Alibaba Cloud
Elastic Compute Service

User Guide

MORE THAN JUST CLOUD | C-コ Alibaba Cloud

Legal disclaimer

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

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

2. No part of this document shall be excerpted, translated, reproduced, transmitted, or disseminat ed 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, adjustment s, 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 intellectu al 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 `cd /d C:/windows` command to enter the Windows system folder.
Italics	It is used for parameters and variables.	`bae log list --instanceid` *Instance_ID*
[] or [a|b]	It indicates that it is a optional value, and only one item can be selected.	`ipconfig [-all|-t]`
{} or {a|b}	It indicates that it is a required value, and only one item can be selected.	`swich {stand | slave}`

Contents

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

To use an ECS instance, follow these steps:

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:

   a. For any type of operating system, use the *Management Terminal* for scenarios involving troubleshooting and maintenance.

   b. 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*.

   c. For Windows OSs, you can *connect to a Windows instance*.

3. *Stop the ECS instance*.

4. *Release the instance*.

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

Prohibitions

- 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.

Suggestions

- 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 no 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 *Limitations*.

3 Limitations

When using ECS, note the following limitations:

- ECS instances do not support virtual application installation or revirtualization (such as installation of VMware). Currently, only *ECS Bare Metal Instance and Super Computing Clusters* supports revirtualization.

- ECS instances do not support sound card applications.

- You cannot mount external hardware devices directly (such as a USB or dongle). Instead, use two-step verification with dynamic passwords.

- ECS does not support IP packet address translation services such as SNAT. Instead, use 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 limitations, the following table details further limitations of ECS and states whether you can submit a ticket to request changing the limitation.

ECS instances

Item	Limitation	Can I submit a ticket?
Instance types for which you can create Pay-As-You-Go instances	Instance types with less than 16 vCPUs	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, e3	No

Block storage

Item	Limitation	Can I submit a ticket?
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 × 5	Yes
Quota of system disks for one instance	1	No
Quota of data disks for one instance	16 (including cloud disks and Shared Block Storage)	No
Quota of instances to which one shared block storage can be attached	8	No
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
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

Item	Limitation	Can I submit a ticket?
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	• 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	• Basic Cloud Disk: 5 GiB−2,000 GiB • SSD Cloud Disk/Ultra Cloud Disk/SSD Shared Block Storage/Ultra Shared Block Storage: 20 GiB−32,768 GiB • Local disk: Dependent on specific disks	No
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 allowed	No
System disk mount points	/dev/xvda	No
Data disk mount points	/dev/xvd[b-z]	No

Snapshots

Item	Limitation	Can I submit a ticket?
Quota of snapshots	Number of elastic block storage devices × 64	No

Images

Item	Limitation	Can I submit a ticket?
Quota of custom images in all regions 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	Limitation	Can I submit a ticket?
Quota of key pairs in all regions for one account	500	No
Instance types supporting key pairs	All instance types except for non-I/O optimized instance types in Generation I	No
Images supporting key pairs	Linux images only	No

Internet bandwidth

Item	Limitation	Can I submit a ticket?
Maximum inbound Internet bandwidth	200 Mbit/s	No
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 group

Item	Limitation	Can I submit a ticket?
Quota of instances/IP for one security group	• Security groups for classic network instances: 1,000 classic network instances • Security groups for VPC instances: 2,000 private	No

Item	Limitation	Can I submit a ticket?
	IP (shared by primary and secondary network cards)	
Quota of authorization rules for one security group	100	No
Quota of security groups in a region for an account	100 (increases with membership levels)	Yes
Quota of security groups to which each elastic network interface 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.	Yes. For more information, see *Request for enabling TCP port 25 .*

ENI

Item	Limitation	Can I submit a ticket?
Quota of ENIs in one region for one account	100 (increases with membership levels)	Yes

Label

Item	Limitation	Can I submit a ticket?
Quota of tags that can be bound to one instance	20	No

API

Item	Limitation	Can I submit a ticket?
Quota of CreateInstance calls	200 times per minute	Yes

> 📋 **Note:**
>
> For more information about limitations of VPC products, see *Limits*.

4 Instances

4.1 Launch template

4.1.1 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 Template**, and then click **Create Template**.
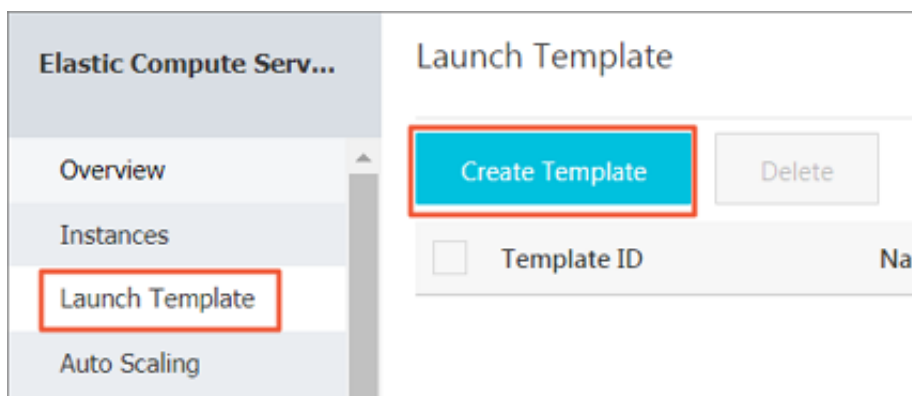
3. Go to the **Launch Template** page and complete the basic configurations and advanced configurations.

> 📋 **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.

5. 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.

3. On the **Preview** page, click **Save as launch template**.

4. In the dialog box that appears, select **Create Template**, enter a template name and description, and then click **Save**.

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

4.1.2 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.

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.
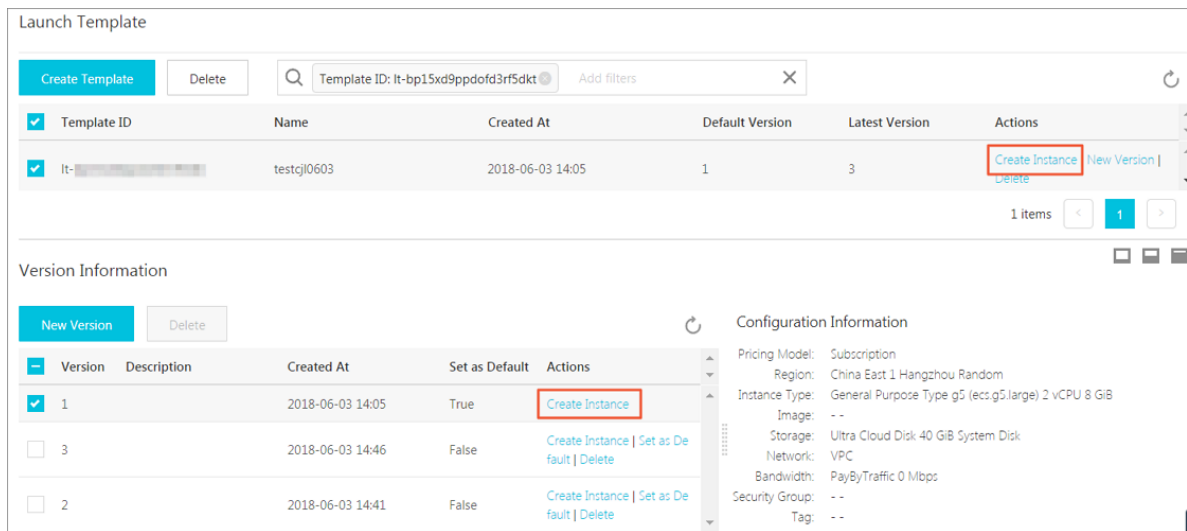

4.1.3 Use a launch template

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 Template**.

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.

5. 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.

4.1.4 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

1. Log on to the *ECS console*.

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

3. Select the target template ID.

4. 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.

5. Click **OK**.

Delete a template

1. Log on to the *ECS console*.

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

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.

4.2 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 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 **Set Display Items** 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**.

Set Display Items   ✕

☑ Operating System   ☐ Monitor   ☑ Zone   ☑ IP Address

☑ Status   ☑ Network Type   ☐ Configuration   ☐ VPC Details

☐ Tags   ☑ Instance Type Family ☑ Billing Method ☑ Automatic Renewal

☐ Key Pairs   ☐ Link Status   ☐ RAM Role   ☐ Stop Instance

                                                                    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 **Instance Disks**, **Shared Block Storage**, **Instance Snapshots**, or **Security Groups** page to view resources related to this instance.

4.3 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.

4.4 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** > **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.

4.5 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** > **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 () 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 Management 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** > **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:

   - 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**.

— 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*

4.6 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**.

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

4.7 Reactivate an instance

For a Pay-As-You-Go instance, in the event of payment failure within 15 days (T+15) after the due date (T), the instance is stopped due to overdue payment and becomes **Expired**. You must open a ticket to settle the payment and reactivate the instance within 30 days (T+30) after the due date (T). Otherwise, the instance is released and the data cannot be recovered.

> **Note:**
>
> If you fail to reactivate the ECS instance within 30 days (T+30) after the due date (T), 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.

You have settled the payment by opening a ticket.

Procedure

To reactivate an instance, follow these steps:

1. Log on to the *ECS console*.

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

3. Select a region.

4. Select the instance to be reactivated, and at the bottom of the instance list, select **More** >
Reactivate.

5. Determine that you reactivate the instance immediately or later at a specified time.

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

4.8 Release an instance

To save cost, we recommend that you release a Pay-As-You-Go instance if you no longer need it.

For a Pay-As-You-Go instance, if the *No fees for stopped instances (VPC-Connected)* feature is
not enabled, charges do not stop until the instance is released.

To release a Pay-As-You-Go instance, you have two options:

- Release now: immediately releases the pay-per-order instance.
- Scheduled Release: Customizes the release of your Pay-As-You-Go instance. The instance
  cannot be released automatically in less than 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 now

To release an instance now, follow these steps:

1. Log on to the *ECS console*.

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

3. Select a region.

4. Set release:

- If you want to release an 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 **Release Now**.

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

Enable automatic release

To enable auto release, follow these steps:

1. Log on to the *ECS console*.

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

3. Select a region.

4. Set release:

- If you want to release an 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 Action: ◯ Release Now   ⦿ Timed Release

Automatic Release
Setting: 🟢⚪

*Release Date: 2018-07-01   📅

*Release Time: 23 : 59

Note:
- The release interval is 5 minutes. The system will stop charging as per the accurate release time which may differ from the specified release time.

Next   Cancel

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:
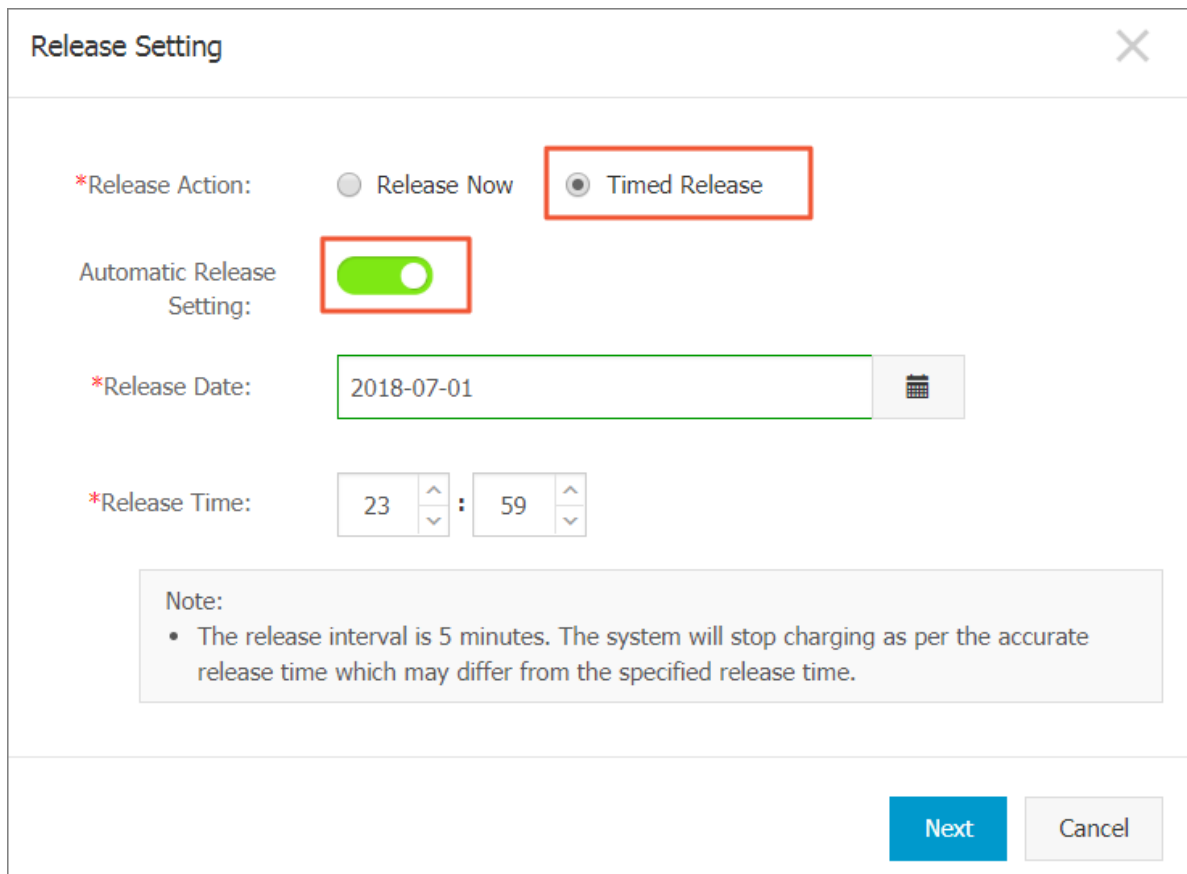
1. Log on to the *ECS console*.

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

3. Select a region.

4. Set release:

    • If you want to disable automatic release for an instance, find the instance, and 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 APIs

DeleteInstance

4.9 Add to or remove from a security group

Add an instance to a security group

You can add an instance to a security group using the console. One ECS instance can be added to up to five security groups.

1. Log on to the *ECS console*.

2. Click **Instances** in the left-side navigation pane.

3. Select a region.

4. Select an instance. Click the instance name or corresponding **Manage** button.

5. Click **Security Groups** in the left-side navigation pane.

6. Click **Add Security Group**. In the displayed dialog box, select the appropriate 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. Click **Instances** in the left-side navigation pane.

3. Select a region.

4. Select an instance. Click the instance name or corresponding **Manage** button.

5. Click **Security Groups** in the left-side navigation pane. You can view the security group list that this instance belongs to.

6. Select the security group to remove from and click **Remove**.

7. Click **OK**.

For use cases of security groups, see *scenarios*.

4.10 User-defined data and metadata

4.10.1 Metadata

The metadata of an instance is the basic information of the ECS instance including Instance ID, IP Address, Operating System, and so on. You can use the metadata to manage and/or 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 instance. ID.

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

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 instance ID.

- 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
dns-conf/nameservers	**DNS configurations for an instance.**	**2016-01-01**
eipv4	EIP address	2016-01-01
hostname	The OS name of 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 marketpalce.	2016-01-01
image/market-place/charge-type	Billing method of the image in the marketplace.	2016-01-01
instance-id	Instance ID	2016-01-01
mac	The MAC address of an instance. When multiple network interface cards exist in an	2016-01-01

Metadata name	Description	Version
dns-conf/nameservers	**DNS configurations for an instance.**	**2016-01-01**
	instance, this metadata indicates the MAC address of eth0.	
network-type	Network type, only applicable for VPC.	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	The 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 ENS 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
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]	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	2016-01-01

Metadata name	Description	Version
dns-conf/nameservers	**DNS configurations for an instance.**	**2016-01-01**
	get the STS credential, `[role-name]` must be replaced with the actual RAM role name you create or you have created. 📋 **Note:** STS The new STS credential is available 30 minutes prior to the expiration of the old one.	
instance/spot/termination-time	The spot instance release time is based on the instance operating system time zone. It is specified in the UTC time standard. 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 the multiple NIC (Network Interface Controller).	2016-01-01
network/interfaces/macs/[mac]/network-interface-id	The unique ID of the NIC, `[mac]` must be replaced with the actual MAC address.	2016-01-01
instance/virtualization-solution	The virtualization solution: 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, which your instance is hosted on.	2016-01-01
instance-identity/document	*Instance identity* Instance identity document.	2016-01-01
instance-identity/pkcs7	Instance identity signature.	2016-01-01

Other data exposing to instance via metadata

- **Maintenance**: On-going *system events*, run the following command in the instance to get data updated:

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

- **Signature**: *Instance identity* document, run the following command in the instance to get data updated:

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

- **Boot behavior configuration**: *Instance user data*, run the following command in the instance to get data updated:

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

4.10.2 User data

User data is provided by ECS for you to customize the startup behaviors of an ECS instance and to pass data into an ECS instance. You can specify instance user data when creating an instance (*RunInstances*) to customize startup behavior for your instance. For example, automatically update software packages, enable services, print logs, install dependencies, initialize Web services, and other actions that configure your instances. Instance user data 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. Others *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 followings:

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 starts to running (**Running**), we first run the instance user data with the administrator or root permission, followed by the initialization or `/etc/init` information.

After you modify the instance user data, whether the modified user data will be run again or not depends on the type of scripts and modules. For example:

- If you configure user data by using a shell script, such as a *user-data script*, we will not run the modified user data.

- If the user data configures modules like Byobu, Set Hostname, and Set Passwords, we will not run the modified user data.

- If the user data configures modules like bootcmd, update_etc_hosts, and yum_add_repo, we will run the modified user data.

  For more information, see the cloud-init documentation *Modules* and pay attention to the module frequency.

Set user data

Assume that you write user data development environment is Windows computer, and you use *Upstart Job* to configure the user data.

1. Use the editor to create a text file, such as NotePad ++.

2. Edit the script related to 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 correct.

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 the Alibaba Cloud OSS to create links. For more information, see OSS *Upload an object* or *Set lifecycle*.

6. Log on to the *ECS Management Console*.

7. See *Step 2. Create an instance* Create a Linux instance.

> **Note:**
>
> The instance must be `VPC-Connected`, and you need to select a *image* that is compliant with 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**, enter your **user data**. If your user data has been encrypted by Base64, click **The text is Base64-encoded**.

8. Waits for creating the instance.

9. After the instance is launched, see *Overview* to 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:

For example, in the `/etc/init` folder, a startup job file `part-001.conf` is generated.

Related API: *RunInstances* + Parameters `UserData`

View user data

You can view instance user data from the server `100.100.100.200` .

1. Connect to the instance.

2. In the instance, run:

- `curl http://100.100.100.200/latest/user-data` View the user data of a Linux instance:

- `Invoke-RestMethod http://100.100.100.200/latest/user-data/` View the user data of a Windows instance:

Related APIs:*DescribeUserdata*

Modify user data

You must stop the instance in advance. 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.

1. Log on to the *ECS Management Console*.

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

3. Select a region.

4. Select the target instance, and in the **Actions** column, click **Sets User Data**.

5. Enter After you fill in the information in the burst window, click **OK**.

> **Note:**
>
> After you modify the user data, whether you want to re-run the modified user data depends on the
> script type and the module type.

Related API: *ModifyInstanceAttribute* + Parameters `Userdata`

Linux instance user data

Linux instance user data can be performed 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, reference the *Metadata* for data sources. The configuration of
Linux instances are automated at boot. For more information, see Cloud-init *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 cannot exceed 16 KB before Base64 encoding. The following are examples of User-Data script:

```
#! /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, start and connect to the instance, and run `cat [file]` to view the results of the user-data script.

```
[root@XXXXX2z ~]# cat output.txt
Hello World. The time is now Mon, 24 Jul 2017 13:03:19 +0800!
```

Cloud config

Cloud Config is the easiest way to implement instance customization data, and its interaction is very friendly. You can use cloud Config to configure services such as updating yum sources, importing SSH keys, installing dependency packages, and so on. The first line of Cloud Config is fixed as `#cloud-config`, and the header cannot have spaces. The file must be yaml syntax valid. 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. See the following 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, start and connect to the instance to view the results.

Include files

The contents of the 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, once 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 Vase64 encoding before being passed in, and the pre-encoding include file cannot exceed 16 KB. See the following include file for example:

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

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

Gzip compressed content

The content of *User-Data Script*, *Cloud Config*, and *Include File* cannot exceed 16 KB. If your script content is more than 16 KB, you can use the Gzip compressed content. Upload the compressed script in available storage service and obtain the link and use the Include file format to render the link. The first line of Gzip compression 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. See the following Gzip compressed content for example:

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

Upstart Job

Upstart service for your init system is required 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. See the following 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 a proprietary utility developed by ECS. We provide Windows instance with 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, start and connect the instance to view the results, A `1.txt`

text file is shown under the `C:\` drive.
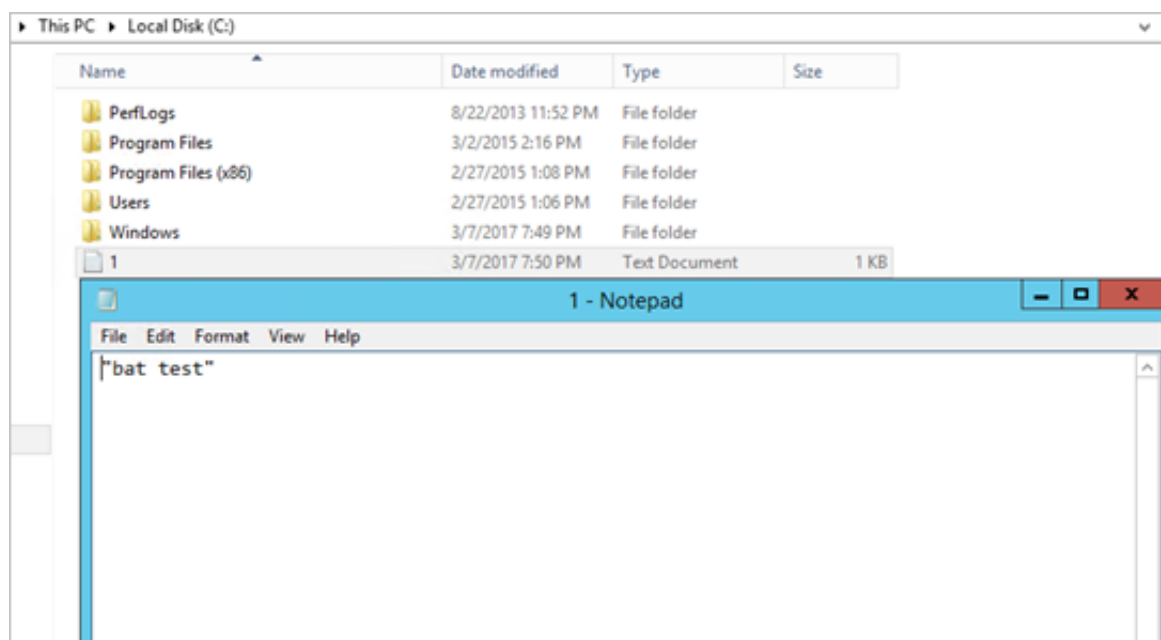
The first line of PowerShell scripts

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.

4.10.3 Instance identity

The instance identity is a part of the *Metadata* that describes and validates an instance. The instance identity enables you fast recognize an instance,

and provides authentication for such as charged software updates, access control, or application activation. The signature of instance identity is encrypted by the *PKCS#7* , which is secure and reliable.

Use cases

You may need the aid of **instance identity** (`instance-identity`) in the following scenarios such as authentication, access grant or instance identifying.

- The 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 *marketplace* is flexible and sometimes free of sales consultant. You can use the instance identity to complete the software activation. For more information, see the *Sample 1. No audience in the signature*.

- When you write sensitive data in the instance, you can use instance identity to make sure whether the server is your instance or not.

- Other scenarios that you want to confirm the source of the target server.

Feature details

The instance identity consists of the dynamically generated **instance identity document** (`document`) and **instance identity signature** (`signature`).

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

Properties	Description	One and only?
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

Properties	Description	One and only?
serial-number	Serial number of the instance	No
zone-id	ID of the zone to which the instance belongs	No
instance-type	Instance types	It changes after you change the instance type.
image-id	Image ID of the instance	It changes after you replace the system disk of the instance.
private-ip	Private IP of the instance	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 PKCS#7, which is digitalized and reliable.

  ── To enhance the security of signature, you can protect it by specifying the `audience` parameter in it. After the `audience` even if someone else gets some information about the identity document and the identity signature, there is a very small probability that your `audience` parameter can be easily acquired and illegally used. The value of the `audience` parameter can be a random string, timestamp, regularly changed data, or output generated by a specific algorithm.

  ── However, 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.

3. 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.

4. Verify the instance identity by using the 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-----
MIIDdzCCAl+gAwIBAgIEZmbRhzANBgkqhkiG9w0BAQsFADBsMRAwDgYDVQQGEwdV
bmtub3duMRAwDgYDVQQIEwdVbmtub3duMRAwDgYDVQQHEwdVbmtub3duMRAwDgYD
VQQKEwdVbmtub3duMRAwDgYDVQQLEwdVbmtub3duMRAwDgYDVQQDEwdVbmtub3du
MB4XDTE4MDIyMzAxMjkzOFoXDTM4MDIxODAxMjkzOFowbDEQMA4GA1UEBhMHVW5r
bm93bjEQMA4GA1UECBMHVW5rbm93bjEQMA4GA1UEBxMHVW5rbm93bjEQMA4GA1UE
ChMHVW5rbm93bjEQMA4GA1UECxMHVW5rbm93bjEQMA4GA1UEAxMHVW5rbm93bjCC
ASIwDQYJKoZIhvcNAQEBBQADggEPADCCAQoCggEBAIJwy5sbZDiNyX4mvdP32pqM
YMK4k7+5lRnVR2Fky/5uwyGSPbddNXaXzwEm+u4wIsJiaAN3OZgJpYIoCGik+9lG
5gVAIr0+/3rZ61IbeVE+vDenDd8g/m/YIdYBfC2IbzgS9EVGAf/gJdtDODXrDfQj
Fk2rQsvpftVOUs3Vpl9O+jeCQLoRbZYm0c5v7jP/L2lK0MjhiywPF2kpDeisMtnD
/ArkSPIlg1qVYm3F19v3pa6ZioM2hnwXg5DibYlgVvsIBGhvYqdQ1KosNVcVGGQa
HCUuVGdS7vHJYp3byH0vQYYygzxUJT2TqvK7pD57eYMN5drc7e19oyRQvbPQ3kkC
```

```
AwEAAaMhMB8wHQYDVR0OBBYEFAwwrnHlRgFvPGo+UD5zS1xAkC91MA0GCSqGSIb3
DQEBCwUAA4IBAQBBLhDRgezd/OOppuYEVNB9+XiJ9dNmcuHUhjNTnjiKQWVk/YDA
v+T2V3t9yl8L8o61tRIVKQ++lDhjlVmur/mbBN25/UNRpJllfpUH6oOaqvQAze4a
nRgyTnBwVBZkdJ0d1sivL9NZ4pKelJF3Ylw6rp0YMqV+cwkt/vRtzRJ31ZEeBhs7
vKh7F6BiGCHL5ZAwEUYe8O3akQwjgrMUcfuiFs4/sAeDMnmgN6Uq8DFEBXDpAxVN
sV/6Hockdfinx85RV2AUwJGfClcVcu4hMhOvKROpcH27xu9bBIeMuY0vvzP2VyOm
DoJeqU7qZjyCaUBkPimsz/1eRod6d4P5qxTj
-----END CERTIFICATE-----
```

Sample 1. No audience in the signature

Assuming that you have published an image in the image market, this sample shows you how to grant access to the instances of your customers.

1. Enter into the target instance after the instance is started.

2. Verify whether the image used by the instance is from the *image market* or not 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

Similarly, assuming that you published an image in the image market, this sample shows you how to grant access to the instances of your customers by specifying an **audience** parameter during the process of validation. To avoid that the instance identity is maliciously acquired and

distorted, you can implement the access control at the application server by combining your audience parameter. The value of the `audience` parameter can be a random string, timestamp, regularly changed data, or output generated by a specific algorithm.

1. Enter into the target instance after the instance is started.

2. Verify whether the image used by the instance is from the *image market* or not 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.

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, you can choose an ideal method to connect to an ECS instance.

Connect to a Linux instance

Choose an ideal method from the following table to create remote connection to your Linux instance.

Internet access	Operating system of the local machine	Connection option
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: • Use an SSH key pair as the credential: *Connect to a Linux instance by using an SSH key pair* • Use a password as the credential: *Connect to a Linux instance by using a password*
Yes	Linux, Mac OS, or other Unix-like OS	Use commands to create remote connection: • Use an SSH key pair as the credential: *Connect to a Linux instance by using an SSH key pair* • Use a password as the credential: *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:

Internet access	Operating system of the local machine	Connection option
		Connect to an instance on a mobile device

Connect to a Windows instance

Choose an ideal method from the following table to create remote connection to your Windows instance.

Internet access	Operating system of the local machine	Connection option
Yes/No	Windows or Unix-like OS	*Connect to an instance by using the Management Terminal*
Yes	Windows	Use mstsc to create remote connection: *Connect to a Windows instance*
Yes	Linux	Use a remote connection tool, such as rdesktop, to create remote connection: *Connect to a Windows instance*
Yes	Mac OS	Use Microsoft Remote Desktop Connection for Mac to create remote connection: *Connect to a Windows instance*
Yes	iOS or Android	Use Microsoft Remote Desktop to create a remote connection: *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.

Specifically, when the remote access software programs that you are using, such as PuTTy, Xshell, or SecureCRT, do not work.

Scenarios

The Management Terminal can be used to:

- Check the status of an ECS instance if it starts slowly.

- Reconfigure the firewall if a remote connection fails because of any 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, use the *Reset Password* feature to set a password.

Procedure

The following figure illustrates how to use the Management Terminal to connect to an ECS instance.
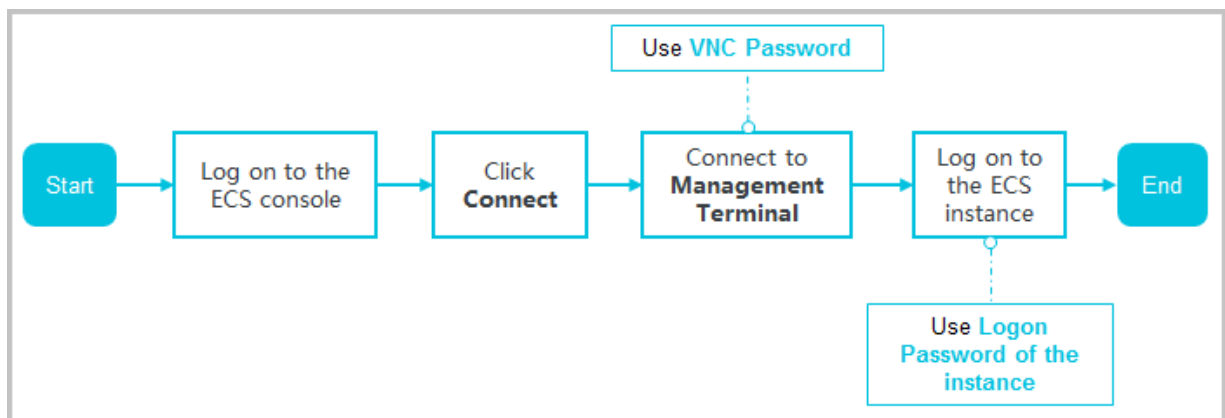
To connect to the ECS instance by using the Management Terminal, follow these steps:

1. Log on to the *ECS console*.

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

3. Select a region.

4. In the instance list, find your instance, and 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 and secure password immediately 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 in case you have forgotten your VNC connection password, follow these steps to connect to the Management Terminal:

  — *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 forget the logon password of your instance, *reset the password*.
  > - The logon password input is invisible.
  > - If you want to do different 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.

- In case you see a black screen, the Linux instance may be in sleep mode. To exit sleep mode, click the 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 forget the logon password of your instance, *reset the password*.

Other Operations

Change the VNC connection password

If you forget the 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 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 the VNC **Connection Password** dialog box or the **Enter VNC Password** dialog box.

3. In the upper-right corner of the **Management Terminal** page, click **Modify Management Terminal Password**.

4. Enter a new password, which must be six characters in length and may contain uppercase letters, lowercase letters, and digits. Special characters are not supported.

5. A new password can be effective in the following events:

   - 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.

Follow these steps:

1. Open the **Management Terminal** page.

2. In the upper-right corner of the **Management Terminal** page, click **Input Commands**.

3. In the **Copy Commands** dialog box, 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 mouse or press any key to activate the screen.

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

- Why am I unable to access the Management Terminal?

  To resolve logon issues, open your browser and connect to the Management Terminal. Press **F12** to open the developer tool. The Management Terminal information can be analyzed to locate errors under the Console tab.

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

  You can access the Management Terminal only if you have IE10 or later versions installed . Only certain versions of Firefox are supported. You can resolve this issue by updating or changing your browser to a recommended version.

  > **Note:**
  >
  > We recommend that you use Google Chrome because it offers the best support for the Management Terminal function.

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

How to use a key pair to log on to a Linux instance depends on the local operating system.

- *Windows OS*

- *Linux OS or other systems supporting SSH commands*

> **Note:**
>
> You can use a password to connect to a Linux instance. For more information, see *Connect to a Linux instance by using a password* and *Connect to an instance by using the Management Terminal*.

Windows OS

In this section, it is demonstrated how to use a key pair to log on to a Linux instance on a Windows system, using the popular SSH tools PuTTY and PuTTYgen as an example.

Prerequisites

- PuTTY and PuTTYgen must have been installed. You can download them at:

  — *PuTTY*

  — *PuTTYgen*

- You must have a Linux instance that has been bound to an instance. You can allocate an SSH key pair when creating an instance or *bind an SSH key pair to an instance*.

- Add the following rule in the security group to enable the access to the TCP Port 22 of the instance. 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

1. Optional. If you are using a key pair generated by Alibaba Cloud, of which the private key is a .pem file, you must convert it to a .ppk file. If your private key is a .ppk file, you can skip this step.

> **Note:**
>
> When you *create an SSH key pair*, download the .pem private key.

a. Start PuTTYgen. In this example, we use PuTTYgen version 0.68.

b. Under **the** > **Type of key to generate option,** select `RSA`.

> **Note:**
>
> **The value of Number of bits in a generated key** can be left as is. The software
>
> automatically update the value based on the imported private key information.

c. Click **Load** to find your `.pem` file.

> **Note:**
>
> By default, PuTTYgen only displays files with an extension of `.ppk`.

d. Select the downloaded private key file from Alibaba Cloud, or the ready private key file, and click **Open**.

e. Click **OK** to close the confirmation dialog box.

f. Click **Save private key**. PuTTYgen displays warning about saving the key without a password. Click **Yes**.

g. Specify the same name for the private key with the key pair, and save the settings. PuTTY automatically adds the .ppk file.

2. Start PuTTY.

3. Select **Connection >** > **SSH >** > **Auth.**Click **Browse…** and select the .ppk file generated in Step 1.
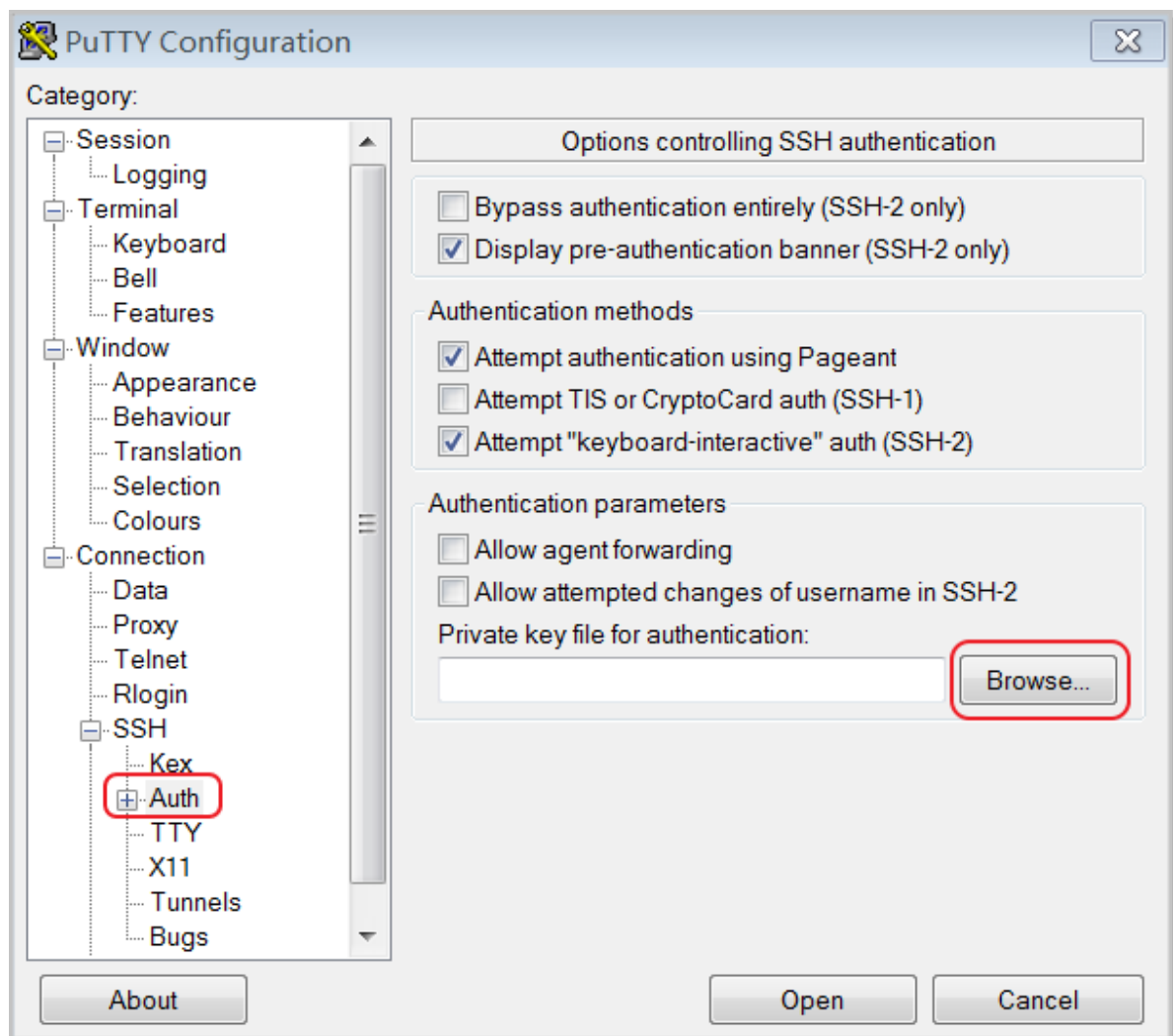
4. Click **Session**.

   - In `Host Name (or IP address)`, enter your account and the public IP address of the instance to be connected to. The format is `root@IP address`.
   - In `Port` enter the port number `22`.
   - For `Connection type`, select `SSH`.
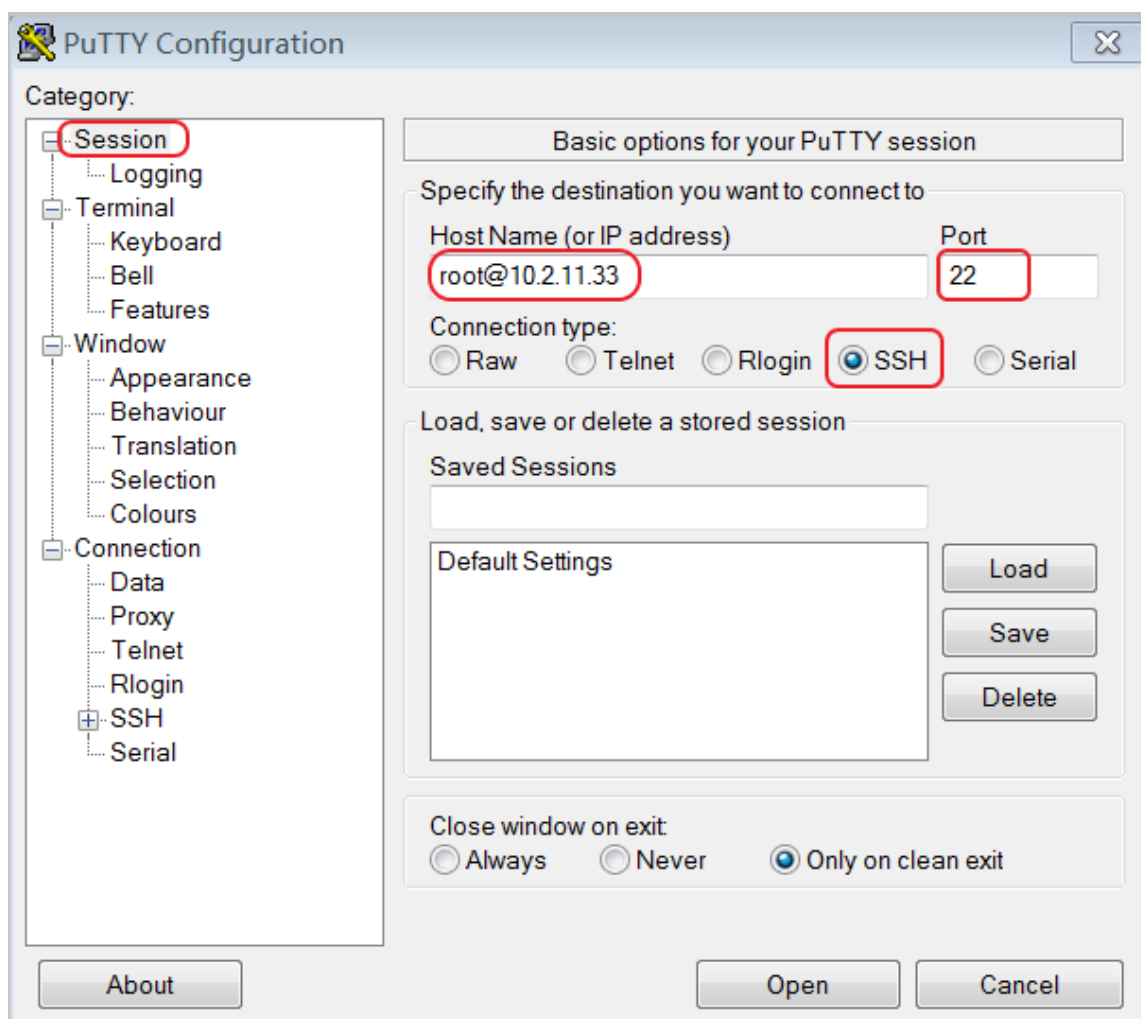
5. Click **Open** to start accessing your Linux instance.

When the window shows `Connection established.`, it indicates you have successfully logged on to the instance using the key pair.

Linux OS or other systems supporting SSH commands

In this section, it is demonstrated how to use a key pair to log on to a Linux instance on a Linux system or a system supporting SSH commands, such as MobaXterm for Windows.

Prerequisites

You must have a Linux instance that has been bound to an SSH key pair. You can *allocate an SSH key pair when creating an instance* , or *bind an SSH key pair to an instance*.

Add the following rule in the security group to enable the access to the TCP Port 22 of the instance. For more information, see *Add security group rules*.

Network Type	NIC	Rule Direction	Authoriza ion Policy	Protocol Type	Port Range	Authoriza ion Type	Authoriza ion Object	Priority
VPC	N/A	Inbound	Allow	SSH(22)	22/22	Address Field Access	0.0.0.0/0	1
Classic	Internet							

Procedure

1. Locate directory of your private key, for example, `/root/xxx.pem`.

   > **Note:**
   >
   > When you *create an SSH key pair*, download the .pem private key. xxx.pem is the private key file.

2. To modify the attributes of the private key, run the command: `chmod 400 [directory of the private key file]`. For example, `chmod 400 /root/xxx.pem`.

3. To connect to the instance, run the command `ssh -i directory of the private key file] root@Internet IP address`. For example, `ssh -i /root/xxx.pem root@10.10.10.100`.

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

Before you begin, make sure the following:

- 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, *reset the password*.
- The instance can access 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 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	Authoriza ion policy	Protocol type	Port range	Authoriza ion type	Authoriza ion 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, you have various options 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**.
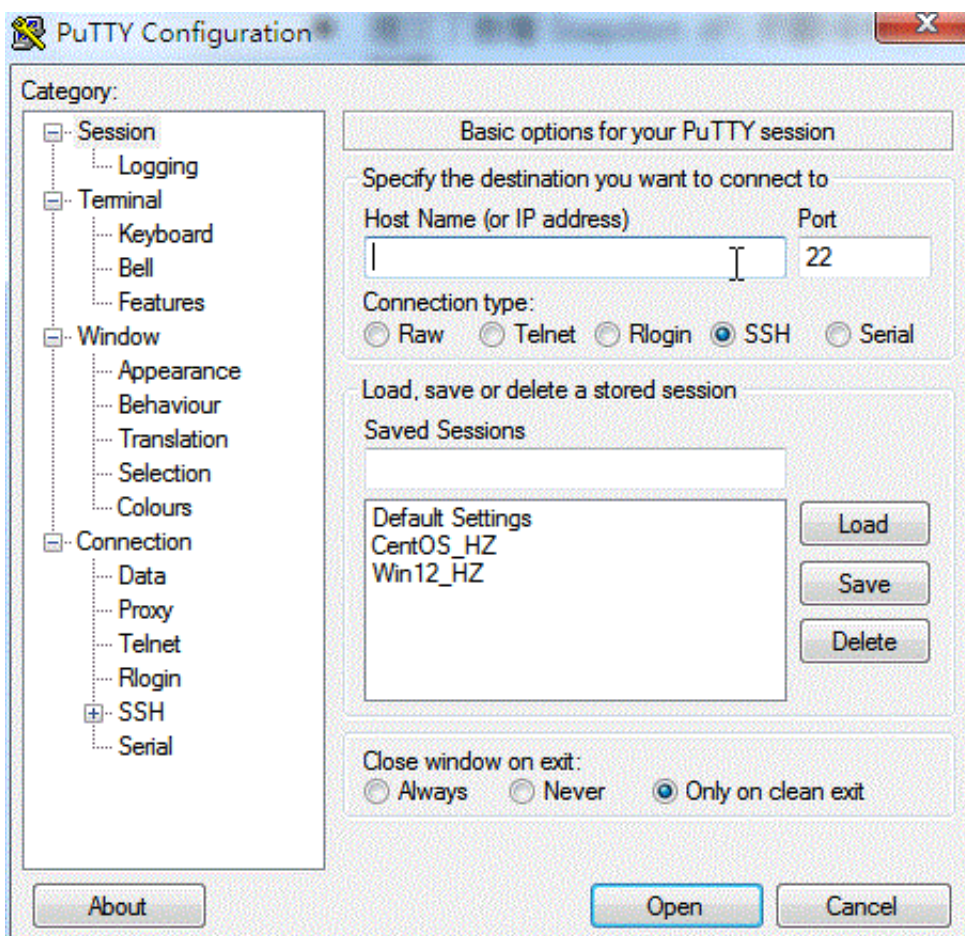
3. Click **Open** to connect, and in the PuTTY Security Alert dialog box, click Yes.

> 📋 **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 may
> occur. For more information, see *PuTTY User Manual*.
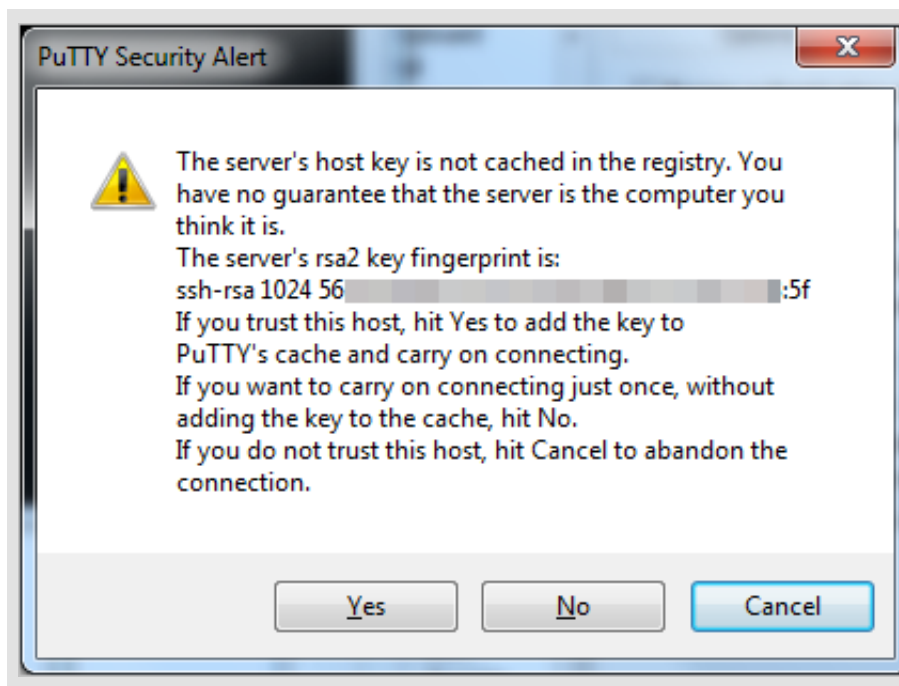
4. As prompted, type the username and password for the Linux instance, and press the Enter key.

> 📋 **Note:**
>
> The password is not displayed on screen.

When you see the following message, you have successfully connected to an instance.

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

Now, you can start working on your instance.

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 press the Enter key.

When you see the following message, you have successfully connected to an instance.

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

Now, you can start working on your instance.

Android or iOS

If your local machine is running Android OS or iOS, you can use various apps to connect to a Linux instance. For more information, see *Connect to an instance on a mobile device*.

Reference

5.5 Connect to a Windows instance

If your Windows instance can access Internet, you can use remote connection tools to connect to it. Otherwise, you can use the *Management Terminal*.

Prerequisites

Before you start, complete the following:

- 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, *reset the password*.
- The instance can access 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 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	Authoriz ation Policy	Protocol Type	Port Range	Authoriza ion Type	Authoriza ion 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, you have various options 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 shortcut key **Windows Logo** + **R** to open the **Run** windows, type `mstsc`, and then press the Enter key.

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

   a. Click the **Show Options** drop-down box.
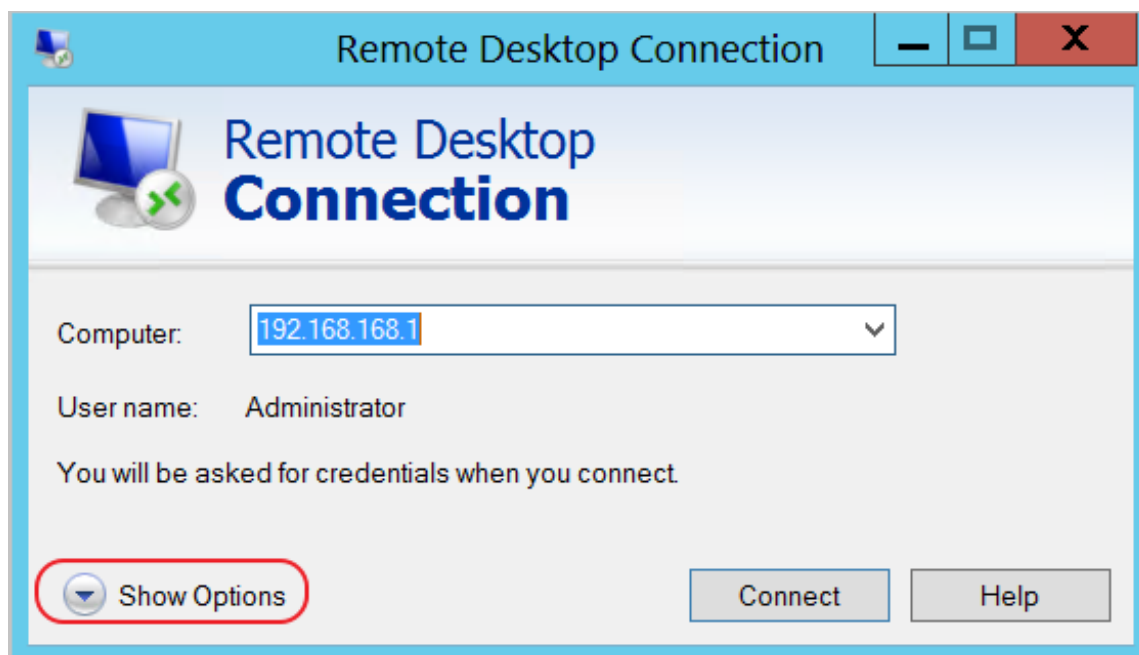
   b. Type the public IP address or EIP address of the instance.

   c. Type the user name. The default user name is `Administrator`

   > 📋 **Note:**
   >
   > If you want to log on to the instance next time without repeating these steps, select **Allow me to save credentials**.
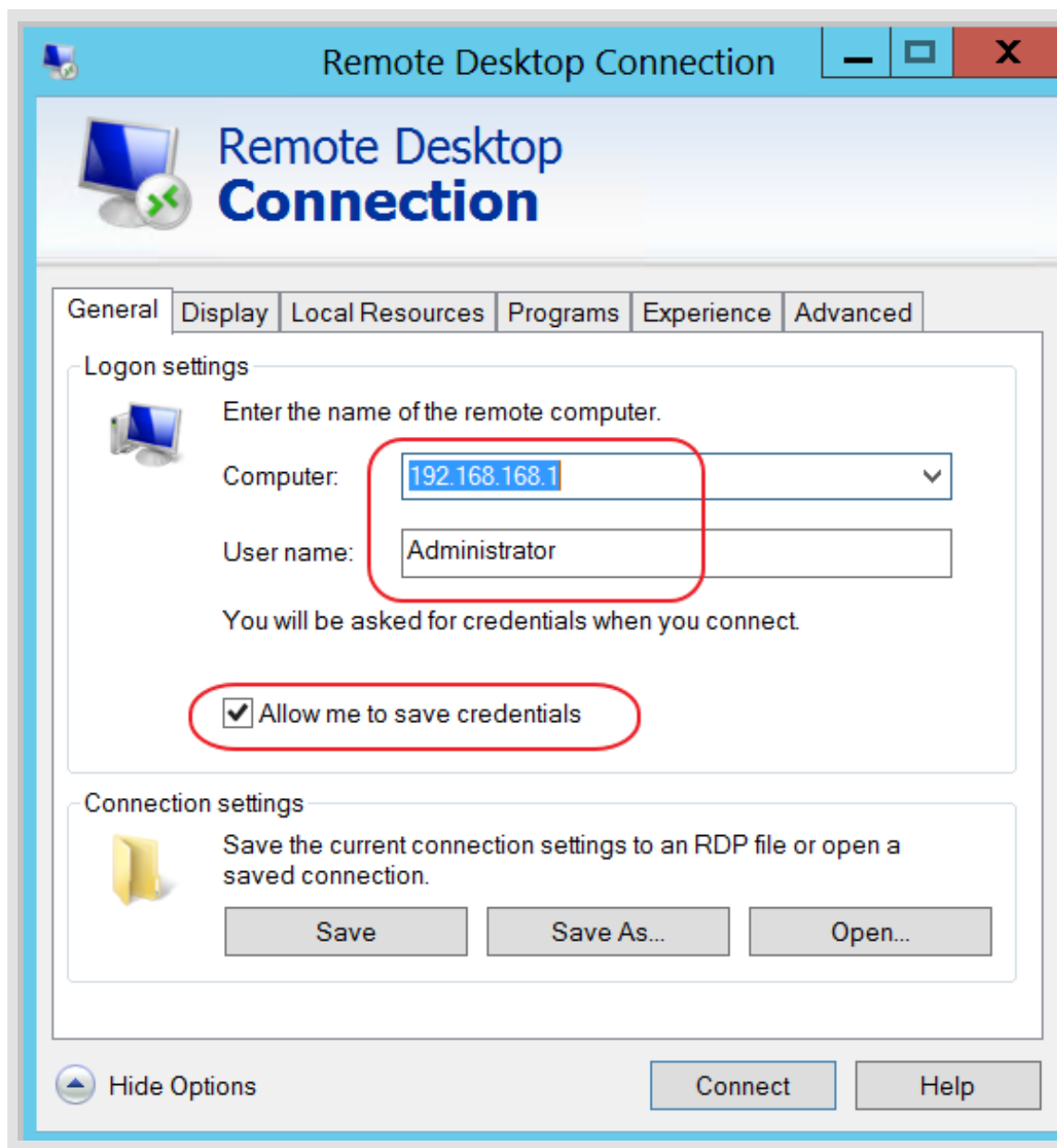
d. Optional. If you want to copy text or files from the local machine to the instance, click the **Local Resources** tab to see options for sharing local computer resources.

- If you want to copy text only, select **Clipboard**.

- If you also want to copy files, select **More** and select drive letters from which you want to copy files to your instance and click OK.
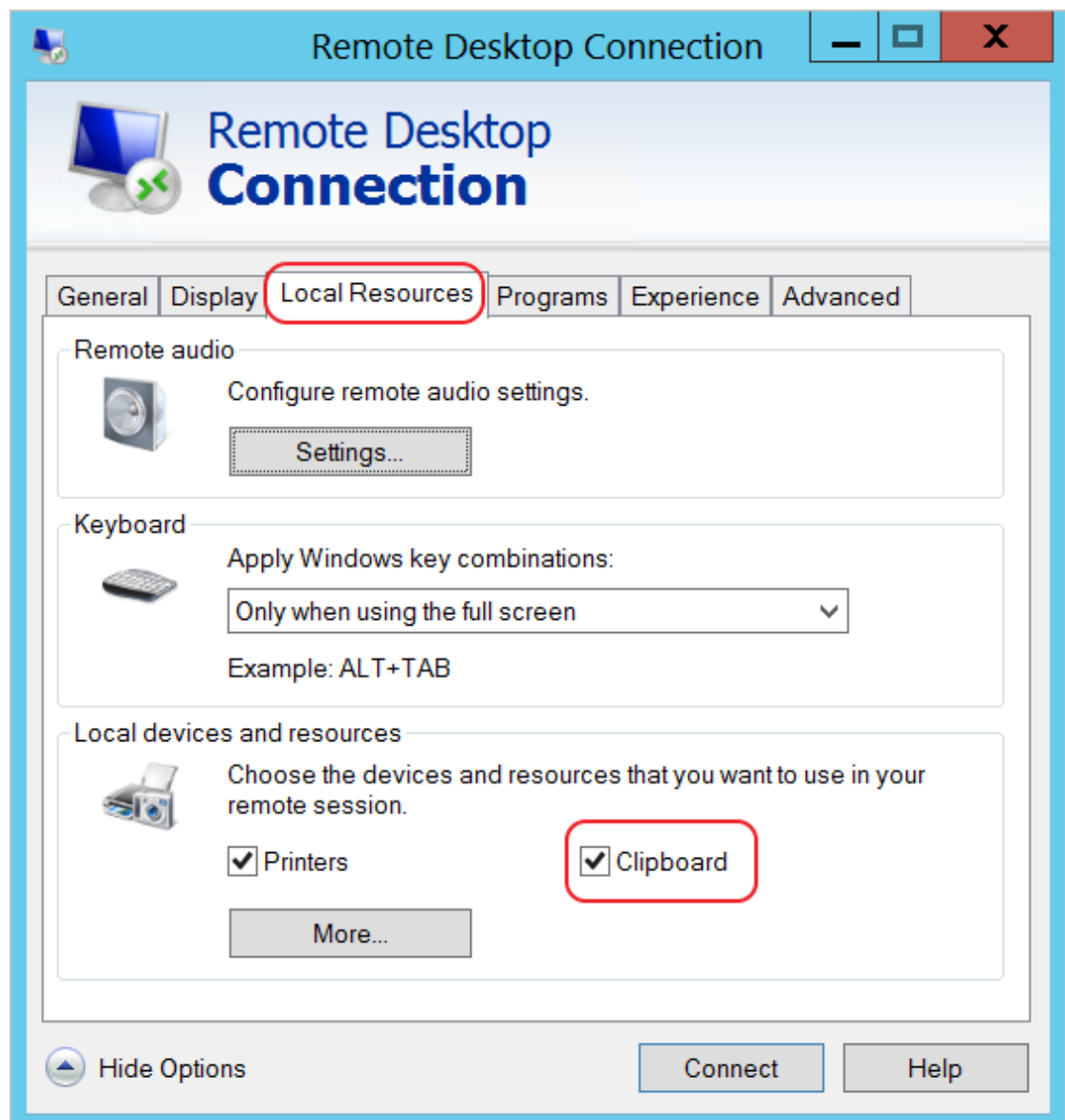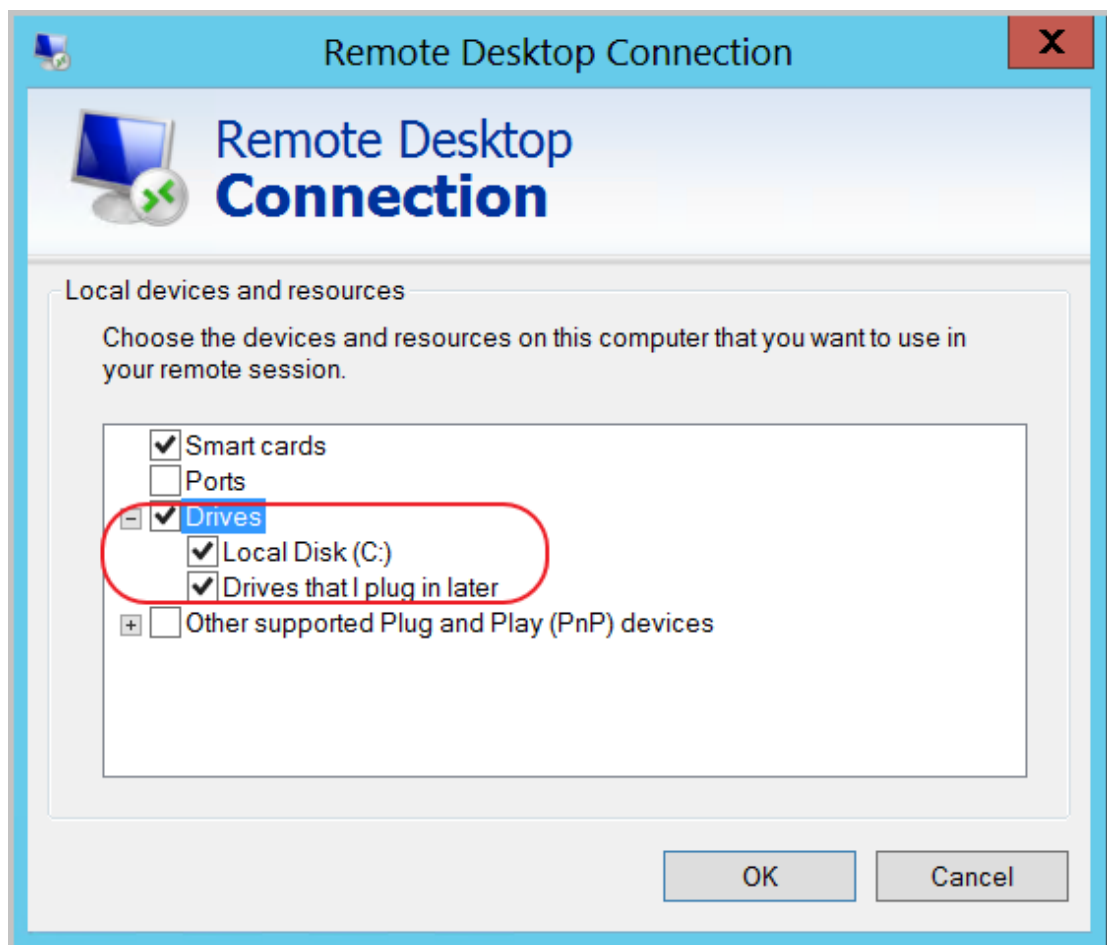
e. Optional. Click the **Display** tab and resize the remote desktop window. Full Screen is

recommended.
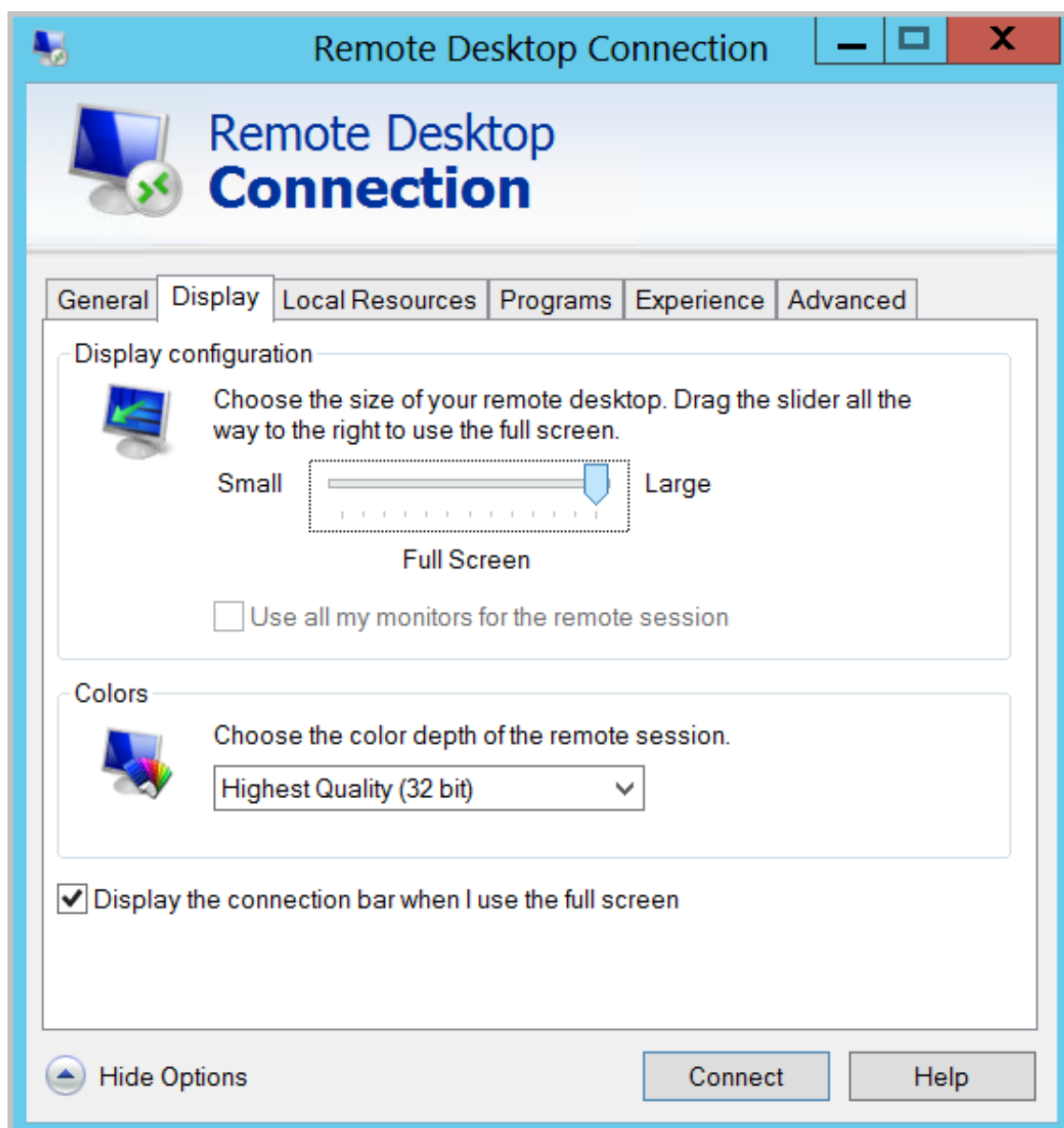
f. Click **Connect**.

Now, you can operate on the instance.

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 parameter descriptions are as follows.

Parameters	Description
-u	The user name. The default user name for 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 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 `-d inc`.
-r	Multimedia reorientation. For example: • Turn on the sound:`--r sound`. • Use a local sound card:-r sound: `-r sound : local`. • Open the U Disk: `-r disk:usb=/mnt/ usbdevice`.
-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 Linux system of a local device maps to a hard disk on a Windows instance. In this way, you can no longer rely on Samba or FTP to transfer files.

For more information about all the 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* in the Microsoft website.

Android or iOS

If your local machine is running Android OS or iOS, you can use various apps to connect to a Linux instance. For more information, see *Connect to an instance on a mobile device*.

5.6 Connect to an instance on a mobile device

This documentation describes how to connect to an ECS instance on a mobile device. The procedure varies with the operating system of your instance.

- *Connect to a Linux instance*: We take SSH Control Lite as an example 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*: We take Microsoft Remote Desktop as an example to describe how to connect to a Windows instance on an iOS or Android device.

Connect to a Linux instance

Prerequisites

Confirm the following before connecting to your instance:

- The instance is **Running** .
- The instance has a public IP address and is accessible from public network.
- 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 *the following security group rules*:

Network type	NIC	Rule direction	Authoriza ion policy	Protocol type	Port range	Authoriza ion type	Authoriza ion object	Priority
VPC	No configurat ion required	Inbound	Allow	SSH(22)	22/22	Address Field Access	0.0.0.0/0	1

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

- You have downloaded and installed the appropriate app:

  — The iOS device has SSH Control Lite installed.

 — The Android device has JuiceSSH installed.

Procedure

For iOS devices, see *Use SSH Control Lite to connect to a Linux instance*. In this example, user name and password are used for authentication.

For Android devices, see *Use JuiceSSH to connect to a Linux instance*. In this example, user name and password are used for the 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.
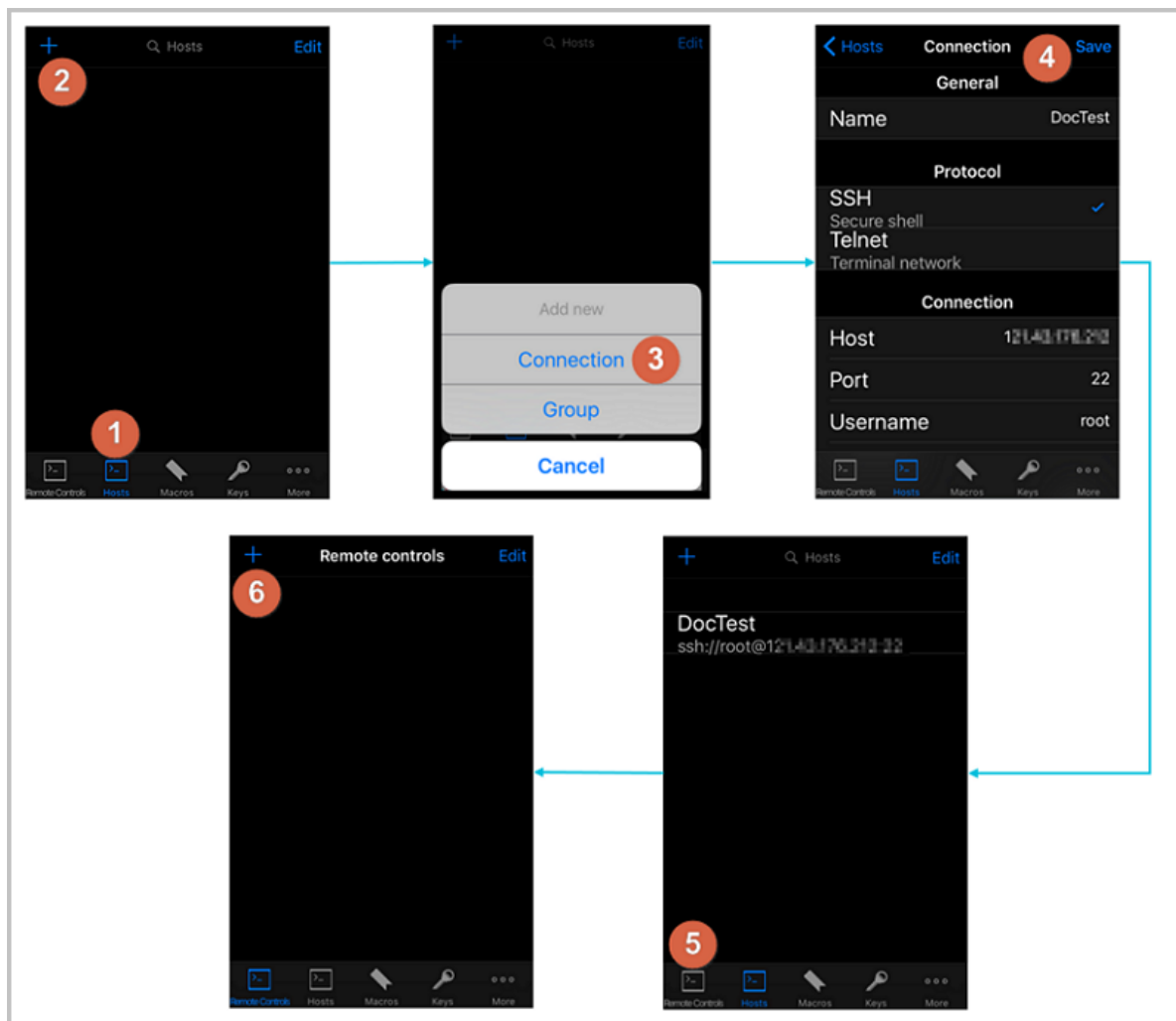
7. On the **New remote** page, tap **Host1**.

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**.
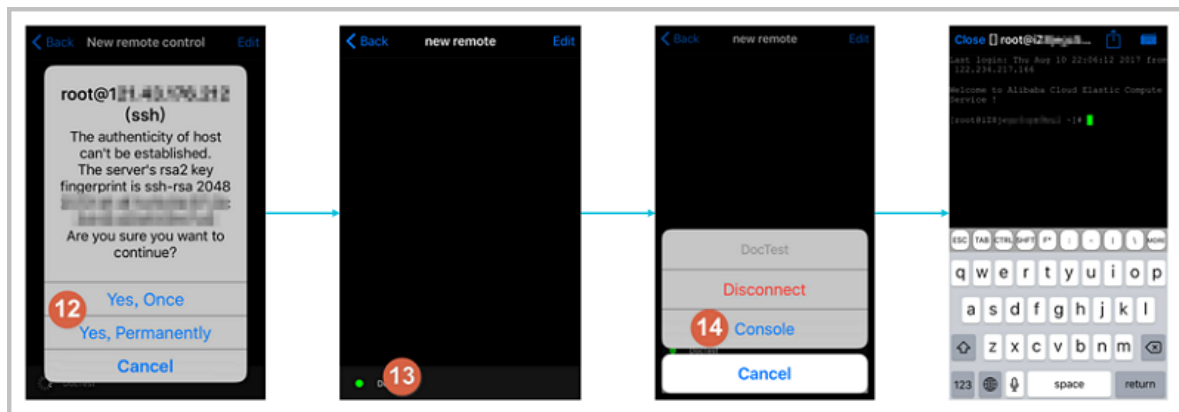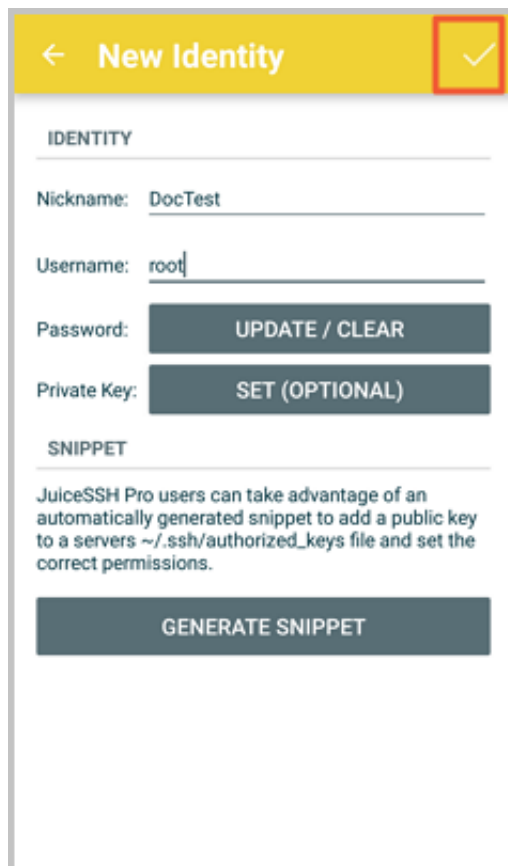
    The following figure shows Steps 7 through 11.
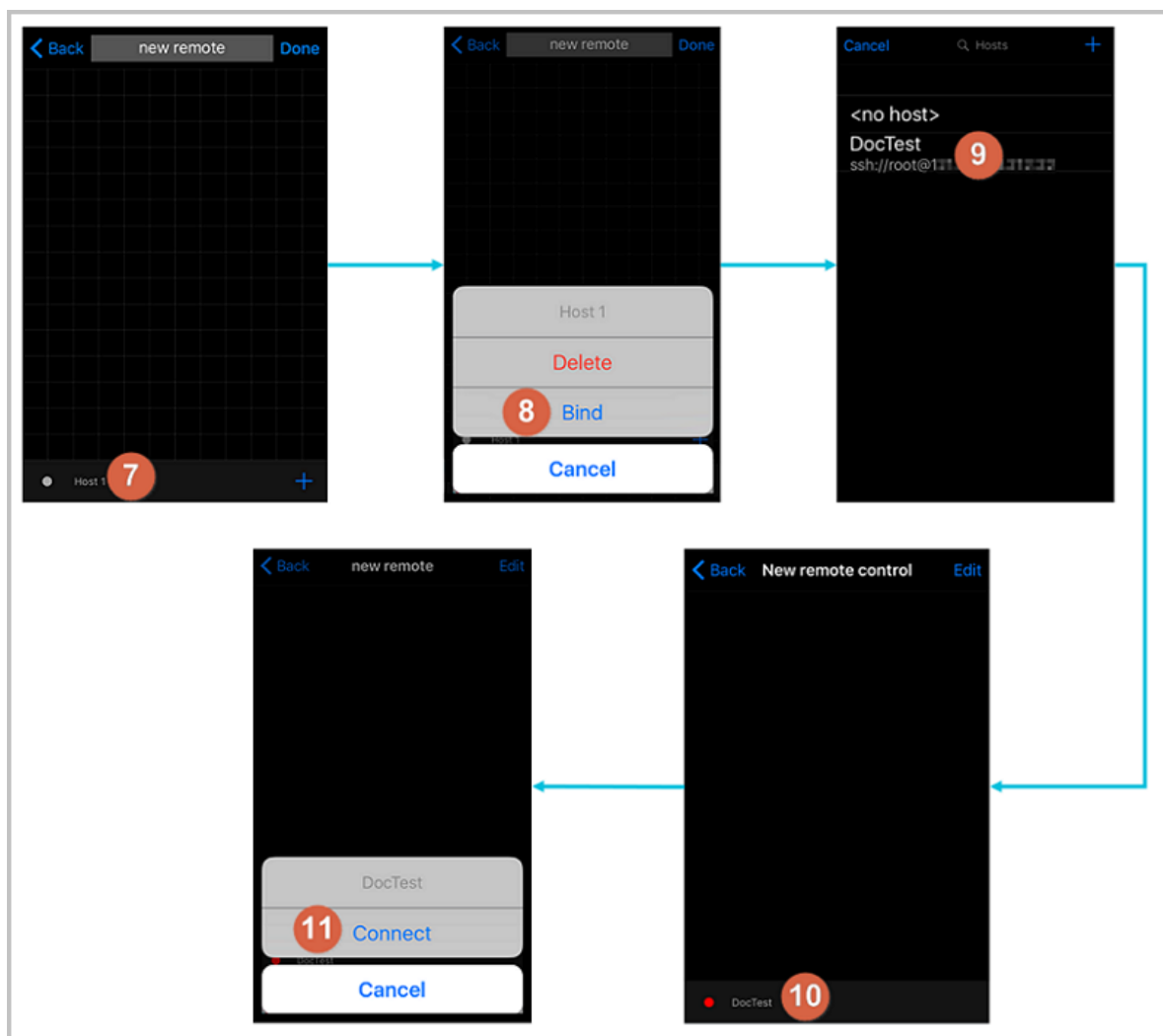
12. In the action sheet, select **Yes, Once** or **Yes, Permanently**. Once the connection is successful, the indicator in front of **DocTest** turns green.

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:

Now, you are connected to the Linux instance.

Use JuiceSSH to connect to a Linux instance

1. Start JuiceSSH, and tap **Connections**.
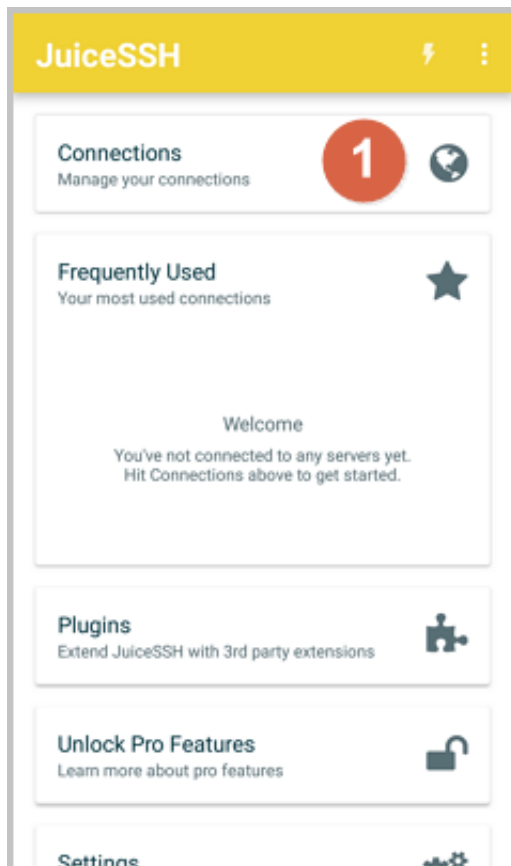
2. Under the **Connections** tab, tap the **+** icon.

3. On the **New Connection** page, add the connection information and 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`I Identity`, follow these steps:

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

        2. On the **New Identity** page, add the following information and tap the [✓] icon. The

            following connection information is required:

            -   `Nickname`: Optional. You may set a nickname to ease management. `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.

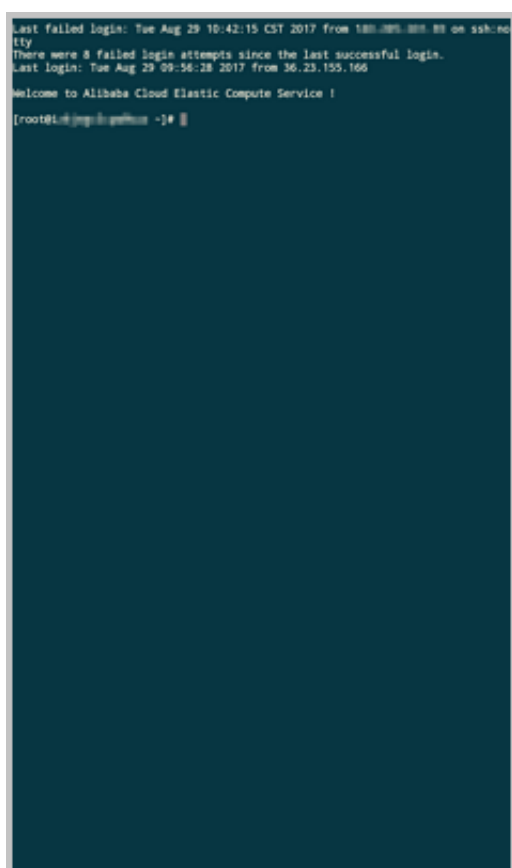- **Port**: Type the port number for SSH protocol. In this example, 22 is used.

4. Confirm the message, and tap **ACCEPT**.
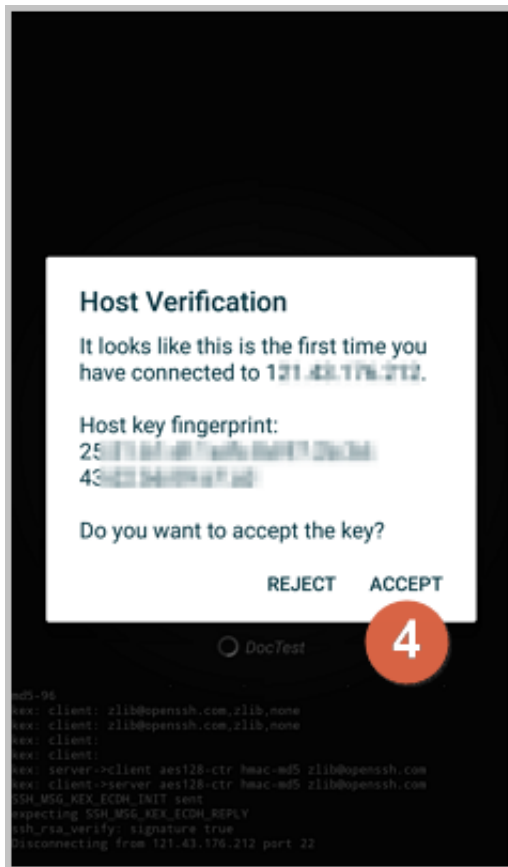
5. (Optional) For the first connection, the app would offer you some tips about font setting and the like. Confirm the message, and tap **OK - I'VE GOT IT!**.

Now, you are connected to the Linux instance.

Connect to Windows instances

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

Prerequisites

Confirm the following before connecting to your instance:

- The instance is **Running**.

- The instance has a public IP address and is accessible from public network.

- 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 configurat	Inbound	Allow	RDP(3389)	3389/ 3389	Address field access	0.0.0.0/0	1

Network type	NIC	Rule direction	Authoriza ion policy	Protocol type	Port range	Authoriza ion type	Authoriza ion object	Priority
	ion required							
Classic	Internet							

- You have downloaded and installed Microsoft Remote Desktop.

  — For iOS devices, download the app from iTunes.

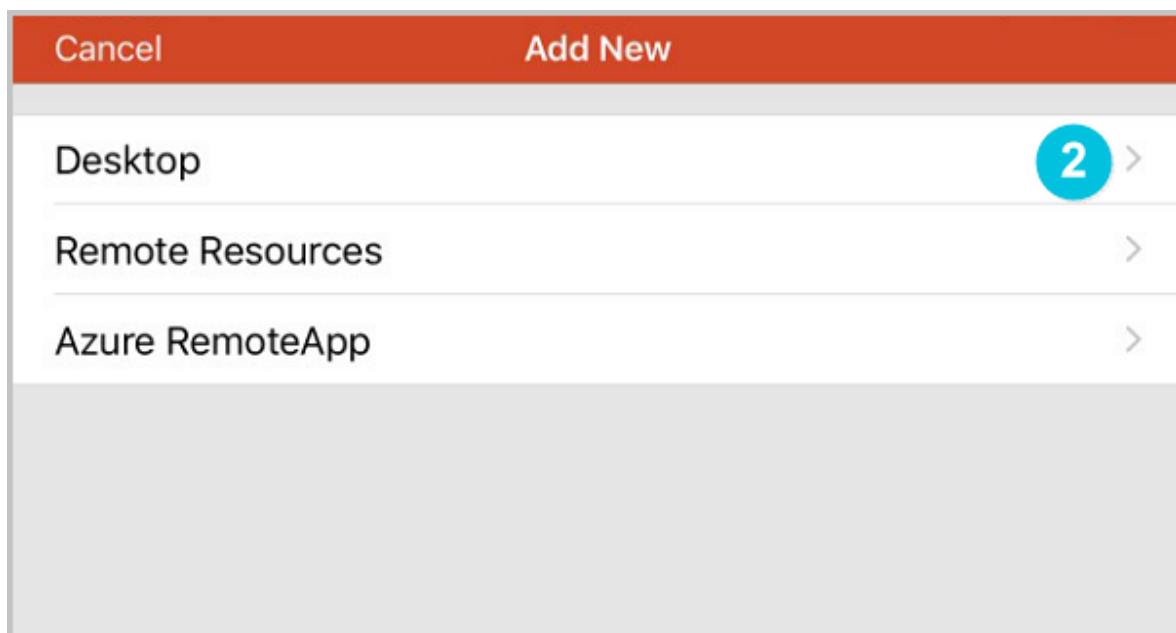 — For Android devices, download the app from Google Play.

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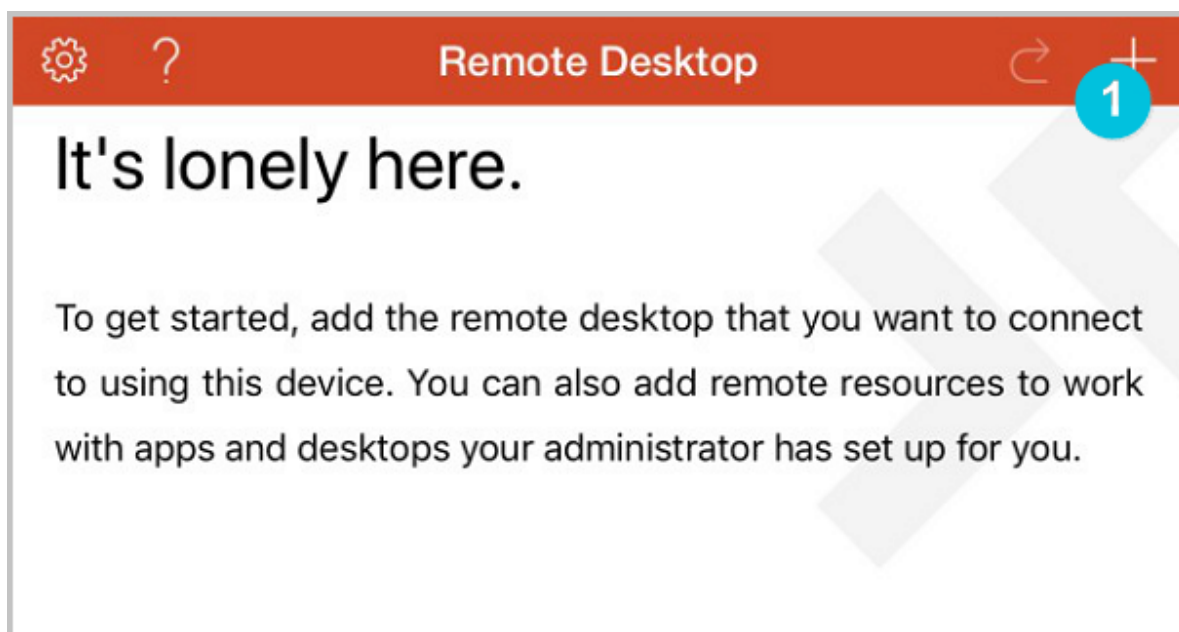
**2.** On the **Add New** page, select **Desktop**.

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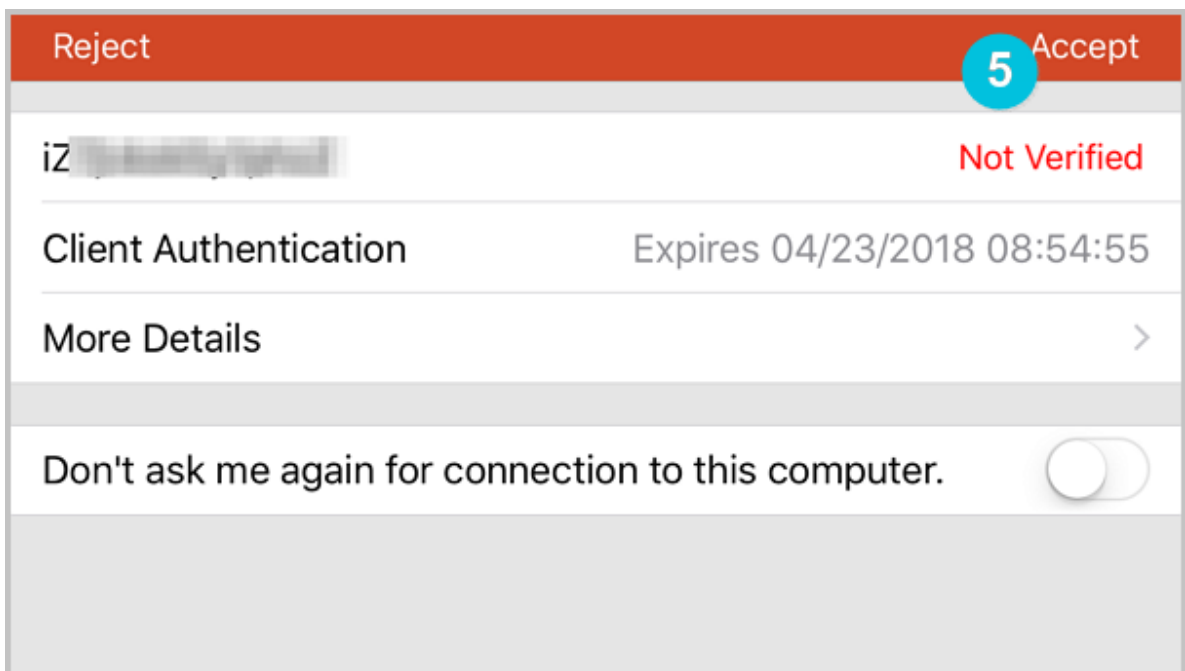
   connection information is required:

   • **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
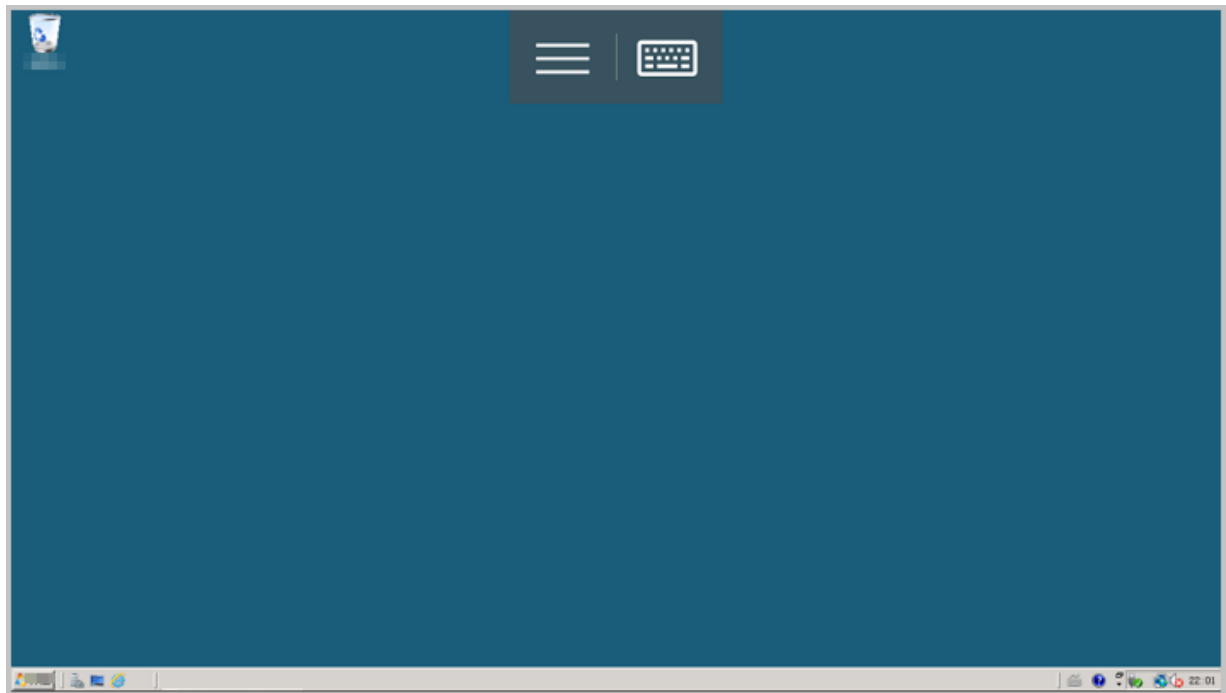
     Windows instance.
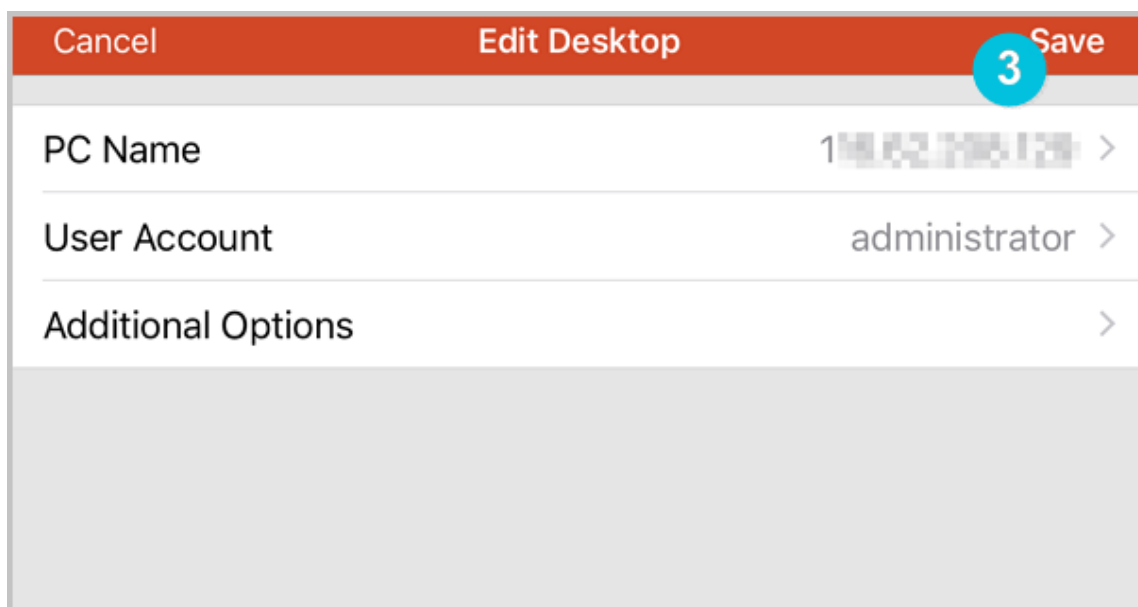
4. On the **Remote Desktop** page, tap the icon of a Windows instance.
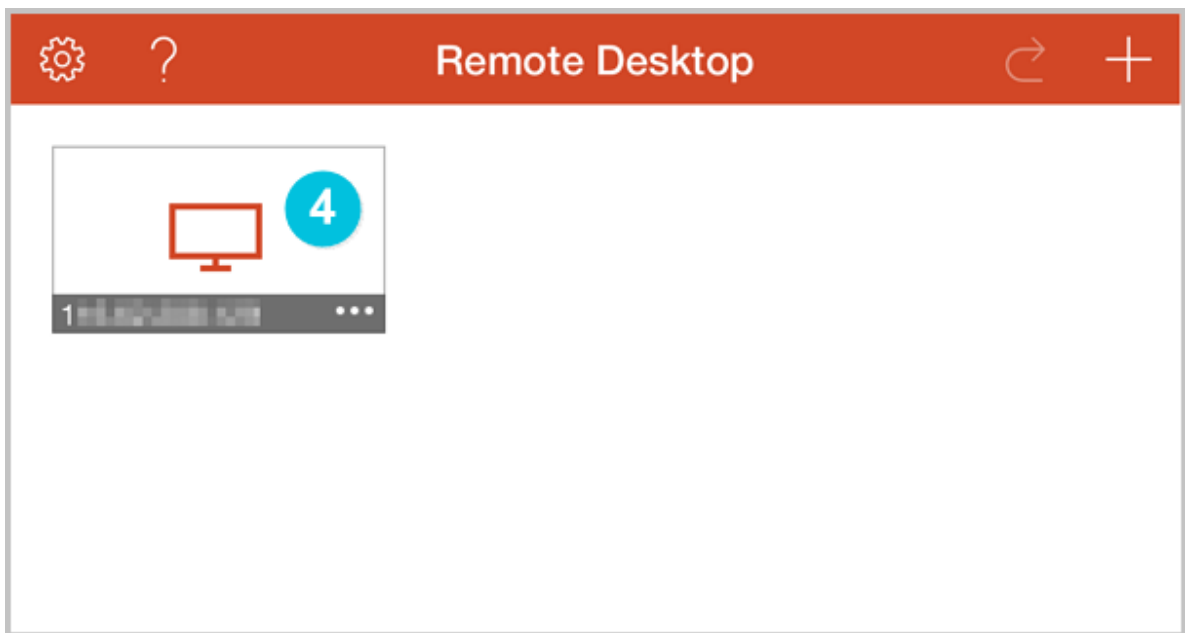
5. On the confirmation page, confirm the message and tap **Accept**.

Now, you are connected to the Windows instance.

6 Cloud disks

6.1 Create a cloud disk

You can create a cloud disk to work as a data disk to expand the storage space 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 *Limitations*.

- 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 and size 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 roll back 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 billing* 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

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** > **Cloud Disks**.

3. In the upper-right corner of the **Disk List** page, click **Create Cloud Disk** to go to the **Create** page.

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. Select a cloud disk category and specify the disk size and the quantity. You can also choose *Create a cloud disk from a snapshot*.

6. Confirm the configuration and the **cost**.

7. Click **Buy Now**, confirm you order, and make the payment.

Go back to the **Cloud Disks** page and refresh it. You can find the new **cloud disk status** is **Available**.

Follow-up operations

Attach a cloud disk

Related APIs

To create a disk after creating an instance: *CreateDisk*

To create a cloud disk when creating an instance: *RunInstances* or *CreateInstance*

6.2 Create a cloud disk from a snapshot

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. This article describes how to create a cloud disk from a snapshot in the ECS console.

Scenarios

If you have 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 impaired instance.

Disk Performance

SSD Cloud Disks and Ultra Cloud Disks that are not created from snapshots can exhibit the maximum performance to its capacity, and no preconditioning is needed. However, for cloud disks created from snapshots, the initial performance decreases because data has 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?*.

Note

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 to work as data disks when creating an ECS instance. Those disks 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 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 *Limitations*.

- 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 and size 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 billing* 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

Before you start, make sure the following:

- You have created a snapshot for your instance, and you make sure the region and zone. For specific actions, see *Create snapshots*.
- because, to attach a cloud disk to an instance, they must be in the same zone of the same region.*Attach a cloud disk* The instance and the coud 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 **Snapshots and Images** > **Snapshots**.
3. In the upper-right corner of the **Disk List** page, click **Create Cloud Disk** to go to the **Create** page.
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 has no influence on this configuration.

   b. Click **Create a disk with snapshot** and select a snapshot.

   c. Specify the size of the cloud disk. The size range is 20 GiB−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 **Overview** and the cost.
7. Click **Buy Now**, confirm you order, and make the payment.

Go back to the **Cloud Disks** page and refresh it. You can find the **status** of the new cloud disk is **Available** when the new disk is created successfully.

Follow-up operations

Attach a cloud disk

Related APIs

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 work as a data disk. You have two
options to attach a cloud disk: attach them on 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 together with 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, but not in the **Locked** status.

  — The instance must not have payment overdue.

- 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. One 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 *Step 2. Create an instance* in *Quick Start*.

Attach a cloud disk on the Instance Disks page

If you want to attach multiple cloud disks to one ECS instance, attach them 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 a region.

4. Find an ECS instance and click its ID to go to the Instance Details page.

5. In the left-side navigation pane, click **Instance Disks**, and on the Disk List page, click **Attach Disk**.

6. In the dialog box, complete the following configurations:

   - **Target Disk**: Select a cloud disk in the **Available** status in the same region and zone.

   - **Release Disk with Instance**: If you select this option, the disk is released when you release its instance.

   - **Delete automatic snapshots when 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 complete data backup, we recommend that you do not select this option.

     Click **OK** , and then **Attach**.

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 different operations to make the disk ready for use. As shown in the following table.

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

Attach a cloud disk on the Disk List page

If you want to attach multiple cloud disks to different ECS instances, attach them on the Disks 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 >** > **Cloud Disks**.

3. Select a region.

4. Find a cloud disk in the **Available** status, and in the **Actions** column, select **More >** > **Attach**.

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 when 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 **Attach**.

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 different operations to make the disk ready for use. As shown in the following table.

Disk content	Operating system of the ECS instance	Follow-up operations
A new empty cloud disk	Linux	*Format and mount data disks for Linux instances*. If the cloud disk is larger than 2 TiB, see *Partition and format data disk more than 2 TB*.
	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 TB*.

Disk content	Operating system of the ECS instance	Follow-up operations
A cloud disk from a snapshot	Linux	Connect to the Linux instance and run the `mount` command to mount the partitions to make the disk ready for use.
	Windows	The cloud disk is ready for use.

Follow-up operations

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

- *#unique_39* 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*.

- If you want to restore the cloud disk to the status at a given time point, you can use its snapshot to *Roll back a cloud disk*.

- If your instance does not need a cloud disk, to reduce the cost, you can *Detach a cloud disk* and *Release a cloud disk*.

Related APIs

AttachDisk

6.4 Partition and format data disk more than 2 TB

If you want to partition and format a data disk more than 2 TB (referred to as a **large data disk** in this article, and a disk smaller than 2 TB is a **small data disk**), you must use the GPT format. This document describes how to partition and format a large data disk in different operating systems.

> **Note:**
> If you want to partition and format a data disk less than 2 TiB, please see *Format and mount data disks 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 tool and file system shown in the following table.

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

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

  Theoretically, this can work. But we recommend that you do not try this practice. Instead, create an empty large data disk, or create large data disk by using snapshots of large data disks, because of the following reasons:

  ― While creating a large data disk by using a snapshot of a small data disk, the system completes expansion at the block device level disk only, but not automatic conversion between the partition format and file system.

  ― If the MBR format is used in the snapshot of the small data disk, neither partition tool mentioned (`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.

- **Effect of data disk snapshots**

  Effect of Data Disk SnapshotsThe volume of data on a large data disk is huge, but the process for creating a snapshot of it is the same as for a small disk data, so the total time required for creating snapshots each day is proportional to the total data volume. Because the total time required to create snapshots is proportional to the total data volume, the more the dirty data is, the longer the snapshot creation time will be.

Windows _ Partition and format an empty large data disk

Consider Windows Server 2008 R2 64-bit system as example 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 detailed operation, 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 of **Server Manager**, select **Storage >** > **Disk Management**.
4. Find the disk that is to be partitioned and formatted (in this example, **Disk 4**). The disk status shows as **Offline**.
5. Right click the blank area around Disk 4, and then click **Online**.

   After going online, Disk 4 is in 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 need 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, in theory, is the maximum volume of NTFS containing $2^{64}-1$ clusters. Actually, in WinXP 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. Distribute 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. After the wizard to create a new simple volume is completed, click **Finish** to close **New Simple Volume Wizard**.

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 created a large data disk by using snapshots of a small data disk, you first need to convert the partition format of data disk from MNR to GPT, and then format the data disk. Data of the original snapshots will not be saved, so 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 disk like this, do the following to partition and format this data disk. The example operating system is Windows Server 2012 R2 64-bit, and we assume capacity of the data disk to be processed is 3 Tib.

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 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 Volume Size: 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, in theory, is the maximum volume of NTFS containing 2^{64}-1 clusters. Actually, in WinXP 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. 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. After the wizard to create a new simple volume is completed, click **Finish** to close **New Simple Volume Wizard**.

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

Linux _ 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 disk normally uses xfs or ext4 file system.

The example operating system is CentOS 7.4 64-bit.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. The expected result is as follows. If you see different returned information, you haven't mounted data disk.

   ```
   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 *Appendix2: Disable lazy init function.*.
> - If capacity of the data disk is 16 TiB, you have to format it by using e2fsprogs in the designated version. See*Appendix1: 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 it shows the new file system information in the returned result, the mount operation was successful and you can use the new file system. After mounting, do not need to restart the instance to use the new file system directly.

```
[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.

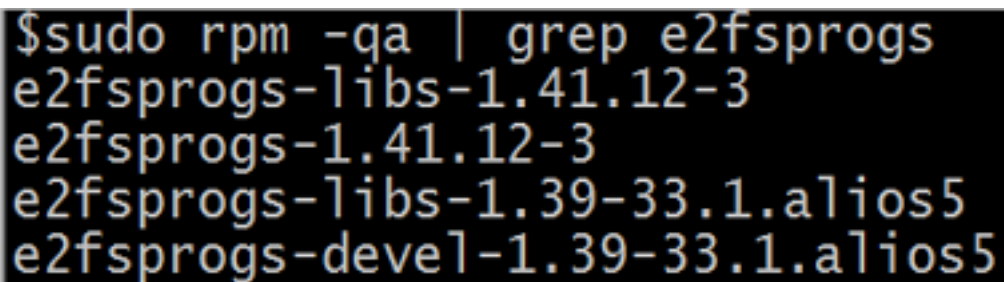You have now successfully partitioned and formatted a 3 TiB data disk.

Appendix 1: Update e2fsprogs

If the disk capacity is 16 TiB, you must use e2fsprogs of version1.42 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.

   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 version1.42.8. You can 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, in the system background, it will initiate a thread to initialize metadata of ext4 file system continuous ly to delay metadata initialization. Therefore, right after formatting a data disk, IOPS can be affected. For example, IOPS performance testing data in data disk will obviously be lower.

If you need to test performance of data disk right after formatting, you need to 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.

You can use the lazy init function according to your needs.

6.5 Detach a cloud disk

When 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.

- On a Windows instance, consider the following:

  — To guarantee the 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
    it offline in **Disk Management** .

- On a Linux instance, consider the following:

  — 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 shows the options available for you to detach a cloud disk in the ECS console.

Scenario	Applicable action
You want to detach one or more cloud disks from one instance.	*Detach cloud disks on the Instance Disk page*

Scenario	Applicable action
You want to detach one specified cloud disk.	*Detach a cloud disks on the Disk List page*

Detach cloud disks on the Instance Disk page

On the Instance Disk 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 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 from the Instance Disks page, follow these steps:

1. Connect to the instance and unmount the partitions. Follow different steps according to the operating system, as shown in the following table.

Operating system	Steps
Linux	Run `umount [partition]`. For example, `umount /dev/vdb1`.
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 a region.

5. Find an instance and click its ID to go to the Instance Details page.

6. In the left-side navigation pane, click**Instance Disks**.

7. Find a cloud disk, in the **Actions** column, select **More** > **Detach**.

    Only the cloud disks that have the following attributes can be detached:

    • **Disk Status** must be **In Use**.

    • **Detachable** must be **Yes**.

    • **Used As** must be **Data Disk**.

8. In the dialog box, click **Confirm Detaching**.

9. Optional. If you want to detach more cloud disks, repeat step 7 and step 8.

When the status of the cloud disk becomes **Available** , the disk is detached.

Detach a cloud disks on the Disk List page

You can detach one specified 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. Follow different steps according to the
operating system, as shown in the following table.

Operating system	Steps
Linux	Run `umount [partition]`. For example, `umount /dev/vdb1`.
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** > **Cloud Disks**.

4. Select a region.

5. Find a cloud disk, in the **Actions** column, select **More** > **Detach**.

Only the cloud disks that have the following attributes can be detached:

- **Disk Status** must be **In Use**.

- **Detachable** must be **Yes**.

- **Used As** must be **Data Disk**.

6. In the dialog box, click **Confirm Detaching**.

When the status of the cloud disk becomes **Available** , the disk is detached.

Related APIs

DetachDisk

Follow-up operations

If you no longer need the disk, you can *release it*.

6.6 Roll back a cloud disk

When errors occur to a cloud disk, if you have *created snapshots* for it, you can use the **Disk Rollback** feature to restore the disk to a healthy status at a given time point.

Note

Before you roll back a cloud disk, consider the following:

- Rolling back a cloud disk is an irreversible action. Once rollback is complete, data cannot be restored. Therefore, proceed with caution.
- 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, ascertain the following:

- *Create a snapshot* for the cloud disk, and no snapshot creation is in progress.
- The cloud disk has not been released.
- 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 to stop the instance in the Keep Instance Fees Apply mode. 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.

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 a region.

4. Find an instance and click its ID to go to the **Instance Details** page.

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

6. Find a snapshot, and in the **Actions** column, click **Disk Rollback**.

7. In the dialog box, click **OK**.

> **Note:**
>
> If you select **Start the instance immediately after the rollback**, the instance starts
>
> automatically after the disk is restored.

Related APIs

ResetDisk

Follow-up operations

If you resize a cloud disk after creating a snapshot, connect to the instance to resize its file system

. For more information, see:

- *Linux _ Resize a data disk*

- *Windows _ Resize a data disk*

6.7 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 (monthly or yearly subscription) instances, upfront

  payment is required for the service to be ready for use. For more information, see *Subscription*.

- For cloud disks created jointly with Pay-As-You-Go instances or separately created are billed

  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	Features	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
	Switch from Pay-As-You-Go to subscription	The system disks and data disks attached to the Pay-As-You-Go instances.	

6.8 Change a system disk (public image)

If you want to change the operating system running on your instance, you can use the Change System Disk feature to complete it.By changing a system disk, the system disk of your instance is replaced with a new cloud disk, which has a new disk ID, and the original system disk is released. If you want to change the operating system running on your instance, you can use the **Change System Disk** feature to complete it. You can replace the OS image with a public image, shared image, custom image, or an image from the image marketplace.

> 📋 **Note:**
>
> Microsoft has terminated technical support for Windows Server 2003. To guarantee your data security, we do not recommend that you continue running Windows Server 2003 on your ECS instance, This image is no longer available on the 2003 system. For more information, refer to *About Alibaba Cloud no longer supports Windows Server 2003 system mirroring*.

After replacing the system tray,

- a new system disk with a new disk ID is assigned to your instance, and the original one is released.
- The cloud type of the system disk cannot be replaced.
- The IP address and the MAC address remain unchanged after the system disk is changed.
- To make sure that your account have enough snapshot quota for the new system disk, you can delete unnecessary snapshots of the original system disk.

This article describes how to replace an existing image with a public image. If you need to use a non-public mirror, refer to *Change the system disk (custom image)*.

Notes

Before you begin, consider the following.

Risks

The risk of replacing the system tray is as follows:

- You have to stop your instance to change its system disk, which may interrupt your business operations.
- After replacement, you must redeploy the business runtime environment on the new system disk. There is a possibility of a long interruption of your business.

- Once you change the system disk, a new system disk, which has a new disk ID, is assigned. It means you cannot use all the snapshots of the original system disk to roll back the new system disk.

> **Note:**
>
> After you replace the system tray, the snapshot that you manually created is not affected, you can still create custom mirrors with these snapshots. If you have applied automatic snapshot policies to the original system disk, and set the auto snapshots to be released with the disk, the snapshot policies cannot work any more and all the auto snapshots of the original system disk are deleted automatically.

Considerations for changing between Windows and Linux

Regions that are not in mainland China do not support replacement between Linux and Windows.

> **Note:**
>
> For instances in those regions, a Linux or Windows edition can be only replaced by another edition of the same operating system type.

After the OS is changed between Windows and Linux, the file systems of the data disks cannot be recognized.

- If you do not have important data on the data disk, we recommend that you reinitialize the disk and format it to a recognizable file system.

- If you have important data on the data disk, follow these tips:

  — Replacing Windows with Linux: Install a software application, such as NTFS-3G, because the NTFS file system cannot be recognized by a Linux OS by default.

  — Replacing Linux with Windows: Install a software application, such as Ext2Read or Ext2Fsd, because ext3, ext4, and xfs cannot be recognized by a Windows OS by default.

When you replace a Windows edition with a Linux edition, two authentication methods are available: a password and an SSH key pair.

Prerequisites

If you want to change the OS to a Linux edition and to use an SSH key pair as the authentication method, create an SSH key pair.

Changing a system disk is so highly risky that it may cause data loss and business interruption. To minimize the impact of this operation, we recommend that you create a snapshot for the system disk.

> **Note:**
>
> • We recommend that you create snapshots at off-peak business hours. It may take about 40 minutes to create a snapshot of 40 GiB. Therefore, leave sufficient time to create a snapshot.
>
> • To create a snapshot, make sure the system disk has sufficient space available. We recommend that at least 1 GiB storage space is reserved. Otherwise, the instance cannot be started after the system disk is changed.

Procedure

Replace your system disk as follows:

1. Log on to the *ECS console*.
2. In the left-side navigation pane, click **Instances**.
3. Select a region.
4. Find an instance, and in the **Actions** column, select **More** > **Stop**and follow the prompts on the page to stop the instance.

> **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 **Notice** dialog box, click **OK**, On the **Stop** dialog box, select **Keep Instance with Fees** and click OK to stop the instance in the Keep Instance, Fees Apply mode. If you use the **No fees for stopped instances (VPC-Connected)** Otherwise, you may not be able to start the instance successfully after changing the system disk.

5. in the **Actions** column, select **More** > **Replacing the system disk**.

6. In the pop-up dialog box, after carefully reading the notes about replacing the system tray, click
   OK, replace the system disk.

7. On the **Change System Disk** page, complete the configurations:

   a. **Image type**: Select **Public Image** and select an image from the drop-down list.

   > 📋 **Note:**
   >
   > If you select an image other than a public image, see *Change the system disk (custom image)*.

   b. **System Disk**: You cannot change the cloud disk category. However, you can change the
   size of the disk to meet the requirements of your system disk. 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	Capacity Limit for capacity expansion (Gib)
Linux (excluding CoreOS) FreeBSD	[Max{20, current size of the system disk}, 500]
CoreOS	[Max{30, current size of the system disk}, 500]
Windows	[Max{40, current size of the system disk}, 500]

> **Note:**
>
> If your instance was renewed for configuration downgrade, you cannot change the system disk size until the next billing cycle.

c. **Security**:

- If the new operating system is a Windows edition, a password is the only authentication method.

- If you are changing the system disk of an I/O optimized instance, and the new operating system is a Linux edition, a password or an SSH key pair can be used as the authentica tion method. Set a password or bind an SSH key pair.

   d. Confirm the **cost**, which includes cost of the image and the system disk. For more information about pricing, see Pricing of Elastic Compute Service.

   e. Confirm the **cost**, which includes cost of the image and the system disk.

   f. Click ECS Service Terms and Product Terms of Service and then click **Confirm to change**.

Log on to the ECS console to monitor the system status. It may take about 10 minutes to change the system disk. 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:

- (Optional) *Set an automatic snapshot policy for the new system disk*. This operation The automatic snapshot policy applied on the old disk automatically fails after a new system disk has been replaced. You need to set up an automatic snapshot policy for the new system disk.

- If the original operating system is a Linux edition, data disks are attached to the instance, and the disks are set to be mounted automatically at startup of the instance, all mounting information is lost. Follow these steps to add new partition and mounting information to the /etc/fstab file. You do not have to partition or format the data disks again. For more information, see *Cite LeftQuick StartCite Right* Linux _ Format and mount a data disk.

    1. Optional. Back up /etc/fstab.

    2. Write new partition information to the /etc/fstab file.

    3. Check new partition information in the /etc/fstab file.

    4. Run `mount` to mount the partitions.

    5. Run `df-h` command to check file system space and usage.

        After the data partitions are mounted, the data disks are ready for use. You do not have to restart the instance.

Related APIs

ReplaceSystemDisk

6.9 Change the system disk (custom image)

By changing a system disk, the system disk of your instance is replaced with a new cloud disk with a new disk ID, and the original system disk is released. If you want to change the operating system running on your instance, you can use the **Change System Disk** feature to complete it. You can replace the OS image with a public image, shared image, custom image, or an image from the image marketplace.

> **Note:**
> Microsoft has terminated technical support for Windows Server 2003. To guarantee your data security, we do not recommend that you continue running Windows Server 2003 on your ECS instance,   and we have stopped providing Windows Server 2003 image. For more information, see *Offline announcement of Windows Server 2003 system image*.

After a system disk is changed,

- A new system disk with a new disk ID is assigned to your instance, and the original one is released.
- The cloud disk category is retained.
- The IP addresses and the MAC address of the instance are retained.

- To make sure that your account have enough snapshot quota for the new system disk, you can delete unnecessary snapshots of the original system disk.

This article describes how to replace an existing image with a non-public image. If you want to use a public image, see *Change a system disk (public image)*.

Note

Before you begin, consider the following.

Risks

Changing a system disk has the following risks:

- You have to stop your instance to change its system disk, which may interrupt your business operations.
- Once you change the system disk, you have to deploy your runtime environment on the new system disk, which may cause prolonged interruption to business operations.
- Once you change the system disk, a new system disk with a new disk ID is assigned. It means you cannot use all the snapshots of the original system disk to roll back the new system disk.

> **Note:**
>
> Changing a system disk has no effect on all the manual snapshots. You can use them to create custom images. If you have applied automatic snapshot policies to the original system disk, and set the auto snapshots to be released with the disk, the snapshot policies cannot work any more and all the auto snapshots of the original system disk are deleted automatically.

Considerations for changing between Windows and Linux

Regions that are not in mainland China do not support replacement between Linux and Windows.

> **Note:**
>
> For instances in those regions, a Linux or Windows edition can be only replaced by another edition of the same operating system type.

After the OS is changed between Windows and Linux, the file systems of the data disks cannot be recognized.

- If you do not have important data on the data disk, we recommend that you reinitialize the disk and format it to a recognizable file system.
- If you have important data on the data disk, follow these tips:

— Replacing Windows with Linux: Install a software application, such as NTFS-3G, because the NTFS file system cannot be recognized by a Linux OS by default.

— Replacing Linux with Windows: Install a software application, such as Ext2Read or Ext2Fsd , because ext3, ext4, and xfs cannot be recognized by a Windows OS by default.

When you replace a Windows edition with a Linux edition, two authentication methods are available: a password and an SSH key pair.

Prerequisites

Before replacing the existing image with a non-public image, complete the following:

- To replace the existing image with a custom image:If you change to a custom mirror:

  — To use an image running on an ECS instance,   create a snapshot for the system disk of the instance, and create a custom image from the snapshot. If both the instances are not in the same region, copy the image.

  — To use an on-premises image, import it in the ECS console or use Packer to create and import an image. The image and the instance must be in the same region.

  — To use an image in a region other than that of the instance, copy the image.

  > 📋 **Note:**
  >
  > When you change a system disk, all the images obtained by using the preceding methods are displayed in the drop-down list of **Custom Image**.

- To use an image owned by other Alibaba Cloud account, share the image.

- If you want to change the OS to a Linux edition and to use an SSH key pair as the authentica tion method, create an SSH key pair.

Changing a system disk is so highly risky that it may cause data loss and business interruption. To minimize the impact of the operation, we recommend that you create a snapshot for the system disk.

> 📋 **Note:**
>
> - We recommend that you create snapshots at off-peak business hours. It may take about 40 minutes to create a snapshot of 40 GiB. to create a snapshot of 40 GiB. Therefore, leave sufficient time to create a snapshot. Creation of a snapshot may reduce the I/O performance of a block storage device, generally it is less than 10%, which results in sharp decrease in I/O speed.

- To create a snapshot, make sure the system disk has sufficient space available. We recommend that at least 1 GiB storage space is reserved. Otherwise, the instance cannot be started after the system disk is changed.

Procedure

1. Log on to the *ECS console*.

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

3. Select a region.

4. Find an instance, and in the **Actions column**, select **More** > **Stop**.

   When the instance is in the **Stopped** status,

5. in the **Actions** column, select **More** > **Change System Disk**.

6. In the dialog box, read the note and click **Yes. Change system disk**.

7. On the **Change System Disk** page, complete the configurations:

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

   b. **System Disk**: You cannot change the cloud disk category. However, you can change the size of the disk to meet the requirements of your system disk. The maximum size is 500 GiB.   The minimum size of the system disk is determined by the current size of the system disk and the image size.

Image	Size limit (GiB)
Linux (excluding CoreOS) + FreeBSD	[Max{20, current size of the system disk}, 500]
CoreOS	[Max{30, current size of the system disk}, 500]
Windows	[Max{40, current size of the system disk}, 500]

   > **Note:**
   >
   > If your instance was renewed for configuration downgrade, you cannot change the system disk size until the next billing cycle.

   c. **Security**:

   - If the new operating system is a Windows edition, a password is the only authentication method.

- If you are changing the system disk of an I/O optimized instance, and the new operating system is a Linux edition, a password or an SSH key pair can be used as the authentication method. Set a password or bind an SSH key pair.

  d. Confirm the cost, which includes cost of the image and the system disk.

  e. Click **Confirm** to change and follow the prompts to complete the order.

Log on to the ECS console to monitor the system status. It may take about 10 minutes to change the system disk. 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:

- Optional. Apply an automatic snapshot policy to the new system disk. The auto-Snapshot policy is bound to the disk ID. The automatic snapshot policy applied on the old disk automatically fails after a new system disk has been replaced. You need to set up an automatic snapshot policy for the new system disk.

- If the original operating system is a Linux edition, data disks are attached to the instance, and the disks are set to be mounted automatically at startup of the instance, all mounting information is lost. Follow these steps to add new partition and mounting information to the /etc/fstab file. You do not have to partition or format the data disks again. For more information, see Linux _ Format and mount a data disk.

  1. Optional. Back up /etc/fstab.

  2. Write new partition information to the /etc/fstab file.

  3. Check new partition information in the /etc/fstab file.

  4. Run `mount` to mount the partitions.

  5. Run `df   -h` to check the file system space and usage.

  After the data partitions are mounted, the data disks are ready for use. You do not have to restart the instance.

Related APIs

ReplaceSystemDisk

6.10 Monitor a cloud disk

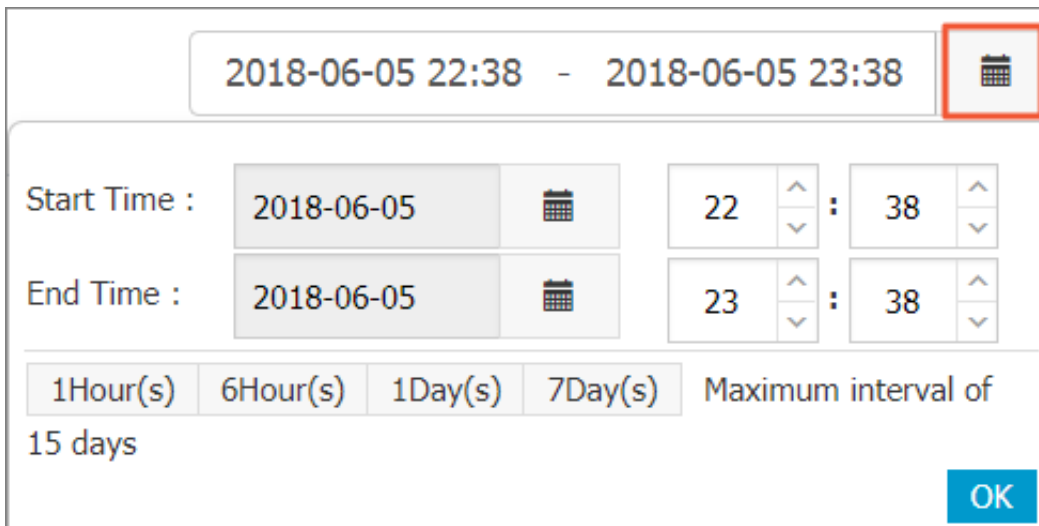When using a cloud disk, consider the following performance parameters:

- IOPS: Indicates Input/Output Operations per Second, which means the amount of write or read operations can be performed in one second. Transaction-intensive applications are sensitive to IOPS.

- Throughput: Measures the data size successfully transferred per second, measured in MBps. Applications that require mass read or write operations are sensitive to throughput.

You can monitor the IOPS and throughput of a cloud disk in the ECS console. Alternatively, if you have installed 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** > **Cloud Disks**.

3. Select a region.

4. Find a cloud disk and click its ID to go to the **Disk Details** page.

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

6. On the **Monitoring Information** page, click the [icon] 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.

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.
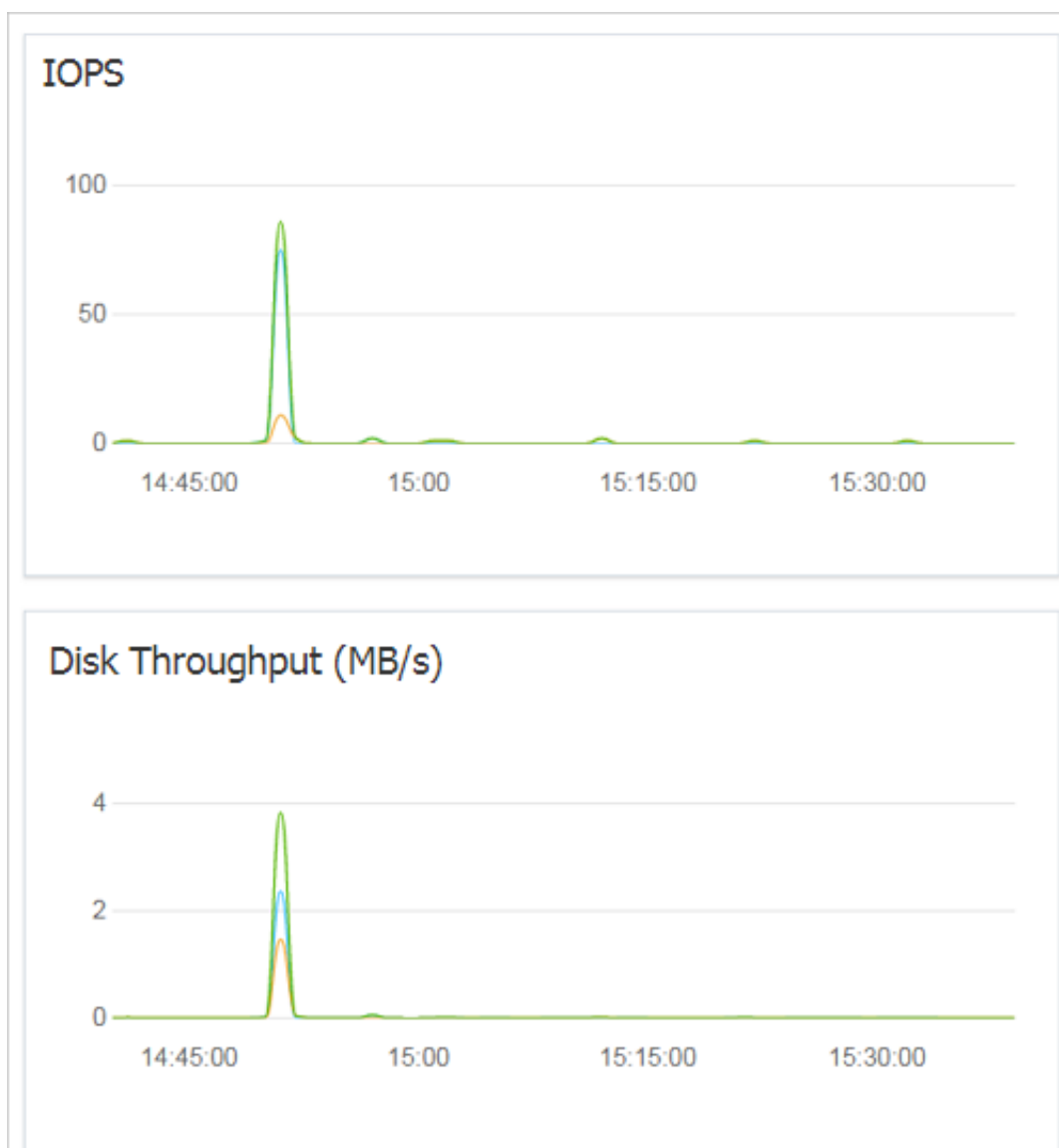
6.11 Release a cloud disk

Release a cloud disk in the Available status if your business does not require it. Otherwise, you are charged for it. Releasing a data disk is a permanent action and is irreversible. After release, the data on the data disk cannot be restored. Proceed with caution.

Note

When releasing a cloud disk, consider the following:

- Only the cloud disks that are in the **Available** status can be released independently. Other cloud disks, such as those used as system disks or those Subscription 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, the automatic snapshots are released together with their cloud disks. However, those created manually are not. You can change the snapshot release configuration when attaching a cloud disk.

> ![note icon] **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 your business no longer require.

- You can have data backed up before releasing a cloud disk. For example, Create a snapshot.

Procedure

1. Log on to the *ECS console*.

2. In the left-side navigation pane, select **Block Storage** > **Cloud Disks**.

3. Select a region.

4. Select the disk that you want to release (in the **Available** status), and in the **Actions** column, select **More** > **Release**.

5. In the Release dialog box, read the note and click **Confirm Release**.

Related APIs

DeleteDisk

7 Images

7.1 Manage custom images

After creating custom images, you can delete custom images that are no longer required, or modify the name and description of the custom images to help you organize and identify them.

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 a region.

4. Find the custom image to be edited.

5. Click the ✏ icon, and enter the image name.
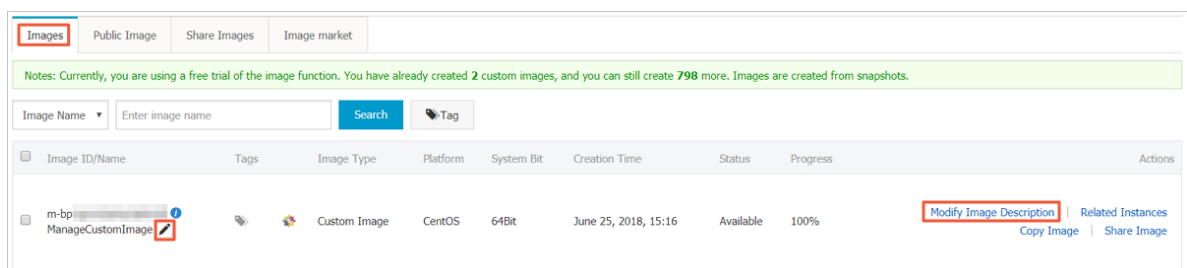
6. Click **Modify Description**, and in the dialog box, edit the new **Custom Image Description**.

7. Click **Save** to complete the description modification.

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 a region.

4. Select one or more custom images that you want to delete, and then click **Delete**.
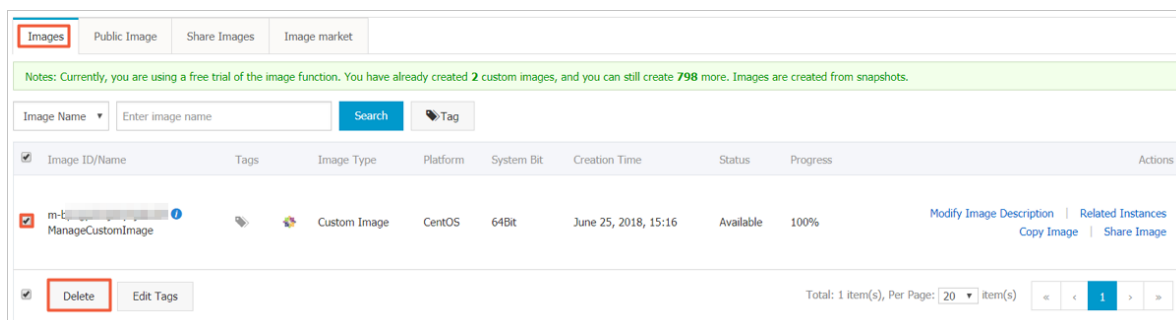
5. In the dialog box that appears, select the method for deleting the custom images:

- **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.
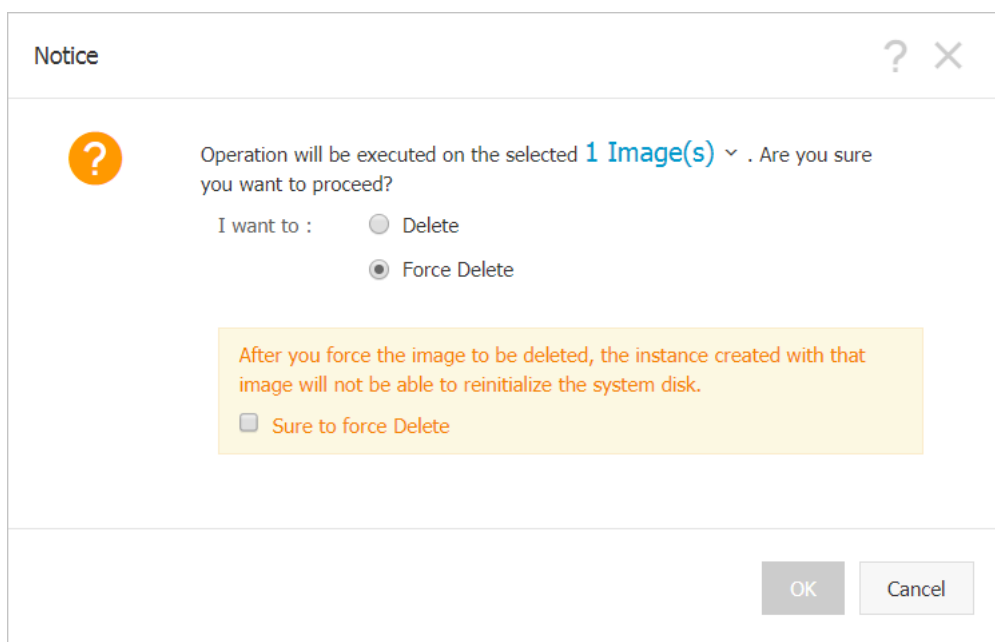
6. Click **OK** to confirm.

Alternatively, you can delete custom images by calling the ECS API `DeleteImage`.

7.2 Copy images

Copying images allows you to deploy an application across regions by running 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

- 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, perform the following steps:

1. Log on to the *ECS console*.
2. Select a 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**. In the **Actions** column, click **Copy Image**.

   > **Note:**
   >
   > If your custom image is larger than 200 GiB, you need to open a ticket to complete the operation. When you click **Copy Image**, you are directed to open a ticket.

5. In the **Copy Image** dialog box, you can find the ID of the selected image. Also, you need to make 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.

You can also use the ECS APIs *CopyImage* and *CancelCopyImage* to perform the operation.

Next Steps

When the image appears as **Available**, you can use it to *create an instance* or *change the system disk*.

You can also view the copied snapshot in the target region.

FAQs

FAQs about copying images

7.3 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 those who share an image.

Attention

You can only share your own custom images, not the images 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 sensitive data and files have been removed 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.

Impacts of deleting shared images

If your custom image has been shared with other accounts, you need to remove all the sharing
relationships for that image before you can delete the image. After deleting a shared custom
image:

- The users who are using the shared image can no longer 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, perform the following steps:

1. Log on to the *ECS console*.

2. Select a region.

3. In the left-side navigation pane, choose **Snapshots and Images** > **Images**.

4. Select a **Custom Image**. In the **Actions** column, click **Share Image**.
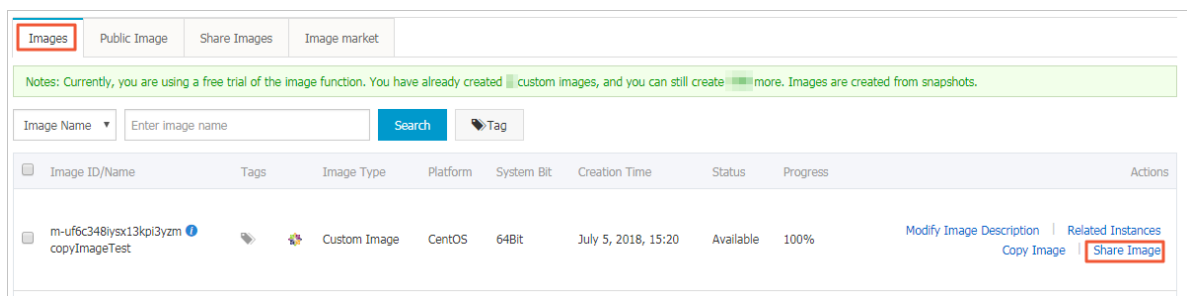
5. In the pop-up dialog box, select **Alibaba Cloud Account ID** in the **Account Type** drop-down
   list. Then, enter the account ID you want to share the image with in the **Account** box. You can
   refer to *Appendix. How to get the account ID* to get the Alibaba Cloud account ID.

> **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. If that account
> has already created an instance by using this shared image, the instance will be unable to *re-
> initialize the system disk*.

6. (Optional) For those with whom you share an image, they 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 *Change the system disk (custom image)* for instances.

Appendix. How to get the account ID

To query your account ID, perform the following steps:

1. Log on to the ECS console.

2. Move the cursor onto the upper right user avatar, such as funCustomer*****@aliyun.com. From the pop-up account menu, click **Security Settings**.

3. On the page that appears, find the account ID.
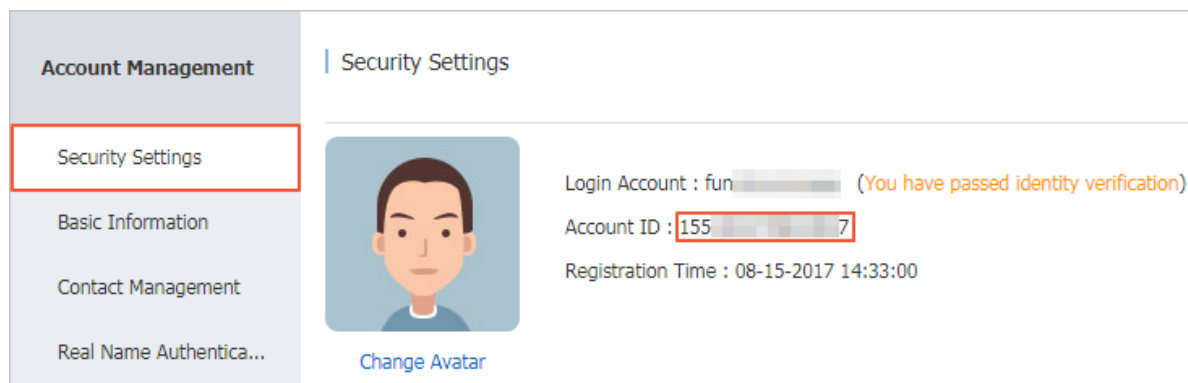
7.4 Import images

7.4.1 Notes for importing images

To guarantee the usability of an imported image and improve the importing efficiency, pay attention to the following before importing an image:

Depending on the operating system, the notes vary for *Windows images* and *Linux images*.

Windows images

Important suggestions

- Verify the integrity of the file system before importing images for Windows.

- Check the remaining space on the system disk to make sure the system disk is not full.

- Disable the firewall and allow access to RDP port 3389.

- The logon password for the administrator account must be 8-30 characters long and must contain uppercase/lowercase letters, numbers, and special characters simultaneously. The special characters can be: ( ) ` ~ ! @ # $ % ^ & * - + = | { } [ ] : ; ' < > , . ? /

- 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 for the importing must be from 40 GiB to 500 GiB.

- Do not modify critical system files.

What is 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, *convert image file format* before importing it.
> It is recommended that you convert the format to VHD, which features 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 is not supported

- The installation of qemu-ga in an image is not supported. Otherwise, some services needed by
   ECS may become unavailable.

- Windows XP, Windows 7 (professional and enterprise editions), Windows 8, and Windows 10.

Linux images

Important suggestions

- Verify the integrity of the file system before importing images for Linux.

- Check the remaining space on the system disk to make sure that the system disk is not full.

- Disable the firewall and allow access to TCP port 22.

- Install the virtualization platform XEN or KVM drives.

- It is recommended to *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 is supported

- Images in RAW, qcow2, or VHD format.

> **Note:**
>
> If you want to import an image in another format, *convert image file format* before importing it. It is recommended that you convert the format to VHD that features smaller transmission capacity.

- The xfs, ext3, and ext4 file systems and MBR partitions.

What is not supported

- Multiple network interfaces.

- IPv6 addresses.

- System disk partitions cannot be adjusted. Currently, only a single root partition is supported.

Issues to note

Depending on whether the Linux system image you are importing is a standard platform image, you need to note different issues.

- The official operating system releases are defined as the *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 in the list of 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:

  — Others 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 Others Linux. 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	• Do not modify `/etc/issue*`. Otherwise, ECS cannot properly identify the system release, leading to system creation failure. • Do not modify `/boot/grub/menu.lst`, or the ECS instance cannot be started. • Do not modify `/etc/fstab`, or the exception partition cannot be loaded, leading to ECS instance start failure. • Do not change `/etc/shadow` to **read only**, or you may be unable to modify the password file, leading to system creation failure. • Do not enable SELinux by modifying `/etc/selinux/config`, or the system may fail to start.	Does not meet the requiremen ts of standard platform images
Requirements for applications	Do not install qemu-ga in an imported image, or some services required by Alibaba Cloud may become unavailable.	Does not meet the requiremen ts of standard platform images

7.4.2 Install cloud-init for Linux images

If you need to create an ECS instance using an existing image, you can import the image to configure the instance. To guarantee successful configuration of the hostname, NTP source, and yum source of the imported Linux image, you are advised to install cloud-init in your on-premise server, virtual machine (VM), or cloud host before import.

Limitations

• Currently, cloud-init supports the following Linux distributions: 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-premise server, VM, or cloud host, make sure that the cloud-init version is later than 0.7.9 because images using an earlier-version cloud-init may cause instance configuration failure in NTP, hostname, and yum.

  1. Log on to your on-premise server, VM, or cloud host.

  2. Run `cloud-init --version` to query the cloud-init version.

     You can use the cloud-init if its version is later than 0.7.9. Otherwise, do as instructed in *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 as an example. If you manage packages using `zypper` or `apt-get`, the installation methods are similar to `yum`.

Install cloud-init

Do as follows to install cloud-init:

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.

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 login 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

- The six library is missing from Python if the following message is displayed during installation.
  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   )
```

- The oauthlib library is missing from Python if the following message is displayed during
  installation. 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.

Next step

You can *import the image to ECS*.

Reference

cloud-init *Alibaba Cloud (AliYun)*

7.4.3 Install virtio driver

To avoid failure in starting the Linux instances created by using the imported images of your
server, virtual machines, or cloud hosts, *Import custom images* an Xen (pv) or virtio driver must
be installed on your on-premises image and configured before importing. Follow these steps to
check whether you must install the driver manually, and then install and configure the virtio driver
for a Linux server if needed.

Images requiring no manual installation

After you import images in *Import custom images*, if the operating systems of your image is listed
in the following, 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.

> 📋 **Note:**
>
> - If VIRTIO_BLK and VIRTIO_NET do not exist in the output, the virtio driver is not built in the kernel, and 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 and parameter **CONFIG_VIR TIO_NET** are y, the virtio driver is already built in the kernel.You can read Notes for importing custom images *Notes for importing images* and import the image *Import custom images*.
>
> - If the values of parameter **CONFIG_VIRTIO_BLK** and and parameter **CONFIG_VIR TIO_NET** are m, continue to step 2.

2. Run `lsinitrd /boot/initramfs-$(uname -r).img | grep virtio`to make sure virtio driver has been complied in the temporary root file system of initramfs or initrd.

```
[root@                    ~]# lsinitrd /boot/initramfs-$(uname -r).img | grep virtio
Arguments: -f --add-drivers ' xen-blkfront xen-blkfront virtio_blk virtio_blk virtio_pci 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@                    ~]#
```

> **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. After reading Notes for importing custom images *Notes for importing images*, you
>   can directly import the image *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

After checking, 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. Take CentOS as an example:

- 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)"
```
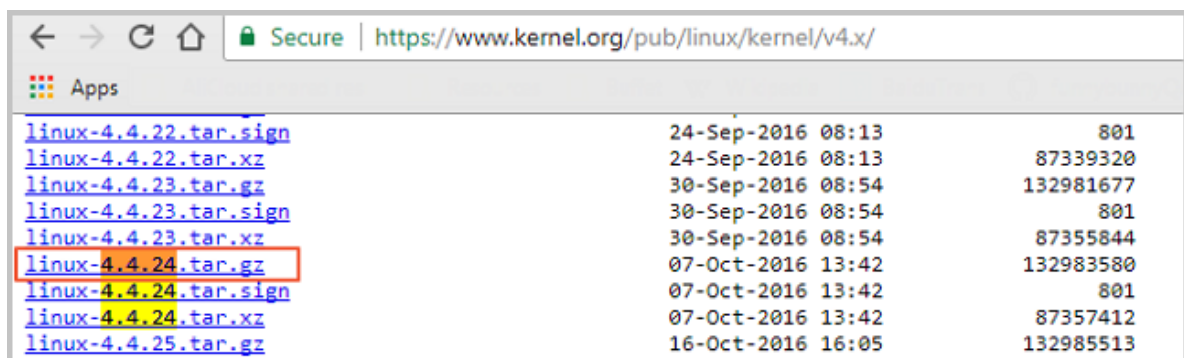
To compile and install virtio driver

Take Redhat server as an example:

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@iZbp1127hr3wi6p2cq9lnbZ ~]# uname -r
4.4.24-2.al7.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*.

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
```
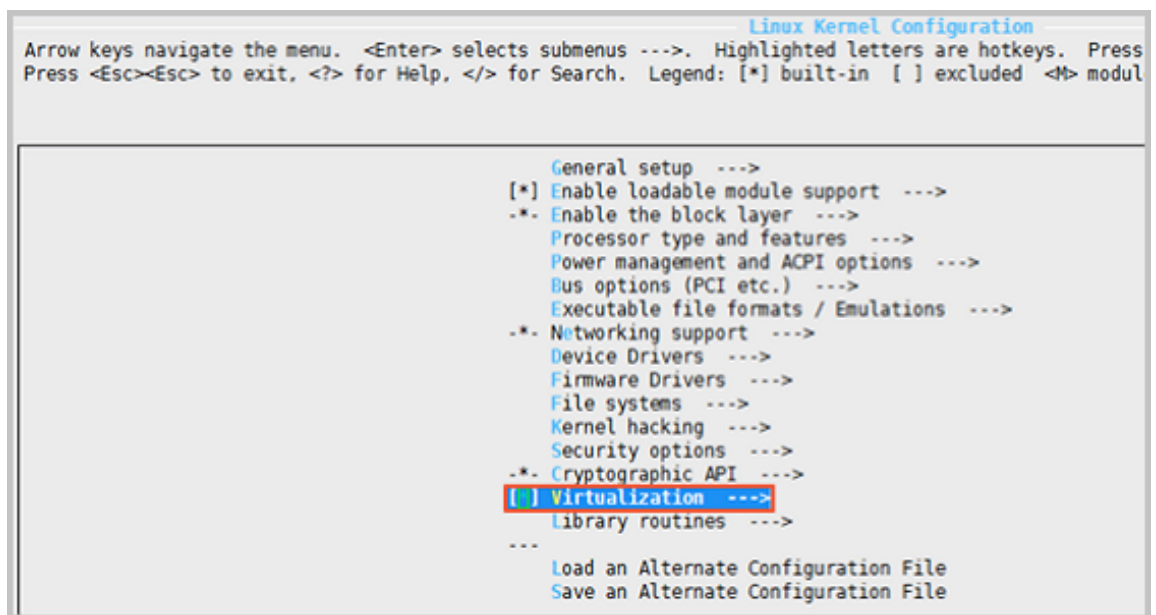
2. Configure the corresponding settings of virtio driver in the following windows:

> **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.

Make sure that you have selected the option of KVM (Kernel-based Virtual Machine).
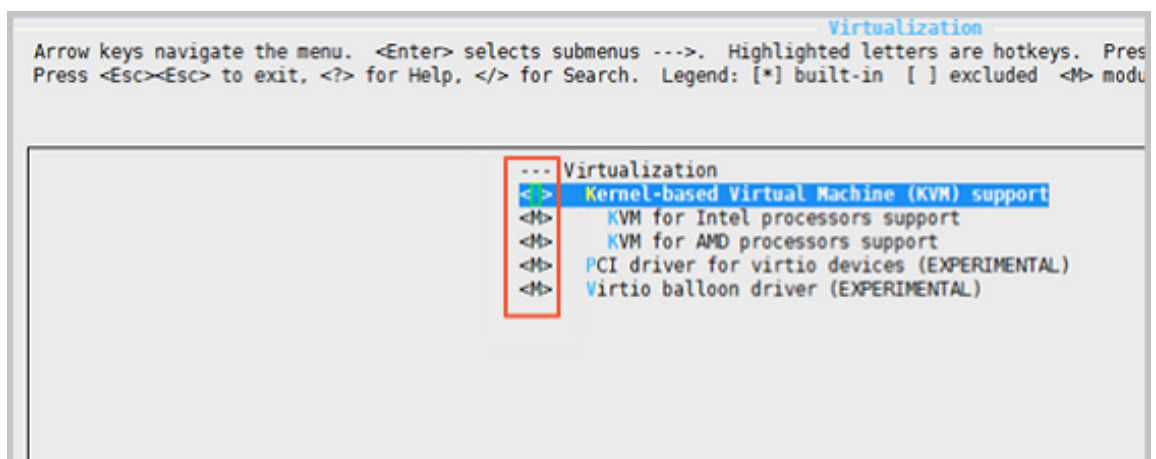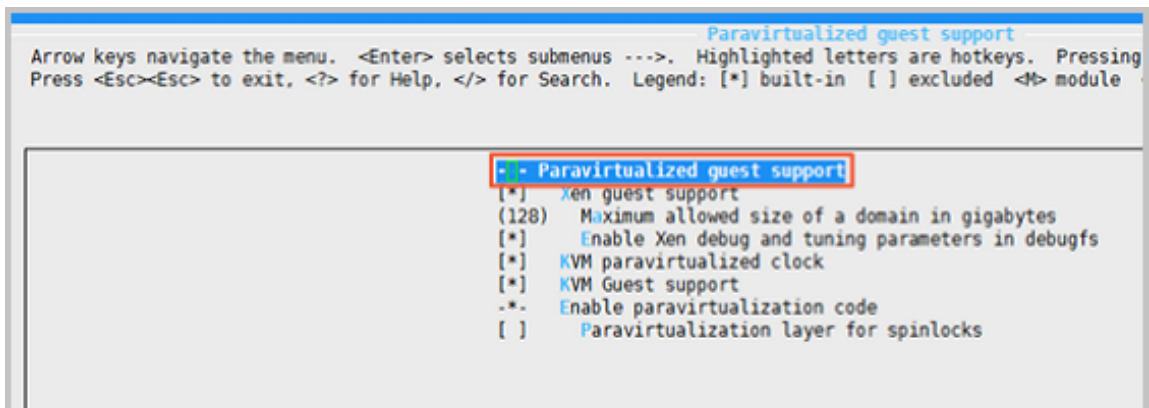


```
Processor type and features  --->
   [*] Paravirtualized guest support  --->
     --- Paravirtualized guest support
  (128)   Maximum allowed size of a domain in gigabytes
  [*]    KVM paravirtualized clock
```

```
[*]    KVM Guest support
```



```
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. 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.

Next steps

7.4.4 Customize Linux images

If your image operating system (OS) is not supported by Alibaba Cloud and cloud-int cannot
be installed, you can select **Customized Linux** when importing a customized image. Alibaba
Cloud regards the customized Linux image as an unrecognized operating system type and 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 by following the instructions provided
in this document 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. You can operate
    on the instance only after changing the password.

  — 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_cus`
   `tom_image` directory. Alibaba Cloud will automatically create an **os.conf** file if there is not any.

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	Host name
password	Password, a Base64-encoded string
eth0_ip_addr	IP address of the eth0 NIC
eth0_mac_addr	MAC address of the eth0 NIC
eth0_netmask	Network mask of the eth0 NIC
eth0_gateway	Default gateway of the eth0 NIC
eth0_route	eth0 intranet route list, in which routes are separated by semicolons (;) by default
eth1_ip_addr	IP address of the eth1 NIC
eth1_mac_addr	MAC address of the eth1 NIC
eth1_netmask	Network mask of the eth1 NIC
eth1_gateway	Default gateway of the eth1 NIC
eth1_route	eth1 internet route list, in which routes are separated by semicolons (;) by default
dns_nameserver	DNS address list, in which addresses are separated by spaces by default

For instances using VPCs

```
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="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	Host name
password	Password, a Base64-encoded string
eth0_ip_addr	IP address of the eth0 NIC
eth0_mac_addr	MAC address of the eth0 NIC
eth0_netmask	Network mask of the eth0 NIC
eth0_gateway	Default gateway of the eth0 NIC
eth0_route	eth0 intranet route list, in which routes are separated by semicolons (;) by default
dns_nameserver	DNS address list, in which addresses are separated by spaces by default

Attentions on script parsing

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 achieve so, you can 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 advised to use `uuid` or `label` to indicate devices in the first partition. The user password is a Base64-encoded string and must be Base64-encoded in the script.

- **Judgement on VPCs and classic networks**: When using the parsing script to judge the network type, you can check whether there is `eth1_route` or other eth1-related configuration

item. Parse and process the instance accordingly after judging 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 advised 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 exitst
 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_a
ddr}
  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
```

```
  inteface_cfg="/etc/sysconfig/network-scripts/ifcfg-${interface}"
  cat << EOF > $inteface_cfg
DEVICE=$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
```

7.4.5 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, convert the format before importing the image. This article 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 mutual 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 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.

FAQ

If errors occur during qemu-img installation and there are no clear prompts about the missing dependent libraries, you can 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

7.4.6 Import custom images

You can import on-premise image files to the ECS environment for deploying your business. Imported custom images appear in your custom images list under the target region. You can use these images to create ECS instances or change system disks.

> 📋 **Note:**

- Importing custom images is a time consuming task. The duration depends on the size of the image file and the number of concurrent tasks, so you need to wait patiently.

> • 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, you should have done the following:

- Learn about the limitations and requirements of importing custom images by referring to *Notes for importing images*, *Customize Linux images*, and *Convert image format*.
- *Sign up for 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, perform these steps:

1. You can use an OSS third-party client, OSS API or OSS SDK to upload the prepared custom image. For how to upload a file larger than 5 GiB, see *OSS Multipart upload*.
2. Log on to *ECS console*.
3. In the left-side navigation pane, choose **Snapshots and Images** > **Images**.
4. On the Images page, click **Import Image**.

5. In the **Import Image** dialog box, click **Confirm Address** as shown

below.
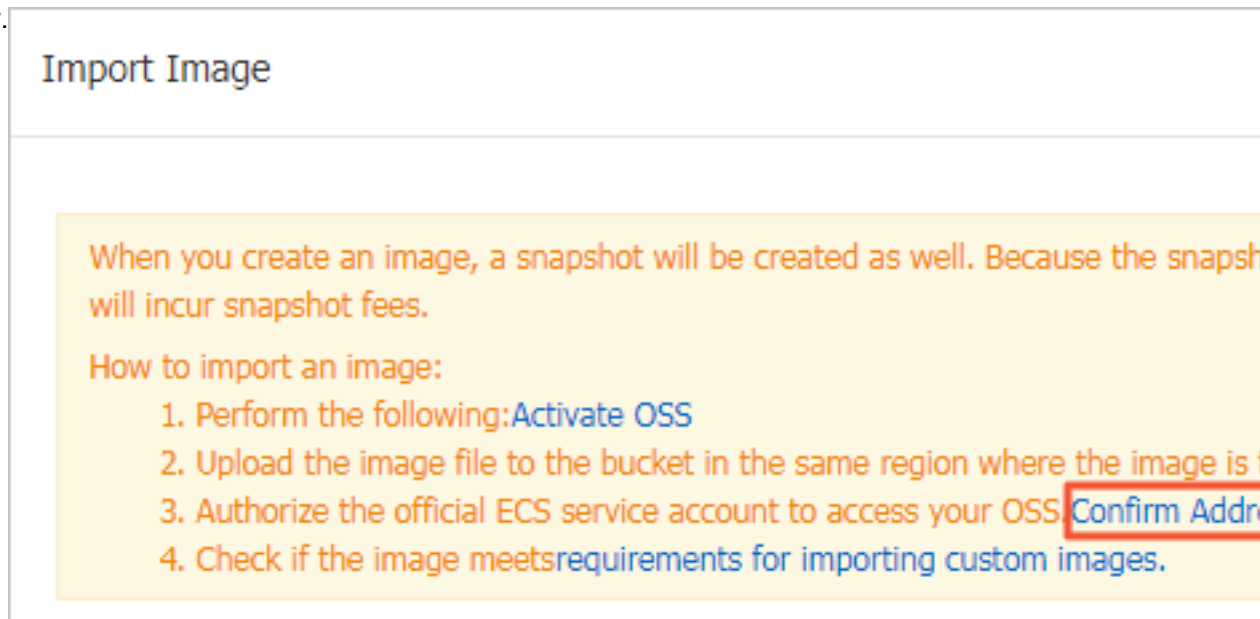
6. In the **Cloud Resource Access Authorization** window, select `AliyunECSImageImport DefaultRole` and `AliyunECSExportDefaultRole`, then click **Confirm Authorization Policy** to allow the ECS service to access your OSS resources.

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 *OSS Download an object*.

- **Image Name**: Specify the name of the custom image file displayed after it is imported. It can be 2 to 128 characters in length. Beginning with upper/lower case letters or Chinese characters, it allows numbers, periods (.), underscores (_), colons (:), and hyphens (-).

- **Operating System**: Select **Windows** or **Linux**, which should be 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 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.

- **System 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**: Fill up the description of the image to facilitate subsequent management.

- **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. After the information is confirmed, click **OK** to create a task to import the image.

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 *Manage Tasks*, then 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*

7.5 Export custom images

You can export custom images for on-premise testing or Apsara stack environments.

> **Note:**
>
> - Exporting custom images is a time-consuming task, so you need to wait patiently. The duration of exporting depends on the size of the image file and the number of export tasks in the queue.
>
> - The exported images are stored in your *OSS* bucket. 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 the custom images that are created by a system disk snapshot from the *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 images to *create an instance by using the wizard*, you need to confirm that the file device recorded in /etc/fstab corresponds to the exported data disk snapshot information.
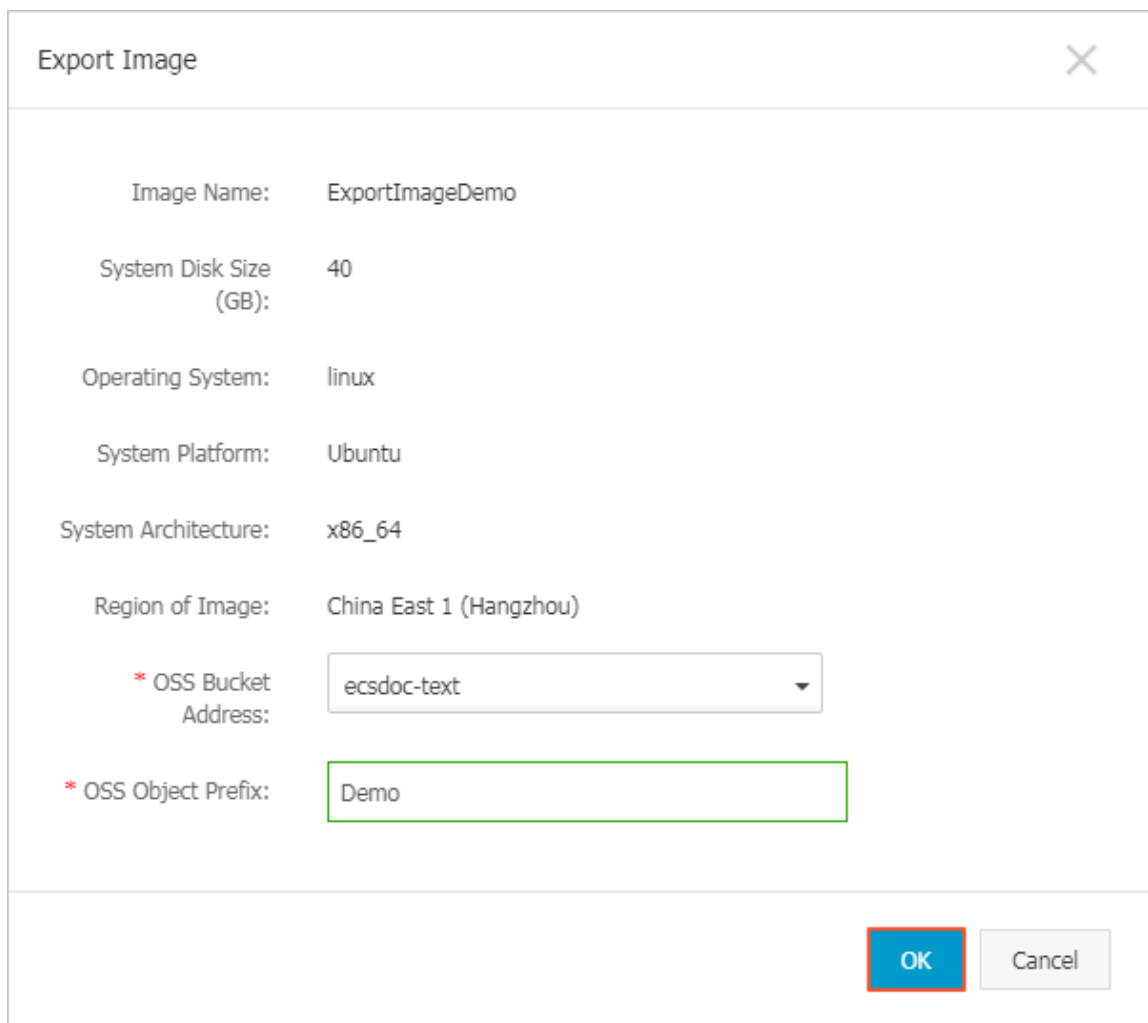
Prerequisites

Before exporting a custom image, you need to do the following:

- 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 *OSS 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 a region.

4. Find the custom image you want to export. In the **Action** column, click **Export Image**.

   a. In the **Export Image** dialog box, click **Confirm Address** shown in the figure below.

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** to export the image.

8. (Optional) Cancel the image export task. Before the task is completed, you can go to the *Manage Tasks* 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 above 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 shapshot 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 defaulted to RAW.

7.6 Open source tools

7.6.1 Create and import on-premise 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.

To create an on-premises image by yourself and then upload it on a cloud platform is a complex process. However, by using Packer, you can create identical on-premises images for multiple platforms from a single source configuration. Follow these steps to create an on-premises image for CentOS 6.9 on an Ubuntu 16.04 server and to upload it to Alibaba Cloud. To create on-premises images for other operating systems, you can **customize your Packer templates** as necessary.

Prerequisites

- You must have the *AccessKey* ready to fill out the configuration file. .

> **Note:**
> The AccessKey has a high level of account privileges. We recommend that you *create a RAM user* and use the RAM account to create *AccessKey* to prevent data breach.

- Before uploading your on-premises images to Alibaba Cloud, you must *sign up for OSS*.

Sample 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
    pdpe1gb 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* configuration file centos.json as needed.

   ```
   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.
```

The running result of the sample 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: XXXXXXXX
```

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

The image file created in the preceding *sample* is for the CentOS 6.9 operating system only. To

create images for other operating systems, you must customize the Packer template.

For example, the following JSON file is customized based on the template 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 }}:{{ .HTTPPort }}/{{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 *sample* 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 `iso_checksum` or `iso_checksum_url` parameter. If you have the `iso_checksum` parameter specified, the `iso_checksum_url` parameter is ignored automatically.

Parameter	Type	Description
iso_checks um_type	String	The type of the checksum specified in iso_checksum. Optional values: • none: If you specify none for iso_checksum_type , the checksuming is ignored, thus none is not recommended. • md5 • sha1 • sha256 • sha512
iso_checksum_url	String	This is a URL pointing 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 at least one of the **iso_checks um** or the **iso_checksum_url** parameter. If you have the iso_checksum parameter specified, the **iso_checks um_url** parameter is ignored automatically.
iso_url	String	This is a URL pointing to the ISO file and containing the installation image. This URL may be an HTTP URL or a file path: • If it is an HTTP URL, Packer downloads the file from the HTTP link and caches the file for running it 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 **disk_image** to `true`.
headless	boolean	By default, Packer starts the virtual machine GUI to build a QEMU virtual machine. If you set **headless** to `True`, a virtual machine without any console is started.

For more information about other optional parameters, see Packer *QEMU Builder*.

Parameters in a Packer provisioner

The provisioner in the preceding *sample* 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

Parameter	Type	Description
		the RAM account to create an AccessKey to prevent data breach.
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 prevent data breach.
region	String	Select the region where you want to upload your on-premises image. In the sample, 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	Your 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: • linux • windows
image_platform	String	Distribution of the image. For example, CentOS.
image_arch itecture	String	The instruction set architecture of the image. Optional values: • i386 • x86_64
format	String	Image format. Optional values: • RAW • VHD

For more information about other optional parameters, 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 Alibaba repository on GitHub *opstools*.

- For more information about Alibaba Cloud and Packer project, visit the Alibaba & Packer repositories on GitHub *packer-provider*.

- For more information about configuration file ks.cfg, see *Anaconda Kickstart* .

7.6.2 Use Packer to create a custom image

Packer is a convenient open-source tool to create custom images. It runs on major operating systems. This document provides information about how to install and use Packer. With Packer, you can easily create a custom image by using only one or two lines of commands.

Prerequisites

You must have the AccessKey ready. For more information, see *Create AccessKey* .

> 📋 **Note:**
>
> The AccessKey has a high level of account privileges. To avoid improper operations and data breach, we recommend that you *Create a RAM user*, and act as a RAM user to *create your AccessKey*.

Step 1. Install Packer

Go to the official *download page of Packer* where you can choose and download the version of Packer for your operating system. Follow these steps or visit the official *installation page of Packer* for how to install Packer.

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. ou 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

Take Windows Server 2012 64-bit as an example:

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, firstly, create a JSON format template file. In the template, specify the *Alibaba Cloud Image Builder* and *Provisioner* for the custom image to be created. Packer has diverse provisioners for you to choose from when configuring the content generation mode of the custom image.In the following alicloud JSON file, we have used the *Shell* provisioner 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:**

You must customize the values of the following parameters.

Parameter	Description
access_key	Your AccessKey ID For more details, see creating an accesskey.
secret_key	Your AccessKey Secret For more information, see *Create AccessKey*.
region	The region of the temporary instance used to create the custom image.
image_name	The custom image's name
source_image	You can retrieve the basic image name from Alibaba Cloud public image list.
instance_type	Type of the temporary instance generated to create the custom image.
internet_charge_type	Internet bandwidth billing method for the temporary instance generated for creating the custom image.

provisioners	Type of *Packer Provisioner* used for creating the custom image

Step 3. Create a custom image by using Packer

Follow these step to specify the Packer template file and create a custom image:

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 sample runs like follows. The sample 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-b1ba08ea94b8
==> alicloud-ecs: Start creating alicloud vpc
--------------------------
==> alicloud-ecs: Provisioning with shell script: /var/folders/3q/
w38xx_js6cl6k5mwkrqsnw7w0000gn/T/packer-shell257466182
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
```

Next steps

You can 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.

8 Cloud Assistant Client

The cloud assistant client is an agent facilitating cloud assistant command invocation on ECS instances. The cloud assistant client does not perform any operations. You can start all the operations and within your controllable range. Instances that are created later than Dec 1, 2017 are pre-installed with the cloud assistant client by default. If your ECS instance was created earlier than Dec 1, 2017, and you want to use cloud assistant service, you can install the cloud assistant client by yourself. This topic illustrates how to install, update, and disable the cloud assistant client in an instance.

Install cloud assistant client

Windows instance

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 instance

Based on the distribution of Linux, to install cloud assistant client, select the most appropriate approach from the following.

- 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 cloud assistant client runs every 1 hour to query update resources for the client. The update process is located at the directory of:

- 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:

- 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. Once you disable the client, *Aliyun* service is also disabled and stopping the instance in the ECS console may fail. We recommend that you perform with caution to disable the client.

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 service.

```
systemctl stop agentwatch
```

```
chkconfig agentwatch off
```

References

Visit the *GitHub aliyun_assist_client* to explore the open source stack of cloud assistant.

You can use the cloud assistant client for the following:

- *Cloud assistant*
- *InvokeCommand*

9 Cloud assistant

9.1 Create commands

You can use cloud assistance commands to perform routine tasks for ECS instances. These tasks include fast execution of automatic maintenance scripts, process polling, resetting of user password, installation and uninstallation of software, application update, and patch installation. Command types can either be Bat or PowerShell for Windows, or Shell for Linux.

Limits

- Within an Alibaba Cloud region, you can create at most 100 cloud assistant commands.
- A script cannot exceed 16 KB after Base64 encoding.

Create commands

To create a command on the ECS Console, take the following steps:

1. Log on to the *ECS console*.

2. In the left-side navigation pane, select **Cloud Assistant**.

3. Select a region.

4. Click **Create Command**, and in the right-side pop-up window.

    a. Input a **command name**, such as HelloECS.

    b. Input a **command description**, such as 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. Modify or paste the contents of your command, such as:

    ```
    echo hello ECS!
    echo root:NewPasswd9! | chpasswd
    echo Remmember your password!
    ```

    e. Determine the **execution path** of the command. The execution paths of Bat and PowerShell commands are by default set to the directory where the cloud assistant client is stored, such as `C:\ProgramData\aliyun\assist\$(version)`. Shell commands are by default in the `/root` directory.

    f. Set the maximum timeout time (in seconds) for commands in an instance. The default value is set to 3600s. When a command you created cannot be run for some reason, the

command times out. After the command times out, the command process will be forcibly terminated.

g. After confirming the command, click **OK**.

You can also use the ESC API *CreateCommand* to create a cloud assistant command.

Next step

Invoke commands

9.2 Run commands

After creating a cloud assistant command, you can run the command on one or more instances. The command execution status and results for each instance are not influenced by the same command being run on other instances. You can also configure the execution interval for the command.

Limits

- In one Alibaba Cloud region, you can run a maximum of 500 cloud assistant commands in a single day.

- You can run a command on a maximum of 50 instances at once.

- The status of the target instance must be **In Progress** (`Running`).

- The target instance must have *cloud assistant client* installed.

- The target instance 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 execute a command on the ECS console, take the following steps:

1. Log on to the *ECS console*.

2. In the left-side navigation pane, select **Cloud Assistant**.

3. Select a region.

4. Search for the Cloud Assistant command you want to run. Select **Execute** from the right-side **Actions**. In the right-side pop-up window:

   a. Click **View Command Content** to confirm the command contents.

   b. Click **Select Instance**. In the pop-up window:

      A. Select one or more instances.

      B. Click [>>] to select an instance.

      > 📋 **Note:**
      >
      > Bat or PowerShell commands can only be selected for Windows instances, and Shell commands can only be selected for Linux instances. All instances must have the Cloud

Assistant Client **installed**. Otherwise, the instance cannot be selected even after you

click the ⟫ icon.

C. Click **OK**.

c. Select **Immediate Execution** or **Scheduled Execution**:

- **Immediate Execution**: The cloud assistant will run the command immediately on the

  instances once.

- **Scheduled Execution**: Use the cron expression to run the command periodically. Fill in

  the **Execution Time**. For more information, see .
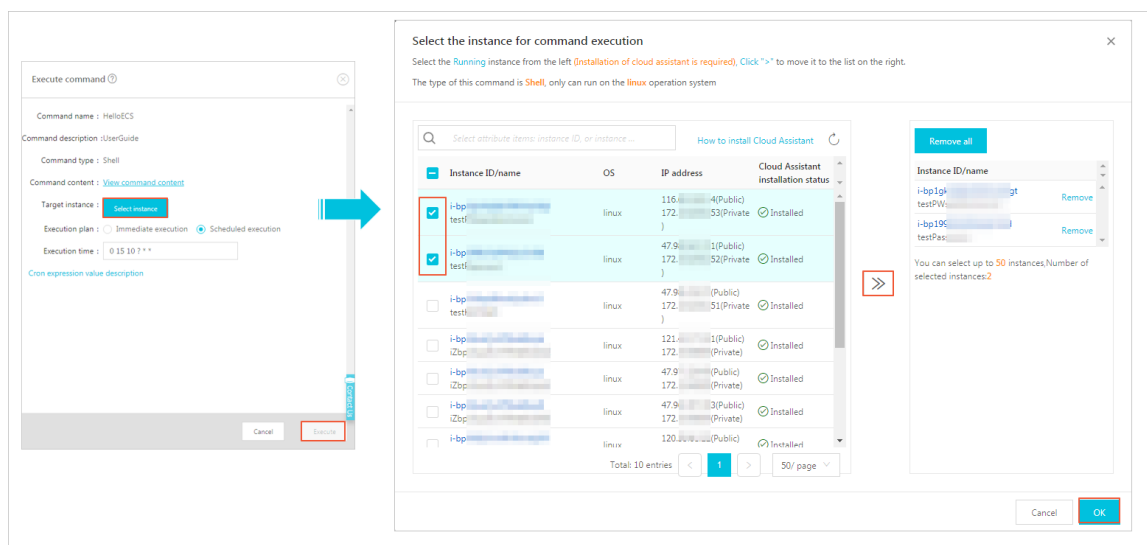
5. Click **Execute**.

You can also use the ECS API *InvokeCommand* to execute a cloud assistant command.

Stop command execution

Prerequisite: Either it must be a periodic command, or the command must have a command

execution status of **Running** (`Running`).

To stop a command on the management console, take the following steps:

1. Log on to the *ECS console*.

2. In the left-side navigation pane, select **Cloud Assistant**.

3. Select a region.

4. In the **Execution Record** area, search for the command you need to stop, and select **Stop**

Command from **Actions**.

Next step

Query execution results and status

9.3 Query execution results and status

There is no difference between running a cloud assistant command on the console and running a command while logged into the instance. In both cases, a command can be run successfully only after all of the command's conditions are satisfied. Cloud assistant commands executed at the same time can provide different command execution results and statuses if the following errors occur: lack of relevant dependencies, network disruptions, command semantic errors, script debugging errors, or abnormal instance statuses. We recommend that you review the command execution results and status after running a command to ensure the target operation has completed properly.

Prerequisites

The command must be run at least once.

Check the results of the command execution

To view command execution result on the ECS Console, you must take the following steps:

1. Log on to the *ECS console*.

2. In the left-side navigation pane, select **Cloud Assistant**.

3. Select a region.

4. In the **Execution Record** area, search for the execution record of the necessary command execution, and select **View Results** from **Actions**.

5. In the pop-up window, 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 command execution status in the ECS Console, you must take the following steps:

1. Log on to the *ECS console*.

2. In the left-side navigation pane, select **Cloud Assistant**.

3. Select a region.

4. In the **Execution Record** area, search for the execution record of the necessary command execution, and then in the **Execution Status** bar view the command execution status.

You can also use the ECS API *DescribeInvocations* to view command execution status.

Invocation status

- Specifically, the invocation status of a command consists of `Running`, `Stopped`, `Finished`, and `Failed`.

- Generally, the invocation status of a command includes **overall invocation status** , **instance invocation status** , and **invocation-record status**. The relationships among various levels are shown in the following figure.
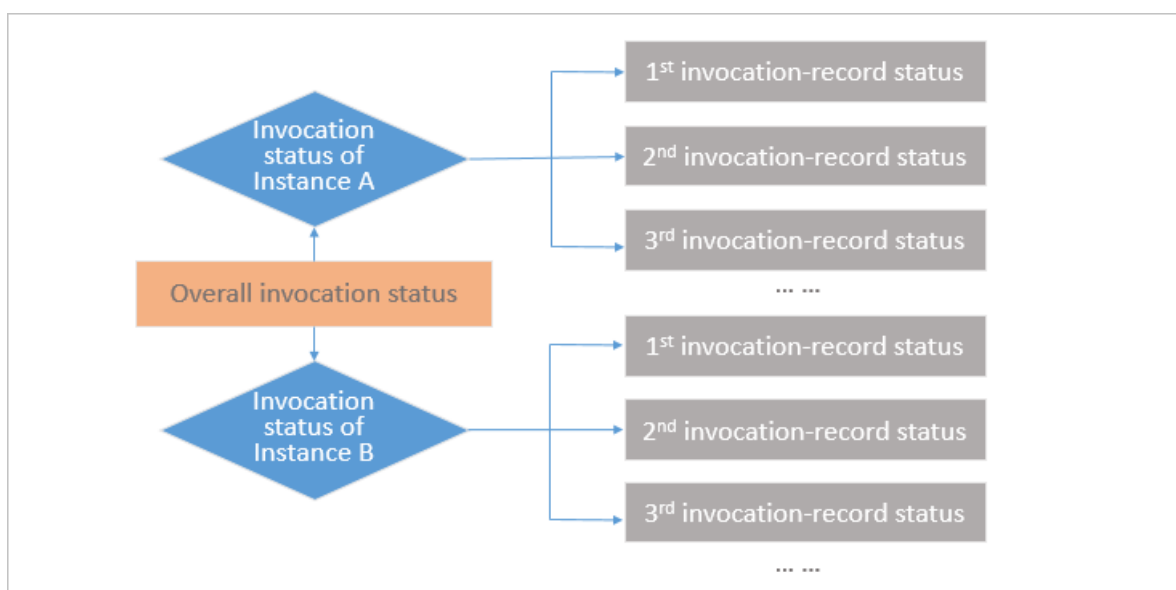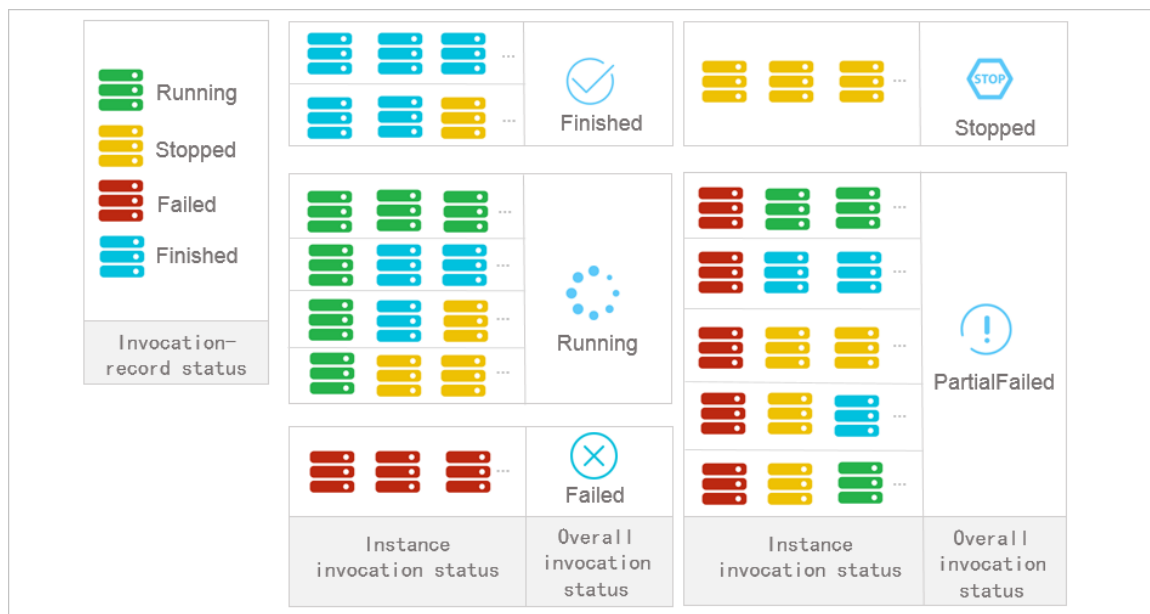

For one-time invocations

- **Overall invocation status**:

  — When the invocation status of all instances are `Finished`, the overall invocation status is displayed as `Finished`.

  — When the invocation status of some instances are `Finished` and those of some others are `Stopped`, the overall invocation status is displayed as. `Finished`

  — When the invocation status of all instances are `Failed`, the overall invocation status is displayed as `Failed`.

  — When the invocation status of all instances are `Stopped`, the overall invocation status is displayed as `Stopped`.

  — When the invocation statuses of all or some instances are `Running`, the overall invocation status is displayed as `Running`.

  — When the invocation statuses of some instances are `Failed`, the overall invocation status is displayed as `PartialFailed`.

  Take three ECS instances as an example. The following picture shows the relationships between the overall invocation status and the instance invocation status during a one-time invocation on multiple instances.

- **Instance invocation status**: The command is invoked only once in a one-time invocation, so the instance invocation status and the invocation-record status are identical.

- **Invocation-record status**:

  — `Running`: Indicates that the command is being executed.

  — `Stopped`: Indicates that the command invocation has been manually stopped by the user.

— `Finished`: Indicates that the command invocation has been completed smoothly. But invocation completion does not indicate invocation success. You can confirm whether the invocation is successful based on the actual `Output` of the command process.

— `Failed`: Indicates that the command process has timed out (`Timeout`) and failed.

For periodical invocations

- **Overall invocation status**: The overall invocation status is always `Running` unless you stop all the scheduled invocation for all instances.

- **Instance invocation status**: The instance invocation status is always `Running` unless you stop the current invocation.

- **Invocation-record status**:

    — `Running`: The command is being executed.

    — `Stopped`: You have stopped the command invocation.

    — `Finished`: The command invocation is complete. However, invocation completion does not guarantee invocation success. You can confirm whether the invocation is successful or not based on the actual `Output` of the command process.

    — `Failed`: The command process is timed out (`Timeout`) and fails.

9.4 Manage commands

After creating cloud assistant commands, you can set the command name and description, clone commands, or delete unnecessary commands to guarantee a sufficient command quota.

Modify the name and description of a command

To set the command name and description in 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 a region.

4. Move the mouse cursor to the command you want to edit, and click the ![icon] icon that appears in the prompted window.

    - **Command name**: Input the new command name.

    - **Command description**: Input the 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 add a new version for an existing cloud assistant command. You can retain all the information of the cloned command as it was previously. Alternatively, you can set a new name, description, type, content, execution path, or timeout time for it. To clone a command in 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 a region.

4. Find the cloud assistant command you want to clone, and from the **Operation** list, click **Clone**.

5. In the **Clone command** dialogue box, complete the following optional steps:

   a. Enter a new **Command name**, such as HelloECS.

   b. Enter a new **Command description**, such as UserGuide.

   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 select **Shell**.

   d. Edit or paste 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 set to 3600 . When a command you created cannot be executed for 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. After you confirm the modification, click **Create**.

Delete commands

Within an Alibaba Cloud region, you can create a maximum of 100 cloud assistant commands. We suggest that you regularly clean your commands to guarantee a sufficient command quota. 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 a region.

4. Locate the cloud assistant command you want to delete:

- To delete a single command, from the **Operation** list, select **Delete**.

- To delete multiple commands, select the target instances, and click **Delete command**.
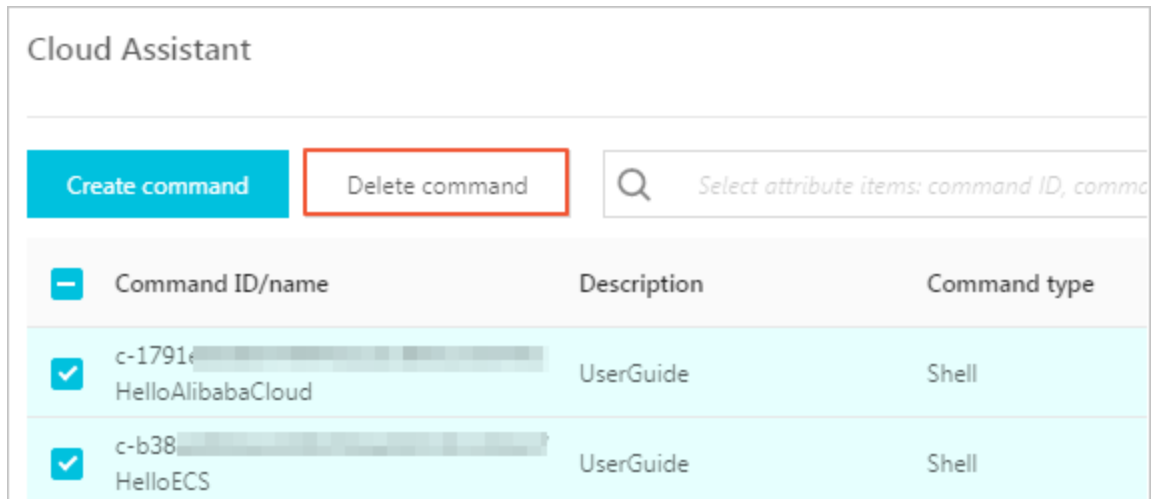
5. In the **Delete command** dialogue box, click **OK**.

You can also use the ECS API *DeleteCommand* to delete commands.

10 Monitoring

10.1 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.
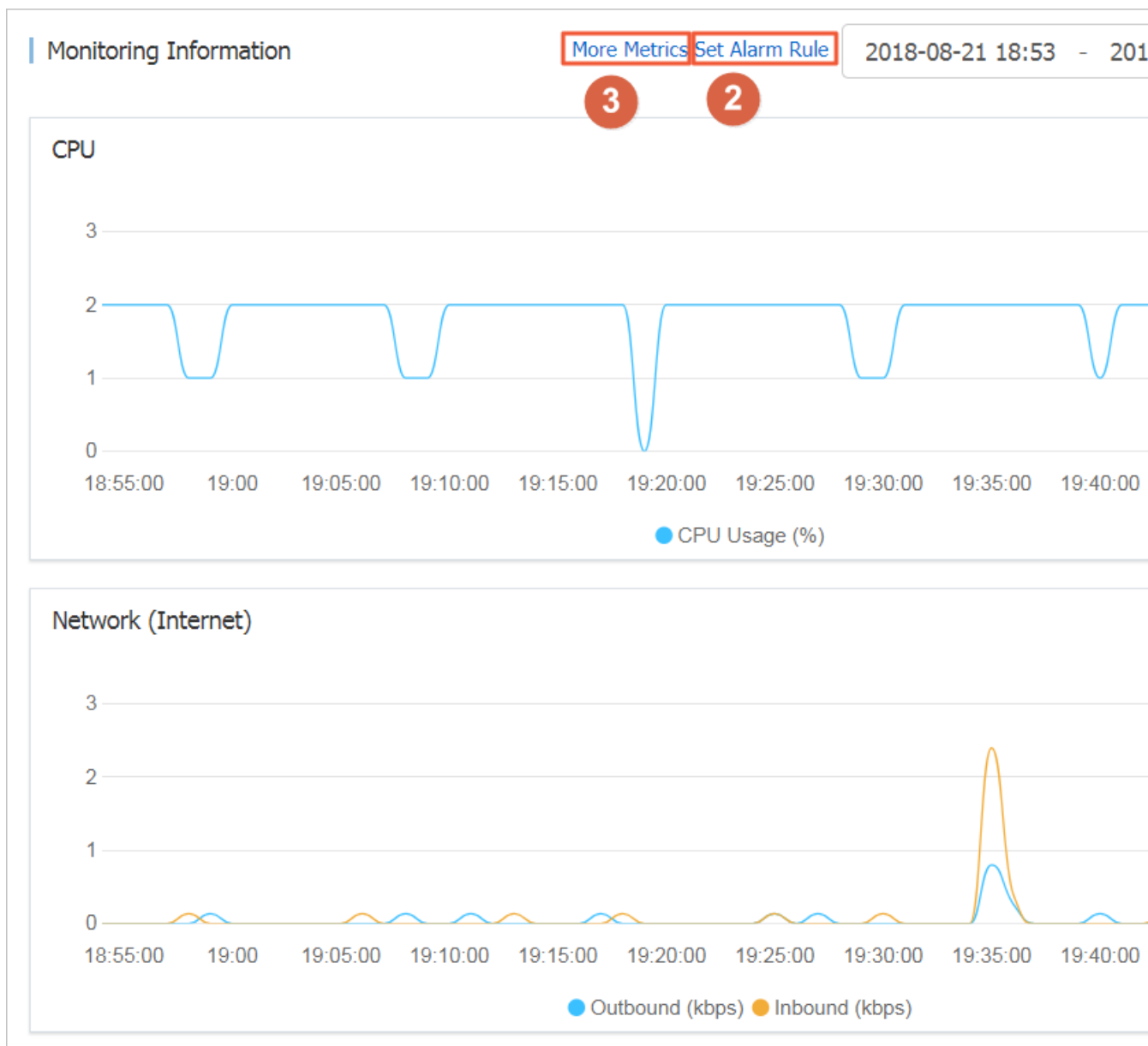
You can also call the ECS API operations *DescribeInstanceMonitorData*, *DescribeDiskMonitorData*, and*DescribeEniMonitorData* to obtain the monitoring data.

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** in the Actions column.

6. To set alarm rules, click **Alarm Rules** in the Actions column.
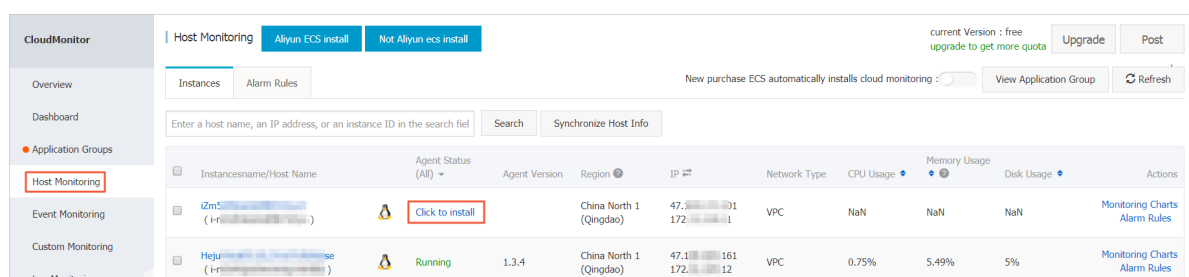
About bandwidth units

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.

10.2 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 c1, c2, m1, m2, s1, s2, s3, and t1 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 system event	An instance restarts after planned system maintenance or security update.	`SystemMaintenance.Reboot`
Unexpected system event	An instance restarts after unexpected system failures.	`SystemFailure.Reboot`
	An instance restarts after unexpected instance failures.	`InstanceFailure.Reboot`

Category	Event type	Parameter
Scheduled restart	An instance restarts after planned system maintenance or security update.	`SystemMaintenance.Reboot`
Unexpected restart	An instance restarts after unexpected system failures.	`SystemFailure.Reboot`
	An instance restarts after unexpected instance failures.	`InstanceFailure.Reboot`
Stop instances	Subscription instances stop due to expiration.	`InstanceExpiration.Stop`
	Pay-As-You-Go instances stop due to overdue payment.	`AccountUnbalanced.Stop`
Release instances	Subscription instances are released after several days of expiration.	`InstanceExpiration.Delete`

Category	Event type	Parameter
	Pay-As-You-Go instances are released due after several days of overdue payment.	AccountUnbalanced.Delete

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 *user operation period*.
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 7 to 15 days before a Subscription instances is stopped, and 1 hour before a Pay-As-You-Go instance is stopped. 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, instances are released 15 days later if no renewal or recharge are made. Later you receive the report of system events.
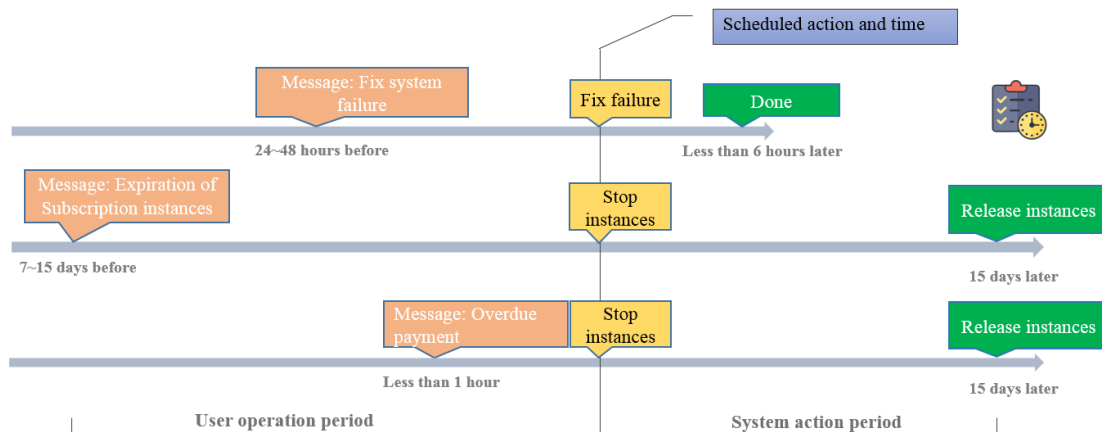
> **Note:**
>
> Only scheduled system events have user operation period. Unexpected system events that are caused by emergency failures or invalid operations do not have user operation periods. Once unexpected system events occur, you will receive notifications, but you cannot take any action. However, 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 **Unsettled events** 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 **Unsettled events** from the navigation pane on the right-side of the **Overview** page.

4. On the **Unsettled events** page, you can see the list of instance IDs, regions, and running status, system events, 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. Log on to the *ECS console*.

2. On the left-side navigation pane, select **Overview**.

3. Select **Unsettled events** from the navigation pane on the right-side of the **Overview** page.

4. Click **All events**, and on the **All events** page, click**Scheduled system event > >
Instances.**You can see the list of instance IDs, event types, and regions, and event status.

API operation: Call *#unique_225* 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 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: • *Restart the instance* in the ECS console. • Call API *RebootInstance*. 📋 **Note:** Instance restart performed in the instance or from the instance list has no effect on this type of system events. We recommend that you *Create snapshots* (*CreateSnapshot*) for the attached disks to back up your data.
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: • 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.

Event type	Parameter	Recommended
A Pay-As-You-Go instance stops due to overdue payment.	AccountUnb alanced.Stop	You can either *recharge your account* or wait for the instances to stop.
A Subscripti on instance is released due to expiration.	InstanceEx piration.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.	AccountUnb alanced.Delete	You can either *recharge your account* or wait for the instances to be released.

10.3 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 can 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 ouput.
- *Phased-out instance types* do not allow you to obtain instance console ouput or screenshots.
- You cannot obtain console ouput or screenshots for instances created before January 1, 2018.

Prerequisites

The instance must be in the **Running** (`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.
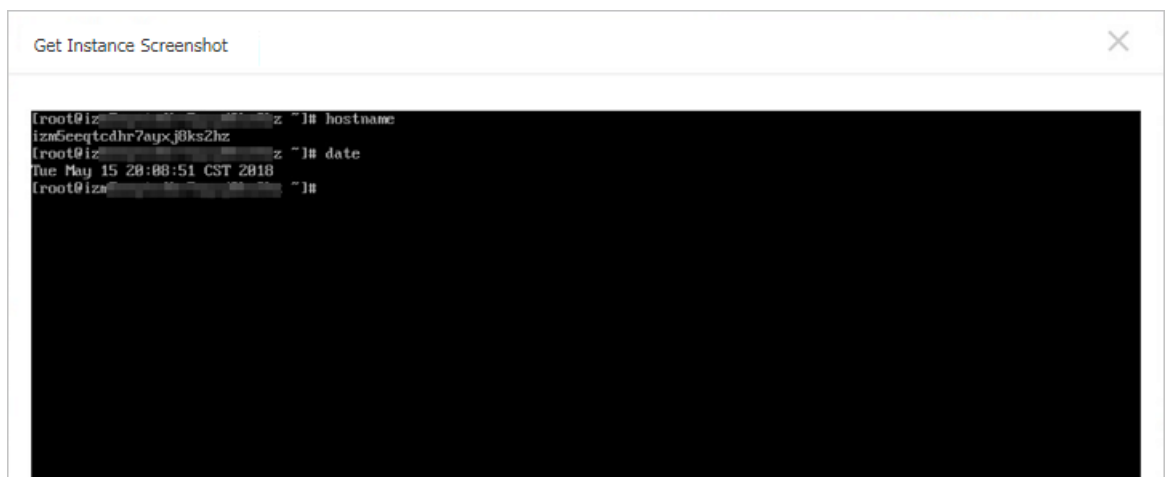
Operation in Instance Details page

1. Log on to the *ECS Management Console*.
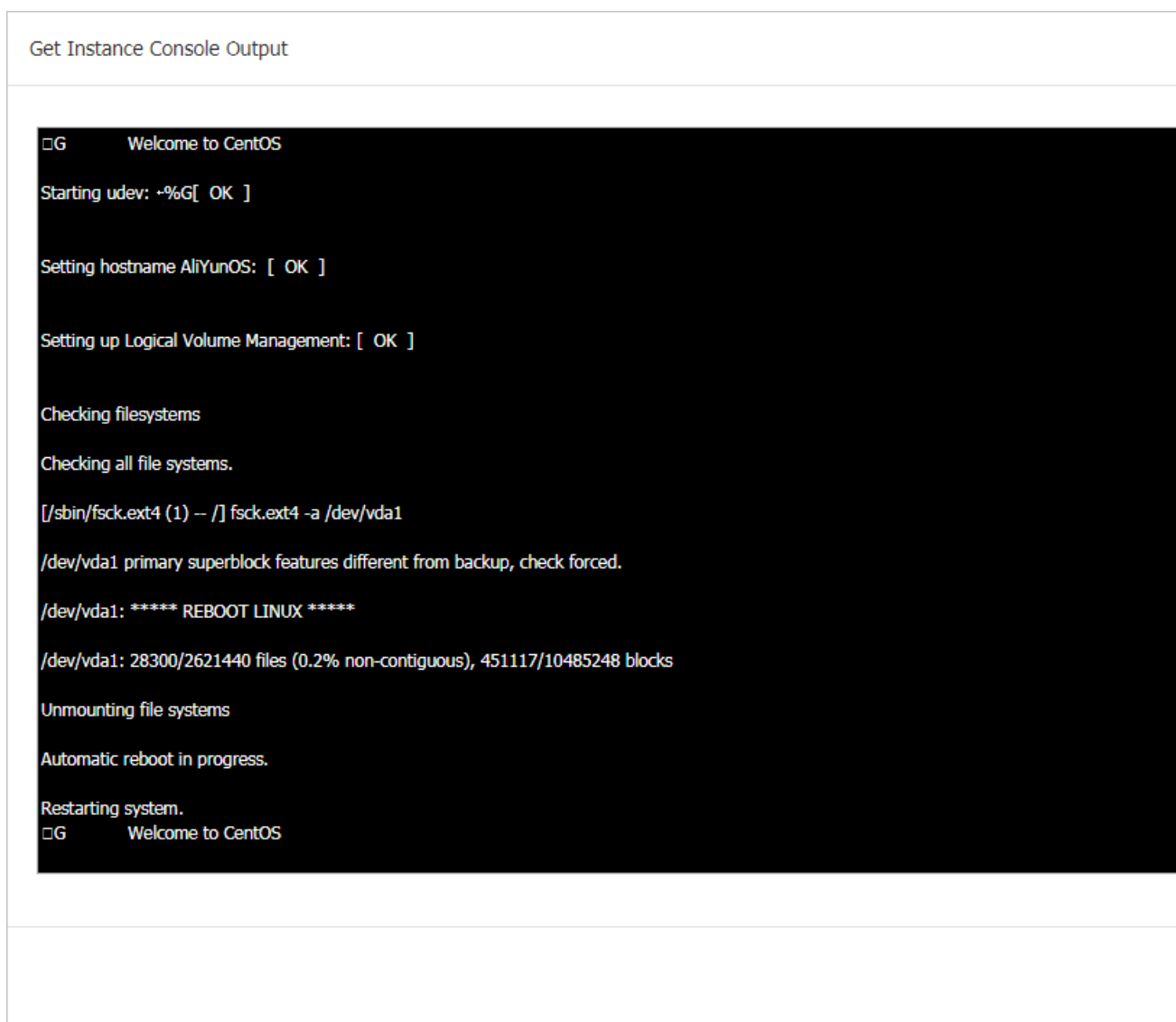2. In the left-side navigation pane, click **Instances**.

3. Select the **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:
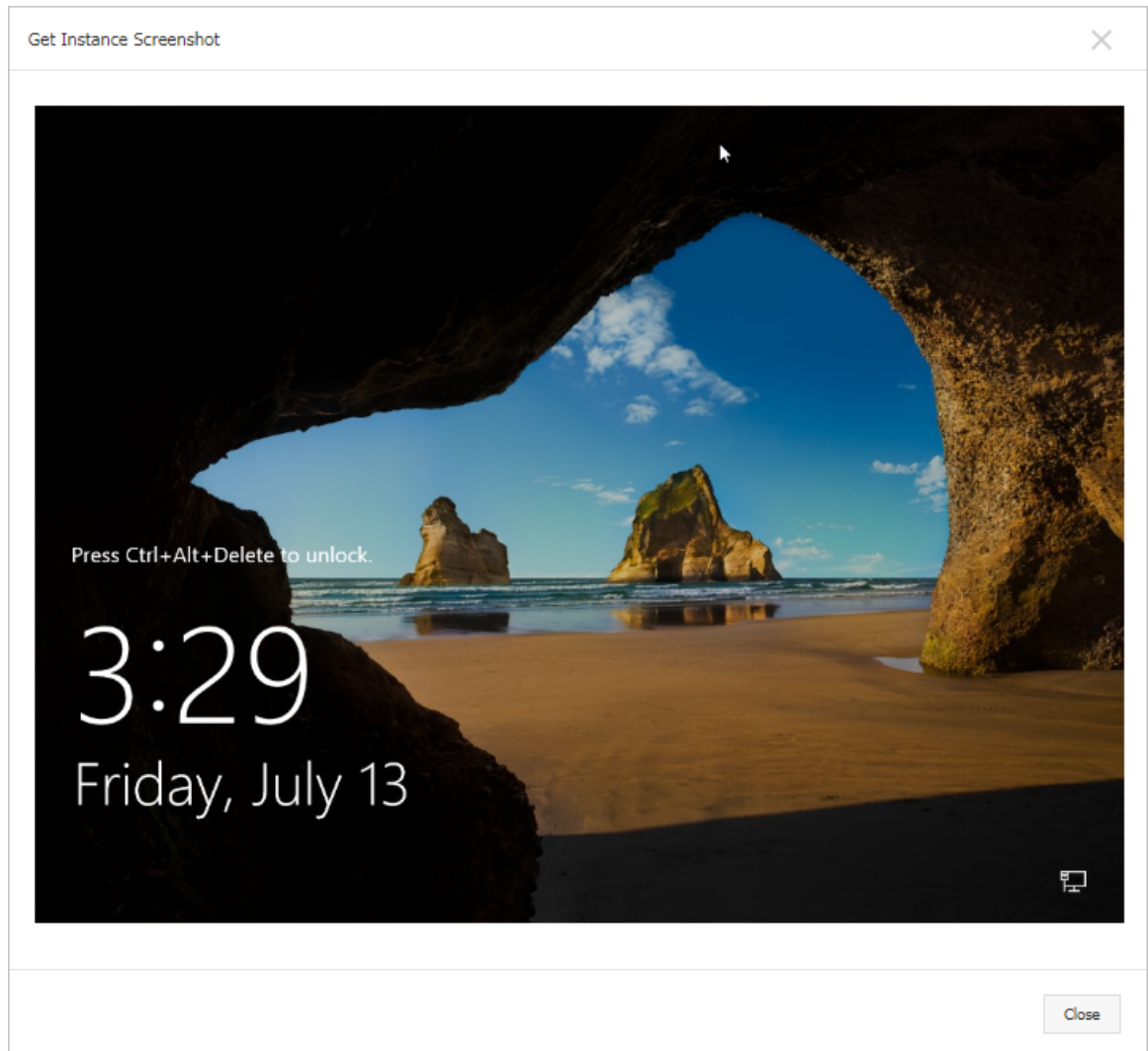
• Linux instance screenshot sample:

- Linux instance console output sample:

Operation in Instances list page

1. Log on to the *ECS Management Console*.

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

3. Select the **region**.

4. Locate the instance to troubleshoot and move to the **Actions** column.

5. 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*

Next step

For other troubleshooting instructions, see *Link testing tool for ping packet loss or ping failure*.