/etc/sysconfig/network-scripts/ifcfg-${interface}"
    cat << EOF > $interface_cfg
DEVICE=$interface

```

```

IPADDR=$ip
NETMASK=$netmask
HWADDR=$mac
ONBOOT=yes
BOOTPROTO=static
EOF
}

config_default_gateway() {
    local gateway=$1
    sed -i "s/^GATEWAY=. */GATEWAY=$gateway/" /etc/sysconfig/network
}

config_route() {
    local interface=$1
    local route=$2
    route_conf=/etc/sysconfig/network-scripts/route-`${interface}
    > $route_conf
    echo $route | sed 's;/;/\n/' | \
    while read line; do
        dst=$(echo $line | awk '{print $1}')
        gw=$(echo $line | awk '{print $2}')
        if ! grep -q "$dst" $route_conf 2> /dev/null; then
            echo "$dst via $gw dev $interface" >> $route_conf
        fi
        if [[ "$dst" == "0.0.0.0/0" ]]; then
            config_default_gateway $gw
        fi
    done
}

##### sysvinit service portal #####

start() {
    if load_os_conf ; then
        config_password
        config_network
        config_hostname
        config_dns
        cleanup
        return 0
    else
        echo "not load $os_conf_file"
        return 0
    fi
}

RETVAL=0

case "$1" in
    start)
        start
        RETVAL=$?
        ;;
    *)
        echo "Usage: $0 {start}"
        RETVAL=3
        ;;
esac

```

```
exit $RETVAL
```

8.5.6 Convert image file format

Only image files in qcow2, RAW, or VHD format can be imported. If you want to import images in other formats, you need to convert the format before importing the image. This topic describes how to use the `qemu-img` tool to convert other image file formats to VHD or RAW. Using `qemu-img`, you can convert RAW, qcow2, VMDK, VDI, VHD (vpc), VHDX, qcow1, or QED, to VHD, or implement conversion between RAW and VHD.

Windows

To install `qemu-img` and convert the image file format, follow these steps:

1. Log on to your server or VM, download [qemu-img](#) and complete the installation. Installation path: `C:\Program Files\qemu`.
2. Perform the following actions to create an environment variable for `qemu-img`:
 - a. Choose **Start > Computer**, then right click **Properties**.
 - b. In the left-side navigation pane, click **Advanced System Settings**.
 - c. In the **System Properties** dialog box, click the **Advanced** tab, and then click **Environment Variables**.
 - d. In the **Environment Variables** dialog box, find the **Path** variable in the **System Variables** part, and then click **Edit**. If the **Path** variable does not exist, click **New**.
 - e. Add a system variable value:
 - In the case of **Edit System Variable**: In the **Variable Value** field, add `C:\Program Files\qemu`. Different variable values are separated with semicolon (;).
 - In the case of **New System Variable**: In the **Variable Name** field, enter `Path`. In the **Variable Value** field, enter `C:\Program Files\qemu`.
3. Open **Command Prompt** in Windows and run the `qemu-img --help` command. If the result is displayed correctly, the environment variable is configured successfully.
4. In the **Command prompt**, run the `cd [directory of the source image file]` command to change the directory. For example, `cd D:\ConvertImage`.
5. Run the `qemu-img convert -f qcow2 -O raw centos.qcow2 centos.raw` command to convert the image file format. Where:
 - `-f` is followed by the source image format.

- `-o` (uppercase is required) is followed by the converted image format, the source file name, and the target file name.

When the conversion is complete, the target file appears in the directory where the source image file is located.

Linux

To install `qemu-img` and convert the image file format, follow these steps:

1. Install `qemu-img`, for example:

- For Ubuntu, run the command: `apt install qemu-img`.
- For CentOS, run the command: `yum install qemu-img`.

2. Run the `qemu-img convert -f qcow2 -O raw centos.qcow2 centos.raw` command to convert the image file format. Where:

- `-f` is followed by the source image format.
- `-o` (uppercase is required) is followed by the converted image format, the source file name, and the target file name.

When the conversion is complete, the target file appears in the directory where the source image file is located.

Troubleshooting

If errors occur during `qemu-img` installation and there are no clear prompts about the missing dependent libraries, run `pip install -r requirements.txt` to install all the dependent libraries based on the libraries shown in the file `requirements.txt` of `cloud-init`.

Next step

[Import custom images](#)

8.5.7 Import custom images

You can import on-premises image files to your ECS environment to create ECS instances or change system disks



Note:

- The time it takes to import an image depends on the size of the image file and the number of concurrent tasks.

- When you import an image, a snapshot is automatically generated. You can view the snapshot information on the **Snapshots** page in the ECS Console. Before the import image task is completed, the status of the snapshot is displayed as **Failed**. When the task is completed, the status is automatically updated to **Successful**. The snapshot capacity is the size of the imported image file, regardless of the system disk size that was set when the image was imported.

Prerequisites

Before importing an image, we recommend that you:

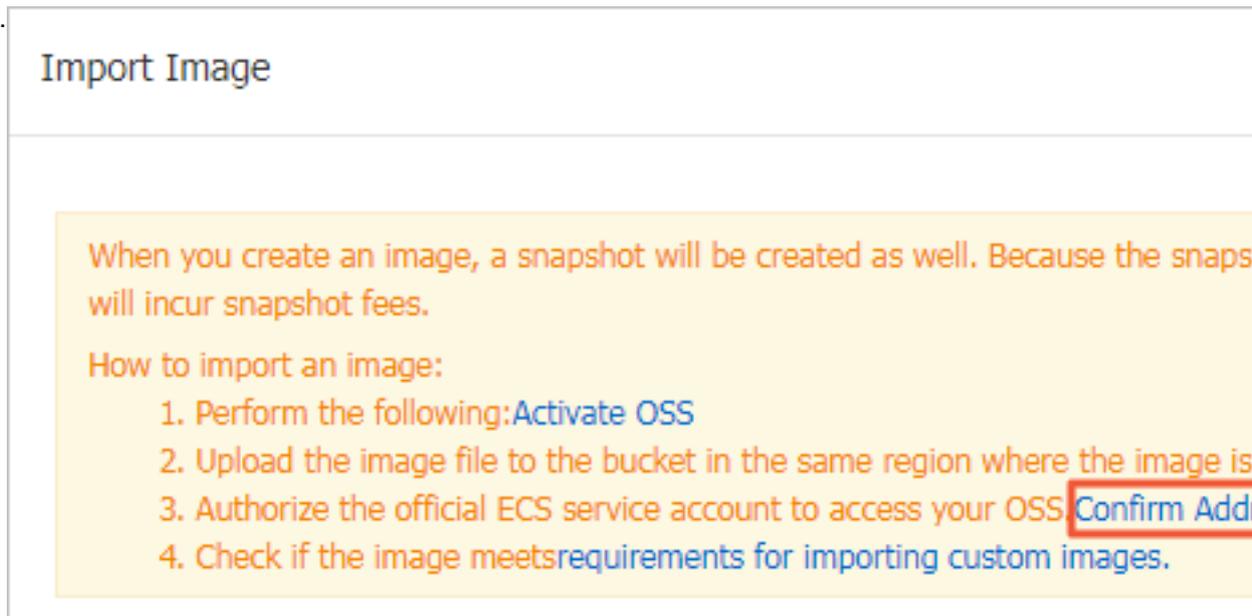
- Review the [notes for importing images](#), [customize Linux images](#), and [convert image format](#) to understand the limitations of importing an on-premises image.
- [Activate OSS](#).
- (Optional) If you are using a RAM sub-account, you need to contact the master account in advance to obtain the permission for the [AliyunECSImageImportDefaultRole](#) role.

Procedure

To import custom images in the ECS console, follow these steps:

1. Use an OSS third-party client, OSS API or OSS SDK to upload the prepared custom image. If the file you want to upload is larger than 5 GiB, see [multipart upload](#).
2. Log on to the [ECS console](#).
3. In the left-side navigation pane, choose **Snapshots and Images > Images**.
4. Click **Import Image**.

5. In the **Import Image** dialog box, click **Confirm Address** as follows.



6. In the **Cloud Resource Access Authorization** window, select `AliyunECSImageImportDefaultRole` and `AliyunECSExportDefaultRole`, then click **Confirm Authorization Policy** to allow the ECS service to access your OSS resources.
7. On the Images page, click **Import Image** again.
8. In the **Import Image** dialog box, enter the following information:
- **Region of Image:** Select the region where the OSS Bucket of the image file to upload is located.
 - **OSS Object Address:** Copy the object address of the image file from the OSS console. For more information, see [download an object](#).
 - **Image Name:** Enter a name for the custom image. The name must be 2 to 128 characters in length and can contain letters, numbers, Chinese characters, periods (.), underscores (_), colons (:), and hyphens (-).
 - **Operating System:** Select **Windows** or **Linux**, that is, the same as that of your image. If you want to import a non-standard platform image, select Linux.
 - **System Disk Size:** The system disk size, which ranges from 40 GiB to 500 GiB.
 - **System Architecture:** Choose **x86_64** for 64 bit operating systems and choose **i386** for 32 bit operating systems.
 - **Platform:** The options depend on the **Operating System** you chose.
 - Windows: Windows Server 2003, Windows Server 2008, and Windows Server 2012.

- Linux: Centos, SUSE, Ubuntu, Debian, FreeBSD, CoreOS, Aliyun, Customized Linux, and Others Linux ([open a ticket](#) to confirm the selected edition is supported).
 - If your image OS is a custom edition developed from Linux kernel, [open a ticket](#) to contact us.
 - **Image Format:** Supports qcow2, RAW, and VHD. Qcow2 or VHD is recommended.
 - **Image Description:** Enter a description of the custom image.
 - **Add Images of Data Disks:** Choose this option if you want to import an image that contains data disks. Supported data disk capacity ranges from 5 GiB to 2,000 GiB.
9. Click **OK**.
- 10.(Optional) You can view the task progress in the image list of the import region. Before the task is completed, you can find the imported custom image through [Tasks](#) management, and, if needed, cancel the import task.

You can also use the ECS API [ImportImage](#) to import a custom image.

Next step

[Create an instance from a custom image.](#)

References

- [Custom images FAQ](#)
- [Create and import on-premise images by using Packer](#)

8.6 Export custom images

You can export custom images for on-premises testing or for Apsara stack environments.



Note:

- The time it takes to export an image depends on the size of the image file and the number of export tasks in the queue.
- Exported images are stored in your [OSS](#) bucket. This means you are billed for the OSS storage and download traffic. For more information, see OSS [billing items](#).

Limitations

Currently, the image export function has the following limitations:

- You cannot export custom images that are created by a system disk snapshot from the [Alibaba Cloud Marketplace](#).

- You can export the custom images that contain four snapshots of data disks at most, and for a single data disk, the maximum volume must be no greater than 500 GiB.
- When using exported custom images to [create an instance by using the wizard](#), you must confirm that the file device recorded in `/etc/fstab` corresponds to the exported data disk snapshot information.

Prerequisites

Before exporting a custom image, you must:

- [Open a ticket](#) to activate the image export feature, and describe the use cases of the exported images in the ticket.
- Activate OSS and make sure that the region where your custom images are located has an available OSS bucket. For more information, see [create a bucket](#).

Procedure

To export a custom image in the ECS console, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, choose **Snapshot & Images > Images**.
3. Select the target region.
4. Find the custom image you want to export and then, in the **Actions** column, click **Export Image**.
 - a. In the **Export Image** dialog box, click **Conform Address**.
 - b. In the **Cloud Resource Access Authorization** window, click **Confirm Authorization Policy** to allow ECS to access your OSS resources.
5. Return to the ECS console homepage. In the **Actions** column of the **Images** page, click **Export Image** again.
6. In the **Export Image** dialog box:
 - Select the OSS bucket in the specified region.
 - Set the prefix of the object name of the exported image. For example, if you set Demo as the prefix, then the exported image file displayed in the OSS bucket is named Demo-[automatically generated file name].
7. Click **OK**.

8. (Optional) Cancel the image export task. Before the task is completed, you can go to the [Tasks](#) management page in the ECS console, find the relevant task in the specified region and cancel the task.

You can also use the ECS APIs [ExportImage](#) and [CancelTask](#) to perform the preceding operations.

Next steps

1. Log on to the [OSS console](#) to query the export result.



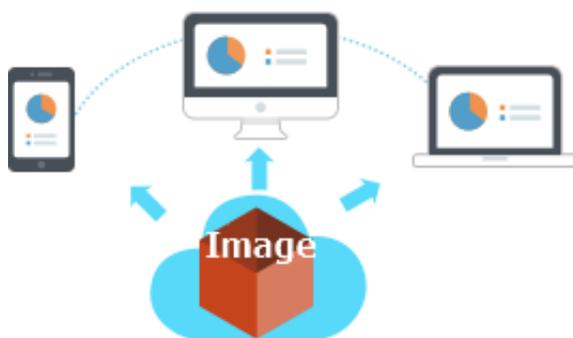
Note:

When an exported custom image contains a data disk snapshot, multiple files appear in your OSS. The file name with `system` indicates a system disk snapshot and the file name with `data` indicates a data disk snapshot. A data disk snapshot has an identifier corresponding to the data disk, which is the mount point of the data disk, such as `xvdb` or `xvdc`.

2. After the custom image is exported successful, [download the object](#) and then download the custom image file. The format of the image file is RAW by default.

8.7 Marketplace images

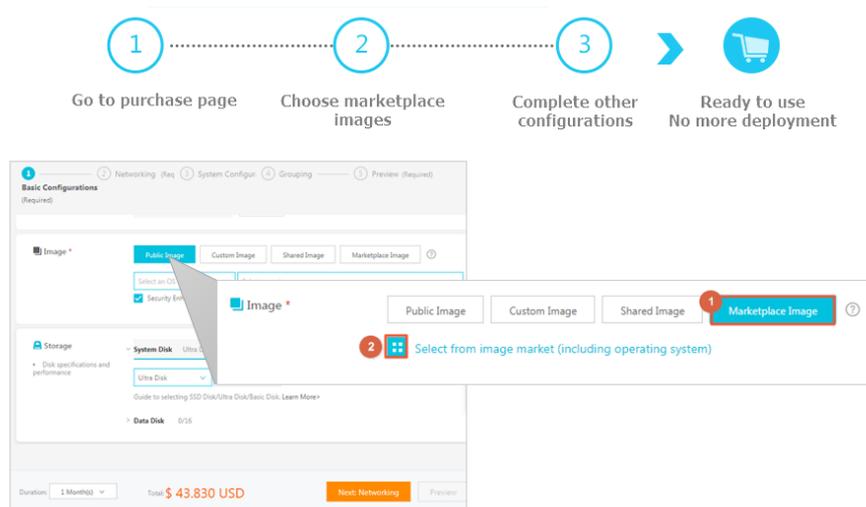
An Alibaba Cloud Marketplace image is equivalent to the installation disk for an Elastic Compute Service (ECS) instance. A Marketplace image allows you to quickly obtain a running environment for ECS instances and any pre-installed software applications. Such an image can be used for site deployment, application development, and visualized management. Marketplace images effectively allow ECS instances to be used out-of-the-box, helping to reduce costs.



Select a Marketplace image when creating an instance

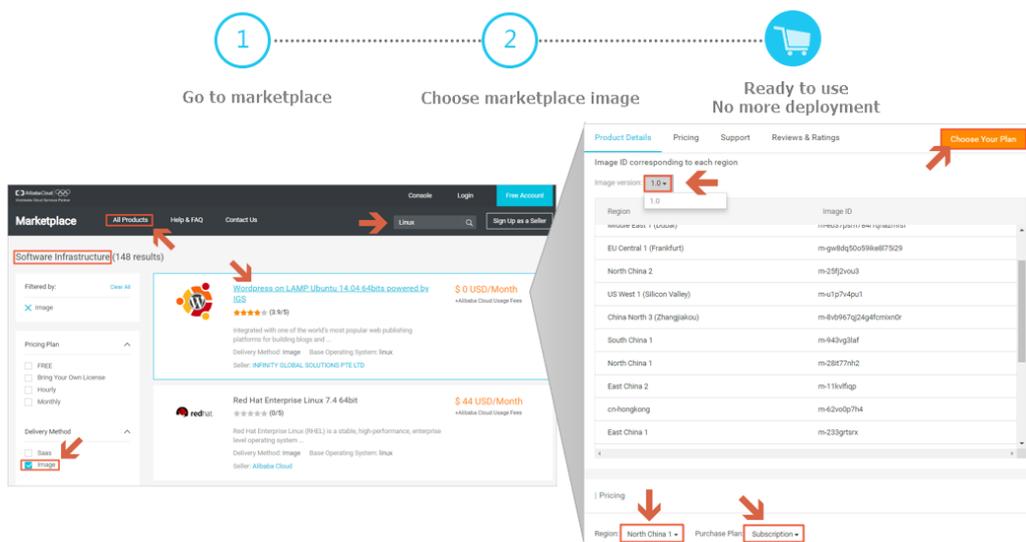
We recommend that you use a Marketplace image if you are new to working with ECS instances. To deploy a Marketplace image, follow these steps:

1. Go to the [ECS purchase](#) page.
2. Select and configure your image. For more information, see [create an Instance](#). Then, on the **Image** configuration page, choose **Marketplace Image > Select from image market (including operating system)**.



Purchase an image from Alibaba Cloud Marketplace and create an instance

1. Go to [Alibaba Cloud Marketplace](#).
2. Select the required image and click **Buy Now**.
3. You may be required to log on to the Alibaba Cloud console before proceeding.



4. Select and configure your image. For more information, see [create an instance](#).

Change the operating system by using the Marketplace image

If you have purchased ECS instances, use an image to deploy the running environment or install software applications as follows:



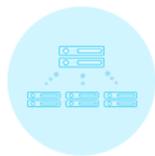
Note:

If you change the image, the data on the system disk will be lost. Therefore, we recommend that you back up your data before changing your operating system. For more information, see [Create snapshots](#).

1. Log on to the [ECS console](#).
2. Stop the target instance.
3. On the **Replace System Disk** page, select **Marketplace Image** in the **Image Type** setting. For more information, see [replace the system disk \(non-public image\)](#).



Go to ECS console



Stop the instance and
replace the system disk



Choose marketplace
images



Reset the ECS instance

8.8 Open source tools

8.8.1 Create and import on-premises images by using Packer

Packer is a convenient open-source tool to create on-premises image files. It runs on the most major operating systems.

By using Packer, you can create identical on-premises images for multiple platforms from a single source configuration. This topic details steps to create an on-premises image for CentOS 6.9 on an Ubuntu 16.04 server and to upload it to Alibaba Cloud. For actual scenarios, you can customize your Packer templates as required.

Prerequisites

- You must have an [AccessKey](#) for the configuration file.



Note:

Do not use the AccessKey of your Alibaba Cloud account. Instead, [create a RAM user](#) and use the RAM account to create the necessary [AccessKey](#) to maintain account security.

- You must [have purchased OSS](#).

Example of creating and importing an on-premises image

1. Run `egrep "(svm|vmx)" /proc/cpuinfo` to check whether your on-premises server or virtual machine supports KVM. If the following output returns, KVM is supported.

```
pat pse36 clflush dts acpi mmx fxsr sse sse2 ss ht tm pbe syscall nx
pdpelgb rdtscp lm constant_tsc art arch_perfmon pebs bts rep_good
nopl xtopology nonstop_tsc aperfmperf tsc_known_freq pni pclmulqdq
dtes64 monitor ds_cpl vmx est tm2 ssse3 sdbg fma cx16 xtpr pdcm
pcid sse4_1 sse4_2 x2apic movbe popcnt tsc_deadline_timer aes xsave
avx f16c rdrand lahf_lm abm 3dnowprefetch epb intel_pt tpr_shadow
vnmi flexpriority ept vpid fsgsbase tsc_adjust bmi1 avx2 smep bmi2
erms invpcid mpx rdseed adx smap clflushopt xsaveopt xsavec xgetbv1
xsaves dtherm ida arat pln pts hwp hwp_notify hwp_act_window hwp_epp
flags          : fpu vme de pse tsc msr pae mce cx8 apic sep mtrr pge
mca cmov
```

2. Run the following commands to install the KVM:

```
sudo apt-get install qemu-kvm qemu virt-manager virt-viewer libvirt-
bin bridge-utils # Install KVM and related dependencies.
sudo virt-manager # Enable virt-manager.
```

If a GUI runs in the VM console window, you have successfully installed the KVM.

3. Install Packer.

To install Packer, see [use Packer to create a custom image](#).

4. Run the following commands to define a Packer template.



Note:

The on-premises image created in the following configuration is for the CentOS 6.9 operating system only. To create images for other operating systems, [customize](#) the configuration file `centos.json` as required.

```
cd /user/local # Switch the directory.
wget https://raw.githubusercontent.com/alibaba/packer-provider/
master/examples/alicloud/local/centos.json # Download file centos.
json that is released by Alibaba Cloud.
wget https://raw.githubusercontent.com/alibaba/packer-provider/
master/examples/alicloud/local/http/centos-6.9/ks.cfg # Download
file ks.cfg that is released by Alibaba Cloud.
mkdir -p http/centos-6.9 # Create a directory.
```

```
mv ks.cfg http/centos-6.9/ # Move file ks.cfg to the http/centos-6.9
directory.
```

5. Run the following commands to create an on-premises image.

```
export ALICLOUD_ACCESS_KEY= SpecifyYourAccessKeyIDHere # Import your
AccessKeyID,
export ALICLOUD_SECRET_KEY= SpecifyYourAccessKeySecretHere # Import
your AccessKeySecret.
packer build centos.json # Create an on-premises image.
```

An example result is as follows.

```
qemu output will be in this color.
==> qemu: Downloading or copying ISO
      qemu: Downloading or copying: http://mirrors.aliyun.com/centos/
6.9/isos/x86_64/CentOS-6.9-x86_64-minimal.iso
.....
==> qemu: Running post-processor: alicloud-import
      qemu (alicloud-import): Deleting import source https://oss-cn-
beijing.aliyuncs.com/packer/centos_x86_64
Build 'qemu' finished.
==> Builds finished. The artifacts of successful builds are:
--> qemu: Alicloud images were created:
cn-beijing: XXXXXXXXX
```

6. Wait for a few minutes, log on to the [ECS console](#) and check your custom image in the image list that is in the corresponding region. In this sample, the region is China North 2 (cn-beijing).

Customize a Packer template

In this example, the following JSON file is customized based on the template used to create an image for the CentOS 6.9.

```
{ "variables": {
  "box_basename": "centos-6.9",
  "build_timestamp": "{{isotime \"20060102150405\"}}",
  "cpus": "1",
  "disk_size": "4096",
  "git_revision": "__unknown_git_revision__",
  "headless": "",
  "http_proxy": "{{env `http_proxy`}}",
  "https_proxy": "{{env `https_proxy`}}",
  "iso_checksum_type": "md5",
  "iso_checksum": "af4a1640c0c6f348c6c41f1ea9e192a2",
  "iso_name": "CentOS-6.9-x86_64-minimal.iso",
  "ks_path": "centos-6.9/ks.cfg",
  "memory": "512",
  "metadata": "floppy/dummy_metadata.json",
  "mirror": "http://mirrors.aliyun.com/centos",
  "mirror_directory": "6.9/isos/x86_64",
  "name": "centos-6.9",
  "no_proxy": "{{env `no_proxy`}}",
  "template": "centos-6.9-x86_64",
  "version": "2.1.TIMESTAMP"
},
  "builders": [
```

```

    {
      "boot_command": [
        "<tab> text ks=http://{{ .HTTPIP }}:{{ .HTTTPort }}/{{user `ks_path`}}<enter><wait>"
      ],
      "boot_wait": "10s",
      "disk_size": "{{user `disk_size`}}",
      "headless": "{{user `headless`}}",
      "http_directory": "http",
      "iso_checksum": "{{user `iso_checksum`}}",
      "iso_checksum_type": "{{user `iso_checksum_type`}}",
      "iso_url": "{{user `mirror`}}/{{user `mirror_directory`}}/{{user `iso_name`}}",
      "output_directory": "packer-{{user `template`}}-qemu",
      "shutdown_command": "echo 'vagrant'|sudo -S /sbin/halt -h -p",
    },
    {
      "ssh_password": "vagrant",
      "ssh_port": 22,
      "ssh_username": "root",
      "ssh_wait_timeout": "10000s",
      "type": "qemu",
      "vm_name": "{{user `template`}}.raw",
      "net_device": "virtio-net",
      "disk_interface": "virtio",
      "format": "raw"
    }
  ],
  "provisioners": [
    {
      "type": "shell",
      "inline": [
        "sleep 30",
        "yum install cloud-util cloud-init -y"
      ]
    }
  ],
  "post-processors": [
    {
      "type": "alicloud-import",
      "oss_bucket_name": "packer",
      "image_name": "packer_import",
      "image_os_type": "linux",
      "image_platform": "CentOS",
      "image_architecture": "x86_64",
      "image_system_size": "40",
      "region": "cn-beijing"
    }
  ]
}

```

Parameters in a Packer builder

QEMU builder is used in the preceding [example](#) to create a virtual machine image. Required parameters for the builder are as follows.

Parameter	Type	Description
iso_checksum	String	The checksum for the OS ISO file. Packer verifies this parameter before starting a virtual machine with the ISO attached. Make sure you specify at least one of the

Parameter	Type	Description
		<code>iso_checksum</code> or <code>iso_checksum_url</code> parameters. If you have specified the <code>iso_checksum</code> parameter, the <code>iso_checksum_url</code> parameter is automatically ignored.
<code>iso_checksum_type</code>	String	The type of the checksum specified in <code>iso_checksum</code> . Optional values: <ul style="list-style-type: none"> • <code>none</code>: If you specify <code>none</code> for <code>iso_checksum_type</code>, the checksumming is ignored. This value is not recommended. • <code>md5</code> • <code>sha1</code> • <code>sha256</code> • <code>sha512</code>
<code>iso_checksum_url</code>	String	A URL that points to a GNU or BSD style checksum file that contains the ISO file checksum of an operating system. It may come in either the GNU or BSD pattern. Make sure you specify either the <code>iso_checksum</code> or the <code>iso_checksum_url</code> parameter. If you specify the <code>iso_checksum</code> parameter, the <code>iso_checksum_url</code> parameter is automatically ignored.
<code>iso_url</code>	String	A URL that points to the ISO file, and contains the installation image. This URL may be an HTTP URL or a file path: <ul style="list-style-type: none"> • If it is an HTTP URL, Packer downloads the file from the HTTP link and caches the file for later. • If it is a file path to the IMG or QCOW2 file, QEMU directly starts the file. If you have the file path specified, set parameter <code>disk_image</code> to <code>true</code>.
<code>headless</code>	boolean	By default, Packer starts the virtual machine GUI to build a QEMU virtual machine. If you set <code>headless</code> to <code>True</code> , a virtual machine without any console is started.

For more information, see Packer [QEMU Builder](#).

Parameters in a Packer provisioner

The provisioner in the preceding [example](#) contains a Post-Processor module that enables automated upload of on-premises images to Alibaba Cloud. Required parameters for the provisioner are as follows:

Parameter	Type	Description
access_key	String	Your AccessKeyID. The AccessKey has a high privilege. We recommend that you first create a RAM user and use the RAM account to create an AccessKey to maintain security of your Alibaba Cloud account.
secret_key	String	Your AccessKeySecret. The AccessKey has a high privilege. We recommend that you first create a RAM user and use the RAM account to create an AccessKey to maintain security of your Alibaba Cloud account.
region	String	Select the region where you want to upload your on-premises image. In this example, the region is cn-beijing. For more information, see regions and zones .
image_name	String	The name of your on-premises image. The name is a string of 2 to 128 characters. It must begin with an English or a Chinese character. It can contain A-Z, a-z, Chinese characters, numbers, periods (.), colons (:), underscores (_), and hyphens (-).
oss_bucket_name	String	The OSS bucket name. If you specify a bucket name that does not exist, Packer creates a bucket automatically with the specified oss bucket name when uploading the image.
image_os_type	String	Image type. Optional values: <ul style="list-style-type: none"> linux windows
image_platform	String	Distribution of the image. For example, CentOS.
image_architecture	String	The instruction set architecture of the image. Optional values: <ul style="list-style-type: none"> i386 x86_64
format	String	Image format. Optional values: <ul style="list-style-type: none"> RAW VHD

For more information, see Packer [Alicloud Post-Processor](#).

Next step

You can use the created image to create an ECS instance. For more information, see [create an instance from a custom image](#).

References

- For more information about how to use Packer, see [Packer](#) documentation.
- For more information about release information, visit the Packer repository on GitHub [packer](#).
- For more information about Alibaba Cloud open source tools, visit the Alibaba Cloud repository on GitHub [opstools](#).
- For more information about Alibaba Cloud and Packer project, visit the Alibaba Cloud & Packer repositories on GitHub [packer-provider](#).
- For more information about configuration file ks.cfg, see [Anaconda Kickstart](#) .

8.8.2 Use Packer to create a custom image

This topic provides information about how to install and use Packer to create a custom image.

Prerequisites

You must have an AccessKey.



Note:

Do not use the AccessKey of your Alibaba Cloud account. Instead, [create a RAM user](#) and use the RAM account to create the necessary [AccessKey](#) to maintain account security.

Step 1. Install Packer

Go to the official [Packer download page](#) where you can choose required version of Packer for your operating system.

To install Packer on a Linux server

1. Connect and log on to the Linux server. If the server you want to connect to is an ECS Linux instance, see [connect to a Linux instance by using a password](#).
2. Run `cd /usr/local/bin` to go to the `/usr/local/bin` directory.



Note:

The `/usr/local/bin` directory is an environment variable directory. You can install Packer to this directory or another directory that has been added to the environment variable.

3. Run `wget https://releases.hashicorp.com/packer/1.1.1/packer_1.1.1_linux_amd64.zip` to download the Packer installer. You can visit the official [download page of Packer](#) to download installers for other versions of Packer.
4. Run `unzip packer_1.1.1_linux_amd64.zip` to unzip the package.
5. Run `packer -v` to verify Packer's installation status. If the Packer version number is returned, you have successfully installed Packer. If error **command not found** is returned, Packer has not been correctly installed.

To install Packer on a Windows server

The following examples uses Windows Server 2012 64-bit:

1. Connect and log on to the Windows server. If the server you want to connect to is an ECS Windows instance, see [connect to a Windows instance](#).
2. Open the official [download page of Packer](#) and select an appropriate Packer installer for 64-bit Windows.
3. Unzip the package to a specified directory and install Packer.
4. Define the directory for Packer in the PATH environment variable.
 - a. Open the **Control Panel**.
 - b. Select **All Control Panel Items > System > Advanced System Settings**.
 - c. Click **Environment Variable**.
 - d. Find **Path** in the system variable list.
 - e. Add the Packer installation directory to the **Variable Value**, such as `C:\Packer` as seen in this example. Separate multiple directories with half-width semicolons (;). Click **OK**.
5. Run `packer.exe -v` in CMD to verify Packer's installation status. If the Packer version number is returned, you have successfully installed Packer. If error **command not found** prompt is returned, Packer has not been correctly installed.

Step 2. Define a Packer template



Note:

To create a custom image by using Packer, you must first create a JSON format template file. In the template, you must specify the [Alibaba Cloud Image Builder](#) and [Provisioner](#) for the custom image to be created. Packer offers a diverse range of provisioners for you to choose from when configuring the content generation mode of the custom image. In the following JSON

file example, the [Shell](#) provisioner is used as an example to illustrate how to define a Packer template.

Create a JSON file named `alicloud` and paste the following content:

```
{
  "variables": {
    "access_key": "{{env `ALICLOUD_ACCESS_KEY`}}",
    "secret_key": "{{env `ALICLOUD_SECRET_KEY`}}"
  },
  "builders": [
    {
      "type": "alicloud-ecs",
      "access_key": "{{user `access_key`}}",
      "secret_key": "{{user `secret_key`}}",
      "region": "cn-beijing",
      "image_name": "packer_basic",
      "source_image": "centos_7_02_64_20G_alibase_20170818.vhd",
      "ssh_username": "root",
      "instance_type": "ecs.n1.tiny",
      "internet_charge_type": "PayByTraffic",
      "io_optimized": "true"
    }
  ],
  "provisioners": [
    {
      "type": "shell",
      "inline": [
        "sleep 30",
        "yum install redis.x86_64 -y"
      ]
    }
  ]
}
```



Note:

Customize the values of the following parameters according to your actual requirements.

Parameter	Description
<code>access_key</code>	Your AccessKey ID. For more details, see create an Accesskey .
<code>secret_key</code>	Your AccessKey Secret. For more information, see create an AccessKey .
<code>region</code>	The region of the temporary instance used to create the custom image.
<code>image_name</code>	The custom image's name
<code>source_image</code>	The name of the basic image name retrieved from Alibaba Cloud public image list.
<code>instance_type</code>	Type of temporary instance generated to create the custom image.

internet_charge_type	The Internet bandwidth billing method for the temporary instance generated for creating the custom image.
provisioners	Type of <i>Packer Provisioner</i> used for creating the custom image

Step 3. Create a custom image by using Packer

To specify the Packer template file and create a custom image, follow these steps:

1. Run `export ALICLOUD_ACCESS_KEY=your AccessKeyID` to import your AccessKey ID.
2. Run `export ALICLOUD_SECRET_KEY=your AccessKeySecret` to import your AccessKey Secret.
3. Run `packer build alicloud.json` to create the custom image.

The following example creates a custom image containing ApsaraDB for Redis and runs as follows:

```
alicloud-ecs output will be in this color.
==> alicloud-ecs: Prevalidating alicloud image name...
alicloud-ecs: Found image ID: centos_7_02_64_20G_alibase_20170818.vhd
==> alicloud-ecs: Start creating temporary keypair: packer_59e44f40-
c8d6-0ee3-7fd8-blba08ea94b8
==> alicloud-ecs: Start creating alicloud vpc
-----
==> alicloud-ecs: Provisioning with shell script: /var/folders/3q/
w38xx_js6cl6k5mwkrqsnw7w0000gn/T/packer-shell1257466182
alicloud-ecs: Loaded plugins: fastestmirror
-----
alicloud-ecs: Total                                     1.3
MB/s | 650 kB 00:00
alicloud-ecs: Running transaction check
-----
==> alicloud-ecs: Deleting temporary keypair...
Build 'alicloud-ecs' finished.
==> Builds finished. The artifacts of successful builds are:
--> alicloud-ecs: Alicloud images were created:
cn-beijing: m-2ze12578be1oa4ovs6r9
```

What to do next

Use this custom image to create an ECS instance. For more information, see [create an instance from a custom image](#).

References

- For more information, visit [packer-provider](#), the Packer repository of Alibaba Cloud Github.
- See the [Packer official documents](#) to learn more about how to use Packer.

9 Security groups

9.1 Typical applications of security group rules

This topic introduces some typical applications of security group rules that are applicable to instances in both classic and VPC networks.

Typical applications of security group rules include:

- [Use SSH to connect to Linux instances remotely](#)
- [Use RDP to connect to Windows instances remotely](#)
- [Ping ECS instances in public network](#)
- [Use ECS instances as Web servers](#)
- [Use FTP to upload or download files](#)

Use SSH to connect to Linux instances remotely

After you create a Linux ECS instance, you can use SSH to connect to the ECS instance remotely. In this scenario, you can add the following security group rules.

Network Types	Network Card Type	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
VPC network	No configuration required	Direction of entry	Allow	SSH (22)	22/22	Address segment access	0.0.0.0/0	1
Classic network	Alibaba Cloud							

Use RDP to connect to Windows instances remotely

After the Windows ECS instance has been created, use RDP to connect to the ECS instance remotely. In this scenario, you can add the following security group rules.

Network Type	Network Card Type	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
VPC network	No configuration required	Direction of entry	Allow	RDP (3389)	3389/3389	Address segment access	0.0.0.0/0	1

Classic network	Public network							
-----------------	----------------	--	--	--	--	--	--	--

Ping ECS instances in public network

After creating the ECS instance, use the ping command to test the communication status between the ECS instances. In this scenario, you can add the following security group rules.

Network Type	Network Card Type	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
VPC network	No configuration required	Direction of entry	Allow	ICMP	-1/-1	Address segment or security group access	Fill it in according to the license type. For details, see add security group rules .	1
Classic network	Public network							

Use ECS instances as Web servers

If you use your instance as a Web server, first install the Web server program on the instance, and then add the following security group rules.



Note:

You must start the Web server program, and verify if port 80 works properly.

Network Type	Network Card Type	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
VPC network	No configuration required	Direction of entry	Allow	HTTP (80)	80/80	Address segment access	0.0.0.0/0	1
Classic network	Public network							

If you cannot access your instance through the `http://public network IP address`, verify if TCP port 80 works properly.

Use FTP to upload or download files

To use FTP to upload/download files to/from the ECS instance, add the following security group rules.



Note:

You must install the FTP program on the instance, and verify if port 20/21 works properly.

Network Type	Network Card Type	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
VPC network	No configuration required	Direction of entry	Allow	Custom TCP	20/21	Address segment access	0.0.0.0/0	1
Classic network	Public network							

9.2 Scenarios

This topic details several scenarios of VPC-Connected and Classic network-connected security groups.



Note:

For more information about how to create a security group and corresponding rules, see [create a security group](#) and [add a security group rule](#).

- [Scenario 1: Enable intranet communication](#)

If you want to copy files between two Classic network-connected ECS instances owned by different accounts or in different security groups, you must first enable intranet communication between both instances by configuring security group rules before you can copy files.

- [Scenario 2: Allow remote connection from specified IP addresses only](#)

If your ECS instance is remotely accessed by hackers, you can modify the port for remote connection and then configure security group rules to allow access from specified IP addresses only.

- [Scenario 3: Allow an instance to access specified IP addresses only](#)

If the security of your ECS instance is compromised, you can configure security group rules to allow the instance access to specified IP addresses.

- [Scenario 4: Allow remote connection to an ECS instance](#)

If you need to remotely access your ECS instance, you can configure a corresponding security group rule.

- [Scenario 5: Allow access to an ECS instance over HTTP or HTTPS service](#)

If you build a website on your instance, you can configure security group rules to enable your users to access the website.

Scenario 1: Enable intranet communication

Security group rules can be used in the following cases to enable intranet communication between ECS instances that belong to different accounts or security groups in the same region:

- Case 1: Instances belong to the same region and to the same account.
- Case 2: Instances belong to the same region but to different accounts.



Note:

For VPC-connected ECS instances,

- If the instances are in one VPC, you can configure their security group rules to enable intranet communication.
- If the instances are in different VPCs, or owned by different accounts in the same region, Express Connect must be used to establish intranet communication. For more information, see [establish an intranet connection between VPCs under different accounts](#).

Case 1: Instances belong to one region and one account

For two instances in the same region and are owned by the same account, if they are in one security group, intranet communication is enabled by default. If they are in different security groups, you must configure security group rules to enable intranet communication according to the network types.

- VPC

If they are in one VPC, add a rule in each security group to authorize shared access between the security groups. The rule must be as follows.

NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Priority	Authorization Type	Authorization Object
N/A	Inbound	Allow	Select the required protocol	Set the required port range	1	Security group access (authorize this account)	Select the Security Group ID on which you want to allow access to the instance

- Classic network

Add a rule in each security group to authorize shared access between the security groups. The rule must be as follows.

NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Priority	Authorization Type	Authorization Object
Intranet	Inbound	Allow	Select the required protocol	Set the required port range	1	Security group access (authorize this account)	Select the Security Group ID on which you want to allow access to the instance

Case 2: Instances belong to one region but different accounts

The following information is for Classic network-connected ECS instances only.

Authorize shared access between security groups. For example:

- User A owns a Classic network-connected ECS instance in the China East 1 region, named Instance A, with the private IP address A.A.A.A. The security group is Group A.
- User B owns a Classic network-connected ECS instance in the China East 1 region, named Instance B, with the private IP address B.B.B.B. The security group is Group B.

- Add a rule in Group A to authorize access of Instance A to Instance B, as shown in the following table.

NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
Intranet	Inbound	Allow	Select the required protocol	Set the required port range	Security group access (authorize other accounts)	Type the account ID of User B and the security group ID of Group B	1

- Add a rule in Group B to authorize access of Instance B to Instance A, as shown in the following table.

NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
Intranet	Inbound	Allow	Select the required protocol	Set the required port range	Security group access (authorize other accounts)	Type the account ID of User A and the security group ID of Group A	1



Note:

To guarantee the security of your instances, when you are configuring an intranet inbound rule for a Classic network-connected security group, **Security Group Access** is the top priority for **Authorization Type**. If you select **Address Field Access**, you must enter an IP address with CIDR prefix, /32, in the format of a.b.c.d/32. Only IPv4 is supported.

Scenario 2: Allow remote connection from specified IP addresses only

If you want to allow remote connection to your instance from the specified public IP addresses, add the following rule. In this example, remote connection to an instance on TCP Port 22 from a specified IP address is allowed.

Network Type	NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
VPC	N/A	Inbound	Allow	SSH(22)	22/22	Address field access	The IP address to allow access, such as 1.2.3.4.	1
Classic network	Internet							

Scenario 3: Allow an instance to access specified IP addresses only

If you want your instance to access specified IP addresses, add the following rules to its security group.

1. Add the following rule to drop any access to all public IP addresses. The priority must be lower than the rule in step 2.

Network Type	NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
VPC	N/A	Inbound	Drop	All	-1/-1	Address field access	0.0.0.0/0	2
Classic network	Internet							

2. Add the following rule to allow access to the specified IP address, with a higher priority than that in step 1.

Network Type	NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
VPC	N/A	Outbound	Allow	Select the required protocol	Set the required port range	Address field access	Type the specified IP	1
Classic network	Internet							

Network Type	NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
							address, such as 1.2.3.4	

After you add the rules, connect to the instance and try to ping it or telnet to the instance from a specified IP address. If the instance can be accessed by the specified IP address, the rule is successfully applied.

Scenario 4: Allow remote connection to an ECS instance

You may want to connect to your instance in the following cases:

- Case 1: Allow remote connection to your instance from the Internet.
- Case 2: Allow remote connection to your instance from intranet.

Case 1: Allow remote connection to your instance from the Internet

To allow remote connection to your instance from the Internet, add the following rule according to the network type and the operating system of your instance.

- VPC

NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
N/A	Inbound	Allow	Windows : RDP(3389)	3389/3389	Address field access	To allow Internet access from any public IP address, type 0.0.0.0/0. To allow Internet access from a specified Internet IP	1
			Linux: SSH (22)	22/22			
			Custom TCP	Customized			

NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
						address, see scenario 2 .	

- Classic network

NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
Internet	Inbound	Allow	Windows : RDP(3389)	3389/3389	Address field access	To allow Internet access from any public IP address, type 0.0.0.0/0. To allow Internet access from a specified Internet IP address, see Scenario 2 .	1
			Linux: SSH(22)	22/22			
			Custom TCP	Customized			

To customize the port for remote connection, see [modify the default remote access port](#).

Case 2: Allow remote connection to your instance from intranet

If you have enabled intranet communication between instances that belong to one region but different accounts, and you want to allow the instances in different security groups to connect to each other, add the following rules as needed.

- To allow a private IP address to connect to an instance.

— VPC

Make sure that an intranet communication has been built between both accounts by using [Express Connect](#). Then you can add any one of the following rules.

NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
N/A	Inbound	Allow	Windows : RDP(3389)	3389/ 3389	Address field access	Specify the private IP address of the peer instance	1
			Linux: SSH (22)	22/22			
			Custom TCP	Customized			

— Classic network

Add any one of the following rules.

NIC	Rule direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
Intranet	Inbound	Allow	Windows : RDP (3389)	3389/ 3389	Address field access	Specify the private IP address of the peer instance. To secure the instance, only an IP address with CIDR prefix, / 32, in the format of a.b.c.d/32,	1
			Linux: SSH (22)	22/22			
			Custom TCP	Customized			

NIC	Rule direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
						is allowed.	

- To allow all the instances in a security group of one account to connect to your instance:

— VPC

Make sure that an intranet communication is built between both accounts by using [Express Connect](#), and then add any one of the following rules.

NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
N/A	Inbound	Allow	Windows : RDP (3389)	3389/ 3389	Security group access (authorize other accounts)	Type the account ID of the peer and the security group ID	1
			Linux: SSH(22)	22/22			
			Custom TCP	Customized			

— Classic network

Add any one of the following rules.

NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
Intranet	Inbound	Allow	Windows : RDP (3389)	3389/ 3389	Security group access (authorize other account)	Type the account ID of the peer and the security group ID	1
			Linux: SSH(22)	22/22			
			Custom TCP	Customized			

Scenario 5: Allow access to an ECS instance over HTTP or HTTPS service

If you have built a website on your instance and want users to visit the site over HTTP or HTTPS, add any one of the following rules.

- VPC

To allow any public IP addresses access your site, add any one of the following rules.

NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
N/A	Inbound	Allow	HTTP(80)	80/80	Address field access	0.0.0.0/0	1
			HTTPS(443)	443/443			
			Custom TCP	Customized, such as 8080/8080			

- Classic network

To allow all public IP addresses to access your site, add any one of the following rules.

NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
Internet	Inbound	Allow	HTTP(80)	80/80	Address field access	0.0.0.0/0	1
			HTTPS (443)	443/443			
			Custom TCP	Customized, such as 8080/8080			



Note:

- If users cannot access your instance by using `http://Public IP address`, [verify if TCP port 80 works properly](#).
- TCP Port 80 is the default port for HTTP services. If you want to use other ports, modify the port in the configuration file of the Web server.

9.3 Default security group rules

This topic introduces the default rules in security groups created manually or by the system.



Note:

Security groups are stateful. This means that, for example, if an outbound packet is allowed then the inbound packets corresponding to this connection are also allowed. For more information about security groups, see security groups.

Security groups created by the system

When you create an ECS instance in a region where you have not created a security group, we recommend that you use the default security group provided by the system.

Such a security group generally only has default rules for access over the ICMP protocol, TCP Port 22 (for SSH), TCP Port 3389 (for RDP), TCP Port 80 (for HTTP), and TCP Port 443 (for HTTPS). Note that the default rules vary with the network type of the security group.

- VPC: The rules apply to both Internet and intranet access. The Internet access of the VPC type instance is realized through the private NIC mapping. Therefore, you cannot see the Internet NIC inside the instance, and you can only set intranet rules in the security group. The security group rules take effect for both intranet and Internet access. The default rules of the default VPC-Connected security group are described in the following table.

NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Priority	Authorization Type	Authorization Object
N/A	Inbound	Allow	Custom TCP (SSH)	22/22	110	Address field access	0.0.0.0/0
			Custom TCP (RDP)	3389/3389			
			All ICMP	-1/-1			
			Custom TCP (HTTP), optional	80/80			
			Custom TCP (443			

			HTTPS), optional				
--	--	--	------------------	--	--	--	--

- Classic network: The default rules of a classic network-connected security group are described in the following table.

NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Priority	Authorization Type	Authorization Object
Internet	Inbound	Allow	Custom TCP (SSH)	22/22	110	Address field access	0.0.0.0/0
			Custom TCP (RDP)	3389/3389			
			All ICMP	-1/-1			
			Custom TCP (HTTP), optional	80/80			
			Custom TCP (HTTPS), optional	443			



Note:

Rules with priority 110 means that they have the lowest priority in the security group. When you manually create a security group, only values from 1 to 100 are valid for priority setting. For more information about the rule priority, see [add security group rules](#).

You can also [add security group rules](#) in the default security group.

Manually created security group

After [creating a security group](#), before you add rules, the following default rules apply to the communication of all the instances in the group over the Internet or intranet:

- Outbound: Allow
- Inbound: Refuse

If your instance has joined such a security group, you can only use the [Management Terminal](#) to connect to an instance. You cannot use other remote connection methods (such as [connecting](#)

to a Linux instance by using a password or connecting to a Windows instance by using remote connection software).

You can also [add security group rules](#) to manually created security groups.

9.4 Create a security group

Each ECS instance must join at least one security group. For more information, see [security group](#).

If you have not created a security group prior to creating an ECS instance, we recommend that you use the default security group provided by the system. For more information, see [default security group rules](#).

Alternatively, you can create a security group to meet your business needs and then add your instances to it. This topic describes how to create a security group.

Prerequisite

To create a VPC-Connected security group, you must [have created a VPC](#).



Note:

You can create a VPC-Connected security group across VSwitches, but not across VPCs.

Procedure

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Networks and Security > Security Groups**.
3. Select the target region.
4. Click **Create Security Group**.
5. In the **Create Security Group** dialog box, complete the following configurations:
 - **Template:** If the instances in the security group are for Web server deployment, select an appropriate template to simplify security group rule configuration.

Setting	Template	Note
To deploy a Web server on the Linux instances in the security group	Web Server Linux	By default, inbound traffic to TCP 80, TCP 443, TCP 22, and ICMP is allowed.
To deploy a Web server on the Windows instances in the security group	Web Server Windows	By default, inbound traffic to port TCP 80, TCP 443, TCP 3389, and ICMP is allowed.

Not for Web server	Custom	After the security group is created, you can add security group rules as needed.
--------------------	---------------	--

- **Security Group Name:** Enter a name for the security group.
- **Description:** Enter a description of the security group for easier management.
- **Network Type:**
 - To create a VPC-Connected security group, select **VPC** and then select the target VPC.
 - To create a Classic network-connected security group, select **Classic**.

Create Security Group ? ×

Template:

* Security Group Name:
2 - 128 characters long. Do not start with a special character or a digit. It can contain the following special characters: ".", "_", and "-".

Description:
It must contain 2-256 characters and it cannot begin with http:// or https://

Network Type:

*VPC: [Create VPC](#)

Inbound Outbound

Authorization Object	Protocol Type	Port Range	Authorization Policy
0.0.0.0/0	TCP	80/80	Allow
0.0.0.0/0	TCP	443/443	Allow
0.0.0.0/0	TCP	22/22	Allow
0.0.0.0/0	ICMP	-1/-1	Allow

6. Click **OK.**

For a new security group without any rules, the following default rules apply to the communication of all the instances in the group over the Internet or intranet:

- Outbound: Allow
- Inbound: Forbid

Follow-up operations

After you create a security, you can [add security group rules](#).

You can also [add/remove an instance to/from a security group](#).

Related API

[CreateSecurityGroup](#)

9.5 Add security group rules

You can add security group rules to enable or disable access to and from the Internet or intranet for ECS instances in the security group.

- VPC: You only need to set inbound and outbound rules, and you do not need to create different rules for the Internet and intranet. The Internet access for VPC instance is realized through private NIC mapping. Therefore, you cannot see the Internet NIC inside the instance, and you can only set intranet rules in the security group. The rules apply to Internet and intranet access.
- Classic network: You must set outbound and inbound rules for the Internet and intranet respectively.

For a new security group without any rules, outbound traffic is allowed and inbound traffic is refused by default, over either the Internet or intranet. Therefore, we recommend that you only set rules to refuse outbound traffic or allow inbound traffic.

Changes to the security group rules automatically apply to ECS instances in the security group.

Prerequisites

You have created a security group. For more information, see [create a security group](#).

You know which Internet or intranet requests need to be allowed or refused for your instance.

Procedure

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Networks and Security** > **Security Groups**.
3. Select the target region.
4. Find the security group to add authorization rules and then, in the **Actions** column, click **Add Rules**.
5. On the **Security Group Rules** page, click **Add Security Group Rule**.



Note:

If you do not need to enable or disable all ports for all protocols, ICMP, or GRE, you can select **Quick Rule Creation**.

Protocol	SSH	telnet	HTTP	HTTPS	MS SQL
Port	22	23	80.	443	1433
Protocol	Oracle	MySQL	RDP	PostgreSQL	Redis
Port	1521	3306	3389	5432	6379

**Note:**

See step 6 for descriptions on each parameter configuration.

6. In the dialog box, set the following parameters:

- **NIC:**

- For a VPC-Connected security group, you do not need to select the NIC.

**Note:**

- If your instances can access the Internet, the rules work for both the Internet and intranet.
- If your instances cannot access the Internet, the rules work for intranet only.

- For a Classic network-connected security group, you must select **Internet** or **Intranet**.

- **Rule Direction:**

- **Outbound:** ECS instances access other ECS instances over intranet networks, or through Internet resources.
- **Inbound:** Other ECS instances in the intranet and Internet resources access the ECS instance.

- **Action:** Select **Allow** or **Forbid**.

**Note:**

Forbid policies discard the data packet without returning a response. If two security group rules overlap except the authorization policy, the **Forbid** rule takes priority over the **Allow** rule.

- **Protocol Type and Port Range:** The port range setting is affected by the selected protocol type. The following table shows the relationship between protocol types and port ranges.

Protocol Type	Port Range	Scenarios
---------------	------------	-----------

All	Shown as -1/-1, indicating all ports. You cannot modify it.	Used in scenarios where both applications are fully and mutually trusted.
All ICMP	Shown as -1/-1, indicating no port restriction. You cannot modify it.	Used to detect the instance network connection status by using <code>ping</code> .
All GRE	Shown as -1/-1, indicating no port restriction. You cannot modify it.	Used for VPN service.
Custom TCP	For custom port ranges, the valid port value is 1–65535 , and the valid port range format is Start Port/End Port . A valid port range format must be used for one port. For example, use 80/80 to indicate port 80.	It can be used to allow or forbid one or several successive ports.
Custom UDP		
SSH	Shown as 22/22. After connecting to the ECS instance, you can modify the port number. For more information, see default remote access port modifications .	Used for SSH to connect to a Linux instance remotely.
TELNET	Shown as 23/23.	Used to remotely log on to instances by using Telnet.
HTTP	Shown as 80/80.	The instance is used as a server for a website or a web application.
HTTPS	Shown as 443/443.	The instance is used as a server for a website or a web application that supports HTTPS.
MS SQL	Shown as 1433/1433.	The instance is used as an MS SQL server.
Oracle	Shown as 1521/1521.	The instance is used as an Oracle SQL server.
MySQL	Shown as 3306/3306.	The instance is used as a MySQL server.

RDP	Shown as 3389/3389. After connecting to the ECS instance, you can modify the port number. For more information, see default remote access port modifications .	Used to remotely connect to Windows instances.
PostgreSQL	Shown as 5432/5432.	The instance is used as a PostgreSQL server.
Redis	Shown as 6379/6379.	The instance is used as a Redis server.

**Note:**

Port 25 is restricted by default and cannot be opened through security group rules.

However, you can submit a ticket to [apply to open TCP port 25](#). For more information, see [introduction to common ECS instance ports](#).

- **Authorization Type** and **Authorization Object**: The authorization object affects the setting of authorization type. The following table shows the relationship between them.

Authorization Type	Authorization Object
Address field access	Use the IP or CIDR block format such as 10.0.0.0 or 192.168.0.0/24. Only IPv4 addresses are supported. 0.0.0.0/0 indicates all IP addresses.
Security group access	<p>Only for intranet access. Authorize the instances in a security group under your account or another account to access the instances in this security group.</p> <ul style="list-style-type: none"> • Authorize this account: Select a security group under your account. Both security groups must be in the same VPC. • Authorize another account: Enter the target security group ID and the account ID. On the Account Management > Security Settings, you can obtain the account ID. <p>For VPC-Connected network instances, security group access works for private IP addresses only. If you want to authorize</p>

Internet IP address access, use address field access.

**Note:**

To guarantee the security of your instance, when you are configuring an intranet inbound rule for a classic network-connected security group, **Security Group Access** is the top priority for **Authorization Type**. If you select **Address Field Access**, and you want to type an IP address in the CIDR format, type an IP address in the format of a.b.c.d/32. Only 32 is the valid CIDR prefix.

- **Priority:** The value range is 1-100. The smaller the value, the higher the priority. For more information, see [ECS security group rule priority explanation](#).

7. Click **OK**.

Security group rules generally take effect immediately.

Verify security group rules

If you have installed a web service on the instance and added a security group rule in a security group, you can allow all IP addresses to have inbound access to TCP port 80 of the instance. Follow these steps according to your instance OS to verify the security group rule.

Linux instances:

For a Linux instance in the security group, follow these steps to verify the security group rule:

1. [Connect to a Linux instance by using a password](#).
2. Run the following command to check whether TCP 80 is being listened.

```
netstat -an | grep 80
```

If the following result returns, web service for TCP port 80 is enabled.

```
tcp        0      0 0.0.0.0:80          0.0.0.0:*
          LISTEN
```

3. Enter `http://public IP address of the instance` into your browser. If access is successful, the rules have been activated.

Windows instances:

For a Windows instance in the security group, follow these steps to verify the security group rule:

1. [Connect to a Windows instance](#).

2. Run the **CMD**, and run the following command to check whether TCP port 80 is being listened.

```
netstat -aon | findstr :80
```

If the following result returns, web service for TCP port 80 is enabled.

```
TCP 0.5.0.0: 80 0.5.0.0: 0 listening 1172
```

3. Enter `http://instance public IP address` into your browser. If access is successful, the rules have been activated.

ECS security group rule priority explanation

The **Priority** value of a security group rule ranges from 1 to 100. A smaller number indicates a higher priority.

ECS instances can belong to different security groups. As a result, instances may have multiple security group rules that have the same protocol types, port ranges, authorization types, and authorization objects. The rule that takes effect depends on the setting of **Priority** and

Authorization Policy:

- If the rules have the same **Priority**, the **Forbid** rule takes effect over the **Allow** rule.
- If the rules have different **Priority**, the rule with the higher priority takes effect first, regardless of the setting of **Authorization Policy**.

Related topics

- [Security group FAQ](#)
- [Security group](#)
- [Default security group rules](#)
- [Implication and matching sequence of the ECS security group rule priority](#)

9.6 View the security group list

You can view security groups in the ECS console. To do so, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Networks and Security > Security Groups**.
3. Select the target region. A list of all security groups in the specified region is displayed.
4. Select the target VPC name or enter a VPC ID in the search box to search for security groups under a VPC.

9.7 Modify security group attributes

You can modify the name and description of a security group at any time. To do so, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Networks and Security > Security Groups**.
3. Select the target region.
4. Modify the attributes of a security group as follows:
 - Modify the name: Hover your mouse over the name of a security group, and then click the pen icon that appears.
 - Modify the name and description: Click **Modify** on the right of the security group, and then enter a new name and description in the dialog box.
5. Click **OK**.

9.8 View the security group rules

You can view the security group rules at any time. To do so, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Networks and Security > Security Groups**.
3. Select the target region.
4. Select a security group, and click **Add Rules**.
5. Depending on the network type, the following information is displayed for security groups s:
 - For VPC, **Inbound** and **Outbound** is displayed.
 - For classic network, **Internet Inbound**, **Internet Outbound**, **Intranet Inbound**, and **Intranet Outbound** is displayed.
6. Click a tab to view the security group rules for that type.

9.9 Delete a security group rule

You can delete security group rules that you no longer require. To do so, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Networks and Security > Security Groups**.
3. Select the target region.
4. Find the security group where you want to delete rules and then, in the **Actions** column, click **Add Rules**.

5. On the security group management page, select the rule direction and find the rule you want to delete.
 - If the security group is for classic network, the rule directions are **Internet Inbound**, **Internet Outbound**, **Intranet Inbound**, and **Intranet Outbound**.
 - If the security group is for VPC, the rule directions are **Inbound** and **Outbound**.
6. In the **Actions** column, click **Delete**.
7. On **Delete Security Group Rule** dialog box, read and confirm the information displayed, and then click **OK**.

9.10 Delete a security group

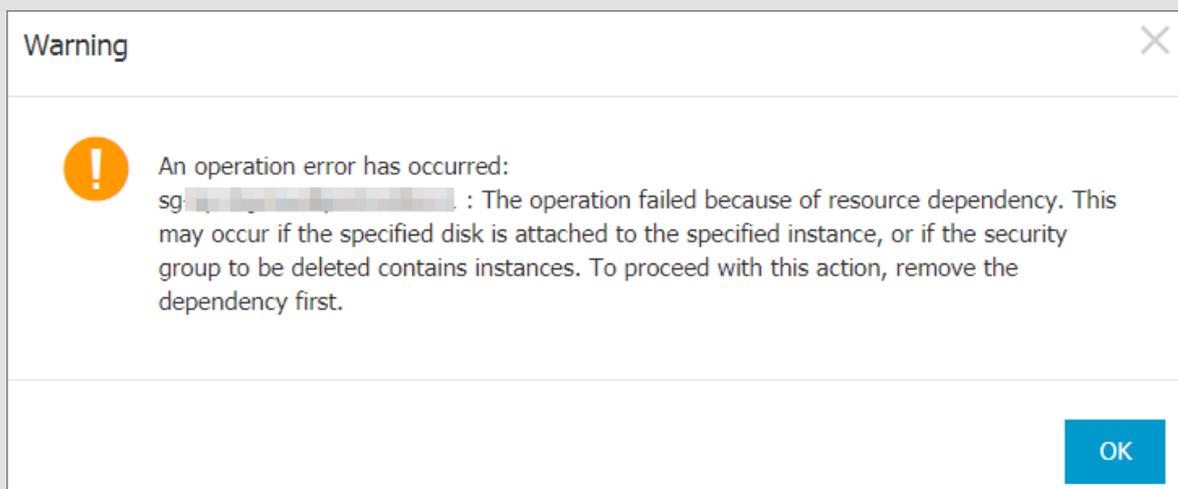
You can delete security groups that you no longer require. Deleting a security group also deletes all its corresponding rules.



Note:

Before you delete a security group, note the following:

- Make sure there are no ECS instances in the security group. For more information on how to move an ECS instance to another security group, see [add to or remove from a security group](#).
- Make sure the security group is not referenced in the rules of another security group. You can delete a security group directly by following the steps described in this document. If the security group is authorized by another security group, error message shown in the following figure appears. If this occurs, you must delete the corresponding authorization rule.



Procedure

To delete a security group, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Networks and Security** > **Security Groups**.
3. Select the target region.
4. Select one or more security groups, and click **Delete**.
5. In the displayed **Delete Security Group** dialog box, click **OK**.

Related APIs

- Delete a security group: [DeleteSecurityGroup](#)
- Query authorization relationships between a security group and another security group: [DescribeSecurityGroupReferences](#)
- Move an ECS instance out of a security group: [LeaveSecurityGroup](#)

9.11 Clone a security group

You can clone a security group across regions and network types.

Scenarios

The following scenarios detail examples as to when you may need to clone a security group:

- You have created a security group, named SG1, in Region A, and you want to apply the same rules of SG1 to ECS instances in Region B. To do so, you can clone SG1 to Region B without creating a new security group in Region B.
- You have a security group in the classic network, named SG2. You want to apply the rules of SG2 to instances in a VPC. To do so, you can clone SG2 and choose VPC as the network type when configuring the cloning. Then, in the VPC network, you will have a new security group that has the same rules as SG2.
- If you want to apply new security group rules to an ECS instance that are running an online business application, we recommend that you clone the security group as a backup before modifying the rules. If the new security group rules are disadvantageous to the online business application, you can restore the rules completely or partly.

Prerequisite

If you want to change the network type of a security group from Classic to VPC, you have to [create a VPC and VSwitch](#) in the target region first.

Procedure

1. Log on to the [ECS console](#).

2. In the left-side navigation pane, select **Networks and Security** > **Security Groups**.
3. Select the target region.
4. Find the target security group and then, in the **Actions** column, click **Clone Security Group**.
5. In the **Clone Security Group** dialog box, set the new security group information:
 - **Target Region:** Select a region suitable for the new security group. Note that only supported regions are displayed in the drop-down list.
 - **Security Group Name:** Specify a new name for the new security group.
 - **Network Type:** Select a network type suitable for the new security group. If **VPC** is selected, you must select a VPC in the drop-down list.

Clone Security Group

Destination Region: China East 1 (Hangzhou) ▼
Only partial regions are supported.

* Security Group Name: sg-
2 - 128 characters long. Do not start with a special character or a digit. It can contain the following special characters: ".", "_", and "-".

Description: System created security group.
It must contain 2-256 characters and it cannot begin with http:// or https://

Network Type: VPC ▼

*VPC: vpc- ▼ Create VPC

OK Cancel

6. Click **OK**.

After successful creation, the **Clone Security Group** dialog box closes automatically. The new security group is displayed on the **Security Groups** page.

9.12 Introduction to common ECS instance ports

The following table lists commonly used ECS instance ports.

Port	Service	Description
21	FTP	A port opened by the FTP service is used for uploading and downloading files.
22	SSH	An SSH port is used to connect to a Linux instance by using a password in command-line mode.
23	Telnet	The Telnet port is used for Telnet to log on to the ECS instance.
25	SMTP	The port that is open to the SMTP service is used for sending mails. To maintain security, Port 25 for ECS instances is restricted by default. Submit a ticket to open it. For details, see apply to open TCP port 25 .
80	HTTP	Provides access to HTTP services, such as IIS, Apache, and Nginx. We recommend that you verify if TCP port 80 works properly .
110	POP3	A port used for the POP3 protocol, which is a protocol for sending and receiving emails.
143	IMAP	A port used for IMAP (Internet Message Access Protocol), which is a protocol for receiving emails.
443	HTTPS	A port used to provide access to the HTTPS service. HTTPS is a protocol that provides encryption and transmission through secure ports.

1433	SQL Server	The TCP port of the SQL Server that is used for external service by SQL Server.
1434	SQL Server	The SQL Server UDP port that is used to return which TCP/IP port SQL Server uses.
1521	Oracle	An Oracle communications port. The port that needs to be released by Oracle SQL is deployed on the ECS instance.
3306	MySQL	The port through which the MySQL database provides external service.
3389	Windows Server Remote Desktop Services	The Windows Server Remote Desktop Services port can be used to connect to a Windows instance .
8080	Proxy port	Similar to 80 port, port 8080 is used by WWW agents to enable web browsing. If you are using port 8080, when you visit a Web site or use a proxy server, you must add <code>:8080</code> after the IP Address: 8080. If you install the Apache Tomcat service, the default service port is 8080.
137, 138, 139	NetBIOS protocol	<ul style="list-style-type: none"> Ports 137 and 138 are UDP ports that are used to transfer files through the network neighbor. The connection entering through the port 139 attempts to obtain the NetBIOS/smb service. <p>NetBIOS protocols are often used for Windows files, printer sharing, and samba.</p>

Some ports cannot be accessed

Issue: An ECS instance attempts to listen for the corresponding port, but the port is not accessible, while other ports can be accessed normally.

Cause: Some operators determine that port numbers 135, 139, 444, 445, 5800, 5900, and related ports, are high-risk ports, which are then blocked by default.

Resolution: We recommend that you change the port to another port number

Related topic

For more information on how to release a service port through a security group, see [add security group rules](#).

9.13 Restore security group rules

Restoring security group rules indicates the process of completely or partially restoring the rules in the original security group to those of a target security group. Specifically:

- **Completely restoring** means moving the rules that do not exist in the target security group from the original security group, and adding the rules that only exist in the target security group to the original security group. After restoration, rules in the original security group are identical with those in the target security group.
- **Partially restoring** means adding the rules that only exist in the target security group to the original security group and ignoring the rules that only exist in the original group.

Limits

Restoring security group rules has the following limits:

- The original security group and the target security group must be in the same region.
- The original security group and the target security group must be of the same network type.
- If any system-level security group rules, of which the priority is 110, exist in the target security group, they are not created during restoration. This means that after restoration, the rules in the original security group may be different. If you need the system-level security group rules, you must manually create the rules and set their priority to 100.

Scenario

If you want to apply new security group rules to an ECS instance that is running an online business application, you can clone the former security group as a backup, and then modify the

rules inside. If the new security group rules affect the online business application, you can choose to fully or partially restore the rules.

Prerequisite

You must have at least one security group of the same network type in the same region.

Procedure

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Networks and Security > Security Groups**.
3. Select the target region.
4. Find the security group you want to restore rules for as the original security group and then, in the **Actions** column, click **Restore Rules**.
5. In the **Restore Rules** dialog box, follow these steps:
 - a. Select the **Target Security Group**: Select a security group as the target security group that must have different rules from the original security group.
 - b. Select a restore **Method**:
 - If you want the original security group to have the same rules as the target security group, select **Completely Restore**.
 - If you want to add the rules that only exist in the target security group to the original security group, select **Partially Restore**.
 - c. In the **Preview** area, preview the restoration result:
 - Rules highlighted in green only exist in the target security group. No matter whether you select **Completely Restore** or **Partially Restore**, these rules are added to the original security group.
 - Rules highlighted in red are the rules that do not exist in the target security group. If **Completely Restore** is selected, the system removes these rules from the original security group. If **Partially Restore** is selected, the rules are retained in the original security group.
 - d. Click **OK**.

The **Restore Rules** dialog box is closed automatically after successful creation. On the **Security Groups** page, find the original security group you restored the rules for and then, in the **Actions** column, click **Add Rules** to enter the **Security Group Rules** page to view the updated security group rules.

9.14 Add to or remove from a security group

Add an instance to a security group

You can add an instance to a security group through the console. Up to five security groups can be added to an instance.

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Instances**.
3. Select the region.
4. Select the target instance by clicking the instance name or clicking its corresponding **Manage** button.
5. Click **Security Groups** in the left-side navigation pane.
6. Click **Add Security Group** and then select the required security group.
7. Click **OK**.

After you add an instance to the security group, the rules apply to the instance automatically.

Remove an instance from a security group

You can remove instances from security groups.

**Note:**

- An instance must be in at least two security groups for this action to be performed.
- Do enough test before this operation to avoid any intranet communication error between instances.

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Instances**.
3. Select the target region.
4. Select the target instance by click the instance name or clicking its corresponding **Manage** button.
5. In the left-side navigation pane, click **Security Groups**.
6. Select the security group to remove from the instance and then click **Remove**.
7. Click **OK**.

For use cases of security groups, see [scenarios](#).

10 Deployment sets

10.1 Create an instance in the deployment set

Once you have an available deployment set, you can learn how to create instances in it by following this article.

Context

Only seven instances can be created in a deployment set in one zone. The number of instances that can be created in a region is seven multiplied by the number of the zones. When you create an instance, you can use a launch template or use the bulk creation feature to avoid the inconvenience caused by the restriction of maximum instances per zone. This topic describes how to create an instance in a deployment set using the ECS console. If you are an API user, you can call [RunInstances](#) and specify the `DeploymentSetId` parameter in the request.

Procedure

1. Log on to [ECS console](#).
2. Select a region.
3. In the left-side navigation pane, choose **Networks and Security > Deployment Sets**.
4. On the Deployment Sets page, locate the target deployment set. You can click **Create Instance** in the Actions column. Alternatively, you can also select the deployment set and then click **Create Instance** in the **Instances** list that appears.
5. In the page that appears, complete the instance configuration on the **Advanced Purchase** tab page. For more information about creating an instance, see [Create an instance by using the wizard](#). You need to note these configurations:
 - **Basic configurations:**
 - **Region:** The instance and the target deployment set must be in the same region.
 - **Zone:** The maximum number of instances in each deployment set zone cannot exceed seven.
 - **Instance:** The type families of the instance that can be created by the deployment set currently support c5, d1, d1ne, g5, hfc5, hfg5, i2, ic5, r5, se1ne, sn1ne, and sn2ne. For more information about instance types and performance, see [Instance type families](#).
 - (Optional) You can specify the number of instances for the **Quantity** field, and you need to consider the number of instances that already exist in the current zone of the same deployment set.

- (Optional) **System Configuration > Sequential Suffix**: After creating instances in bulk, you can add sequential suffixes to the Instance name and host name. The sequential suffix increments from 001 to 999.
 - **Grouping > Deployment Sets**: Select the target deployment set.
 - (Optional) **Preview > Save as launch template**: Saves the configuration as a launch template that you can use to quickly create an instance the next time. For more information, see [Instance launch template](#).
6. Check the settings you have configured, and then click **Create Instance**.
 7. In the left-side navigation pane, click **Deployment Sets**. You can view the deployment set that you have created.

What's next

After creating an instance, you can perform the following tasks:

- View and manage instances in the deployment set. For more information, see the related topics in the *Cite LeftUser GuideCite Right*.
- Move an instance between deployment sets:
 1. Log on to the [ECS console](#).
 2. In the left-side navigation pane, click **Instances**.
 3. Select a region.
 4. Locate the target instance, and make sure that the instance is in the **Stopped** or **Running** status.
 5. In the **Actions** column, choose **More > Instance Settings > Change Deployment Set**.
 6. In the **Change Deployment Set** dialog box that appears, select the target deployment set, and configure the **Force Change** settings:
 - **Yes**: This option allows the system to move the instance to a new host and to restart the instance in the **Running** or **Stopped** status.
 - **No**: This option does not allow the system to move the instance to a new host. Instead, the system only attaches the target deployment set to the current host machine. This may cause the change of deployment set to fail.
 7. Click **OK**.

10.2 Manage deployment sets

After creating a deployment set, you can modify the deployment set name and description, or remove deployment sets that are no longer required to ensure that the usage limit is not exceeded.

Edit deployment set information

To change the name or description of a deployment set in the ECS console, follow these steps:

1. Log on to [ECS console](#).
2. Select a region.
3. In the left-side navigation pane, choose **Networks and Security > Deployment Sets**.
4. Locate the deployment set that needs to be modified.
5. Edit the information using either of the following methods:
 - Hover the cursor over the **Deployment Set Name** column, click the  icon that appears, and then enter the deployment set name and description.
 - In the Actions column of the target deployment set, click **Edit**, and enter the deployment set name and description.
6. Click **OK**.

You can also use the [ModifyDeploymentSetAttributes](#) API operation to modify the deployment set name and description.

Delete deployment sets

**Note:**

If a deployment set already includes an instance, you cannot delete the deployment set.

To delete one or more deployment sets in the ECS console, follow these steps:

1. Log on to [ECS console](#).
2. Select a region.
3. In the left-side navigation pane, select the **Network and Security > Deployment Sets**.
4. Select one or more deployment sets that need to be deleted, hover the cursor over the **Actions** menu, and then click **Delete Deployment Set**.
5. Click **OK** to delete the deployment set.

You can use the [DeleteDeploymentSet](#) API operation to delete deployment sets.

10.3 Create deployment sets

A deployment set is a strategy that controls the distribution of instances and enables you to design the disaster tolerance functionality and availability of the business as you create an instance.

Context

You can use the deployment set to spread the Elastic Compute Service (ECS) instances involved in your business across different physical servers to guarantee high availability and disaster tolerance at the level of physical machines. For more information, see [Deployment sets](#). This topic describes how to create a deployment set in the ECS management console. If you are an API user, you can use [CreateDeploymentSet](#).

Procedure

1. Log on to [ECS console](#).
2. Select a region.
3. In the left-side navigation pane, choose **Networks and Security** > **Deployment Sets**.
4. On the Deployment Sets page, click **Create Deployment Set**.
5. On the **Create Deployment Set** page, enter a **Name** and **Description** for the deployment set. Currently, only **High availability** is supported for the **Strategy** option. For more information about the deployment set strategy, see [Deployment set strategy](#).

What's next

After creating a deployment set, you can perform the following tasks:

- [Create an instance in the deployment set](#).
- Add an ECS instance to the deployment set:
 1. Log on to the [ECS console](#).
 2. In the left-side navigation pane, click **Instances**.
 3. Select a region.
 4. Locate the target instance, and make sure that the instance is in the **Stopped** or **Running** status.
 5. In the **Actions** column, choose **More** > **Instance Settings** > **Change Deployment Set**.
 6. In the **Change Deployment Set** dialog box that appears, select the target deployment set, and configure the **Force Change** settings:
 - **Yes**: This option allows the system to move the instance to a new host and to restart the instance in the **Running** or **Stopped** status.

- **No:** This option does not allow the system to move the instance to a new host. Instead, the system only attaches the target deployment set to the current host machine. This may cause the change of deployment set to fail.

7. Click **OK**.

11 Key pairs

11.1 Create an SSH key pair

Limits

- The *SSH key pair*, known as as key pair, applies to Linux instances only.
- Currently, only 2048-bit RSA key pairs are supported.
 - Of the key pair components, Alibaba Cloud holds the public key.
 - After creating the key pair, you must download and securely store the private key of the key pair for future use.
 - The private key is in the unencrypted PEM-encoded PKCS#8 format.
- An Alibaba Cloud account can have a maximum of 500 key pairs per region.

Create an SSH key pair

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Networks and Security** > **SSH Key Pair**.
3. Select the target region.
4. On the **SSH Key Pairs** page, select the target region and click **Create SSH Key Pair**.
5. On the **Create SSH Key Pair** page, enter a name for the key pair, and select **Auto-Create SSH Key Pair**.



Note:

The specified key pair name must be unique under the Alibaba Cloud account. That is, the name cannot be the same as that of an existing key pair or of a key pair that has been attached to an instance but has since been deleted. Otherwise, the error message "The key pair already exists" appears.

Create Key Pair [Return to keypair list](#)

*Key Pair Name:
2 - 128 characters long. Do not start with a special character or a digit. It can contain the following special characters: ".", "_", and "-".

*Creation Type: Automatically Create a Key Pair Import an Existing Key Pair

Download the private key immediately after creation. This is the only chance for you to download the private key.

6. Click **OK**.

**Note:**

After a key pair is created, you must download and securely store the private key for future use. If you do not have the private key, you cannot log on to the ECS instance.

After creating the key pair, you can view the information, including key pair **Name** and **Fingerprint** in the key pair list.

What to do next

After creating an SSH key pair, you can [attach or remove it](#) to or from an ECS instance.

11.2 Import an SSH key pair

If you use a third-party key generation tool to generate an RSA key pair, you can import its public key into Alibaba Cloud. For information about supported types of imported key pairs, see [SSH key pairs](#).

**Note:**

To guarantee the security of your instance, store the private key of the key pair securely and do not import the private key to Alibaba Cloud.

To import an SSH key pair, you must have generated a key pair, and the public key to be imported into Alibaba Cloud must be Base64-encoded.

To import an SSH key pair, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Networks and Security** > **SSH Key Pair**.
3. Click **Create SSH Key Pair**.
4. Select the target region.
5. On the **Create SSH Key Pair** page, enter a name for the key pair, and select **Import SSH Key Pair**, and then enter the **Public Key**.

**Note:**

The specified key pair name must be unique under the Alibaba Cloud account. That is, the name must not be the same as that of an existing key pair or a key pair that was bound to an instance but has since been deleted. Otherwise, the error message "The key pair already exists" appears.

Create SSH Key Pair [← Back to SSH Key Pair List](#)

*SSH Key Pair Name:
The name can be 2 to 128 characters in length and can contain periods (.), underscores (_), and hyphens (-). It cannot start with a special character or number.

*Creation Type: Auto-Create SSH Key Pair Import SSH Key Pair

*Public Key: 

(Base64 Encoded) [Show Example](#)

Tag:

6. Click **OK**.

After creation, you can view the information, including the key pair **Name** and **Fingerprint**, in the key pair list.

11.3 Attach or remove an SSH key pair

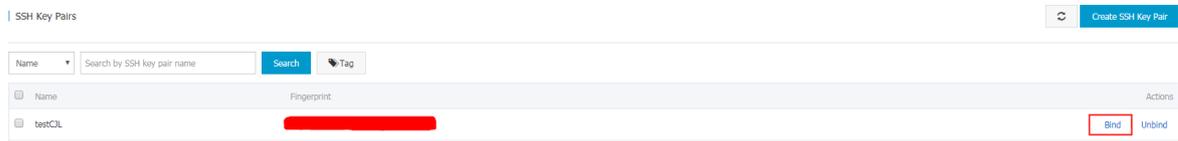
Limits

- Each ECS instance can be attached to only one SSH key pair.
- Except for the non-I/O-optimized instances of Generation I, all the Linux instances listed in the [instance type families](#) support the authentication method of SSH key pair.
- If your ECS instance is in the **Running** status when you attach it to an SSH key pair, you must, [restart it](#) after the attachment.
- If the ECS instance has already attached to an SSH key pair, after the new key pair is attached, the new key automatically replaces the original key.
- If you use password-based authentication for Linux logon, the password authentication feature is automatically disabled after the key pair is attached.
- After an SSH key pair is removed, you must [reset the instance password](#) for successful connection.

Attach an SSH key pair

To attach an SSH key pair to an ECS instance, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Networks and Security** > **SSH Key Pair**.
3. Select the target region.
4. Select the target key pair, and then click **Bind** in the **Actions** column.



5. In the **Bind** dialog box, select one or more instances from the **Select Instance** box, and then click the icon > to move them to the **Selected** column.

**Note:**

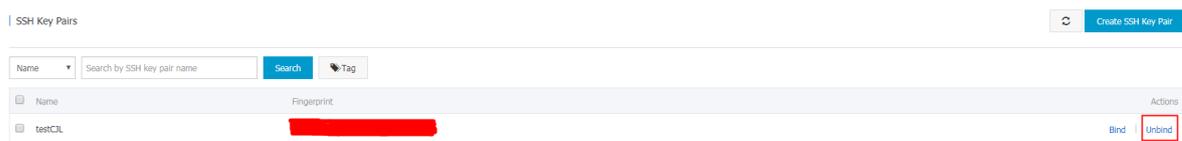
In the **Select Instance** box, the instances with gray names are either Windows instances or non-I/O-optimized instances of Generation I, for which SSH key pairs are not supported.

6. Click **OK**.

Remove an SSH key pair

To remove an SSH key pair from an ECS instance, follow these steps:

1. Log on to the [ECS console](#).
2. Select the target region.
3. In the left-side navigation pane, click **Networks and Security** > **SSH Key Pair**.
4. Select the target SSH key pair, and then click **Unbind** in the **Actions** column.



5. In the **Unbind** dialog, select one or more instances from the **Select Instance** box, and then click the icon > to move them to the **Selected** column.
6. Click **OK**.

11.4 Delete an SSH key pair

If you no longer require a key pair, you can delete it. Note that deleting a key pair is irreversible. Existing instances that have used the key pair are not affected. However, the deleted key pair name remains associated to the instance. Exercise caution when performing this action.

**Note:**

- If you delete a key pair that is still attached to an instance, its name cannot be used to create or import a key pair again. Otherwise, the error message "The key pair already exists".
- If you delete a key pair that had never been attached to an instance, its name can be used create or import a key pair again.

To delete one or more key pairs, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Networks and Security > SSH Key Pair**.
3. Select one or more key pairs.
4. Click **Delete**.

12 Cloud assistant

12.1 Cloud assistant client

The cloud assistant client is an agent that facilitates cloud assistant command invocation on ECS instances. The cloud assistant client does not perform any operations. Instead, the cloud assistant client allows you to start all operations on ECS instances under your account. By default, instances created later than Dec 1, 2017 are pre-installed with the cloud assistant client. If an ECS instance was created earlier than Dec 1, 2017 and you want to use cloud assistant service, you can install the cloud assistant client manually. This topic describes how to install, update, and disable the cloud assistant client in an ECS instance.

Install cloud assistant client

Windows

1. [Connect to a Windows instance](#).
2. [Download](#) the cloud assistant client.
3. Double-click the client file and follow the instructions to install the client.

Linux

Based on the distribution of Linux, to install cloud assistant client, select the required method from the following installation options.

- Install the RPM package:
 1. [Connect to a Linux instance by using a password](#).
 2. Run `wget https://repo-aliyun-assist.oss-cn-beijing.aliyuncs.com/download/aliyun_assist.rpm` to download the RPM package of the cloud assistant client.
 3. Run `rpm -ivh aliyun_assist.rpm` to install the cloud assistant client.
- Install the DEB package:
 1. Connect to your Linux instance.
 2. Run `wget https://repo-aliyun-assist.oss-cn-beijing.aliyuncs.com/download/aliyun_assist.deb` to download the DEB package of the cloud assistant client.
 3. Run `dpkg -i aliyun_assist.deb` to install the cloud assistant client.
- Install with the compilation file of source code:

1. Connect to your Linux instance.
2. Run `git clone https://github.com/aliyun/aliyun_assist_client` to download the cloud assistant client source code.
3. Enter the source code directory.
4. Run `cmake .` to generate the compilation file.
5. Run `make` to start compilation.
6. Run `cmake_install.sh` to install the cloud assistant client.

Update the cloud assistant client

The update process of the cloud assistant client runs one time every hour to query update resources for the client. Depending on the OS installed in your instance, the update process is located at the following directory:

- Windows instance: `C:\ProgramData\aliyun\assist\${version}\aliyun_assist_update`
- Linux instance: `/usr/local/share/aliyun-assist/${version}/aliyun_assist_update`

Generally, the update process is one of the startup items in the instance. However, you can disable the update process as follows:

- Windows instance: Run `rename aliyun_assist_update` in CMD or PowerShell.
- Linux instance: Run `chmod a-x aliyun_assist_update`.

Disable the cloud assistant client

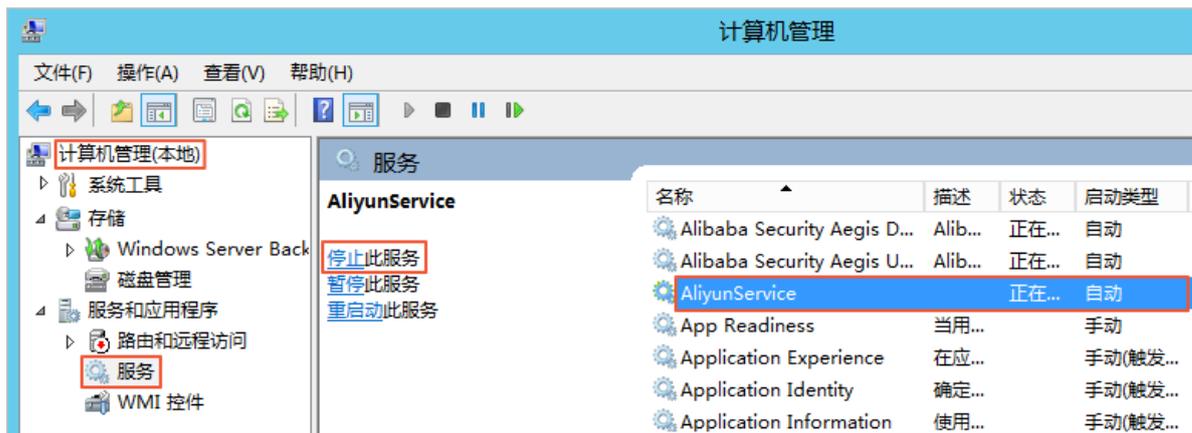


Note:

The cloud assistant client is managed by the *Aliyun* service. If you disable the client, the *Aliyun* service is also disabled, which means stopping an instance in the ECS console may fail. Exercise caution before performing this action.

Windows instance

1. [Connect to a Windows instance](#).
2. Select **Computer Management > Services and Applications > ServicesAliyunService**
3. Click **Stop the service**.



Linux instance

1. [Connect to a Linux instance by using a password.](#)
2. Run the following commands to disable the cloud assistant client.

```
systemctl stop agentwatch
chkconfig agentwatch off
```

What to do next

You can visit the [GitHub aliyun_assist_client](#) to explore the open source stack of cloud assistant.

You can also use the cloud assistant client for the following actions:

- [Cloud assistant](#)
- [InvokeCommand](#)
- [Automatically manage instances](#)

12.2 Manage commands

After creating cloud assistant commands, you can set a name and description for a command, clone commands, and delete commands you no longer require.

Modify the name and description of a command

To create or modify the name and description of a cloud assistant command, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Cloud Assistant**.
3. Select the target region.
4. Hover your mouse over the target command and click the  icon that appears. Then, edit the following as required:

- **Command Name:** Enter a new command name.
- **Command Description:** Enter a new command description.

5. Click **OK**.

You can also use the ECS API [ModifyCommand](#) to modify command information.

Clone a command

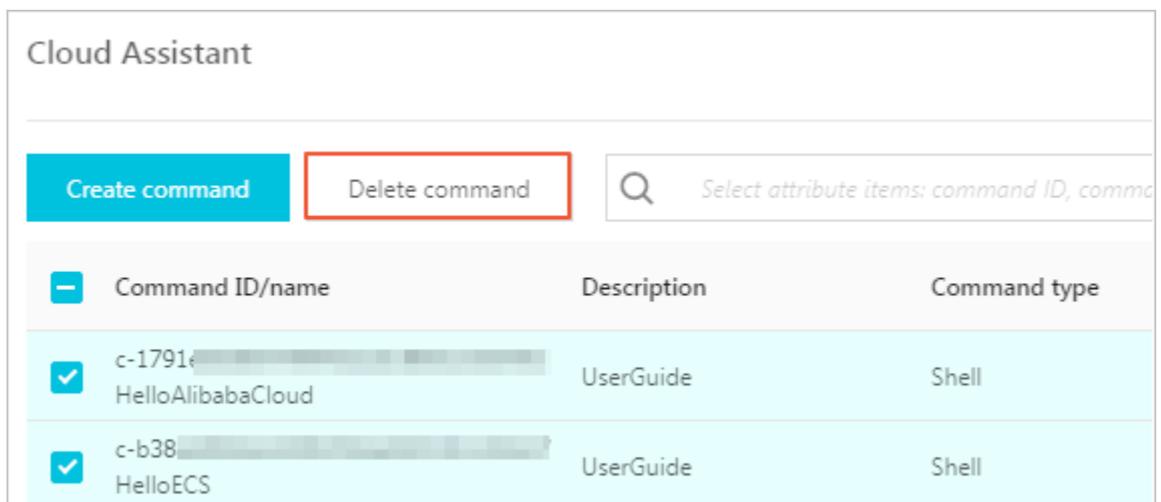
The clone command is equivalent to adding a new version of an existing cloud assistant command . You can retain all the information of the original cloned command, or set a new name, description , type, content, execution path, or timeout time for it. To clone a command, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Cloud Assistant**.
3. Select the target region.
4. Find the target cloud assistant command and from the **Operation** list, click **Clone**.
5. In the **Clone Command** dialog box, complete the following as required:
 - a. Enter a new **Command Name**.
 - b. Enter a new **Command Description**.
 - c. Click the icon  to replace the command type from the drop-down list. For Windows instances, you can select **Bat** or **Power Shell**. For Linux instances, you can only select **Shell**.
 - d. Enter or paste the new command content.
 - e. Determine a new command **Execution Path**. The default execution path for Bat or PowerShell commands is the directory where the cloud assistant client is installed, such as `C:\ProgramData\aliyun\assist\$(version)`. The default execution path for Shell commands is the `/root` directory.
 - f. Configure the timeout time, in seconds for the command. The default value is 3600. If a command you created cannot be executed within the amount of time set by this parameter , the command times out. When the timeout time of the command expires, the command process will be forcibly terminated.
 - g. Confirm your modification settings and then click **Create**.

Delete commands

Within each Alibaba Cloud region, you can create a maximum of 100 cloud assistant commands. We suggest that you regularly review your commands to check if any commands are unnecessary and can be deleted to . To delete a command on the ECS console, perform the following steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Cloud Assistant**.
3. Select the target region.
4. Depending on your requirements, perform the following actions:
 - To delete a single command, from the **Operation** list, select **Delete**.
 - To delete multiple commands, select the target instances and then click **Delete Command**.



5. In the **Delete Command** dialog box, click **OK**.

You can also use the ECS API [DeleteCommand](#) to delete commands.

12.3 Create commands

You can use cloud assistance commands to perform routine tasks for ECS instances, such as execution of automatic maintenance scripts, process polling, resetting of user passwords, installation and removal of software, application updates, and patch installations. Command types can either be Bat or PowerShell for Windows, or Shell for Linux.

Limits

- Each Alibaba Cloud region can support up to 100 cloud assistant commands.
- A script cannot exceed 16 KB after being Base64 encoded.

Create commands

To create a command on the ECS Console, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Cloud Assistant**.
3. Select the target region.
4. Click **Create Command** and then perform the following actions:
 - a. Enter a **Command Name**, for example, HelloECS.
 - b. Enter a **Command Description**, for example, UserGuide.
 - c. Click the  icon, and select command type from the drop-down list. For Windows instances, you can select either **Bat** or **PowerShell**. For Linux instances, you must select **Shell**.
 - d. Enter or paste the contents of your command, such as:

```
echo hello ECS!  
echo root:NewPasswd9! | chpasswd  
echo Remember your password!
```

- e. Determine the **Execution Path** of the command. By default, the execution paths of Bat and PowerShell commands are set to the directory where the cloud assistant client is stored, such as `C:\ProgramData\aliyun\assist\$(version)`. Additionally, shell commands are in the `/root` directory by default.
- f. Set the value for the maximum timeout time (in seconds) for commands in an instance. The default value is 3600. If a command times out, the command process will be forcibly terminated.
- g. Confirm the command details and then click **Create**.

Create command ?

* Command name : ✓

Command description :

* Command type : ▾

* Command content :

```
1 echo hello ECS!  
2 echo root:NewPasswd9! | chpasswd  
3 echo Remember your password!
```

Execution path ?:

Timeout ?: Second

The timeout range can be set to 0-86400 seconds (24 hours).The

You can also use the ESC API [CreateCommand](#) to create a cloud assistant command.

What to do next

[Invoke commands](#)

12.4 Run commands

You can run a cloud assistant command on one or more instances. The command execution status and corresponding results for cloud assistant commands run on each instance do not

impact the execution results of other instances. You can also configure an execution interval for each cloud assistant command as required.

Limits

- You can run a maximum of 500 cloud assistant commands in each Alibaba Cloud region per day.
- You can run a command on a maximum of 50 instances at a time.
- The status of the target instance or instances must be **Running**.
- The target instance or instances must have *cloud assistant client* installed.
- The target instance or instances network type must be *VPC-Connected*.
- The period for running cloud assistant commands cannot be less than 10 seconds.
- The scheduled time for periodic command execution is set to China Standard Time (UTC +08 :00) based on the system time obtained from the ECS instances. Make sure that the time or time zone of your ECS instance is consistent with your own expectations.

Run commands

To run a cloud assistant command on the ECS console, follow these steps:

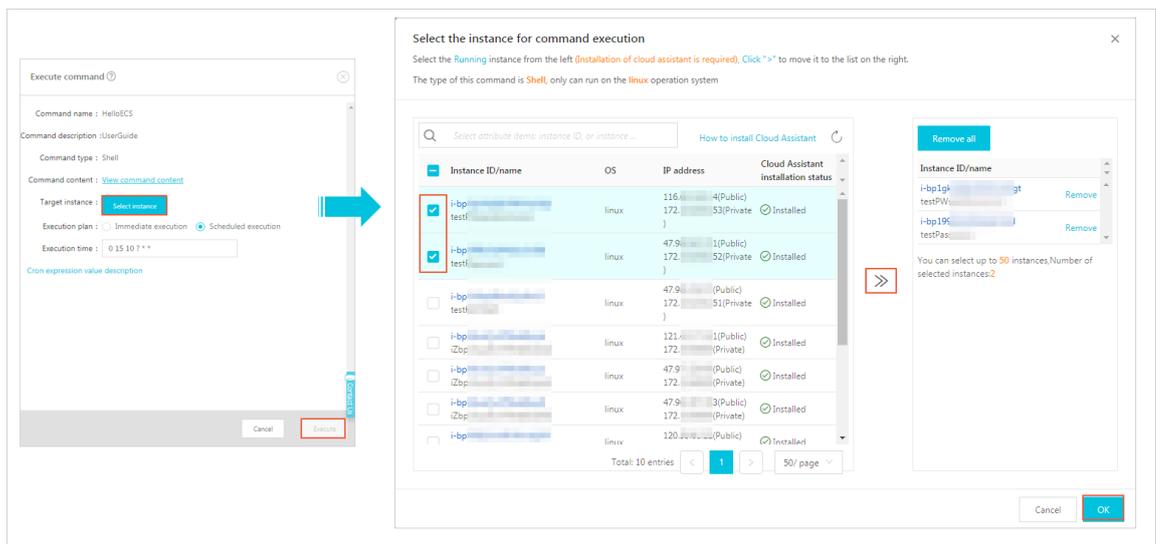
1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Cloud Assistant**.
3. Select the target region.
4. Search for the Cloud Assistant command you want to run, and then select **Execute** from the **Operation** column. In the pop-up window that is displayed, configure the following parameters:
 - a. Click **View command content** to confirm the command contents.
 - b. Click **Select Instance** and perform the following actions:
 - A. Select one or more instances.
 - B. Click  to add the selected instance or instances.
 - c. Click **OK**.
- c. Select **Immediate Execution** or **Scheduled Execution**:



Note:

For Windows instances, you can select Bat or PowerShell commands. For Linux instances, you can only select Shell commands. All instances must have the cloud assistant client **installed**. Otherwise, it cannot be selected.

- **Immediate Execution:** The cloud assistant will run the command immediately on the instances once.
- **Scheduled Execution:** The cron expression will be used to run the command periodically. Enter the required **Execution Time**. For more information, see [Cron expression value description](#).



5. Click Execute.

You can also use the ECS API [InvokeCommand](#) to execute a cloud assistant command.

Stop command execution

Prerequisite: The command must be a periodic command, or the command must be in **Running** status.

To stop a command on the ECS console, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Cloud Assistant**.
3. Select the target region.
4. In the **Execution Record** area, search for the command you want to stop, and select **Stop Command** from the **Operation** column.

Execution status	Command execution ID	Command ID/name	Command type	Periodical execution	Execution frequency	Target instance	Operation
In progress	t-d8d4c7...	c-c4f214e50 HelloECS	Shell	Yes	0 15 10 ? * *	1	View result Stop execution
Execution completed	t-eb5869...	c-c4f214e50 HelloECS	Shell	No		1	View result
In progress	t-52f274...	c-4295d46c5 HelloECS	Shell	No		1	View result Stop execution

What to do next

[Query execution results and status.](#)

12.5 Query execution results and statuses

We recommend that you review the results and status of a command execution after running a command to ensure the target operation has completed properly.

Prerequisite

The command has been run at least once.

Check the results of the command execution

To view command execution results on the ECS console, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Cloud Assistant**.
3. Select the target region.
4. In the **Execution Record** area, search for the execution record of the target command execution, and select **View Results** from the **Operation** column.
5. In the pop-up window that is displayed, select an execution record and click  to expand the command execution record.

You can also use the ECS API [DescribeInvocationResults](#) to view command results.

View command execution status

To view the status of a command execution in the ECS console, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Cloud Assistant**.
3. Select the target region.
4. In the **Execution Record** area, search for the execution record of the target command execution, and view the status of the command execution in the **Execution Status** column.

Execution status	Command execution ID	Command ID/name	Command type	Periodical execution	Execution frequency	Target instance	Operation
In progress	t-d8d4c7...	c-c4f214e50 HelloECS	Shell	Yes	0 15 10 ? * *	1	View result Stop execution
Execution completed	t-eb5869...	c-c4f214e50 HelloECS	Shell	No		1	View result
In progress	t-52f274...	c-4295d46c5 HelloECS	Shell	No		1	View result Stop execution

You can also use the ECS API [DescribeInvocations](#) to view command execution status.

Lifecycle of command invocation

A command may have the following status when running on an instance.

Table 12-1: The status of commands executed on an instance

Command status	API status	Description
Being executed	Running	The command is being executed.
Stopped	Stopped	You stop a command during its execution.
Execution finished	Finished	The command invocation is finished. However, this does not mean the invocation is successful. You can check whether the invocation is successful by checking the actual output (<code>Output</code>) of the command process.
Execution failed	Failed	The command invocation is not finished for some reason when the timeout time (<code>Timeout</code>) is reached.

To facilitate the management of bulk or periodical execution, you can manage the lifecycle of command execution from the perspectives of **overall invocation status**, **instance invocation status**, and **invocation record status**. The relationships among various levels are shown in the following figure.

Figure 12-1: Relationships among the invocation status

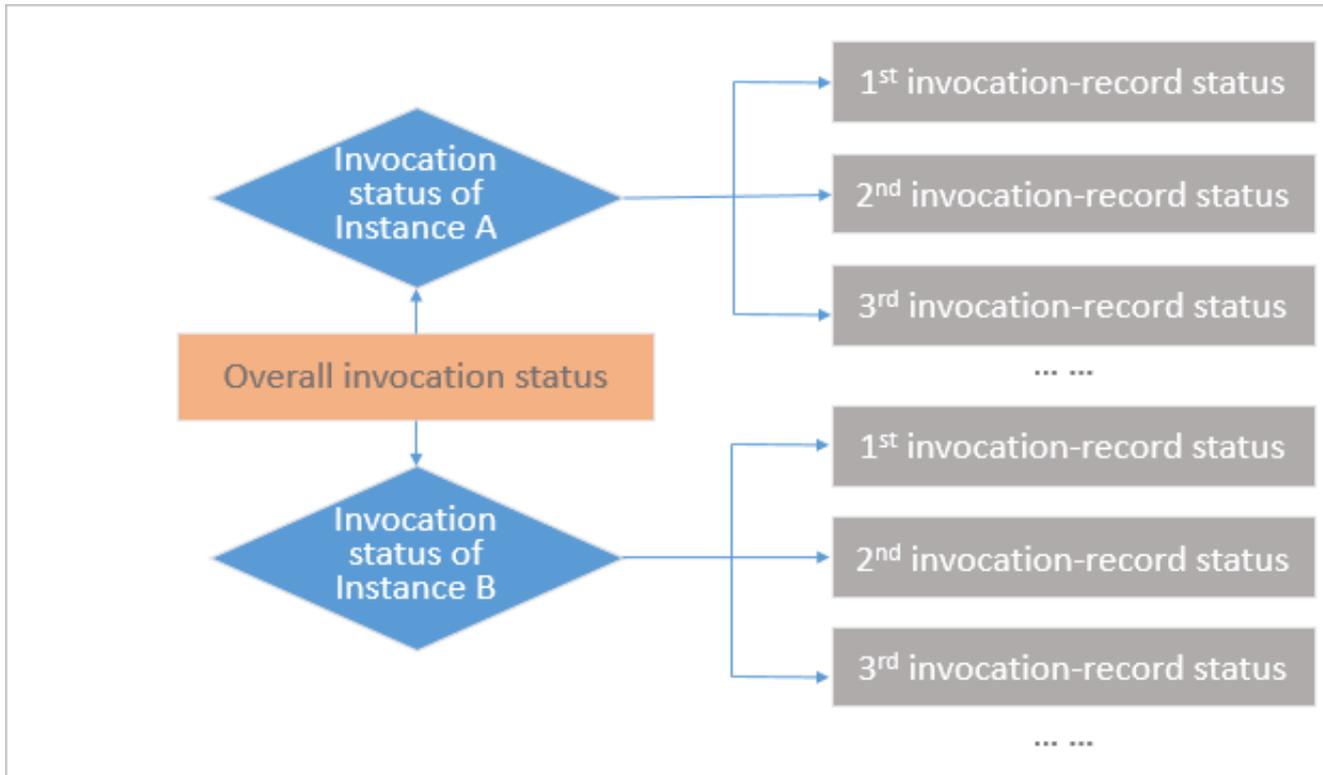


Table 12-2: Lifecycle of one-time batched execution

Status	Command invocation	Displayed status
Overall invocation status	The invocation status of all instances are Finished (Finished).	Finished
	The invocation status of some instances are Finished (Finished), while the status of other instances are Stopped (Stopped).	
	The invocation status of all instances are Failed (Failed).	Failed
	The invocation status of all instances are Stopped (Stopped).	Stopped
	The invocation status of all instances are Running (Running), or that of some instances is Running (Running).	Running
	The invocation status of some instances are Failed (Failed).	Partially failed
Instance invocation status	One-time batched execution is a one-off operation, so the instance invocation status is the same as the invocation record status.	

Status	Command invocation	Displayed status
Invocation record status	See the table The status of commands executed on an instance .	

Take three ECS instances for example. The following figure shows the relationships between the overall invocation status and the instance invocation status during a one-time execution on multiple instances.

Figure 12-2: Lifecycle of one-time batched execution

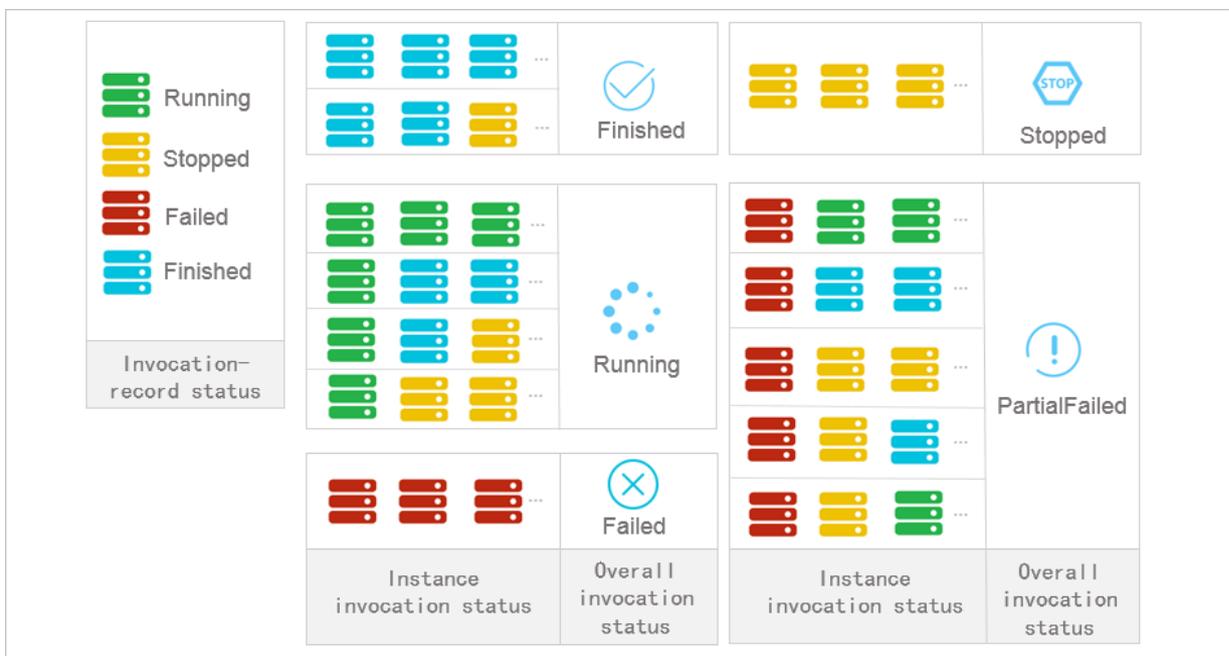


Table 12-3: Lifecycle of periodical execution

Status	Description
Overall invocation status	The overall invocation status remains Running (Running) unless you stop the invocation on all instances.
Instance invocation status	The instance invocation status remains Running (Running) unless you stop the invocation.
Invocation record status	See the table The status of commands executed on an instance .

13 Elastic Network Interfaces

13.1 Attach an ENI when creating an instance

You can attach an ENI when creating an ECS instance in the ECS console. For more information about instance creation, see [create an instance](#). This document focuses on the notes for attaching an ENI during ECS instance creation.

Note the following configurations when attaching an ENI during ECS instance creation:

- Basic configurations:
 - Region: ENIs are supported in all regions.
 - Instance type: Select an instance type that supports ENI. The selected instance type must be I/O optimized.
 - Image: Only the following types of image can automatically recognize ENIs without any additional configuration. For other images, you must configure the ENI to enable the created instance to recognize it.
 - Centos 7.3 64-bit
 - Centos 6.8 64-bit
 - Windows Server 2016 Data Center Edition 64-bit
 - Windows Server 2012 R2 Data Center Edition 64-bit
 - Networking:
 - Network: Select **VPC**, and then select a created VPC and a VSwitch.
 - ENI: Click **Add ENI** to attach an ENI, and then select a VSwitch for the ENI.



Note:

- You are only allowed to attach a maximum of two ENIs when creating an instance in the console. One of them is the primary ENI, which is attached automatically, and the other one is a secondary ENI.
- After the instance is started, you can attach more secondary ENIs to the instance based on the instance type in the console or by using the [AttachNetworkInterface](#) API.

If you want to keep the secondary ENI that is created in this way, detach it from the instance before you release the instance.

13.2 Create an ENI

You can create an elastic network card individually. You can create an ENI in the ECS console, and then [attach it to an instance](#).

This document describes how to create an ENI in the ECS console.

Limits

To create an ENI, you have the following limits:

- Each ENI must be in a VSwitch of a VPC.
- Each ENI must be in one security group.

Prerequisites

Before you create an ENI, finish the following operations:

- Create a VPC and then create a VSwitch in the VPC.
- Create a security group in the same VPC.

Procedure

To create an ENI, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Networks and Security** > **ENI**.
3. Select a region.
4. Click **Create ENI**.
5. In the **Create ENI** dialog box, finish the following configurations:
 - a. **Network Interface Name**: Specify a name for the ENI.
 - b. **VPC**: Select a VPC. When you attach an ENI to an instance, they must be in the same VPC.

**Note:**

In addition, after an ENI is created, you cannot change the VPC.

- c. **VSwitch**: Select a VSwitch. When you attach an ENI to an instance, they must be in the same zone, but they do not have to be in the same VSwitch.

**Note:**

In addition, after an ENI is created, you cannot change the VSwitch.

- d. **Primary Private IP:** Specify an IPv4 address as the private IP address of the ENI. The IPv4 address must be available in the CIDR block of the specified VSwitch. If you do not specify one, a private IP address is automatically assigned to your ENI after the ENI is created.
- e. **Security Group:** Select a security group in the selected VPC.
- f. **Description:** Give a brief description for the ENI for easing further management.
- g. Click **OK**.

In the Network Interfaces page, refresh the table. When the new ENI is in the **Available** status, it is created successfully.

What to do next

After you create an ENI, you can

- [attach an ENI to an instance](#).
- [modify attributes of the ENI](#).
- [delete the ENI](#).

13.3 Attach an ENI to an instance

You can attach an ENI to an instance.

This topic describes how to attach an ENI to an instance in the ECS console.

Limits

To attach an ENI to an instance, you have the following limits:

- You can only attach a secondary ENI to an instance.
- The ENI must be in the **Available** status.
- The instance must be in the **Stopped** or **Running** status.
- You can only attach an ENI to a VPC-Connected ECS instance, and they must be in the same VPC.
- The VSwitches of the ENI and the instance can be different, but they must be in the same zone.
- An ENI can be attached to an I/O optimized ECS instance only.
- An ENI can only be attached to one VPC-Connected ECS instance at a time. However, a VPC-Connected ECS instance can be associated with multiple ENIs. For more information about the maximum number of ENIs that can be attached to one instance, see [instance type families](#).

Prerequisites

Before you attach an ENI to an instance, finish the following operations:

- [Create an ENI](#).
- Make sure the ENI is in the **Available** status.
- Make sure your instance can be associated with secondary ENIs, and the instance is in the **Stopped** or **Running** status. For the number of ENIs that can be attached to each instance type, see the [instance type families](#).

Procedure

To attach an ENI to an instance, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Networks and Security > ENI**.
3. Select a region.
4. Find an available ENI, and in the **Actions** column, click **Bind to Instance**.
5. In the **Bind to Instance** dialog box, after selecting an instance, click **OK**.

In the **Network Interfaces** page, refresh the table. When the selected ENI is in the **Bound** status, it is successfully attached to the instance.

What to do next

After an ENI is attached to an instance, you can:

- [Detach the ENI from an instance](#), and then [delete the ENI](#).
- [Modify attributes of the ENI](#).
- [Configure the ENI](#), if the ENI cannot be automatically recognized by the operating system of your instance.

13.4 Detach an ENI from an instance

You can detach a secondary ENI, but not the primary ENI, from an instance.

Limits

To detach a secondary ENI from an instance, you have the following limits:

- The secondary ENI must be in the **Bound** status.
- The instance must be in the **Stopped** or **Running** status.

Prerequisites

Your ENI *is attached to an instance*. Before you detach an ENI from an instance, the instance must be in the **Stopped** or **Running** status.

Procedure

To detach a secondary ENI from an instance, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Networks and Security > ENI**.
3. Select a region.
4. Find an ENI in the **Bound** status, and in the **Actions** column, click **Unbind**.
5. In the **Unbind** dialog box, confirm the information, and then click **OK**.

In the Network Interfaces page, refresh the table. When the selected ENI is in the **Available** status, it is successfully detached from the instance.

What to do next

After an ENI is detached from an instance, you can:

- [Attach the ENI to another instance](#).
- [Delete the ENI](#).
- [Modify attributes of the ENI](#).

13.5 Modify attributes of an ENI

You can modify attributes of a secondary ENI, but not the primary ENI, of an instance. You can modify the following attributes of an ENI:

- The name of the ENI.
- The security group associated with the ENI. One ENI must be associated with at least one security group. However, it cannot be associated with more than five security groups.
- Description of the ENI.

You can modify attributes of an ENI when it is in the **Available** or the **Bound** status. This document describes how to modify attributes of an ENI in the ECS console.

Prerequisite

Before you modify attributes of an ENI, [create an ENI](#).

Procedure

To modify attributes of an ENI, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Networks and Security > ENI**.
3. Select a region.
4. Find an ENI, and in the **Actions** column, click **Modify**.
5. In the **Modify** dialog box, modify the following optional configurations:
 - **Network Interface Name**: Specify a new name for the selected ENI.
 - **Security Group**: Select more security groups for the ENI, or remove security groups. Retain at least one security group.
 - **Description**: Give a brief description for the ENI.

After you finish the modification, click **OK**.

13.6 Delete an ENI

If you do not require an ENI, you can delete it. But you can only delete a secondary ENI, but not the primary ENI of an instance.

After an ENI is deleted,

- the primary private IP address of the ENI is released automatically,
- and the ENI is automatically removed from all associated security groups.

An ENI will be deleted along with an instance if you did not detach it from the instance before you release the instance.

Limits

You can only delete an ENI in the **Available** status.

Prerequisite

If an ENI is *attached to an instance*, [detach it from the instance](#).

Procedure

To delete an ENI, follow these steps:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Networks and Security > ENI**.
3. Select a region.

4. Find an available ENI, and in the **Actions** column, click **Delete**.
5. In the dialog box, click **OK**.

In the **Network Interfaces** page, refresh the table. If the ENI disappears, it is deleted successfully.

13.7 Configure an ENI

If your instance is running one of the following images, you do not have to configure the Elastic Network Interfaces (ENI) manually to have them recognized by the OS.

- Centos 7.3 64-bit
- Centos 6.8 64-bit
- 64-bit Windows Server 2016 data center Edition
- Windows Server 2012 R2 Data Center Edition 64-bit64-bit Windows Server 2012 R2 data center Edition

If your instance is running none of the preceding images, and you want to attach an ENI to your instance, you must manually configure the ENI to be recognizable. If your instance does not use these images, however, if you want to attach a flexible network card to an instance, You need to manually configure the elastic network card. This document uses an instance running CentOS 7.2 64-bit as an example to introduce how to configure an ENI to make the interface recognizable.

Prerequisite

You have attached an elastic network card to an ECS instance.

Procedure

To configure the ENI, follow these steps:

1. Use the [DescribeNetworkInterfaces](#) interface or log on to the ECS console to obtain the following attributes of the ENI: primary private IP address, subnet mask, the default route, and the MAC address. To obtain these attributes in the ECS console, follow these steps: MAC address. Do the following on the console.
 - a. Log on to the [ECS console](#).
 - b. Find a network interface, and obtain its primary private IP address, subnet mask, default route, and MAC address. Locate the primary and private IP address, mask address, default route, and MAC for each network interface. Example:

```
eth1 10.0.0.20/24 10.0.0.253 00: 16: 12: E7: 27
```

```
eth2 10.0.0.21/24 10.0.0.253 00:16:12:16:EC
```

2. [Connect to the ECS instance.](#)

3. Run the command to generate the config file `cat /etc/sysconfig/network-scripts/ifcfg-[network interface name in the OS]`.



Note:

- Pay attention to the relation between network interface name in the OS and the MAC address.
- Pay attention to the relation between network interface name in the OS and the MAC address. The default route must be set to `DEFROUTE=no`. Other editions must have the same configuration. Note that running the `ifup` command may change the active default route configuration after configuring the network interface.
- Example:

```
# cat /etc/sysconfig/network-scripts/ifcfg-eth1
DEVICE=eth1
BOOTPROTO=dhcp
ONBOOT=yes
TYPE=Ethernet
USERCTL=yes
PEERDNS=no
IPV6INIT = No
PERSISTENT_DHCLIENT = Yes
HWADDR=00:16:3e:12:e7:27
DEFROUTE=noDefroute = No
```

4. Follow these steps to start the network interface:

- a. Run the `ifup [network interface name in the OS]` command to start the `dhclient` process, and initiate a DHCP request. Example:

```
# ifup eth1
# ifup eth2
```

- b. After a response is received, run the `ip a a` command to check the IP allocation on the network interfaces, which must match with the information displayed on the ECS console.

Example:

```
# ip a
1: lo: mtu 65536 qdisc noqueue state UNKNOWN qlen 1
link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
inet 127.0.0.1/8 scope host loInet 125.0.0.1/8 Scope host Lo
valid_lft forever preferred_lft forever
2: eth0: mtu 1500 qdisc pfifo_fast state UP qlen 10002: eth0: MTU
1500 qdisc glasstate up qlen 1000
link/ether 00:16:3e:0e:16:21 brd ff:ff:ff:ff:ff:ff
Inet 10.0.0.19/24 BRD glasscope Global Dynamic eth0
```

```

valid_lft 31506157sec preferred_lft 31506157secValid_lft
31506157sec preferred_lft 31506157sec
3: eth1: MTU 1500 qdisc glasstate up qlen 1000
link/ether 00:16:3e:12:e7:27 brd ff:ff:ff:ff:ff:ff
inet 10.0.0.20/24 brd 10.0.0.255 scope global dynamic eth1Inet 10.
0.0.20/24 BRD glasscope Global Dynamic eth1
Valid_lft 31525994sec preferred_lft 31525994sec
4: eth2: MTU 1500 qdisc glasstate up qlen 1000
Link/ether 00: 16: Rye: 12: 16: ec brd ff: FF: FF
inet 10.0.0.21/24 brd 10.0.0.255 scope global dynamic eth2
valid_lft 31526009sec preferred_lft 31526009sec

```

5. Set the metric for each network interface in the route table. In this example, set the metric parameters of eth1 and eth2 as follows.

```

eth1: gw: 10.0.0.253 metric: 1001
eth2: gw: 10.0.0.253 metric: 1002

```

- a. Run the following command to set the metric parameters.

```

# Ip-4 route add default via glasdev eth1 metric 1001
# ip -4 route add default via 10.0.0.253 dev eth2 metric 1002

```

- b. Run the `route -n` command to check whether the configuration is successful or not.

Example:

```

# route -n
Kernel IP routing table
Destination Gateway Genmask Flags Metric Ref Use Iface
0.0.0.0 10.0.0.253 0.0.0.0 UG 0 0 0 eth0
0.0.0.0 10.0.0.253 0.0.0.0 UG 1001 0 0 eth1
0.5.0.0 10.0.0.253 ug ub1002 0 0 eth2
10.0.0.0 0.5.0.0 255.25.25.0 u 0 0 0 eth0
10.0.0.0 0.0.0.0 255.255.255.0 U 0 0 0 eth1
10.0.0.0 0.5.0.0 255.25.25.0 u 0 0 0 eth2
169.254.0.0 0.0.0.0 255.0.0.0 U 1002 0 0 eth0
169.254.0.0 0.0.0.0 255.255.0.0 U 1003 0 0 eth1
169.254.0.0 0.0.0.0 255.255.0.0 U 1004 0 0 eth2169.254.0.0 0.0.0
255.0.0 U 1004 0 0 eth2

```

6. Follow these steps to build a route table:



Note:

We recommend that you use the metric value as the route table name.

- a. Run the command to build a route table.

```

# ip -4 route add default via 10.0.0.253 dev eth1 table 1001
# Ip-4 route add default via glasdev eth2 table 1002

```

- b. Run the command to check whether the route table is built successfully or not.

```

# ip route list table 1001
default via 10.0.0.253 dev eth1
# ip route list table 1002

```

```
default via 10.0.0.253 dev eth2
```

7. Configure policy routing.

a. Run the following command .

```
# ip -4 rule add from 10.0.0.20 lookup 1001  
# ip -4 rule add from 10.0.0.21 lookup 1002
```

b. Run `ip rule list` to view routing rules.

```
# ip rule list  
0: from all lookup local  
32764: from 10.0.0.21 lookup 1002  
32765: from 10.0.0.20 lookup 1001  
32766: from all lookup main  
32767: from all lookup default
```

At this point, you have completed the configuration of the elastic network card.

14 Tags

14.1 Limits

You can bind tags to the following resources in the ECS console: ECS instance, storage, snapshot, image, and security group.

Tags have the following limits:

- Each tag has a key-value pair.
- You can bind 20 tags to an instance at most. You can bind 20 tags at most to an instance at a time.
- Every tag key of a resource must be unique. A tag with the same key as an existing one will be overwritten.
- Tag information is not shared across regions. For example, tags created in China East 1 (Hangzhou) are invisible to China East 2 (Shanghai).
- If a tag is unbound and no longer bound to any other resource, the tag will be automatically deleted.

14.2 Add a tag to resources

If your account maintains various types of resources that are associated with each other in different ways, you can bind tags to the resources to categorize and manage the resources in a unified manner.

You can bind 20 tags to a resource at most. You can bind/unbind 20 tags at most for the resource each time.

Take the following steps to bind resources with tags:

1. Log on to the [ECS console](#).
2. Select the resource type in the left-side navigation pane for the binding operation, such as **Instance**, **Disks**, **Snapshots**, **Images**, and **Security Groups**.
3. Select a region.
4. Select the resources in the resource list to bind tags.
5. Click **Edit Tag** at the bottom of the resource list.

**Note:**

If the selected resources are instances, select **More > Instance Settings > Edit Tag**.

6. In the dialog box,

- If the selected resource has already been created a tag, click the existing tag, and select the available tags.
- Click **Create** and set **Key** and **Value** if no tags are available for the selected resource.



Note:

- **Key** is required whereas **Value** is optional.
- **Key** cannot start with aliyun, http://, or https://. The key is case-insensitive and can contain up to 64 characters.
- **Value** cannot start with http:// or https://. The value is case-insensitive and can contain up to 128 characters. It can be empty.
- Any tag **Key** of a resource must be unique. A tag with the same key as an existing one will be overwritten.
- **Available Tags** and **Create** are grayed out if the selected resources are already bound with 20 tags. You must unbind some tags before binding new tags.

7. Click **Confirm**.

To check if tags are successfully bound, use the **Edit Tag** function of the resource or click **Tags** in the left-side navigation pane of the ECS console. You can click **Tags** with a tag symbol at the top of the resource list to filter resources.

14.3 Delete a tag

You can unbind a tag from the resource if the tag is no longer applicable to resource management. After a tag is unbound and is no longer bound to any other resource, the tag will be automatically deleted.

- The Delete Tags function unbinds one or more tags from an instance at a time.



Note:

Currently, this function is only available for instances. It is unavailable for other resource types.

- The Edit Tags function unbinds tags one by one.



Note:

You can unbind 20 tags from a resource each time.

Unbind tags from instances using the tag deletion function

Currently, the Delete Tags function is only available for instances.

See the following steps to delete tags:

1. Log on to the [ECS console](#).
2. Click **Instances** from the left-side navigation pane.
3. Select a region.
4. Select the instances from which you want to unbind tags from the instance list.

**Note:**

You can also filter instances by tag and select the expected instance.

5. Click **More > Instance Settings > Delete Tag**.
6. In the **Delete Tag** dialog box, enter the **Tag Key** and **Tag Value** of the tags you want to unbind.
7. Click **OK** to complete tag unbinding.

To check whether the tags are successfully unbound, use the **Edit Tag** function of the instance or click **Tags** in the left-side navigation pane of the ECS console.

Unbind tags from resources using the tag edit function

The Edit Tags function unbinds one or more tags from a resource.

See the following steps to unbind tags:

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select the resource type for the unbinding operation, such as **Instances**, **Disks**, **Snapshots**, **Images**, or **Security Groups**.

**Note:**

The block storage function is now in beta. For more information, see [block storage FAQ](#) Learn more.

3. Select a region.
4. In the resource list, select the resource from which you want to unbind tags.

**Note:**

You can also filter resources by tag and select the expected resource.

5. Click **Edit Tag** at the bottom of the resource list.

6. In the **Edit Tags** dialog box, click the deletion icon next to a tag.
7. Click **Confirm** to complete tag unbinding.

To check whether the tags are successfully unbound, use the Edit Tags function of the resource or click **Tags** in the left-side navigation pane of the ECS console.

14.4 Filter resources by tags

After you bind the tags to the resources, use any of the following methods to filter resources by tags.

Filter resources in resource lists

Follow these steps to filter resources:

1. Log on to [ECS console](#).
2. In the left-side navigation pane, select the resource type you want to view, such as **Instances**, **Disks**, **Snapshots**, **Images**, or **Security Groups**.
3. Select a region.
4. Click **Tag** at the top of the resource list.
 - Click a key to filter out the resources that are bound with this key, which may have multiple values.
 - Click a key and value to filter out the resources that are bound with this key-value pair (tag).

The console returns a list of resources that are bound with the key or a key-value pair.

Filter resources by tags

Follow these steps to filter resources:

1. Log on to [ECS console](#).
2. Click **Tags** in the left-side navigation pane.
3. Select a region.
4. Enter a key in the search box and click **Search**.

The console returns a list of resources that are bound with the key.

15 Monitoring

15.1 Monitoring

Monitoring the status of your ECS instances helps you guarantee that can always quickly access your websites and applications, process data, and render videos. Alibaba Cloud provides data monitoring, visualization of monitoring data, and real-time alerts to make sure that your ECS instances are running without interruption.

Details

You can monitor your ECS instances by using the ECS monitoring service or CloudMonitor. ECS provides CPU utilization, network traffic, and disk I/O monitoring for a specified instance. In CloudMonitor, you can monitor the instances by using a wider range of metrics with finer granularity. For more information about CloudMonitor, see [host monitoring metrics](#). Some of the metrics provided by the ECS monitoring service are shown as follows.

- CPU utilization: The percentage of allocated ECS compute units that are currently in use on the instance. Higher percentage indicates higher CPU load of the instance. You can view the CPU utilization in the ECS console or in the CloudMonitor console. You can also obtain the data by calling the ECS API operations or after connecting to the specified instance through [remote connection](#). The following shows how to view the CPU utilization of different ECS instances after you connect the instance.
 - Windows instance: View the CPU utilization in the **Task Manager**. You can sort the tasks by CPU utilization to find the process that is consuming the CPU of the specified ECS instance.
 - Linux instance: Run the `top` command to view the CPU utilization. To locate the process that is consuming the CPU of the specified ECS instance, press **Shift + P** to sort the tasks by CPU utilization.
- Network traffic: The bandwidth usage for the inbound and outbound traffic of the ECS instance in kbps. ECS provides data connection monitoring, while CloudMonitor can monitor Internet and internal network traffic. If the outbound traffic reaches 1,024 kbps and the outbound bandwidth limit is 1 Mbps, the outbound bandwidth for the specified ECS instance is fully utilized.

ECS monitoring service

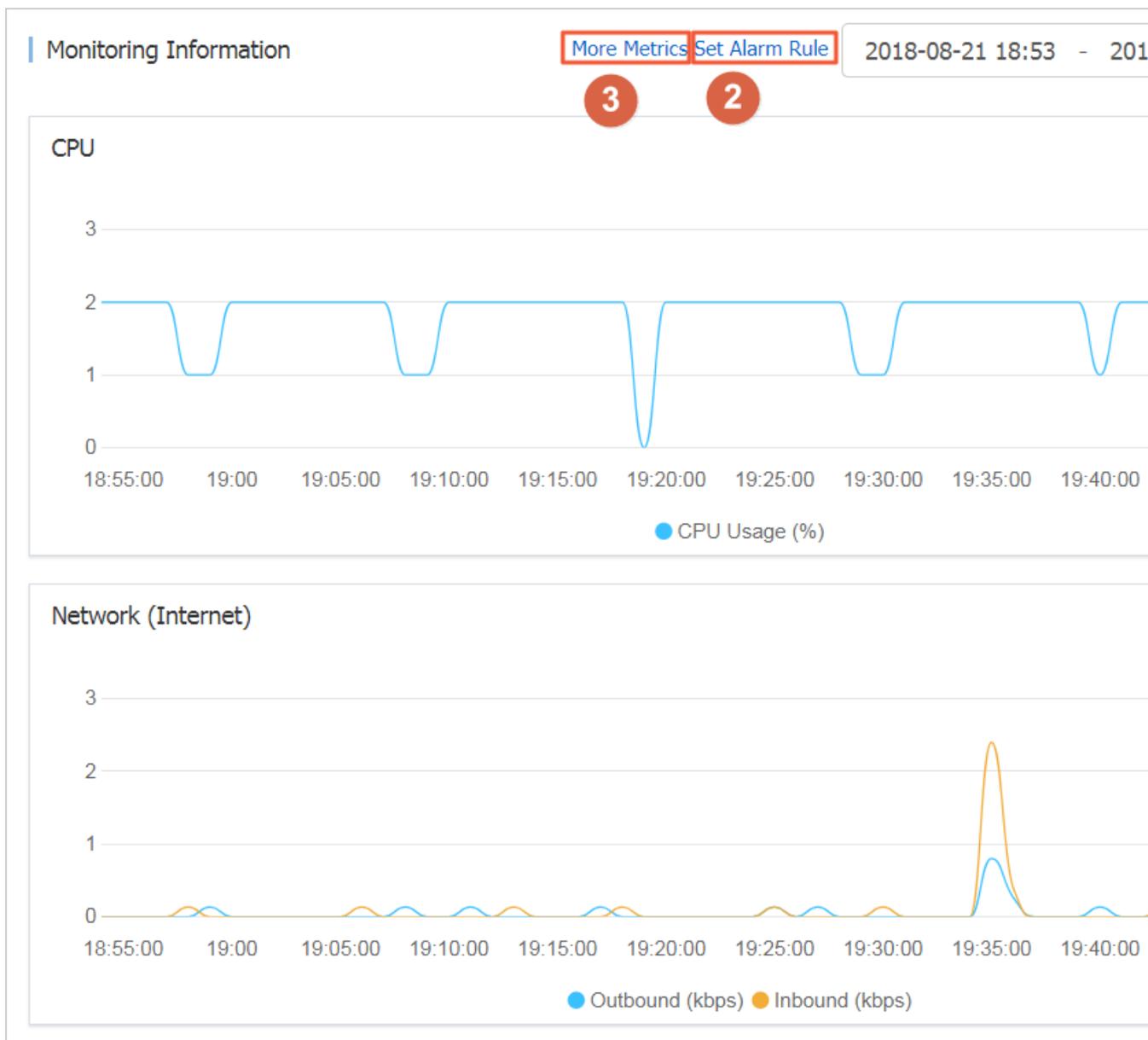
To view the monitoring data in the ECS console, follow these steps.

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Instances**.
3. Select a region.
4. Find the target instance and click the instance ID.
5. On the **Instance Details** page, you can view the **Monitoring Information** including CPU utilization and network traffic.
 - a. Click  to specify the **Start Time** and **End Time**.

**Note:**

The Start Time and End Time you specify affects the granularity of the data display. Smaller sampling intervals result in finer granularity of data displayed. For example, the average values shown will be different when you select a sampling interval of 5 and 15 minutes.

- b. (Optional) Click **Set Alarm Rule** and you will be directed to the CloudMonitor console. Then you can specify the CPU utilization and network traffic alarm rules. For more information about the metrics, see [overview of alarm services](#).
- c. (Optional) Click **More Metrics** and you can view more monitoring data in the CloudMonitor console. It takes a few minutes for the monitoring data to update.



You can also call the ECS API operations [DescribeInstanceMonitorData](#), [DescribeDiskMonitorData](#), and [DescribeEniMonitorData](#) to obtain the monitoring data.

The monitoring metrics in ECS are listed as follows. The sampling interval for each metric is 1 minute.

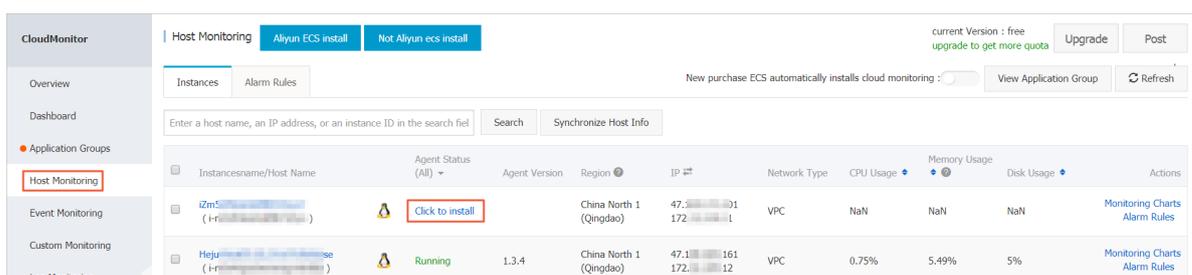
Metric	Description
Instance	Instance ID.
CPU Usage	The percentage of allocated ECS compute units that are currently in use on the instance.
Intranet inbound traffic	The internal network traffic to your instance. Unit: kbits.

Metric	Description
Intranet outbound traffic	The internal network traffic from your instance. Unit: kbits.
Intranet bandwidth	The internal network traffic of the instance per unit time. Unit: kbits/s.
Public network inbound traffic	The Internet traffic to the instance. Unit: kbits.
Public network outbound traffic	The Internet traffic from the instance. Unit: kbits.
Public network bandwidth	Internet traffic of the instance per unit time. Unit: kbits/s.
Disk read IOPS	The number of disk read operations per second.
Disk write IOPS	The number of disk write operations per second.
Disk read BPS	The number of bytes read from disk per second. Unit: Byte/s.
Disk write BPS	The number of bytes written to disk per second. Unit: Byte/s.

CloudMonitor

CloudMonitor provides one-stop and out-of-the-box monitoring solutions for enterprises in the cloud. It offers host monitoring service for your ECS instances. For more information about CloudMonitor, see [introduction to Host monitoring](#). To view the monitoring data of your ECS instance in the CloudMonitor console, follow these steps.

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, click **Host Monitoring**.
3. Find your target instance.
4. (Optional) If your instance has not been installed with the CloudMonitor agent, click **Click to install**.
5. To obtain the monitoring data, click **Monitoring Charts** from the **Actions** column.
6. To set alarm rules, click **Alarm Rules** from the **Actions** column.



About bandwidth units

Differences between Kb and KB

- A bit (b) is the smallest unit of data in a computer. A bit has a single binary value, either 0 or 1. Eight bits forms a Byte. For example, 0101 0010. 1 Byte = 8 bits (1B = 8b).
- If K or k indicates kilo, one Kb equals one thousand bits, while a kilobyte (KB) equals 1,024 bytes.

In the ECS monitoring service, network traffic is measured in kbps, which is kilobit per second. Kbps indicates network speed, which is the number of kilobits transmitted per second. The unit bps is usually omitted when bandwidth is described. For example, the full form of 4M in the bandwidth scenario is 4 Mbps.

Relations between bandwidth and download speed

- Common misunderstanding: Bandwidth is equivalent to download speed.
- In theory, if a network bandwidth is 1 Mbps, the download speed can reach 125 KB/s. Download unit conversions are as follows: 1 KB = 8 Kb, 1 Mbps = 125 KB/s, 1 kbps = 1,000 bps.

However, some applications running on the instance consume a small amount of bandwidth, such as remote desktop programs. Therefore, the actual download speed is usually between 100-110 KB/s.

15.2 System events

System events are scheduled and recorded maintenance events of your ECS resources. System events occur when security updates, invalid operations, expiration of Subscription instances, overdue payment, or unexpected failures are detected in your ECS instances. Your instances will start, restart, stop, or be released when system events occur.

Routine maintenance versus system events

ECS instances are the core component used to establish your applications. After you select and start ECS instances, initiate configuration, and start to deploy applications, the health of the ECS instance is crucial to your business. To guarantee the backend performance and security of ECS, we perform routine maintenance for the physical servers. When we scan for the hardware and software faults or potential risks on the physical servers, we live-migrate your instances to healthy servers. This is routine maintenance. Unlike system events, you do not receive any notification and also, your instances are not impacted, while the routine maintenance is in progress.

Once system events occur, you are notified about the default actions and the time scheduled to perform these actions on your instances. For planned system events, information such as the impact of the event on the instance and the expected execution point is told in advance. To prevent impact on your business, we recommend that you back up the data and distribute incoming traffic before handling system events. You can query the system events history for the last week later, for further analysis of faulty diagnosis and faulty replay.

Limits

Phased-out instance types, including but not limited to sn2, sn1, t1, s1, s2, s3, m1, m2, c1, c2, c4, ce4, cm4, n1, n2 and e3, do not support system events. For more information, see [instance type families](#).

Event types

The following table describes the types of ECS system events.

Category	Event type	Parameter
Scheduled restart	An instance restarts after planned system maintenance or security update.	<code>SystemMaintenance.Reboot</code>
Unexpected restart	An instance restarts after unexpected system failures.	<code>SystemFailure.Reboot</code>
	An instance restarts after unexpected instance failures.	<code>InstanceFailure.Reboot</code>
Stop instances	Subscription instances stop due to expiration.	<code>InstanceExpiration.Stop</code>
	Pay-As-You-Go instances stop due to overdue payment.	<code>AccountUnbalanced.Stop</code>
Release instances	Subscription instances are released after several days of expiration.	<code>InstanceExpiration.Delete</code>
	Pay-As-You-Go instances are released due after several days of overdue payment.	<code>AccountUnbalanced.Delete</code>

Event status

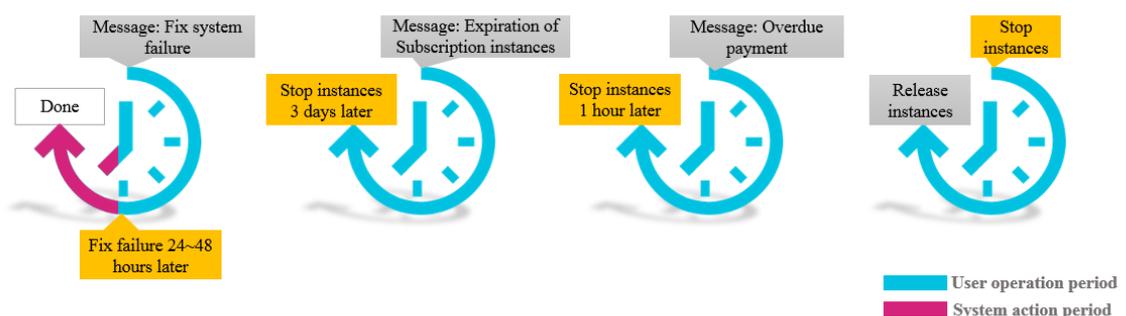
The following table describes the status of a system event during its lifecycle.

Status	Status attribute	Description
Scheduled	Intermediate status	The system event is scheduled but not performed.
Avoided	Stable status	You have taken the actions in advance within the <i>user operation period</i> .
Executing	Intermediate state	The response plan of the system event is being performed.
Executed	Stable status	The system event has been fixed.
Canceled	Stable status	ECS cancels the scheduled system event.
Failed	Stable status	The system event is not fixed.

System event periods

System events observe the following two periods:

- User operation period:** The period between initiation and scheduled time of system events. Normally, you receive a notification from 24 to 48 hours before a system failure event is fixed, from 3 days before a Subscription instances is stopped, and 1 hour before a Pay-As-You-Go instance is stopped. Instances are released 15 days later if no renewal or recharge are made. During this period, you can choose the recommended methods to handle system events in advance. You can also wait until the default actions are triggered.
- System action period:** Generally, if you wait until we take the default action, system events are automatically fixed within 6 hours after the system action period begins at a scheduled time. Later you receive the report of system events.



Note:

Only scheduled system events have user operation period. Unexpected system events that are caused by emergency failures or invalid operations do not have user operation periods.

Once unexpected system events occur, you will receive notifications, but you cannot take any action. However, you can query the system events history for fault diagnosis, cause analysis, or data recovery.

View system events

If a system event is scheduled, the **Pending Tasks** button in the ECS console shows a highlighted tag to remind you to check the event.

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, select **Overview**.
3. Select **Pending Tasks** from the navigation pane on the right-side of the **Overview** page.
4. On the **Pending Tasks** page, you can see the list of instance ID, region, status, system tasks, recommended user operations, and buttons for operations. Optionally, you can choose recommended user operations under the **Actions** column to handle the system events.

API operation: Call [DescribeInstancesFullStatus](#) to view system events.

View system events history

On the All events page, you can query the system events history within the last week for faulty diagnosis and faulty replay.

1. [ECS console](#)
2. On the left-side navigation pane, select **Overview**.
3. Select **Pending Tasks** from the navigation pane on the right-side of the **Overview** page.
4. Click **All Tasks**, and on the **All Tasks** page, click **System Tasks** > **Instances** to see the list of instance ID, event type, region, and task status.

API operation: Call [DescribeInstanceHistoryEvents](#) to view system events history.

System event suggestions

System events make you perceptible to underlying components of Alibaba Cloud ECS. You can optimize the O&M of instances based on system events. We recommend that you use the following actions to handle system events.

Event type	Parameter	Recommended
An instance restarts after pending system maintenance.	SystemMaintenance.Reboot	Use either of the following methods at a convenient time within the user operation period: <ul style="list-style-type: none"> • Restart the instance in the ECS console.

Event type	Parameter	Recommended
		<ul style="list-style-type: none"> Call API RebootInstance. <div style="border: 1px solid gray; background-color: #f0f0f0; padding: 5px; margin-top: 10px;">  Note: Instance restart performed in the instance or from the instance list has no effect on this type of system events. </div> <p>We recommend that you create snapshots (CreateSnapshot) for the attached disks to back up your data.</p>
An instance restarts after unexpected system failures.	SystemFailure.Reboot	When you receive the notification, your instances are being restarted. We recommend that you verify the recovery of instances and applications after the event.
An instance restarts after unexpected instance failures.	InstanceFailure.Reboot	When you receive the notification, your instances are being restarted. We recommend that you: <ul style="list-style-type: none"> Verify the recovery of instances and applications. Analyze the cause of instance crashes to prevent potential events.
A Subscription instance stops due to expiration.	InstanceExpiration.Stop	You can either renew the instances or wait for the instances to stop.
A Pay-As-You-Go instance stops due to overdue payment.	AccountUnbalanced.Stop	You can either keep sufficient balance of your credit card or wait for the instances to stop.
A Subscription instance is released due to expiration.	InstanceExpiration.Delete	You can either renew the instances or wait for the instances to be released.
A Pay-As-You-Go instance is released due to overdue payment.	AccountUnbalanced.Delete	You can either keep sufficient balance of your credit card or wait for the instances to be released.

15.3 Console output and screenshot

ECS instances are virtualized cloud-based services that cannot be connected to any display devices and prohibit mobile snapshots. However, the console output of instances are cached at the time of the last startup, restart, or shutdown event. Moreover, you can obtain instance screenshots in real time. We recommend that you use these features to analyze and troubleshoot instance faults, such as operating system exception diagnosis, abnormal reboots, or unable to connect to instances.

Limits

- Instances running Windows Server image do not allow you to obtain console output.
- [Phased-out instance types](#) do not allow you to obtain instance console output or screenshots.
- You cannot obtain console output or screenshots for instances created before January 1, 2018.

Prerequisite

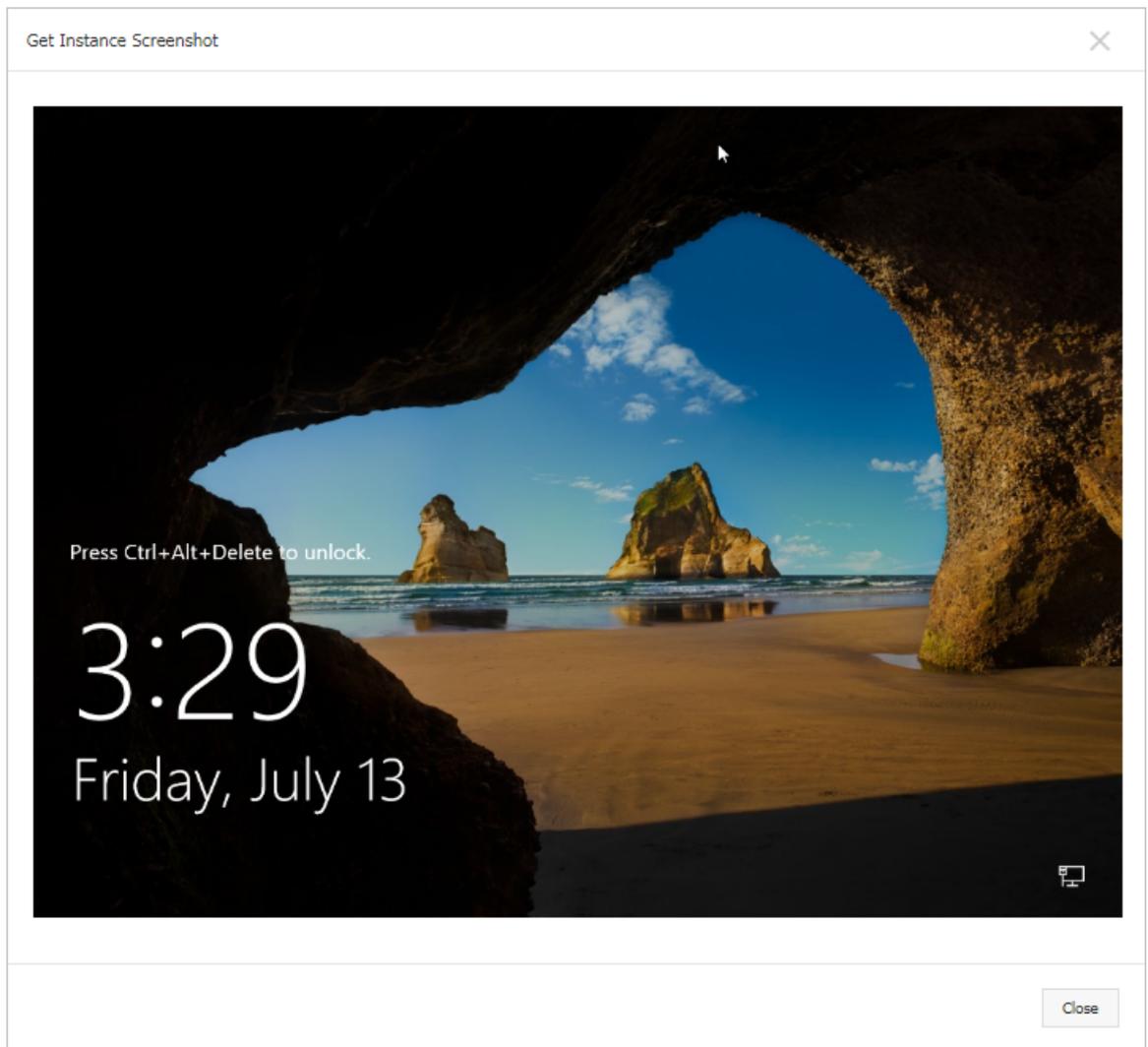
The instance must be in the **Running** status. For more information, see [overview](#).

Procedure

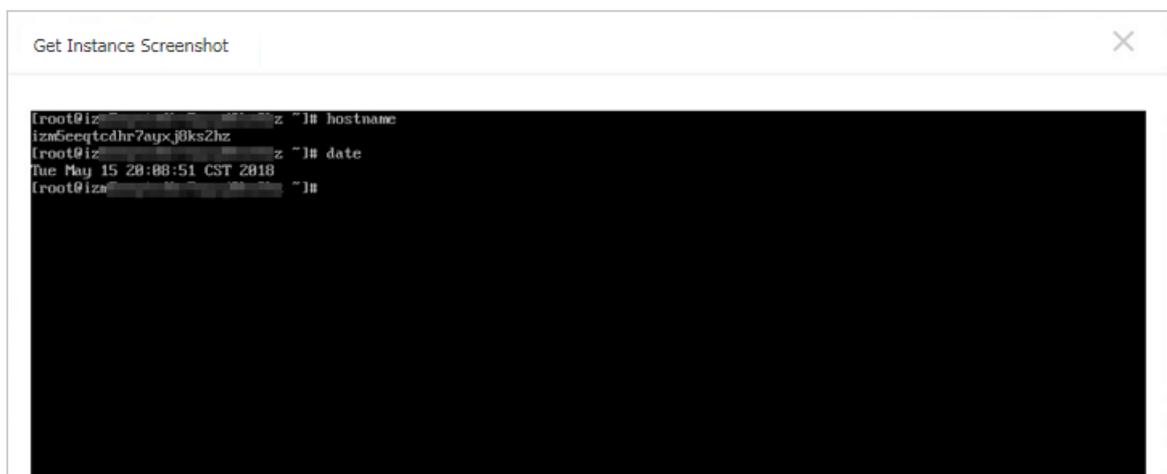
You can view instance console output and screenshot from the **Instance Details** page, the **Instances** list page, or by calling API.

Operation in Instance Details page

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Instances**.
3. Select a region.
4. Select and click the instance to troubleshoot and go to the **Instance Details** page.
5. Click **More > Get Instance Screenshot** to view the screenshot. Alternatively, click **More > Get Instance Console Output** to monitor the root console.
6. Check the instance screenshot or console output.
 - Windows instance screenshot sample:



- Linux instance screenshot sample:



- Linux instance console output sample:

```
Get Instance Console Output

□G      Welcome to CentOS

Starting udev: +%G[ OK ]

Setting hostname AliYunOS: [ OK ]

Setting up Logical Volume Management: [ OK ]

Checking filesystems

Checking all file systems.

[/sbin/fsck.ext4 (1) -- /] fsck.ext4 -a /dev/vda1

/dev/vda1 primary superblock features different from backup, check forced.

/dev/vda1: ***** REBOOT LINUX *****

/dev/vda1: 28300/2621440 files (0.2% non-contiguous), 451117/10485248 blocks

Unmounting file systems

Automatic reboot in progress.

Restarting system.
□G      Welcome to CentOS
```

Operation in Instances list page

1. Log on to the [ECS console](#).
2. In the left-side navigation pane, click **Instances**.
3. Select a region.
4. Locate the instance to troubleshoot.
5. In the **Actions** column, click **More > Operations and Troubleshooting > Get Instance Screenshot** to view the screenshot. Alternatively, click **More > Operations and Troubleshooting > Get Instance Console Output** to monitor the root console.
6. Check the instance screenshot or console output.

API operations

- Instance screenshots: [GetInstanceScreenshot](#)

- Instance console output: [GetInstanceConsoleOutput](#)

What to do next

For other troubleshooting instructions, see [Link testing tool for ping packet loss or ping failure](#).