阿里云
云服务器 ECS

用户指南

文档版本：20180907

法律声明

阿里云提醒您在阅读或使用本文档之前仔细阅读、充分理解本法律声明各条款的内容。如果您阅读或使用本文档，您的阅读或使用行为将被视为对本声明全部内容的认可。

1. 您应当通过阿里云网站或阿里云提供的其他授权通道下载、获取本文档，且仅能用于自身的合法合规的业务活动。本文档的内容视为阿里云的保密信息，您应当严格遵守保密义务；未经阿里云事先书面同意，您不得向任何第三方披露本手册内容或提供给任何第三方使用。
2. 未经阿里云事先书面许可，任何单位、公司或个人不得擅自摘抄、翻译、复制本文档内容的部分或全部，不得以任何方式或途径进行传播和宣传。
3. 由于产品版本升级、调整或其他原因，本文档内容有可能变更。阿里云保留在没有任何通知或者提示下对本文档的内容进行修改的权利，并在阿里云授权通道中不时发布更新后的用户文档。您应当实时关注用户文档的版本变更并通过阿里云授权渠道下载、获取最新版的用户文档。
4. 本文档仅作为用户使用阿里云产品及服务的参考性指引，阿里云以产品及服务的”现状“、“有缺陷”和“当前功能”的状态提供本文档。阿里云在现有技术的基础上尽最大努力提供相应的介绍及操作指引，但阿里云在此明确声明对本文档内容的准确性、完整性、适用性、可靠性等不作任何明示或暗示的保证。任何单位、公司或个人因为下载、使用或信赖本文档而发生任何差错或经济损失的，阿里云不承担任何法律责任。在任何情况下，阿里云均不对任何间接性、后果性、惩戒性、偶然性、特殊性或刑罚性的损害，包括用户使用或信赖本文档而遭受的利润损失，承担责任（即使阿里云已被告知该等损失的可能性）。
5. 阿里云网站上所有内容，包括但不限于著作、产品、图片、档案、资讯、资料、网站架构、网站画面的安排、网页设计，均由阿里云和/或其关联公司依法拥有其知识产权，包括但不限于商标权、专利权、著作权、商业秘密等。非经阿里云和/或其关联公司书面同意，任何人不得擅自使用、修改、复制、公开传播、改变、散布、发行或公开发表阿里云网站、产品程序或内容。此外，未经阿里云事先书面同意，任何人不得为了任何营销、广告、促销或其他目的使用、公布或复制阿里云的名称（包括但不限于单独为或以组合形式包含”阿里云”、Aliyun”、“万网”等阿里云和/或其关联公司品牌，上述品牌的附属标志及图案或任何类似公司名称、商号、商标、产品或服务名称、域名、图案标示、标志、标识或通过特定描述使第三方能够识别阿里云和/或其关联公司）。
6. 如若发现本文档存在任何错误，请与阿里云取得直接联系。

通用约定

格式	说明	样例
	该类警示信息将导致系统重大变更甚至故障，或者导致人身伤害等结果。	 **禁止：** 重置操作将丢失用户配置数据。
	该类警示信息可能导致系统重大变更甚至故障，或者导致人身伤害等结果。	 **警告：** 重启操作将导致业务中断，恢复业务所需时间约10分钟。
	用于补充说明、最佳实践、窍门等，不是用户必须了解的内容。	 **说明：** 您也可以通过按**Ctrl** + **A**选中全部文件。
>	多级菜单递进。	设置 > 网络 > 设置网络类型
粗体	表示按键、菜单、页面名称等UI元素。	单击 **确定**。
`courier` 字体	命令。	执行 `cd /d C:/windows` 命令，进入Windows系统文件夹。
斜体	表示参数、变量。	`bae log list --instanceid` *`Instance_ID`*
[]或者[a|b]	表示可选项，至多选择一个。	`ipconfig` *`[-all|-t]`*
{}或者{a|b}	表示必选项，至多选择一个。	`swich` *`{stand | slave}`*

目录

1 Elastic Network Interfaces

1.1 Attach an ENI when creating an instance

You can attach an ENI when creating an ECS instance in the ECS console. For more information about instance creation, see *create an instance*. This document focuses on the notes for attaching an ENI during ECS instance creation.

Note the following configurations when attaching an ENI during ECS instance creation:

- Basic configurations:
  - Region: ENIs are supported in all regions.
  - Instance type: Select an instance type that supports ENI. The selected instance type must be I/O optimized.
  - Image: Only the following types of image can automatically recognize ENIs without any additional configuration. For other images, you must configure the ENI to enable the created instance to recognize it.
    - Centos 7.3 64-bit
    - Centos 6.8 64-bit
    - Windows Server 2016 Data Center Edition 64-bit
    - Windows Server 2012 R2 Data Center Edition 64-bit
  - Networking:
    - Network: Select **VPC**, and then select a created VPC and a VSwitch.
    - ENI: Click **Add ENI** to attach an ENI, and then select a VSwitch for the ENI.

      Note:

      - You are only allowed to attach a maximum of two ENIs when creating an instance in the console. One of them is the primary ENI, which is attached automatically, and the other one is a secondary ENI.
      - After the instance is started, you can attach more secondary ENIs to the instance based on the instance type in the console or by using the *AttachNetworkInterface* API.

If you want to keep the secondary ENI that is created in this way, detach it from the instance before you release the instance.

1.2 Create an ENI

You can create an ENI in two ways.

- *Attach an ENI when creating an instance*. In this way, you can attach a maximum of 2 elastic network cards, one of which is the main network card, one is an auxiliary network card. A secondary network card created in this way, if it is not *detached* from the instance, releases along with the instance.
- Create a separate elastic network card. Create a good elastic network card that can be *attached* to an instance. The elastic network card created by this method can only be used as a secondary network card.

This document describes how to create an ENI in the ECS console.

Limits

To create an ENI, you have the following limits:

- Each ENI must be in a VSwitch of a VPC.
- Each ENI must be in one security group.

Prerequisites

Before you create an ENI, finish the following operations:

- Create a VPC and then create a VSwitch in the VPC.
- Create a security group in the same VPC.

Procedure

To create an ENI, follow these steps:

1. Log on to the *ECS console*.
2. In the left-side navigation pane, select**Networks & Security** > **Network Interfaces**.
3. Select a region.
4. Click **Create**.
5. In the **Create** dialog box, finish the following configurations:
   a. Network Interface Name: Specify a name for the ENI.
   b. VPC: Select a VPC. When you attach an ENI to an instance, they must be in the same VPC.

> **Note:**
> In addition, after an ENI is created, you cannot change the VPC.

c. **VSwitch**: Select a VSwitch. When you attach an ENI to an instance, they must be in the same zone, but they do not have to be in the same VSwitch.

> **Note:**
> In addition, after an ENI is created, you cannot change the VSwitch.

d. IP: Specify an IPv4 address as the private IP address of the ENI. The IPv4 address must be available in the CIDR block of the specified VSwitch. If you do not specify one, an private IP address is automatically assigned to your ENI after the ENI is created.

e. SecurityGroup: Select a security group in the selected VPC.

f. Description: Give a brief description for the ENI for easing further management.

g. Click **OK**.

In the Network Interfaces page, refresh the table. When the new ENI is in the **`Available`** status, it is created successfully.

Follow-up operations

After you create an ENI, perform the following operations:

- *Attaching an ENI to an instance*.
- *Modifying attributes of the ENI*.
- *Deleting the ENI*.

1.3 Attach an ENI to an instance

You can attach an ENI to an instance.

This document describes how to attach an ENI to an instance in the ECS console.

Limits

To attach an ENI to an instance, you have the following limits:

- You can only attach a secondary ENI to an instance.
- The ENI must be in the **`Available`** status.
- The instance must be in the **`Stopped`** or **`Running`** status.
- You can only attach an ENI to a VPC-Connected ECS instance, and they must be in the same VPC.
- The VSwitches of the ENI and the instance can be different, but they must be in the same zone .

- An ENI can be attached to an I/O optimized ECS instance only.
- An ENI can only be attached to one VPC-Connected ECS instance at a time. However, a VPC-Connected ECS instance can be associated with multiple ENIs. For more information about the maximum number of ENIs that can be attached to one instance, see *Instance type families*.

Prerequisites

Before you attach an ENI to an instance, finish the following operations:

- *Create an ENI*.
- Make sure the ENI is in the **`Available`** status.
- Make sure your instance can be associated with secondary ENIs, and the instance is in the **`Stopped`** or **`Running`** status. For the number of ENIs that can be attached to each instance type, see the *Instance type families*.

Procedure

To attach an ENI to an instance, follow these steps:

1. Log on to the *ECS Management Console*.
2. In the left-side navigation pane, select**Networks & Security >** > **ENI**.
3. Select a region.
4. Find an available ENI, and in the Operations column, click **Attach**.
5. In the **Bind instance** dialog box, after selecting an instance, click **OK**.

In the Network Interfaces page, refresh the table. When the selected ENI is in the **`InUse`** status, it is successfully attached to the instance.

Follow-up operations

After an ENI is attached to an instance, you can perform the following operations:

- *Detaching the ENI from an instance*, and then *Deleting the ENI*
- *Modifying attributes of the ENI*
- *Configuring the ENI*, if the ENI cannot be automatically recognized by the operating system of your instance.

1.4 Detach an ENI from an instance

You can detach a secondary ENI, but not the primary ENI, from an instance.

Limits

To detach a secondary ENI from an instance, you have the following limits:

- The secondary ENI must be in the **InUse** status.
- The instance must be in the **Stopped** or **Running** status.

Prerequisites

Your ENI *is attached to an instance*. Before you detach an ENI from an instance, the instance must be in the **Stopped** or **Running** status.

Procedure

To detach a secondary ENI from an instance, follow these steps:

1. Log on to the *ECS console*.
2. In the left-side navigation pane, select**Networks & Security** > **Network Interfaces**.
3. Select a region.
4. Find an ENI in the InUse status, and in the Actions column, click **Detach**.
5. In the **Detach** dialog box, confirm the information, and then click **OK**.

In the Network Interfaces page, refresh the table. When the selected ENI is in the **Available** status, it is successfully detached from the instance.

Follow-up operations

After an ENI is detached from an instance, you can perform these operations:

- *Attaching the ENI to another instance*.
- *Deleting the ENI*.
- *Modifying attributes of the ENI*.

1.5 Modify attributes of an ENI

You can modify attributes of a secondary ENI, but not the primary ENI, of an instance. You can modify the following attributes of an ENI:

- The name of the ENI.
- The security group associated with the ENI. One ENI must be associated with at least one security group. However, it cannot be associated with more than five security groups.
- Description of the ENI.

You can modify attributes of an ENI when it is in the **`Available`** or the **`InUse`** status. This document describes how to modify attributes of an ENI in the ECS console.

Prerequisites

Before you modify attributes of an ENI, *create an ENI*.

Procedure

To modify attributes of an ENI, follow these steps:

1. Log on to the *ECS console*.
2. In the left-side navigation pane, select**Networks & Security** > **Network Interfaces**.
3. Select a region.
4. Find an ENI, and in the Actions column, click **Modify**.
5. In the **Modify** dialog box, modify the following optional configurations:
   - Network Interface Name: Specify a new name for the selected ENI.
   - SecurityGroup: Select more security groups for the ENI, or remove security groups. Retain at least one security group.
   - Description: Give a brief description for the ENI.

   After you finish the modification, click **OK**.

1.6 Delete an ENI

If you do not require an ENI, you can delete it. But you can only delete a secondary ENI, but not the primary ENI of an instance.

After an ENI is deleted,

- the primary private IP address of the ENI is released automatically,
- and the ENI is automatically removed from all associated security groups.

An ENI will be deleted along with an instance if you did not detach it from the instance before you release the instance.

Limits

You can only delete an ENI in the **`Available`** status.

Prerequisites

If an ENI is *attached to an instance*, *detach it from the instance*.

Procedure

To delete an ENI, follow these steps:

1. Log on to the *ECS console*.
2. In the left-side navigation pane, select**Networks & Security** > **Network Interfaces**.
3. Select a region.
4. Find an available ENI, and in the Operations column, click **Delete**.
5. In the dialog box, click **OK**.

In the Network Interfaces page, refresh the table. If the ENI disappears, it is deleted successfully.

1.7 Configure an ENI

If your instance is running one of the following images, you do not have to configure the Elastic Network Interfaces (ENI) manually to have them recognized by the OS.

- Centos 7.3 64-bit
- Centos 6.8 64-bit
- 64-bit Windows Server 2016 data center Edition
- Windows Server 2012 R2 Data Center Edition 64-bit64-bit Windows Server 2012 R2 data center Edition

If your instance is running none of the preceding images, and you want to attach an ENI to your instance, you must manually configure the ENI to be recognizable.If your instance does not use these images, however, if you want to attach a flexible network card to an instance, You need to manually configure the elastic network card. This document uses an instance running CentOS 7.2 64-bit as an example to introduce how to configure an ENI to make the interface recognizable.

Prerequisite

You have attached an elastic network card to an ECS instance.

Procedure

To configure the ENI, follow these steps:

1. Use the *DescribeNetworkInterfaces* interface or log on to the ECS console to obtain the following attributes of the ENI: primary private IP address, subnet mask, the default route, and the MAC address. To obtain these attributes in the ECS console, follow these steps: MAC address. Do the following on the console.

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

b. Find a network interface, and obtain its primary private IP address, subnet mask, default route, and MAC address.Locate the primary private IP address, mask address, default route, MAC for each network cardAddress. Example

```
eth1 10.0.0.20/24 10.0.0.253 00: 16: 12: E7: 27
eth2 10.0.0.21/24 10.0.0.253 00: 16: 12: 16: EC
```

2. *Connect to the ECS Instance*.

3. Run the command to generate the confi`cat /etc/sysconfig/network-scripts/ifcfg-[network interface name in the OS]`.

Note:

- Pay attention to the relation between network interface name in the OS and the MAC address.
- Pay attention to the relation between network interface name in the OS and the MAC address. The default route must be set to `DEFROUTE=no` . Other editions must have the same configuration. Note that running the `ifup` command may change the active default route configuration after configuring the network interface.
- Example:

```
# cat /etc/sysconfig/network-scripts/ifcfg-eth1
DEVICE=eth1
BOOTPROTO=dhcp
ONBOOT=yes
TYPE=Ethernet
USERCTL=yes
PEERDNS=no
IPV6INIT = No
PERSISTENT_DHCLIENT = Yes
HWADDR=00:16:3e:12:e7:27
DEFROUTE=noDefroute = No
```

4. Follow these steps to start the network interface:

a. Run the `ifup [network interface name in the OS]` command to start the dhclient process, and initiate a DHCP request. Example

```
# ifup eth1
```

```
# ifup eth2
```

b. After a response is received, run the `ip a` a command to check the IP allocation on the network interfaces, which must match with the information displayed on the ECS console. Example:

```
# ip a
1: lo: mtu 65536 qdisc noqueue state UNKNOWN qlen 1
link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
inet 127.0.0.1/8 scope host loInet 125.0.0.1/8 Scope host Lo
valid_lft forever preferred_lft forever
2: eth0: mtu 1500 qdisc pfifo_fast state UP qlen 10002: eth0: MTU
1500 qdisc glasstate up qlen 1000
link/ether 00:16:3e:0e:16:21 brd ff:ff:ff:ff:ff:ff
Inet 10.0.0.19/24 BRD glasscope Global Dynamic eth0
valid_lft 31506157sec preferred_lft 31506157secValid_lft
31506157sec preferred_lft 31506157sec
3: eth1: MTU 1500 qdisc glasstate up qlen 1000
link/ether 00:16:3e:12:e7:27 brd ff:ff:ff:ff:ff:ff
inet 10.0.0.20/24 brd 10.0.0.255 scope global dynamic eth1Inet 10.
0.0.20/24 BRD glasscope Global Dynamic eth1
Valid_lft 31525994sec preferred_lft 31525994sec
4: eth2: MTU 1500 qdisc glasstate up qlen 1000
Link/ether 00: 16: Rye: 12: 16: ec brd ff: FF: FF
inet 10.0.0.21/24 brd 10.0.0.255 scope global dynamic eth2
valid_lft 31526009sec preferred_lft 31526009sec
```

5. Set the metric for each network interface in the route table. In this example, set the metric parameters of eth1 and eth2 as follows.

```
eth1: gw: 10.0.0.253 metric: 1001
eth2: gw: 10.0.0.253 metric: 1002
```

a. Run the following command to set the metric parameters.

```
# Ip-4 route add default via glasdev eth1 metric 1001
# ip -4 route add default via 10.0.0.253 dev eth2 metric 1002
```

b. Run the`route -n` command to check whether the configuration is successful or not. Example:

```
# route -n
Kernel IP routing table
Destination Gateway Genmask Flags Metric Ref Use Iface
0.0.0.0 10.0.0.253 0.0.0.0 UG 0 0 0 eth0
0.0.0.0 10.0.0.253 0.0.0.0 UG 1001 0 0 eth1
0.5.0.0 10.0.0.253 ug ub1002 0 0 eth2
10.0.0.0 0.5.0.0 255.25.25.0 u 0 0 0 eth0
10.0.0.0 0.0.0.0 255.255.255.0 U 0 0 0 eth1
10.0.0.0 0.5.0.0 255.25.25.0 u 0 0 0 eth2
169.254.0.0 0.0.0 255.0.0 U 1002 0 0 eth0
169.254.0.0 0.0.0.0 255.255.0.0 U 1003 0 0 eth1
```

```
169.254.0.0 0.0.0.0 255.255.0.0 U 1004 0 0 eth2169.254.0.0 0.0.0
255.0.0 U 1004 0 0 eth2
```

6. Follow these steps to build a route table:

Note:

We recommend that you use the metric value as the route table name.

   a. Run the command to build a route table.

```
# ip -4 route add default via 10.0.0.253 dev eth1 table 1001
# Ip-4 route add default via glasdev eth2 table 1002
```

   b. Run the command to check whether the route table is built successfully or not.

```
# ip route list table 1001
default via 10.0.0.253 dev eth1
# ip route list table 1002
default via 10.0.0.253 dev eth2
```

7. Configure policy routing.

   a. Run the following command .

```
# ip -4 rule add from 10.0.0.20 lookup 1001
# ip -4 rule add from 10.0.0.21 lookup 1002
```

   b. 运行命令 `ip rule list` View routing rules.

```
# ip rule list
0: from all lookup local
32764: from 10.0.0.21 lookup 1002
32765: from 10.0.0.20 lookup 1001
32766: from all lookup main
32767: from all lookup default
```

At this point, you have completed the configuration of the elastic network card.

2 Tags

2.1 Limits

You can bind tags to the following resources in the ECS console: ECS instance, storage, snapshot , image, and security group.

Tags have the following limits:

- Each tag has a key-value pair.
- You can bind 20 tags to an instance at most. You can bind 20 tags at most to an instance at a time.
- Every tag key of a resource must be unique. A tag with the same key as an existing one will be overwritten.
- Tag information is not shared across regions. For example, tags created in China East 1 (Hangzhou) are invisible to China East 2 (Shanghai).
- If a tag is unbound and no longer bound to any other resource, the tag will be automatically deleted.

2.2 Add a tag to resources

If your account maintains various types of resources that are associated with each other in different ways, you can bind tags to the resources to categorize and manage the resources in a unified manner.

You can bind 20 tags to a resource at most. You can bind/unbind 20 tags at most for the resource each time.

Take the following steps to bind resources with tags:

1. Log on to the*ECS console*.
2. Select the resource type in the left-side navigation bar for the binding operation, such as Instance, Cloud Disks, Snapshot, Image, and Security Groups.
3. Select a region.
4. Select the resources in the resource list to bind tags.
5. Click Edit Tags at the bottom of the resource list.

Note:

Choose **More** > **Edit Tags at the bottom of the resource list if the selected resources are Instance.** .

6. In the edit label dialog box,
   - If the selected resource has already been created a label, click the existing label, and select the available labels.
   - Click **Create** and set Key and Value if no tags are available for the selected resource: Note when entering:
     - Key is mandatory whereas Value is optional.
     - Key cannot start with aliyun, http://, or https://. The key is case-insensitive and can contain up to 64 characters.
     - Value cannot start with http:// or https://. The value is case-insensitive and can contain up to 128 characters. It can be empty.
     - Any tag Key of a resource must be unique. A tag with the same key as an existing one will be overwritten.
     - Available Tags and Create are grayed out if the selected resources are already bound with 20 tags. You need to unbind some tags before binding new tags.
7. Click **Confirm**.

To check if tags are successfully bound, use the Edit Tags function of the resource or click Tags in the left-side navigation bar of the ECS console. You can click Tags with a tag symbol at the top of the resource list to filter resources.

2.3 Delete a tag

You can unbind a tag from the resource if the tag is no longer applicable to resource management . After a tag is unbound and is no longer bound to any other resource, the tag will be automatically deleted.

- The Delete Tags function unbinds one or more tags from an instance at a time.

  Note:
  Currently, this function is only available for instances. It is unavailable for other resource types.

- The Edit Tags function unbinds tags one by one.

  Note:
  You can unbind 20 tags from a resource each time.

Unbind tags from instances using the tag deletion function

Currently, the Delete Tags function is only available for instances.

See the following steps to delete tags:

1. Log on to the *ECS console*.
2. Click Instance in the left-side navigation pane.
3. Select a region.
4. Select the instance(s) from which you want to unbind tags in the instance list.

> **Note:**
> You can also filter instances by tag and select the expected instance.

5. Choose **More** > **Delete Tags**.
6. In the **Delete Tags** dialog box, enter the **Tag Key** of the tags you want to unbind.
7. Click **OK** to complete tag unbinding.

To check whether the tags are successfully unbound, use the Edit Tags function of the instance or click Tags in the left-side navigation pane of the ECS console.

Unbind tags from resources using the tag edit function

The Edit Tags function unbinds one or more tags from a resource.

See the following steps to unbind tags:

1. Log on to the *ECS console*.
2. In the left-side navigation pane, select the resource type for the unbinding operation, such as Instance, Cloud Disks, Snapshots, Images, or Security Groups.

> **Note:**
> The block storage function is now in beta. For more information, see *block storage FAQ* Learn more.

3. Select a region.
4. In the resource list, select the resource from which you want to unbind tags.

> **Note:**
> You can also filter resources by tag and select the expected resource.

5. Click **Edit Tags** at the bottom of the resource list.

6. In the **Edit Tags** dialog box, click the deletion icon next to a tag.
7. Click **Confirm** to complete tag unbinding.

To check whether the tags are successfully unbound, use the Edit Tags function of the resource or click **Tags** in the left-side navigation pane of the ECS console.

2.4 Filter resources by tags

After you bind the tags to the resources, use any of the following methods to filter resources by tags.

Filter resources in resource lists

See the following steps to filter resources:

1. Log on to *ECS console*.
2. In the left-side navigation pane, select the resource type you want to view, such as Instances, Cloud Disks, Snapshots, Images, or Security Groups.
3. Select a region.
4. Click **Tag** at the top of the resource list.
   - Click a key to filter out the resources that are bound with this key, which may have multiple values.
   - Click a key and value to filter out the resources that are bound with this key-value pair (tag).

The console returns a list of resources that are bound with the key or a key-value pair.

Filter resources by tags

See the following steps to filter resources:

1. Log on to *ECS console*.
2. Click **Tags** in the left-side navigation pane.
3. Select a region.
4. Enter a key in the search box and click **Search**.

The console returns a list of resources that are bound with the key.

3 Access Control Ram

If you purchased multiple cloud server ECs instances, there are multiple users in your organizati on who need to use these instances. If these users share your cloud account key, the following issues exist:

- Your key is shared by multiple people and the risk of leakage is high;
- You cannot restrict the user's access rights, and it is a security risk that misoperations can occur.

Access Control RAM (Resource Access Management) is a resource access control service provided by Ali cloud. Through Ram, which allows you to centrally manage your users, such as employees, systems, or applications), and the permissions that control which resources in your name can be accessed by the user.

Access Control Ram helps you manage user access control over resources. For example, to strengthen network security control, you can attach an authorization policy to a group, policy: if the user's raw IP If the address does not come from an enterprise network, it denies such users from requesting access to the ECS resources in your name.

You can set different permissions for different groups, such:

- Sysadmins: This group needs to create and manage the ECS mirror, instance, snapshot, security group, and so on. You have attached an authorization policy to the sysadmins group that grants the group members to execute all ECs The permission for the operation.
- Developers: The group only needs to use the permissions of the instance. You can attach an authorization policy to the developers group that grants group membership calls Description instances, startinstance, stopinstance, createinstance, and deleteinstance Permissions.
- If a developer's job responsibilities change and become a system manager, you can easily move it from the developers group to the sysadmins group.

For more information on access control Ram, refer to the product documentation for Ram.

4 Monitoring

4.1 监控

您可以從多個維度監控Elastic Compute Service 執行個體的健全狀態，確保執行個體的暢通運行。

您可以從以下兩個入口監控 ECS 執行個體的運行情況：

- 執行個體詳情
- Cloud Monitor

執行個體詳情

1. 登入 *ECS管理主控台*。
2. 單擊左側導覽列中的 執行個體。然後選擇頁面頂部的地域。
3. 找到要監控的執行個體，單擊該執行個體名稱。
4. 在 執行個體詳情 頁面，您可以看到監控資訊，包括 CPU 使用率和網路的出網和入網情況。
5. 您可以在頁面右側、監控圖的右上方設定要查看的時間段。

> 说明：
> 由於在展示的時候彙總操作不一樣，比如 5 分鐘和 15 分鐘的平均值就會顯示不一樣的結果，所以選擇時間段的長短會影響顯示的精度。選擇時間範圍越小，顯示結果越精細。

下面是對監控資訊的解析：

- CPU：顯示的監控資料是伺服器 CPU 使用的百分比。百分比越高，說明伺服器 CPU 的負載越高。
  - Windows 執行個體可以在伺服器上用工作管理員來查看 CPU 的使用方式，按 CPU 排序，可以找出是什麼程式佔用伺服器的CPU資源。
  - Linux 執行個體可以通過 top 命令查看 CPU 的使用方式。登入伺服器，在命令列執行命令 top，然後鍵盤輸入shift+p，讓 top 按 CPU 使用大小排序，可以查看目前佔用 CPU 的進程。
- 網路：顯示的監控資料是伺服器的公網流量，單位是 kbps，1MB 頻寬=1024 kbps。監控資料可以查看出伺服器和入伺服器兩個方向的流量。1 MB 的頻寬，如果出網流量達到 1024 kbps，頻寬基本就是跑滿了。

Cloud Monitor

1. 在管理主控台，開啟 產品與服務 > **Cloud Monitor**；或者在執行個體的 執行個體詳情 頁面，單擊 設定報警規則。
2. 單擊左側導覽列中的 **Elastic Compute Service**，然後選擇項要監控的執行個體名稱。
3. 單擊 點擊安裝，您可以監控執行個體的作業系統。單擊 監控圖表，可以查看各種基礎參數；單擊 報警規則，可以設定報警規則。

更多關於Cloud Monitor的詳細資料，請參考Cloud Monitor的產品文檔。

背景知識：頻寬單位的換算

首先解釋 Kb 和 KB 的區別：

- 電腦中的資訊是二進位來表示，每個 0 或 1 被稱作一個 bit（位），用小寫 b 表示；
- 8 個 bit 為 1 個 byte (位元組)，用大寫 B 表示， 即，1B = 8b；
- 大寫 K 或小寫 k 表示千，即千個位（Kb）或千個位元組（KB）。

ECS 控制台中看到的出網入網監控資訊中，ps指的是/s，即每秒。kbps指的是網路速度，也就是每秒鐘傳送多少個千位的資訊。一般描述頻寬時常常把 bps 省略掉，例如：頻寬為 4M，完整的寫法應為 4Mbps。

誤區：頻寬是多少，下載速度就是多少。

正確的換算關係，以 1Mbps 頻寬為例：

1KB=8Kb (1 Byte=8 bits)

1Mbps=125KB/s

1Mbps 頻寬的下載速率，理論上為 125KB/s，而系統中一些應用程式（包括遠端連線）會佔用少量頻寬，因此實際中速率為 100KB/s−110KB/s 是比較常見的。

4.2 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`
	Pay-As-You-Go instances are released due after several days of overdue payment.	`AccountUnbalanced.Delete`

Event status

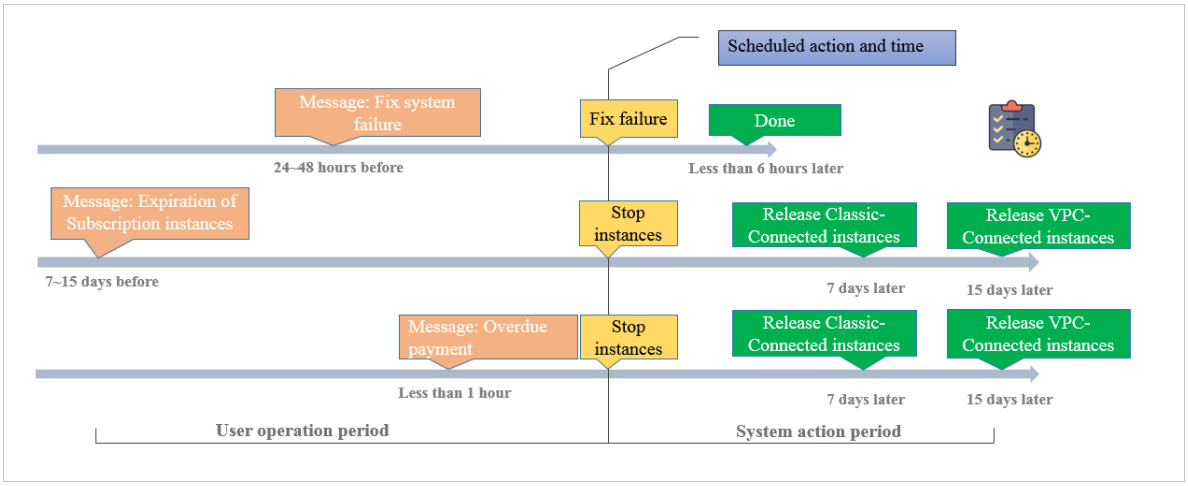
The following table describes the status of a system event during its lifecycle.

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, classic network-connected instances are released 7 days later and VPC-Connected 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_35* 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.

Event type	Parameter	Recommended
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.
A Pay-As-You-Go instance stops due to overdue payment.	AccountUnbalanced.Stop	You can either *recharge your account* or wait for the instances to stop.
A Subscription instance is released due to expiration.	InstanceExpiration.Delete	You can either *renew the instances* or wait for the instances to be released.
A Pay-As-You-Go instance is released due to overdue payment.	AccountUnbalanced.Delete	You can either *recharge your account* or wait for the instances to be released.

4.3 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

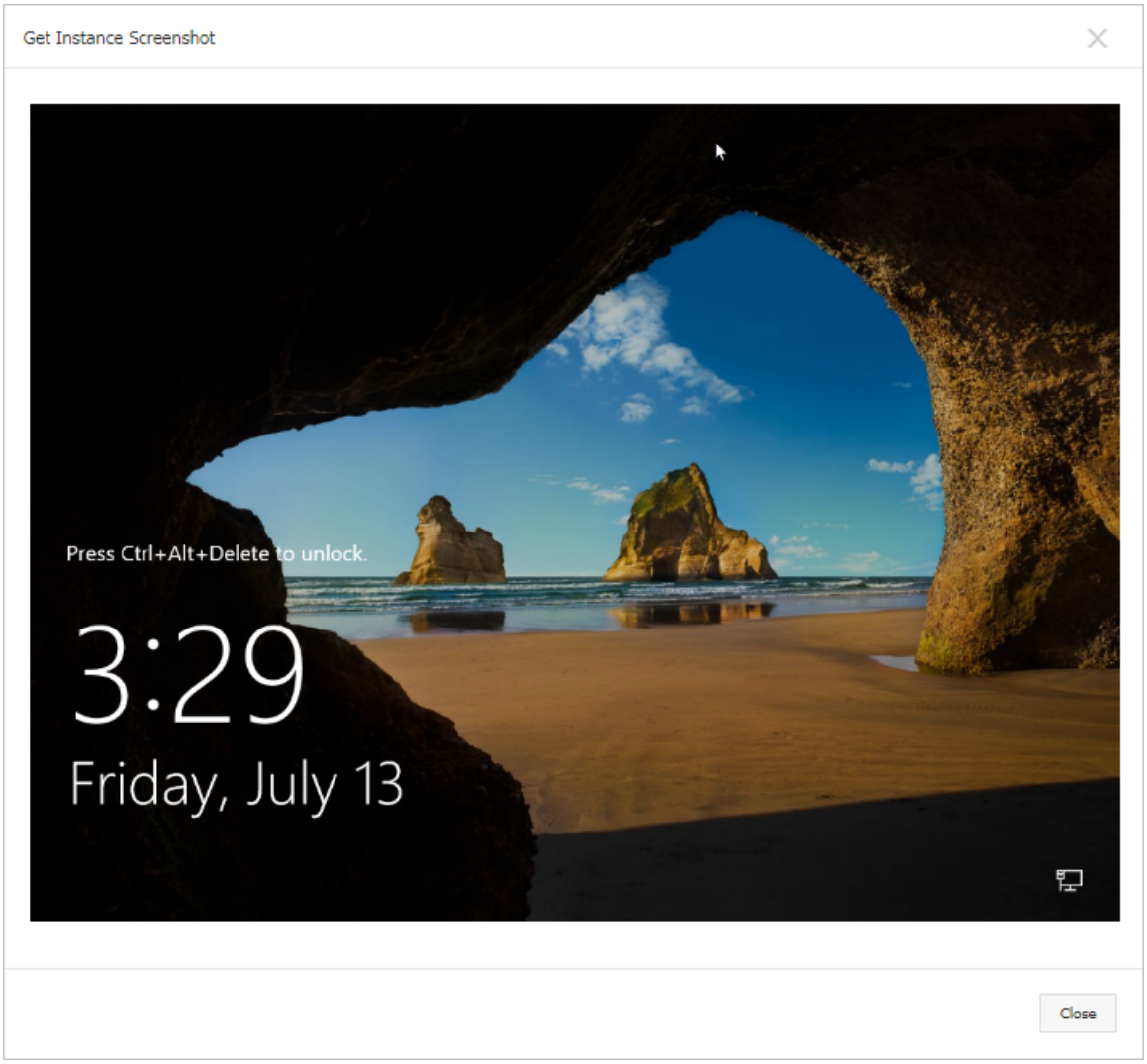
The instance must be in the **Running** (`Running`) status. For more information, see *Overview*.

Procedure

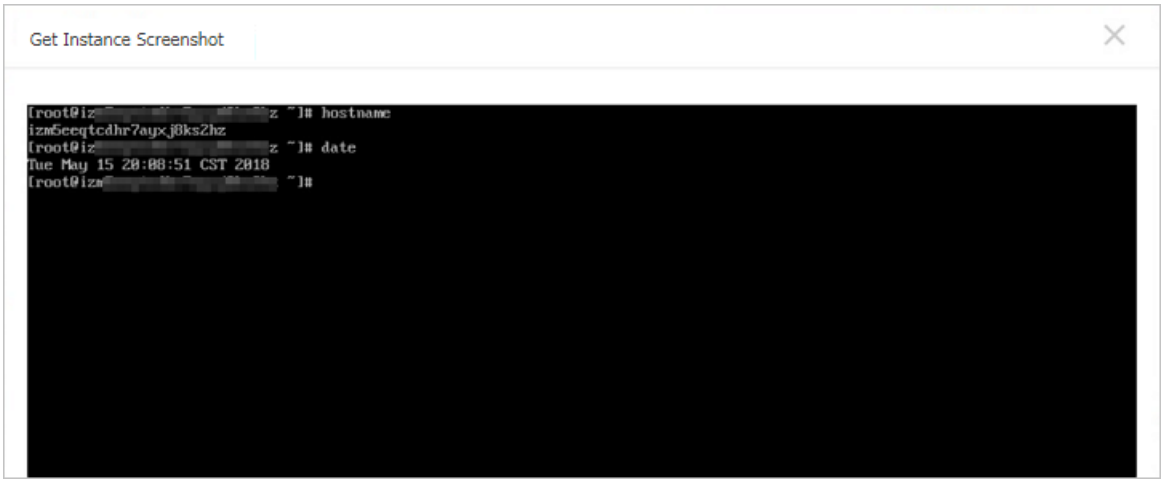
You can view instance console output and screenshot from the Instance Details page, the Instances list page, or by calling API.

Operation in Instance Details page

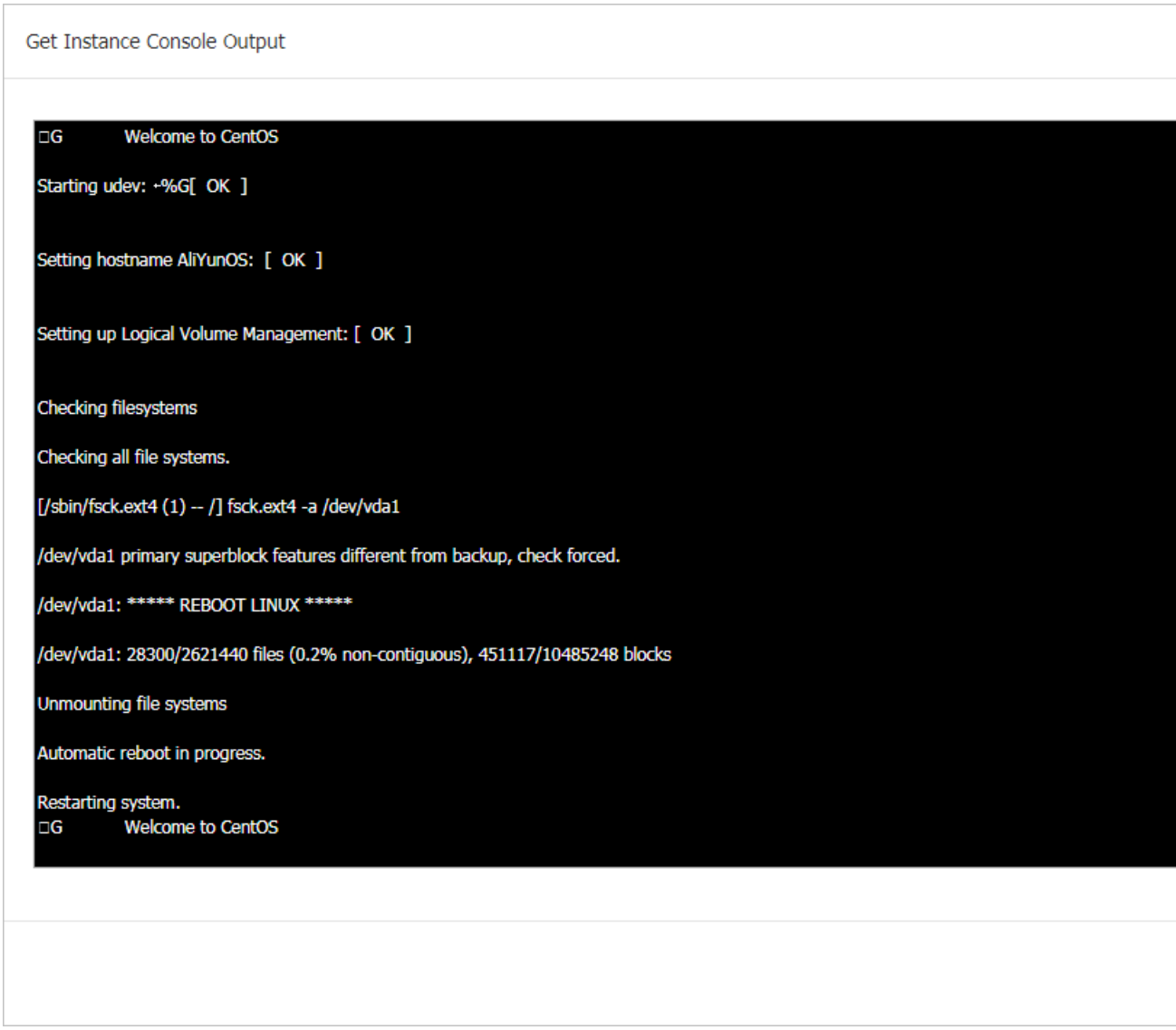
1. Log on to the *ECS Management Console*.
2. In the left-side navigation pane, click **Instances**.
3. Select the **region**.
4. Select and click the instance to troubleshoot and go to the **Instance Details** page.
5. Click **More** > **Get Instance Screenshot** to view the screenshot. Alternatively, 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 .

- *Unable to connect Windows instances*
- *Unable to connect to Linux instances*

5 雲助手

5.1 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:

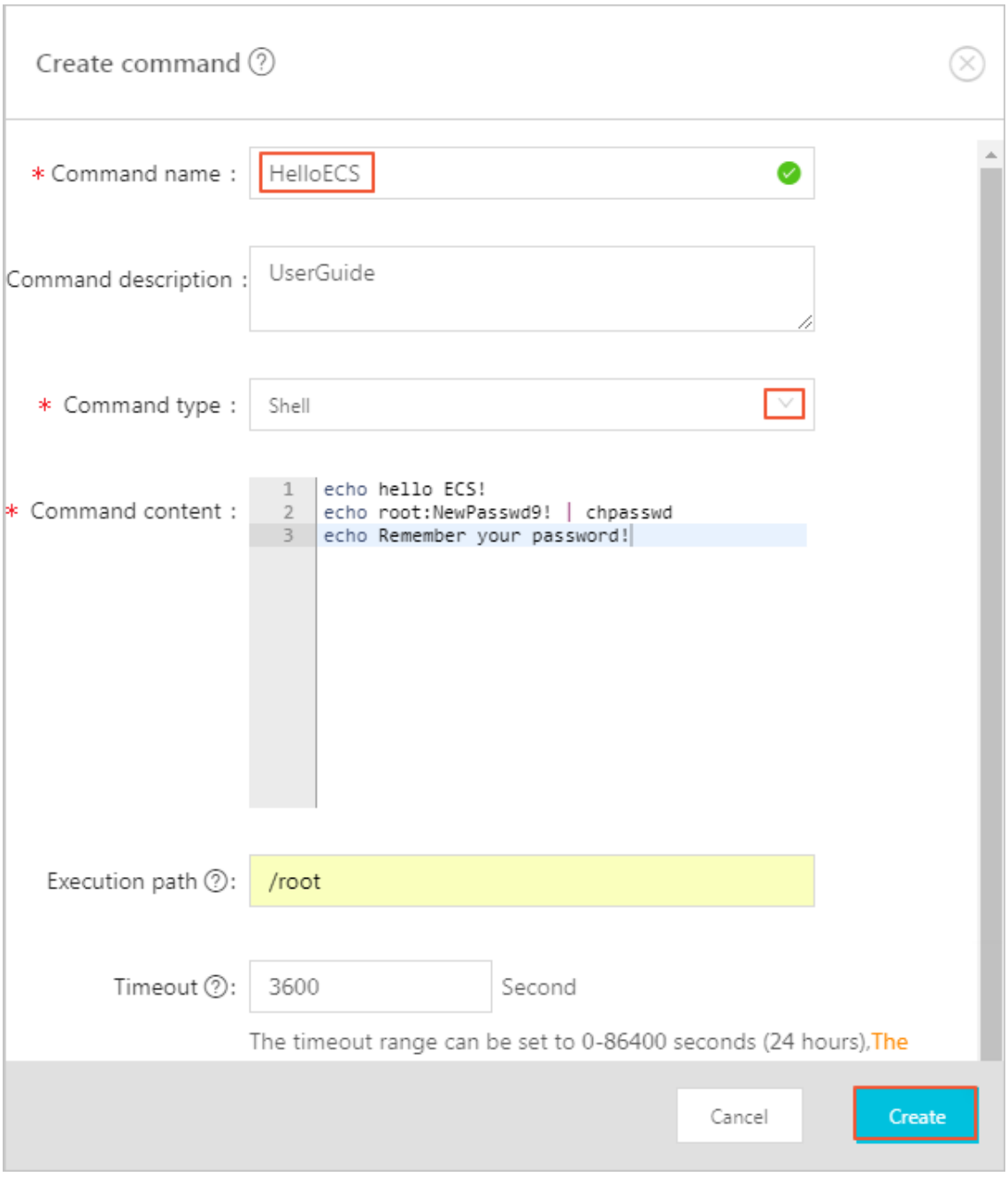
1. Log on to the ECS Console *ECS console* .
2. From the left-side navigation bar, select **Cloud Assistant**.
3. Select a region.
4. Click **Create Command**, and in the right-side pop-up window.
   a. Input a **command name**, such as HelloECS.
   b. Input a **command description**, such as UserGuide.
   c. Click the <x id="1"> 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

执行命令

5.2 執行命令

建立雲助手命令後，您可以在一台或者多台執行個體中執行命令。多台執行個體間的執行狀態和執行結果互相不影響，您還可以為命令設定執行循環。

使用限制

- 在一個阿里雲地域下，您每天最多能執行500次雲助手命令。
- 一次執行命令操作最多可以選擇50台執行個體。
- 目標執行個體的狀態必須處於**執行中**（`Running`）。
- 目標執行個體必須已安裝雲助手用戶端。
- 目標執行個體的網路類型必須是*Virtual Private Cloud*。
- 周期性雲助手命令設定的間隔不能小於10秒。
- 周期執行的時間設定基準為UTC +08:00，且該時間以執行個體的系統時間為準，您需要確保您的ECS執行個體的時間或者時區與您預期的時間一致。

執行命令

在管理主控台上執行命令的步驟如下所示：

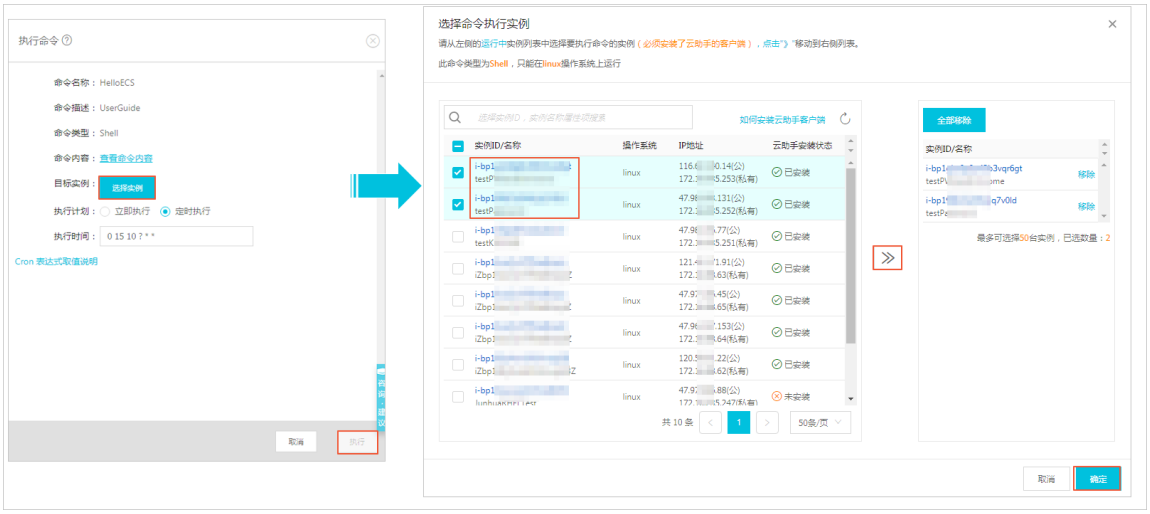
1. 登入 *ECS管理主控台* 。
2. 在左側導覽列，單擊**雲助手**。
3. 選擇地域。
4. 找到需要執行的雲助手命令，在右側**操作**列表中單擊**執行**，在右側彈窗中：
   a. 單擊**查看命令內容**確認命令內容。
   b. 單擊**選擇執行個體**，在彈窗中：
      A. 勾選一台或多台執行個體。
      B. 單擊 選中執行個體。

      > 说明：
      > Bat或者PowerShell命令只能選擇Windows執行個體，Shell命令只能選擇Linux執行個體，並且執行個體必須已安裝雲助手用戶端。否則單擊 表徵圖後無法選中執行個體。

      C. 單擊**確定**。

c. 選擇立即執行或者定時執行：

- 立即執行：雲助手命令在執行個體中即可執行一次。
- 定時執行：使用Cron運算式為設定命令為周期任務。填寫執行時間，更多詳情，請參閱 *Cron運算式取值說明*。



5. 單擊執行。

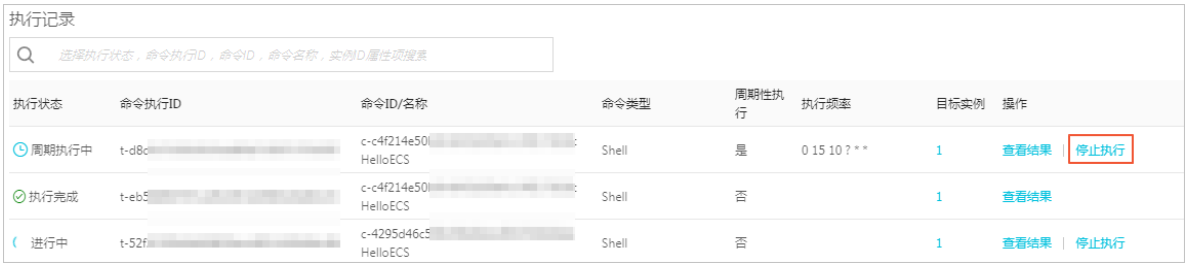
您也可以使用ECS API *InvokeCommand*執行雲助手命令。

停止執行命令

前提條件：命令執行狀態必須為進行中（Running）或者是周期命令。

在管理主控台上停止命令的步驟如下所示：

1. 登入 *ECS管理主控台* 。
2. 在左側導覽列，選擇雲助手。
3. 選擇地域。
4. 在執行記錄區域，找到需要停止的命令，在操作欄中單擊停止執行。



下一步

查詢執行結果與狀態

5.3 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:

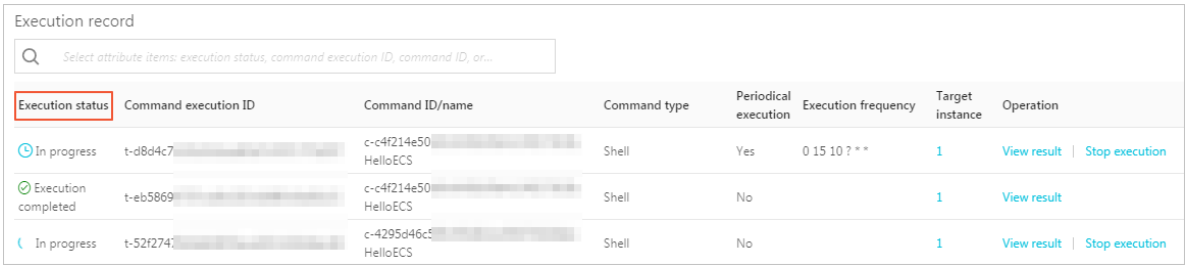
1. Log on to the ECS Console*ECS console* .
2. In the left-side navigation bar, select **Cloud Assistant**.
3. Select a region.
4. In the **Execution Record** area, search for the execution record of the necessary command execution, and select **View Results** from **Actions**.
5. In the pop-up window, 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 *ECS console* .
2. In the left-side navigation bar, 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.

5.4 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. Select **Cloud Assistant** from the left-side navigation pane.
3. Select a region.
4. Move the mouse cursor to the command you want to edit, and click the 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 *#unique_58* 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:

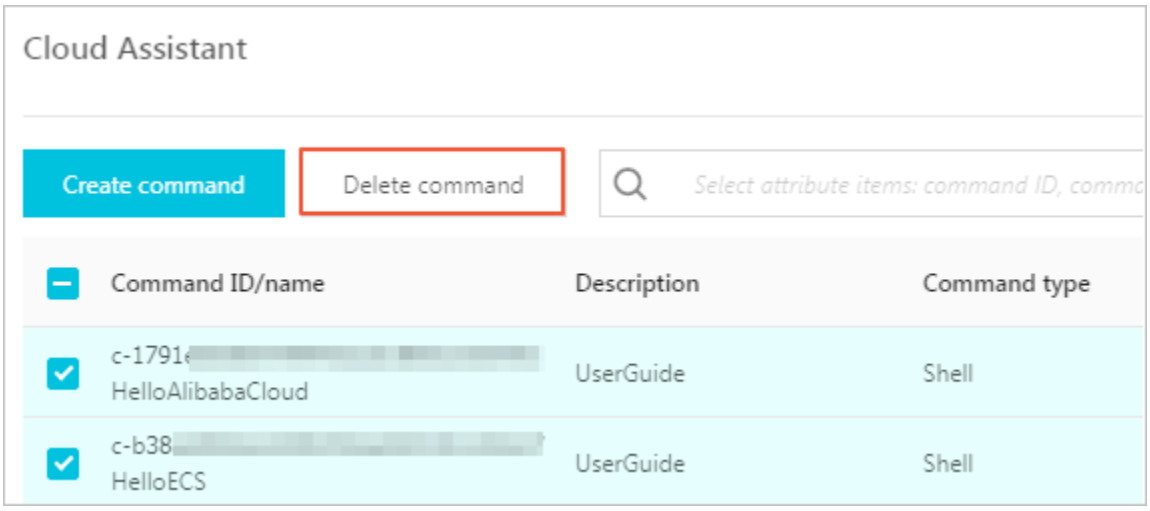
1. Log on to the *ECS console* .
2. Select **Cloud Assistant** from the left-side navigation pane.
3. Select a region.
4. Find the cloud assistant command you want to clone, and from the **Operation** list, click **Clone**.

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. Select **Cloud Assistant** from the left-side navigation pane.
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.

6 Quick reference

When using ECS, you may encounter various issues, such as connecting to the instance, resizing the disk, upgrading or downgrading the instance configurations, and using snapshots or images. This article provides you with a quick reference to popular features of ECS resources.

Watch 2 minutes of video to learn more about common operations.

Go Cloud

Operation instructions and limits

To guarantee proper operation of your ECS instance, You must carefully read all the *ECS operation instructions* and *Limits* before you use it.

Create and manage ECS instances

Basic operations

To use an ECS instances, follow these steps:

1. *Create an ECS instance*.
2. Connect to the ECS instance. Use different methods according to its operating system:
   a. Use the *Management Terminal* regardless of the operating system. Generally, this method is used for troubleshooting and maintenance.
   b. For Linux or Unix-like OS: *Connect to a Linux instance by using a password*, or *Connect to a Linux instance by using an SSH key pair*.
   c. For Windows OS: *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 if the configurations cannot meet your business needs.

- Prepaid instance: *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*

- *Change EIP Internet bandwidth*

If the current operating system does not meet your needs, you can *change the operating system*.

Billing

You can *switch from Pay-As-You-Go to subscription*.

Elaborate management of and control over ECS instances

You can use the following features to elaborate 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

When the capacity of the system disks or data disks cannot meet your business needs, you can *increase 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

When errors occur to data on a cloud disk, you can use a snapshot to *roll back a cloud disk* of the disk to 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 creating snapshots automatically.
2. *View a snapshot chain*.
3. To save space occupied by snapshots, *delete unnecessary snapshots*.

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 ease environment 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. Using custom images can simplify environment deployment.

You can own 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 images* across different accounts.
- *Import custom images*
- *Create and import on-premise images by using Packer*

You can *export custom images* to back up the environment and *delete custom images*.

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.

When new security group rules impair the 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 use tags to group resources to improve efficiency. To use tags, follow these steps:

1. *Add a tag to resources*.
2. *Filter resources by tags*.
3. *Delete a tag*.

7 ECS operation instructions

To guarantee proper operation of your ECS instance, you must take the considerations outlined in this section into account before use.

Prohibitions

- Alibaba Cloud prohibits you from using your instance for flow-through services. Any violation results in punishment up to shutdown and lockout of instance, and termination of services.
- Alibaba Cloud prohibits you from using your instance for click farming, advertising, or fictitious transactions.
- Alibaba Cloud prohibits you from activating SELinux.
- Alibaba Cloud prohibits you from uninstalling hardware related drivers.
- Alibaba Cloud prohibits you from arbitrarily modifying the MAC address of the network adapter.

Suggestions

- For an ECS with more than 4 GiB RAM, we recommend that you use a 64-bit OS, because a 32-bit OS supports a maximum of 4 GiB RAM. Currently available 64-bit systems include:
  - Aliyun Linux
  - CoreOS
  - CentOS
  - Debian
  - FreeBSD
  - OpenSUSE
  - SUSE Linux
  - Ubuntu
  - Windows
- 32-bit Windows OS supports CPUs with up to 4 cores.
- A minimum of 2 GiB RAM is needed for buiding a website on a Windows instance, and 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, you must enable auto-start of service applications upon OS boot.
- For I/O-optimized instances, do not stop the aliyun-service process.

- We do not recommmend that you update the kernel and the OS. For more information, see *How to avoid Linux instance startup failure after kernel upgrade*.

Windows restrictions

- Do not close the built-in shutdownmon.exe process, which may delay the restart of your Windows server.
- Do not rename, delete, or disable the Administrator account.
- We do not recommend that you use virtual memory.

Linux restrictions

- Do not modify the content of the default /etc/issue file. Otherwise, if you create a custom image of the ECS instance and create a new ECS instance based on the image, the new instance cannot start because the operating system edition cannot be recognized.
- Proceed with caution when modifying permissions of the directories in the root partition, such as `/etc`, `/sbin`, `/bin`, `/boot`, `/dev`, `/usr` and `/lib`. Improper modification of permissions may cause errors. Such modifications may cause system errors.
- Do not rename, delete, or disable the Linux root account.
- Do not compile or perform any other operations on the Linux kernel.
- We do not recommend that you use swap partition.
- Do not enable the NetWorkManager service. This service conflicts with the internal network service of the system and causes network errors.

For more information, see *Limits*.

8 使用限制

使用Elastic Compute Service有下列限制：

- 暫不支援虛擬化軟體安裝和再進行虛擬化（如安裝使用VMware）。目前，僅 *ECS Bare Metal Instance#神龍#*和*Super Computing Cluster#SCC#* 支援再虛擬化。
- 暫不支援音效卡應用。
- 不支援直接載入外接硬體裝置（如硬體加密狗、隨身碟、外接硬碟、銀行U key等），您可以嘗試軟加密狗或者動態口令二次驗證等。
- 暫不支援SNAT等IP包地址轉換服務。您可以使用自己搭建VPN或者代理方式來實現。
- 暫不支援多播協議。如果需要使用多播，建議改為使用單播點對點方式。
- 目前Log Service（LOG）不支援32位Linux雲伺服器。您可以參考 服務入口 查看支援Log Service的地域（Region）；參考Logtail采集簡介 查看支援Log Service的雲伺服器系統。

除了以上所列限制外，Elastic Compute Service還有如下表所示的限制。

ECS執行個體

限制項	普通使用者限制描述	例外申請方式（例外上限）
建立ECS執行個體的使用者限制	實名認證	沒有例外
建立隨用隨付資源的限制	賬戶餘額、代金券和信用度之和不得小於100元	提交工單
可以建立隨用隨付執行個體的規格	vCPU核心數少於16（不含16）的執行個體規格	提交工單
一個帳號在每個地域的隨用隨付執行個體配額	最少50 vCPU（使用者會員等級不同會相應自動提升）	提交工單
一個帳號在每個地域的搶佔式執行個體配額	最少50 vCPU（使用者會員等級不同會相應自動提升）	提交工單
一個帳號在每個地域的執行個體啟動模板數量	最多30個	沒有例外
一個執行個體啟動模板中的版本數量	最多30個	沒有例外

限制項	普通使用者限制描述	例外申請方式（例外上限）
隨用隨付轉預付費	以下執行個體規格（族）不支援：t1、s1、s2、s3、c1、c2、m1、m2、n1、n2、e3	沒有例外
預付費轉隨用隨付	• 每月最少 5000 vCPU * 小時 • 每月有最大退款額度限制，根據使用者會員等級不同，最大退款額度不同	沒有例外

區塊存放裝置

限制項	普通使用者限制描述	例外申請方式（例外上限）
建立隨用隨付雲端碟的使用者限制	使用者必須實名認證，而且賬戶餘額、代金券和信用度之和不得小於100元	沒有例外
一個帳號在所有地域的隨用隨付雲端碟配額	使用者帳號下所有地域的隨用隨付執行個體數量 * 5	提交工單
單一實例系統盤數量	1	沒有例外
單一實例資料盤數量	16塊（包括雲端碟和共用區塊存放裝置）	沒有例外
單塊共用區塊存放裝置允許同時掛載的執行個體數量	8	沒有例外
單個帳號在全地域的共用區塊存放裝置配額	10	提交工單
單塊普通雲端碟容量	5 GiB ~ 2000 GiB	沒有例外
單塊SSD雲端碟容量	20 GiB ~ 32768 GiB	沒有例外
單塊高效雲端碟容量	20 GiB ~ 32768 GiB	沒有例外
單塊SSD本地碟容量	5 GiB ~ 800 GiB	沒有例外
單一實例SSD本地碟總容量	1024 GiB	沒有例外
單塊NVMe SSD本地碟容量	1456 GiB	沒有例外
單一實例NVMe SSD本地碟總容量	2912 GiB	沒有例外
單塊SATA HDD本地碟容量	5500 GiB	沒有例外

限制項	普通使用者限制描述	例外申請方式（例外上限）
單一實例SATA HDD本地碟總容量	154000 GiB	沒有例外
單塊SSD共用區塊存放裝置	32768 GiB	沒有例外
單一實例SSD共用區塊存放裝置總容量	128 TiB	沒有例外
單塊高效共用區塊存放裝置	32768 GiB	沒有例外
單一實例高效共用區塊存放裝置總容量	128 TiB	沒有例外
單塊ESSD雲端碟	32768 GiB	沒有例外
系統盤單盤容量限制	• Windows 40 GiB ~ 500 GiB • Linux（不包括CoreOS）+ FreeBSD：20 GiB ~ 500 GiB • CoreOS：30 GiB ~ 500 GiB	沒有例外
資料盤單盤容量限制	• 普通雲端碟：5 GiB ~ 2000 GiB • SSD雲端碟/高效雲端碟/SSD共用區塊存放裝置/高效共用區塊存放裝置：20 GiB ~ 32768 GiB • 本地碟：以各本地碟的容量為準	沒有例外
本地碟執行個體是否可以自行掛載新的本地碟	不允許	沒有例外
本地碟執行個體是否支援變更配置	僅允許變更頻寬	沒有例外
系統盤掛載點範圍	/dev/xvda	沒有例外
資料盤掛載點範圍	/dev/xvd[b-z]	沒有例外

快照

限制項	普通使用者限制描述	例外申請方式（例外上限）
快照額度	彈性區塊存放裝置保有數量 * 64	沒有例外

鏡像

限制項	普通使用者限制	例外申請方式（例外上限）
一個帳號在一個地域的保有自訂鏡像配額	最少100個	提交工單
單個鏡像最多可共用的使用者數量	50	提交工單
鏡像與執行個體規格的限制	4 GiB及以上記憶體的執行個體規格不能使用32位鏡像	沒有例外

金鑰組

限制項	普通使用者限制描述	例外申請方式（例外上限）
一個帳號在每個地域的金鑰組配額	500	沒有例外
支援通行金鑰對的執行個體規格	僅系列I的非I/O優化執行個體不支援	沒有例外
支援通行金鑰對的鏡像類型	僅Linux支援	沒有例外

公網頻寬

限制項	普通使用者限制描述	例外申請方式（例外上限）
公網入頻寬可選範圍	200 Mbit/s	沒有例外
公網出頻寬可選範圍	• 預付費：200 Mbit/s • 隨用隨付：100 Mbit/s	提交工單，最高配置400 Mbit/s
單一實例更換分配的公網IP地址的限制	新建執行個體6小時內可以更換公網IP地址。一個執行個體最多可以更換3次	沒有例外

安全性群組

限制項	普通使用者限制描述	例外申請方式（例外上限）
單個安全性群組內的執行個體/IP配額	• 經典網路類型執行個體的安全性群組：1000個經典網路類型執行個體 • VPC類型執行個體的安全性群組：2000個私網IP（主網卡和輔助網卡共用此配額）	沒有例外
單個安全性群組的授權規則配額	100	沒有例外
一個帳號在一個地域的安全性群組配額	最少100個（使用者會員等級不同會相應自動提升）	提交工單
一個執行個體下每個彈性網卡所屬的安全性群組配額	5	提交工單
通信埠	公網出方向的STMP預設通信埠25，預設受限，而且不能通過安全性群組規則開啟	提交工單申請解封，請參考 *TCP 25* 通信埠控制台解封申請

彈性網卡

限制項	普通使用者限制描述	例外申請方式（例外上限）
一個帳號在一個地域的彈性網卡配額	最少100個（使用者會員等級不同會相應自動提升）	提交工單

標籤

限制項	普通使用者限制描述	例外申請方式（例外上限）
單個執行個體允許綁定的標籤數量	10	沒有例外

API

限制項	普通使用者限制描述	例外申請方式（例外上限）
CreateInstance調用次數	一分鐘內最多200次	提交工單

说明：

專有網路（VPC）的產品限制請參見 使用限制。

9 Connect to instances

9.1 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*

9.2 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.

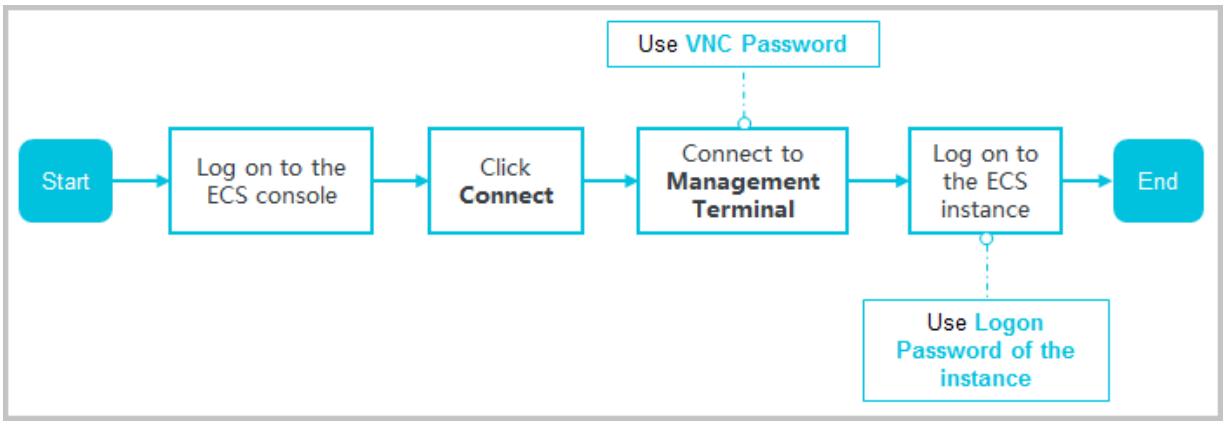
Note:
The Management Terminal can be used to connect to an instance even if no public IP address is assigned to your instance.

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.

9.3 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:

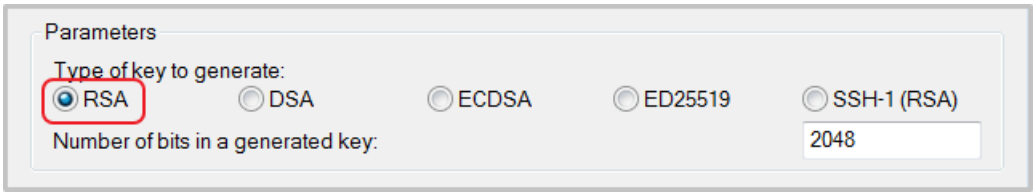
When you *create an SSH key pair*, download the .pem private key.

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**.

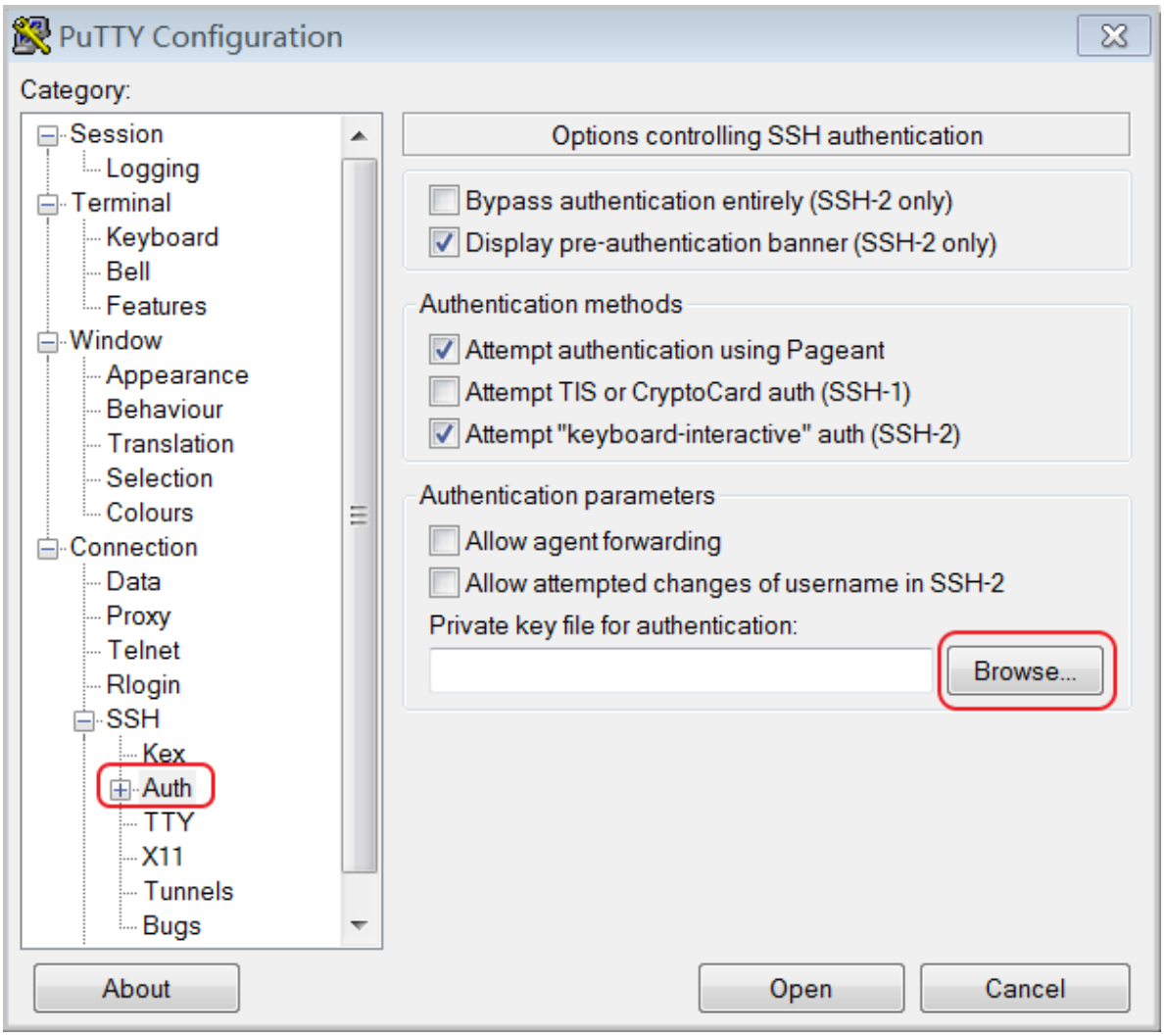
e. Click **OK** to close the confirmation dialog box.

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.

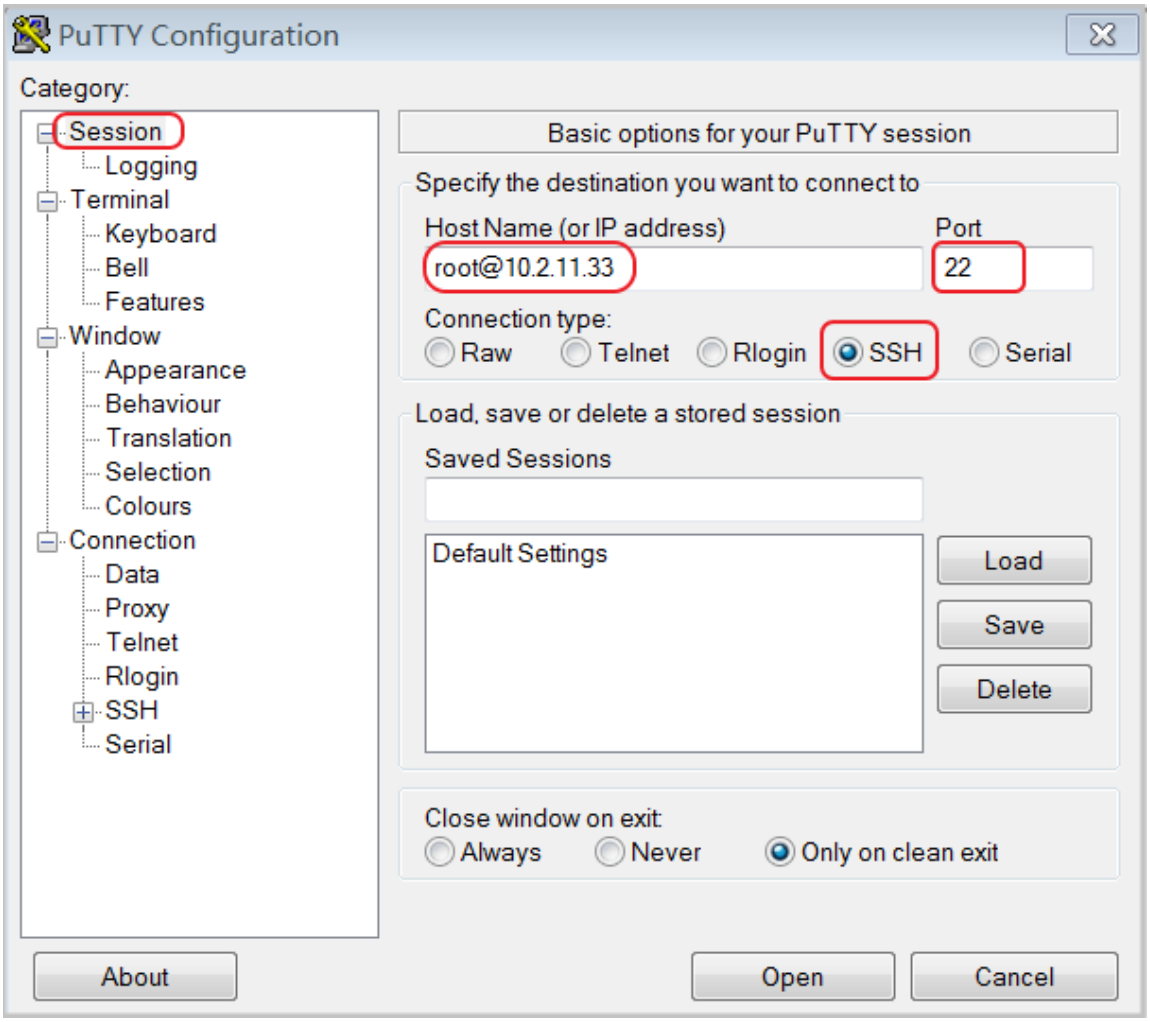
2. Start PuTTY.

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	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. 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`.

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

Procedure

Based on the operating system of your local machine, 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

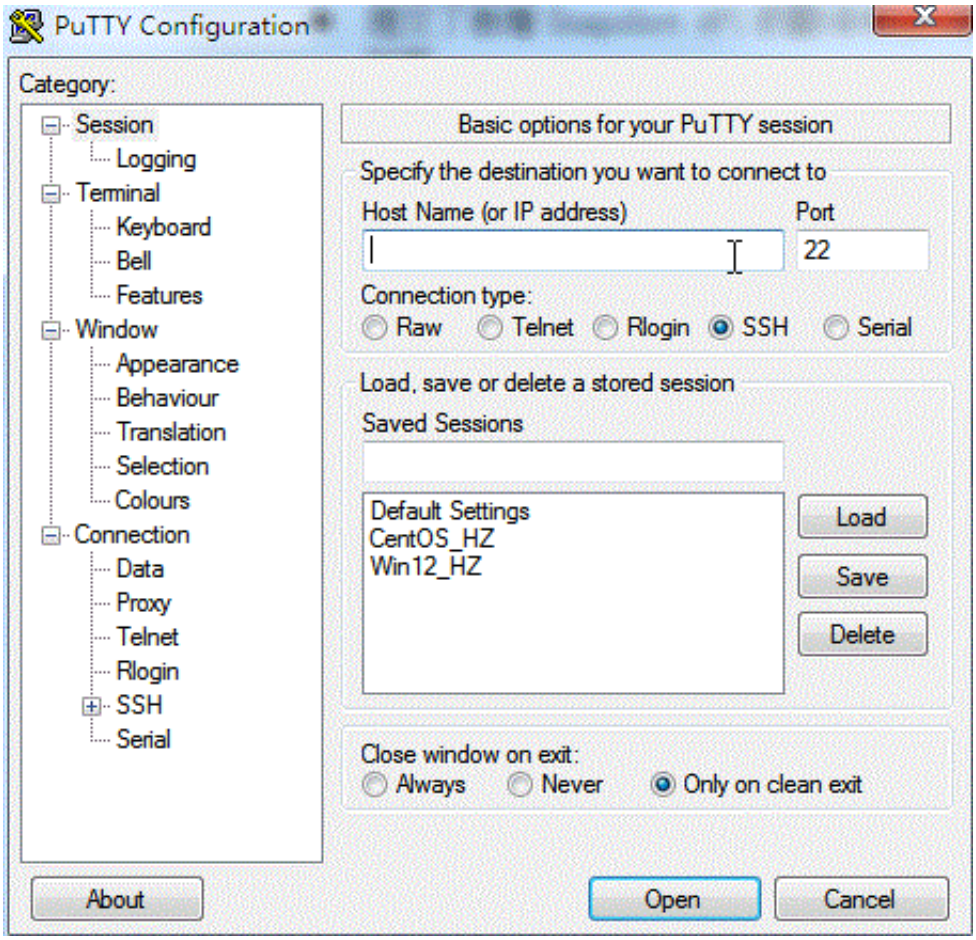
If your local machine is running Windows OS, you can use a remote connection tool, such as PuTTY, to connect to a Linux instance. In this article, we use PuTTY as an example to describe how to connect to a Linux instance by using the password authentication method. Before you start, download *PuTTY*.

Note:
To do so, you can also watch the video: *How the small assistant family can connect to a Linux instance remotely*.

Follow these steps to connect to a Linux instance:

1. Start putty.exe.
2. In the left-side navigation pane, click **Session**, and configure the following parameters:
   - **Host Name**: Type the public IP address or EIP address of the instance.

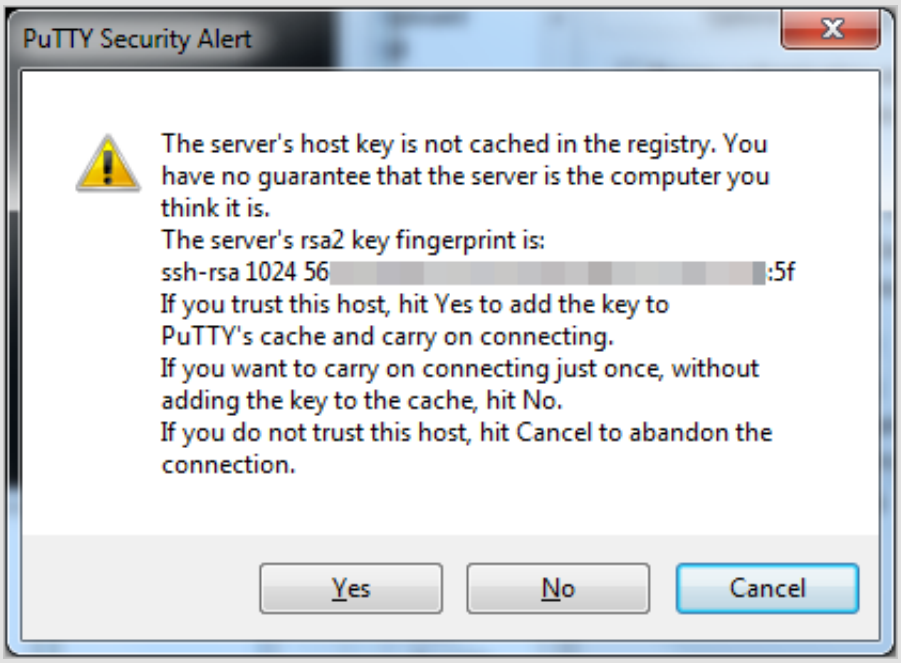
- **Port**: Type 22.
- **Connection Type**: Select SSH.
- (Optional) **Saved Session**: If you do not want to repeat the configurations during the next logon, add a name for the session, and click **Save**.



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

Note:

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



4. 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

The connection failed, and you can refer to this document for troubleshooting issues: *unable to connect to the Linux instance*.

You can run a script to install a graphical desktop on an instance running CentOS. For more information, see *Automatic installation tool for Linux instance*.

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

Procedure

Based on the operating system of your local machine, you have various options to connect to a Windows instance:

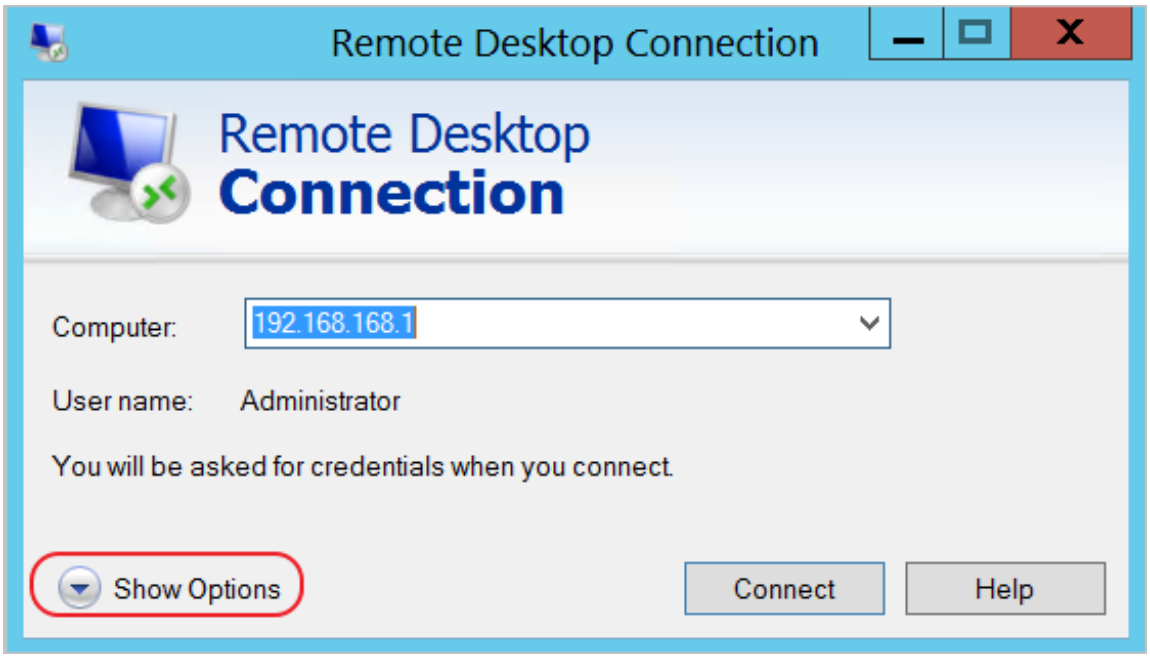
- *Windows OS*
- *Linux*
- *Local devices use Mac OS Operating System*
- *Android or iOS*

Windows OS

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

> **Note:**
> To do so, you can also watch the video: How the small assistant family can connect to Windows instance remotely.

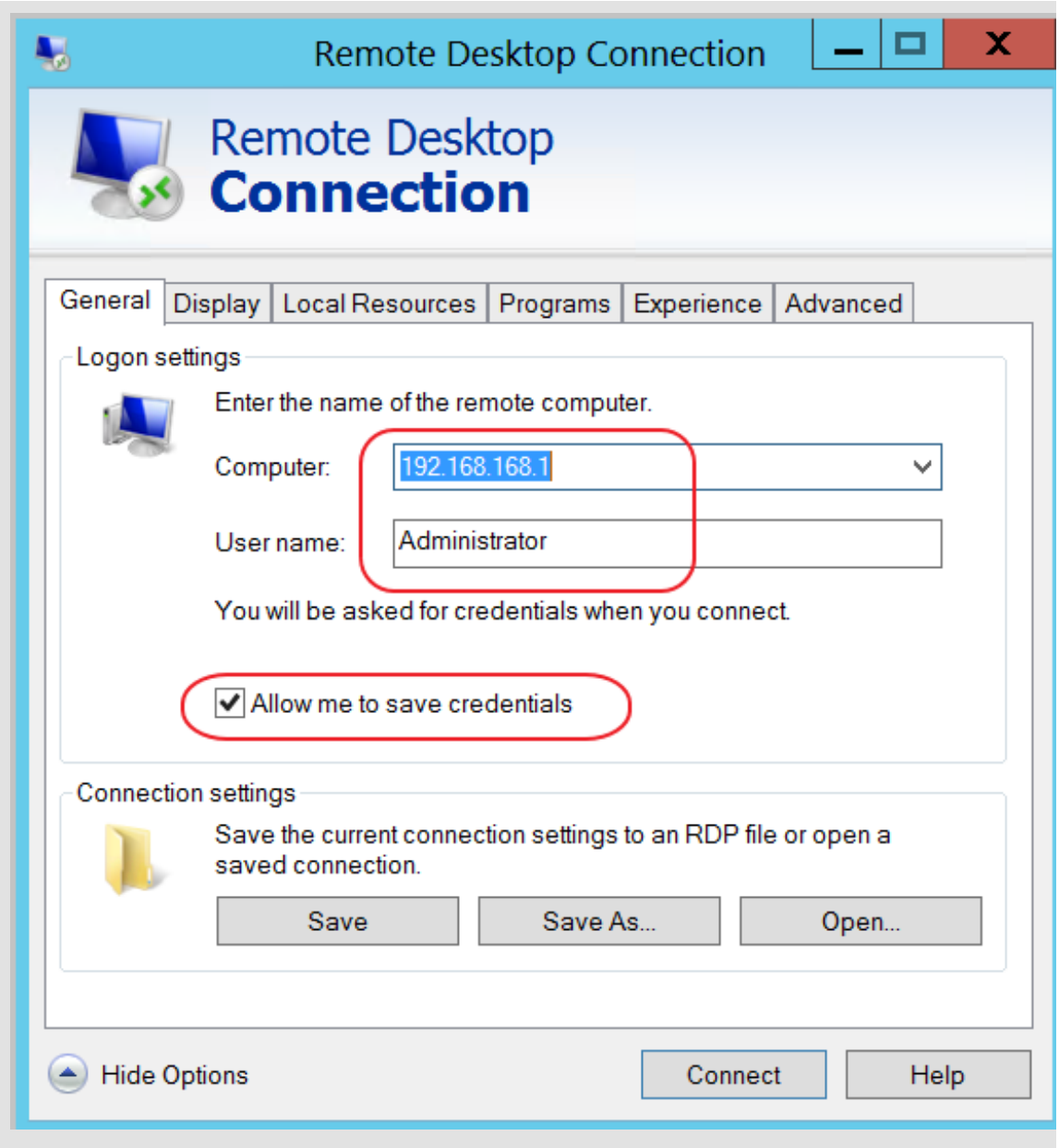
1. Use any one of the following methods to start **mstsc**:
   - Select **Start** > **icon** > **Remote Desktop Connection**.
   - Click the **Start** icon and search for mstsc.
   - Press the shortcut key **Windows Logo** + **R** to open the **Run** windows, type **`mstsc`**, and then press the Enter 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.

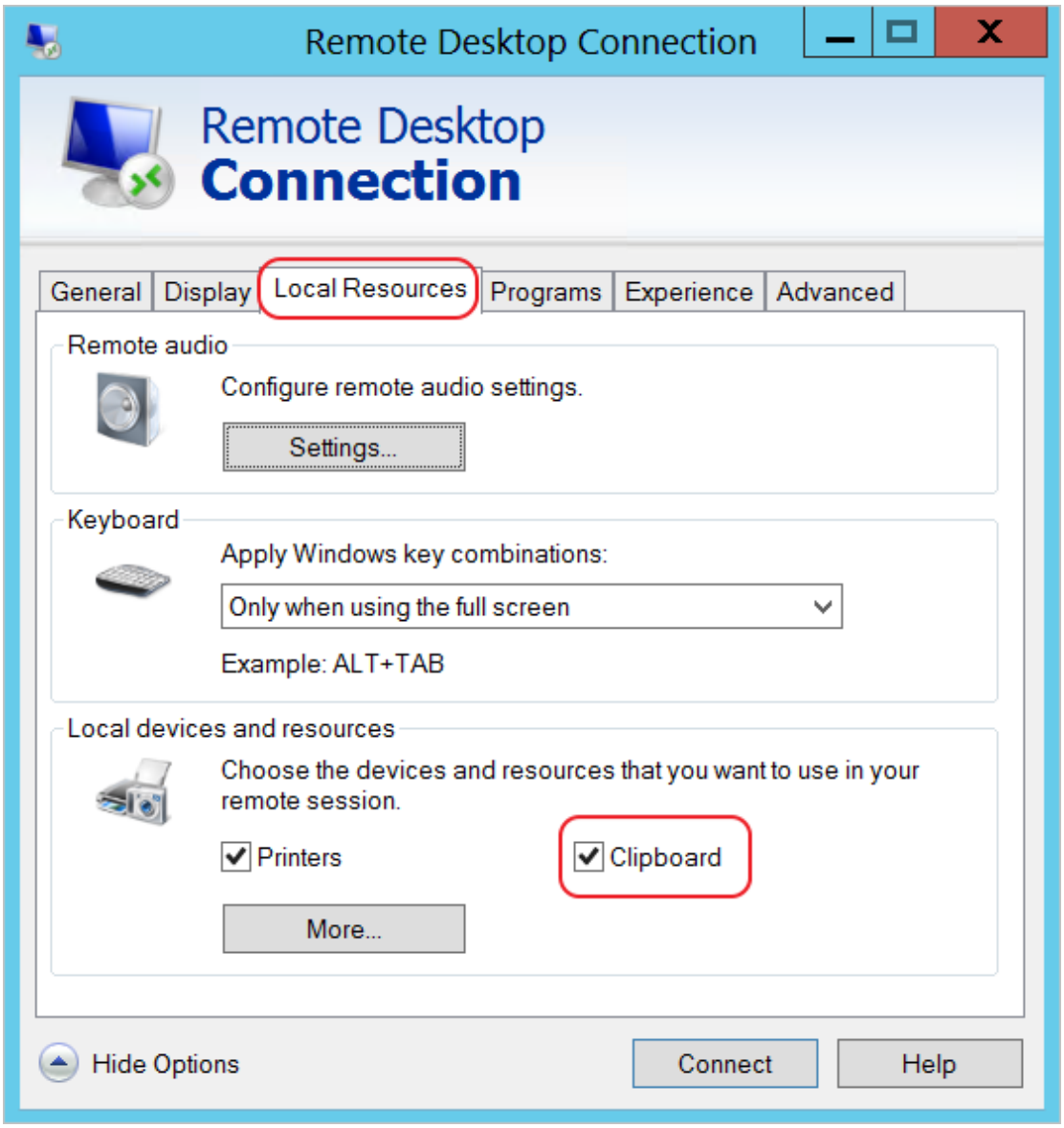
c. Type the user name. The default user name is **`Administrator`**

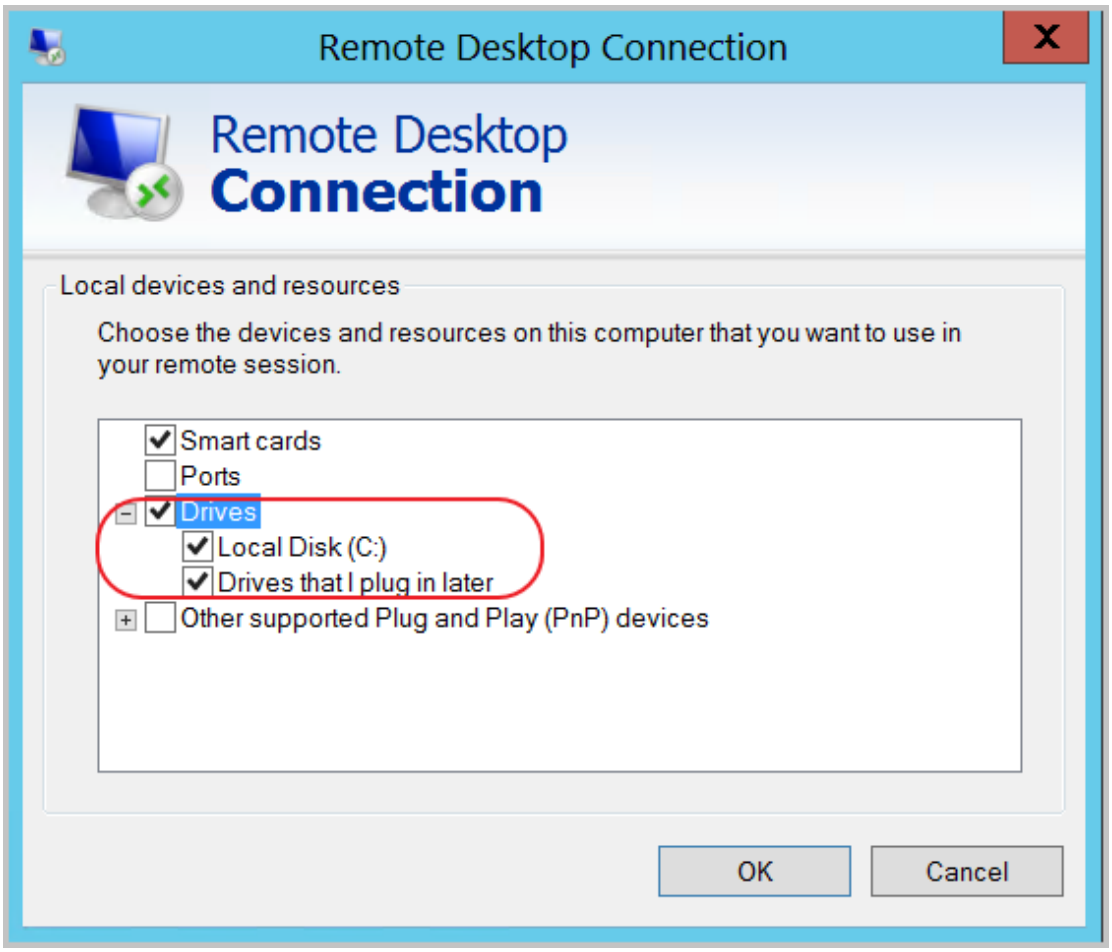
> **Note:**
> If you want to log on to the instance next time without repeating these steps, select **Allow me to save credentials**.



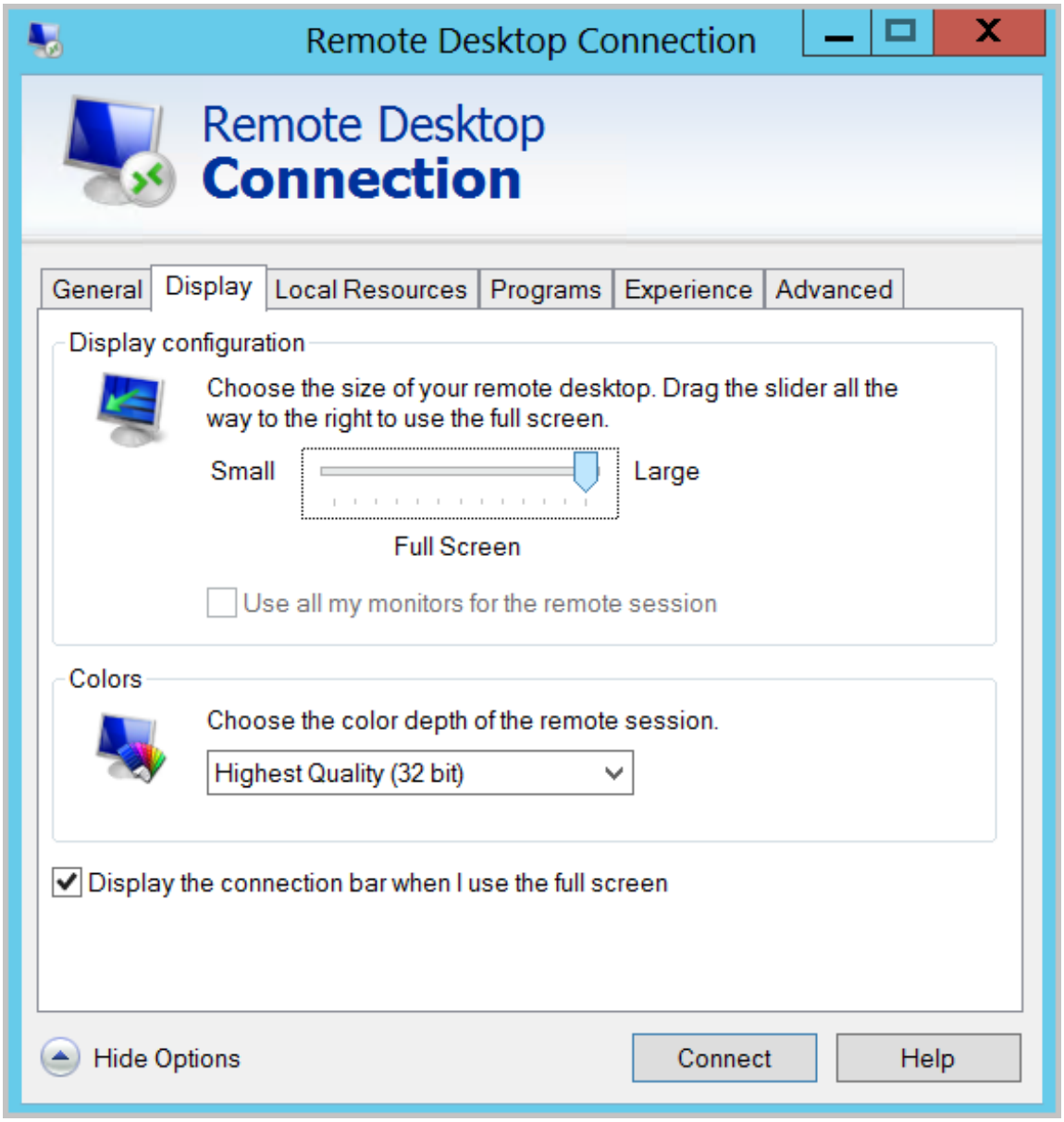
d. Optional. If you want to copy text or files from the local machine to the instance, click the **Local Resources** tab to see options for sharing local computer resources.

- 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.


Remote Desktop Connection
Remote Desktop
Connection
General
Display
Local Resources
Programs
Experience
Advanced
Remote audio
Configure remote audio settings.
Settings...
Keyboard
Apply Windows key combinations:
Only when using the full screen
Example: ALT+TAB
Local devices and resources
Choose the devices and resources that you want to use in your remote session.
Printers
Clipboard
More...
Hide Options
Connect
Help



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*.

Local devices use Mac OS Operating System

When connecting windows instances from Mac OS, must first download install Microsoft Remote on Mac App Store Desktop Connection for Mac. If you only have a Chinese apple account, you can download microsoft's official Microsoft remote from the hockeyapp. Desktop for Mac beta. This software applies only to Mac OS 10.10 and later systems.

This section is Microsoft Remote Desktop for Mac beta (MRD beta) for example, how to use Mac Connect windows instances on OS:

- *First connection*
- *Connect again*

First connection

Your first MRD on Mac OS When beta connects windows instance, follow these steps:

1. Start MRD beta.
2. Click get started.

3. In the quick connect window, enter the public or EIP address of Windows instance, and click Connect.

4. In the pop-up dialog box, enter your login information:
   - User name: Enter Administrator. The Default User Name For Windows instances is Administrator.
   - Password: Enter the instance login password.

5. In the pop-up dialog box, click Continue.

At this point, you have successfully logged on Windows instance desktop.

Connect again

MRD on Mac OS for the second time and later When beta connects windows instance, follow these steps:

1. Start MRD beta.
2. Click Add desktop, and on the pop-up add Desktop Dialog box, set PC name and select how to connect later (User Account), and click Save.
3. Select the instance icon.
4. In the toolbar, select  > **Connect**.
5. In the pop-up dialog box, enter your login information:
   - User name: Enter Administrator. The Default User Name For Windows instances is Administrator.
   - Password: Enter the instance login password.
6. In the pop-up dialog box, click Continue.

At this point, you have successfully logged on Windows instance desktop.

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 Links

Connection failed, you can refer to this document for troubleshooting issues: you cannot connect windows instance.

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

- You have downloaded and installed the appropriate app:
  - The iOS device has SSH Control Lite installed.
  - The Android device has JuiceSSH installed. You can get it from various Android app markets.

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.

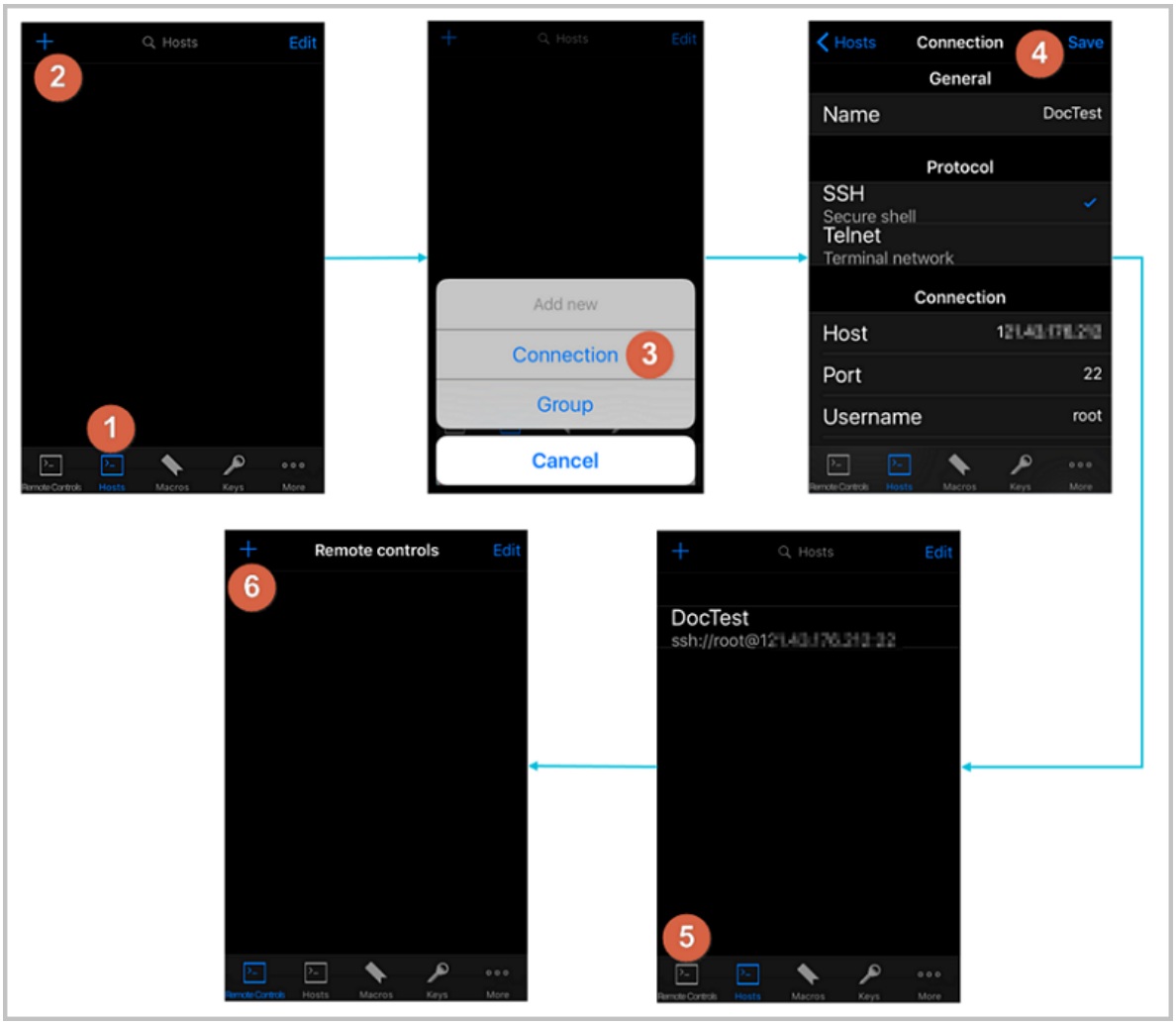
For Android devices, see *Use JuiceSSH to connect to a Linux instance*. In this example, user name and password are used for the authentication.

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.

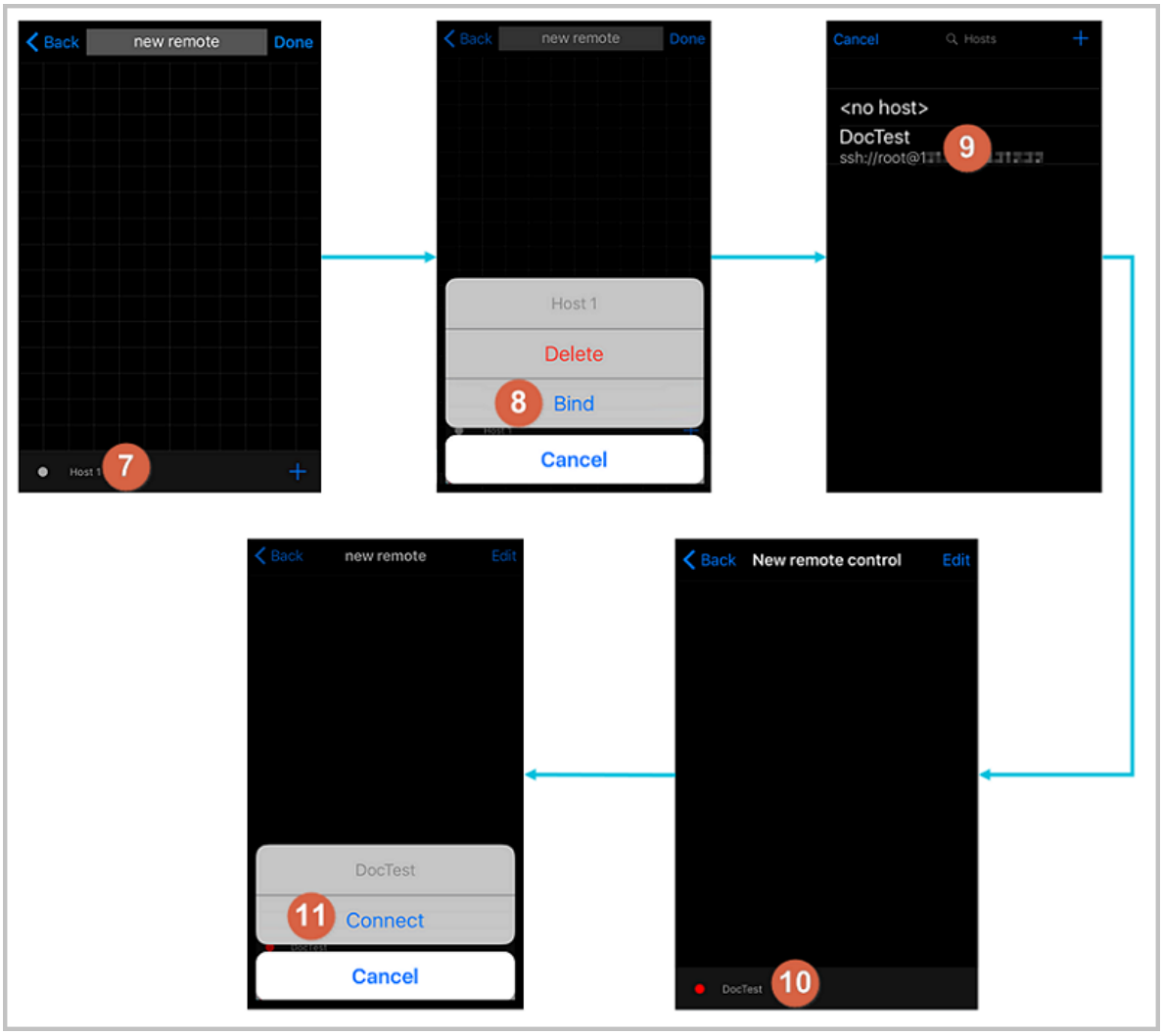
   The following figure shows Steps 1 through 6.



7. On the **New remote** page, tap **Host1**.

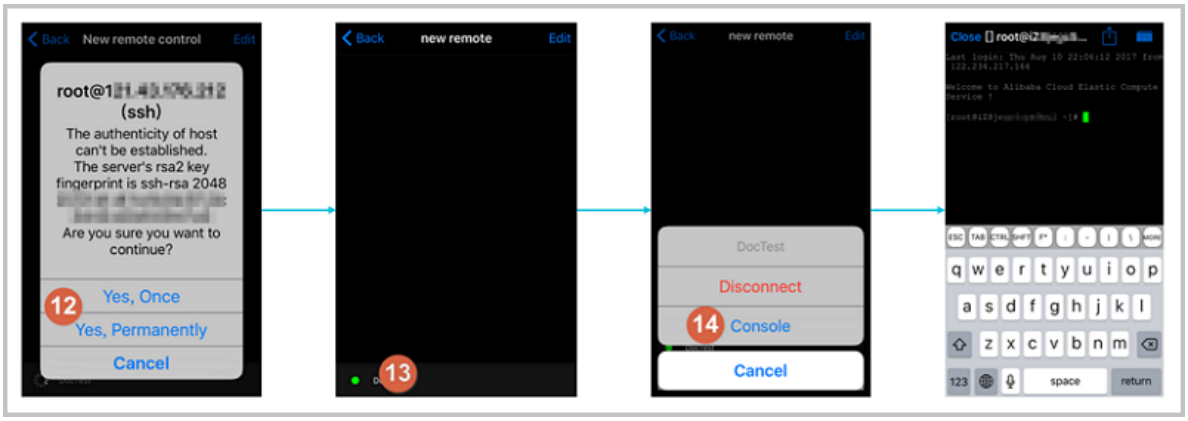
8. In the action sheet, tap **Bind**.
9. Select the new Linux instance. In this example, select `DocTest`.
10. On the **New remote** page, tap **Done** to switch it to the **Edit** mode, and then tap **DocTest**.
11. In the action sheet, tap **Connect**.

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.

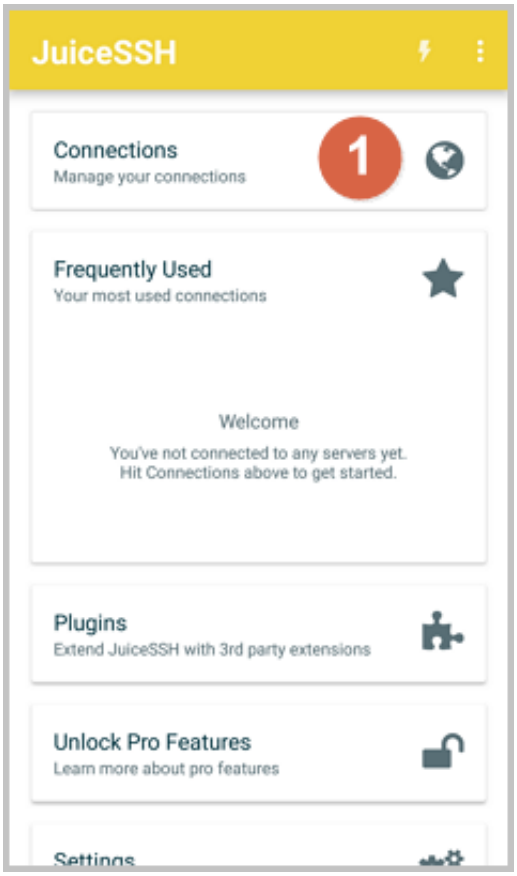
The following figure shows Steps 12 through 14:



Now, you are connected to the Linux instance.

Use JuiceSSH to connect to a Linux instance

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



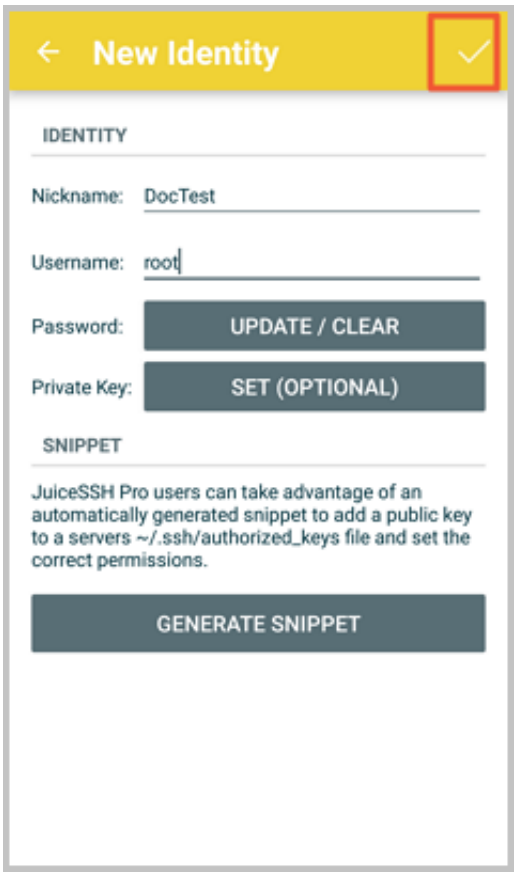
2. Under the **Connections** tab, tap the **+** icon.

3. On the **New Connection** page, add the connection information and tap the icon. The following connection information is required:

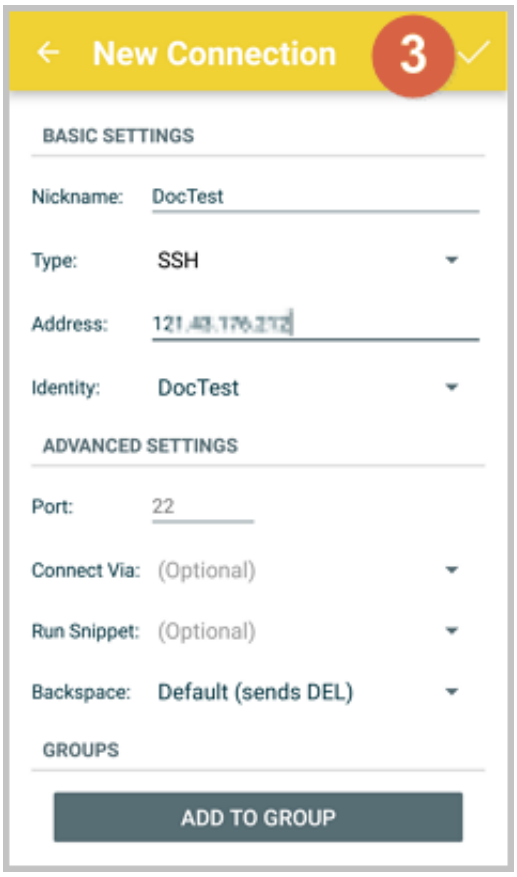
   - **`Nickname`**: Specify the name of the connection session. `DocTest` is used in this example.
   - **`Type`**: Use the default value SSH.
   - **`Address`**: Type the public IP address of the Linux instance to connect to.
   - To set**`I Identity`**, follow these steps:

     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:

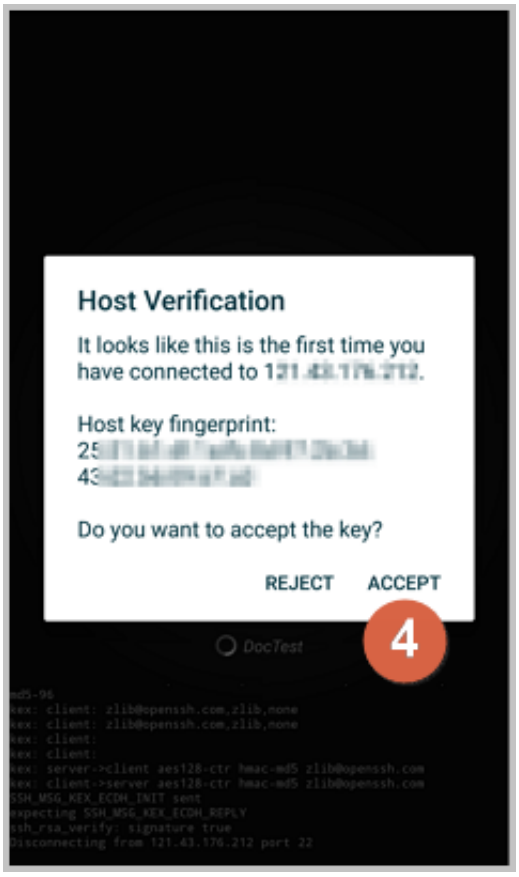
        - **`Nickname`**: Optional. You may set a nickname to ease management. `DocTest` is used in this example.
        - **`Username`**: Type `root` for the user name.
        - **`Password`**: Tap **SET(OPTIONAL)**, and type the logon password of the instance.



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

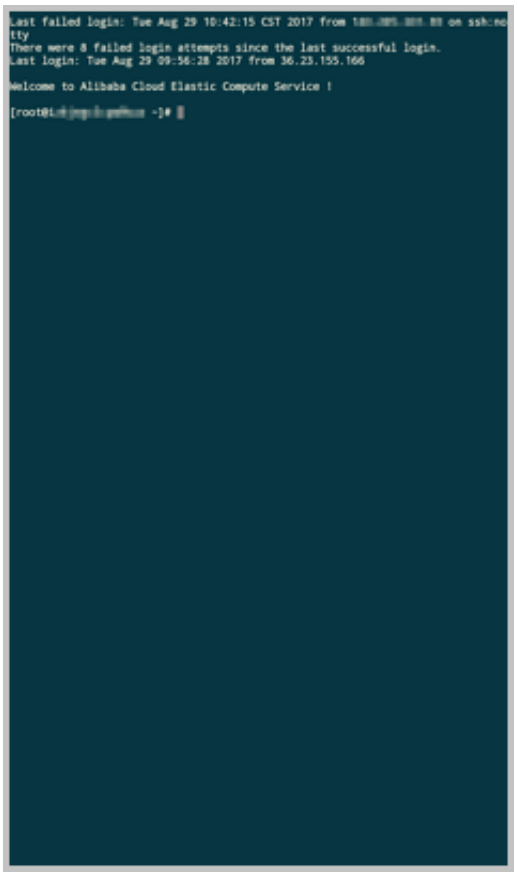


4. Confirm the message, and tap **ACCEPT**.



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

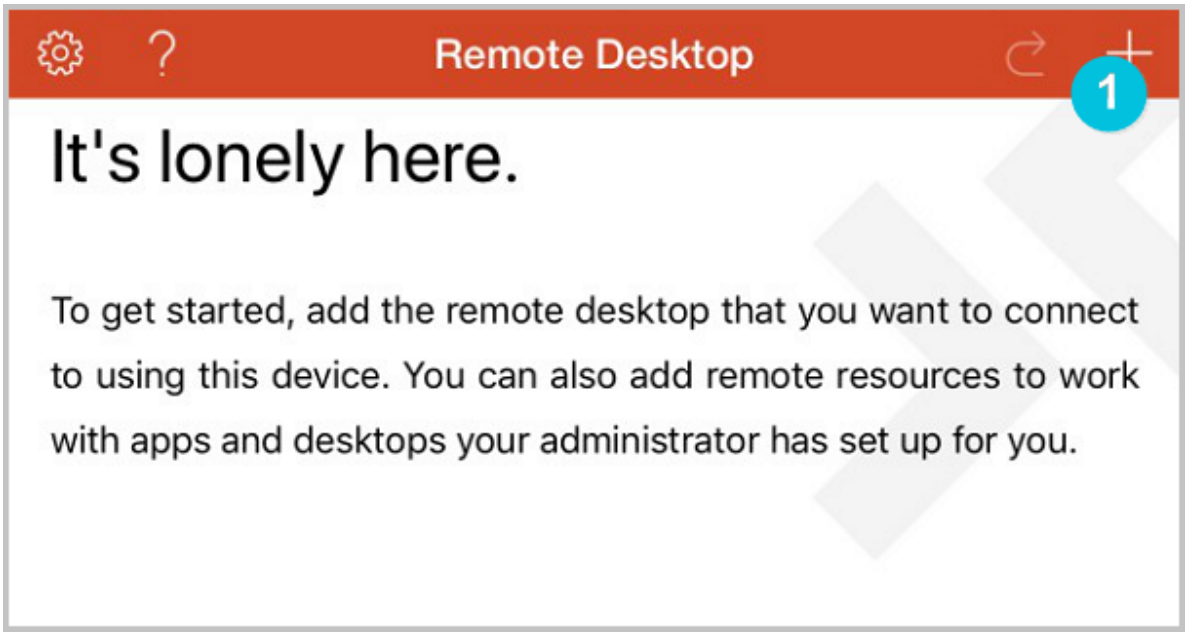
Network type	NIC	Rule direction	Authorization policy	Protocol type	Port range	Authorization type	Authorization object	Priority
	ion required							
Classic	Internet							

- 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.

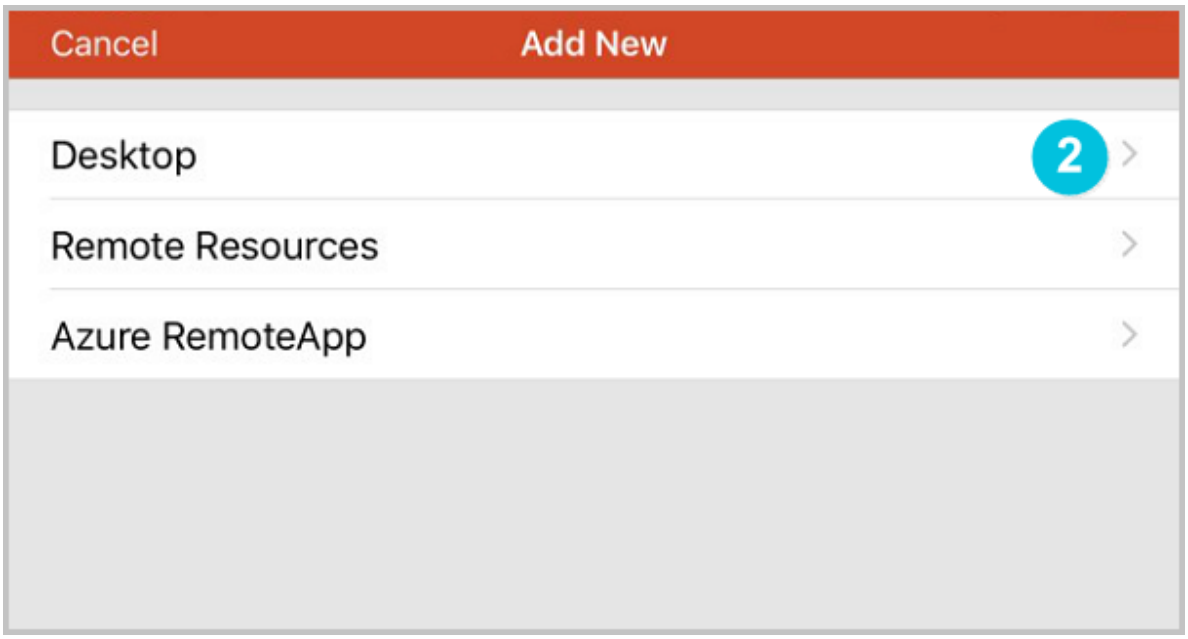
Procedure

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.

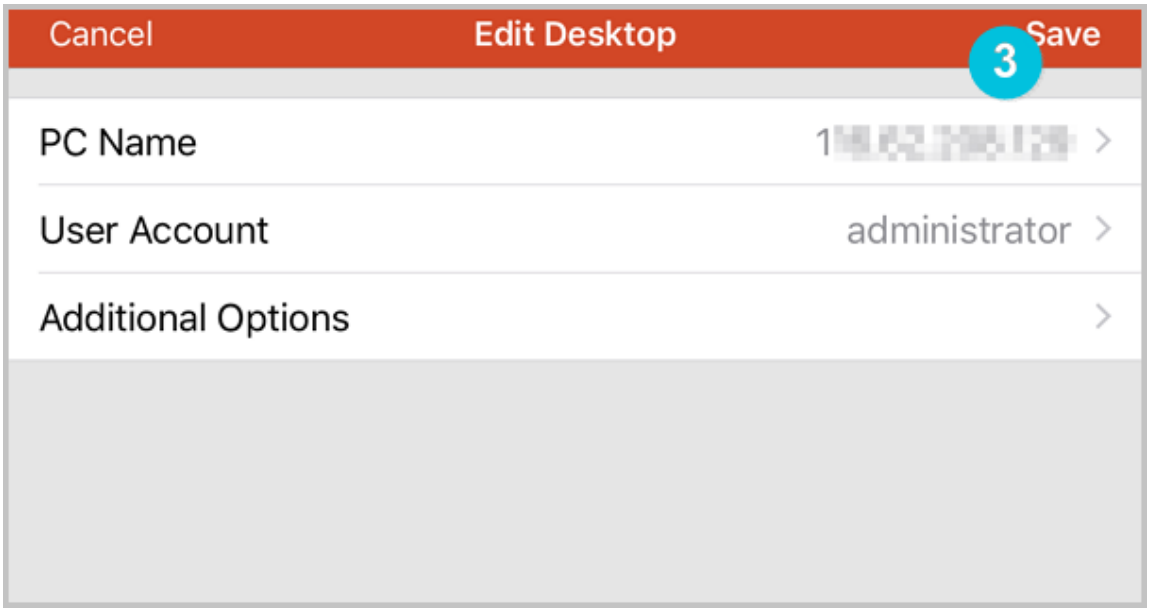


2. On the **Add New** page, select **Desktop**.

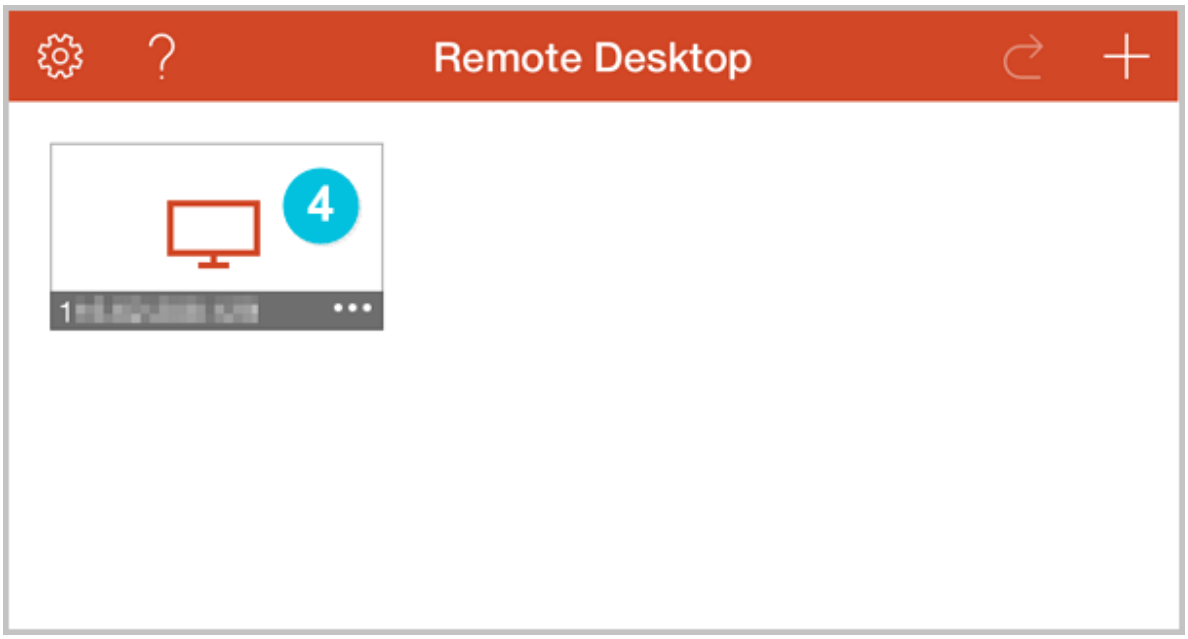


3. On the **Edit Desktop** page, type the connection information and tap **Save**. The following connection information is required:

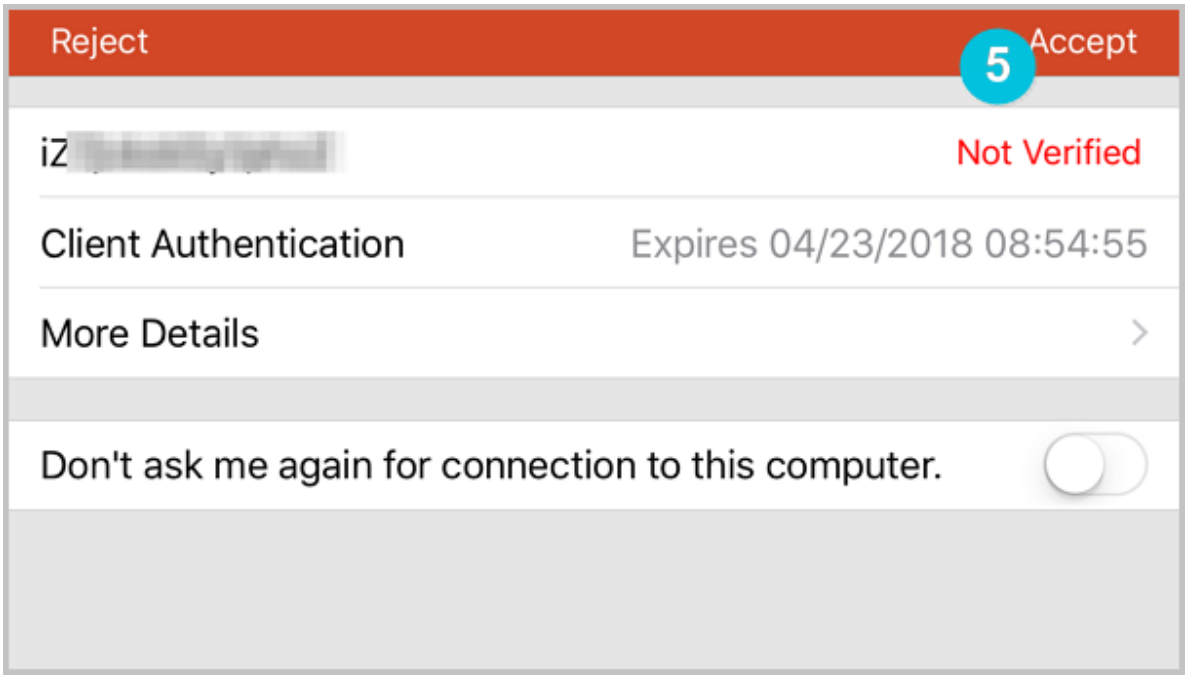
   - **`PC Name`**: Type the public IP address of the Windows instance to connect to.
   - **`User Account`**: Type the account name `administrator` and the logon password of the Windows instance.



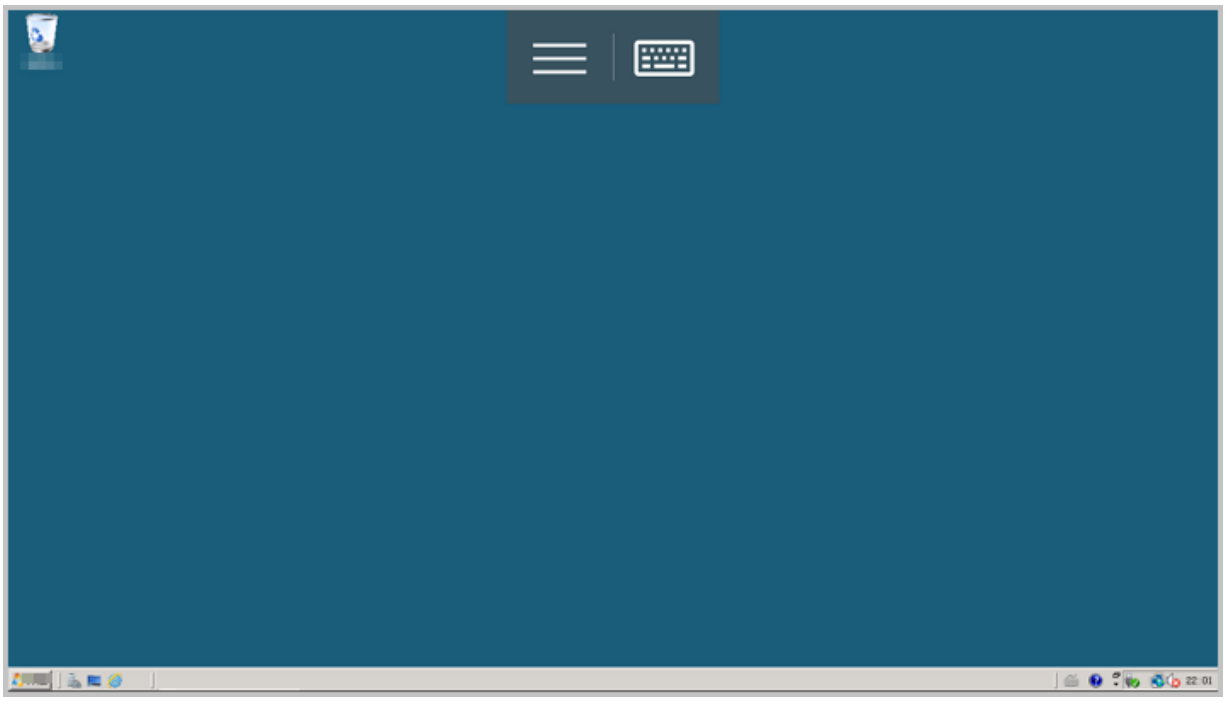
4. On the **Remote Desktop** page, tap the icon of a Windows instance.



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



Now, you are connected to the Windows instance.



10 Instances

10.1 Create an instance

10.1.1 Create an instance by using the wizard

This article describes how to create an instance using the console wizard. If you want to create a custom image from a snapshot of a system disk, and then use the custom image to create an ECS instance.

Prerequisites

- Before creating the ECS instance, you have completed the *preparation work*.
- To bind an SSH key pair when creating a Linux instance, you need to *create an SSH key pair* in the target region.
- To set the user-defined data, you need to prepare the *user data*.
- To authorize an instance to play a role, you need to *create an instance RAM role and grant it permissions*.

Procedure

1. Log on to the *ECS console*.
2. In the left-side navigation pane, click **Instances**.
3. On the **Instances** page, click **Create Instance**.
4. Follow these steps to finish **Basic Configurations**:
   a) Select a **billing method**: **Subscription**, **Pay-As-You-Go** or *Preemptible Instance*.

   Note:

   - To create an instance that is charged by week, you must select **Subscription** and set **Duration** to **1 week** at the bottom of the page. For more information about the billing methods, see *Billing method comparison*.
   - For how to create preemptible instances, see *Create preemptible nstances*.

   b) Select a region and zone. By default, a zone is assigned randomly. You can also select an applicable one. For more information about regions and zones, see *regions and zones*.

   Note:

   - After an instance is created, you cannot change its region and zone.

- Some instance type families are not supported in all regions. For more information, see *Create a compute optimized instance with GPUs*, *Create an f1 instance*, *Create an f2 instance*, *Create an SCC server instance*, and *Create an EBM instance*.

c) Select an instance type and specify the quantity of instances. The available instance type families are determined by the selected region. For the application scenarios of each instance type, see *instance type families*.

Note:

- The quota of Pay-As-You-Go or preemptible instances for your account is shown on the page.
- If you want to create a preemptible instance, you must specify the bidding price for one preemptible instance in this step.
- To use Elastic Network Interfaces (ENIs), select an enterprise-level instance type with no less than two vCPU cores or an entry-level instance type with no less than four vCPU cores. For more information about the maximum number of ENIs that can be attached to one instance, see *Instance type families*.
- To use an SSD Cloud Disk, select an I/O-optimized instance.

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

Note:

- To use an SSH key pair, select a Linux image.
- To use user-defined data, select an image as instructed in *user data*.

e) Select storage devices:

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

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

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

  Note:

  The data disks added here have the following features:

  - The billing method is the same as that of the instance.
  - A Subscription data disk has to be released along with the instance, but a Pay-As-You-Go data disk can be set to being released along with the instance.

- If you have selected an instance type family that has local disks (such as i1, d1, or d1ne), the local disk information is displayed. You cannot specify the quantity or category of local disks, which is determined by the selected instance type. For information about the local disks corresponding to various instance types with local disk, see *instance type families*.

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

   a) Select a network:

   - **VPC**: Must select a VPC and a VSwitch. If you do not have a VPC and a VSwitch, you can use the default ones.
   - **Classic network**: If you purchased the ECS instance for the first time after June 14, 2017, 17:00 (UTC + 8), you can no longer select a classic network.

   b) Configure the Internet bandwidth:

   - To assign a public IP address to the instance, select **Assign public IP**. Then, select **PayByTraffic** or **PayByBandwidth** as the network billing method and specify the bandwidth. For public IP addresses assigned in this way, you cannot unbind them

from the instance. For more information about network billing, see *Billing of network bandwidth*.

- If your instances do not need to access the Internet or your VPC instances *use an Elastic IP (EIP) address to access the Internet*, you do not need to assign a public IP. You can unbind an EIP address from an instance.

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

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

> **Note:**
> By default, the ENI is released along with the instance. You can detach it from the instance in the *ECS console* or by using the *DetachNetworkInterface* interface.

6. (Optional.) Click **Next: System Configurations** to finish the following configuration:
   - Select and set logon credentials. You can choose to *set the credentials after creating an instance* or do it now. Select a credential based on the image:
     - Linux: You can select a password or SSH key pair as a logon credential.
     - Windows: You can only select a password as a logon credential.
   - Specify the instance name, which is displayed in the ECS console, and the host name, which is displayed inside the guest operating system.
   - Set the advanced options:
     - Instance RAM role: Assign a RAM role to the instance.
     - UserData: Customize the startup behaviors of an instance or pass data into an instance.
7. (Optional) Click **Next: Grouping** to manage instances by group. You can add tags to instances to simplify future management. If you are an enterprise user who has enabled Resource Management Services and created resource groups, you can manage instances by resource group.
8. Confirm the order:
   - In the **Selected Configurations** area, confirm all the configurations. You can click the edit icon to re-edit the configuration.
   - (Optional) If the billing method is **Pay-As-You-Go**, you can **set the automatic release time**.

- (Optional) If the billing method is **Subscription**, you can set the duration and select whether to enable **auto renewal**.
- Confirm the configuration costs. The billing methods for an instance and Internet bandwidth determine the displayed cost information, as shown in the following table.

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

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

9. Click **Create Instance**.

Result

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

What's next

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

10.1.2 Create an instance of the same configuration

Context

To meet your growing business demands of having more ECS instances of the same configurations, use the **Buy Same Type** feature.

Procedure

1. Log on to the *ECS console*.
2. Select a region.
3. In the left-side navigation pane, click **Instances**.
4. Find your ECS instance, and in the **Actions** column, select **More** > **Buy Same Type**.
5. On the **Buy Same Type** page, confirm the selected configurations in the **Overview** section. If you want to change other configurations, select **View More** to change the billing method, security group, network billing method, bandwidth, logon credential, or instance name.
6. To purchase a Subscription ECS instance, you can change the **Purchase Time**.
7. Set the **Amount**.
8. Read and confirm the **ECS Service Terms and Product Terms of Service**.

9. To purchase a Subscription instance, click **Create Order**. To purchase a Pay-As-You-Go instance, click **Activate**.

10.1.3 使用自訂鏡像建立執行個體

如果您要建立一個執行個體，與現有的某個執行個體擁有完全相同的作業系統、應用程式和資料，您可以使用這個執行個體建立自訂鏡像，並根據這個自訂鏡像建立執行個體。採用這種方法可以提高工作或交付效率。

背景信息

- 如果自訂鏡像與要建立的執行個體在同一個地域，您需要使用以下任一方法建立自訂鏡像：
  - 匯入鏡像
  - 使用執行個體建立自訂鏡像
  - 使用快照建立自訂鏡像
- 如果自訂鏡像與要建立的執行個體在不同的地域，您需要將自訂鏡像複製到目標地域。詳細資料，請參見 複製鏡像。
- 如果鏡像屬於不同的帳號，必須先完成鏡像共用。詳細資料，請參見 共用鏡像。

操作步驟

1. 登入 *ECS*管理主控台
2. 在左側導覽列中，單擊 執行個體。
3. 在 執行個體列表 頁的右上方，單擊 建立執行個體。
4. 在建立執行個體頁面，按 建立*ECS*執行個體 配置資訊，但是需要注意以下配置：
   - 地域：必須選擇鏡像所在的地域。
   - 鏡像：選擇 自訂鏡像 或者 共用鏡像，並在下拉式清單中選擇需要的鏡像。

     说明：
     如果您選擇的自訂鏡像中包含了一個或多個資料盤快照，系統會自動根據這些快照建立相同數量的雲端碟作為資料盤，每個雲端碟大小與對應的快照相同。您可以增加雲端碟容量，但不能縮小。
5. 確認訂單。

10.1.4 Create an instance of ga1

Image description

The GPU visually calculates the ga1 specification Family instance using AMD's s5150 series GPU. Ali cloud and AMD work together to optimize the GPU drivers that you need to use Pre-installed-driven images in the **mirror market** , respectively:

- Centos Version 7.3 comes pre-installed amd GPU drives
- Ultimate pre-installed amd GPU driver
- Amd gpu driver pre-installed Windows 2008 r2
- Amd gpu driver pre-installed Windows 2008 r2

Create an instance

You can create a ga1 specification Family instance as described in *Create an ECS instance*. When selecting a configuration, you need to note the following:

- **Network**: select **VPC**. Because the current GPU-rendered ga1 Spec family instances only support proprietary networks (VPCs).
- **Instance**: Select **Series III** , **GPU Visually computed ga1**.
- **image**: Select **image market**, and click **Select (including operating system) from the mirror market**. Enter `GPU` or `AMD in the pop-up box of the mirror market` Search the image.

> **Note:**
> We recommend you purchase or subscribe to these images. You can find an instance later when you create it from **purchased image** or **subscribed image** .

Precautions:

- The driver used for GPU visualization computing of ga1 instances is of the optimized version driven by the cooperation between Alibaba Cloud and AMD. It is currently available only through image output provided by Alibaba Cloud. No driver download links are provided, and the customer is not currently supported to install the driver on his own.
- If the GPU driver-related components are uninstalled or deleted and the driver does not work properly, you need to *replace the system disk* to restore the GPU functions.

Warning:
This operation will cause data loss.

- When creating a GPU visualization to calculate a ga1 instance, selecting other images will cause the instance driver to fail to work properly. Users need to *replace the system disk* to reselect the pre-installed AMD GPU-driven image.
- For Windows systems, a **remote connection** to the Alibaba cloud cannot be used, and the **management terminal** always displays a black screen or stay in the startup interface. Please enter a system via another protocol, such Windows Remote Desktop Connection (RDP).
- The Remote Desktop Connection (RDP) protocol, which comes with Windows, does not support DirectX, OpenGL, and other related applications. You need to install VNC services and clients, or other supported protocols, such as PCOIP and XenDeskop HDX 3D.

10.1.5 建立GPU計算型執行個體

建立執行個體

您可以按照 建立*ECS*執行個體 的描述建立gn4、gn5或gn5i執行個體，建立時需要注意以下配置。

- 地域：不同的執行個體規格類型系列供應的地域資訊不同。如下所示：
  - gn4：華北2（可用性區域A）、華東2（可用性區域B）、華南1（可用性區域C）
  - gn5：華北2（可用性區域C、E）、華北5（可用性區域A）、華東1（可用性區域G、F）、華東2（可用性區域D、B、E）、華南1（可用性區域D）、香港（可用性區域C、B）、亞太東南1（可用性區域B、A）、亞太東南2（可用性區域A）、亞太東南3（可用性區域A）、亞太東南5（可用性區域A）、美國西部1（可用性區域B、A）、美國東部1（可用性區域B、A）、歐洲中部1（可用性區域A）

    說明：
    如果您要在gn5執行個體上部署NGC（NVIDIA GPU CLOUD）環境，選擇地域時請參見 在*gn5*執行個體上部署*NGC*環境。
  - gn5i：華北2（可用性區域C、E、A）、華東1（可用性區域B）、華東2（可用性區域D、B）、華南1（可用性區域A）

  如果ECS建立頁面顯示的地域和可用性區域資訊與上述描述不符，以ECS建立頁面上顯示的資訊為準。
- 鏡像：

- 如果您需要安装GPU驱动和CUDA库，可以选择以下任一种方式：
  - 选择 公共镜像 中的CentOS 64位（目前提供的所有版本都支援）、Ubuntu16.04 64位或SUSE Linux Enterprise Server 12 SP2 64位镜像，並選擇 **自動安裝GPU驅動**。再選擇需要的CUDA库和GPU驱动的版本。

    说明：

    - 您可以根據您的業務需要選擇合適的GPU驅動版本。如果是新業務系統，建議您在下拉式功能表中選擇最新的GPU驅動版本。
    - 如果選擇 **自動安裝GPU驅動**，**系統配置** 的 **進階選項** 中會自動生成 **執行個體自訂資料**，即自動安裝CUDA庫和GPU驅動的shell指令碼或直譯式程式。執行個體第一次啟動後，cloud-init會自動執行指令碼或直譯式程式，自動安裝GPU驅動。更多資訊，參見 自動安裝*GPU*驅動指令碼或直譯式程式注意事項。

  - 選擇 **鏡像市場**，並搜尋 `NVIDIA`，在搜尋結果中選擇需要的鏡像。目前只支援CentOS 7.3和Ubuntu 16.04。
- 如果gn4、gn5或gn5i執行個體要用於深度學習，可以選擇預裝深度學習架構的鏡像：選擇 **鏡像市場**，並搜尋 深度學習，在搜尋結果中選擇需要的鏡像。目前只支援Ubuntu 16.04和CentOS 7.3。
- 除上述以外的其他鏡像，執行個體建立完成後，自行 下載並安裝*GPU*驅動。

- 執行個體：選擇 **異構計算GPU/FPGA > GPU計算型**，按需求選擇合適的執行個體規格。
- 網路：選擇 **專有網路**。
- 公網頻寬：根據您的實際需要選擇頻寬。

  说明：

  如果使用Windows 2008 R2鏡像，GPU驅動安裝生效後，您不能使用控制台的 遠端連線 功能串連gn4、gn5或gn5i執行個體，所以，您必須選擇 **分配公網IP地址**，或者建立執行個體後綁定EIP 。
- 登入憑證：根據實際需求設定登入憑證。

  说明：

建議您不要選擇 建立後設定。執行個體建立成功後，GPU驅動安裝成功之前，如果您需要登入執行個體，必須重設密碼或者綁定SSH金鑰組，需要重啟執行個體使修改生效，而重啟操作會導致GPU驅動安裝失敗。

- 執行個體自訂資料：如果選擇了 自動安裝**GPU**驅動，這裡會顯示自動安裝CUDA庫和GPU驅動的shell指令碼或直譯式程式。請您仔細閱讀指令碼或直譯式程式內容和注意事項。

查看自動安裝GPU驅動進程

如果您選擇了 自動安裝**GPU**驅動，執行個體建立完成後，您可以 遠端連線執行個體，通過安裝日誌 `/root/nvidia_install.log`查看GPU驅動的安裝進程。

说明：

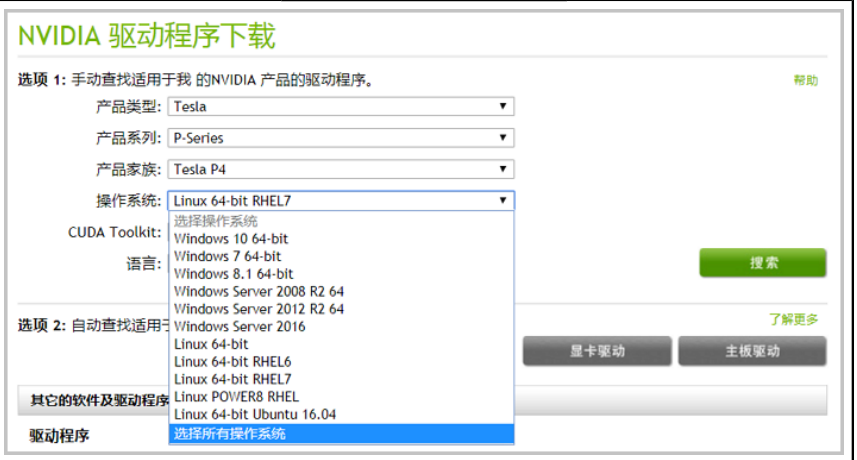
GPU驅動安裝完成前，您不能操作GPU，也不能安裝其他GPU相關軟體，以免自動安裝失敗。

下載並安裝GPU驅動

如果使用沒有預裝GPU驅動的鏡像，您必須為執行個體安裝GPU驅動。操作步驟如下：

1. 獲取GPU驅動安裝包：
   a. 進入 *NVIDIA* 官網。
   b. 手動尋找適用於執行個體的驅動程式，並單擊 搜尋。篩選資訊說明如下表所示。

	gn4	gn5	gn5i
產品類型	Tesla	Tesla	Tesla
產品系列	M-Class	P-Series	P-Series
產品家族	M40	Tesla P100	Tesla P4
作業系統	根據執行個體的鏡像選擇對應的版本。如果下拉式清單中沒有顯示伺服器作業系統，請單擊下拉式清單底部的 選擇所有作業系統。		

	gn4	gn5	gn5i
			

c. 確認無誤後，單擊 下載 按鈕。

2. 安裝GPU驅動：

- Windows執行個體：直接雙擊安裝GPU驅動。
- Linux執行個體：按以下步驟安裝驅動

1. 下載並安裝kernel對應版本的kernel-devel和kernel-header包。
2. 運行以下命令，確認已經完成下載並安裝kernel-devel和kernel-header包：

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

以CentOS 7.3為例，如果出現以下類似資訊，表示已經完成安裝。

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

3. 按NVIDIA官網GPU驅動下載頁的 其他資訊 描述安裝GPU驅動。

以Linux 64-bit Ubuntu 14.04為例：

安装**GRID**驅動

如果gn5或gn5i執行個體需要支援OpenGL圖形顯示，必須安裝GRID驅動，具體操作，請參見 在*gn5*或*gn5i*執行個體中安裝*GRID*驅動。

注意事項

遠端連線功能

對於Windows 2008 R2及以下版本，GPU驅動安裝生效後，控制台的 遠端連線 功能不可用，管理終端 會始終顯示黑屏或停留在啟動介面。請您通過其他協議進入系統，如Windows自帶的遠端連線（RDP）。

Windows自帶的遠端連線（RDP）協議不支援DirectX、OpenGL等相關應用，您需自行安裝VNC服務和用戶端，或其他支援的協議，例如PCOIP、XenDeskop HDX 3D等。

自動安裝**GPU**驅動指令碼或直譯式程式

關於自動安裝GPU驅動的shell指令碼或直譯式程式，注意事項如下：

- 該指令碼或直譯式程式會自動下載並安裝NVIDIA GPU的驅動和CUDA庫。
- 因執行個體規格的內網頻寬和vCPU核心數不同，實際自動安裝時間為4.5分鐘 ~ 10分鐘不等。安裝GPU驅動時，您不能操作GPU，也不能安裝其他GPU相關軟體，以免自動安裝失敗。
- 自動安裝結束後，執行個體自動重啟，使驅動生效。

- 指令碼或直譯式程式會自動開啟GPU驅動的 **Persistence Mode**，並將該設定添加到系統自啟動指令碼或直譯式程式中，確保執行個體重啟後還能預設開啟該模式。該模式下GPU驅動工作更穩定。
- 更換作業系統 時，您需要注意以下資訊：
  - 如果原來的鏡像是Ubuntu16.04 64位或SUSE Linux Enterprise Server 12 SP2 64位，換成其他鏡像後，無法自動安裝GPU驅動。
  - 如果原來的鏡像是CentOS的某個版本，換成其他版本的CentOS鏡像後，GPU驅動能正常安裝。
  - 如果換成其他不支援自動安裝GPU驅動指令碼或直譯式程式的鏡像，無法自動安裝GPU驅動。
- 安裝過程中會生成相應的安裝日誌，日誌存放路徑為 `/root/nvidia_install.log`。您可以通過日誌查看驅動安裝是否成功。如果失敗，您可以通過日誌查看失敗原因。

10.1.6 Create an instance of a bid

You can create a bid instance on the ECS console. For specific actions and instructions for creating an instance, refer to the Personal Edition ECs quick start. This article mainly introduces several matters that need to be taken into account when creating a spot instance.

- Billing method: You should select an instance of bidding.
- Bid strategy: Set up the highest bid example, set the maximum hour price you would like to pay for an instance specification, when your bid is higher than the current market rate, the instance runs. Eventually, it will be traded according to the market price. You need to enter the highest possible instance price.
- Select or enter the number of tables you want to purchase.
- Click, buy now, and if the price you are giving is higher than the current deal, an instance is created for you based on your settings.

After the bid instance has been created successfully, you can log on to the ECS console, view the bid instance information that you just created in the list of instances. The bid instance is marked By quantity-spot instance. When you enter the instance details page, When you pay for informatio n, the section displays Bid policy.

10.1.7 Create an f1 instance

This article describes how to create an f1 instance.

Procedure

Follow the steps described in *create an ECS instance*, but before doing so, consider the following:

- **Region**: Select **China East 1 (Hangzhou)** > **China East 1 Zone F**.
- **Instance Type**: Select **Heterogeneous Computing** > **FPGA** > **Compute**. And select the appropriate F1 instance type.
- **Image**: In the **image market**, select **FaaS F1 image**.

  Note:
  You can find quartus17.0, vcs2017.3, dcp sdk in the *opt* directory.
- **Network**: Select **VPC**, and select a created VPC and VSwitch.

After an f1 instance is created, *connect to the instance*, and run the following command to check whether the License is configured.

```
echo $LM_LICENSE_FILE #To check whether the variable is set.
```

Best practices

See best practices of f1 instances:

- *Use OpenCL on an f1 instance*
- *Using f1 RTL (Register Transfer Level)*

10.1.8 Create an f2 instance

This article describes how to create an f2 instance.

Procedure

Follow the steps described in *create an ECS instance*, but before doing so, consider the following:

- **Region**: Select **China East 1 (Hangzhou)** > **China East 1 Zone B**.
- **Instance Type**: Select **Heterogeneous Computing** > **FPGA Compute**.
- **Image**: In the **image market**, select FA**FaaS F2 image**.
- **Network**: Select **VPC**, and select a created VPC and VSwitch.

Best practice

Use OpenCL on an f2 instance

10.1.9 Create an f3 instance

This article describes how to create an f3 instance.

> **Note:**
> Due to limited computing resources, we recommend that you use instances with four cores or more for testing, for example, instance type family g5-ecs.g5.2xlarge (8 vCPU core, 32 GiB). Create an f3 instance when you need to download the image to the FPGA chipset.

Prerequisite

The f3 instance type family is currently available for testing by invited users. *Open a ticket* to request a free f3 instance test.

Procedure

For more information about how to create an f3 instance, see *create an ECS instance*. When you create an f3 instance, follow these guidelines:

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

  > **Note:**
  > - During the testing phase, f3 instances are available for free. Other ECS resources including cloud disks, public network bandwidth, and snapshots will incur usage fees.
  > - f3 instances are not available as preemptible instances.

- **Region**: Select **China East 2 (Shanghai)**.
- **Instance Type**: Select **Heterogeneous Computing** > **FPGA Compute**, and then select your required instance type.
- **Image**: Click **Shared Image**, and then select the specified image.

  > **Note:**
  > We have provided a Xilinx image for development use. At present, the image can only be retrieved through image sharing.

- **System Disk**: We recommend that you allocate a 200 GiB Ultra Cloud Disk for the system image.
- **Network**: Select **VPC**.

Best practices

For the best practices of f3 instances, see *use RTL compiler on an f3 instance*.

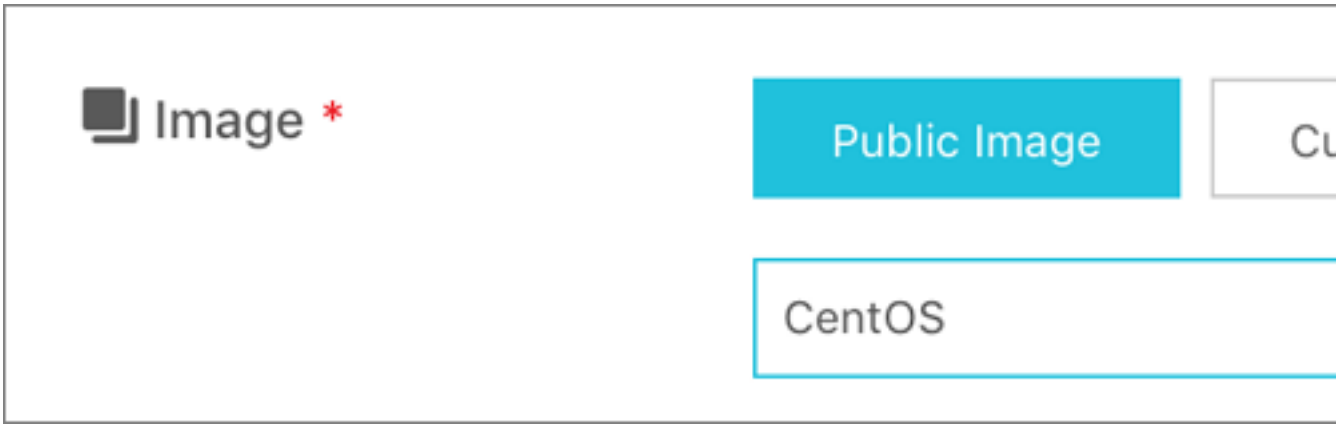
10.1.10 Create an SCC server instance

An SCC is based on ECS Bare Metal (EBM) Instance. With the help of the high-speed interconnectivity of RDMA (Remote Direct Memory Access) technology, SCC greatly improves network performance and increases the acceleration ratio of large-scale clusters. Therefore, SCC has all the advantages of EBM Instances, and provides high-quality network performance featuring high bandwidth and low latency. For more information, see *ECS Bare Metal Instance and Super Computing Clusters*.

This article describes some consideration when you create an SCC instance. For more information about creating an SCC instance, see *create an ECS instance*.

You must consider the following when creating an SCC instance:

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



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

10.1.11 Create an EBM instance

You can see *create an ECS instance* to create an EBM instance. When creating an EBM instance, consider the following:

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

Operating system	Image
Linux	• CentOS 7.2/7.3/7.4/6.9/6.8 64-bit • Ubuntu 14.04/16.04 64-bit • Debian 8.9/9.2 64-bit • OpenSUE 42.3 64-bit • SUSE Linux Enterprise Server 12 SP2 64-bit • Aliyun Linux 17.1 64-bit
Windows	• 2016 Data Center Edition 64-bit Chinese Edition • 2016 Data Center Edition 64-bit English Edition • 2012 R2 Data Center Edition 64-bit Chinese Edition • 2012 R2 Data Center Edition 64-bit English Edition

- **Storage**: EBM instances support up to 16 data disks. You can add a data disk here, or you can *add a disk* after the instance has been created, and *mount the data disk*.
- **Network**: Supports VPC only.

10.2 Check instance information

Through the console, you can check all the ECS instances you own. You can check the:

- On the *View all ECS instances under your account on the Overview page* or *View all ECS instances under your account on the Overview page* page, all ESC instances in all regions and their status under your account can be viewed.

- View all the ECS instances in a specified region on the Instance List page For details, see *View the information of ECS instances on the Instance List page*.
- Detailed information of any ECS instance on its **Instance Details** page. For details, see *View details of an ECS instance on Instance Details page*.

View all ECS instances under your account on the Overview page

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

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

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

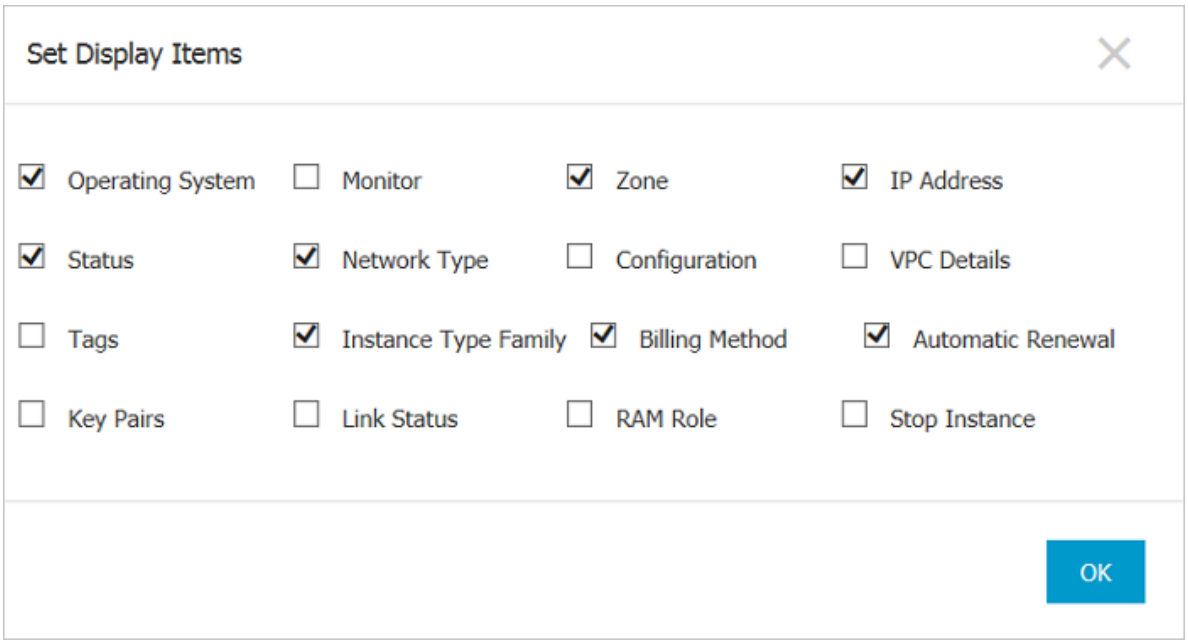
View the information of ECS instances on the Instance List page

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.

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. In the dialog box of **Set Display Items** , select the instance information to be displayed and click **OK**.



View all ECS instances under your account on the Overview page

On the **Instance Overview** page of the **Resource Overview**, you can visualize all instances in a single region from multiple perspectives. You can filter instances based on network type (including proprietary networks and classic networks) or *Add a tag to resources* and export all statistics in one click.

Follow these steps to go to the **Instance Overview** page:

1. Log on to the *ECS Console*.
2. In the left navigation bar, click **Overview**.
3. Above **Common Actions**, click **Resource Overview**.

4. In the left navigation bar, click **Instance Overview**.

Instance overview page information

On the **Examples Overview** page, you can view and export instance information by region, network type, and tag, including status, payment type, availability zone, instance types, number of mirrored instances, and the number of instances created in the last month. The exported instance information is a CSV file that includes all instance information displayed on the current page.

See the last 30 days expired instance

If you have a prepaid instance, you can also view instances that have expired in the last 30 days , or even instances that will expire in different periods in the future, to provide you with more basis for renewals and budgets. As shown in the figure below, one instance will expire in the next 15 days to 30 days.

See instance type distribution

You can use **instance type** distribution graph which shows the proportion of instance types for different architectures or levels, to help you determine if the current instance type ratio is appropriate. For example, in a large mature enterprise, if the proportion of entry-level instances is too high, it may not be appropriate; or all the entry-level instances are Burstable instances(t5 instances), you may need to re-examine whether the configuration is suitable for the current business demand for CPU.

View the instance image distribution

You can use **instance image distribution** graph to view the distribution of the mirror, which is easy to manage. For example, if you deploy the same application on multiple instances, these instances should use the same image as much as possible. At this time, you can use **Instance Image Distribution**graph to confirm the image distribution.

View details of an ECS instance on Instance Details page

You can navigate to the **Instance Details** page to view detailed information of an 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 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**, the **Shared Block Storage**, the **Instance Snapshots**, the **Security Groups**, or the **Security** page to view resources related to this instance.

10.3 Change the operating system

Use the management console to convert the instance OS to your preferred OS.

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 wan to use a public image, see *Change a system disk (public image)*.

Note:
Regions outside of mainland China do not currently support transition between Linux and Windows OSs. If your instance is hosted in one of these regions, you are not allowed to change the operating system between Windows and Linux. You can only change the version of Windows OS, or replace one Linux OS with another Linux OS.

10.4 Change configurations

10.4.1 Overview of configuration changes

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

Upgrade or downgrade instance configurations

You can only upgrade or downgrade the configurations of vCPU and memory (实例规格族) simultaneously by changing instance types. The methods to change an instance type vary according to the billing method of the instance:

- Prepaid:

- Upgrade: See 预付费实例升级配置 . The new configurations are effective after you 重启实例 in the console or by using the *RebootInstance* interface. Applies to year-to-month instances and paid-per-week instances.
- Downgrade: See 续费降配 . You can downgrade the configuration of an instance when you renew the instance. The new configuration takes effect after you 重启实例 in the ECS console within the first seven days of the new billing cycle. Renewal Scaledown can only be used for Subscription instances.

- Pay-As-You-Go:

  See 按量付费实例变更实例规格 . You must stop the instance to use this feature.

  Note:
  Stopping an instance disrupts your business traffic. Proceed with caution.

Adjust Internet bandwidth

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

Billing method of instances	Do you want to upgrade your bandwidth permanently?	Effective immediately	Available feature	Available feature
Subscription	Yes	Yes	预付费实例升级配置	Only applicable to VPC-Connected ECS instances to which no EIP addresses are attached or classic network -connected ECS instances . The Internet and intranet IP addresses remain unchanged after you upgrade your configurations.

Billing method of instances	Do you want to upgrade your bandwidth permanently?	Effective immediately	Available feature	Available feature
Subscription	No	Yes	临时升级带宽	Only applicable to VPC-Connected ECS instances to which no EIP addresses are attached or classic network -connected ECS instances . Temporarily adjust bandwidth at any time during the current life cycle of the instance , bandwidth automatically drops back to its original value after completing the task. The Internet and intranet IP addresses remain unchanged after you upgrade your configurations.
Prepaid	Yes	Effective from next billing cycle	续费降配	Adjust bandwidth in the new billing cycle. Renewal Scaledown can only be used for Subscription instances. • If the bandwidth of an instance

Billing method of instances	Do you want to upgrade your bandwidth permanently?	Effective immediately	Available feature	Available feature
				is charged for a fixed bandwidth, the public network bandwidth can only be reduced. The Internet and intranet IP addresses remain unchanged after you upgrade your configurations. • If the bandwidth of the instance is based on the Usage flow meter, you can increase or decrease the peak of the public network bandwidth. When the Internet bandwidth is set to 0 Mbit/s, the Internet IP address of a VPC-Connected instance is released in the new billing cycle, but that of a classic network-connected ECS

Billing method of instances	Do you want to upgrade your bandwidth permanently?	Effective immediately	Available feature	Available feature
				instance is retained.
Pay-As-You-Go or Subscription	Yes	Yes	按量付费实例修改公网带宽	Only applicable to those *VPC-Connected instances* that are bound to. You can adjust the Internet bandwidth on an EIP address at any time.
Pay by volume	Yes	Yes	按量付费实例修改公网带宽	Applies only to a VPC type ECs instance that has a public network IP address assigned to it, or to a classic network type ECs instance. You can adjust bandwidth at any time in the current lifecycle of an instance.

Assign a public IP address

Assign a public IP address to an ECS instance while 步骤 *2#创建ECS实例* . If you skip it, you can even assign after an ECS instance is created. However, the feature is only available for Prepaid instances. For more information, see the following table.

Feature	Effective immediately	Description
预付费实例升级配置	Yes	Only applicable to VPC-Connected ECS instances to which no EIP addresses are attached or classic network-

Feature	Effective immediately	Description
		connected ECS instances. Set the Internet bandwidth to a non-zero value to assign a public IP address.
临时升级带宽	Yes	Only applicable to VPC-Connected ECS instances to which no EIP addresses are attached or classic network-connected ECS instances. Set the Internet bandwidth to a non-zero value to assign a public IP address.
续费降配	Effective from next billing cycle	Only applicable to VPC-Connected ECS instances to which no EIP addresses are attached or classic network-connected ECS instances. Select use flow meter fee, and then set the peak bandwidth to a non-zero value, assign the public network IP address to the instance.

Change public network bandwidth Billing

Depending on how the instance is billing, you can change the public network bandwidth billing in different ways:

- Only applicable to VPC-Connected ECS instances to which no EIP addresses are attached or classic network-connected ECS instances.
  - Charge by use flow meter-> charge by fixed bandwidth: You can use 预付费实例升级配置 Function, effective immediately. Applies to examples of package year, package month, and pay by week.
  - Charge by fixed bandwidth-> use flow meter fee: You can use 续费降配 Function to change the billing method of public network bandwidth at the same time. The change takes effect after entering the new billing cycle. Renewal Scaledown can only be used for Subscription instances.

- Pay by volume: applies to a VPC type ECs instance or a classic network type ECs instance that has been assigned a public network IP address. You can use to change the public network bandwidth billing at any time. Effective immediately.

10.4.2 預付費執行個體升級配置

當前 預付費（包括訂用帳戶執行個體和按周付費執行個體）的執行個體規格無法滿足您的業務需求時，您能使用 **升級配置** 功能更新執行個體規格。同時，您也能使用這個功能完成以下操作：

- **轉換資料盤計費方式**：將隨用隨付資料盤轉為訂用帳戶資料盤。不能變更系統盤的計費方式。
- **永久升級基礎頻寬**：您能修改公網頻寬，適用於經典網路類型ECS執行個體和未綁定EIP的VPC類型ECS執行個體。如果您在建立執行個體時沒有購買公網頻寬，即沒有分配公網IP地址，可以使用這個功能為執行個體分配一個公網IP地址。
- **變更公網頻寬計費方式**：如果您當前的公網頻寬按使用流量計費（即 **流量頻寬**），可以使用這個功能將計費方式改為按固定頻寬計費（即 **固定頻寬**）。

費用

升級配置後，您需要為當前計費周期的剩餘時間補差價。

限制

使用 **升級配置** 功能有以下限制：

- 僅適用於預付費執行個體，包括訂用帳戶執行個體和按周付費執行個體。
- 您只能升級執行個體規格（包括執行個體vCPU核心數和記憶體容量），不能單獨升級其中一個配置。
- 部分執行個體規格不支援升級，詳見 執行個體規格類型系列。
- 僅經典網路執行個體和未綁定EIP的VPC類型ECS執行個體能通過這個功能修改公網頻寬，或修改公網頻寬的計費方式。
- 只能將資料盤的計費方式從隨用隨付轉為訂用帳戶，但是不能轉換系統盤的計費方式。
- 在當前計費周期內，如果您已經執行過 續費降配 操作，只有進入新的計費周期後，您才能升級配置。
- 升級執行個體規格，或者經典網路執行個體公網頻寬首次從0 Mbps升級到一個非零值後，您必須在控制台或使用API *RebootInstance* 重啟執行個體，新配置才能生效。

操作步驟

1. 登入 *ECS管理主控台*。

2. 在左側導覽列中，單擊 執行個體。
3. 選擇地域。
4. 選中需要升級配置的訂用帳戶執行個體或按周付費執行個體，並在 操作 列中，單擊 升降配。
5. 在 升降配嚮導 對話方塊中，選擇 升級配置，並單擊 繼續。
6. 在 升級配置 頁面上，完成以下任一個操作：
   - 選擇 執行個體規格。

     说明：
     執行個體規格能否升級以及支援升級到哪種規格，以介面上顯示的資訊為準。
   - 如果執行個體挂载云盘，您能選擇是否將資料盤的計費方式 轉換為訂用帳戶。
   - 如果執行個體為經典網路類型ECS執行個體或者未綁定EIP的VPC類型ECS執行個體，您能修改頻寬：選中 是否永久升級基礎頻寬，並設定升級後的頻寬。

     说明：
     如果您建立執行個體時沒有購買公網頻寬，即沒有分配公網IP地址，可以在這裡將公網頻寬設為一個非零值，從而分配公網IP地址。
   - 如果您當前的公網頻寬按使用流量計費（即 流量頻寬），可以使用這個功能將計費方式改為按固定頻寬計費（即 固定頻寬）。

7. 確認價格後，單擊 確定升配，並按頁面提示完成升配操作。
8. 升級執行個體規格，或者經典網路類型ECS執行個體公網頻寬首次從0 Mbps升級到一個非零值後，您必須在控制台或使用API *RebootInstance* 重啟執行個體，新配置才能生效。

> 说明：
> VPC類型ECS執行個體公網頻寬首次從0 Mbps變為一個非零值，不需要重啟執行個體。

10.4.3 隨用隨付執行個體變更執行個體規格

使用執行個體時，如果您發現執行個體配置超出或不能滿足您的應用需求，您可以變更執行個體規格，即記憶體和CPU配置。本文描述如何變更隨用隨付執行個體的執行個體規格。如果是預付費執行個體，請參見 升降配概述。

> 说明：
> 變更執行個體規格需要停止執行個體，會造成您的業務中斷。建議您在非業務高峰期時執行該操作。

使用限制

變更隨用隨付執行個體的規格有以下限制：

- 2次變更操作之間的間隔不得少於5分鐘。
- 不支援跨系列變更執行個體規格，比如系列I執行個體規格不能變更為系列II或系列III的執行個體規格。
- 系列III中，以下執行個體規格類型系列內或規格類型系列之間不支援變更：
  - GPU執行個體規格類型系列，包括：gn5、gn4、gn5i和ga1。
  - FPGA執行個體規格類型系列，包括f1。
  - 大資料型執行個體規格類型系列，包括d1和d1ne。
  - 本地SSD型執行個體規格類型系列，包括i1和i2。
- 系列III中，可以按下表所示變更執行個體規格。

系列III變配	ecs.sn1ne	ecs.sn2ne	ecs.mn4	ecs.se1ne	ecs.cm4	ecs.c4	ecs.se1	ecs.ce4	ecs.xn4	ecs.e4	ecs.n4
ecs.sn1ne	Y	Y	—	Y	Y	Y	Y	Y	—	—	—
ecs.sn2ne	Y	Y	—	Y	Y	Y	Y	Y	—	—	—
ecs.mn4	—	—	Y	—	—	—	—	—	Y	Y	Y

系列III 變配	ecs.sn1ne	ecs.sn2ne	ecs.mn4	ecs.se1ne	ecs.cm4	ecs.c4	ecs.se1	ecs.ce4	ecs.xn4	ecs.e4	ecs.n4
ecs.se1ne	Y	Y	—	Y	Y	Y	Y	Y	—	—	—
ecs.cm4	Y	Y	—	Y	Y	Y	Y	Y	—	—	—
ecs.c4	Y	Y	—	Y	Y	Y	Y	Y	—	—	—
ecs.se1	Y	Y	—	Y	Y	Y	Y	Y	—	—	—
ecs.ce4	Y	Y	—	Y	Y	Y	Y	Y	—	—	—
ecs.xn4	—	—	Y	—	—	—	—	—	Y	Y	Y
ecs.e4	—	—	Y	—	—	—	—	—	Y	Y	Y
ecs.n4	—	—	Y	—	—	—	—	—	Y	Y	Y

- 系列II中，可以按下表所示變更執行個體規格。

系列Ⅱ變配	ecs.n2	ecs.e3	ecs.n1	ecs.sn2	ecs.sn1
ecs.n2	Y	Y	Y	—	—
ecs.e3	Y	Y	Y	—	—
ecs.n1	Y	Y	Y	—	—
ecs.sn2	—	—	—	Y	Y
ecs.sn1	—	—	—	Y	Y

- 系列I內所有執行個體規格均可以互相變更配置。

说明：

以上表格中，Y表示執行個體規格類型系列之間允許變更配置，—表示執行個體規格類型系列之間不允許變更配置。

前提條件

您已經停止執行個體。

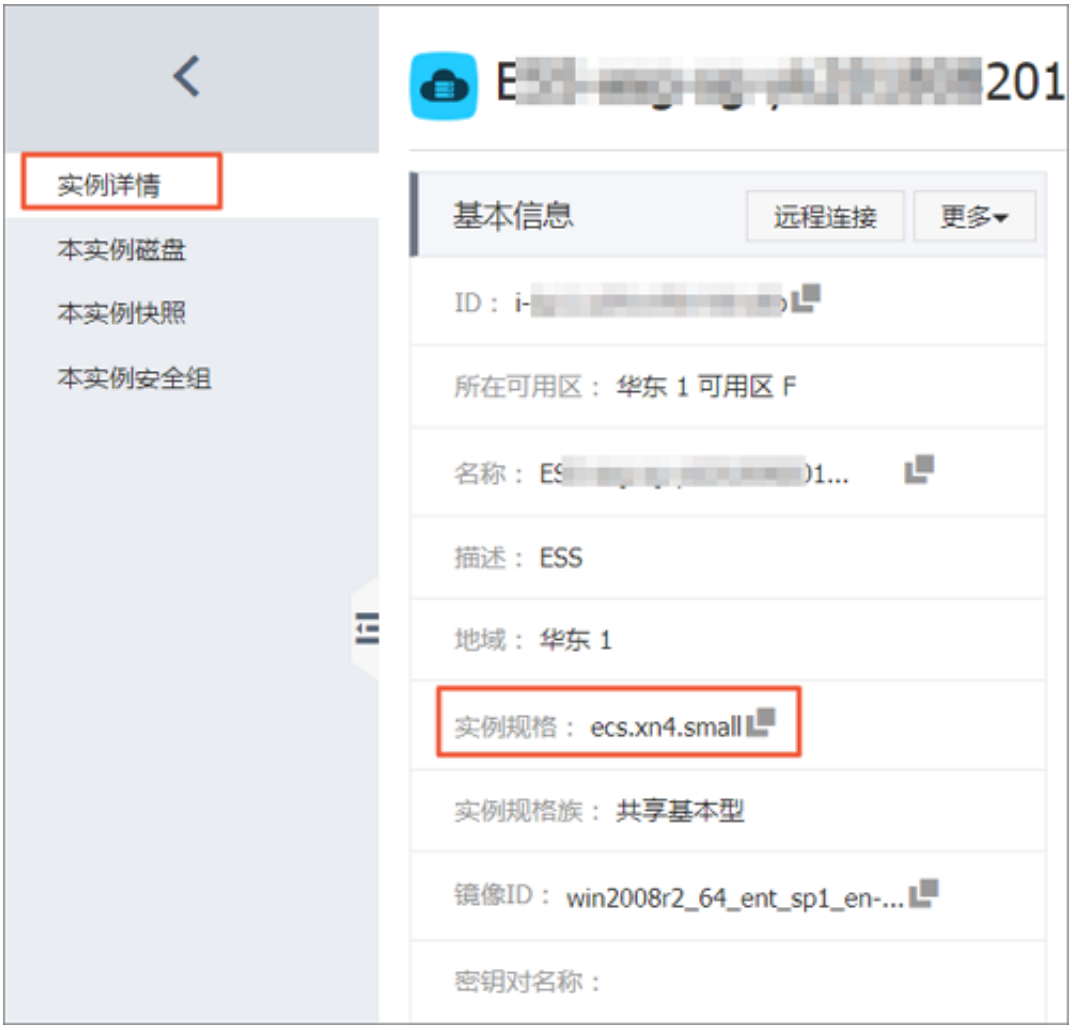
操作步驟

按以下步驟變更隨用隨付執行個體的記憶體和vCPU配置：

1. 登入 *ECS管理主控台*。
2. 在左側導覽列中，單擊 **執行個體**。
3. 選擇地域。
4. 選中需要變更配置的隨用隨付執行個體，並在 **操作** 列中，單擊 **更改執行個體規格**。
5. 在 **更改執行個體規格** 對話方塊中，選擇需要的執行個體規格，再單擊 **確定**。

> **说明：**
> 您可以在搜尋方塊裡輸入執行個體規格資訊即時篩選執行個體規格。

變更完成後，變更立即生效。您可以在 **執行個體詳情** 頁的 **基本資料** 部分查看執行個體規格資訊，如下圖所示。



之後，啟動執行個體恢複服務。

10.4.4 臨時升級頻寬

臨時升級頻寬 是指在預付費（包括訂用帳戶和按周付費）執行個體當前計費周期內，因某些業務原因在指定時間段內臨時提高公網頻寬，業務結束時自動降回到原來的公網頻寬。該功能不會提升執行個體的基礎公網頻寬。

您可以使用 **臨時升級頻寬** 功能完成以下任務：

- 臨時提升公網頻寬。

  > **说明：**
  > 基礎公網頻寬是指您在建立執行個體時購買的公網頻寬，或者通過 預付費執行個體升級配置 購買的公網頻寬。

- 如果 步驟 *2#建立ECS執行個體* 時未購買公網頻寬：通過這個功能臨時購買公網頻寬，會為執行個體分配公網IP地址。臨時頻寬升級結束後，分配的公網IP地址會自動釋放。

使用 **臨時升級頻寬** 功能時，需要注意：

- 臨時升級頻寬不會改變公網IP地址或私有IP地址。
- 在頻寬升級結束時間之前，如果手动续费 或自动续费 了執行個體，系統按基礎公網頻寬值續費。

使用限制

臨時升級頻寬功能有以下使用限制：

- 僅支援預付費執行個體，包括訂用帳戶執行個體和按周付費執行個體。
- 僅能臨時提升按固定頻寬計費的公網頻寬，不能提升按使用流量計費的公網頻寬。
- 該功能只能用於升級公網頻寬，不能降低公網頻寬。
- 如果在當前計費周期內已經執行過 續費降配 操作，只有進入新的計費周期後，您才能再臨時升級頻寬。
- 只能在當前計費周期內臨時升級頻寬，即 **頻寬升級結束時間** 不能晚於當前計費周期的到期時間。
- 如果您將經典網路執行個體的公網頻寬首次從0 Mbit/s臨時升級到一個非零值，必須在控制台或使用API *RebootInstance* 重啟執行個體，才能使新配置生效。

操作步驟

1. 登入 *ECS管理主控台*。
2. 在左側導覽列中，單擊 執行個體。
3. 選擇地域。
4. 選中訂用帳戶執行個體或按周付費執行個體，並在 操作 列中，單擊 升降配。
5. 在 升降配嚮導 對話方塊中，選擇 頻寬臨時升級，並單擊 繼續。
6. 在 頻寬臨時升級 頁面上，設定目標頻寬以及頻寬升級的適用時間段。

7. 單擊 去支付，並按照頁面提示完成升級操作。
8. 如果您將經典網路執行個體的公網頻寬首次從0 Mbit/s臨時升級到一個非零值，必須 在控制台或使用API *RebootInstance* 重啟執行個體，才能使新配置生效。

> 说明：
> VPC類型ECS執行個體的公網頻寬首次從0 Mbit/s變為一個非零值，不需要重啟執行個體。

相關操作

- 如果您需要提升基礎公網頻寬，可以 預付費執行個體升級配置。
- 如果您想要降低公網頻寬，可以 續費降配。
- 如果您需要一直保留分配給VPC類型ECS執行個體的公網IP地址，可以 公網*IP*轉換為*Elastic IP Address*。

10.4.5 Change EIP Internet bandwidth

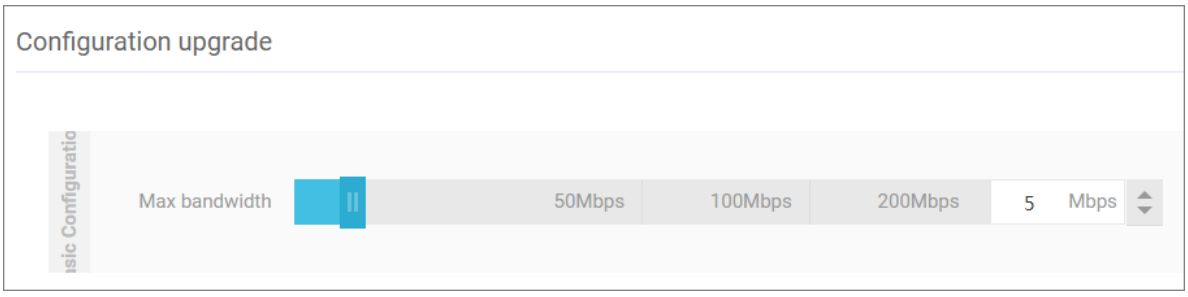
When using a pay-per-order instance, if you find that public network bandwidth does not meet or exceed your business requirements, you can select different ways to modify the public network bandwidth based on the type of network and the public network IP category of the instance, as shown in the following table.

Network type	Public Network IP category	Available feature
VPC	Elastic public network IP address	**Change configurations** > *Change bandwidth*
	Allocated public network IP address	**More** > *Change bandwidth by volume instance*
Classic network	Allocated public network IP address	**More** > *Change bandwidth by volume instance*

Change bandwidth

If the instance is a dedicated network (VPC) Paying instance, and the elastic public network IP (EIP) address is bound, you can change the network bandwidth of an EIP address by following 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 pay-per-order instance that is bound to the elastic public network IP address, and in the **action** column, click **Change configurations**.
5. In the **Change configurations wizard** dialog box, select **Change bandwidth**, and click **Continue**.
6. On the **Upgrade** page, set the new peak bandwidth.



7. Click **Activate**.

Change bandwidth by volume instance

Whether it's a proprietary network (VPC) or a classic network, if your pay-per-sum instance is assigned a public network IP address, you can **change the bandwidth using the on-order instance** to change public network bandwidth.

Note:

After a successful change, the second operation cannot be performed in 5 minutes.

Follow these steps to change public network bandwidth:

1. Log on to the *ECS Management Console*.
2. In the left-hand navigation bar, click **Instances**.
3. Select a region.
4. Locate the pay per volume instance, and in the **action** column, select **More** > **Change bandwidth by volume instance**.

   Note:
   You can also select multiple pay-per-order instances and, at the bottom of the list, select **More** > **Change bandwidth of Pay-As-You-Go instance**.
5. On the **Change bandwidth** by volume instance page, click **Bulk change**.
6. In the **Change bandwidth** dialog box, you can complete the following settings:
   - (Optional) change the billing method for public network bandwidth: Select `Fixed bandwidth`, or `Traffic bandwidth`.
   - (Optional) select the new public network bandwidth value.

     Note:
     If the public network bandwidth is set to 0 Mbps here, after the change is successful:
     - The public IP address of the private network instance is released immediately.
     - Classic Network instances no longer provide public network access, but the public network IP address is retained.

     When the modification is complete, click **OK**.
7. On the **Change bandwidth** by volume instance page, click **OK**.

The new public network bandwidth setting takes effect immediately after the change has been completed.

10.4.6 Instance type families that support upgrading instance types

This article describes the instance type families that support upgrading instance types.

Impacts

Upgrading instance types has the following impacts:

- Classic network instances:
  - For *phased-out instance types*, when a non-I/O optimized instance is upgraded to an I/O optimized instance, changes will be made to the private IP address, the driver name, and the software authorization code. For Linux instances, Basic Cloud Disks (`cloud`) will be recognized as `xvda` or `xvdb`, while Ultra Cloud Disks (`cloud_efficiency`) and SSD Cloud Disks (`cloud_ssd`) as `vda` or `vdb`.
  - For *available instance types*, changes will be made to the private IP address of the instance.
- VPC instances:

  For *phased-out instance types*, when a non-I/O optimized instance is upgraded to an I/O optimized instance, changes will be made to the driver name and the software authorization code. For Linux instances, Basic Cloud Disks (`cloud`) will be recognized as `xvda` or `xvdb`, while Ultra Cloud Disks (`cloud_efficiency`) and SSD Cloud Disks (`cloud_ssd`) as `vda` or `vdb`.

Instance type families that support upgrading instance types

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

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

Source instance type family	Target instance type (family)
g5, r5, c5, ic5	• g5, r5, c5, ic5 • sn1ne, sn2ne, se1ne, c4, cm4, ce4, hfc5, hfg5, re4, t5, n4, mn4, xn4, e4
sn1ne, sn2ne, se1ne	• sn1ne, sn2ne, se1ne

Source instance type family	Target instance type (family)
	• c4, cm4, ce4, hfc5, hfg5, g5, r5, c5, ic5, re4, t5, n4, mn4, xn4, e4
se1	• se1 • sn1, sn2, n1, n2, e3, sn1ne, sn2ne, se1ne, c4, cm4, ce4, hfc5, hfg5, g5, r5, c5, ic5, re4, t5, n4, mn4, xn4, e4
n4, mn4, xn4, e4	• n4, mn4, xn4, e4 • sn1, sn2, se1, n1, n2, e3, sn1ne, sn2ne, se1ne, c4, cm4, ce4, hfc5, hfg5, g5, r5, c5, ic5, re4, t5
re4	• re4 • sn1ne, sn2ne, se1ne, c4, cm4, ce4, hfc5, hfg5, g5, r5, c5, ic5, t5, n4, mn4, xn4, e4, ecs.se1.14xlarge
hfc5, hfg5	• hfc5, hfg5 • sn1ne, sn2ne, se1ne, c4, cm4, ce4, g5, r5, c5, ic5, re4, t5, n4, mn4, xn4, e4
gn4	gn4
gn5i	gn5i
gn6v	gn6v
t5	• t5 • sn1ne, sn2ne, se1ne, c4, cm4, ce4, hfc5, hfg5, g5, r5, c5, ic5, re4, n4, mn4, xn4, e4
t1, s1, s2, s3, m1, m2, c1, c2	• t1, s1, s2, s3, m1, m2, c1, c2 • sn1, sn2, se1, n1, n2, e3, sn1ne, sn2ne, se1ne, c4, cm4, ce4, hfc5, hfg5, g5, r5, c5, ic5, re4, t5, n4, mn4, xn4, e4
n1, n2, e3	• n1, n2, e3 • sn1, sn2, se1, sn1ne, sn2ne, se1ne, c4, cm4, ce4, hfc5, hfg5, g5, r5, c5, ic5, re4, t5, n4, mn4, xn4, e4
sn1, sn2	• sn1, sn2

Source instance type family	Target instance type (family)
	• se1, n1, n2, e3, sn1ne, sn2ne, se1ne, c4, cm4, ce4, hfc5, hfg5, g5, r5, c5, ic5, re4, t5, n4, mn4, xn4, e4
c4, ce4, cm4	• c4, ce4, cm4 • sn1ne, sn2ne, se1ne, hfc5, hfg5, g5, r5, c5, ic5, re4, t5, n4, mn4, xn4, e4

10.5 Reset an instance password

If you did not specify a logon password for an instance at the time of creation, or the password is lost, specify a new password in the ECS console. This article describes how to use the Reset Password feature to specify a new logon password.

Note:
You must restart an instance after its password is reset, which may impact the service. To reduce the impact, we recommend that you reset the password when the related service is not busy.

Prerequisites

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 a region.
4. According to the number of instances to be operated, do the following:
   - To reset the password for one instance, find the instance, and in the **Actions** column, select **More** > **Reset Password**.
   - To reset the passwords for multiple instances, select the instances, and under the instance list, click **Reset Password**.
5. In the **Reset Password** dialog box, specify a new valid password, and click **Submit**.
6. (Optional), for an Alibaba account, obtain and enter the verification code sent to your phone, and then click **OK**.

> **Note:**
> This step is optional for a*Subaccount* depending on the authorizations.

7. Operate different actions to make the password effective according to the instance status:
   - **Running**: *Restart an instance* in the console.
   - **Stopped**: Start the instance in the console.

Related APIs

To modify the password: *Step 3: Connect to an instance ModifyInstanceAttribute*.

10.6 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 a **Running** status.

Prerequisite

The instance must be in a **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 a region.
4. Find an instance to be started, and in the **Actions** column, select **More** > **Start**. If you want to start multiple **Stopped** instances, select them, and under the instance list, click **Start**.
5. In the **Start Instance** dialog box, read the note and click **OK**.

The instance is in a **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 a **Stopped** status.

> **Note:**
> Stopping an instance interrupts your business operations. Proceed with caution.

If you stop a Prepaid () instance part of the way through its billing cycle, the prepaid bill for that cycle will not be affected. If the auto-renewal service is activated, you will continue to be 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 keep being billed for the instance or not. However, you will continue to be 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 will stop only after you *Release an instance*.

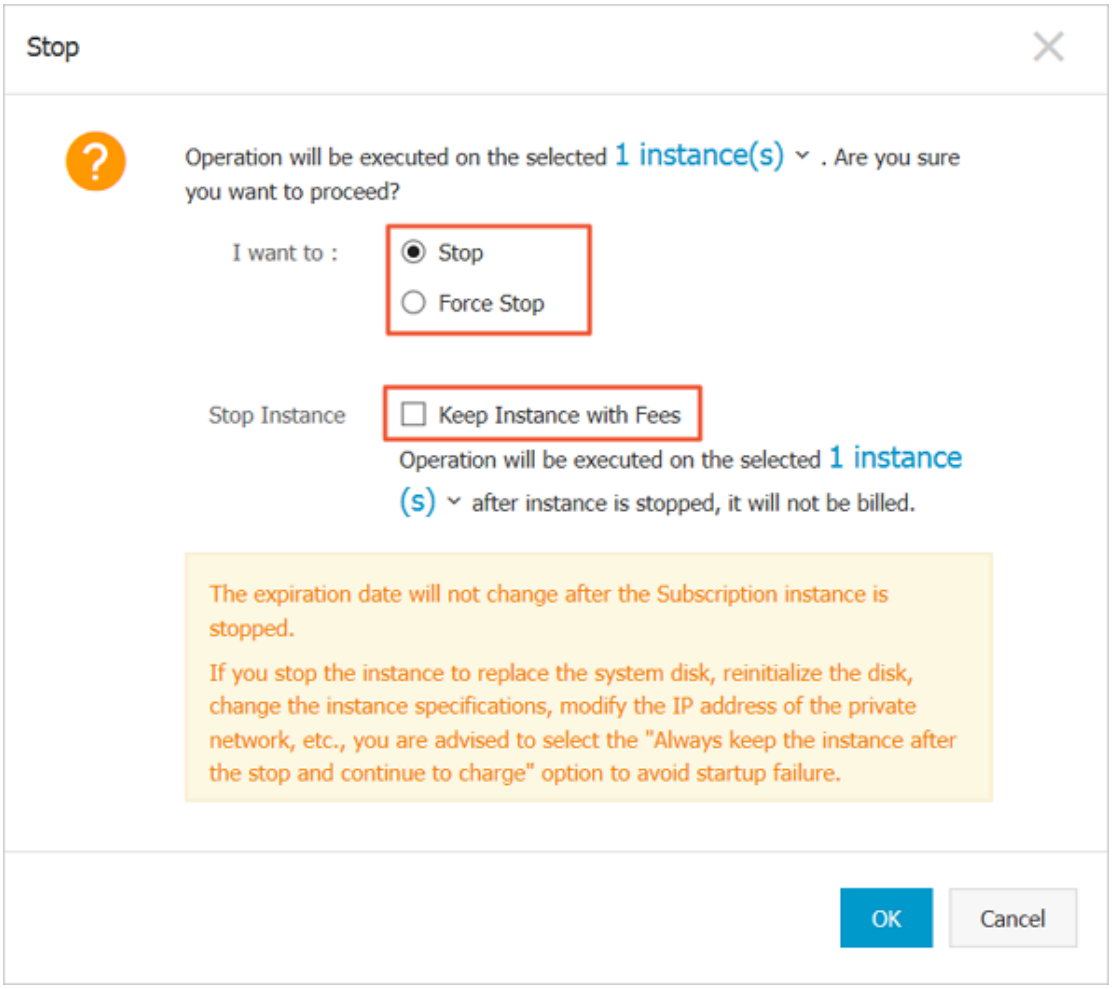
Prerequisites

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 a region.
4. Find an instance to be stopped, and in the **Actions** column, select **More** > **Stop**. If you want to stop multiple **Running** instances, select them, and under the instance list, click **Stop**.
5. According to the billing method and network type of the instance, apply suitable actions:
   - Pre-paid 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, and read the **Notice**, in the **Stop Instance** dialog box, select **Stop Method** (Stop or Force Stop), and select **Stop 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 Method** (Stop or Force Stop).

  Note:
  To disable the **No Fees for Stopped Instances (VPC-Connected)** feature, see *Disable the feature*.

6. (Optional) If you are using an Alibaba Cloud master account, you need to obtain and enter a mobile phone verification code, and then click **OK**.

   Note:
   Create a RAM user Depending on the permissions granted, you may not need to perform this step.

Once the instance is successfully stopped, the instance enters a **Stopped** status. For a VPC-Connected Pay-As-You-Go instance, if you choose not to keep the instance, **Stop Instance, No**

Fees is shown in the **Stop Mode** column of the instance list. Otherwise, **Keep Instance, Fees Apply** is shown. For other ECS instances, the **Stop Mode** column shows no information.

Related APIs

Start instance: *StartInstance*

Stop instance: *StopInstance*

10.7 Restart an instance

Instances can be restarted from within or through the management console.

Note:

- Only instances in the Running status can be restarted.
- Restarting an instance may disrupt your business traffic. Proceed with caution.

Procedure

1. Log on to the *ECS console*.
2. Click **Instances** in the left-side navigation pane.
3. Select your desired region.
4. Select the desired instance. You can select multiple instances, as long as they are all in the Running status.
5. Click **Restart**.
6. In the displayed dialog box, click **Restart**, and then click **OK**.

10.8 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**.

The classic network instance must clear the bill within 7 days after the cost is down, otherwise, the bill will re-boot, the instance is released and all data cannot be recovered.

After a VPC instance is down within 15 days, it will enter from the expiration at any time. Tax Recovery in progress. Enter tax recovery Before, you can clear the bill to re-boot. If the reboot is successful at this time, all resources are retained and will not be affected. Instance into tax recovery After the status, you can still clear the bill and re-boot, however, you may fail to boot again, please try again or submit a job solution after a period of time. If the bill cannot be cleared within 15 days, the instance will be released, all data cannot be recovered.

Note:
If you don't re-boot your account after closing your bill, the VPC instance is released automatically after 15 days from the date of its debt, classic Network instances are automatically released seven days from the date of the loss, and the data cannot be recovered.

Prerequisites

The pay per volume instance is in a State that has expired, or is in the form of a fee recovery.

Your account has been charged, and the balance of the account is not less than 100 RMB. For value-filling operations, please refer to, financial documents How to charge a value.

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.

10.9 釋放執行個體

為了節省費用，如果您不再需要某個隨用隨付執行個體，應該儘快釋放執行個體。

對於隨用隨付執行個體，如果您未開啟 隨用隨付執行個體停機不收費，停止執行個體仍會計費，只有釋放後才停止計費。

對於預付費執行個體，計費周期到期後執行個體會自動釋放，執行個體到期前，您只能申請 退款 提前釋放執行個體。

目前，您可以選擇以下任一種方式釋放執行個體：

- 立即釋放：立即釋放隨用隨付執行個體。
- 定時釋放：開啟自動釋放功能，設定自動釋放執行個體的時間。最早只能預約30分鐘後自動釋放，時間精確到分鐘。每一次設定都會覆蓋前一次設定。

立即釋放

按以下步驟立即釋放執行個體。

1. 登入 *ECS管理主控台*。
2. 在左側導覽列中，單擊 **執行個體**。
3. 選擇地域。
4. 設定釋放：
   - 如果您要釋放單個執行個體：找到需要釋放的執行個體，在 **操作** 列中，選擇 **更多 > 釋放設定**。
   - 如果您要釋放多個執行個體：根據 **付費方式** 篩選隨用隨付執行個體，選擇需要釋放的多個執行個體，在列表下方，單擊 **釋放設定**。
5. 在彈出的視窗中，選擇 **立即釋放**。
6. 單擊 **下一步**，並單擊 **確認**。
7. 輸入您收到的手機驗證碼，單擊 **確認**。

開啟自動釋放

按以下步驟開啟自動釋放：

1. 登入 *ECS管理主控台*。
2. 在左側導覽列中，單擊 **執行個體**。
3. 選擇地域。
4. 設定釋放：
   - 如果您要釋放單個執行個體：找到需要釋放的執行個體，在 **操作** 列中，選擇 **更多 > 釋放設定**。
   - 如果您要釋放多個執行個體：根據 **付費方式** 篩選隨用隨付執行個體，選擇需要釋放的多個執行個體，在列表下方，單擊 **釋放設定**。
5. 在彈出的視窗中，選擇 **定時釋放**。
6. 開啟自動釋放開關，並設定釋放的日期和時間。最早只能設定在30分鐘後自動釋放執行個體。

7. 單擊 下一步，然後單擊 確認。
8. 輸入您收到的手機驗證碼，然後單擊 確認。

關閉自動釋放

如果不需要自動釋放您的隨用隨付執行個體，您可以關閉自動釋放功能。

按以下步驟關閉自動釋放：

1. 登入 *ECS管理主控台*。
2. 在左側導覽列中，單擊 執行個體。
3. 選擇地域。
4. 釋放設定：
   - 如果您要關閉單個執行個體的自動釋放：找到需要釋放的執行個體，在 操作 列中，選擇 更多 > 釋放設定。
   - 如果您要關閉多個執行個體的自動釋放：根據 付費方式 篩選隨用隨付執行個體，選擇需要釋放的多個執行個體，在列表下方，單擊 釋放設定。
5. 在彈出的視窗中，選擇 定時釋放。
6. 關閉自動釋放開關。

7. 單擊 下一步，然後單擊 確認。

相關API

DeleteInstance

10.10 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 ECS Management 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 your desired region.
4. Select the desired 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:

- Note that an instance must be in at least two security group for this action to be performed
- , and you have done 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 your desired region.
4. Select the desired 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*.

10.11 Change IP addresses

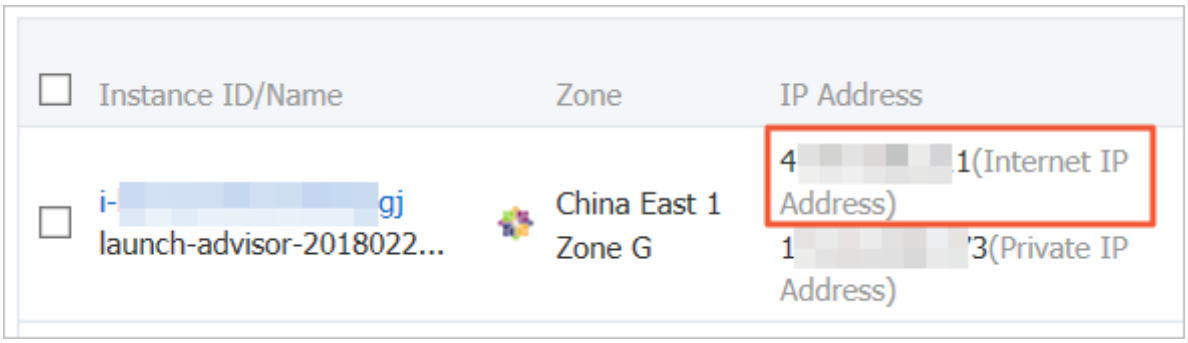
10.11.1 Change public IP address

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

Limits

Following are the limits to change the public IP address of an ECS instance:

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



Note:

- If the public network IP address is not assigned at the time of creation of the instance, after the instance is created successfully, you can assign the public IP address by upgrading or downgrading the network bandwidth configuration. For more information, see *Overview of configuration changes*.
- If the public network IP address is not assigned during the creation of a Pay-As-You-Go instance, after the instance is created successfully, public IP address cannot be assigned. You can only *bind an elastic IP (EIP) address*.

- The instance must be in the **Stopped** status.
- The instance has been existing for less than six hours.

Note:

After six hours, for a VPC instance, you can *Convert public IP address to EIP address* convert the public IP address to an EIP address, but you cannot change the public IP address of an instance in the Classic network.

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

Prerequisites

Before you change the public IP address of an instance, stop the instance.

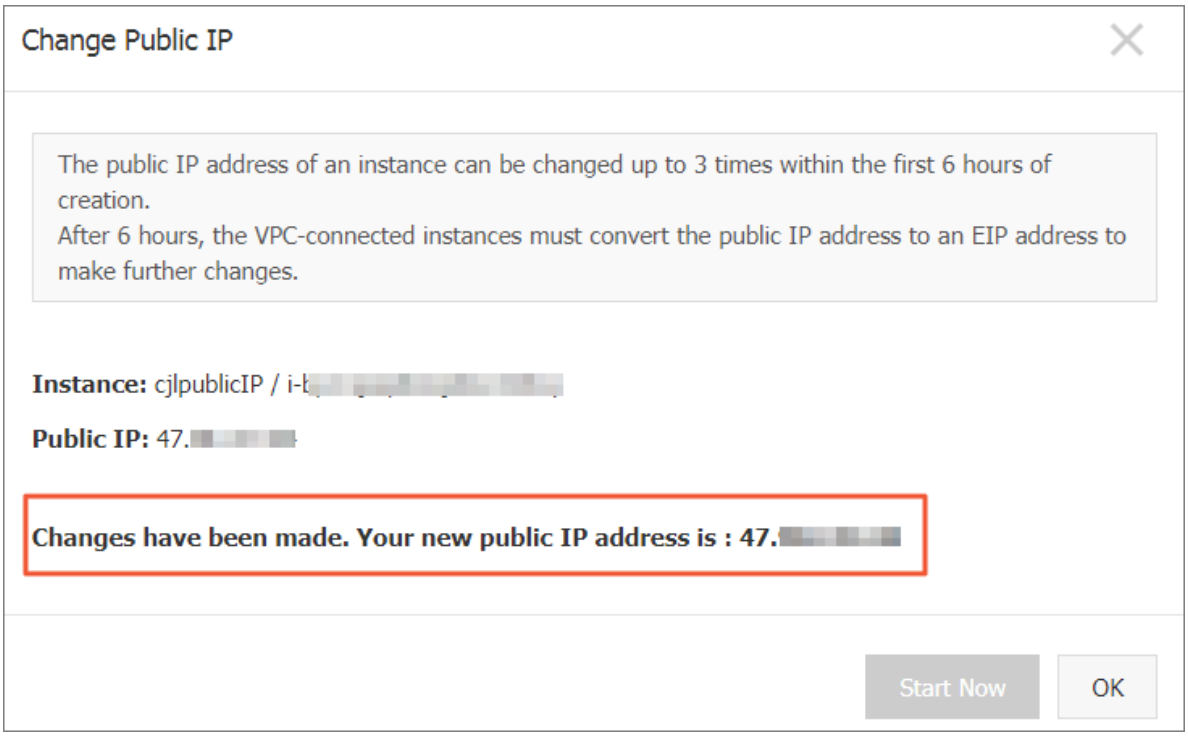
Procedure

To change the public IP address, follow these steps:

1. Log on to the *ECS console*.
2. In the left-side navigation pane, click **Instances**.
3. Select a region.
4. Find the instance to change the public IP address. In the **Actions** column, select **More >** > **Replace public IP**.

   Note:
   If your instance has been existing for more than six hours, you cannot get the **Replace public IP** option in the More drop-down menu.
5. On the **Replace public IP** dialog box, click **Start to change**.

   When a new public IP address shows on the dialog box,



6. click **OK**.

Related operations

You can *Change the private IP of an ECS instance*.

10.11.2 Convert public IP address to EIP address

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

Limits

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

- The action is irreversible.
- Only a VPC instance assigned a public IP address is supported.
- Only a VPC instance in the **Stopped** or **Running** status is supported.
- Only a VPC instance that does not have any inactivated specification changes is supported.
- Only a VPC instance that is not within the last 24 hours of its life cycle is supported.
- You can only use this feature to convert a public IP address to an EIP address.

Note

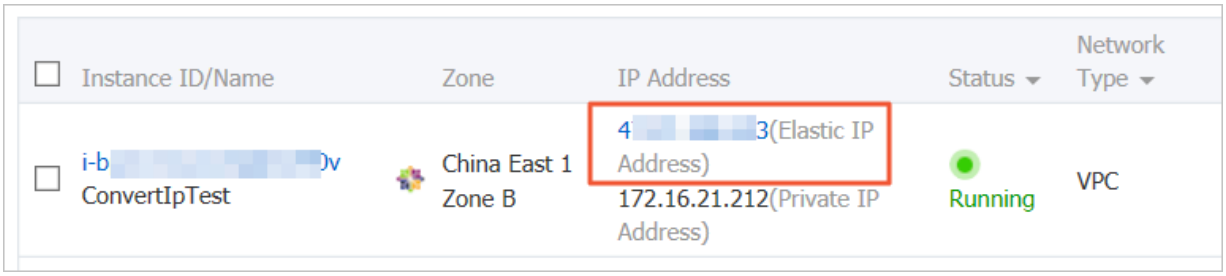
- The conversion has no effect on the Internet access of the VPC instance. It does not cause transient traffic interruption.
- The billing method of the public traffic remains unchanged.
- After conversion, the EIP address is charged separately. For more information about billing of EIP addresses, see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

Procedure

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

1. Log on to the *ECS Console*.
2. In the left-side navigation pane, click **Instances**.
3. Select a region.
4. Find a **VPC instance** to convert the public IP address, in the **Actions** column, select **More** > **Convert to EIP**.
5. In the **Convert the public IP address to EIP** dialog box, read the note and click **OK**.
6. Refresh the instance list.

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



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

Follow-up operations

After the public IP address is converted to an EIP address, you can unbind the EIP address from the instance and bind it to another instance. Besides, you can release the EIP address. For more information, see *EIP User Guide*.

Related APIs

You can use the *ConvertNatPublicIpToEip* interface to convert a public IP address to an EIP address. Currently, only SDK 4.3.0 or a higher version supports this interface.

Download the latest version.

10.11.3 Change the private IP of an ECS instance

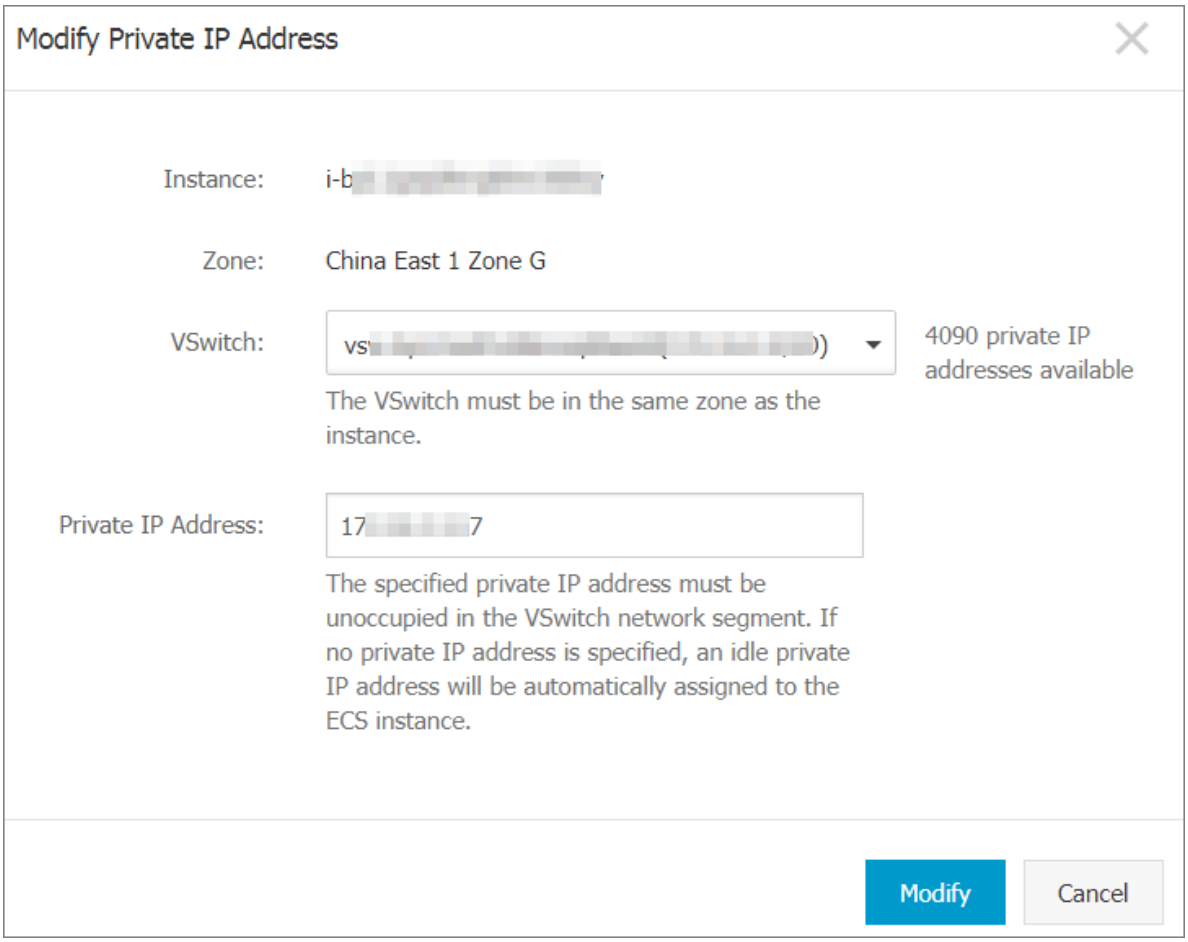
After creating a VPC ECS instance, you can change the private IP address and also can change the VSwitch of the ECS instance.

Procedure

1. Log on to the *ECS console*.
2. On the left-side navigation pane, click **Instances**. Click a region and then click the ID of the target ECS instance.
3. In the **Actions** column, click **More >** > **Stop**.
4. When the instance is stopped, click the instance ID to go to the **Instance Details** page.
5. In the **Configuration Information** panel, click **More >** > **Modify Private IP Address**.
6. In **Modify Private IP** Address dialog, select a VSwitch, and then click **Modify**.

   Ensure the zone of the selected VSwitch and the current VSwitch is the same.

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



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

10.12 User-defined data and metadata

10.12.1 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 pre-encoding user data 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:

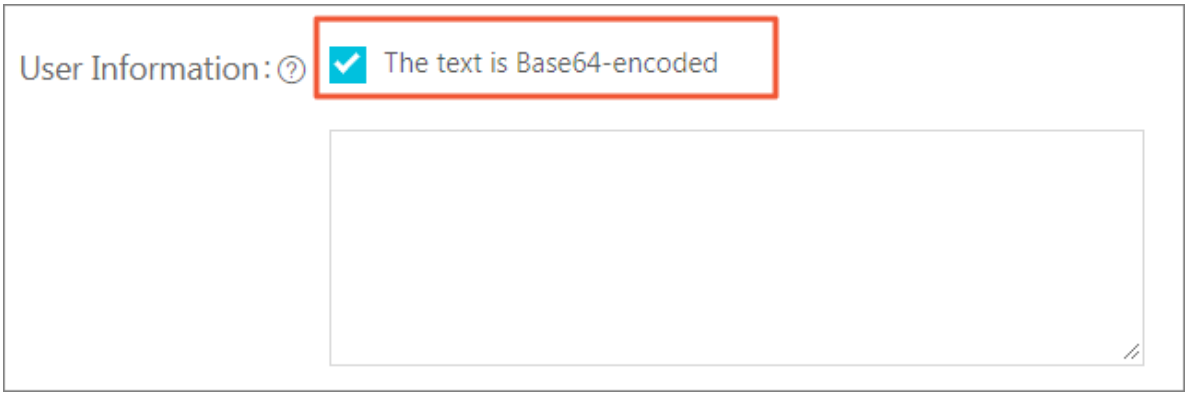
The first line must meet the format requirements of the instance user data script, such as `#! /bin/sh`, `#cloud-config`, `#upstart-job`, `[bat]` and `[powershell]`. For more information , see *Linux instance user data* and *Windows instance user data*.

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` .

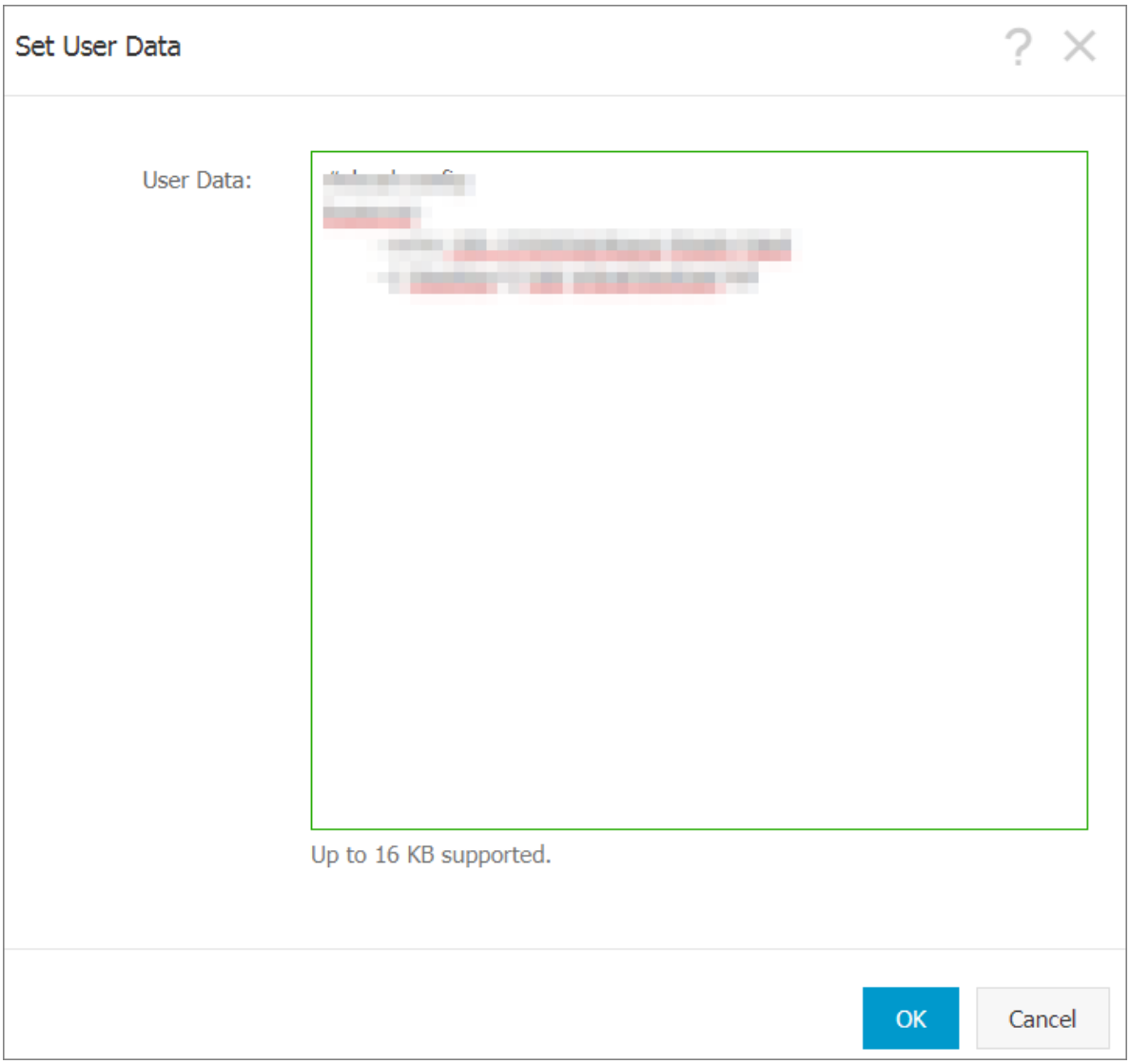
1. Connect to the instance.
2. In the instance, run:
   - `curl http://100.100.100.200/latest/user-data` View the user data of a Linux instance:
   - `Invoke-RestMethod http://100.100.100.200/latest/user-data/` View the user data of a Windows instance:

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.

```
[redacted]    localhost localhost.localdomain localhost4 localhost4.localdomain4
::1           localhost localhost.localdomain localhost6 localhost6.localdomain6
192.168.1.130 us.archive.ubuntu.com
~
~
```

Include files

The contents of 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

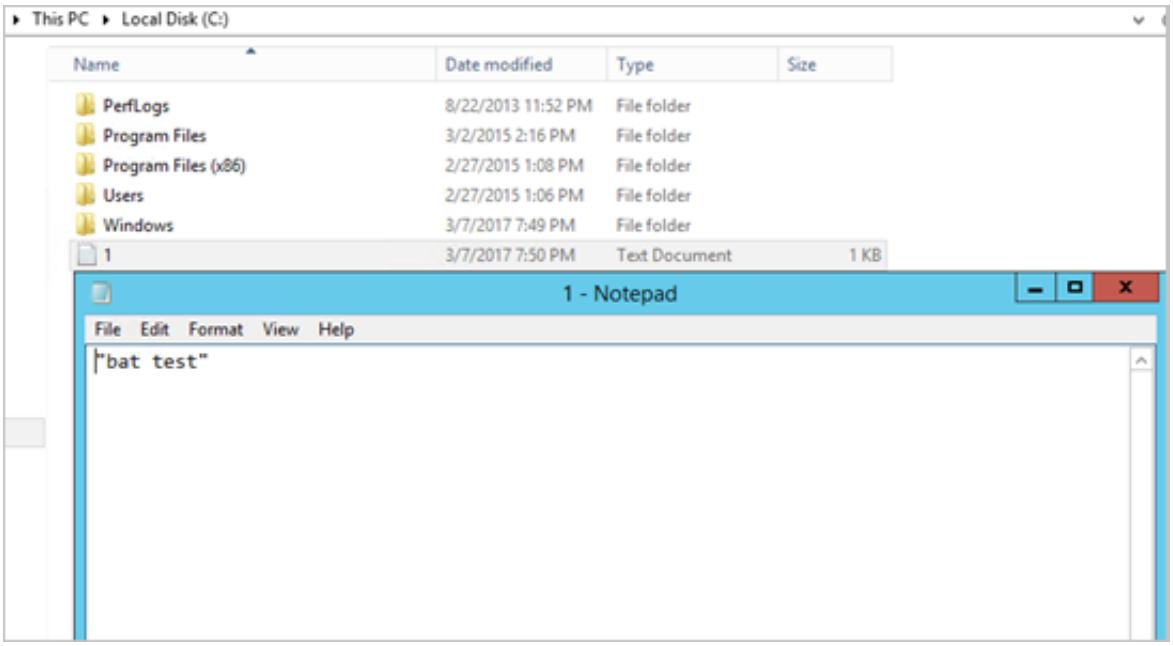
before being passed in, and the pre-encoding user data cannot exceed 16 KB. Only SBC case characters are allowed. You can write Bat script or PowerShell script to configure the instance user data.

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*.

10.12.2 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

Connect to a Linux instance by using a password.

- Run `curl http://100.100.100.200/latest/meta-data/` to access the root directory of the metadata.
- 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

Connect to a Windows instance.

- Use PowerShell to run `Invoke-RestMethod http://100.100.100.200/latest/meta-data/` to get the metadata.
- 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.**	**2016/1/1**
eipv4	EIP address	2016/1/1
hostname	The OS name of an instance.	2016/1/1
image-id	ID of the image that is selected at the time of instance creation .	2016/1/1
image/market-place/product-code	Product code of the image in the marketpalce.	2016/1/1
image/market-place/charge-type	Billing method of the image in the marketplace.	2016/1/1
instance-id	Instance ID	2016/1/1
mac	The MAC address of an instance. When multiple network interface cards exist in an instance, this metadata indicates the MAC address of eth0.	2016/1/1
network-type	Network type, only applicable for VPC.	2016/1/1
ntp-conf/ntp-servers	The address of a NTP server.	2016/1/1
owner-account-id	The aliuid of the instance owner.	2016/1/1
private-ipv4	Private IP address.	2016/1/1
public-ipv4	Public network IP address.	2016/1/1
public-keys	The list of all public keys of the current instance.	2016/1/1
region-id	The region where the instance is located.	2016/1/1
zone-id	Zone ID of the zone where the ENS instance is located.	2016/1/1
serial-number	The serial number of an instance.	2016/1/1

Metadata name	Description	Version
dns-conf/nameservers	**DNS configurations for an instance.**	**2016/1/1**
source-address	The source of Yum/apt, only applicable for a Linux instance.	2016/1/1
kms-server	Activate the server, only applicable for a Windows instance.	2016/1/1
wsus-server/wu-server	Update the server, only applicable for a Windows instance.	2016/1/1
wsus-server/wu-status-server	The server that monitors the update status of an instance, only applicable for a Windows instance.	2016/1/1
vpc-id	ID of the VPC that an instance is in.	2016/1/1
vpc-cidr-block	The CIDR block of the VPC that an instance is in.	2016/1/1
vswitch-cidr-block	The CIDR block of the VSwitch that an instance is in.	2016/1/1
vswitch-id	ID of the VSwitch that an instance is in.	2016/1/1
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 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	2016/1/1

Metadata name	Description	Version
dns-conf/nameservers	**DNS configurations for an instance.**	**2016/1/1**
	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.	2016/1/1
network/interfaces/macs	The MAC address list of the multiple NIC (Network Interface Controller).	2016/1/1
network/interfaces/macs/[mac]/network-interface-id	The unique ID of the NIC, **[mac]** must be replaced with the actual MAC address.	2016/1/1
instance/virtualization-solution	The virtualization solution: ECS Virt 1.0 / 2.0	2016/1/1
instance/virtualization-solution-version	The internal Build version.	2016/1/1
instance/last-host-landing-time	The latest update time of the physical server, which your instance is hosted on.	2016/1/1
instance-identity/document	*Instance identity* Instance identity document.	2016/1/1
instance-identity/pkcs7	Instance identity signature.	2016/1/1

10.13 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
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.

Properties	Description	One and only?
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

Prerequisite: The instance identity is verified by using the OpenSSL commands. Make sure that you have the OpenSSL configured in your instance. For Windows instances, see *How to install OpenSSL in ECS Windows*. For Linux instances, visit *https://www.openssl.org/source* to download and install 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.

```
[root@LocalHost ~]# curl http://100.100.100.200/latest/dynamic/instance-identity/document
{"region-id":"cn-hangzhou","instance-id":"i-bp1b39[illegible]","serial-number
":"52ae1af6-64aa-407f-88fd-[illegible]","private-ipv4":"172.[illegible].174","mac":"
00:16:3f:[illegible]","image-id":"centos_7_04_64_20G_alibase_201701015.vhd","zone-i
d":"cn-hangzhou-g","owner-account-id":"[illegible]","instance-type":"ecs.g5
.large"}[root@LocalHost ~]# curl http://100.100.100.200/latest/dynamic/instance-identity
/pkcs7?audience=test
MIIDMwYJKoZIhvcNAQcCoIIDJDCCAyACAQExCzAJBgUrDgMCGgUAMIIBbQYJKoZIhvcNAQcBoIIBXgSC
BSsOAwIaBQAwDQYJKoZIhvcNAQEBBQAEggEAUOghxG3i3hKgjPggPX6NInYNPOZJusp//fy15Pr+gZoq
LgvxonLzOXxBG1yy1aEzjb2b2zUFZGfjuNDbk1kcVSgq3kKEbpBCXF2GNm9FaS54r0szTcw9HkNPskla
CqR9Z6LvBB/sPMTz8i3dY8pu/zhiZOulHdpvKCYYP8Q89sN+QVlsS2eQDXnqNSFBi/QS/hZOz83gIuac
H6+rWXojBf3Hs7bfT4YdmNBhaTpxq8R3wl6rWvTq3W58ULraHgMzq/Kn9e1SCnSAiqETLj7i60As7h/h
d/DebOVoF0QANiXAIvdhLci4KK1rKJ0HOW4BzOy44s/jc1flASYIsAQRug==[root@LocalHost ~]#
```

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.

10.14 Instance RAM roles

10.14.1 What is the RAM role of an instance

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

You can assign a 角色 to an ECS instance to allow applications hosted on that instance to access other cloud services by using a temporary STS (Security Token Service) credential. In this way, one side guarantees that the AccessKey and allows you to use the fine-grained access control in virtue of RAM.

Background

Typically, the applications within an ECS instance use the AccessKey of the **user account** or 用户, including AccessKeyID and AccessKeySecret, to access various cloud services on the Alibaba Cloud platform.

To meet the requirements of the call, you need to cure the AccessKey directly in the instance, such as in the configuration file. However, if Alibaba Cloud writes the AccessKey into the instance for calling purposes, for example, writing in the configuration files, the exposed AccessKey leads to problems such as excessive permission, data breaches, and maintenance complexities. Thus, Alibaba Cloud has designed the instance RAM role to solve the complexities.

Benefits

The instance RAM role enables you to:

- Associate a 角色 to an ECS instance.
- Access other cloud services securely, such as OSS, SLB, or ApsaraDB for RDS, by using the STS credential from the applications within the ECS instance.
- Assign roles that have different policies for different ECS instances, and let these instances have restrictive access level to other cloud services to obtain fine-grained access control.
- Maintain the access permission of the ECS instances efficiently only by modifying the policy of the RAM role, without manually changing the AccessKey.

Free of charge

ECS does not charge any additional fee on instance RAM role.

Limits

The instance RAM role has the following limits:

- The instance RAM role is only applicable to VPC instances.
- One ECS instance can only be authorized to one instance RAM role.

How to use an instance RAM role

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

- 通过控制台使用实例 *RAM* 角色
- 通过 *API* 使用实例 *RAM* 角色

References

- For a list of cloud services that support the STS credential, see RAM document 支持 *RAM* 的云服务.
- See *Access other cloud products by using the instance RAM role* for instruction on how to access other cloud services.

10.14.2 Use the instance RAM role in the console

Limits

The instance RAM role has the following limits:

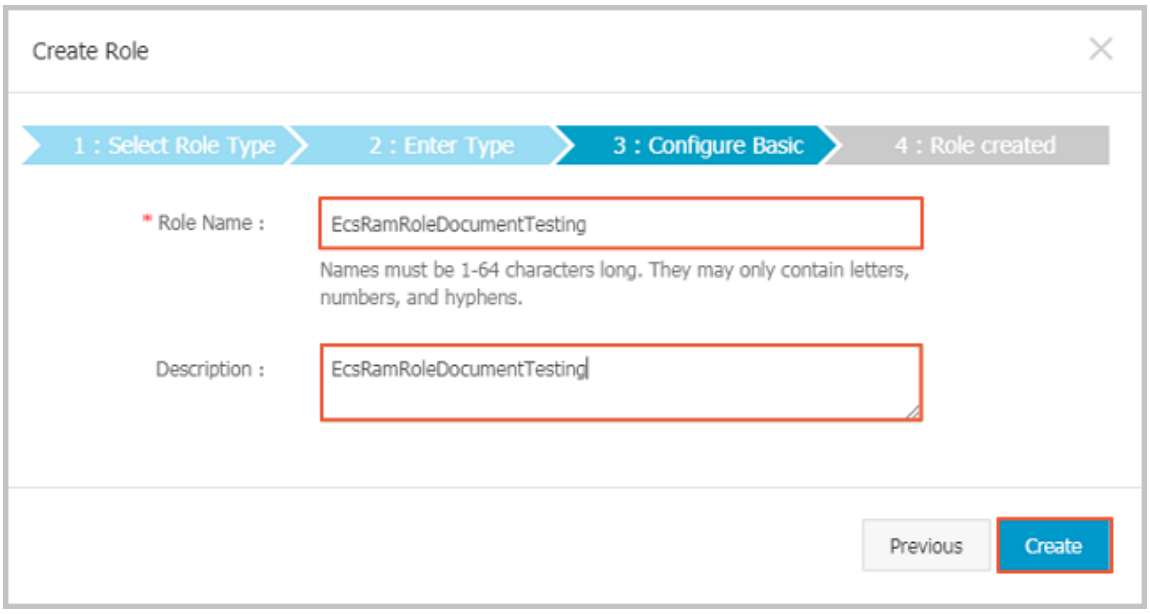
- The instance RAM role is only applicable to VPC instances.
- One instance RAM role can be bound to one instance at a time.
- After an instance RAM role is attached to an ECS instance, if you want to access other cloud services, such as OSS, SLB, or ApsaraDB for RDS, from the applications within the ECS instance, you must *Metadata* of the instance RAM role by using metadata. See *6. (Optional). Obtain the authorization credential*.
- Before using this feature, the RAM user must be *7. (Optional). Authorize a RAM user to use the instance RAM role*.

Prerequisites

You must have activated the RAM service. See *Activation method* to activate the RAM service.

1. Create an instance RAM role

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

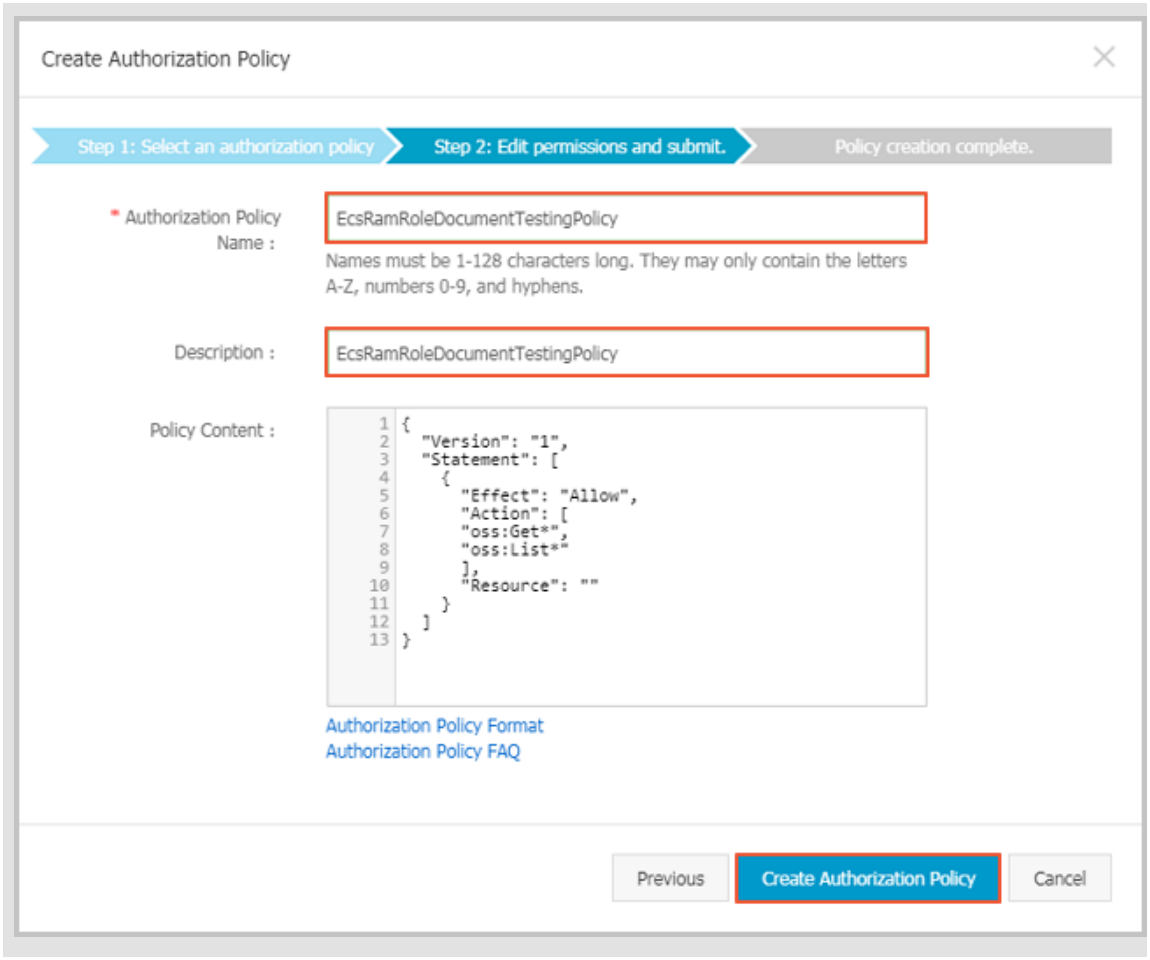


d. Click **Create** to create the instance RAM role.

2. Authorize the instance RAM role

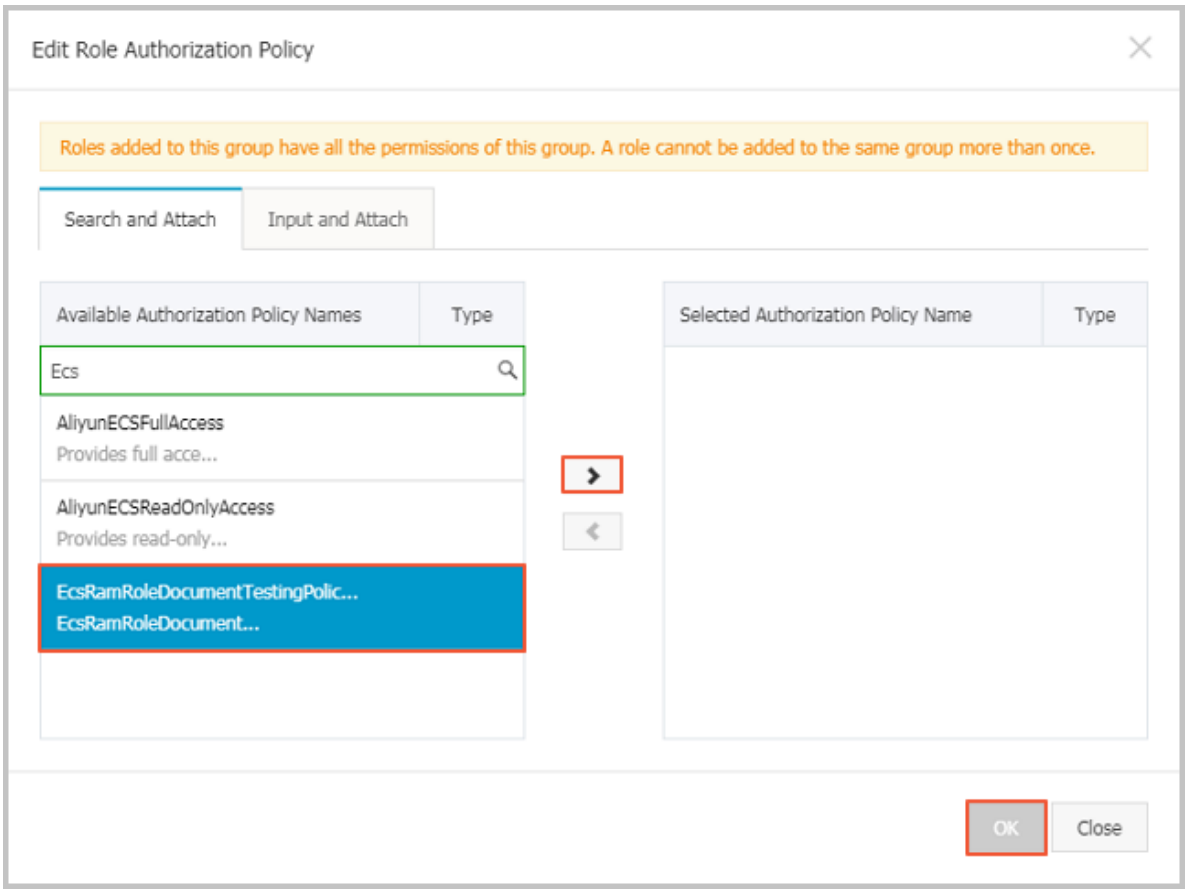
1. Log on to the *RAM console*.
2. On the left-side navigation pane, click **Policy**.
3. Click **Create Authorization Policy**.
4. In the dialog box:
   a. Select **Blank Template** for **authorization policy template**.
   b. Enter the **Authorization Policy Name** and **Policy Content**, for example,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

> **Note:**
> For more information about how to write the authorization policy by using the JSON language, see *Policy syntax structure*.



c. Click **Create Authorization Policy** to complete authorization.

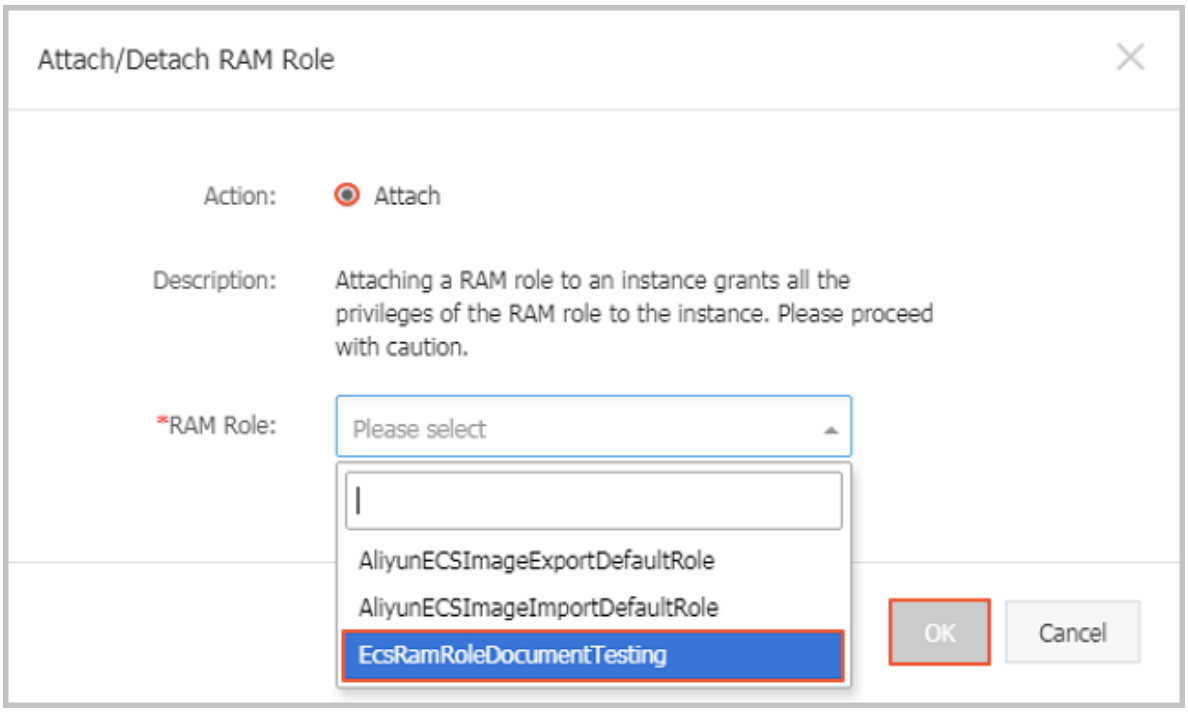
5. On the left-side navigation pane, click **Roles**.
6. On the **Roles** page, select the created role, for example, EcsRamRoleDocumentTesting, and click **Authorize**.
7. Enter the **authorization policy name** and click it, for example, EcsRamRoleDocumentTestingPolicy.
8. Click the icon **>** to select the policy name, and click **OK**.



3. Attach an instance RAM role

Method 1: Attach an instance RAM role in the console

1. Log on to the *ECS console*.
2. On the left-side navigation pane, click **Instances**.
3. Select a region.
4. Find the target ECS instance and select **More >** > **Attach/Detach RAM Role**.
5. Select Attach for Action, select the created role, for example, EcsRamRoleDocumentTesting, and click **OK** to attach the instance RAM role.



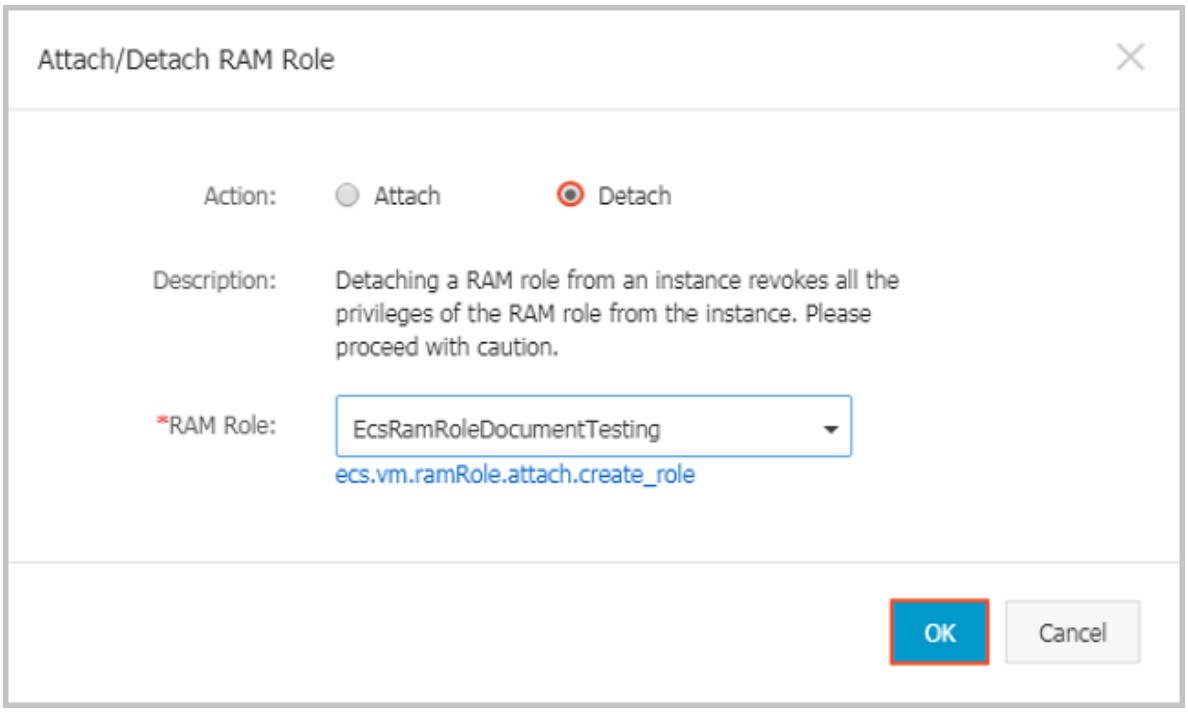
Method 2: Attach an instance RAM role when creating the ECS instance

1. Log on to the *ECS console*.
2. On the left-side navigation pane, click **Instances**.
3. Click **Create Instance**.
4. See *Step 2. Create an instance* relevant information about the instance setting, and in **RAM Roles**, select the created instance RAM role, such as EcsRamRoleDocumentTesting.

When an instance is created, it has the permissions granted in the instance RAM role policy.

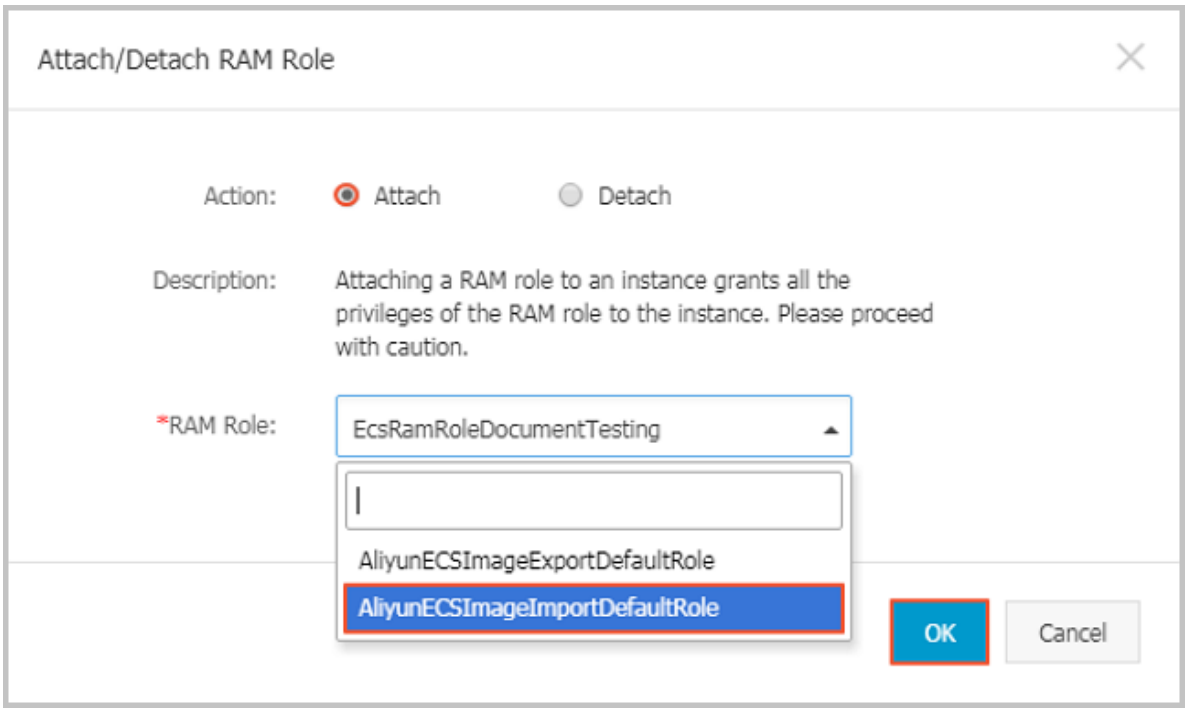
(Optional). Detach an instance RAM role

1. Log on to the *ECS console*.
2. On the left-side navigation pane, click **Instances**.
3. Select a region.
4. Select an ECS instance, and select **More >** > **Attach/Detach RAM Role**.
5. Select **Detach** for **Action**, and click **OK** to detach the instance RAM role.



5. (Optional). Replace an instance RAM role

1. Log on to the *ECS console*.
2. On the left-side navigation pane, click **Instances**.
3. Select a region.
4. Select an ECS instance, and select **More >** > **Attach/Detach RAM Role**.
5. Select **Attach** for **Action**, select another instance RAM role in the list of RAM Role, and click **OK** to replace the current RAM role.



6. (Optional). Obtain the authorization credential

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

1. Connect and log on to your ECS instance.
2. Obtain the STS credential of the instance RAM role, for example, EcsRamRoleDocumentTesting:
   - Linux instance: run `curl http://100.100.100.200/latest/meta-data/Ram/security-credentials/EcsRamRoleDocumentTesting`.
   - Windows instance: see *Metadata*.
3. Get the credential. Return example:

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

```
"Code" : "Success"
```

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

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

Log on to the RAM console and authorize a RAM user *Attach policies to a RAM user* to complete the authorization, see the following code snippet as an authorization policy example:

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

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

    "Resource": "*"


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

The parameter **[ECS RAM Action]** indicates the action that a ram user can be authorized. For more information, see *Authorization rules*.

References

- You can also *Use the instance RAM role by calling APIs*.
- You may want to *Access other cloud products by using the instance RAM role*.

10.14.3 Use the instance RAM role by calling APIs

Limits

The instance RAM role has the following limits:

- The instance RAM role is only applicable to VPC instances.
- One instance RAM role can be bound to one instance at a time.

- After an instance RAM role is attached to an ECS instance, if you want to access other cloud services, such as OSS, SLB, or ApsaraDB for RDS, from the applications within the ECS instance, you must obtain the authorization credential of the instance RAM role by using *Metadata*. See *5. (Optional). Obtain the on-demand authorization credential*.
- If you are using an instance RAM role through a RAM user sub-account, you need to use a cloud account*6. (Optional). Authorize a RAM user to use the instance RAM role*.

Prerequisites

Before using this feature, the RAM user must be authorized to use the instance RAM role. See *Activation method* to activate the RAM service.

1. Create an instance RAM role

1. Call the CreateRole *CreateRole* to create an instance RAM role.
2. Set the parameter RoleName, for example, EcsRamRoleDocumentTesting.
3. Set the **AssumeRolePolicyDocument**as follows:

```
"Statement": [

"Action": "sts:AssumeRole",
"Effect": "Allow",
"Principal": {
"Service": [
"ecs.aliyuncs.com"

}


"Version": "1"
```

2. Authorize the instance RAM role

1. Call the CreatePolicy to *CreatePolicy* create an authorization policy.
2. Set the parameter RoleName, for example, set it to EcsRamRoleDocumentTestingPolicy.
3. Set the **PolicyDocument**as follows.

```
"Statement": [

"Action": [
"oss:Get*",
"oss:List*"

"Effect": "Allow",
"Resource": "*"


"Version": "1"
```

4. Call the *AttachPolicyToRole* to authorize the role policy.

5. Set **PolicyType** to Custom.
6. Set the parameter **PolicyName**, for example, EcsRamRoleDocumentTestingPolicy.
7. Set the parameter **RoleName**, for example, EcsRamRoleDocumentTesting.

Attach the instance RAM role

1. Call the *AttachInstanceRamRole* to attach an instance RAM role to an ECS instance.
2. Set the parameters **RegionId** and **InstanceIds** to specify an ECS instance.
3. Set the parameter **RamRoleName**, for example, EcsRamRoleDocumentTesting.

4. (Optional). Detach an instance RAM role

1. Call the *DetachInstanceRamRole* to detach an instance RAM role.
2. Set the parameters **RegionId** and **InstanceIds** to specify an ECS instance.
3. Set the parameter **RamRoleName**, for example, EcsRamRoleDocumentTesting.

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

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

1. Obtain the STS credential of the instance RAM role, for example, EcsRamRoleDocumentTesting:
   - Linux instance: run `curl http://100.100.100.200/latest/meta-data/Ram/security-credentials/EcsRamRoleDocumentTesting` .
   - Windows instance: see *Metadata*.
2. Get the credential Token. Return example:

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

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

Note:

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

Log on to the RAM console and *Attach policies to a RAM user* authorize a RAM user to complete the authorization, see the following code snippet as an authorization policy example:

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

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

"Resource": "*"


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

The parameter **[ECS RAM Action]** indicates the action that a RAM user can be authorized. See *Authorization rules*.

References

- You can also *Use the instance RAM role in the console*.
- For instruction on how to access other cloud services, see *Access other cloud products by using the instance RAM role*.
- APIs related to the instance RAM role include:
  - *CreateRole*: Create an instance RAM role
  - *ListRoles*: Query the list of instance RAM roles
  - *CreatePolicy*: Create an instance RAM role policy
  - *AttachPolicyToRole*: Authorize an instance RAM role policy
  - *AttachInstanceRamRole*: Attach an instance RAM role
  - *DetachInstanceRamRole*: Detach an instance RAM role
  - *DescribeInstanceRamRole*: Query an instance RAM role

10.15 Launch template

10.15.1 Create a template

You can create a template using the following methods:

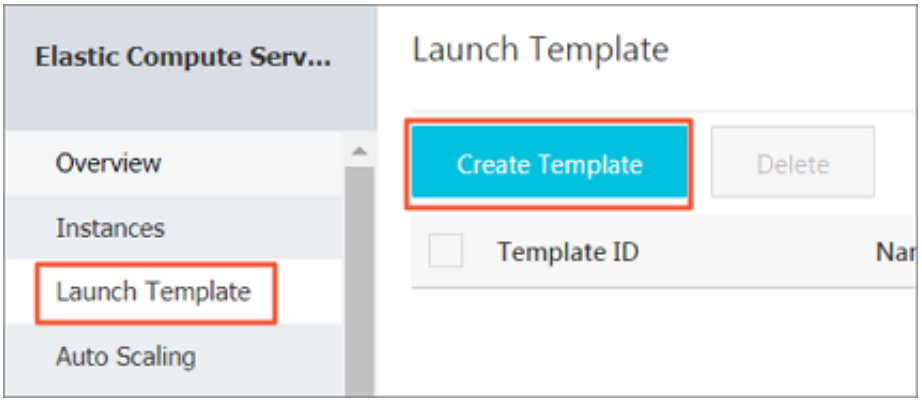
- If you do not want to create instances at this time, you can still *create a template in the ECS console* create templates using the ECS console, and then create instances using your required template in one click when needed.
- If you want to create an instance and save its configuration information, *create a template on the ECS buy page*.

Note:

- In each region, one user account can only create a maximum of 30 launch templates.
- 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 once you have created it.

Create a template in the ECS console

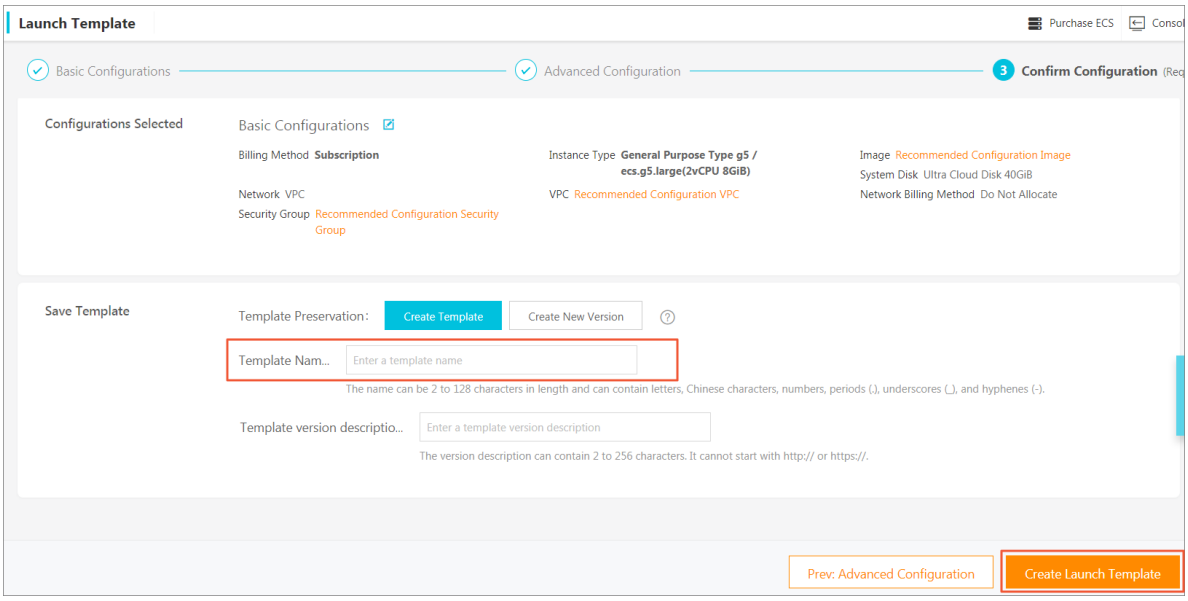
1. Log on to the *ECS console*.
2. In the left-side navigation pane, select **Launch Template**, and then click **Create Template**.



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. On the **Confirm Configuration** page, we recommend that you configure the required parameters so that you can create instances in one click later. You can also leave the parameter settings unchanged.

5. In the **Activated** dialog box, click **View Template** 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. 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, select **Create Template**, enter a template name and description, and then click **Save**.
5. In the **Activated** dialog box, click **View Template** to view the template you have created.

10.15.2 Create a template version

One template can have multiple versions. The default version number of a newly created template is 1. You can create additional versions based on this template. The version number automatically increments as you create a new version. You cannot customize the version number. 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.
- The version cannot be modified once you have created it.

You can create a version using the following methods:

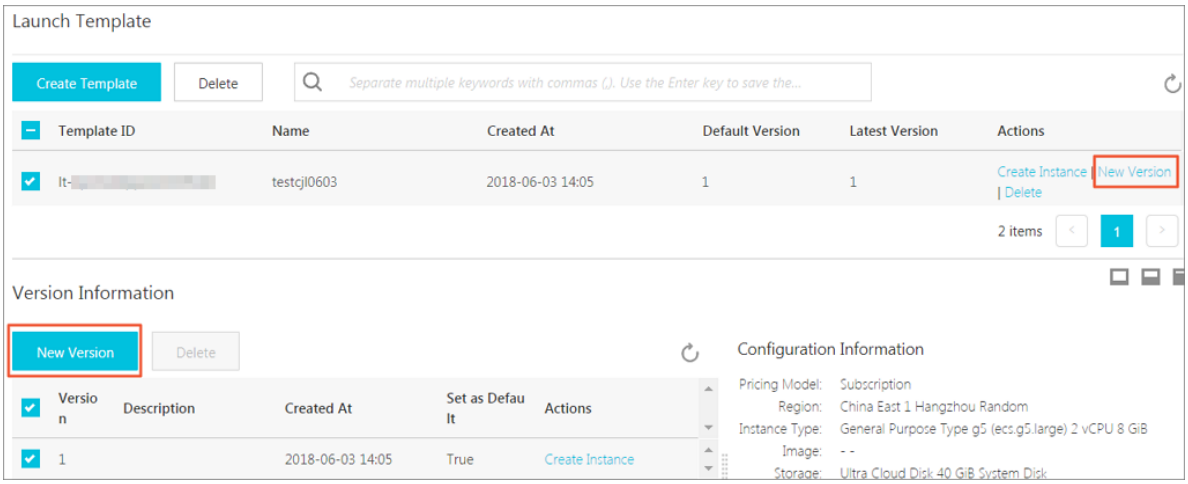
- If you do not want to create an instance now, you can still create multiple versions of a template *Create an instance using the ECS console* and create instances using your specified version later.
- If you want to create an instance and save the configuration information, go to the ECS buy page *Create an instance on the ECS buy page* to create versions of a template.

Prerequisites

You have already *Create 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.

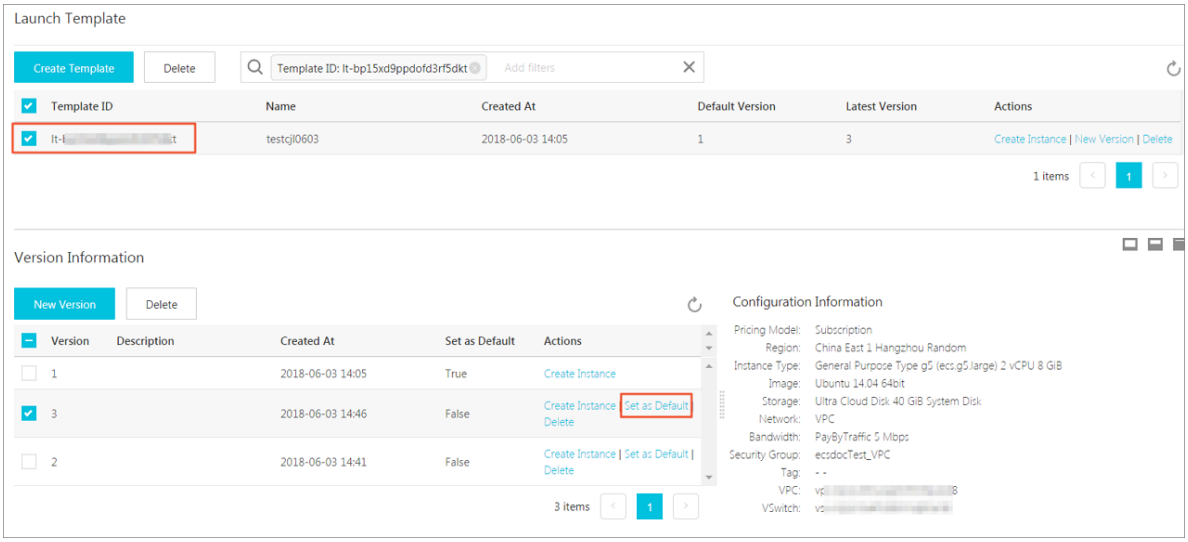
5. On the **Confirm Configuration** page, select **Create New Version**, and then select a template to save the version.
6. Click **Create Launch Template**.
7. In the dialog box that appears, click **View New Version** to view the version you have created.

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 to view its version details.
2. Locate the version you want to set as default, and then click **Set as Default** in the **Actions** column.



10.15.3 Delete a template or version

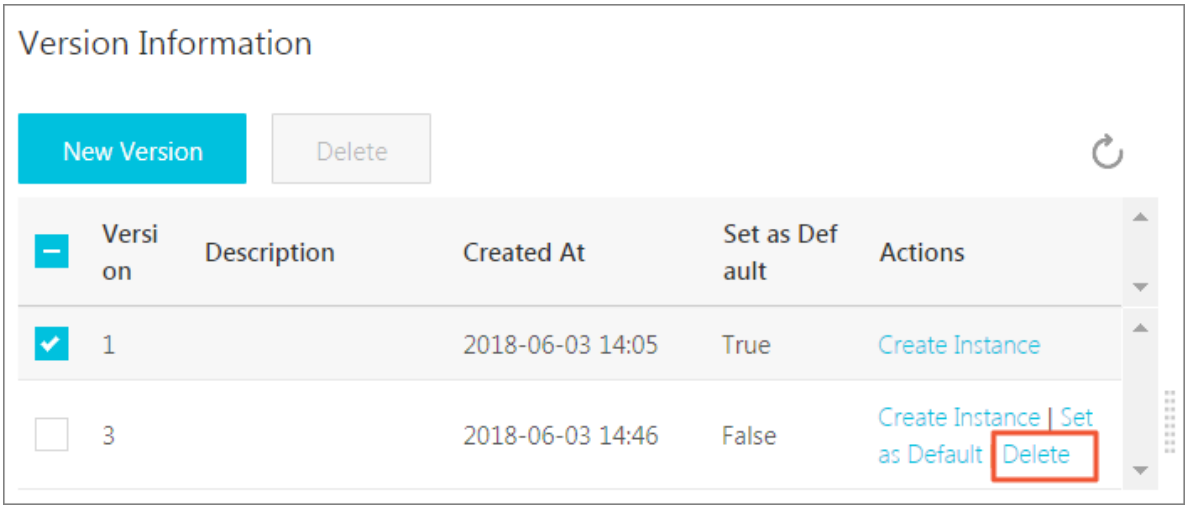
You can delete templates and versions using the ECS console. Once you delete a template, all the versions of that template are also deleted.

Delete a version

1. Log on to the *ECS console*.

2. In the left-side navigation pane, click **Launch Template**.
3. Select a template ID and check its version details.
4. In the **Version Information** area, locate the version you want to delete, and in **Actions** column, click **Delete**.

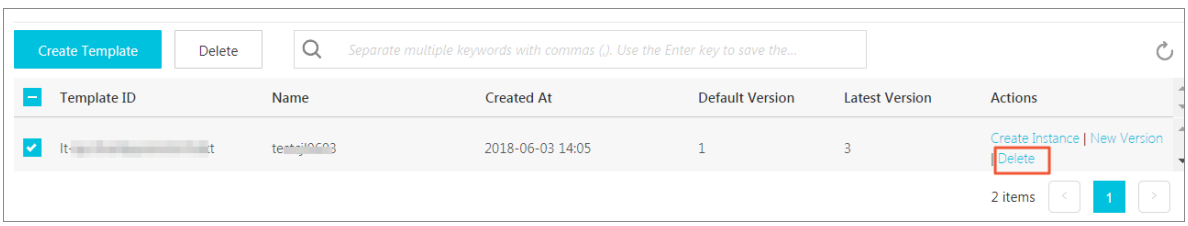
> **Note:**
> You cannot delete the default template version. You only can delete non-default versions. If the version you want to delete is the default version, change it to a non-default version, and then delete it. If all versions in a single template are not needed, we recommend that you delete the template.



5. In the dialog box that appears, 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. In the dialog box that appears, click **OK**.

> **Note:**

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

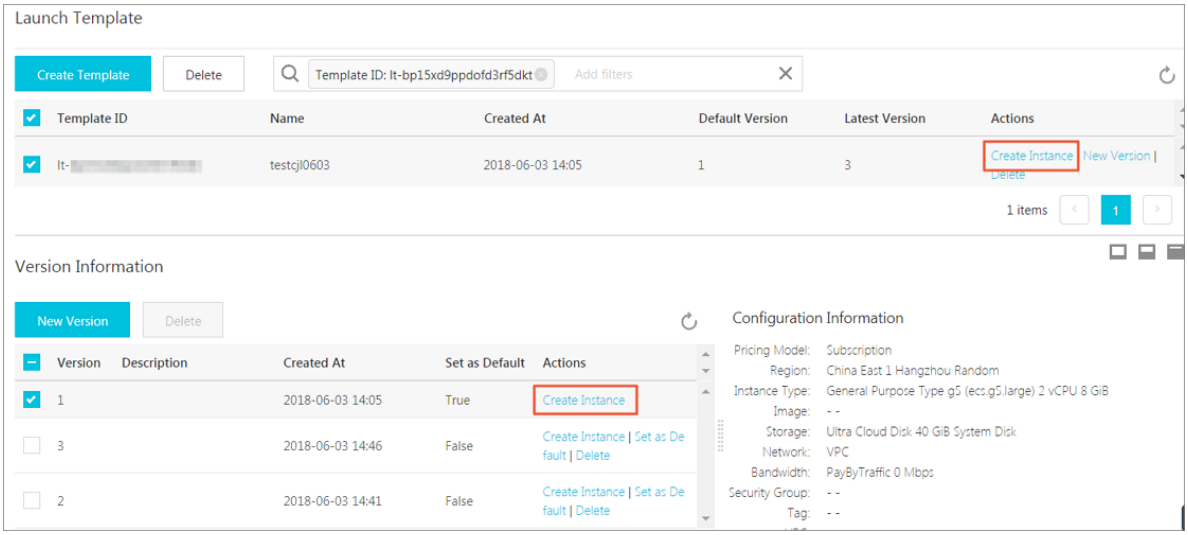
10.15.4 Use a launch template

Prerequisites

You have completed the *Create a template* or *Create a template version* step.

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, click **Create Instance** in the **Actions** column.



4. On the **ECS buy page**, select the template and version. Verify the configurations when they are displayed.

Note:
If you want to modify the configurations, or the selected template does not have the required parameters, you can click the Change settings icon to modify the configurations.

5. If you want to create an instance using the Subscription billing method, choose 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.

11 Cloud disks

11.1 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 *Limits*.
- Currently, you cannot merge multiple cloud disks. After cloud disks are created, they are independent from each other, and you cannot merge their space by formatting. We recommend that you determine the number 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 prepaid instance, use the *Upgrade configurations of prepaid 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.

- Your account balance must be more than RMB 100 yuan or an equivalent voucher or coupon. Because a cloud disk separately created is billed as Pay-As-You-Go.
- 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.
- Choose whether to encrypt the disk. For more information, see *ECS disk encryption*.

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. Choose whether to **Encrypt** it, and specify the number of disks you want to buy. 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 videos

You can watch the video *Attach a disk to a Windows instance*.

Related APIs

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

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

11.2 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 *Limits*.

- Currently, you cannot merge multiple cloud disks. After cloud disks are created, they are independent from each other, and you cannot merge their space by formatting. We recommend that you determine the number 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 prepaid instance, use the *Upgrade configurations of prepaid instances* feature.
  - If it is attached to a Pay-As-You-Go 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*.
- Your account balance must be more than RMB 100 yuan or an equivalent voucher or coupon. Because a cloud disk separately created is billed as Pay-As-You-Go.
- 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.

      Note:
      If the selected snapshot is a dense snapshot, the **encryption** item is automatically selected. Otherwise, you cannot select **Encryption**. For more information, see *ECS disk encryption*.

   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*

11.3 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*.
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_93* 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

Related documents

You can also watch the video *Attach a cloud disk on a Windows ECS instance* to check how to create, attach, and format the data disk.

11.4 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.

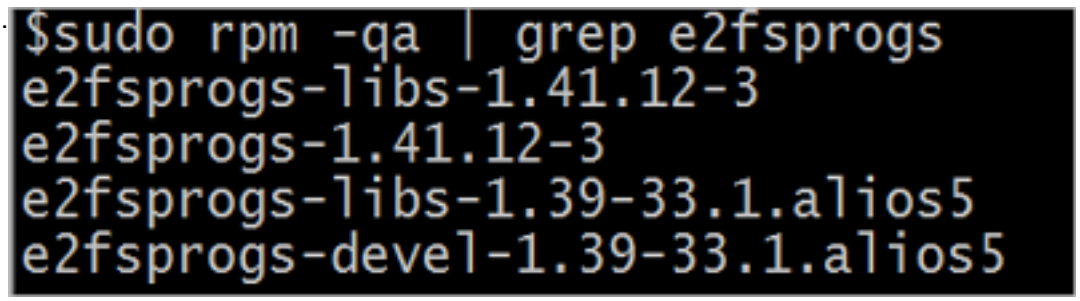
Appendix 1: Update e2fsprogs

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 continuously 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.

11.5 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*
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*.

11.6 Resize cloud disks

11.6.1 Overview

To meet the expanded business needs, you can increase the size of a cloud disk when it is used as a system disk or a data disk. To resize the disks, use different features:

- For a system disk: **Change System Disk**
- For a data disk: **Resize Disk**

Size limits of cloud disks

The size limits of cloud disks for resizing vary between system disks and data disks.

System disks

By using the **Change system disk** feature, you can keep the system disk size unchanged or increase the size only, but not reduce the size. For example, before changing, the system disk of a CentOS instance is of 35 GiB, so it must be equal to or greater than 35 GiB **after changing**. 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	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]

Data disk

By using the **Resize Disk** feature, you can keep the data disk size unchanged or increase the size only, but not reduce the size. The following table lists the capacity limit of a data disk after resizing, which is determined by the cloud disk types.

Cloud disk type	Current capacity	Capacity after resizing
Basic Cloud Disk	Any	2000 GiB
SSD Cloud Disk or Ultra Cloud Disk	equal or less than 2048 GiB	2048 GiB
SSD Cloud Disk or Ultra Cloud Disk	> 2048 GiB	Cannot be resized

Operations

You can perform the following tasks:

- To increase the size of the system disk of an ECS instance, see *Increase system disk size*.
- To resize a data disk attached to a Windows instance, see *Windows _ Resize a data disk*.
- To resize a data disk attached to a Linux instance, see *Linux _ Resize a data disk*.

11.6.2 Increase system disk size

To meet the growing business demands, you can increase the size of the system disk of your ECS instance by using **Change System Disk** feature. This article introduces how to increase the size of a system disk while keeping the operating system and environment intact.

> **Note:**
> You can change the operating system while increasing the size of a system disk. For more information, see *Change the operating system*.

Notes

Before you begin, consider the following.

Risks

The risk of replacing the system disk 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.
- After the system disk is changed, a new cloud disk with a new disk ID is assigned, and the old one is released. Therefore, you cannot roll back the system disk by using any snapshot of the released cloud disk.

> **Note:**
> After the system disk is changed, you can still use those manually created snapshots of the released disk to create custom images. If you have applied an automatic snapshot policy to the old system disk and set the automatic snapshots to release when the disk is released, you must apply the policy to the new disk. Besides, all the automatic snapshots of the old disk are released.

Limits and recommendations

When changing the system disk, you must consider the following:

- After the system disk is changed, your instance is assigned a new cloud disk as the system disk, with a new disk ID, and the old one is released.
- You cannot replace the Cloud Type of the system disk.
- After expansion, the minimum capacity is as much as before, while the maximum capacity is 500 GiB. The capacity of the system disk cannot be reduced.
- You cannot increase the size of the system disk that runs Windows 2003.
- If your Subscription instance has been *renewed for configuration downgrade*, you cannot modify the system disk capacity until you enter the next billing cycle.

- The IP address and the MAC address remain unchanged after the system disk is changed.
- We recommend that you create a snapshot for the system disk before you change the disk. Consider the following when creating the snapshot:
  - 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.
  - Make sure the system disk has enough available storage space when creating a snapshot, at least 1 GiB. Otherwise, the system may fail to start after the system disk is changed.
- To make sure you have enough quota for automatic snapshots of the new system disk, you can delete the unnecessary snapshots of the old system disk. For more information, see *Delete snapshots or automatic snapshot policies*.

Procedure

If you want to increase the size of the system disk while keeping the operating system and environment intact, follow these steps:

1. Log on to the *ECS console*.
2. In the left-side navigation pane, click **Instances**.
3. Select a region.
4. Find an instance to change the system disk, click the instance ID to go to the **Instance Details** page.
5. Follow these steps to create a snapshot of the system disk:
   a. In the left-side navigation pane, click **Instance Disks**.
   b. Locate the system disk, find **Actions**, and click **Create Snapshot**.

   > **Note:**
   > For more information about the limits or note for creating a snapshot, see *Create snapshots*.
6. Follow these steps to create a custom image by using the snapshot:
   a. In the left-side navigation pane, click **Instance Snapshots** to check the creation status and progress. When the progress is 100% and the status is **Success**, in the **Actions** column, click **Create Custom Image**.

Note:

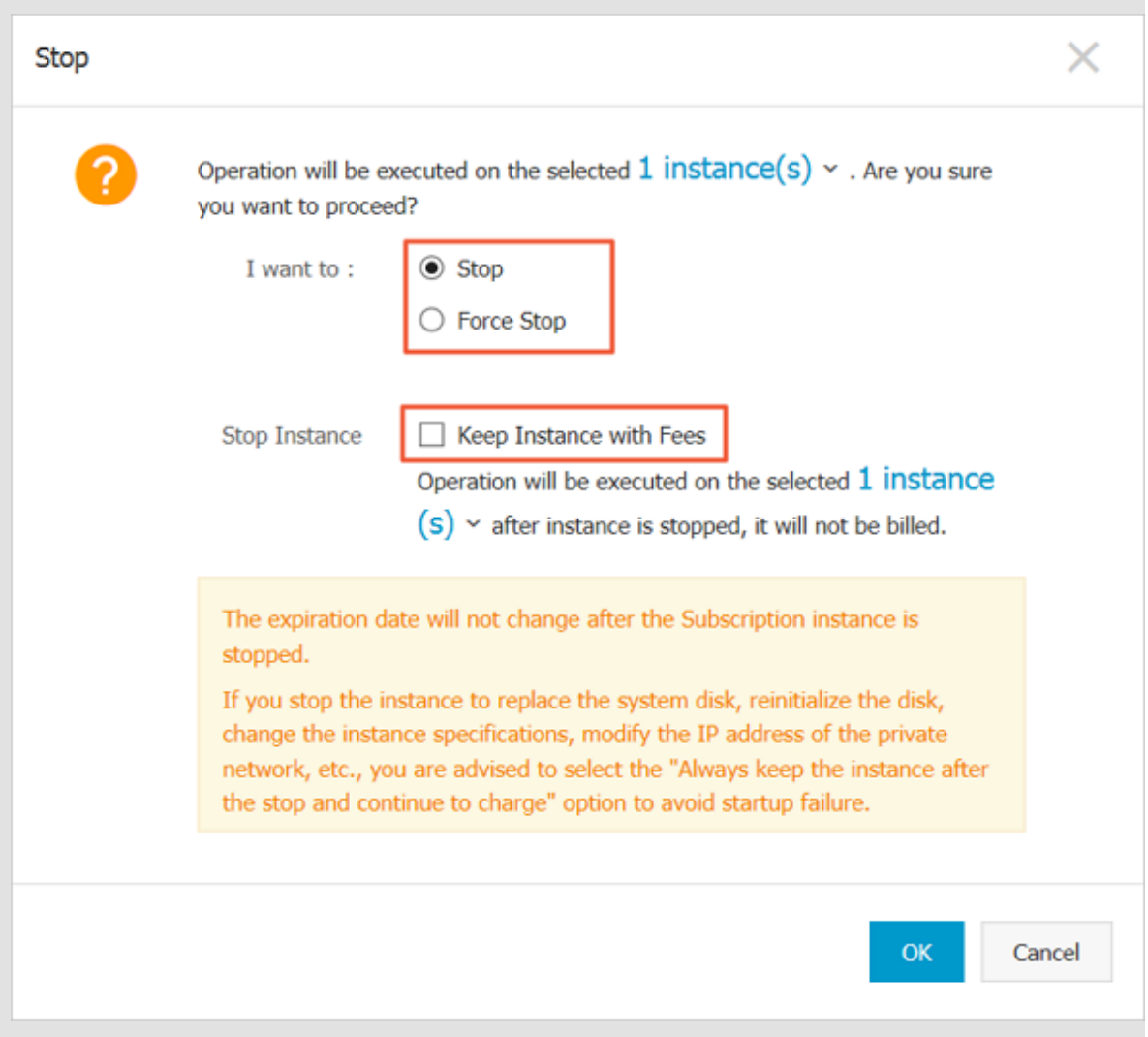
- For more information about the limits or note for creating a custom image, see *Create a custom mirror using a snapshot*.
- The custom image is displayed in the dropdown ist of the **Custom Image** l on the Replace System Disk page.

b. Go back to the **Instances**. In the left-side navigation pane, select **Snapshots & Images** > **Image**to check the creation status and progress of the custom image.

7. When the progress is 100% and the status is **Available**, in the left-side navigation pane, click **Instances**.
8. In the Instance List, find the instance, and in the **Actions** column, select **More** > **Stop**.

Note:

For a Pay-As-You-Go VPC-Connected ECS instance, if the *No fees for stopped instances (VPC-Connected)* feature is enabled, in the **Notice** dialog box, click **OK**, On the 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)** you may not be able to start the instance successfully after changing the system disk.



9. When the instance is in the **Stopped** status, in the Actions column, select**Action More** > **Replacing the system tray**.
10. In the pop-up dialog box, after carefully reading the notes about replacing the system tray, click **OK, replace the system disk**.
11. On the **Change System Disk** page, complete the configurations:
    a. **Image type**: Click the **Custom Image** tab and select the created custom image in the drop-down list.
    b. **System disk**: Specify a new size for the system disk according to your business needs. The maximum size is 500 GiB. 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:
You cannot modify the Cloud Type of the system tray.

c. **Security**:

- If a Windows image is used, set a logon password.
- If a Linux image is used and the instance is I/O optimized, you can choose to set a password or bind an SSH key pair for logon.

d. Confirm **Instance Cost**: For more information about pricing, see *Pricing of Elastic Compute Service*.

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

Go back to the ECS console to check the status of the process. It may take a few minutes to process the change. After the system disk is changed, the instance starts automatically.

Follow-up operations

After the system disk is changed, you may have to perform the following:

- If your instance is running a Linux image and any data disk has been attached to the instance and set to automatically mount the file systems at the beginning, the mount information is lost while changing the system disk. Therefore, you must write the new partition and mounting information to the `/etc/fstab` file on the new system disk and mount the file systems. You must not partition or format the data disk again. Follow these steps to mount the file systems. For more information about the commands, see *Linux _ Format and mount a data disk*:

  1. (Optional) Make backup of the `/etc/fstab file`.
  2. Write the new partition and mounting information to the `/etc/fstab` file.
  3. Check the new partition information in the `/etc/fstab` file.
  4. Mount the file systems.

5. To view disk space and usage: run the command `df -h`.

   After mounting, do not need to restart the instance to use the new file system directly.

- *Apply automatic snapshot policies to disks*. Optionally, apply an automatic snapshot policy to the new system disk. The link between an automatic snapshot policy ID and a disk ID is broken after the system disk is changed. You need to set up an automatic snapshot policy for the new system disk.

11.6.3 Linux _ Resize a data disk

As your business grows, the current capacity of your data disks may not be able to meet your data storage needs. You can use the **Resize Disk** feature to resize your data disks as necessary.

Note:

- Resize the data disks that are attached to an instance only when the instance is in the **Running** or **Stopped** status. **You must restart the instance in the ECS console to apply the changes. This action causes your instance to stop working and may cause your business to be interrupted, so please proceed with caution.**
- We recommend that you manually create a snapshot to back up your data before resizing your data disk.
- You can resize a data disk when the data disk is either in the **Available** status or in the **In Use** status.
- If you have renewed a Subscription ECS instance for configration downgrade (*Renew for configuration downgrade*), during its current billing cycle, you cannot resize the attached Subscription cloud disks, including its data or system disks.
- If a snapshot is being created for a data disk, you cannot resize the data disk.
- You can resize data disks, but not system disks or local disks.

This example uses a data disk of the ultra cloud disk type and an ECS instance running 64-bit CentOS 7.3 to describe how to resize data disk and extend the available capacity.

To resize a data disk, follow these steps:

Step 1. Increase the size of a data disk in the ECS console

Step 2. Log on to the instance to resize the file system

Step 1. Increase the size of a data disk in the ECS console

To increase the size of a data disk in the ECS console, follow these steps:

1. Log on to the *ECS console*.
2. In the left-side navigation pane, select **Block Storage** > **Cloud Disks**.

   Note:
   If the data disk you want to resize has been attached to an instance, in the left-side navigation pane, click **Instances**, find the corresponding instance, go to the instance details page, and click **Instance Disks**.
3. Select a region.
4. Find the disk to be resized, and in the **Actions** column, select **More** > **Resize Disk.**.
5. On the **Resize Disk** page, set **Capacity after resizing** (In this example, 30 GiB). The capacity after resizing must be larger than the current capacity.
6. When the cost is calculated, click **Confirm to resize**.

   Note:
   After the resizing, you can view the new disk size in the console. However, if the data disk is attached to an ECS instance, you must *Restart an instance* restart the instance in the ECS console to view the new disk size when you log on to the instance.

After the disk size is increased,

- If the data disk is attached to an instance, *Step 2. Log on to the instance to resize the file system*.
- If the data disk is not attached to an instance, attach the disk to an instance in the console (*Attach a cloud disk*) first, and then proceed depending on the data disk:
  - If it is a new data disk, which has not been formatted, format it. For more information, see *Format and mount data disks for Linux instances*.
  - If it has been formatted and partitioned, *Step 2. Log on to the instance to resize the file system*.

Step 2. Log on to the instance to resize the file system

After the disk size is increased, you must log on to the instance to resize the file system.

In this example, the data disk is attached to a Linux instance running the 64-bit CentOS 7.3. The data disk before resizing has only one primary partition (/dev/vdb1, ext4 file system), the mount point of the file system is `/resizetest`, and after resizing is completed, the data disk still has only one primary partition.

1. *Connect to a Linux instance by using a password*.
2. Run the `umount [file system name]` command to unmount the primary partition.

```
umount /dev/vdb1
```

> **Note:**
> Run the `df -h` command to check whether the unmounting is successful. If you do not see the /dev/vdb1 information, unmounting is successful. The following is the sample output.

```
[root@iXXXXXX ~]# df -h
Filesystem Size Used Avail Use% Mounted on
/dev/vda1 40G 1.5G 36G 4% /
devtmpfs 487M 0 487M 0% /dev
tmpfs 497M 0 497M 0% /dev/shm
tmpfs 497M 312K 496M 1% /run
tmpfs 497M 0 497M 0% /sys/fs/cgroup
tmpfs 100M 0 100M 0% /run/user/0
```

3. Run the `fdisk` command to delete the original partition and create a new partition:

> **Note:**
> If you use the `parted` tool to manipulate partitions, you cannot use it in conjunction with `fdisk`. Otherwise, this results in an inconsistent first sector of the partition. Instructions on how to use the `parted` tool can be found *here*.

   a. Run the `fdisk -l` command to list the partition details and record the final size of the partition and its first sector before resizing.
   b. Run the `fdisk [device name of data disk]` command to go to `fdisk` . In this example, the device name is `/dev/vdb`.
   c. Type `d` and press the Enter key to delete the original partition.

> **Note:**
> Deleting a partition does not cause loss of data in the data disk.

   d. Type `d` and press the Enter key to start creating a new partition.
   e. Type `p` and press the Enter key to create a primary partition. In this example, you are creating a single-partition data disk, so it is sufficient to create one primary partition.

> **Note:**
> If you want to create more than four partitions, create at least one extended partition, that is, type `e`.

f. Type the partition number and press the Enter key. In this example, only one partition is to be created, so type `1`.

g. Type a number for the First sector: For data consistency, the number for the First sector must be identical with that of the original partition. In this example, press the Enter key to use the default value of 1.

> **Note:**
> If you find that the First sector is not identical with the recorded one, you may have used the `parted` tool for partitioning. In that case, stop the current `fdisk` operation and use *`parted`* to start over again.

h. Type a number for the last sector: Because only one partition is to be created in this example, press the Enter key to use the default value.

i. Type `wq` and press the Enter key to start partitioning.

```
[root@iXXXXXX ~]# fdisk /dev/vdb
Welcome to fdisk (util-linux 2.23.2).
Changes will remain in memory only, until you decide to write them
.
Be careful before using the write command.
Command (m for help): d
Selected partition 1
Partition 1 is deleted
Command (m for help): n
Partition type:
p primary (0 primary, 0 extended, 4 free)
e extended
Select (default p):
Using default response p
Partition number (1-4, default 1):
First sector (2048-62914559, default 2048):
Using default value 2048
Last sector, +sectors or +size{K,M,G} (2048-62914559, default
62914559):
Using default value 62914559
Partition 1 of type Linux and of size 30 GiB is set
Command (m for help): wq
The partition table has been altered!
Calling ioctl() to re-read partition table.
Syncing disks.
```

> **Note:**
> If you are using the `parted` tool, type `p` in the `parted` window to list the current partition details. If any partition is displayed, use rm + serial number to delete the original partition table, then run the `unit s` command to specify the start unit, calculated by the number of sectors, and finally run the `mkpart` command to create it, as shown in the following figure.

```
[root@          ~]# parted /dev/xvdb
GNU Parted 3.1
Using /dev/xvdb
Welcome to GNU Parted! Type 'help' to view a list of commands.
(parted) p
Model: Xen Virtual Block Device (xvd)
Disk /dev/xvdb: 5369MB
Sector size (logical/physical): 512B/512B
Partition Table: gpt
Disk Flags:

Number  Start  End  Size  File system  Name  Flags

(parted) unit s
(parted) mkpart primary ext3 56 5369MB
Warning: The resulting partition is not properly aligned for best performance.
Ignore/Cancel? i
(parted) p
Model: Xen Virtual Block Device (xvd)
Disk /dev/xvdb: 10485760s
Sector size (logical/physical): 512B/512B
Partition Table: gpt
Disk Flags:

Number  Start  End        Size       File system  Name     Flags
 1      56s    10485726s  10485671s  ext3         primary
```

4. For some operating systems, the file system may be automatically mounted to the mount point after partitioning. We recommend that you run the `df -h` command to check the file system space and usage. Run the `umount [file system name]` to unmount the file system again.
5. Check the file system and resize the file system.

```
e2fsck -f /dev/vdb1 # check the file system
resize2fs /dev/vdb1 # resize the file system
```

Note:

- Running the `e2fsck` command is time-consuming because the system needs to check and revise the file system metadata during that process, so be patient.
- Properly running the `e2fsck` command and the `resize2fs` command does not cause data loss.

The following is the sample output.

```
[root@iXXXXXX ~]# e2fsck -f /dev/vdb1
e2fsck 1.42.9 (28-Dec-2013)
Pass 1: Checking inodes, blocks, and sizes
Pass 2: Checking directory structure
Pass 3: Checking directory connectivity
Pass 4: Checking reference counts
```

```
Pass 5: Checking group summary information
/dev/vdb1: 11/1835008 files (0.0% non-contiguous), 159218/7339776
blocks
[root@iXXXXXX ~]# resize2fs /dev/vdb1
resize2fs 1.42.9 (28-Dec-2013)
Resizing the filesystem on /dev/vdb1 to 7864064 (4k) blocks.
The filesystem on /dev/vdb1 is now 7864064 blocks long.
```

6. Mount the resized file system to the original mount point (in this example, */resizetest*).

```
mount /dev/vdb1 /resizetest
```

7. Run the `df -h` command to check file system space and usage. If the correct information about the resized file system is displayed, the mounting is successful and the resized file system is ready for use.

Note:

After the mounting is completed, you can use the resized file system without restarting the instance.

The following is the sample output.

```
[root@iXXXXXX ~]# df -h
Filesystem Size Used Avail Use% Mounted on
/dev/vda1 40G 1.5G 36G 4% /
devtmpfs 487M 0 487M 0% /dev
tmpfs 497M 0 497M 0% /dev/shm
tmpfs 497M 312K 496M 1% /run
tmpfs 497M 0 497M 0% /sys/fs/cgroup
tmpfs 100M 0 100M 0% /run/user/0
/dev/vdb1 30G 44M 28G 1% /resizetest
```

11.6.4 Windows _ Resize a data disk

As your business grows, the current capacity of your data disks may not be able to meet your data storage needs. You can use the **Resize Disk** function to resize your data disks as necessary.

Note:

- We recommend that you manually create a snapshot to back up your data before resizing a data disk.
- You can resize a data disk when the data disk is either in the **Available** status or in the **In Use** status.
- If a snapshot is being created for a data disk, you cannot resize the data disk.

- If you have renewed a Subscription ECS instance for configration downgrade (*Renew for configuration downgrade*) during its current billing cycle, you cannot resize the attached Subscription cloud disks, including its data or system disks.
- You can resize data disks, but not file system.
- You can resize data disks, but not system disks or local disks.
- Resize the data disks that are attached to the instance only when the instance is in the **Running** (**Running**) or **Stopped** (**Stopped**) status. The changes are applied when you restart the instance in the ECS console. This action stops your instance from working and interrupts your business. Hence, **proceed with caution**.

This example uses a data disk of the ultra cloud disk type and an ECS instance running 64-bit Windows Server 2008 R2 Enterprise Edition to show how to resize a data disk and extend the available capacity. In this example, the current disk capacity is 20 GiB, and we resize it to 24 GiB.

To resize a data disk, follow these steps:

Step 1. Resize a data disk in the ECS console

Step 2. Log on to the instance to enable the extended storage space

Step 1. Resize a data disk in the ECS console

To resize a data disk in the ECS console, follow these steps:

1. Log on to the *ECS console*.
2. In the left-side navigation pane, choose **Block Storage** > **Cloud Disks**.

Note:

If the data disk you want to resize is attached to an instance, click **Instances** in the left-side navigation pane, find the instance, go to the Instance Details page, and then click. **Instance Disks**.

3. Select a region.
4. Find the disk to be resized, and in the **Actions** column, choose **More** > **Resize Disk**.
5. On the **Resize Disk** page, set **Capacity after resizing**. In this example, 24 GiB. The capacity after resizing must be larger than the current capacity.
6. When the cost is calculated, click. **Confirm to resize**.

Note:

If your data disk is attached to an instance, restart the instance in the ECS console *Restart an instance* to make the disk resize take effect.

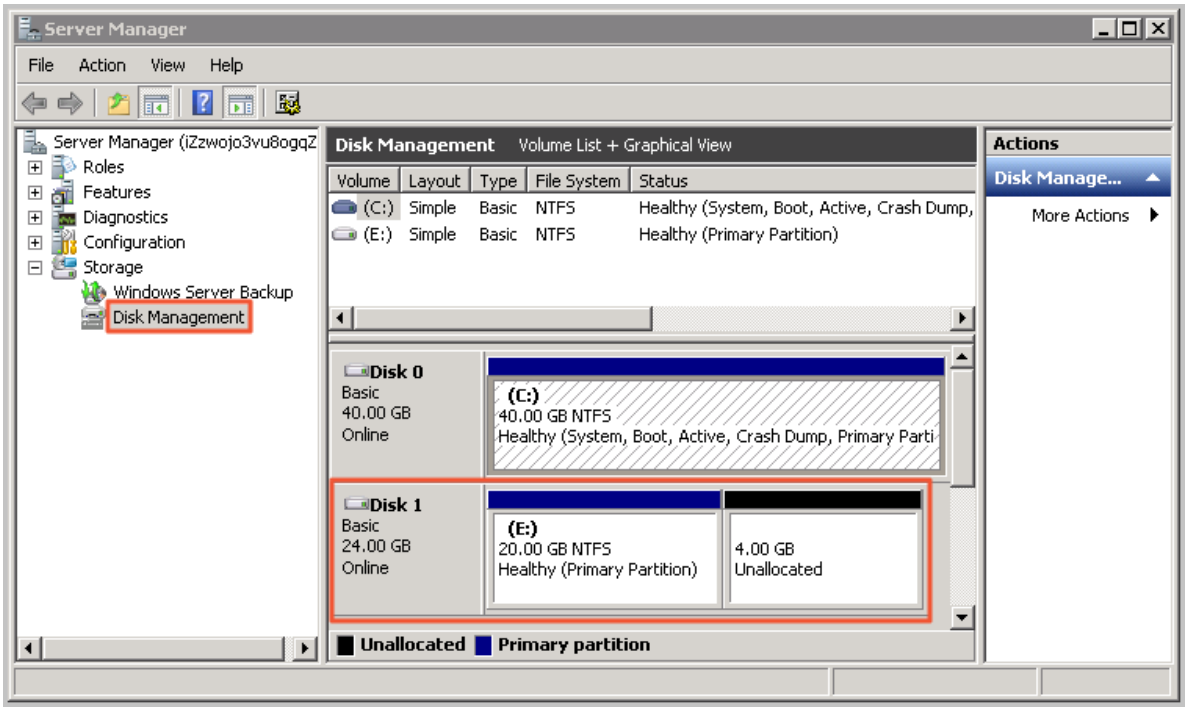
Once the data disk resizing completes, you can do the following:

- If the data disk is attached to an instance, *Step 2. Log on to the instance to enable the extended storage space*.
- If the data disk is not attached to an instance, attach the disk to an instance in the console (x id="1"/>) first, and then proceed depending on the data disk:
  - If it is not formatted or partitioned, format and mount the data disk. For more information, see *Format a data disk for Windows instances*.
  - If it is formatted and partitioned, *Step 2. Log on to the instance to enable the extended storage space*.

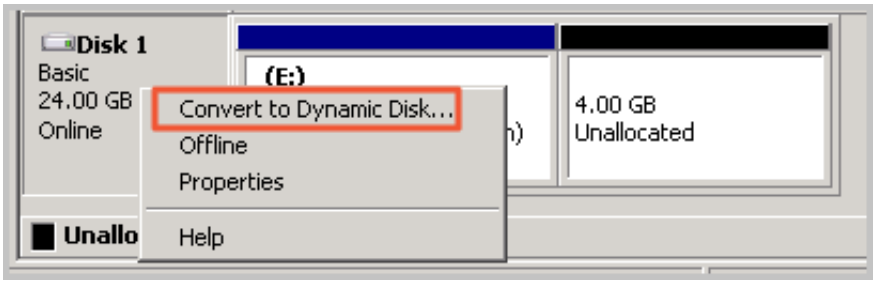
Step 2. Log on to the instance to enable the extended storage space

To resize a data disk within the instance, follow these steps:

1. *Connect to a Windows instance*.
2. On the Windows Server desktop, click the Server Manager icon  .
3. In the left-side navigation pane of **Server Manager**, choose **Storage** > **Disk Management**. In the disk management area, you can see the relationship between the new and the original data disk spaces. In this example, **Disk 1** is the resized data disk.



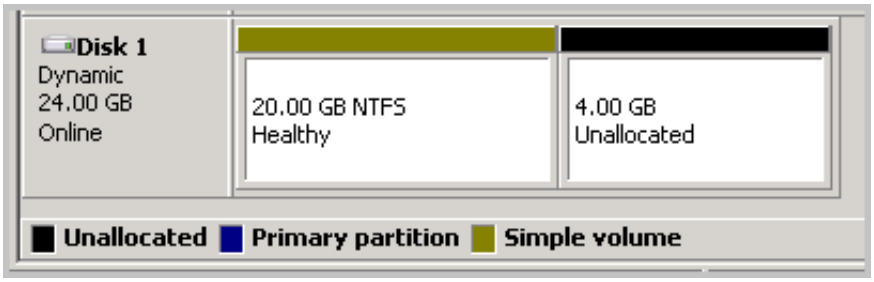
4. Right click **Disk 1**, select **Convert to Dynamic Disk**, and follow the wizard to convert a basic disk to a dynamic disk.



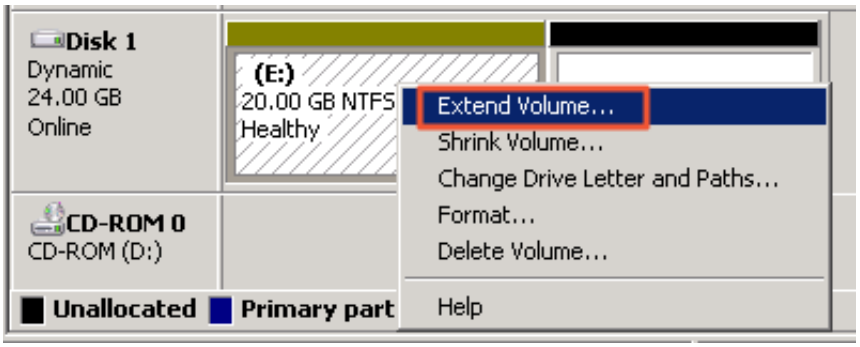
Note:

Converting a basic disk into a dynamic disk unmounts the disk from the system. Applications installed on the data disk, if any, are temporarily unavailable during the conversion process. The conversion process does not cause any data loss.

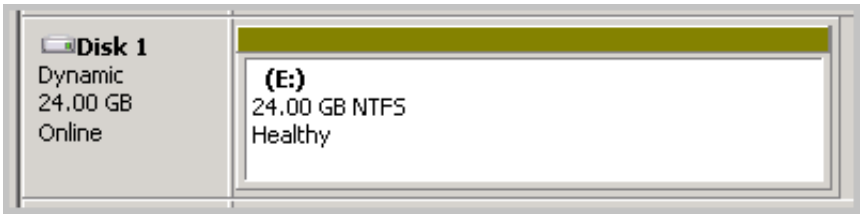
After the conversion, the **Disk 1** information shows in the Disk Manager as follows.



5. Right click an empty area of the simple volume of Disk 1, and select **Extend Volume**.



6. Follow the **Extend Volume Wizard** to extend the volume. When the wizard is complete, the new data disk space is automatically merged into the original volume and the **Disk 1** information showed in the Disk Manager as follows.



Note:
On Windows Server 2003, the extended storage space is added to the data disk but it is displayed as a separate volume in Disk Manager. On Windows Server 2003, one separate volume is created for each expansion and is not merged into the original volume, which does not affect the availability of the extended storage space.

You have resized a data disk successfully and the extended storage space is ready for use.

11.8 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. If you want to keep this part of the data, see *synchronizing data after rolling back the disks* .

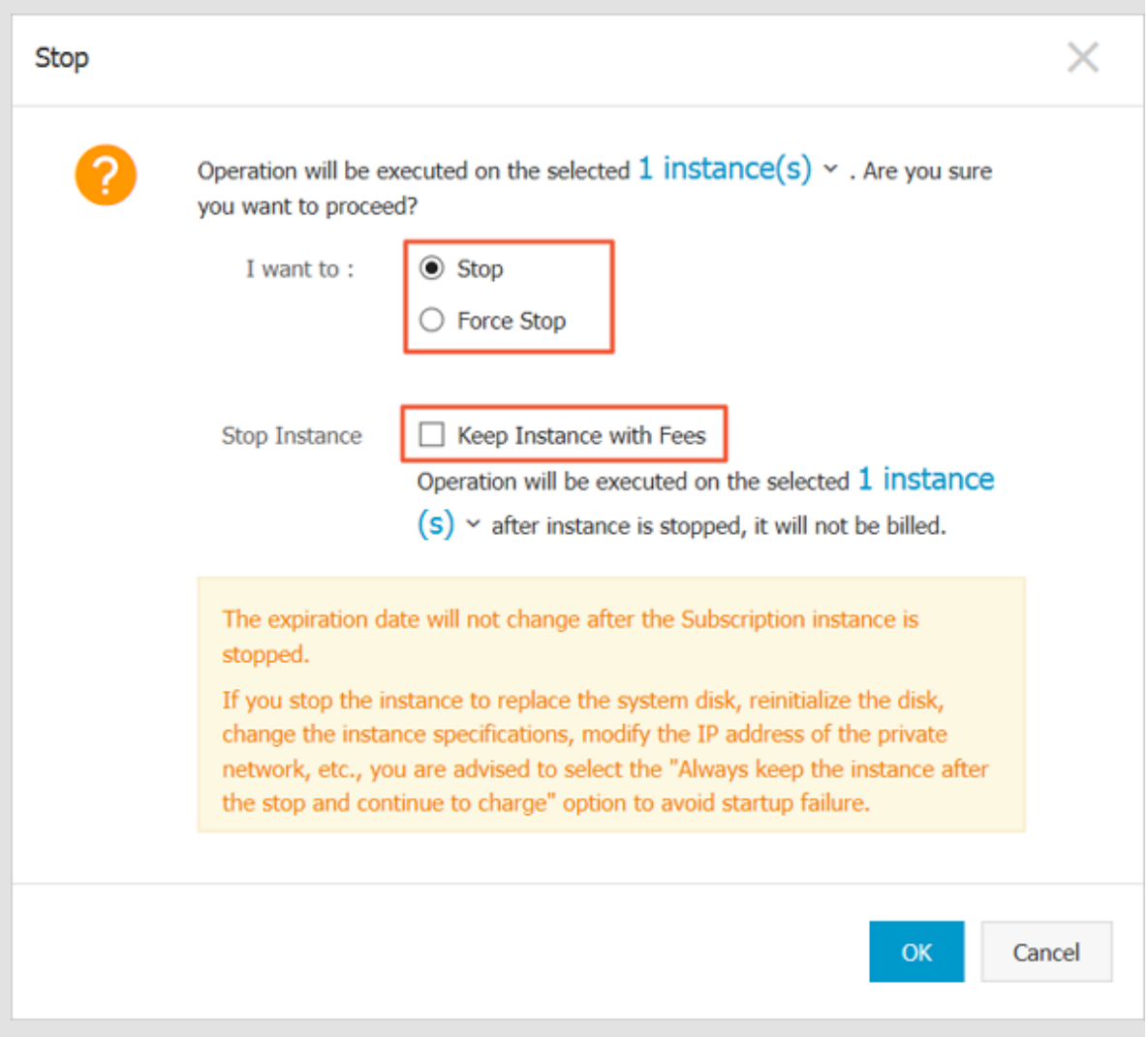
- After a system disk is restored, the logon password or the SSH key pair of the ECS instance is retained.

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*

11.9 Convert billing methods of cloud disks

The billing method of a cloud disk depends on how it is created:

- For cloud disks created with Subscription (monthly, yearly or weekly 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	Effective data	Suitable for
Subscription —> Pay-As-You-Go	*Renew for configuration downgrade*	Effective from the next billing cycle	Subscription (monthly or yearly) cloud disks attached to Subscription instances. The billing method of the system disk cannot be changed. **Note:** Subscription (weekly) instances do not support downgraded configurations. You cannot change the billing method of Subscription

Conversion of billing methods	Features	Effective data	Suitable for
			(monthly or yearly) cloud disks attached to the weekly subscription instances.
Pay-As-You-Go —> Subscription	*Upgrade configurations*	Effective immediately	Pay-As-You-Go data disks attached to Subscription instances. The billing method of the system disk cannot be changed.
	Switch from Pay-As-You-Go to subscription		The system disks and data disks attached to the Pay-As-You-Go instances.

11.10 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, see *Offline announcement of Windows Server 2003 system image*.

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

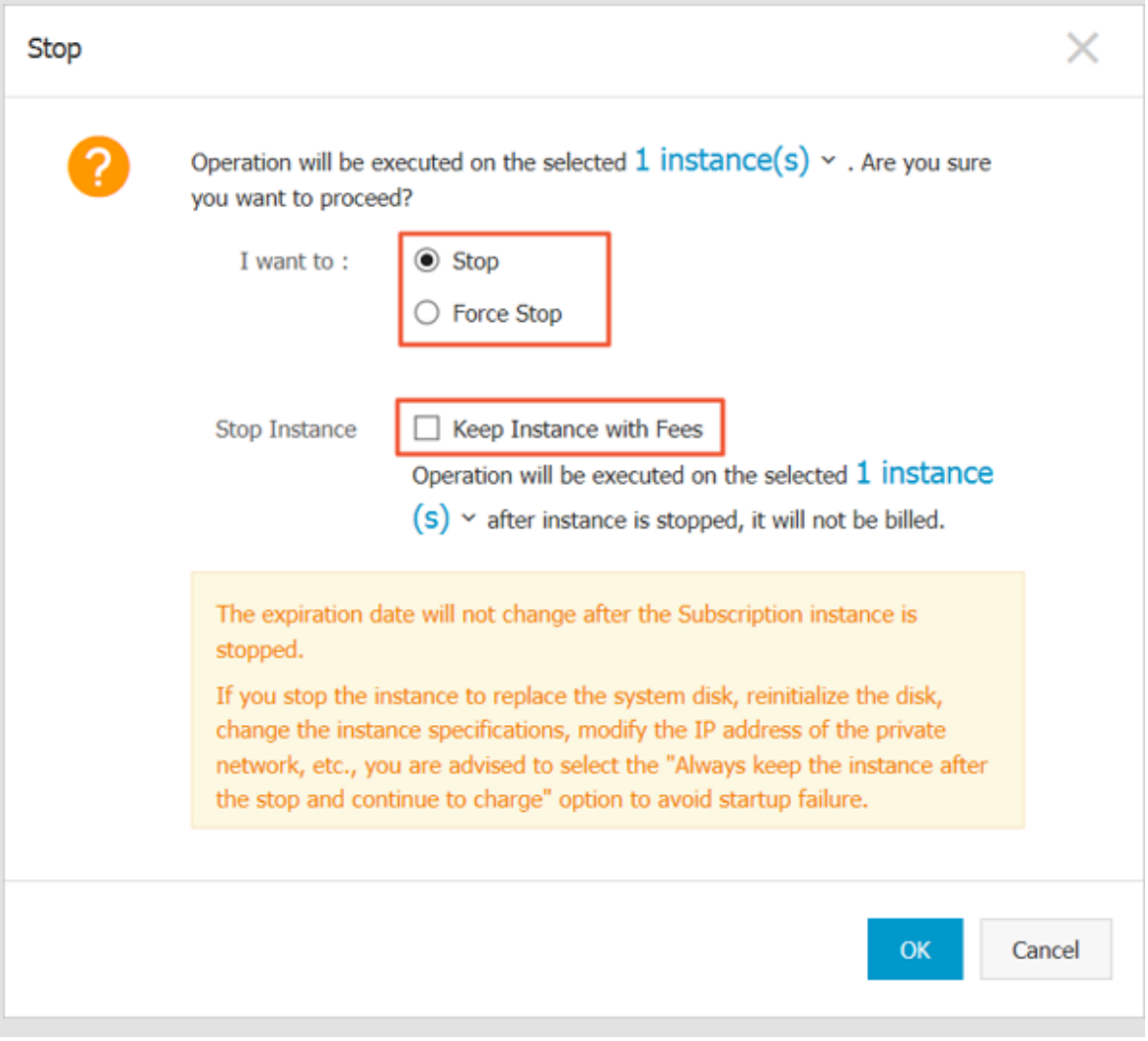
Replace your system disk as follows:

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)*.

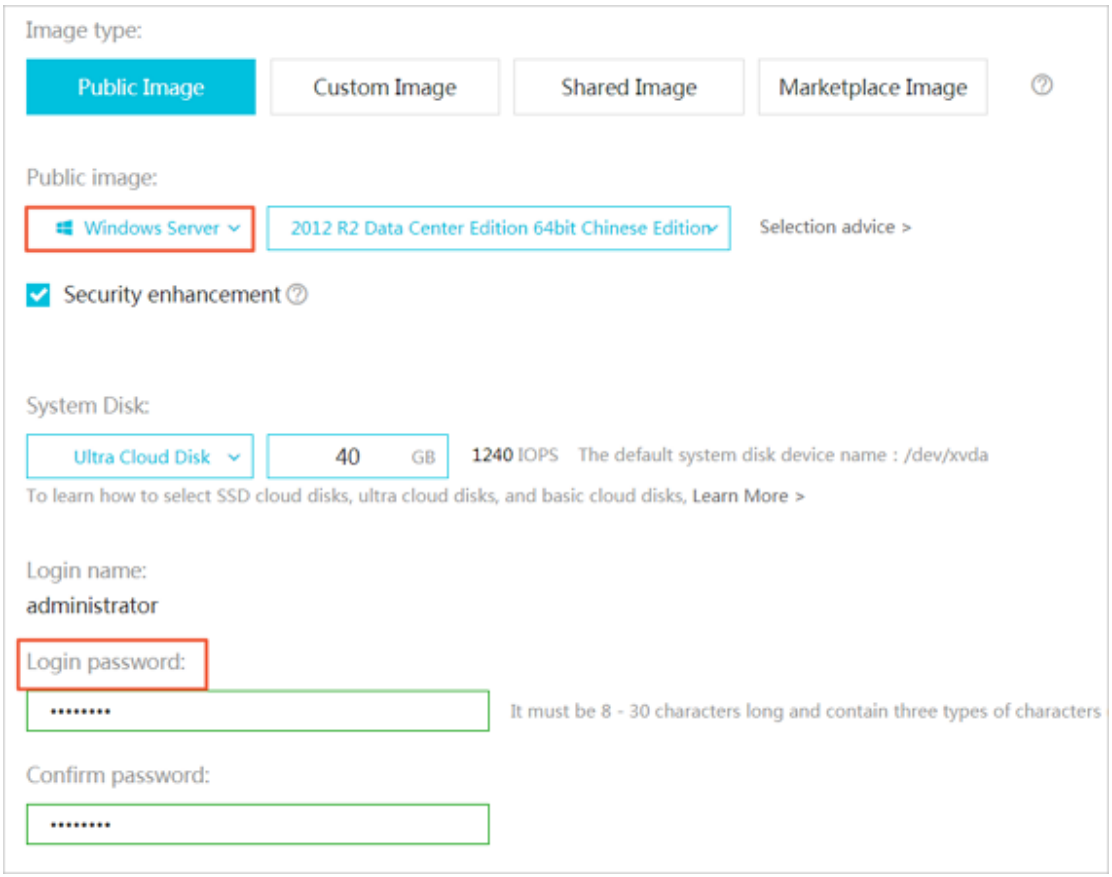
b. **System Disk**: You cannot change the cloud disk category. However, you can change the size of the disk to meet the requirements of your system disk. The maximum size is 500 GiB. The size limit for changing is determined by the image and the current size of the system disk, as displayed in the following 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]

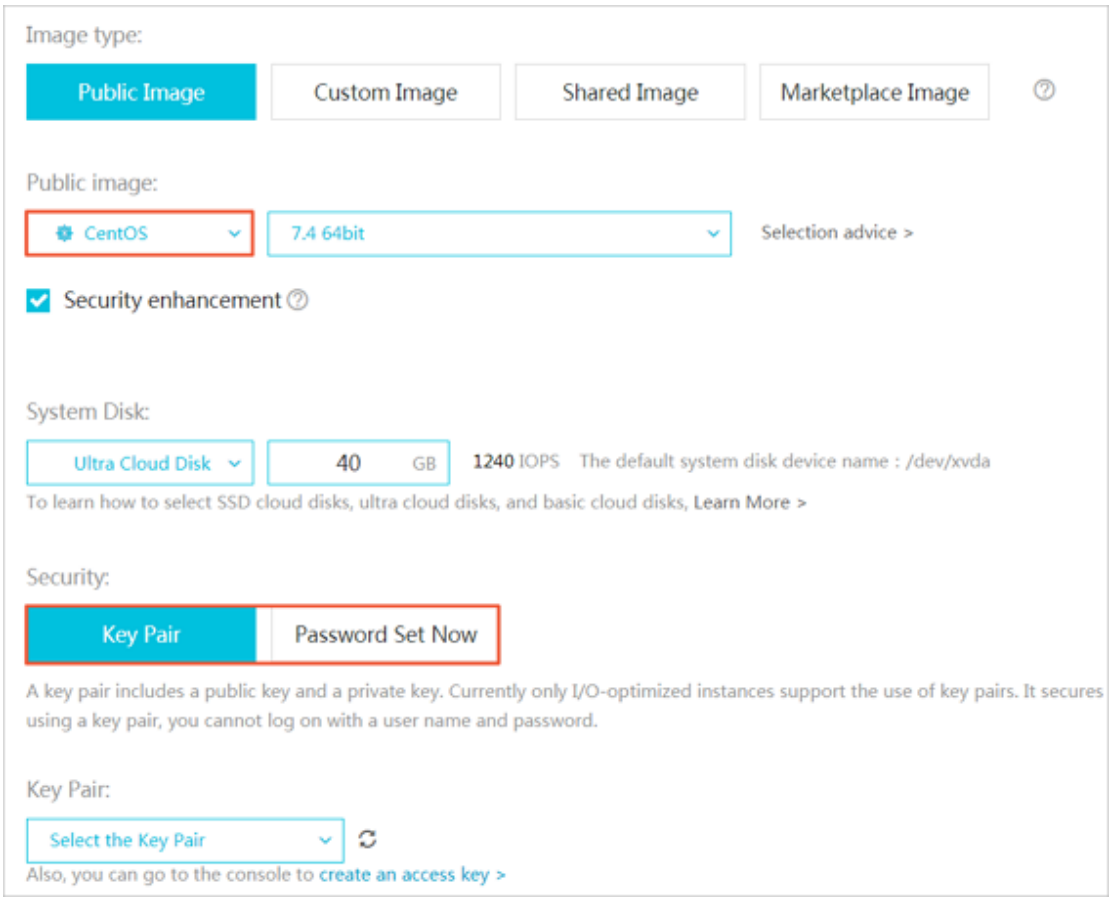
Note:
If your instance was renewed for configuration downgrade, you cannot change the system disk size until the next billing cycle.

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 authentication method. Set a password or bind an SSH key pair.



d. Confirm **Configuration fee**: at present, all public images in China station are not charged, the cost of the configuration here refers to the cost of the system tray. System Tray prices can be found on the *cloud Product Price* page.

> **Note:**
> System Disk price Description: the system tray is sold at the starting capacity, and the starting capacity is priced from the starting capacity, after exceeding the starting capacity, 1 per increase Gib charges at Linear Charge price. Start capacity of the system disk, package annual monthly price in **Yuan/month**, units of pay price per volume are **Yuan/hour**.

e. 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

11.11 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 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*. If you select a custom image that comes from a mirror market, you may be charged for it, also included in the configuration fee. For mirrored billing information, please refer to billing overview.

> **Note:**
> System Disk price Description: System Tray is sold at the starting capacity, priced at the starting price, and exceeds the starting capacity, 1 per increase Gib charges at Linear Charge price. Start capacity of the system disk, package annual monthly price in Yuan/ month, units of pay price per volume are Yuan/hour.

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 automatica lly 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 informatio n 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

11.12 Monitor a cloud disk

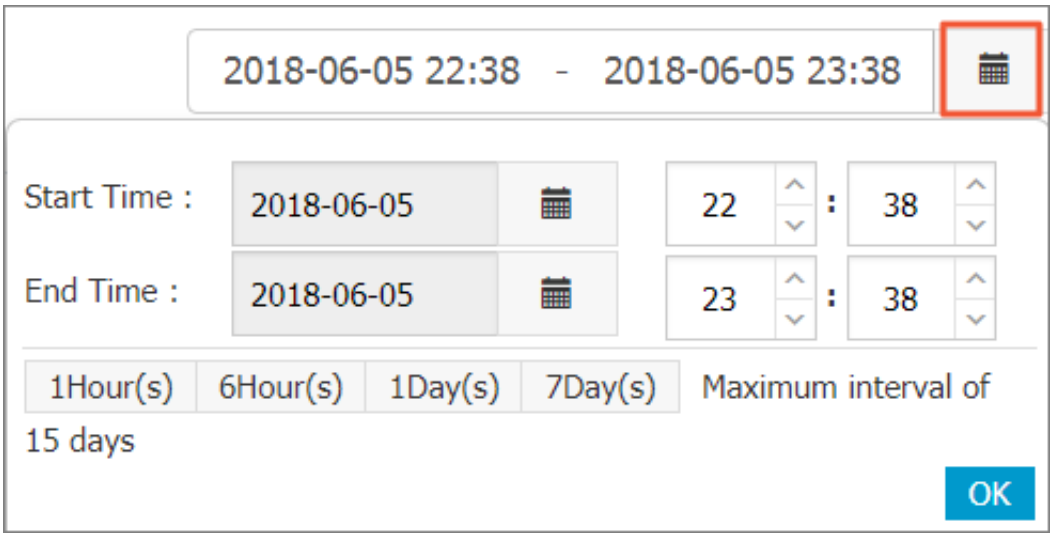
When using a cloud disk, consider the following performance parameters:

- 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:

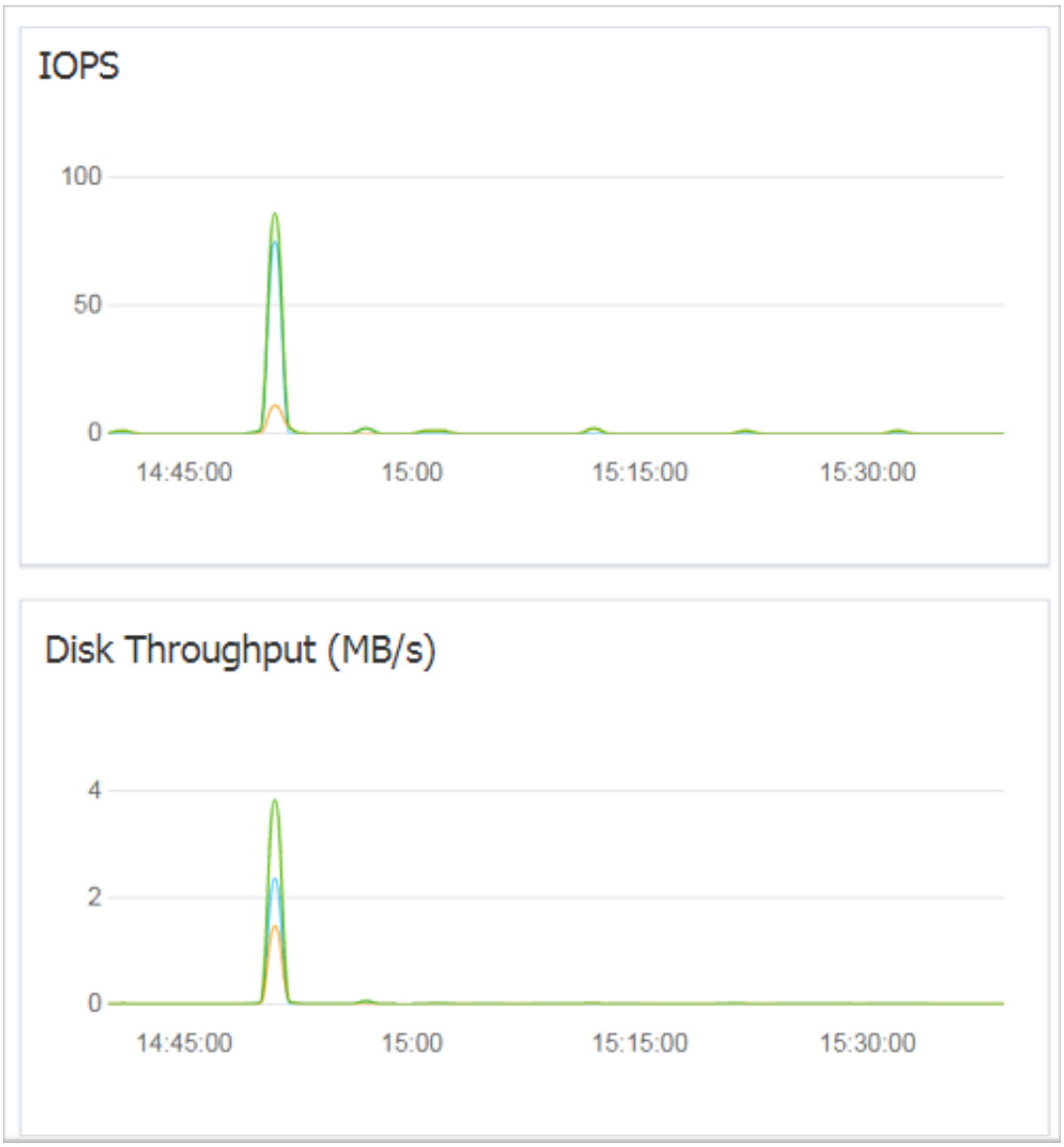
1. Log on to the *ECS console*.
2. In the left-side navigation pane, select **Block Storage** > **Cloud Disks**.
3. Select a region.
4. Find a cloud disk and click its ID to go to the **Disk Details** page.
5. In the left-side navigation pane, click **Disk Monitoring**.
6. On the **Monitoring Information** page, click the  icon 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.



11.13 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:
  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

12 Snapshots

12.1 Create snapshots

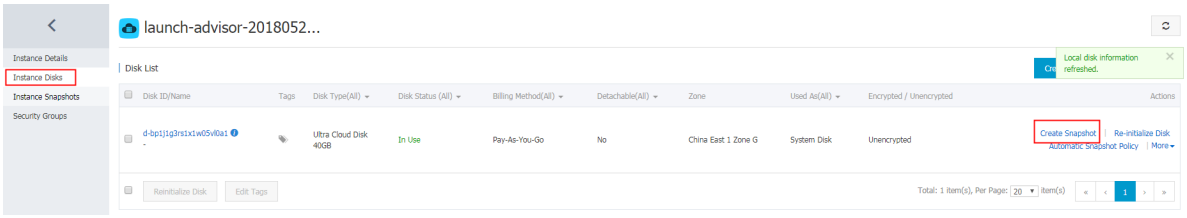
You can create instance snapshots to save the system state from a certain point in time for data backup or to create images.

Note:

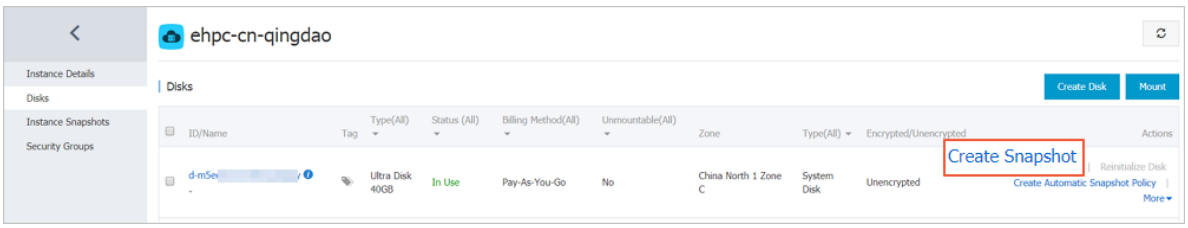
- Creation of the first snapshot will take relatively longer than subsequent snapshots due to the first snapshot being a full snapshot. However, depending on the amount of changed data since previous snapshots, the length of time for each snapshot creation may vary.
- Creating snapshots of a disk may reduce disk performance.
- We recommended that you not create snapshots during peak traffic hours.
- Manually created snapshots, unlike automatic snapshots, are retained until they are manually deleted.

Procedure

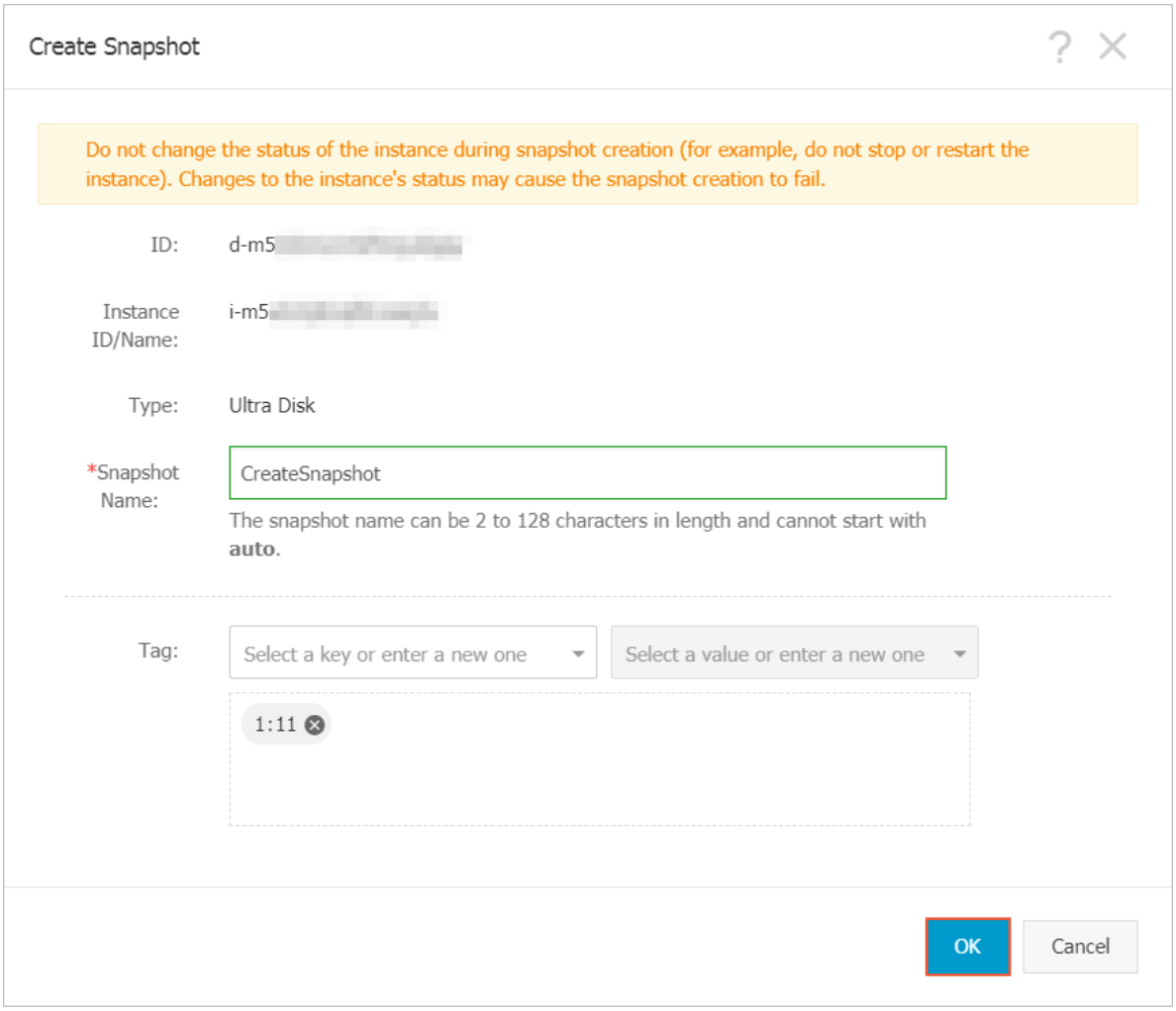
1. Log on to the *ECS console*.
2. Select a region. In the left-side navigation pane, click Instances, and click **Manage**.



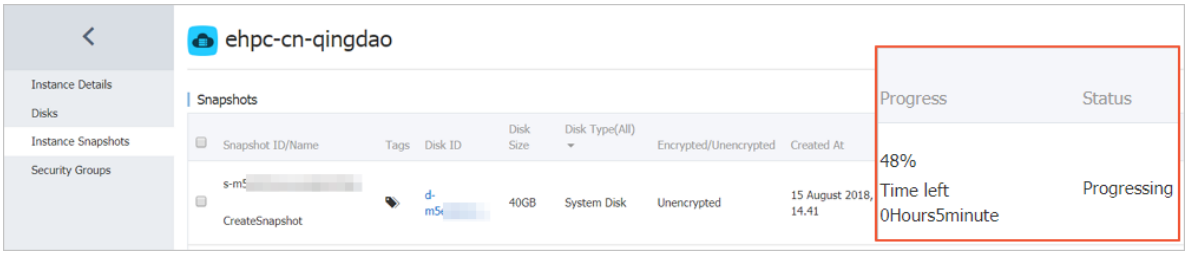
3. In the left-side navigation pane, click **Instance Disks**, and click **Create Snapshot** for the target disk. You can select only one disk at a time, either system disk or data disk.



4. Enter the name for the snapshot, and click **OK**.



5. To view the snapshots, click **Instance Snapshots** from the left-side navigation pane. You can see the progress and status of the snapshot.



12.2 Create and delete an automatic snapshot policy

An automatic snapshot policy is a set of defined parameters for automatically creating snapshots.

Note:

- Starting in March 28, 2017, the snapshot service starts charging fees. For more information on Snapshot charges, see the .*snapshot commercialization FAQ*.

- Avoid business peak hours when you set the automatic snapshot creation time and repeated date, because creating a snapshot may slightly impact the performance of the disk.
- You can create a maximum of 100 automatic snapshot policies in a region.

Prerequisite

If you want to modify an automatic snapshot policy, you must first create an automatic snapshot policy.

Procedure

To create an automatic snapshot policy, follow these steps:

1. Log on to the *ECS console*.
2. In the left-side navigation pane, select **Snapshots and Images** > **Automatic Snapshot Policy**.
3. On the **Create Automatic Snapshot Policy** page,
   - if you want to create a policy, click **Create Policy** at the upper-right corner.
   - if you want to modify a policy, find the policy that you want to modify, and click **Modify Policy** in the **Actions** column.
4. In the **Create Policy** or **Modify Policy** dialog box, define the automatic snapshot policy:
   - Enter a policy name.
   - Select a time after **Executed At** to specify the time of day for automatically creating snapshots. There are 24 time points available between 00:00 and 23:00.
   - Specify the **Execution Frequency**. There are seven repetition days available.
   - Set a period after **Keep Snapshots** to defines the number of days a snapshot can be retained. This parameter can be set between 1–65535 days, or permanently. By default, it is set to 30 days. You can also choose to keep automatic snapshots always.

     Note:
     When the number of snapshots reaches the limit, the system automatically removes the oldest automatic snapshots created. Manually created snapshots are not affected.
5. Click **OK**.

Follow-up operations

You can *apply automatic snapshot policies to disks*.

Related APIs

- *CreateAutoSnapshotPolicy*: Creates automatic snapshot policies.
- *DescribeAutoSnapshotPolicyEx*: Queries automatic snapshot policies.
- *ModifyAutoSnapshotPolicyEx*: Modifies automatic snapshot policies.

12.3 Apply automatic snapshot policies to disks

You can apply an automatic snapshot policy to disks according to your business needs.

Automatic snapshots are named in the format of auto_yyyymmdd_1, for example, auto_20140418_1.

Note:

- Creating snapshots may disturb read and write operations on your disk. We recommend that you set the creation time of automatic snapshots to periods when service load is low to reduce effects on your service.
- Automatic snapshot policies cannot be applied to basic cloud disks when they are not in use.
- Snapshots that are manually created do not conflict with automatic snapshots. However, if an automatic snapshot is being created on a disk, you must wait for it to finish before manually creating a snapshot.

You can apply an automatic snapshot policy to a disk in two ways:

- From the Cloud Disks menu: For applying an automatic snapshot policy to a specific disk.
- From the Snapshots and Images menu: For applying a unified automatic snapshot policy to several or all disks.

From the Cloud Disks menu

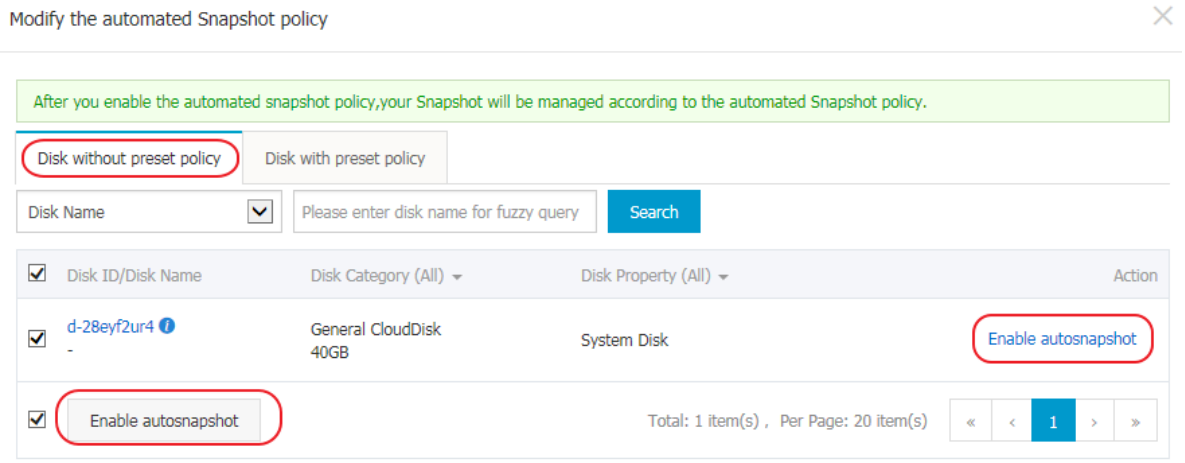
To apply an automatic snapshot policy through the Cloud Disks menu, follow these steps:

1. Log on to the *ECS console*.
2. Select a region.
3. In the left-side navigation pane, click Cloud Disks.
4. Select the disk for which you want to execute the policy and click Automatic Snapshot Policy.
5. Enable the automatic snapshot function and select the desired snapshot policy.
6. Click **OK**.

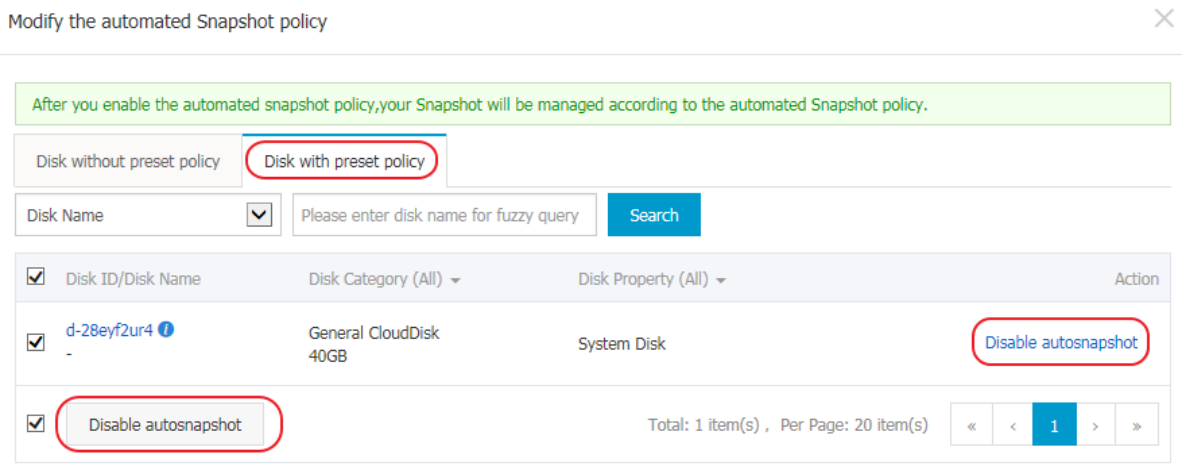
From the Snapshots and Images menu

To apply or disable an automatic snapshot policy, follow these steps:

1. Log on to the *ECS console*.
2. Select a region. You can see a list of all automatic snapshot policies in the region.
3. In the left-side navigation pane, select **Snapshots and Images** > **Automatic Snapshot Policies**.
4. Select the automatic snapshot policy you want to apply and click **Apply Policy**.
5. To enable an automatic snapshot policy, select **Disks without Policy Applied** to view the disks. Find the disk for which you want to enable the policy, and then click **Apply Policy** after it. Alternatively, after selecting multiple disks, click **Apply Policy** at the lower-left corner.



6. To disable the automatic snapshot policy, select the **Disks with Policy Applied** tab to view the disks. Select the disk for which you want to disable the policy, and then click **Disable Policy** after it. Alternatively, after selecting multiple disks, click **Disable Policy** at the lower-left corner.



12.4 Delete automatic snapshots when releasing disks

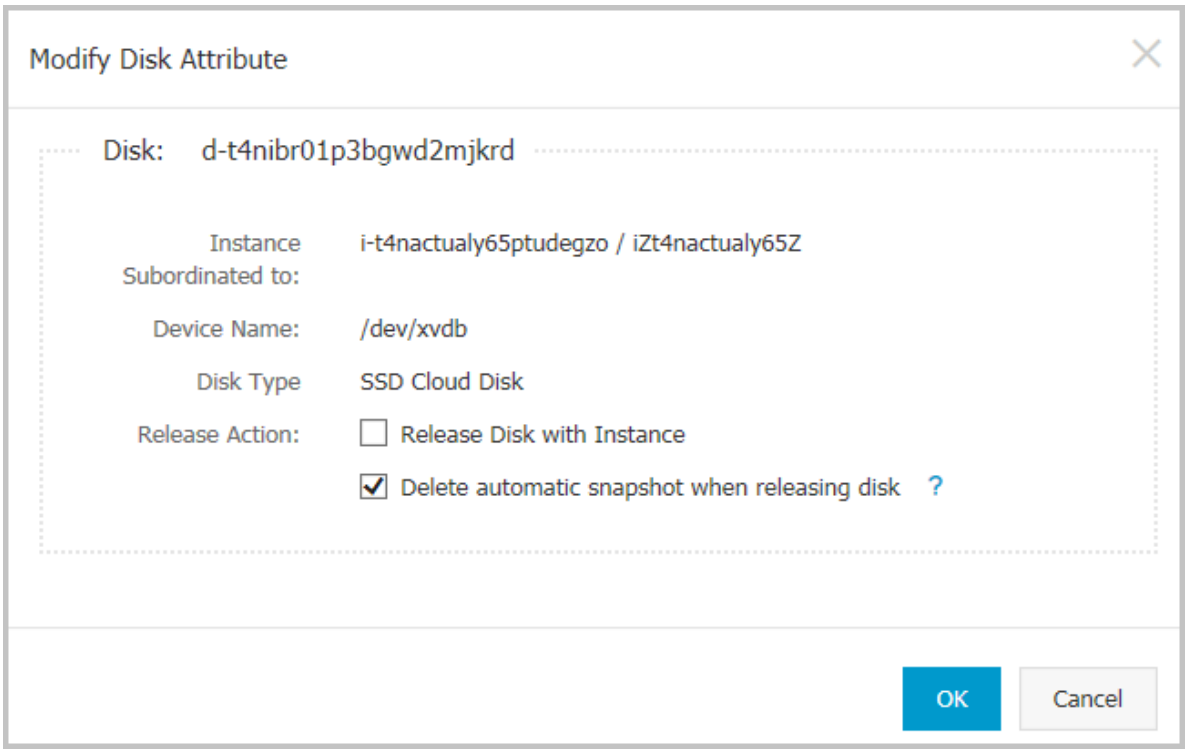
The automatic snapshots of cloud disks are not released along with the cloud by default. However, you can change the disk property so that automatic snapshots are released when you:

- *Change the system disk (custom image)*: The previous system disks are released. If an automatic snapshot has been set up to release with the cloud disks, the automatic snapshots of the previous system disks are automatically deleted.
- *Detach a cloud disk*.

Procedure

Follow these steps:

1. Log on to the *ECS console*.
2. Select a region.
3. In the left-side navigation pane, click **Block Storage** > **Disks**.
4. Select the disk that you want to configure, and in the **Actions** column, click **More** > **Modify Attributes**.
5. In the **Modify Disk Type** dialog box, select **Delete Automatic Snapshots while Releasing Disk**, and then click **OK**.



Related API

ModifyDiskAttribute

12.5 Delete snapshots or automatic snapshot policies

When you no longer need a snapshot, or you have reached your snapshot quota, you can delete snapshots to free up space.

Note:

- After a snapshot is deleted, it cannot be restored. So proceed with caution.
- If a snapshot has been used to create a custom image, you must delete the associated image before you can delete the snapshot.

Delete snapshots

1. Log on to the *Container Service console*.
2. In the left-side navigation pane, select **Snapshots and Images** > **Snapshots**. Select a region.
3. Select the snapshots you want to delete.
4. Click **Delete** at the lower-left corner, and then click **OK**.

Delete snapshot policies

1. Log on to the *ECS console*.
2. In the left-side navigation pane, select **Snapshots and Images** > **Automatic Snapshot Policies**. Select a region.
3. Find the snapshot policy you want to delete, and in the **Actions** column, click **Delete Automatic Snapshot Policy**.
4. In the dialog box, confirm information and click **OK**.

12.6 View a snapshot chain

On March 28, 2017, Alibaba Cloud began to commercialize the snapshot service. After commercialization, the snapshot service will charge for the snapshot capacity, for detailed pricing information, *block storage for detailed pricing information* . For more information about Snapshot commercialization, see the *snapshot commercialization FAQ*.

The snapshot service fee is related to the snapshot size. This article describes how to view the snapshot size on a single disk or under a region.

View the snapshot size on a single disk

When you create snapshots of an elastic block storage device, such as a cloud disk or a shared block storage device, you can view the snapshot size of the device by using the **Snapshot Chains** feature in the ECS console.

A snapshot chain is composed of all the snapshots of an elastic block storage device. After you create a snapshot, the device has a snapshot chain. The chain has the identical ID with that of the disk. A snapshot chain provides the following information:

- Snapshot nodes: Each snapshot node of the chain represents one snapshot of the device.
- Snapshot size: The storage space occupied by all snapshots of the device.
- Snapshot quota: Each device has up to 64 snapshots, including those created manually or automatically.

Prerequisite

You have *created snapshots*.

Procedure

To view the total size of all the snapshots of an elastic block storage device, follow these steps:

1. Log on to the *ECS console*.
2. Select a region.
3. In the left-side navigation pane, select **Snapshots and Images** > **Snapshots** .
4. Find the disk ID of the snapshot. The disk should have at least one snapshot.
5. In the left-side navigation pane, click **Snapshot Chains**.
6. View the size of all snapshots on the disk according to the disk ID found in step 5. You can view the total number and size of snapshots of the disk in the list.

In the **Actions** column, click **Details** to go to the **Snapshot Chain Details** page. On the page, you can see all snapshots of the disk, which you can use to *roll back a cloud disk* or *create a custom image by using a snapshot*.

View the snapshot size under a region

Follow these steps:

1. Log on to the *ECS console*.
2. Select a region.
3. In the left-side navigation pane, select **Snapshots and Images** > **Snapshots**.

You can view the total number and size of snapshots of the disk in the list.

13 Images

13.1 Open source tools

13.1.1 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.

13.1.2 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.
iso_checksum_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_checksum** or the **iso_checksum_url** parameter. If you have the iso_checksum parameter specified, the **iso_checksum_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:

Parameter	Type	Description
		• 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 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 automatica lly with the specified oss_bucket_name when uploading the image.
image_os_type	String	Image type. Optional values:

Parameter	Type	Description
		• linux • windows
image_platform	String	Distribution of the image. For example, CentOS.
image_architecture	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

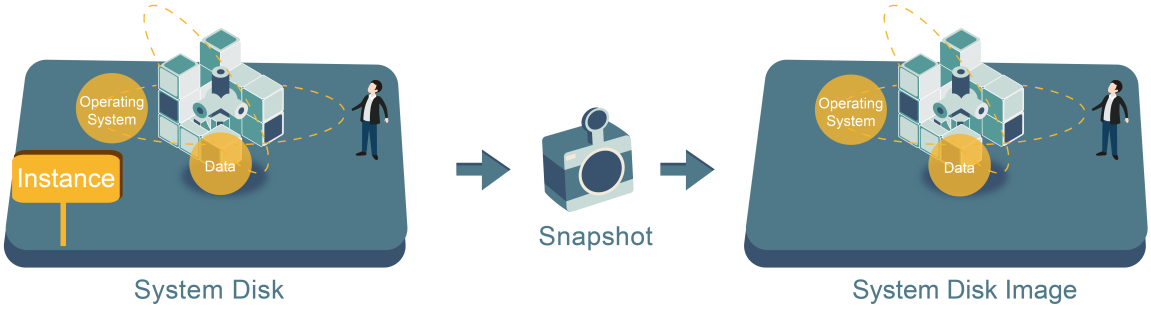
- For more information about how to use Packer, see *Packer* documentation.
- For more information about release information, visit the Packer repository on GitHub *packer*.
- For more information about Alibaba Cloud open source tools, visit Alibaba repository on GitHub *opstools*.
- For more information 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* .

13.2 Create custom image

13.2.1 Create a custom image by using a snapshot

Custom images help you run ECS instances effectively by allowing you to create multiple ECS instances with identical OS and environment data to meet scaling requirements.

Custom images are based on ECS disk snapshots. You can set up identical or different configurations for ECS instances that are created from images.



You can use a snapshot to create a custom image, including the operating system and data environment of the snapshot in the image. You can then use the custom mirror to create multiple instances with the same operating system and data environment, replicating instances easily.

You can also use an instance to create an image. See *create a custom image by using an instance*.

To enhance the security of creating custom images from snapshots, operation, *Alibaba Cloud custom mirror security recommendations*.

Note:

- Custom images cannot be used across regions.
- You can change the operating system of an instance created from a custom image. The custom image can still be used after the operating system is changed. See *change the system disk (custom image)*.
- You can upgrade the instance created from a custom image, including upgrading the CPU, memory, bandwidth, and disks.
- Custom images are independent from billing methods. Both Subscription and Pay-As-You-Go billing methods work. Custom images created from Subscription instances can used for creating Pay-As-You-Go instances. The opposite is also true.
- If the ECS instance used for creating a custom image expires, or the data is erased (that is, the system disk used for the snapshot expires or is released), the custom image and the ECS instances created from the custom image are not affected. However, automatic snapshots are cleared when an ECS instance is released.

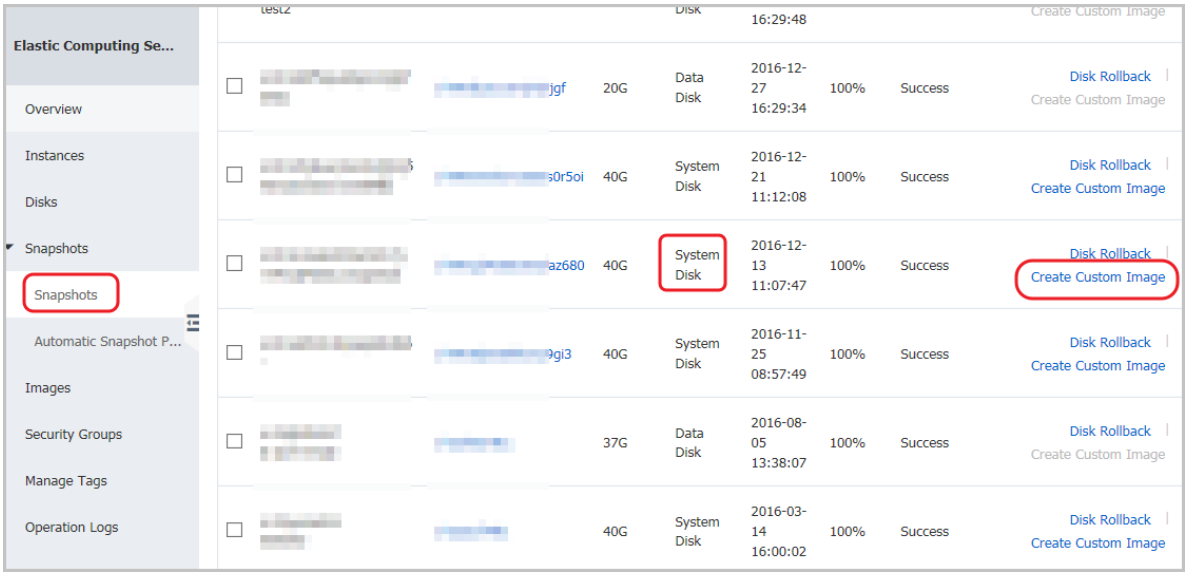
Considerations for Linux instances

- Do not load data disk information in the `/etc/fstab` file. Otherwise, instances created using this image cannot start.

- We recommend that you **`umount`** all data disks before creating a custom image, and then use a snapshot to create a custom image. Otherwise, ECS instances that are created based on this custom image may not start.
- Do not upgrade the kernel or operating system version.
- Do not change the system disk partitions. The system disk only supports single root partitions.
- Check the available space of the system disk to make sure that the system disk is not full.
- Do not modify critical system files such as `/sbin`, `/bin`, `/lib`, and so on.
- Do not modify the default logon user name root.

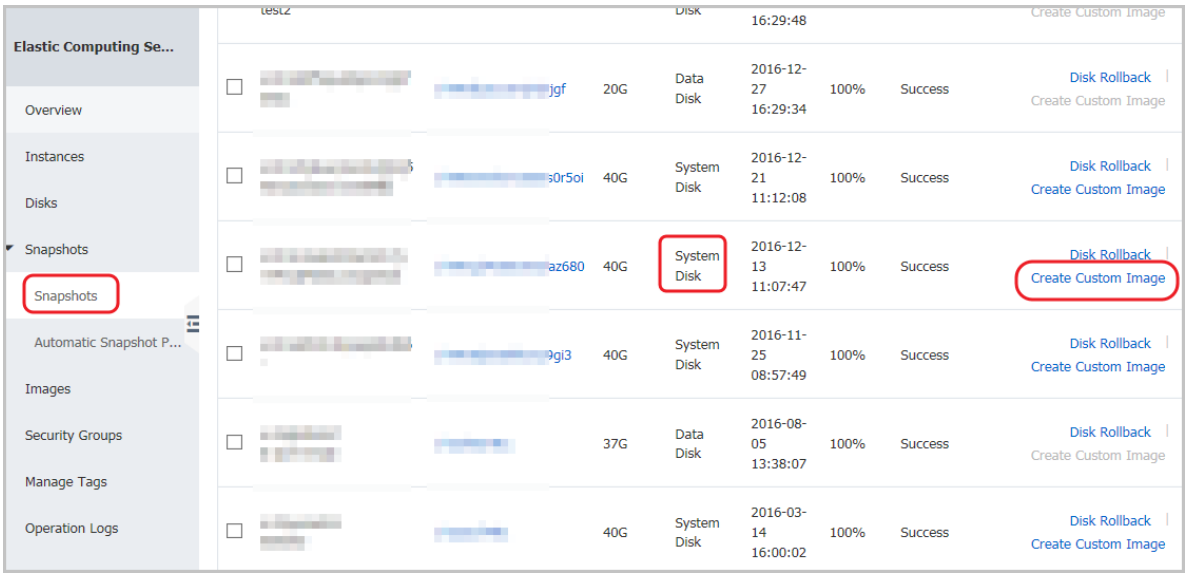
Procedure

1. Log on to the *ECS console*.
2. Select a region.
3. In the left-hand navigation pane, click **Instances**.
4. Find the target instance, and click the instance ID, or click **Manage** in the **Actions** column.
5. In the left-hand navigation pane, click **Instance Snapshots**. Find the target system disk, and click **Create Custom Image** in the **Actions** column.



The snapshot must be created from system disks. Data disks cannot be used to create custom images.

You can also click **Snapshots and Images** > **Snapshots**, and select a snapshot created from a system disk to **Create Custom Image**.

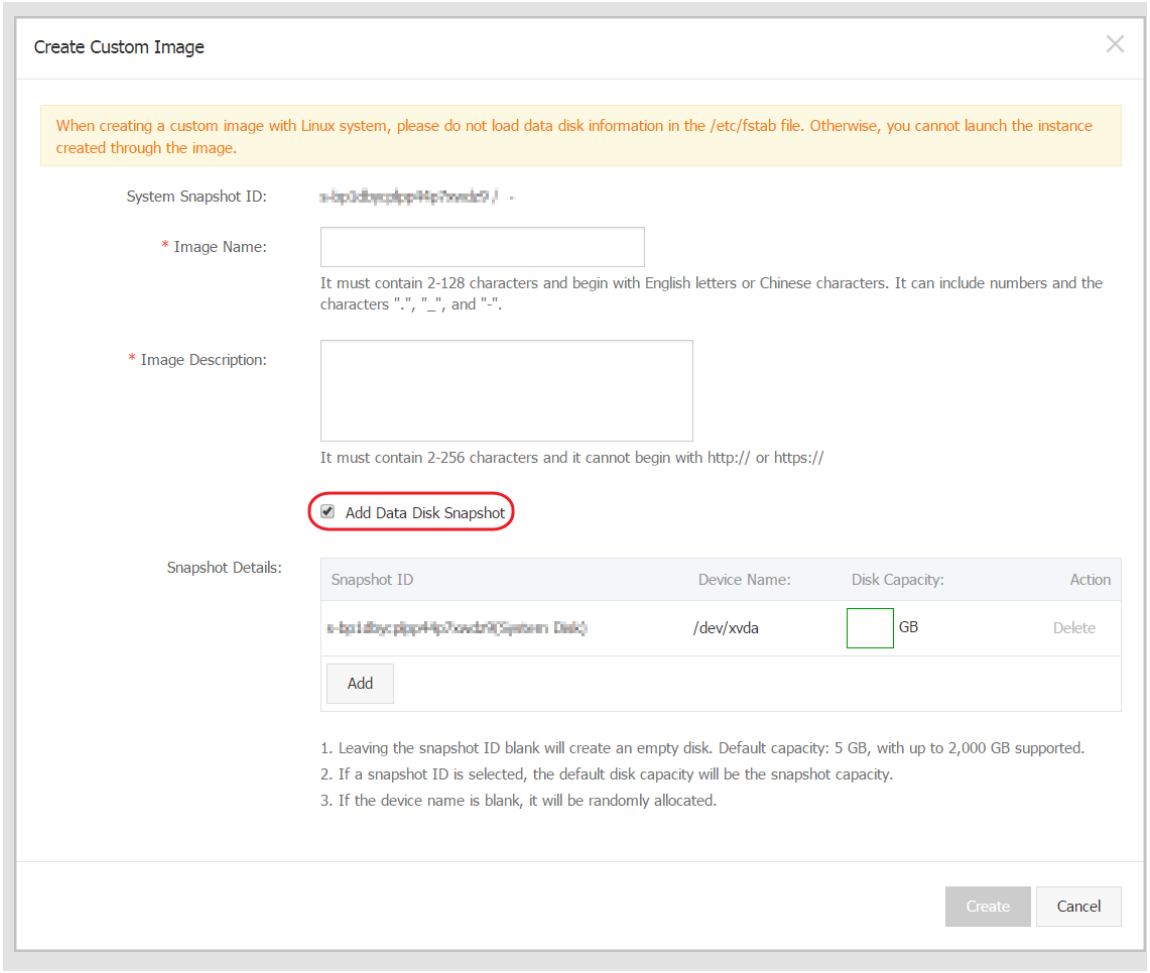


6. In the **Create Custom Image** dialog box, complete the following:

   - Confirm the snapshot ID.
   - Specify the name and description of the custom image.
   - Optional. Check **Add Data Disk Snapshot**, select multiple snapshots of data disks for the image, and click **Add** to add a data disk.

     Note:

     - Remove sensitive data from the data disk before creating a custom image to guarantee data security.
     - If the snapshot disk capacity is left blank, an empty disk is created with the default capacity of 5 GiB.
     - If you select available snapshots, the disk size is the same as the size of these snapshots.



7. Click **Create**. The custom image is successfully created. In the left-side navigation pane, select **Snapshots and Images** > **Images** to view the images you have created.

Linux instance image FAQ

How to `umount` a disk and delete disk table data?

If `/dev/hda5` is attached to `/mnt/hda5`, run any of the following three commands to detach the file system.

```
umount/dev/hda5
umount/mnt/hda5
umount/dev/hda5/mnt/hda5
```

`/Etc/fstab` is an important configuration file in Linux. It contains the details of mounting the file system and storage devices upon startup. If you do not want to mount a specified partition when starting the instance, delete the corresponding lines from `/etc/fstab`. For example, you can delete the following statement to disconnect xvdb1 upon startup: `/dev/xvdb1 /leejd ext4 defaults 0 0`.

How to determine whether a data disk is detached and a custom image can be created?

You must make sure that the statement line for automatically attaching mounting data disk has been deleted from the fstab file.

Use the **`mount`** command to view the information of all mounted devices. Make sure that the execution results do not contain the information of the data disk partition.

Relevant configuration files

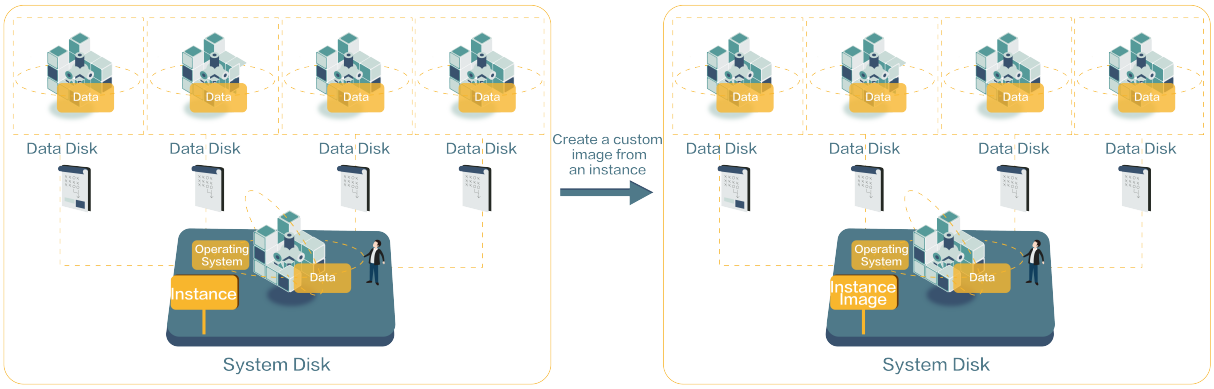
Before creating an image, make sure that the key configuration files listed in the following table have not been modified. Otherwise, the new instance cannot start.

Configuration file	Description	Risks if modified
`/etc/issue*`, `/etc/*-release`, and `/etc/*_version`	For system release version	Modifying /etc/issue* makes the system release version unidentifiable, and cause instance creation failure.
`/boot/grub/menu.lst` and `/boot/grub/grub.conf`	For system startup	Modifying /boot/grub/menu.lst results in kernel loading failure, and the system is unable to start.
/etc/fstab	For partitions upon startup	Modifying it causes partition mounting failure, and the system is unable to start.
/etc/shadow	For system passwords	If this file is set to read-only, the password file cannot be edited, and instance creation fails.
/etc/selinux/config	For system security policies	Modifying /etc/selinux/config and enabling SELinux results in start failure.

13.2.2 Create a custom image by using an instance

You can create a custom image based on an ECS instance, namely, you can fully copy all its disks and pack the data into an image.

During this process, snapshots are automatically created for all disks of the instance, including the system disk and data disks. All the created snapshots compose a new custom image. See the following picture.



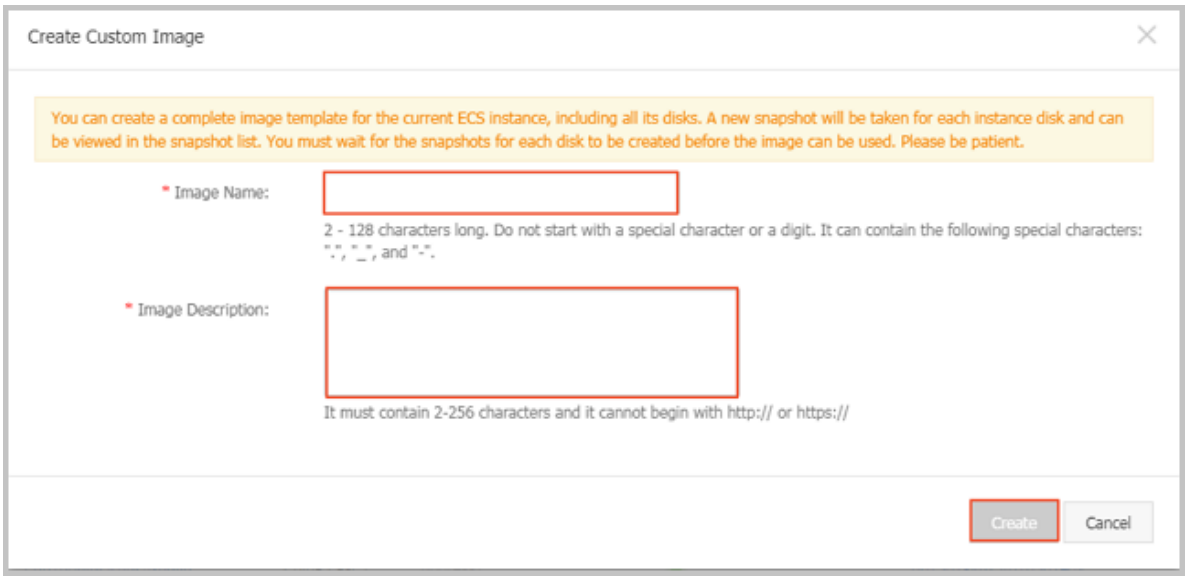
In addition, you can create a custom image based on the snapshot. See *create a custom image by using a snapshot*.

Prerequisites

- To guarantee data security, make sure that you have deleted all the confidential data in the ECS instance before creating a custom image.
- During creation, do not change the status of the instance. Do not stop, start, or restart the instance.
- If your custom image contains data disks, new data disks along with the ECS instance are created together. The data on the data disk duplicates the data disk snapshot in your custom image according to the mount device.
- You can export custom images that contain data disks.
- You cannot use a custom image which contains data disks to replace the system disk.

Procedure

1. Log on to the *ECS console*.
2. Select a region.
3. Click **Instances** from the left-side navigation pane.
4. Find the target instance, and click **More** > **Disk and Image** > **Create Custom Image**.
5. Specify the name and description.
6. Click **Create**.



The image is available after all snapshots of all disks have been created.

Follow-up operation

After creating the custom image, you may want to *create a custom image by using a snapshot*.

13.3 Copy custom images

Copying an image is a process in which a custom image is copied from one region to another region. Copying images across regions allows you to deploy a backup image system, or an identical application environment, in different regions. Copying images is allowed among all the regions supported by Alibaba Cloud. The speed of the process of copying the snapshot between regions depends on the network status and concurrent requests quantity.

Precautions

When an image is copied, the snapshot is generated at the target region, and then a custom image based on the snapshot is generated in the target region. Therefore, you are charged for the data transferring between regions. Currently, such traffic is for free. The specific charge date is subject to the official website announcement.

Procedure

To copy an image, follow these steps:

1. Log on to the *ECS console*.
2. Select a region.
3. In the left-side navigation pane, click **Snapshots and Images** > **Images**.
4. Select the custom image you want to copy, and in the **Actions** column, click **Copy Image**.

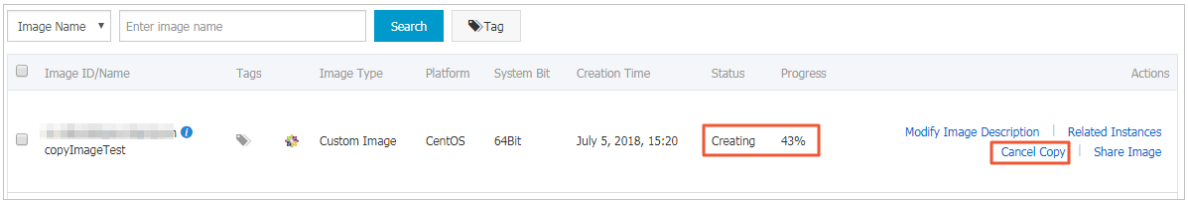
Note:

If your custom image is larger than 200 GB, when you click **Copy Image**, you are directed to open a ticket to complete copying the image.

5. In the **Copy Image** dialog box, the ID of the selected image is displayed, and you have to complete the configurations:
   a. Select the target region. Currently, copying images is only allowed between regions in mainland China.
   b. **Custom Image** and **Custom Image Description**: Specify a name for the image to be displayed in the target region, and give a brief description of the image to ease future management.
   c. Click **OK**.
6. Click the target region and check the progress. When 100% progress is displayed, the image is copied successfully.

Note:

When the progress is not 100% and the status of the image is **Creating**, you can click **Cancel Copy** to cancel the copying process. After the process is canceled, the image information is removed from the target region.



Next step

After the copied image is ready, it is in the **Available** status, you can use the custom image to *create an ECS instance* or *change a system disk*.

In the *ECS console*, check the snapshot for creating the custom image in the **Snapshot**.

FAQ

FAQs about copying images

13.4 Share images

You can share your custom images with other users. Through the ECS console or ECS API, you can query images shared by other accounts with your own account, and select images shared by other accounts to create ECS instances. and replace system disk.

Before sharing an image, make sure that no confidential data is accessible on the disks to be shared.

Note:

The integrity or security of images is not guaranteed. Make sure that you use only images shared by trusted accounts. Before using shared images to create ECS instances, log on to the ECS instances to which the shared images belong and verify that the images are secure and complete .

Precautions

Limits

- One image can be shared with a maximum of 50 accounts.
- Shared images do not count towards your image quota.
- Shared images can only be used to create instances in the same region as the source image.
- Only image owners can share images with other accounts.

Impact of deleting shared images

- You can delete a custom image even you have shared it with other accounts. Before deleting the shared image, however, you must unassociate it from other accounts.
- If you delete an account that has shared a custom image, the users who are using the shared image can no longer find the image through the ECS console or ECS API, or use the image to create ECS instances and replace system disks.
- Deleting shared custom images may cause system disk re-initialization to fail for ECS instances created from these images.

You can share your custom images with other users. Through the ECS console or ECS API, you can query images shared by other accounts with your own account, and select images shared by other accounts to create ECS instances and replace system disk..

Procedure

1. Log on to the *ECS console*. In the left-side navigation pane, click Images. Select a region. Select the Custom Image you want to share. Click **Share Image**.
2. In the displayed dialog box, select the Account Type and enter the account ID you want to share the image with. To obtain the account ID, logg on to Security Settings
3. of the Alibaba Cloud console and click**Account Management** > **Security Settings** > **Account ID**.
4. View accounts using your shared images.

Creating ECs instances using shared images

Note:

The integrity or security of images is not guaranteed. Make sure that you use only images shared by trusted accounts. If you delete an account that has shared a custom image, the users who are using the shared image can no longer find the image through the ECS console or ECS API, or use the image to create ECS instances.

Cancel the sharing of an image

You can cancel the sharing of an image to specific accounts at any time. After you cancel the sharing, the user is unable to query and use the image.

Note:

Any instances using the image, including instances of other accounts using the shared image, will not be able to reinitialize the system disk.

1. Log on to the *ECS console*.
2. Select a region.
3. In the left-side navigation pane, click **Images**. Select the
4. image you want to cancel sharing. The image type must be Custom Image. Click **Share Image**.
5. A list of the accounts using the selected image is displayed. Click **Unshare** next to the account with which you want to stop sharing the image.

View the shared images

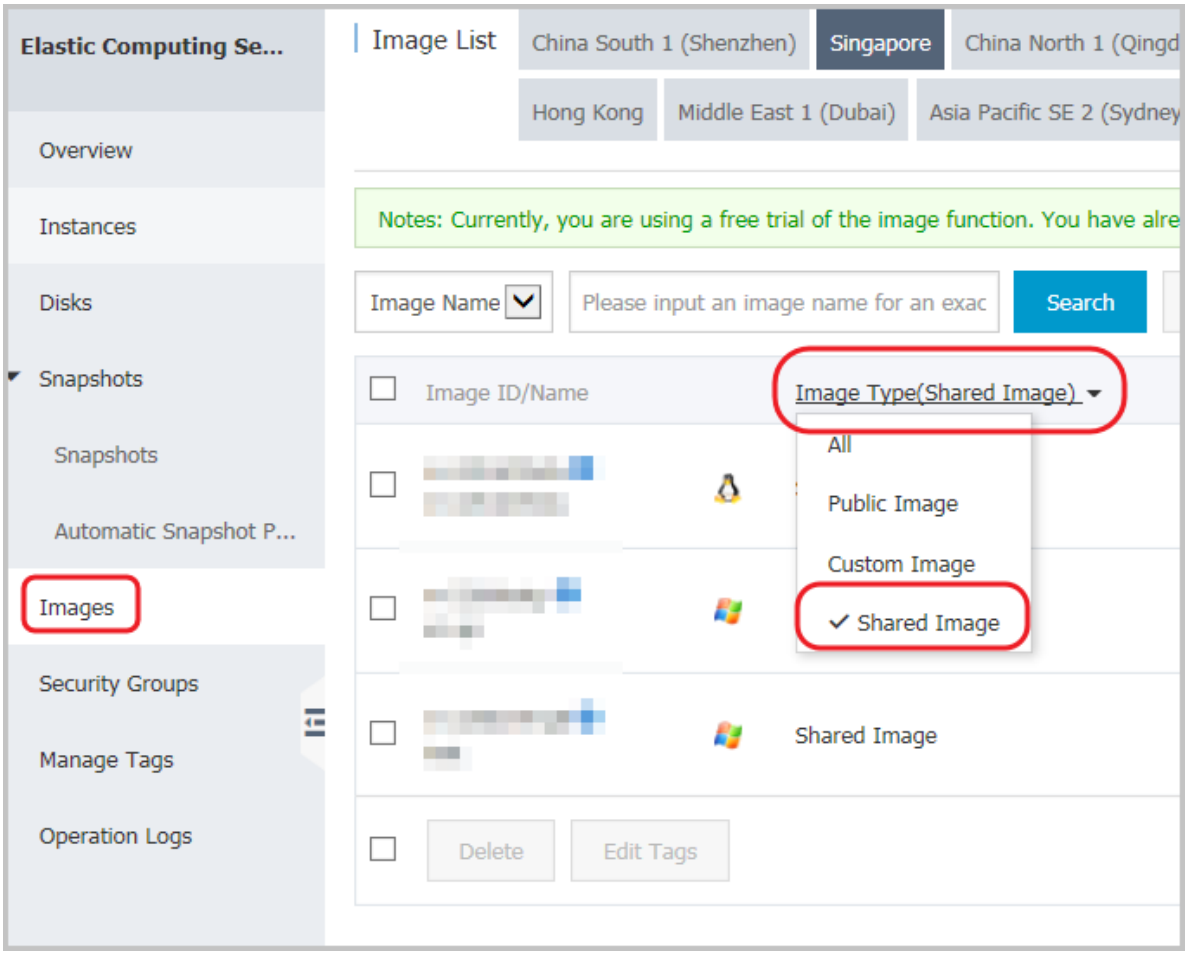
You can view which accounts are using your shared images. To view accounts using your shared images, perform the following:

1. Log on to the *ECS console*.
2. Select a region.
3. In the left-side navigation pane, click **Images**. You can see the list of images.
4. Click Images, you can see the list of images.
5. select image you want to vie. The image type must be **Custom Image**. Click Click **Share Image**.
6. A list of the accounts using the selected image is displayed.

View the shared images you are using

You can view a list of the shared images from other accounts that you are using. To view a list of the shared images you are using, perform the following:

1. Log on to the *ECS console*.
2. Select a region.
3. In the image type dropdown, select the **Image Type**. as **Shared Image**, A list of the shared images you are using will be displayed.



13.5 Import images

13.5.1 Image compliance tool

ECS allows you to create instances from imported custom images. The imported custom images can be created based on your offline server, virtual machine, or a cloud host on any cloud platform. The images you import must meet certain requirements. For more information, see *Notes for importing images*. To reduce the time required for creating images and instances, we recommend that you use the **image compliance tool** of ECS (hereinafter referred to as **compliance tool**) to create images that comply with the relevant standards. The compliance tool can detect non-compliance of various configuration indicators and locations based on a given server environment, generate TXT and JSON detection reports, and offer possible solutions.

Limits

The compliance tool currently supports Linux images only, such as Ubuntu, CentOS, Debian, RedHat, SUSE Linux Enterprise Server (SLES), OpenSUSE, FreeBSD, CoreOS, and other Linux versions.

Sample

The following sample use a CentOS 7.4 64-bit server.

1. Log on to your server, virtual machine, or cloud host on any cloud platform.
2. *Download* the compliance tool.
3. Run `image_check` with root permissions to guarantee that the compliance tool can read configuration files under permission control.

```
chmod +x image_check
sudo image_check –p [destination path]
```

> **Note:**
> You can use `-p [destination path]` to specify the path where detection reports are generated. If you do not set this parameter, reports are generated in the compliance tool path by default.

4. Wait for the compliance tool to detect the system configuration.

```
Begin check your system...
 The report is generating.
 ----------------------------------------
 The information you need to enter when you import your image to the
 Alibaba Cloud website:
 Current system: CentOS # System information 1: Server operating
system
```

```
 Architecture: x86_64 # System information 2: System architecture
 System disk size: 42 GB # System information 3: Server system disk
capacity
 ----------------------------------------
 # Detection item
 Check driver [ OK ]
 Check shadow file authority [ OK ]
 Check security [ OK ]
 Check qemu-ga [ OK ]
 Check network [ OK ]
 Check ssh [ OK ]
 Check firewall [ OK ]
 Check filesystem [ OK ]
 Check device id [ OK ]
 Check root account [ OK ]
 Check password [ OK ]
 Check partition table [ OK ]
 Check lvm [ FAILED ]
 Check lib [ OK ]
 Check disk size [ OK ]
 Check disk use rate [ WARNING ]
 Check inode use rate [ OK ]
 ----------------------------------------
 15 items are OK
 1 items are failed
 1 items are warning
 ----------------------------------------
 The report is generated: /root/image_check_report_2018-05-14_18-18-
10.txt
 Please read the report to check the details
```

5. View the detection report. The report is generated in the format of `image_check_report_date_time.txt` or `image_check_report.json`.

Detection items

The compliance tool detects the following server configuration items to ensure that the ECS instances created from your custom image are fully functional.

Detection item	Non-compliance	Suggestion
driver	The ECS instance cannot start normally.	Install a virtualization driver. For example, *install a virtio driver*
/etc/shadow	You cannot modify the password file, so you cannot create an ECS instance from the custom image.	Do not use the `chattr` command to lock the /etc/shadow file.
SElinux	The ECS instance cannot start normally.	Do not modify /etc/selinux/config to start SELinux.
qemu-ga	Some of the services required by ECS are unavailable, and the instance is not fully functional.	Uninstall qemu-ga.

Detection item	Non-compliance	Suggestion
network	Network functions of the ECS instance are unstable.	Disable or delete the Network Manager and enable the network service. For the latest Linux versions, we recommend that you use the Network Manager and run the nmcli command to configure the network.
ssh	You cannot *connect* to the ECS instance from the console.	Enable the SSH service and do not set PermitRootLogin.
firewall	The system does not automatically configure your ECS instance environment.	Disable the firewall iptables, firewalld, IPFilter (IPF), IPFireWall (IPFW), or PacketFilter (PF).
file system	You cannot *resize the disk*.	The XFS, Ext3, and Ext4 file systems are used, and the Ext2, UFS, and UDF file systems are allowed. The Ext4 file system does not support 64-bit features.
root	You cannot use your username and password to remotely connect to the ECS instance.	Reserve the root account.
passwd	You cannot add users for the ECS instance.	Retain or reinstall the passwd command.
Partition table	The ECS instance cannot start normally.	Use MBR partitioning.
Logical Volume Manager (LVM)	The ECS instance cannot start normally.	Switch to another partitioning service.
/lib	The ECS instance cannot be automatically configured.	The /lib and /lib64 files cannot be stored in absolute paths. Modify the storage paths of /lib and /lib64 to their relative paths.
system disk	N/A	Increase the system disk capacity. The optimal system disk capacity is 40 GiB to 500 GiB. When you import images, configure the system disk capacity based on the virtual file size of images, instead of the usage capacity of images.

Detection item	Non-compliance	Suggestion
disk_usage	You cannot install the necessary drivers or services for the ECS instance.	Make sure that sufficient disk space is available.
inode usage	You cannot install the necessary drivers or services for the ECS instance.	Make sure that sufficient inode resources are available.

The compliance tool provides a detection result `OK`, `FAILED`, or `WARNING` based on detection items.

- `OK`: The detection items all comply with requirements.
- `FAILED`: The detection items do not comply with requirements. The ECS instance created from the custom image cannot start normally. We recommend that you rectify the non-compliant items and recreate the image to improve instance startup efficiency.
- `WARNING`: The detection items do not comply with requirements. The ECS instance created from the custom image can start normally, but ECS cannot use valid methods to configure your instance. You can choose to immediately rectify the non-compliant items or temporarily neglect the items and create an image.

Output items

The compliance tool provides detection reports in both TXT and JSON formats after it detects the system environment. You can use `-p [destination path]` to specify the path where detection reports are generated. If you do not specify this parameter, reports are generated in the compliance tool path by default.

- Reports in TXT format are named `image_check_report_date_time.txt`. The reports include server configuration information and detection results. The following example uses a CentOS 7.4 64-bit server.

```
The information you need to input when you import your image to
Alibaba Cloud Website:
  Current system is: CentOS #Server operating system
  Architecture: x86_64 #System architecture
  System disk size: 42 GB #Server system disk capacity
  ----------------------------------------
   Check driver #Detection item name
   Pass: kvm drive is exist #Detection result
   Alibaba Cloud supports kvm virtualization technology
```

```
    We strongly recommend installing kvm driver.
```

- Reports in JSON format are named `image_check_report.json`. The reports include server configuration information and detection results. The following example uses a CentOS 7.4 64-bit server.

```
"platform": "CentOS", \\Server operating system
  "os_big_version": "7", \\Operating system version number (major)
  "os_small_version": "4", \\Operating system version number (minor)
  "architecture": "x86_64", \\System architecture
  "system_disk_size": "42", \\Server system disk capacity
  "version": "1.0.2", \\Compliance tool version
  "time": "2018-05-14_19-18-10", \\Detection time
  "check_items": [{
      "name": "driver", \\Detection item name
      "result": "OK", \\Detection result
      "error_code": "0", \\Error code
      "description": "Pass: kvm driver exists.", \\Description
      "comment": "Alibaba Cloud supports kvm virtualization
technology. We strongly recommend installing kvm driver."
  }]
}
```

Next steps

1. View *Notes for importing images*.
2. *Install the virtio driver.*
3. (Optional) *Convert the image file format.*
4. *Import custom images.*
5. *Create an instance from a custom image.*

13.5.2 Notes for importing custom images

To guarantee the usability of an imported image to improve the efficiency of importing an image, pay attention to the followings before importing an image. Depending on the operating system of the custom image, the notes vary for *Linux operating systems* and *Windows operating systems*.

Linux operating systems

Restrictions

When importing a Linux image, considering the following:

- Does not support multiple network interfaces.
- IPv6 address is not allowed.
- The password must be 8 to 30 characters in length. It must contain uppercase and lowercase letters, numbers, and special symbols.
- You must install the XEN and KVM virtualization platform drivers.

- The firewall is disabled, and port 22 is enabled by default.
- DHCP must be enabled in the image.
- We recommend that you *Install cloud-init* to guarantee the successful configuration of Hostname, NTP, and Yum sources.

Notes

If you want to import a Linux image, you must pay attention to the notes listed in the table.

Item	Standard operating system	Non-standard platform Mirroring
Definition	The official distribution editions of operating systems supported by Alibaba Cloud, includes: • Alibaba Cloud • CentOS 5,6,7 • CoreOS 681.2.0+ • Debian 6,7 • Freebag • OpenSUSE 13.1 • Redhat • SUSE Linux 10,11,12 • Ubuntu 10,12,13,14	The non-standard operating system refers to either of the followings: • The operating system that are not included in the list of operating systems that are currently supported by ECS. • A standard operating system that fails to comply with the requirements for a standard operating system in terms of critical system configuration files, system basic environment, and applications. If you want to use an image of a non-standard operating system, you are only allowed to choose: • Customized Linux: set-up edition mirror. If you import an image of this type of operating system, Alibaba Cloud conducts necessary network or password configuration according to the pre-defined configuration norms. For more information, see *Configure Customized Linux images*. • Others Linux: ECS identifies all of these images as other system types. If you import an image of such operating system, Alibaba Cloud does not process any of the created instance. After instance creation is complete, you must connect to the instance by using the **Connect** feature in the ECS console and then manually configure the IP address, the router, and the password.

Item	Standard operating system	Non-standard platform Mirroring
System critical Profile	• Do not modify `/etc/issue*`. If it is modified, the distribution of the system cannot be properly recognized and the system creation fails. • Do not modify `/boot/grub/menu.lst`. If it is modified, the system may fail to start up. • Do not modify `/etc/fstab`. If it is modified, an exception may occur preventing partitions from being loaded, leading to system startup failure. • Do not modify `/etc/shadow` to **read-only**. If it is modified, the password file cannot be modified and the system startup fails. • Do not enable SELinux by modifying `/etc/selinux/config`. If it is modified, the system may fail to start up.	Fails to comply with the requirements of a standard operating system.
Requirements for system basic environments	• Do not adjust the partition of the system disk. Currently only a single root partition is supported. • Check the remaining space on the system tray to make sure that the system tray is not full. • Do not modify critical system files, such as `/sbin`, `/bin`, `/lib*` or /lib*. • Before importing an image, confirm the integrity of the file system.	Does not meet standard platform mirroring requirements.

Item	Standard operating system	Non-standard platform Mirroring
	• File system: File systems of xfs, ext3, and ext4 for Linux images are supported. MBR is used.	
Applications	Do not install qemu-ga in an imported image. If it is installed , some of the services that Alibaba Cloud needs may become unavailable.	Does not meet standard platform mirroring requirements
File format	Currently, images in qcow2, RAW and VHD formats are supported. If you want to import images in other formats, *convert the image file format* before importing the image. We recommend that you import images in a qcow2 or VHD format.	
File size	Setting the system disk size when importing an image: We recommend that you configure the system disk size for importing based on the virtual file size (not the usage) of the image. The size of the disk for importing must be between 40 GiB to 500 GiB.	

Windows operating systems

When importing a Windows image, pay attention to the following notes.

Restrictions

- The password must be 8 to 30 characters in length and must contain uppercase and lowercase letters, numbers, and special symbols.
- The firewall is disabled, and port 3389 is enabled by default.

Distribution editions of Windows operating system

You are allowed to import the following distribution editions of Windows operating system:

- Microsoft Windows Server 2016
- Microsoft Windows Server 2012 R2 (standard edition)
- Microsoft Windows Server 2012 (standard edition, data center edition)
- Microsoft Windows Server 2008 R2 (standard edition, data center edition, enterprise edition)
- Microsoft Windows Server 2008 (Standard Edition, Data Center Edition, Enterprise Edition)
- Microsoft Windows Server 2003 with Service Pack 1 (SP1) (standard edition, data center edition, enterprise edition)

Note:

Windows XP, Windows 7 (both Professional Edition and Enterprise Edition), Windows 8, and Windows 10 are not supported.

Requirements on the basic system environment

- Supports multi-partition system disks.
- Check the remaining space on the system tray to make sure that the system tray is not full.
- Do not modify critical system files.
- Verify the integrity of the file system before importing.
- File system: Only NTFS file system and MBR is supported.

Applications

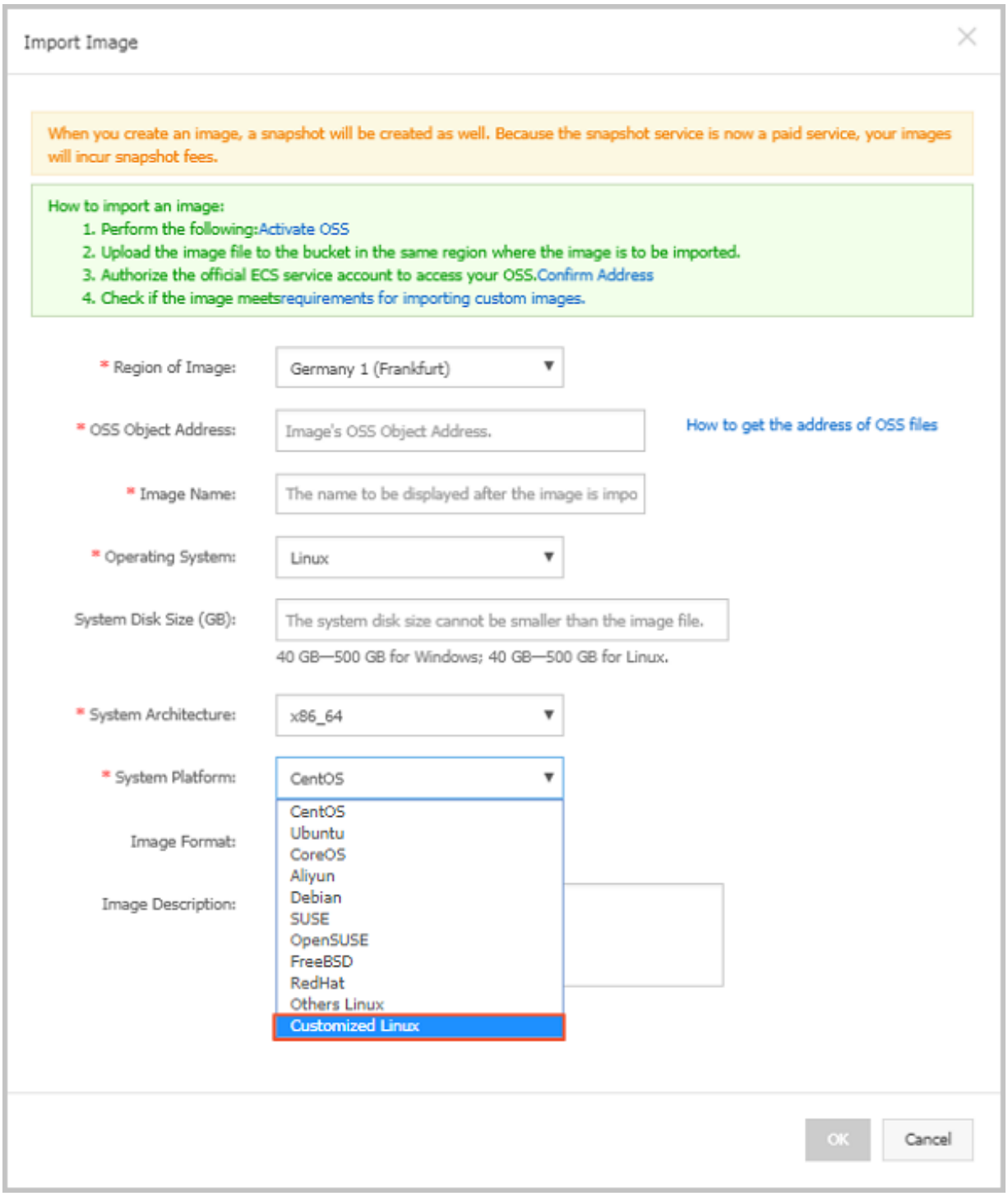
Do not install qemu-ga in an imported image. If it is installed, some of the services that Alibaba Cloud needs may become unavailable.

Size and format

- Currently, images in RAW and VHD formats are supported, and you can *open a ticket* to apply for importing a qcow2 image. If you want to import images in other formats, *convert image file format* and import it again. We recommend that you import images in a VHD format, which has a smaller transmission capacity.
- Setting the system disk size when importing an image: We recommend that you configure the system disk size for importing based on the virtual disk size rather than the usage of the image. The size of the disk for importing must be between 40 GiB to 500 GiB.

13.5.3 Configure Customized Linux images

Customized Linux images are Linux images that currently ECS cannot provide. If you want to run Customized Linux image on your ECS instances, you can import the Customized Linux image in the *ECS console* and configure **Customized Linux** according to this tutorial, see the following image.



A Customized Linux image is treated as an unknown operating system. ECS configures the Customized Linux image by writing the necessary configuration information, such as SSH configuration, to the instance, and then runs a predefined script at the instance startup to process the configuration information.

Limits an procedures

Limits

- The first primary partition of the custom image must be writable.
- The first primary partition type of the custom image must be FAT32, EXT2, EXT3, EXT4, or UFS.
- The virtual file size of the image must be greater than 5 GB.

Procedures

1. Create the directory `aliyun_custom_image` in the root directory of the first primary partition.
2. When the Customized Linux instance starts up, ECS either writes the instance-related configuration information to the os.conf file in the `aliyun_custom_image` directory or, if this file does not exist, a new file with the name as **os.conf** is automatically created.
3. The image must contain a pre-defined startup script to parse various configuration parameters in the **os.conf** file and implement the configuration. For more information, see *Configure the parsing script* t and *Script sample*.

Security requirements

The basic security requirements for Customized Linux are as follows:

- No high-risk vulnerabilities that can be exploited remotely.
- If a default initial password exists for **VNC**, the password must be changed at the first log on. Access of the default VNC account is denied until the password is changed.
- No default initial password for SSH. A random initial password must be generated and distributed by the Alibaba Cloud ECS control system.

Samples of os.conf configuration file

Sample for the classic network-connected instances

```
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 description of the parameters is as follows.

Parameter	Description
hostname	The host name
password	The password, in the format of a Base64-encoded string
eth0_ip_addr	The IP address of eth0 NIC
eth0_mac_addr	The MAC address of eth0 NIC
eth0_netmask	The network mask of eth0 NIC
eth0_gateway	The default gateway of eth0 NIC
eth0_route	The route list (intranet route list) of eth0, separated with semicolons by default
eth1_ip_addr	The IP address of eth1 NIC
eth1_mac_addr	The MAC address of eth1 NIC
eth1_netmask	The network mask of eth1 NIC
eth1_gateway	The default gateway of eth1 NIC
eth1_route	The route (default Internet route) list of eth1, separated with semicolons by default
dns_nameserver	DNS address list, separated with spaces by default

Sample for VPC-Connected instances

```
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 parameter definitions are as follows.

Parameter	Description
hostname	The host name
password	The password, in the format of a Base64-encoded string
eth0_ip_addr	The IP address of eth0 NIC
eth0_mac_addr	The MAC address of eth0 NIC

Parameter	Description
eth0_netmask	The network mask of eth0 NIC
eth0_gateway	The default gateway of eth0 NIC
eth0_route	The route list of eth0, separated with semicolons by default
dns_nameserver	DNS address list, separated with spaces by default

Configure the parsing script

For an optimized Customized Linux configuration, we recommend you predefine the script in the image. When you create an instance, ECS writes information related to the configuration parameters to the os.conf file in the `aliyun_custom_image` directory, in the first primary partition. The script then reads the relevant configuration information from the os.conf file and implements the configuration. During script creation, pay attention to the following:

- Rules for configuring parameters: As stated in *Samples of os.conf configuration file*, the number of configuration parameters and rules for some configuration parameter values for instances are different for VPC and classic networks.
- Boot start: The script must be set to boot automatically at system startup.
- Configuration file path: When you use the image to create an instance, the default device name assigned to the first primary partition varies between I/O-optimized and non-I/O-optimized instances. We recommend you use the `uuid` or `label` in the script to identify the device in the first primary partition. The user password is a Base64-encoded string, and must be entered the same way in the script.
- Identify VPC or classic network: You must determine the instance network type in the script as either VPC or classic network. The easiest method is to identify whether `eth1_route` or other eth1-related configuration items exist.
- Configuration differences between the VPC and classic network:
  - For a VPC instance, the default Internet route is configured in the `eth0_route` parameter in the os.conf file.
  - For a classic network instance, the default route is configured in the `eth1_route` parameter, and the intranet route is configured in the `eth0_route` parameter.

  Therefore, it is necessary in the script to determine the network type of the instance, and then have a specific analysis and processing.

- Configuration optimization: During the life cycle of an instance, the os.conf file must be executed only once, so we recommend that you delete the os.conf configuration file after the script runs successfully. If the script fails to read the os.conf configuration, no configuration is executed.
- Process a custom image: The custom image created on the Customized Linux instance must include the boot script. When you create an instance using the custom image, ECS writes the os.conf configuration when the instance is started for the first time. The script then implements the related configuration items when it detects the configuration.
- Modify related configuration: When the instance configuration information is changed in the ECS console or by an API, ECS writes the relevant information to the os.conf file. Changes are implemented when the script runs again.

Script sample

See the following script example for a CentOS image.

- The script is for reference only. You must modify it based on the actual operating system type to get a valid script.
- Before you use the script, make sure that the script passes the debugging process.
- The script must be configured to run at boot automatically, for example, by putting the script in the `/etc/init.d/` directory.

```
#!/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
```

13.5.4 BugInstall cloud-init

If you need to create ECS instances by using existing images, you can importing them to Alibaba Cloud ECS. To guarantee the successful configuration of the hostname, NTP source, and yum source of the imported image, we recommend that you install cloud-init in your on-premises server , virtual machine, or cloud host before importing an image.

Limits

- Currently, cloud-init supports the Linux distributions of CentOS, Debian, Fedora, FreeBSD, Gentoo, RHEL (Red Hat Enterprise Linux), SLES (SUSE Linux Enterprise Server), and Ubuntu.
- The AliYun datasource support is present in cloud-init since 0.7.9. If your on-premises server , virtual machine, or cloud host already has cloud-init installed, make sure that the version is later than 0.7.9.

  1. Connect to your on-premises server, virtual machine or cloud host.
  2. Run `cloud-init --version` to confirm the version.

     If the version is later than 0.7.9, you can skip this tutorial and start to make your image. Otherwise, follow the tutorial to install cloud-init.

Prerequisites

Make sure that you have installed the following programs. We use yum as an example to describe the installation. If you manage packages by using zypper or apt-get, the installation methods are similar.

- git: Downloads the source code package of cloud-init.

  Command: `yum install git`
- Python2.7: The basis of running and installing cloud-init.

  Command: `yum install python`
- pip: Installs certain Python libraries on which cloud-init depends.

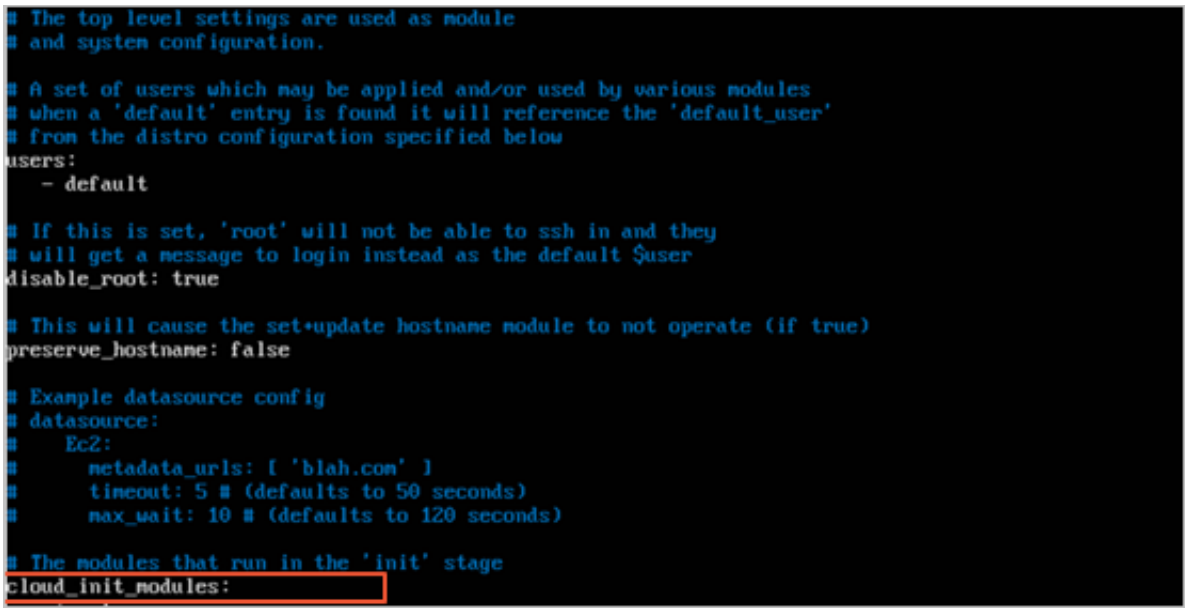
  Command: `yum install python-pip`

We use `yum` as an example to describe the installation. If you manage packages by using `zypper` or `apt-get`, the installation methods are similar to `yum`.

Install cloud-init

Follow these steps to install cloud-init:

1. Connect to your on-premises server, virtual machine or cloud host.
2. Run `git clone https://git.launchpad.net/cloud-init` to download the cloud-init project.
3. Run `cd cloud-init` to change the directory to cloud-init.
4. Run `python setup.py install` to install setup.py, which is the installation file of cloud-init.
5. Run `vi /etc/cloud/cloud.cfg` to modify configuration file cloud.cfg.



Change the preceding content of cloud_init_modules `cloud_init_modules` 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 missing libraries may vary depending on the operating system. You can use pip to install the missing libraries. After you install the missing libraries, install setup.py again.

Library six or library oauthlib is missing

- If the following message appears, it indicates the six library is missing from Python. Run `pip install six` to install the six library.

```
File "/root/cloud-init/cloudinit/log.py", line 19, in <module>
      import six
  ImportError: No module named s  )
```

- If the following message appears, it means the oauthlib library is missing from Python. Run `pip install oauthlib` to install the oauthlib library.

```
File "/root/cloud-init/cloudinit/url_helper.py", line 20, in <module
>
      import oauthlib.oauth1 as oauth1
  ImportError: No module named oauthlib.oaut  )
```

No library is specified when an error occurs during installation

If no dependency library is specified according to the error output, you may `pip install -r requirements.txt` to install all the dependency libraries listed in file requirements.txt of cloud-init.

Next step

Make an image and *Import custom images*.

Reference

cloud-init *Alibaba Cloud (AliYun)*

13.5.5 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 tutorial describes how to use

qemu-img tool to converts custom image file format, such as RAW, Qcow2, VMDK, VDI, VHD (vpc), VHDX, qcow1, or QED to vhd or raw format.

Install qemu-img

You can use different methods to install qemu-img and convert the image file format based on operating system of your computer:

- *Windows operating system*
- *Linux operating system*

Windows operating system

Follow these steps to install qemu-img and convert the image file format:

1. Download and install qemu-img from *https://qemu.weilnetz.de/w64/*. Installation path: *C:\Program Files\qemu*.
2. Perform the following to create an environment variable (For Windows 7):
   a. Choose **Start** > **Computer**, 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 click **Environment Variables**.
   d. In the **Environment Variables** dialog box, in the **System variables**, find the **Path** variable, and click **Edit**. If the **Path** variable does not exist, click **New**.
   e. Add a variable value:
   - In the **Edit System Variable**: In the **Variable** value, add `C:\Program Files\qemu`. Different variable values are separated with semicolon (;).
   - In the **New System Variable**: In the **Variable** name, enter `Path`. In the **Variable** value, enter `C:\Program Files\qemu`.
3. Open **Command** Prompt in Windows and run the `qemu-img --help` command. If it is displayed successfully, the installation was successful.
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. In **Command** prompt, run the `qemu-img convert -f raw -O qcow2 centos.raw centos.qcow2` command to convert the image file format.

   The command parameters are described as follows:

- **-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 operating system

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 raw -O qcow2 centos.raw centos.qcow2` command to convert the image file format.

   The command parameters are described as follows:

   - **-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.

13.5.6 Import custom images

You can import image files to the ECS environment to create custom images. You can then use these images to create ECS instances or change system disks.

> **Note:**
>
> When you import an image, a snapshot is created, because the snapshot service has already started charging for it. 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.

Prerequisite

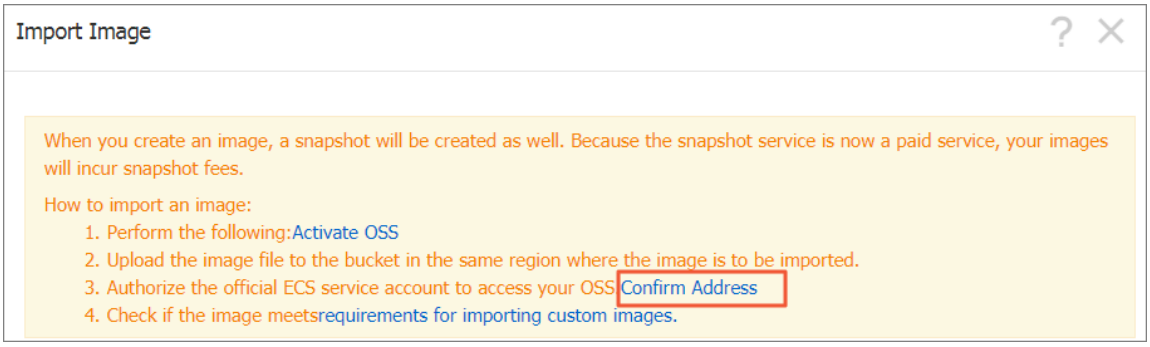
Before importing an image, you should have done the following:

- For the limits and requirements of importing custom images, See *Notes for importing custom images*, *Configure Customized Linux images*, and *Convert image file format*
- *Sign up for OSS*.

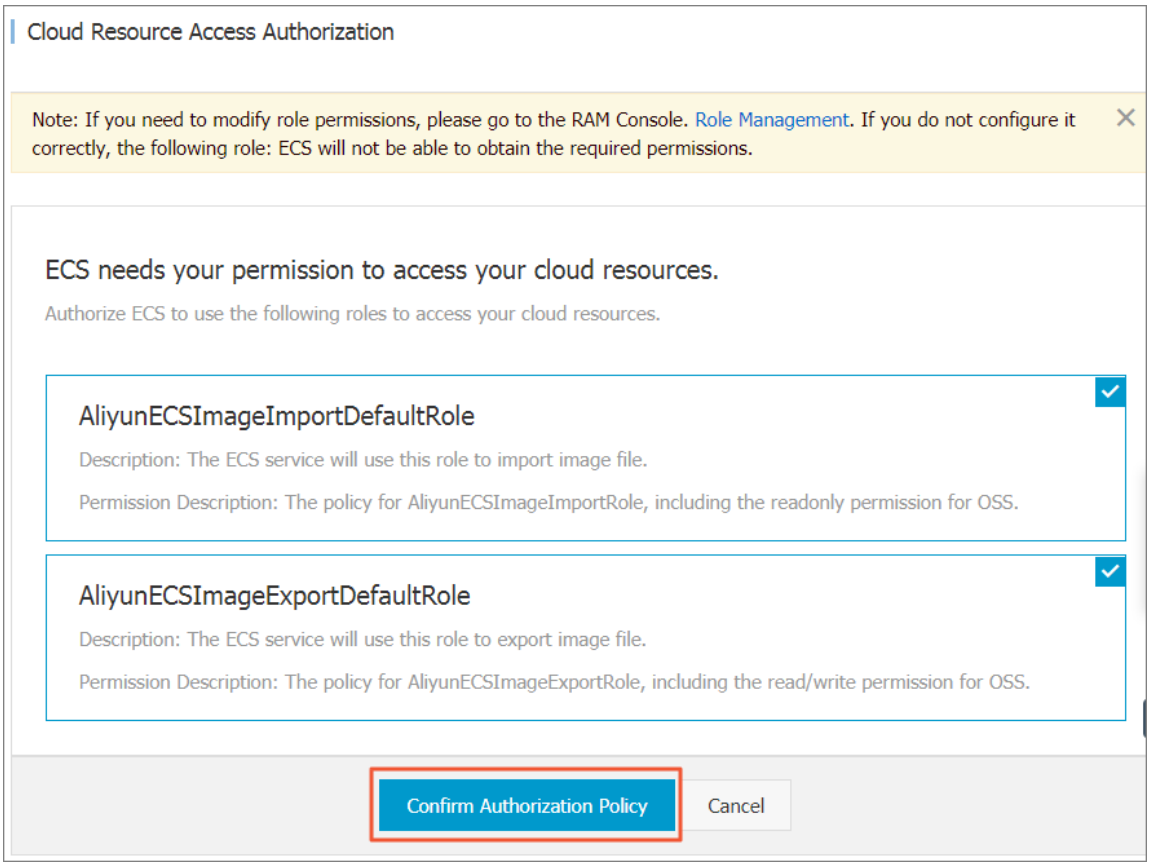
- You can only import an image file to a region from OSS in the same region. The image and the OSS must belong to one account.
- You can use an OSS third-party tool client, OSS API or OSS SDK, to upload the file to a bucket in the same region as the ECS custom image to import. See *Multipart upload* to upload an image file that is larger than 5 GiB.

Procedures

1. Log on to the OSS console, *get object URL*.
2. Follow these steps to authorize the ECS service to access your OSS resources:
   a. Log on to the *ECS console*.
   b. In the left-side navigation pane, choose **Snapshots and Images** > **Images**.
   c. Click **Import Image**.
   d. On the third items of How to import an image. Click **Confirm Address**.



   e. Click Confirm Authorization Policy on the **Cloud Resource Access Authorization** page. Go back to the ECS console.



3. In the left-side navigation pane, choose **Snapshots and Images** > **Images**.
4. Click **Import Image**,
5. Enter the following information in the **Import Image** pop-up window:

   - **Region of image**: Select the region where you want to deploy the application.
   - **OSS Object Address**: Copy the object address taken from the OSS console.
   - **OSS Object Address**: Copy the object address taken from the OSS console. It can be 2 to 128 characters in length. Begins with lower case Latin letters or Chinese characters. Allows numbers, periods (.), underscores (_), colons (:), and hyphens (-).
   - **Operating System** : Supported OS releases are Windows or Linux. If you want to import a non-standard platform image, select Linux.
   - **System Disk size**: The system disk size range is 40 Gib-500 GiB.
   - **System Architecture**: Choose **x86_64** for 64 bit operating systems and choose **i386** for 32 bit operating systems and choose.
   - **System Platform**: The system platform Depends on the **Operating System** you choosed. Available options:

     - Windows: Windows Server 2003, Windows Server 2008, and Windows Server 2012.

  - Linux: Centos, Suse, Ubuntu, Debian, FreeBSD, CoreOS, Aliyun, Customized Linux, and Others Linux. (Linux only) *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 format. VHD format is recommended. Note that you cannot use qemu-image to create VHD images.
- **Image Description**: The description of the image.
- (Optional) If you want to import an image that contains data disks, choose **Add Images of Data Disks**, and follow the page prompts to set information. Supported data disk capacity range is 5 GiB-2000 GiB.
- After the information is confirmed, click **OK** to create a task to import the image.

Note:

- It usually takes 1 to 4 hours to import an image. The duration of the task depends on the size of your image file and the amount of concurrent tasks. You can view the task progress in the image list of the import region.
- We create snapshots for you when importing images, you can check the **Snapshot** list for progress monitoring. Before the import image task is complete, the status of the snapshot is displayed as **Failed**. When the task completes, the status is automatically updated **Available**.

You can find and cancel the image import task in the *task manager*.

Next steps

After you import the custom image, you may want to *create instances from a custom image*.

See also

- *Images*
- *Custom images FAQ*
- *Export custom images*
- *ECS custom image operation practice*
- *Create and import on-premises image by using Packer*
- *Use Packer to create a custom image*
- *Copy custom images*
- *Share images*

13.5.7 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.

```
[root@izbp1[illegible] ~]# grep -i virtio /boot/config-$(uname -r)
CONFIG_VIRTIO_VSOCKETS=m
CONFIG_VIRTIO_VSOCKETS_COMMON=m
CONFIG_VIRTIO_BLK=m
CONFIG_SCSI_VIRTIO=m
CONFIG_VIRTIO_NET=m
CONFIG_VIRTIO_CONSOLE=m
CONFIG_HW_RANDOM_VIRTIO=m
CONFIG_DRM_VIRTIO_GPU=m
CONFIG_VIRTIO=m
# Virtio drivers
CONFIG_VIRTIO_PCI=m
CONFIG_VIRTIO_PCI_LEGACY=y
CONFIG_VIRTIO_BALLOON=m
CONFIG_VIRTIO_INPUT=m
# CONFIG_VIRTIO_MMIO is not set
```

Note:

- If VIRTIO_BLK and VIRTIO_NET do not exist in the output, the virtio driver is not built in the kernel, 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_VIRTIO_NET** are y, the virtio driver is already built in the kernel.You can read Notes for importing custom images *Notes for importing custom images* and import the image *Import custom images*.
- If the values of parameter **CONFIG_VIRTIO_BLK** and and parameter **CONFIG_VIRTIO_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@izbp1[illegible] ~]# lsinitrd /boot/initramfs-$(uname -r).img | grep virtio
Arguments: -f --add-drivers ' 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@izbp1[illegible] ~]#
```

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 custom 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

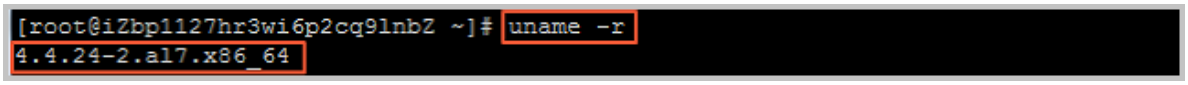
```
echo -e 'xen-blkfront\nvirtio_blk\nvirtio_pci\nvirtio_console' >> \
/etc/initramfs-tools/modules
mkinitramfs -o /boot/initrd.img-$(uname -r)"
```

To compile and install virtio driver

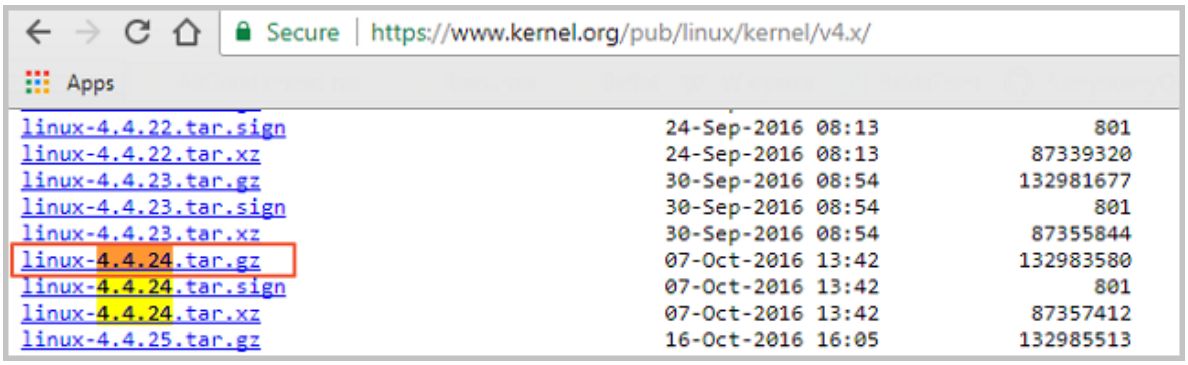
Take Redhat server as an example:

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.



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.

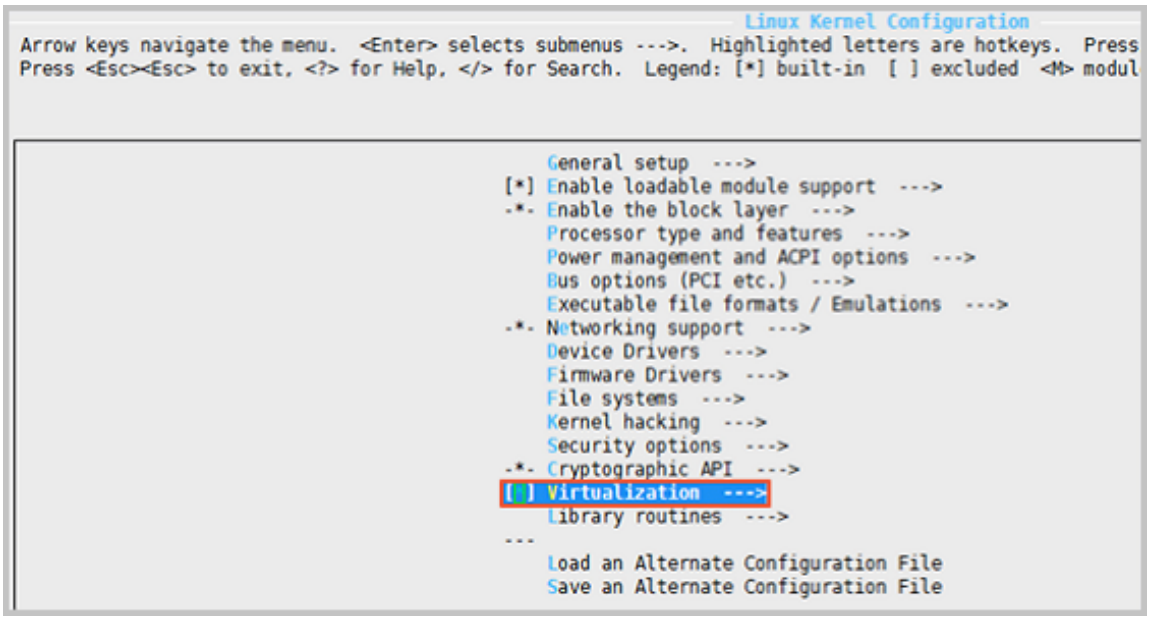
```
make mrproper
 symvers_path=$(find /usr/src/ -name "Module.symvers")
 test -f $symvers_path && cp $symvers_path .
 cp /boot/config-$(uname -r) ./.config
 make menuconfig
```

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

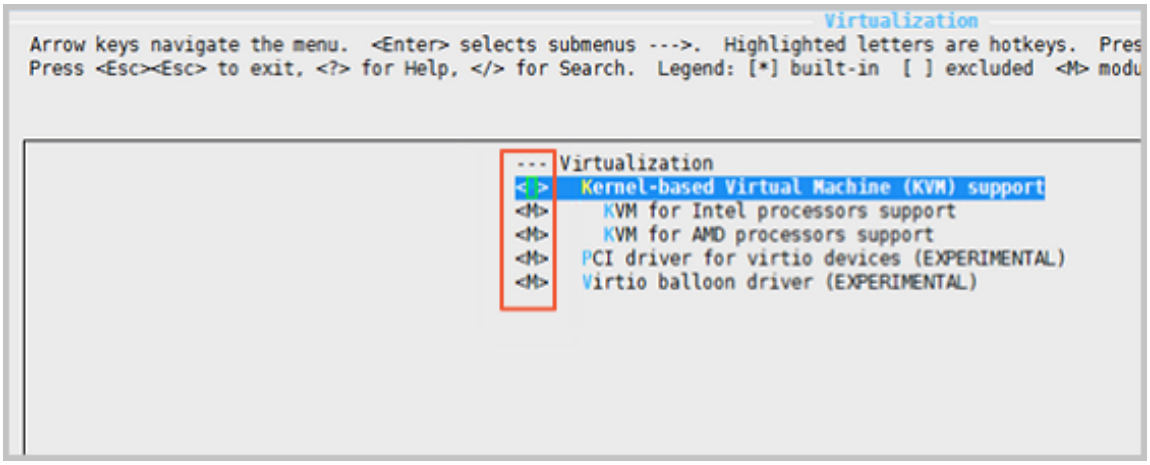
Note:

Select *to build the driver in the kernel, select `m` to compile it as a module.

   a. Press the space bar to select Virtualization.

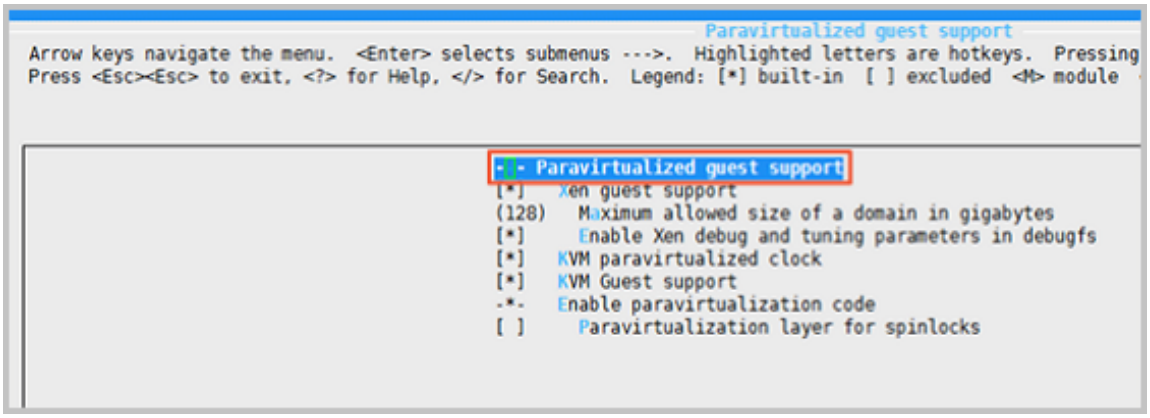


Make sure that you have selected the option of KVM (Kernel-based Virtual Machine).




```
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

After compiling the virtio driver, You can *Migrate your server to Alibaba Cloud by using Cloud Migration Tool*.

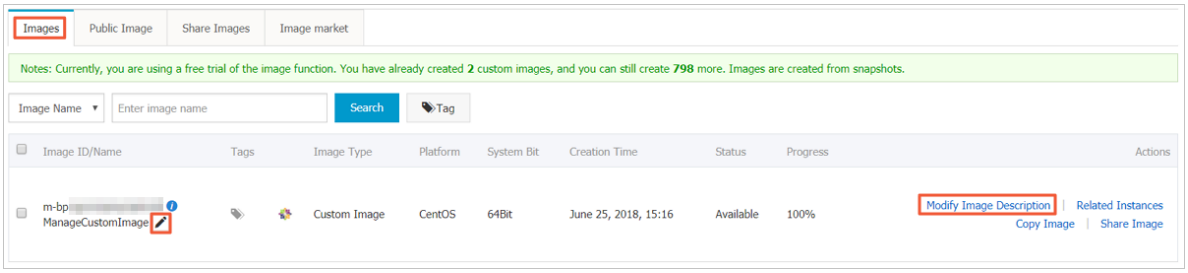
13.6 Manage custom images

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.



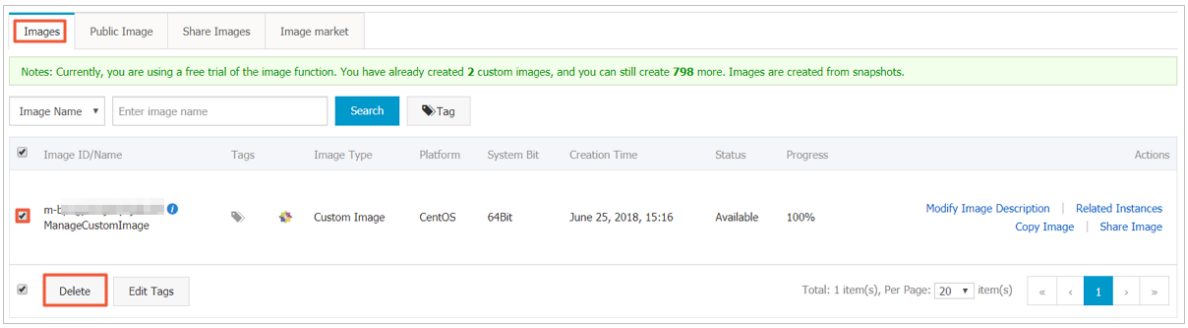
6. Click **Modify Description**, and in the dialog box, edit the new **Custom Image Description**.
7. Click **Save** to complete the description modification.

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:

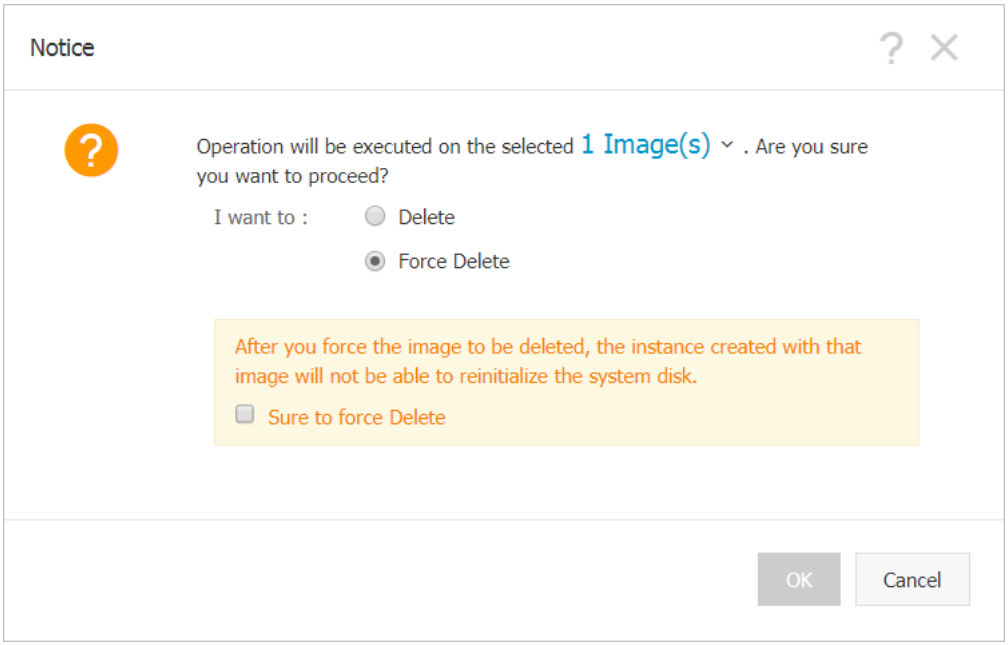
1. Log on to the *ECS console*.
2. In the left-side navigation pane, select **Snapshots and Images** > **Images**.
3. Select a region.
4. Select one or more custom images that you want to delete, and then click **Delete**.



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.

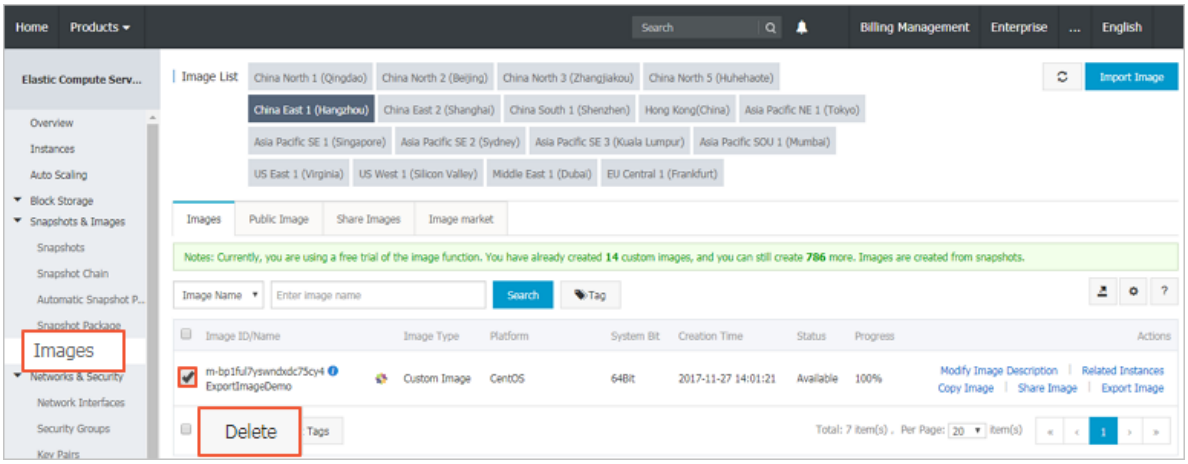
Alternatively, you can delete custom images by calling the ECS API ***DeleteImage***.

13.7 Delete custom images

You can delete a custom image if you no longer need it.

Procedure

1. Log on to the *ECS console*.
2. Select a region.
3. In the left-side navigation pane, select **Images** .
4. Select the image you want to delete. The image type must be **Custom Image**. Click **Delete**.



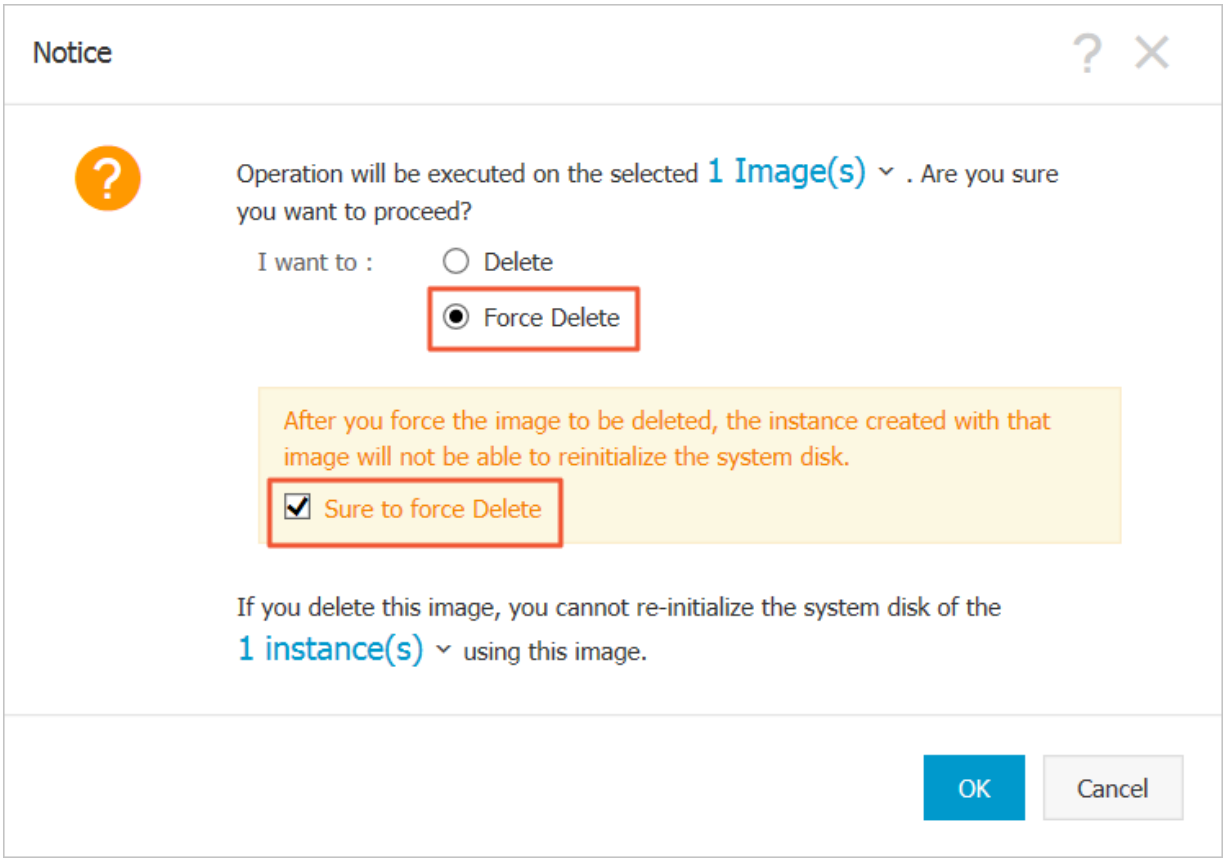
5. In the dialog box, click **OK**.

The custom images has been removed.

FAQ

Can I delete a specified custom image from which an ECS instance is created?

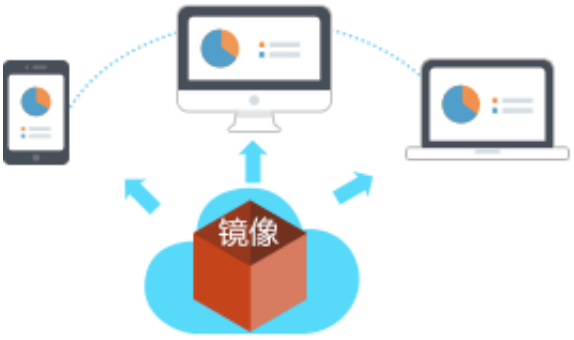
Yes, you can **Force Delete** the image. However, you cannot *#unique_93* the instance after the custom image is deleted. You will also be prompted about the warning.



13.8 Cloud market Images

A mirror is the disk on which the cloud server is loaded. In the past, after purchasing a cloud server, you had to configure your own environment, install software, and it was tricky and time consuming.

Now through the mirror market, you only need to do one simple operation, you can quickly get the cloud server run environments or software applications that are pre-installed in the image, to meet the personalized needs of the site, application development, visual management, and so on, let the cloud server be ready for use, save time and convenience.



Mirror deployment of cloud servers vs manual deployment of cloud servers

Mirrored Installation Method

- Cloud service has not been purchased
- Cloud server already exists

Get a "mirror + cloud server" combo package with one-click deployment

If you are a user of the new cloud server, it is recommended to obtain this combo package.

Mirroring allows you to deploy a cloud server in one click as much as you need to run the environment and personalize the software applications, let your cloud server plug in and run immediately, don't worry about buying any more problems.

Method 1: Buy a cloud server.

On the ECS purchase page, you can complete the selection and configuration of the mirror. The specific process is as follows:

Method 2: the software market chooses the image, and the cloud server is configured by clicking enter.

You can also view and select the desired mirroring services in the software marketplace in the cloud marketplace, then make a matching purchase for the cloud server. The specific process is as follows:

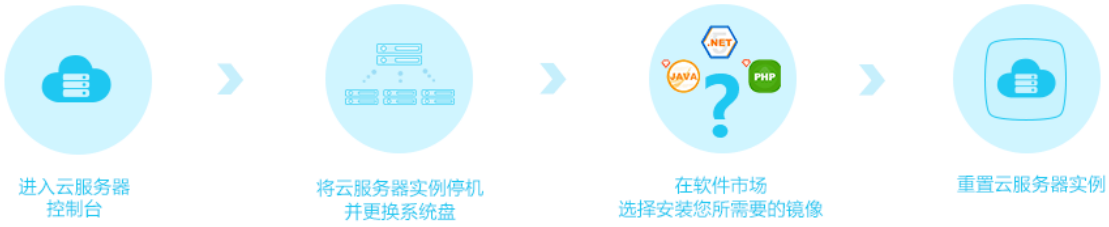
Cloud server already exists, using mirrored deployment

If you have purchased a cloud server, you need to use a mirror to deploy the running environment , or install the software application as follows:

Note:

Replacing a mirror causes the system disk data to be lost, please back up the data before you confirm the replacement.

1. Stop the cloud server instance in the ECS console.
2. Choose replace system tray to use the desired mirror.



13.9 Export custom images

You can export custom images to a local device for test purposes or to offline private stack. This topic describes the limits of the image export function, and provides instructions on how to export images in the ECS console.

Note:

Note that exported images are stored in your OSS bucket, which must be in the same region as the custom images. You are biiled for the data used of for OSS storage and downloading.

Limits

Currently, the image export function has the following constraints and restrictions:

- The image export function is usable before it is *whitelisted*.
- 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 less than 500 GB.
- The default format of exported image files is RAW.

Precautions

- When you export images that contain data disk snapshot, the snapshot and image files are shown in your OSS bucket.
  - A snapshot of the system disk has a system in the file name.

- A snapshot of the data disk has data in the file name. The data disk snapshot will have the identity corresponding to the data disk, which is the mount point of the data disk, such as xvdb or xvdc.

- When using exported images to *Create an instance of the same configuration*, you need to confirm that the file device recorded in /etc/fstab records corresponds to the exported data disk snapshot information.

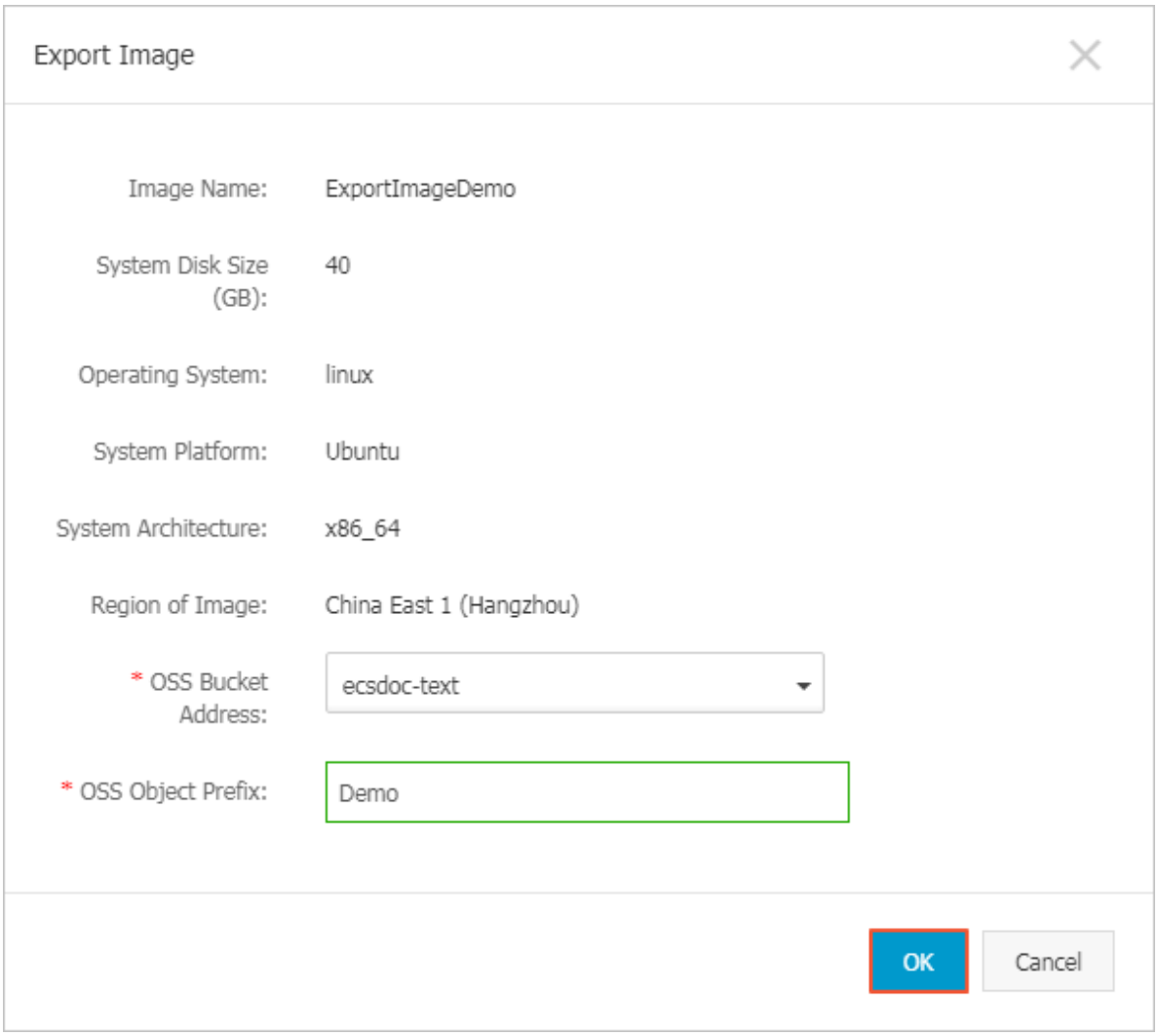
Prerequisites

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

- *Open a ticket* to activate the image export feature. 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. See *Create a bucket*. to create an OSS bucket.

Procedure

1. Log on to the *ECS console*.
2. Select a region.
3. (Optional) Authorize the ECS service to access your OSS bucket:
   a. Choose **Snapshots & Images** > **Images**in the left-side navigation pane.
   b. Find the custom image you want to export. In the **Action** column, click **Export** Image.
   c. In the **Export Image** dialog box, click **Confirm Address** in Step 3 of the prompt message.
   d. In the **Cloud Resource Access Authorization** window, click **Confirm Authorization Policy**. Return to the ECS console homepage.
4. In the left-side navigation pane, choose **Snapshot & Images** > **Images**.
5. Find the custom image you want to export, In the **Action** column, click **Export Image**.
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.



The duration of exporting depends on the size of the image file and the number of other export tasks in the queue. Be patient. You can go to the ***Manage Tasks*** page in the ECS console to query the task progress based on the task ID. When the **Task Status** is Task **Completed**, the image is successfully exported.

To cancel the export task, go to the ***Manage Tasks*** page and find the task.

Next step

- To query the export result, log on to the *OSS console*.
- To download the exported image file, log on to the OSS console *Get object URL* and Get object URL.

14 Security groups

14.1 Typical applications of security group rules

This article introduces the typical applications of security group rules. It applies to instances in classic and VPC network.

The typical applications listed in this article include:

- *Use SSH to connect to Linux instances remotely*
- *Use RDP to connect to Windows instances remotely*
- *Ping ECS instances in public network*
- *Use ECS instances as Web servers*
- *Use FTP to upload or download files*

Use SSH to connect to Linux instances remotely

After you create a Linux ECS instance, you can use SSH to connect to the ECS instance remotely . Add the following security group rules.

Network Types	Network Card Type	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
VPC network	No configuration required	Direction of entry	Allow	SSH (22)	22/22	Address segment access	0.5.0.0/0	1
Classic network	Alibaba Cloud							

Use RDP to connect to Windows instances remotely

After the Windows ECS instance has been created, use RDP to connect to the ECS instance remotely. Add the following security group rules.

Network Type	Network Card Type	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
VPC network	No configurat	Direction of entry	Allow	RDP (3389)	3389/ 3389	Address segment access	0.5.0.0/0	1

	ion required							
Classic network	Public network							

Ping ECS instances in public network

After creating the ECS instance, use Ping program to test the communication status between the ECS instances. Add the following security group rules.

Network Type	Network Card Type	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
VPC network	No configuration required	Direction of entry	Allow	ICMP	-1/-1	Address segment or security group access	Fill it in according to license type, see *add security group rules*.	1
Classic network	Public network							

Use ECS instances as Web servers

If you use your instance as a Web server, install the Web server program on the instance, and add the following security group rules.

Note:

You need to start the Web server program before checking that the 80 ports are working properly. For detailed operation, refer to the Documentation: check if the TCP 80 port is working properly.

Network Type	Network Card Type	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
VPC network	No configuration required	Direction of entry	Allow	HTTP (80)	80/80	Address segment access	0.5.0.0/0	1
Classic network	Public network							

If you are unable to access your instance via http: // public network IP address, please refer to check if the TCP 80 port is working properly.

Use FTP to upload or download files

To use FTP to upload/download files to/from the ECS instance, add the following security group rules.

Note:

You need to install the FTP server program on the instance before checking that port 20/21 is working properly. To install the FTP server program, you can refer to the documentation: the configuration and use of the FTP service under the cloud server ECS.

Network Type	Network Card Type	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
VPC network	No configuration required	Direction of entry	Allow	Custom TCP	20/21	Address segment access	0.5.0.0/0	1
Classic network	Public network							

14.2 Scenarios

This article introduces several common scenarios of VPC-connected and Classic network-connected security groups.

Note:

For more information about how to create a security group and its rules, see *create a security group* and *add a security group rule*.

- *Scenario 1: Enable intranet communication*

  Example: If you want to copy files between two Classic network-connected ECS instances owned by different accounts or in different security groups, you can enable intranet communication between both instances by configuring security group rules and then copy files.

- *Scenario 2: Allow remote connection from specified IP addresses only*

Example: When your ECS instance is compromised by hackers as a zombie, you can modify the port for remote connection, and configure security group rules to allow access from specified IP addresses only.

- *Scenario 3: Allow an instance to access specified IP addresses only*

  Example: When your ECS instance is compromised by hackers as a zombie and scan or send packets maliciously, you can configure security group rules to allow the instance to access to specified IP addresses.

- *Scenario 4: Allow remote connection to an ECS instance*

  Example: You can connect to an ECS instance by configuring a security group rule.

- *Scenario 5: Allow access to an ECS instance over HTTP or HTTPS service*

  Example: If you build a website on your instance, you can configure security group rules to enable your users to access the website.

Scenario 1: Enable intranet communication

Security group rules can be used in the following cases to enable intranet communication between ECS instances that belong to different accounts or security groups in the same region:

- Case 1: Instances belong to one region and one account.
- Case 2: Instances belong to one region but different accounts.

> **Note:**
> For VPC-connected ECS instances,
>
> - If they are in one VPC, you can configure their security group rules to enable intranet communication.
> - If they are in different VPCs, or owned by different accounts in the same region, Express Connect is the only option to establish intranet communication. For more information, see *establish an intranet connection between VPCs under different accounts*.

Case 1: Instances belong to one region and one account

For two instances in one region but owned by one account, if they are in one security group, intranet communication is enabled by default. If they are in different security groups, you must configure security group rules to enable intranet communication according to the network types.

- VPC

If they are in one VPC, add a rule in their security groups respectively to authorize the security groups to access each other. The rule must be as follows.

NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Priority	Authorization Type	Authorization Object
N/A	Inbound	Allow	Select the required protocol	Set the required port range	1	Security group access (authorize this account)	Select the Security Group ID on which you want to allow access to the instance

- Classic network

  Add a rule in their security groups respectively to authorize the security groups to access each other. The rule must be as follows.

NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Priority	Authorization Type	Authorization Object
Intranet	Inbound	Allow	Select the required protocol	Set the required port range	1	Security group access (authorize this account)	Select the Security Group ID on which you want to allow access to the instance

Case 2: Instances belong to one region but different accounts

For Classic network-connected ECS instances only.

Authorize the security groups to access each other. For example:

- User A owns a Classic network-connected ECS instance in the China East 1 region, named Instance A, with the private IP address A.A.A.A. The security group is Group A.

- User B owns a Classic network-connected ECS instance in the China East 1 region, named Instance B, with the private IP address B.B.B.B. The security group is Group B.
- Add a rule in Group A to authorize access of Instance A to Instance B, as shown in the following table.

NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
Intranet	Inbound	Allow	Select the required protocol	Set the required port range	Security group access (authorize other accounts)	Type the account ID of User B and the security group ID of Group B	1

- Add a rule in Group B to authorize access of Instance B to Instance A, as shown in the following table.

NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
Intranet	Inbound	Allow	Select the required protocol	Set the required port range	Security group access (authorize other accounts)	Type the account ID of User A and the security group ID of Group A	1

Note:

To guarantee the security of your instances, when you are configuring an intranet inbound rule for a Classic network-connected security group, **Security Group Access** is the top priority for **Authorization Type**. If you select **Address Field Access**, you must enter an IP address with CIDR prefix, `/32`, in the format of a.b.c.d/32. Only IPv4 is supported.

Scenario 2: Allow remote connection from specified IP addresses only

If you want to allow remote connection to your instance from the specified public IP addresses, add the following rule. In this example, we allow remote connection to an instance on TCP Port 22 from a specified IP address.

Network Type	NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
VPC	N/A	Inbound	Allow	SSH(22)	22/22	Address field access	The IP address to allow access, such as 1.2.3.4.	1
Classic network	Internet							

Scenario 3: Allow an instance to access specified IP addresses only

If you want your instance to access specified IP addresses, add the following rules in its security group.

1. Add the following rule to drop any access to all public IP addresses. The priority must be lower than the rule in step 2.

Network Type	NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
VPC	N/A	Inbound	Drop	All	-1/-1	Address field access	0.0.0.0/0	2
Classic network	Internet							

2. Add the following rule to allow access to the specified IP address, with a higher priority than that in step 1.

Network Type	NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
VPC	N/A	Outbound	Allow	Select the required protocol	Set the required port range	Address field access	Type the specified IP address,	1
Classic network	Internet							

Network Type	NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
							such as 1.2.3.4	

After you add the rules, connect to the instance and try to `ping` or `telnet` the instance from the specified IP address and other IP addresses. If the instance can be accessed by the specified IP address, it means the rules work.

Scenario 4: Allow remote connection to an ECS instance

You may want to connect to your instance in the following cases:

- Case 1: Allow remote connection to your instance from the Internet.
- Case 2: Allow remote connection to your instance from intranet.

Case 1: Allow remote connection to your instance from the Internet

To allow remote connection to your instance from the Internet, add the following rule according to the network type and the operating system of your instance.

- VPC

NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
N/A	Inbound	Allow	Windows: RDP(3389)	3389/3389	Address field access	To allow Internet access from any public IP address, type 0.0.0.0/0. To allow Internet access from a specified Internet IP address,	1
			Linux: SSH (22)	22/22			
			Custom TCP	Customized			

NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
						see *Scenario 2*.	

- Classic network

NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
Internet	Inbound	Allow	Windows: RDP(3389)	3389/3389	Address field access	To allow Internet access from any public IP address, type 0.0.0.0/0. To allow Internet access from a specified Internet IP address, see *Scenario 2*.	1
			Linux: SSH(22)	22/22			
			Custom TCP	Customized			

To customize the port for remote connection, see *Modify the default remote access port*.

Case 2: Allow remote connection to your instance from intranet

If you have enabled intranet communication between instances that belong to one region but different accounts, and you want to allow the instances in different security groups to connect to each other, add the following rules as needed.

- To allow a private IP address to connect to an instance.
  - VPC

Make sure that intranet communication has been built between both accounts by using *Express Connect*, and then add any one of the following rule.

NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
N/A	Inbound	Allow	Windows: RDP(3389)	3389/3389	Address field access	Specify the private IP address of the peer instance	1
			Linux: SSH (22)	22/22			
			Custom TCP	Customized			

- Classic network

Add any one of the following rules.

NIC	Rule direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
Intranet	Inbound	Allow	Windows: RDP (3389)	3389/3389	Address field access	Specify the private IP address of the peer instance. To secure the instance, only an IP address with CIDR prefix, /32, in the format of a.b.c.d/32,	1
			Linux: SSH (22)	22/22			
			Custom TCP	Customized			

NIC	Rule direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
						is allowed.	

- To allow all the instances in a security group of one account to connect to your instance:

  - VPC

    Make sure that intranet communication is built between both accounts by using *Express Connect*, and then add any one of the following rules.

NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
N/A	Inbound	Allow	Windows: RDP (3389)	3389/3389	Security group access (authorize other accounts)	Type the account ID of the peer and the security group ID	1
			Linux: SSH(22)	22/22			
			Custom TCP	Customized			

  - Classic network

    Add any one of the following rules.

NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
Intranet	Inbound	Allow	Windows: RDP (3389)	3389/3389	Security group access (authorize other account)	Type the account ID of the peer and the security group ID	1
			Linux: SSH(22)	22/22			
			Custom TCP	Customized			

Scenario 5: Allow access to an ECS instance over HTTP or HTTPS service

If you have built a website on your instance and expect your users to visit the site over HTTP or HTTPS service, add any one of the following rules.

- VPC

  To allow all public IP addresses to access your site, add any one of the following rules.

NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
N/A	Inbound	Allow	HTTP(80)	80/80	Address field access	0.0.0.0/0	1
			HTTPS(443)	443/443			
			Custom TCP	Customized, such as 8080/8080			

- Classic network

  To allow all public IP addresses to access your site, add any one of the following rules.

NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Authorization Type	Authorization Object	Priority
Internet	Inbound	Allow	HTTP(80)	80/80	Address field access	0.0.0.0/0	1
			HTTPS (443)	443/443			
			Custom TCP	Customized, such as 8080/8080			

Note:

- If your users cannot access your instance by using `http://Public IP address`, *verify if TCP port 80 works properly*
- TCP Port 80 is the default port for HTTP service. If you want to use other ports, modify the port in the configuration file of the Web server.

14.3 Default security group rules

This article introduces the default rules in the security groups created manually or by the system.

Note:

Security groups have status. If an outbound packet is allowed, inbound packets corresponding to this connection are also allowed. For more information about security groups, see *security groups*.

Security groups created by the system

When you create an ECS instance in a region where you have not created a security group, use the default security group provided by the system.

Such a security group only has the default rules for access over the ICMP protocol, TCP Port 22 (for SSH), TCP Port 3389 (for RDP), TCP Port 80 (for HTTP), and TCP Port 443 (for HTTPS). The default rules vary with the network type of the security group.

- VPC: The rules apply to both Internet and intranet access. The Internet access of the VPC type instance is realized through the private NIC mapping. So, you cannot see the Internet NIC inside the instance, and you can only set intranet rules in the security group. The security group rules take effect for both intranet and the Internet. The default rules of the default VPC-connected security group are shown in the following table.

NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Priority	Authorization Type	Authorization Object
N/A	Inbound	Allow	Custom TCP (SSH)	22/22	110	Address field access	0.0.0.0/0
			Custom TCP (RDP)	3389/ 3389			
			All ICMP	-1/-1			
			Custom TCP (HTTP), optional	80/80			
			Custom TCP (	443			

			HTTPS), optional				

- Classic network: The default rules of a classic network-connected security group are shown in the following table.

NIC	Rule Direction	Authorization Policy	Protocol Type	Port Range	Priority	Authorization Type	Authorization Object
Internet	Inbound	Allow	Custom TCP (SSH)	22/22	110	Address field access	0.0.0.0/0
			Custom TCP (RDP)	3389/3389			
			All ICMP	-1/-1			
			Custom TCP (HTTP), optional	80/80			
			Custom TCP (HTTPS), optional	443			

Note:
Rules with priority 110 means that they have the lowest priority in the security group. When you manually create a security group, only values from 1 to 100 are valid for priority setting. For more information about the rule priority, see *add security group rules*.

To meet your business needs, you can *add security group rules* in the default security group.

Manually created security group

After *creating a security group*, before you add rules, the following default rules apply to the communication of all the instances in the group over the Internet or intranet:

- Outbound: Allow
- Inbound: Refuse

If your instance has joined such a security group, you can use the *Management Terminal* only to connect to an instance, rather than using any remote connection methods like *connecting to*

a Linux instance by using a password or *connecting to a Windows instance by using remote connection software*.

To meet your business needs, you can *add security group rules* in the manually created security groups.

14.4 Create a security group

Each ECS instance must join at least one security group. For more information, see *security group*.

If you do not have a security group when your create an ECS instance, you can use the default one provided by the system. For more information, see *default security group rules*.

Alternatively, you can create a security group to meet your business needs and add your instances to it. This article describes how to create a security group.

Prerequisite

To create a VPC-connected security group, you must *manage a VPC*.

Note:

You can create a VPC-connected security group across VSwitches, but not across VPCs.

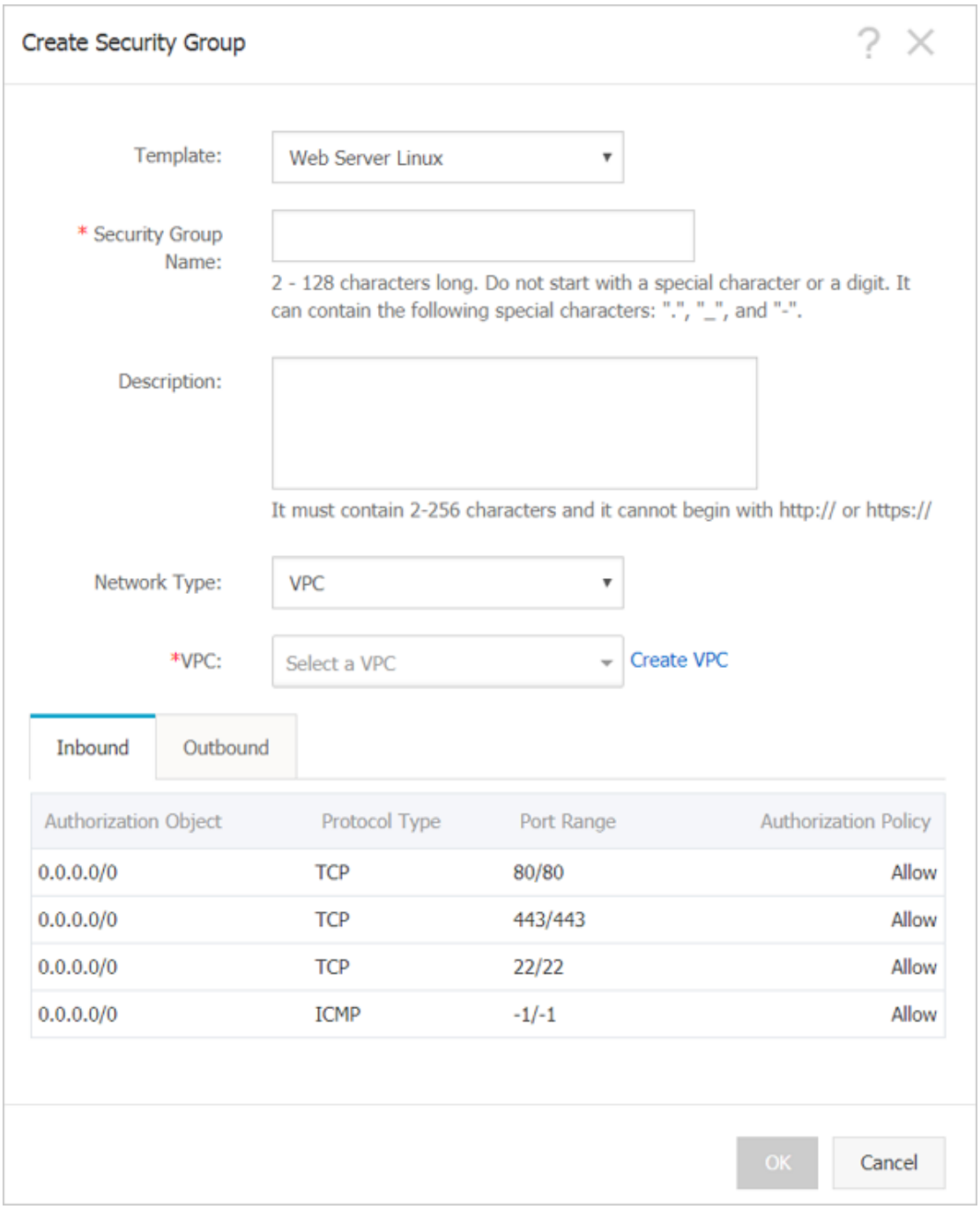
Procedure

1. Log on to the *ECS console*.
2. In the left-side navigation pane, select **Networks and Security >** > **Security Groups**.
3. Select a region.
4. Click **Create Security Group**.
5. In the **Create Security Group** dialog box, complete the following configurations:
   - **Template**: If the instances in the security group are for Web server deployment, select proper template to simplify security group rule configuration.

Setting	Template	Note
To deploy a Web server on the Linux instances in the security group	**Web Server Linux**	By default, inbound traffic to TCP 80, TCP 443, TCP 22, and ICMP is allowed.
To deploy a Web server on the Windows instances in the security group	**Web Server Windows**	By default, inbound traffic to port TCP 80, TCP 443, TCP 3389, and ICMP is allowed.

Not for Web server	**Custom**	After the security group is created, *add security group rules* to meet your business needs.

- **Security Group Name**: Specify a valid security group name.
- **Description**: Give a brief description to the security group for future management.
- **Network Type**:
  - To create a VPC-connected security group, select **VPC** and then a specific VPC.
  - To create a classic network-connected security group, select **Classic**.



6. Click **OK** to save the rule.

For a new security group without any rules, the following default rules apply to the communication of all the instances in the group over the Internet or intranet:

- Outbound: Allow
- Inbound: Forbid

Follow-up operations

After you create a security, *add security group rules*.

To meet your business needs, *add/remove an instance to/from a security group*.

Related API

CreateSecurityGroup

14.5 Add security group rules

You can add security group rules to enable or disable access to and from the Internet or intranet for ECS instances in the security group:

- VPC: You only need to set inbound and outbound rules. Also, you do not need to create different rules for the Internet and intranet. The Internet access for VPC instance is realized through private NIC mapping. So, you cannot see the Internet NIC inside the instance, and you can only set intranet rules in the security group. The rules apply to Internet and intranet access at the same time.
- Classic network: It is required to set outbound and inbound rules for the Internet and intranet respectively.

For a new security group without any rules, outbound traffic is allowed and inbound traffic is refused by default, over either the Internet or intranet. Therefore, we recommend that you only need rules to refuse outbound traffic or allow inbound traffic.

Changes to the security group rules automatically apply to ECS instances in the security group.

Prerequisites

You have created a security group. For more information, see *create a security group*.

You know which Internet or intranet requests need to be allowed or refused for your instance.

Procedure

1. Log on to the *ECS console*.
2. In the left-side navigation pane, select **Networks and Security** > **Security Groups**.
3. Select a region.
4. Find the security group to add authorization rules, and in the **Actions** column, click **Add Rules**.
5. On the **Security Group Rules** page, click **Add Security Group Rule**.

> **Note:**
> If you do not need to enable or disable all ports for all protocols, ICMP, or GRE, you can select **Quick Rule Creation**.

Protocol	SSH	telnet	HTTP	HTTPS	MS SQL
Port	22	2, 3	80.	443	1433
Protocol	Oracle	MySQL	RDP	PostgreSQL	Redis
Port	1521	3306	3389	5432	6379

Note:
See step 6 for descriptions on each parameter configuration.

6. In the dialog box, set the following parameters:
   - **NIC**:
     - For a VPC-connected security group, you can skip selecting the NIC.

       Note:
       - If your instances can access the Internet, the rules work for both the Internet and intranet.
       - If your instances cannot access the Internet, the rules work for intranet only.
     - For a classic network-connected security group, you must select **Internet** or **Intranet**.
   - **Rule Direction**:
     - **Outbound**: ECS instances access other ECS instances over intranet networks, or through Internet resources.
     - **Inbound**: Other ECS instances in the intranet and Internet resources access the ECS instance.
   - **Action**: Select **Allow** or **Forbid**.

     Note:
     Forbid policies discard the data packet without returning a response. If two security group rules overlap except the authorization policy, the **Forbid** rule takes priority over the **Allow** rule.
   - **Protocol Type** and **Port Range**: The port range setting is affected by the selected protocol type. The following table shows the relationship between protocol types and port ranges.

Protocol Type	Port Range	Scenarios

All	Shown as -1/-1, indicating all ports. You cannot set it.	Used in scenarios where both the applications are fully and mutually trusted.
All ICMP	Shown as -1/-1, indicating no port restriction. You cannot set it.	Used to detect the instance network connection status by using `ping`.
All GRE	Shown as -1/-1, indicating no port restriction. You cannot set it.	Used for VPN service.
Custom TCP Custom UDP	For custom port ranges, the valid port value is 1–65535, and the valid port range format is Start Port/End Port. A valid port range format must be used for one port. For example, use 80/80 to indicate port 80.	It can be used to allow or forbid one or several successive ports.
SSH	Shown as 22/22. After connecting to the ECS instance, you can modify the port number. For more information, see *Server default remote port modifications*.	Used for SSH to connect to a Linux instance remotely.
TELNET	Shown as 23/23.	Used to remotely log on to instances by using Telnet.
HTTP	Shown as 80/80.	The instance is used as a server for a website or a web application.
HTTPS	Shown as 443/443.	The instance is used as a server for a website or a web application that supports the HTTPS protocol.
MS SQL	Shown as 1433/1433.	The instance is used as an MS SQL server.
Oracle	Shown as 1521/1521.	The instance is used as an Oracle SQL server.
MySQL	Shown as 3306/3306.	The instance is used as a MySQL server.

RDP	Shown as 3389/3389. After connecting to the ECS instance, you can modify the port number, in particular, see *default remote access port modifications*.	Used to remotely connect to Windows instances.
PostgreSQL	Shown as 5432/5432.	The instance is used as a PostgreSQL server.
Redis	Shown as 6379/6379.	The instance is used as a Redis server.

Note:

Port 25 is disabled by default, and cannot be available by adding security group rules.

To enable Port 25, *Apply to open TCP port 25*. For more common about other ports, see *Introduction to common ECS instance ports*.

- **Authorization Type** and **Authorization Object**: The authorization object affects the setting of authorization type. The following table shows the relationship between them.

Authorization Type	Authorization Object
Address field access	Use the IP or CIDR block format such as 10.0.0.0 or 192.168.0.0/24. Only IPv4 addresses are supported. 0.0.0.0/0 indicates all IP addresses. For format introduction of CIDR, see *ECS instance subnet partition and mask representation*.
Security group access	Only for intranet access. Authorize the instances in a security group under your account or another account to access the instances in this security group. • Authorize this account: Select a security group under your account. Both security groups must be in the same VPC. • Authorize another account: Enter the target security group ID and the account ID. On the **Account Management** > **Security Settings**, you can obtain the account ID.

	For VPC network instances, security group access works for private IP addresses only. If you want to authorize Internet IP address access, use address field access.

Note:
To guarantee the security of your instance, when you are configuring an intranet inbound rule for a classic network-connected security group, **Security Group Access** is the top priority for **Authorization Type**. If you select **Address Field Access**, and you want to type an IP address in the CIDR format, type an IP address in the format of a.b.c.d/32. Only 32 is the valid CIDR prefix.

- **Priority**: The value range is 1-100. The smaller the value, the higher the priority. For more information, see *ECS security group rule priority explanation*.

7. Click **OK**.

Security group rules usually take effect immediately. A little delay is still possible.

Verify security group rules

If you have installed a web service on the instance and added a security group rule in a security group: allow all IP addresses to have inbound access to TCP port 80 of the instance. Follow these steps according to your instance OS to verify the security group rule.

Linux instances:

For a Linux instance in the security group, follow these steps to verify the security group rule:

1. *Connect to a Linux instance by using a password*.
2. Run the following command to check whether TCP 80 is being listened.

```
netstat -an | grep 80
```

If the following result returns, web service for TCP port 80 is enabled.

```
tcp        0      0 0.0.0.0:80                  0.0.0.0:*
        LISTEN
```

3. Enter `http://public IP address of the instance` in the browser address bar. If access is successful, the rules have been activated.

Windows instances:

For a Windows instance in the security group, follow these steps to verify the security group rule:

1. *Connect to a Windows instance*.
2. Run the **CMD**, and run the following command to check whether TCP Whether 80 is being listened.

   ```
   netstat -aon | findstr :80
   ```

   If the following result returns, web service for TCP port 80 is enabled.

   ```
   TCP 0.5.0.0: 80 0.5.0.0: 0 listening 1172
   ```

3. Enter `http: // instance public IP address` in the browser address bar. If access is successful, the rules have been activated.

ECS security group rule priority explanation

The **Priority** value of a security group rule ranges from 1 to 100. A smaller number indicates a higher priority.

ECS instances can belong to different security groups. As a result, instances may have multiple security group rules that have the same protocol types, port ranges, authorization types, and authorization objects. The rule that takes effect depends on the setting of **Priority** and **Authorization Policy**:

- If the rules have the same **Priority**, the **Forbid** rule takes effect, and the **Allow** rule does not.
- If the rules have different **Priority**, the rule with higher priority takes effect first, regardless the setting of **Authorization Policy** .

Related topics

- *FAQ about security groups*
- *Security group*
- *Default security group rules*
- *Implication and matching sequence of the ECS security group rule priority*
- *securityGroups-01*
- *securityGroups-01*
- *securityGroups-01*

14.6 View the security group list

You can view the security groups in the ECS console at any time. To view the security groups list, follow these steps:

1. Log on to the *ECS console*.
2. In the left-side navigation pane, click **Networks and Security** > **Security Groups**.
3. Select a region. A list of all the security groups in the specified region is displayed.
4. You can select or enter a VPC ID in the filter input box to search the security groups under this VPC.

14.7 Modify security group attributes

You can modify the name and description of a security group at any time. To modify the name and description of a security group, follow these steps:

1. Log on to the *ECS console*.
2. In the left-side navigation pane, click **Networks and Security** > **Security Groups**.
3. Select a region to display all the security groups in this region.
4. You have two options to modify the attributes of a security group:
   - Modify the name: Hover the cursor over the name of a security group, and then click the pen icon that appears.
   - Modify the name and description: Click **Modify** on the right of the security group, and then enter a new name and description in the dialog box.
5. Click **OK**.

14.8 View the security group rules

You can view the security group rules at any time. To view the security group rules, follow these steps:

1. Log on to the *ECS console*.
2. In the left-side navigation pane, click **Networks and Security** > **Security Groups**.
3. Select a region.
4. Select a security group, and click **Add Rules**.
5. The following information is displayed for security groups of classic network and VPC:
   - For VPC, **Inbound** and **Outbound** can be seen.

- For classic network, **Internet Inbound**, **Internet Outbound**, **Intranet Inbound**, and **Intranet Outbound** can be seen.

6. Click a tab to view the security group rules for that type.

14.9 Delete a security group rule

You can delete security group rules if you no longer need them. To delete rules in a security group , follow these steps:

1. Log on to the *ECS console*.
2. In the left-side navigation pane, click **Networks and Security** > **Security Groups**.
3. Select a region.
4. Find the security group where you want to delete rules, and in the **Actions** column, click **Add Rules**.
5. On the security group management page, select the rule direction and find the rule you want to delete.
   - If the security group is for classic network, the rule directions are **Internet Inbound**, **Internet Outbound**, **Intranet Inbound**, and **Intranet Outbound**.
   - If the security group is for VPC, the rule directions are **Inbound** and **Outbound**.
6. In the **Actions** column, click **Delete**.
7. On **Delete Security Group Rule** dialog box, read and confirm the notes, and then click **OK**.

You have successfully deleted a security group rule.

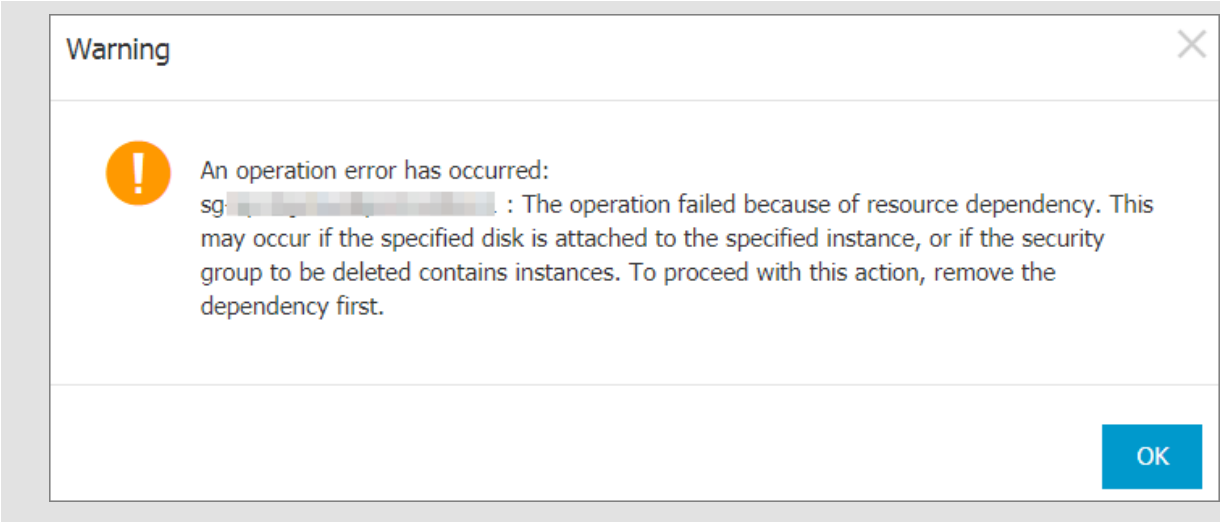
14.10 Delete a security group

You can delete security groups if you no longer need them. Deleting a security group also deletes all its rules.

Note:

When you want to delete a security group, you must make sure:

- There are no ECS instances in the security group. For more information on how to move and ECS instance out of a security group, see *add to or remove from a security group*.
- It is not referenced in the rules of another security group. You can delete a security group directly by following the steps described in this document. If the security group is authorized by another security group, error message shown in the following figure appears. You must delete the corresponding authorization rule.



Procedure

To delete a security group, follow these steps:

1. Log on to the *ECS console*.
2. In the left-side navigation pane, click **Networks and Security** > **Security Groups**.
3. Select a region to display the list of all security groups in this region.
4. Select one or more security groups, and click **Delete**.
5. In the displayed **Delete Security Group** dialog box, click **OK**.

Related APIs

- Delete a security group: *DeleteSecurityGroup*
- Query authorization relationships between a security group and another security group: *DescribeSecurityGroupReferences*
- Move an ECS instance out of a security group: *LeaveSecurityGroup*

14.11 Clone a security group

You can clone a security group across regions and network types.

Scenarios

You may need to clone a security group in the following scenarios:

- You have created a security group, named SG1, in Region A, and you want to apply the same rules of SG1 to ECS instances in Region B. Then you can clone SG1 to Region B without creating a new security group in Region B.
- You have a security group in the classic network, named SG2. You want to apply the rules of SG2 to instances in a VPC. You can clone SG2 and choose VPC as the network type when

configuring the cloning. Then in VPC network, you have a new security group that has the same rules as SG2.

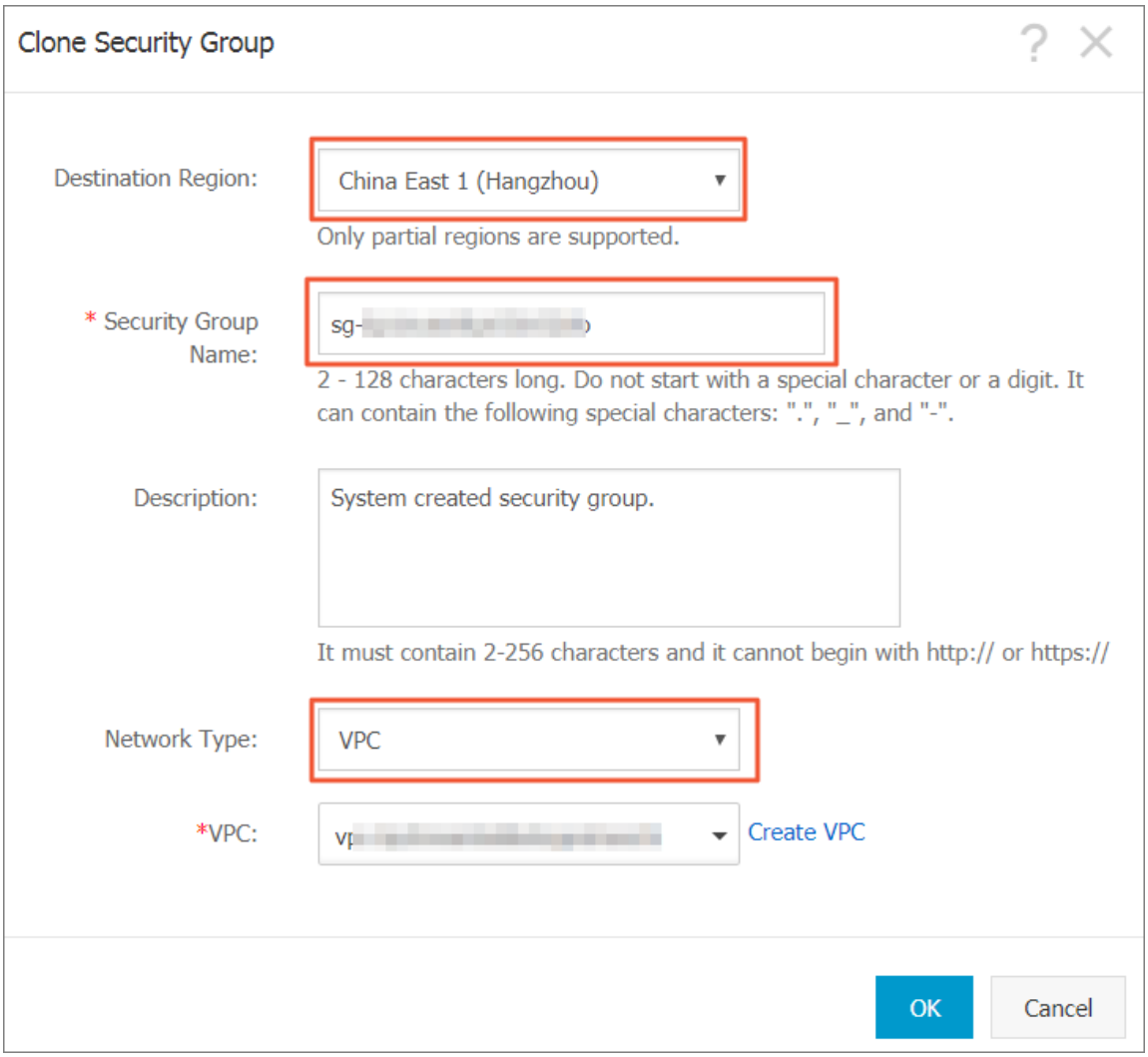
- If you want to apply new security group rules to an ECS instance that are running an online business application, we recommend that you clone the security group as a backup before modifying the rules. If the new security group rules are disadvantageous to the online business application, you can restore the rules completely or partly.

Prerequisite

If you want to change the network type of a security group from Classic to VPC, you have to *create a VPC and VSwitch* in the target region first.

Procedure

1. Log on to the *ECS console*.
2. In the left-side navigation pane, select **Networks and Security** > **Security Groups**.
3. Select a region.
4. Find the target security group, and in the **Actions** column, click **Clone Security Group**.
5. In the **Clone Security Group** dialog box, set the new security group information:
   - **Target Region**: Select a region suitable for the new security group. Not all regions are supported now. The supported regions are displayed in the drop-down list.
   - **Security Group Name**: Specify a new name for the new security group.
   - **Network Type**: Select a network type suitable for the new security group. If **VPC** is selected, you have to select a VPC in the drop-down list.



6. Click **OK**.

After successful creation, the **Clone Security Group** dialog box closes automatically. The new security group is displayed on the **Security Groups** page.

14.12 Introduction to common ECS instance ports

The following table lists commonly used ECS instance ports.

Port	Service	Description
21	FTP	A port opened by the FTP service is used for uploading and downloading files.
22	SSH	An SSH port is used to *connect to a Linux instance by using a password* in command-line mode.

23	Telnet	The Telnet port is used for Telnet to log on to the ECS instance.
25	SMTP	The port that is open to the SMTP service is used for sending mails. Based on security concerns, ECS instance Port 25 is restricted by default. Open a job request to unseal. See *request to unseal TCP 25 Port* .
80	HTTP	Provides access to HTTP services, such as IIS, Apache, and Nginx. You can *verify if TCP port 80 works properly* for port troubleshooting.
110	POP3	Used for the POP3 protocol , which is the protocol for sending and receiving emails.
143	IMAP	Used for IMAP (Internet Message Access Protocol), which is the protocol for receiving emails.
443	HTTPS	Used to provide access to the HTTPS service. HTTPS is a protocol that provides encryption and transmission through secure ports.
1433	SQL Server	The TCP port of the SQL Server is used for external service by SQL Server.
1434	SQL Server	SQL Server UDP port is used to return which TCP/IP port SQL Server uses.
1521	Oracle	Oracle communications port . The port which needs to be released by Oracle SQL is deployed on the ECS instance.

3306	MySQL	The port through which the MySQL database provides external service.
3389	Windows Server Remote Desktop Services	Windows Server Remote Desktop Services port can be used to *connect to a Windows instance*.
8080	Proxy port	As with 80 port, port 8080 is commonly used in WWW agent service to achieve web browsing. If you are using port 8080, when you visit a Web site or use a proxy server, you must add `:8080` after the IP Address: 8080. After you install the Apache Tomcat service, the default service port is 8080.
137, 138, 139	NetBIOS protocol	• 137 and 138 are UDP ports that are used to transfer files through the network neighbor. • The connection entering through the port 139 attempts to obtain the NetBIOS/smb service. NetBIOS protocols are often used for Windows files, printer sharing, and samba.

Some ports cannot be accessed

Problem: The ECS instance listens for the corresponding port, but the port is not accessible in some areas, while other ports can be accessed normally.

Analysis: Some operators judge ports 135, 139, 444, 445, 5800, 5900, and so on as high-risk ports, so they are blocked by default.

Solution: We recommend that you change the port to other non-high-risk ports for business operation.

Related topic

For more information on how to release a service port through a security group, see *add security group rules*.

14.13 Restore security group rules

Restoring security group rules indicates the process of completely or partially restoring the rules in the original security group to those of a target security group. Specifically:

- **Completely restoring** means moving the rules that do not exist in the target security group from the original security group and adding the rules that only exist in the target security group to the original security group. After restoration, rules in the original security group are identical with those in the target security group.
- **Partially restoring** means adding the rules that only exist in the target security group to the original security group and ignoring the rules that only exist in the original group.

Limits

Restoring security group rules has the following limits:

- The original security group and the target security group must be in the same region.
- The original security group and the target security group must be of the same network type.
- If any system-level security group rules, of which the priority is 110, exist in the target security group, they are not created during restoration. After restoration, the rules in the original security group may be different from what is expected. If you need the system-level security group rules , you have to manually create the rules and set their priority to 100.

Scenario

If you want to apply new security group rules to an ECS instance that is running an online business application, you can clone the former security group as a backup, and then modify the rules inside. If the new security group rules affect the online business application, you can restore the rules fully or partially.

Prerequisite

You must own at least one security group of the same network type in the same region.

Procedure

1. Log on to the *ECS console* .
2. In the left-side navigation pane, select **Networks and Security** > **Security Groups**.

3. Select a region.
4. Find the security group you want to restore rules for as the original security group, and in the **Actions** column, click **Restore Rules**.
5. In the **Restore Rules** dialog box, follow these steps:
   a. Select the **Target Security Group**: Select a security group as the target security group that must have different rules from the original security group.
   b. Select a restore **Method**:
      - If you want the original security group to have the same rules as the target security group, select **Completely Restore**.
      - If you only want to add the rules that only exist in the target security group to the original security group, select **Partially Restore**.
   c. In the **Preview** area, preview the restoration result:
      - Rules highlighted in green only exist in the target security group. No matter whether you select **Completely Restore** or **Partially Restore**, these rules are added to the original security group.
      - Rules highlighted in red are the rules that do not exist in the target security group. If **Completely Restore** is selected, the system removes these rules from the original security group. If **Partially Restore** is selected, the rules are retained in the original security group.
   d. Click **OK**.

The **Restore Rules** dialog box is closed automatically after successful creation. On the **Security Grougs** page, find the original security group you restored the rules for. In the **Actions** column, click **Add Rules** to enter the **Security Group Rules** page to view the updated security group rules.

15 Key pairs

15.1 Create an SSH key pair

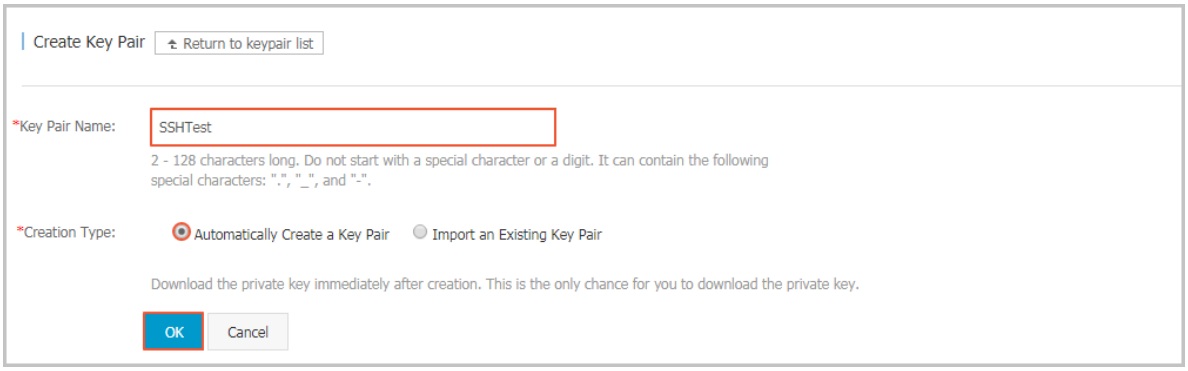
Limits

- The *SSH key pair*, abbreviated as key pair, applies to Linux instances only.
- Currently, only 2048-bit RSA key pairs are supported.
  - Alibaba Cloud holds the public key of the key pair.
  - After creating the key pair, you must save and keep the private key of the key pair for further use.
  - The private key follows the unencrypted PEM-encoded PKCS#8 format.
- An Alibaba Cloud account can have a maximum of 500 key pairs per region.

Create an SSH key pair

1. Log on to the *ECS console*.
2. In the left-side navigation pane, select **Networks and Security** > **SSH Key Pair**.
3. Select a region.
4. On the **SSH Key Pairs** page, select a region, and click **Create SSH Key Pair**.
5. On the **Create SSH Key Pair** page, enter a name for the key pair, and select **Auto-Create SSH Key Pair**.

   Note:
   The specified key pair name must be unique. It cannot be the same as that of the existing key pairs or a key pair that was bound to the instance before being deleted. Otherwise, an error message “The key pair already exists” appears.



6. Click **OK** to create the key pair.

> **Note:**
> After a key pair is created, you must download and save the private key for further use. If you do not have the private key, you cannot log on to the ECS instance.

After creating the key pair, you can view the information, including **Key Pair Name** and **Key Pair Fingerprint** in the key pair list.

Follow-up operations

After creating an SSH key pair, you can *bind or unbind it* to an ECS instance.

15.2 Import an SSH key pair

If you prefer to use another key generation tool, you can use it to generate an RSA key pair and then import its public key into Alibaba Cloud. See *SSH key pairs* for the supported types of imported key pairs.

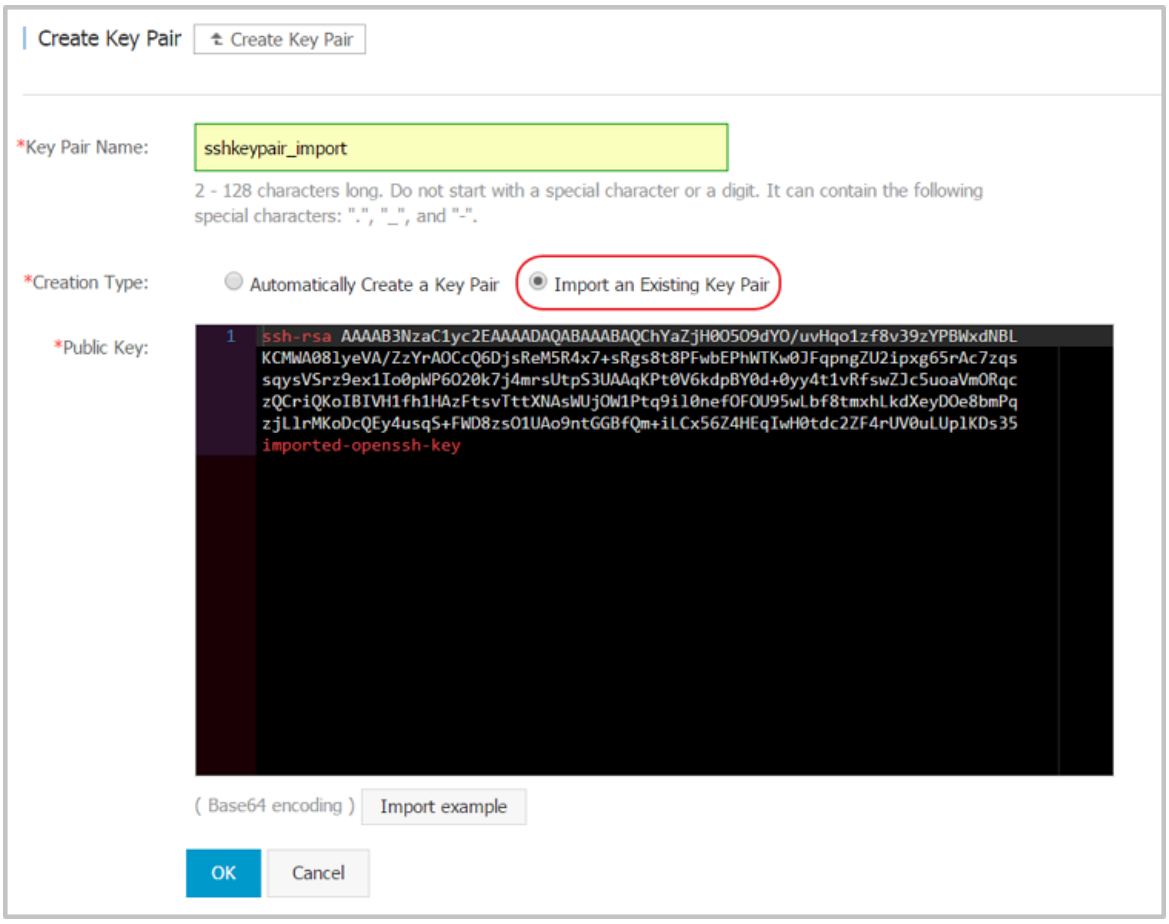
> **Note:**
> To guarantee your instance security, keep the private key of the key pair secure and do not import the private key to Alibaba Cloud.

To import an SSH key pair, you must have a key pair that has been generated using another tool. The public key to be imported into Alibaba Cloud must be Base64-encoded.

To import an SSH key pair, follow these steps:

1. Log on to the *ECS console*.
2. In the left-side navigation pane, select **Networks and Security** > **SSH Key Pair**.
3. Click **Create SSH Key Pair**.
4. Select a region.
5. On the **Create Key Pair** page, enter a name for the key pair, and select **Import SSH Key Pair**, and then enter the **Public Key**.

> **Note:**
> The specified key pair name must not be the same as that of an existing key pair or a key pair that was bound to an instance before being deleted. Otherwise, an error message “The key pair already exists” appears.



6. Click **OK**.

After creation, you can view the information, including **Key Pair Name** and **Key Pair Fingerprint**, in the key pair list.

15.3 Bind or unbind an SSH key pair

Limits

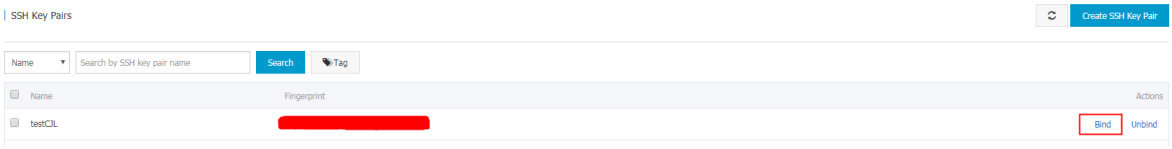
- An ECS instance can bind only one SSH key pair.
- Except for the non-I/O-optimized instances of Generation I, all the Linux instances listed in the *instance type families* support the authentication method of SSH key pair.
- When your ECS instance is in the **Running** status, *restart it* after you bind an SSH key pair to the instance.
- If the ECS instance has already bound an SSH key pair, after the new key pair is bound, the new key automatically replaces the original key.
- If you use password-based authentication for Linux logon, the password authentication feature is automatically disabled after the key pair is bound.

- After an SSH key pair is unbound, you must *reset the instance password* for successful connection.

Bind an SSH key pair

To bind an SSH key pair to an ECS instance, follow these steps:

1. Log on to the *ECS console*.
2. In the left-side navigation pane, click **Networks and Security** > **SSH Key Pair**.
3. Select a region.
4. Select a key pair, and then click **Bind** in the **Actions** column.



5. In the **Bind** dialog box, select one or more instances from the **Select Instance** box, and then click the icon **>** to move them to the **Selected** column.

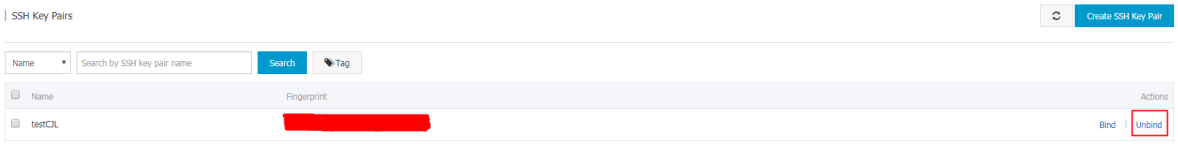
> **Note:**
> In the **Select Instance** box, the instances with gray names are either Windows instances or non-I/O-optimized instances of Generation I, for which SSH key pairs are not supported.

6. Click **OK**.

Unbind an SSH key pair

To unbind an SSH key pair from an ECS instance, follow these steps:

1. Log on to the *ECS console*.
2. Select a region.
3. In the left-side navigation pane, click **Networks and Security** > **SSH Key Pair**.
4. Select an SSH key pair, and then in the **Actions** column, click **Unbind**.



5. In the **Unbind** dialog, select one or more instances from the **Select Instance** box, and then click the icon **>** to move them to the **Selected** column.
6. Click **OK**.

15.4 Delete a SSH key pair

If you no longer require a key pair, you can delete it. Note that a deleted key pair is not recoverable. Existing instances that have used the key pair are not affected, and the deleted key pair name remains associated to the instance.

Note:

- If you delete a key pair that is still bound to an instance, its name is not available for you to create or import a key pair again. Otherwise, an error message “The key pair already exists” appears when you are using the same name to create or import a key pair.
- If you delete a key pair that is not bound to an instance, its name is still available for you to create or import a key pair again.

Follow these steps to delete one or more key pairs:

1. Log on to the *ECS console*.
2. In the left-side navigation pane, click **Networks and Security** > **SSH Key Pair**.
3. Select one or more key pairs.
4. Click **Delete**.