Elastic Compute Service

Quick Start

Quick Start

Overview

The Overview part of the Quick Start describes the operating procedure, the purpose, and the target readers of the Quick Start.

Operating procedure

Follow the steps below to purchase and use an ECS instance:

1. Select configurations.
2. Create an instance.
3. Connect to an instance.
4. For a Linux instance, format and mount a data disk. For a Windows instance, format a data disk.

Purpose of the Quick Start

The Quick Start describes how to quickly create an instance, connect to an instance remotely, and format and mount a data disk. It is designed to be the one-stop process to walk you through the creation and purchase of instances, as well as remote logon and quick environment deployment.

Elastic Compute Service instance, used and abbreviated as **ECS instance** in this document, may be also called Alibaba Cloud Elastic Compute Service, Elastic Compute Service, or ECS.

Target readers

The Quick Start is a reference for anyone who wants to get started with an ECS instance on the ECS console.

For API users, refer to API documentation.

Step 1. Select a configuration scheme

Alibaba Cloud recommends the following configuration schemes, which can meet the requirements

of most users.

- **Entry level**: 1-core CPU, 1GB memory and 1Mpbs bandwidth. This scheme is applicable to personal sites at their initial stage with relatively low traffic.
- **Advanced**: 1-core CPU, 2GB memory and 1Mpbs bandwidth. This scheme is applicable to the websites with moderate traffic, simple development environments and codebase.
- **General**: 2-core CPU, 4GB memory and 1Mbps bandwidth. This scheme is applicable to enterprise operations, parallel computing applications and common data processing. It can meet the requirements of 90% cloud computing users.
- **Ideal**: 4-core CPU, 8GB memory and 1Mbps bandwidth. This scheme is applicable to services requiring higher computing performance, for example, enterprise operations, batch processing, distributed analysis, and applications.

These recommended configuration schemes are only for reference when you start using an ECS. For subscription users, if the configuration is found to be excessively high or low during use, you can modify the configuration at any time to upgrade or degrade the configuration scheme.

For more information about web hosting, see Web Application Hosting.

After confirming the configuration scheme, you can start to create an ECS instance.

> This document describes how to create an ECS instance by using an existing image. 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, see Create an instance by using a custom image in the *User Guide* of Elastic Compute Service.

Prerequisites

Before you start creating an ECS instance, complete the following tasks:

- Signing up for Alibaba Cloud.
- Adding a payment method.
- Real-name registration for creating an ECS instance in a region inside mainland China.
- Creating at least one VPC and VSwitch for creating an ECS instance of the VPC network.
- Creating a security group that have the rules meeting your business needs for creating an ECS instance of the Classic network.

  > To improve the user experience on network, Alibaba Cloud initiated the default VPC network strategy that VPC network is set as the default network type since 12:00am (UTC+8) June 16, 2016. If the switch has not started in your region, you can still create ECS instances of the Classic network. Otherwise, the VPC network is the only option.

- Creating at least one SSH key pair for creating a Linux instance authenticated by using key pairs.

- Writing the instance custom data for customizing the startup behaviors of an instance or passing in custom data.

Procedure

To create an ECS instance, follow these steps:

Log on to the ECS console.

On the left-side navigation pane, click **Instances**.

On the **Instance List** page, click **Create Instance**.

On the purchase page, you can choose **Starter Package**, which is designed mainly for new starters. For more information about this package, see General Package Terms & Conditions . This section describes how to create an ECS instance by using the **Advanced Purchase** option.

i. **Choose Pricing Model**: You can choose either **Subscription** or **Pay-As-You-Go** as the billing method of your instance. Refer to Purchase an ECS instance in the *Purchase Guide* for the differences between both billing methods.

Choose the Datacenter Region and Zone: The region and zone cannot be changed after the instance is created. When choosing a region and zone, consider the following:

i. Generally, if the region where your instance is located is close proximity to your customers, they will experience shorter network latency and faster download speed when using your service.
ii. Some features, such as the number of zones, instance types, storage types, and network service pricing, vary by region. Select an appropriate region to meet your business needs.
iii. ECS instances in regions outside mainland China do not support interchange between Linux and Windows systems.
iv. If you are creating multiple instances for an application,
    - If it requires shorter network latency, we recommend that you create the instances in the same zone.
    - If it has higher requirements for disaster recovery, we recommend that you create the instances in different zones of one region.
v. Instances in different regions cannot communicate with each other over intranet.

Instance Type: The availability of instance types varies by region. See **Instance generations and type families** in *Product Introduction* for scenarios of each instance type.

i. If you want to create an ECS instance based on gn4 or gn5, refer to **Create a gn4 instance** or **Create a gn5 instance** in *User Guide* for details.
ii. When used with SSD cloud disks, I/O optimized instance delivers improved storage performance.
iii. To use the Windows OS for website building and Web environment deployment, you must select the instance type with **at least 2 GiB** memory.
iv. The instance type with 1 core and 1 GiB of memory does not support running MySQL.

Choose Network Type, including:

i. **Network Type**: By default, the network type is Virtual Private Cloud (VPC). You can select your own VPC and VSwitch. If you do not have one, use the default VPC and VSwitch.

If you can select **Classic Network**, you have to select a security group that meets your business needs.

ii. **Network Billing Type**: Billing by **Data Transfer**. With this billing method, charges are determined by the amount of the data transferred to an instance every hour (usually calculated by GB). The traffic price varies by region. Check the **Internet traffic fee** on the bottom of the page.
iii. **Network Bandwidth Peak**: To help prevent high charges from sudden traffic spikes, you can specify a peak bandwidth for the Instance.

Network type	Internet access?	Network bandwidth peak
VPC	Yes	• If no Elastic IP (EIP) address is used, set the peak bandwidth to a non-zero value and an

		Internet IP address will be bound to your instance. The address cannot be changed or unbound. • If an Elastic IP (EIP) address is used, set the peak bandwidth to 0Mbps.
	No	Set the peak bandwidth to 0Mbps.
Classic	Yes	Set the peak bandwidth to a non-zero value.
	No	Set the peak bandwidth to 0Mbps.

If you place your ECS instance behind a Server Load Balancer instance, you do not have to purchase bandwidth for the ECS instance. However, we recommend that you purchase a small amount of bandwidth for the ECS instance for you to connect to it to perform management tasks.

Choose the Operating System.

i. **Public Image**: It contains the image of operating system officially provided by Alibaba Cloud. On top of this, you will need to install the related software and configure the application environment based on

your specific requirements.

ii. **Custom Image**: It is generated based on the user system snapshot, including the initial system environment, application environment, and related software configuration. Selecting custom images to create ECS can save your time for repeated configuration.

iii. **Shared Image**: It is a custom image shared by an Alibaba Cloud account. Alibaba Cloud does not guarantee the integrity and security of such shared images, and you shall bear any risks associated with using these shared images.

iv. **Marketplace Image**: It is a custom image provided by either Alibaba Cloud or ISVs on the Alibaba Cloud Marketplace. Marketplace images contain the operating system image and software applications. To meet additional business needs, you need to configure the application environment.

v. Bear the following points in mind when selecting an image:

- Not all images support instance custom data. Refer to Instance custom data for supported images.
- An image based on **Windows OS**:
  - By default, the Windows 2003/2008 system allows remote connection of up to 2 sessions. For more connections, you need to buy Remote Desktop Licensing (RD Licensing) service from Microsoft separately at your expense.
  - Windows OS is suitable for running Windows programs, such as .net.
  - Windows OS supports SQL Server and other databases, which must be independently installed.
- An image based on **Linux OS**:
  - Linux OS is the most popular server operating system built with robust security and stability.
  - Linux OS is suitable for high-performance Web server and other server applications, and supports common programming languages such as PHP and Python, as well as MySQL and other databases (must be independently installed).
  - CentOS is recommended.
  - SSH key pairs are only supported by Linux OS.

Choose Storage: You have to select cloud disks for the system disk and/or data disks.

i. Cloud disk types vary by region. Refer to Disk parameters and performance test for differences of cloud disk features.

ii. **System Disk** is required. It is for installing the operating system.

iii. **Data Disk** is optional. The data disks that you add here will be released with the instance and cannot be detached from the instance. You can create a cloud disk independently.
iv. You can add up to 16 data disks, blank or from a snapshot, based on business needs.

If you selected an instance type with ephemeral storage, such as i1, d1, and d1ne, the ephemeral storage configuration will be displayed in this section. You cannot change the configuration. For more information about ephemeral storage of instance types, see Instance generations and type families in *Product Introduction*.

Security Setting. You can set authentication method based on the operating system.

Operating system	Authentication method
Linux	SSH key pairs or password
Windows	Password
i. You have two choices: Set authentication method now or later. ii. If you want to set the authentication method later, after the instance is created, reset the password or bind an SSH key pair in the ECS console.	

Set User Data, optional. If your instance type and image meet the requirements for the instance custom data, you can add the data.

Purchase Plan. You have to configure:

i. **Subscription Type**: If you are creating an ECS of the Subscription billing method, you have to set **Subscription Type** to either **1 Month** or **1 Year**. If you do not want to manually renew your instance after expiration, select **Auto-renew** to activate automatic renewal. For more information, refer to Auto-renewal in *Purchase Guide*.
ii. **Instance Name**: We recommend that you specify a name for the instance for efficient management.
iii. **Number of Instances**: You can create up to 10 instance of the Pay-As-You-Go billing method at the same time, but no quota for instances of the Subscription billing method.

Overview and cost. Check the overview and cost information to make sure that the selected configuration details are correct.

Click either **Add to Cart** (if you decide to continue shopping) or **Buy Now** (if you want to confirm the purchase).

On the **Confirm order** page, confirm the order information, and then,

i. For an instance of the Subscription billing method, click **Place Order**, make payment, and then activate the instance.
ii. For an instance of the Pay-As-You-Go billing method, click **Activate**, and then activate the instance.

When the instance is activated, you can go to the ECS console to check the instance details, such as the instance name, Internet IP address, and internal IP address (for Classic network) or private IP address (for VPC network).

After the procedure

After you create an instance, for security of your instance, we recommend that you perform security compliance inspection and configuration on:

- Linux instances: Refer to Harden operating system security for Linux in *Security Advisories*.
- Windows instances: Refer to Harden operating system security for Windows in *Security Advisories*.

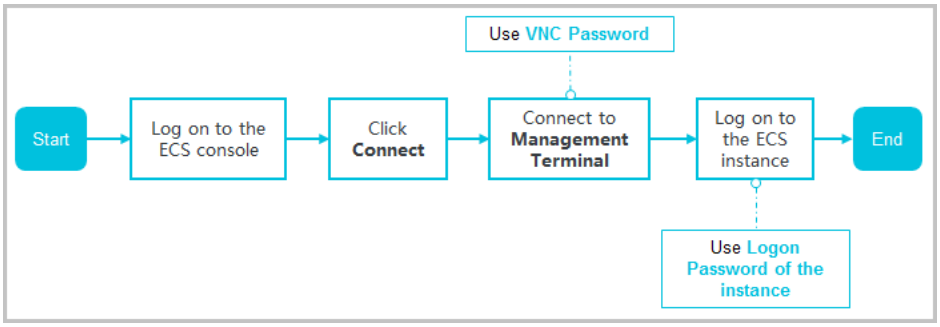
Alibaba Cloud provides you with the **Management Terminal**, which allows you to connect to your ECS instance by using a web browser, even though you do not purchase bandwidth for Internet access.

You can use other methods to connect to your ECS instance:

- To connect to a Linux instance, you can use an SSH client from Windows or SSH commands from Linux or UNIX by using SSH key pairs or password authentication.
- To connect to a Windows instance, you can use access tools from local computer.

Operating procedure

The following flowchart illustrates how to use a web browser to connect to an ECS instance.



To connect to the ECS instance by using a web browser, follow the steps:

Log on to the **ECS console**.

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

Select a region.

In the instance list, find your instance, and in the **Action** column, click **Connect**.

In the **Management Terminal** page, follow the instructions to connect to the **Management Terminal**:

- If you connect to the **Management Terminal** for the first time, follow these steps:
  a. On the pop-up **VNC Connection Password** dialog box, copy the password.

     The password appears only once. Record and securely store the password immediately. If a new password is required, change the password.

  b. Click **Close** to close the **VNC Connection Password** dialog box.
  c. On the pop-up **Enter VNC Password** dialog box, paste the VNC connection password that you copied, and then click **OK** to connect to the **Management Terminal**.

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** to connect to the **Management Terminal**.

If you lose the VNC connection password, follow these steps to connect to the **Management Terminal**:

a. Change the password.
b. On the upper left corner of the **Management Terminal** page, select **Send**

Remote Command > **Connect to Management Terminal**.

c. On the pop-up **Enter VNC Password** dialog box, enter the new password and click **OK** to finish connection.

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

- To log on to a Linux instance, enter the user name (**root**) and logon password.
  - If you lose your logon password of your instance, reset the password.
  - The logon password input is invisible.
  - If you want to do different operations within the instance, on the upper left corner of the **Management Terminal** page, select **Send Remote Command** > **CTRL + ALT + Fx**, of which **Fx** can be any one from **F1** to **F10**, to switch the interfaces for different operations.
  - If a black screen is displayed, the Linux instance may be in sleep mode. To exit sleep mode, click the mouse or press any key.
- To log on to a Windows instance, on the upper left corner of the **Management Terminal** page, select **Send Remote Command** > **CTRL+ALT+DELETE**. The Windows logon interface will then be displayed. Enter the username (**Administrator**) and password to log on.

  If you lose your logon password of your instance, reset the password.

Change the VNC connection password

If you lose the VNC connection password, follow these steps to change the password.

Notice: If the instance that you are connecting to is not I/O optimized, you need restart your instance to make the new VNC connection password take effect after you change it. The restart operation will stop the work of your instance and interrupt your business. Therefore, you must be cautious of changing the password.

Open the **Management Terminal** page.

Close the **VNC Connection Password** dialog box or **Enter VNC Password** dialog box.

On the upper right corner of the **Management Terminal** page, click **Modify Management Terminal Password** to start changing the password.

Enter a new password, which must be 6 characters in length and may contain uppercase letters, lowercase letters, and digits. Special characters are not supported.

Make the new password take effect:

- For an I/O-optimized instance, the new password takes effect immediately.
- For a non-I/O-optimized instance, restart the instance on the ECS console. Restarting directly within the instance will not apply the new password.

If you select a data disk when creating an instance, you need to format the data disk before using it.

You can also configure multiple data disk partitions based on service requirements. We recommend you to use the built-in system tool for partitioning.

Note: For ECS, either running Windows or Linux, only the partitions on the data disk, but not on the system disk, can be subdivided into multiple partitions. If you use a third-party tool to forcibly subdivide the partition on the system disk, some unknown risks, such as system crash and data loss, may occur.

In this article, the example instance has the following configurations:

- Non-I/O optimized:

  The only difference between I/O Optimized and non-I/O Optimized instances is that the later has an additional x in its device name, for example, xvdb for a non-I/O Optimized instance and vdb for an I/O Optimized instance.

- Linux OS: Redhat, CentOS, Debian, or Ubuntu, at your choice
- SSD cloud disk

Prerequisites

- At least one data disk has been attached to the Linux ECS instance. Refer to Attach a data disk in *User Guide* for details.
- You have connected to your instance. Refer to Connect to a Linux instance for details.

Operating procedure

To format and mount a data disk:

Run the **fdisk -l** command to view the data disk. **Note**: Before the data disk is partitioned and formatted, you cannot view the data disk by running the **df –h** command. In the following example, a 5GB data disk needs to be mounted.
If you do not find /dev/xvdb after running the **fdisk -l** command, it indicates that your instance does not have a data disk. Therefore, mounting is not required. In this case, you can skip this chapter.

```
[root@xxxx ~]# fdisk -l

Disk /dev/xvda: 42.9 GB, 42949672960 bytes
255 heads, 63 sectors/track, 5221 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes
Disk identifier: 0x00078f9c

Device Boot Start End Blocks Id System
/dev/xvda1 * 1 5222 41940992 83 Linux

Disk /dev/xvdb: 5368 MB, 5368709120 bytes
255 heads, 63 sectors/track, 652 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes
Disk identifier: 0x00000000
```

Run the following command to partition the data disk.

```
fdisk /dev/xvdb
```

Enter the commands **n**, **p**, **1** in sequence as prompted, press the Enter key twice, and then enter the **wq** command. The partitioning will begin.

```
[root@xxx ~]# fdisk /dev/xvdb
Device contains neither a valid DOS partition table, nor Sun, SGI or OSF disklabel
Building a new DOS disklabel with disk identifier 0x33eb5059.
Changes will remain in memory only, until you decide to write them.
After that, of course, the previous content won't be recoverable.

Warning: invalid flag 0x0000 of partition table 4 will be corrected by w(rite)

WARNING: DOS-compatible mode is deprecated. It's strongly recommended to
switch off the mode (command 'c') and change display units to
sectors (command 'u').

Command (m for help): n
Command action
e extended
p primary partition (1-4)
p
Partition number (1-4): 1
First cylinder (1-652, default 1):
Using default value 1
Last cylinder, +cylinders or +size{K,M,G} (1-652, default 652):
Using default value 652

Command (m for help): wq
The partition table has been altered!
```

```
Calling ioctl() to re-read partition table.
Syncing disks.
```

Run the **fdisk -l** command to view the new partition. A new partition is created, for example, /dev/xvdb1 shown in the below.

```
[root@xxx ~]# fdisk -l

Disk /dev/xvda: 42.9 GB, 42949672960 bytes
255 heads, 63 sectors/track, 5221 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes
Disk identifier: 0x00078f9c

Device Boot Start End Blocks Id System
/dev/xvda1 * 1 5222 41940992 83 Linux

Disk /dev/xvdb: 5368 MB, 5368709120 bytes
255 heads, 63 sectors/track, 652 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes
Disk identifier: 0x33eb5059

Device Boot Start End Blocks Id System
/dev/xvdb1 1 652 5237158+ 83 Linux
```

Run the following command to format the new partition. The formatting time depends on the size of the data disk. You can also choose other file formats, for example, ext4.

```
mkfs.ext3 /dev/xvdb1
```

Run the following command to write the new partition information.

```
echo '/dev/xvdb1 /mnt ext3 defaults 0 0'>> /etc/fstab
```

Upon completion, run the cat /etc/fstab command to view the information.
Note: Ubuntu 12.04 does not support barrier. Therefore, the correct command for the system is as follows:

```
echo '/dev/xvdb1 /mnt ext3 barrier=0 0 0'>>/etc/fstab
```

To mount the data disk to a folder separately, for example, to store webpages separately, modify the /mnt part in the above command.

Run the **mount /dev/xvdb1 /mnt** command to mount the new partition. Then, run the **df -h** command to view the partition. If data disk information is displayed, the new partition has been mounted successfully and can be used.

```
 [root@xxx ~]# mount /dev/xvdb1 /mnt
[root@xxx ~]# df -h
Filesystem Size Used Avail Use% Mounted on
/dev/xvda1 40G 1.5G 36G 4% /
tmpfs 498M 0 498M 0% /dev/shm
/dev/xvdb1 5.0G 139M 4.6G 3% /mnt
```

Note: ECS does not support installation and deployment of virtualization software, for example, KVM, Xen, and VMware.

To format a data disk on a Linux instance, refer to Linux _ Format and mount a data disk.

If you have already attached a data disk to your ECS instance, you must log on to the instance to partition and format the data disk for use. This document describes how to create a single-partition data disk using a new data disk and format the partition. You can also create multiple partitions on a data disk according to your business needs.

Warning:

- Disk partitioning and formatting are high-risk operations, so please proceed with caution. This document describes how to deal with a new blank data disk. If you have data on a data disk, make sure that you have created a snapshot of the data disk to avoid any possible data loss.
- ECS only supports partitioning data disks, not system disks. Forcibly partitioning a system disk using a third-party tool may cause unknown risks such as system crashes or data loss.

Prerequisites

For a separately purchased data disk, you must attach it to an instance in the console before partitioning and formatting.

A data disk purchased along with the instance can be partitioned and formatted without being attached.

Procedure

This example describes how to partition and format a 20 GB data disk on the 64-bit Windows Server

2012 R2.

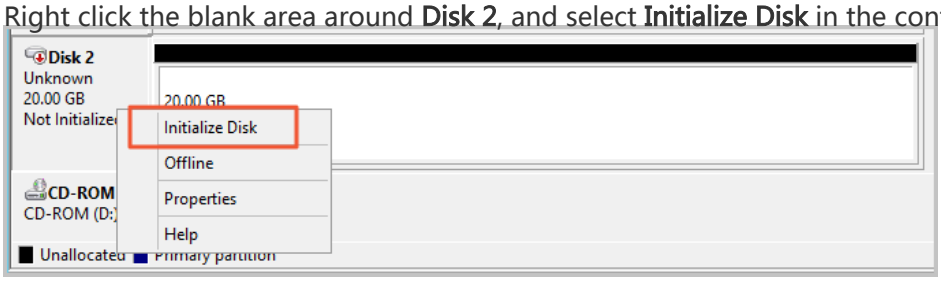
Connect to an instance.

On the Windows Server desktop, right click the Start icon, and select **Disk Management**. The data disk not formatted and partitioned (in this example, **Disk 2**) is in the **Offline** status.

Right click the blank area around **Disk 2**, and select **Online** in the context menu.

After going online, the status of **Disk 2** is displayed as **Not Initialized**.

Right click the blank area around **Disk 2**, and select **Initialize Disk** in the context menu.



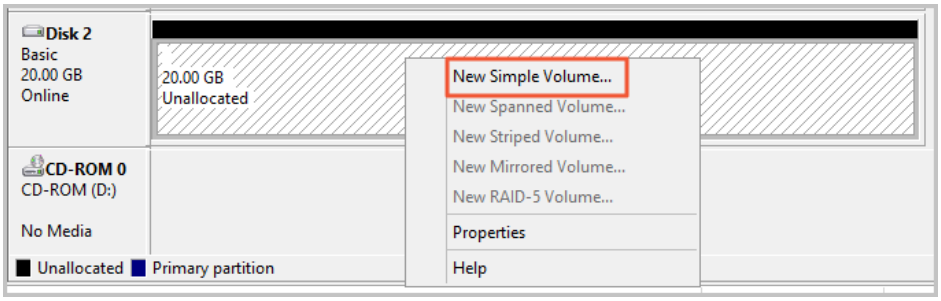
In the **Initialize Disk** dialog box, select **Disk 2**, and select a partitioning method:

- **MBR** is still the most common partitioning method. However, this method only supports a maximum partition size of 2 TB and creation of a maximum of four primary partitions. If you want to divide a disk into more than four partitions, you need to take a primary partition as an extended partition and create logical partitions in it.

  GPT is a new partitioning method, which is not recognizable to earlier versions of Windows. The maximum partition size that GPT can support is determined by the operating system and the file system. On the Windows operating system, GPT supports creating up to 128 primary partitions.

  In this example, we select the MBR partitioning method, and click **Yes**.

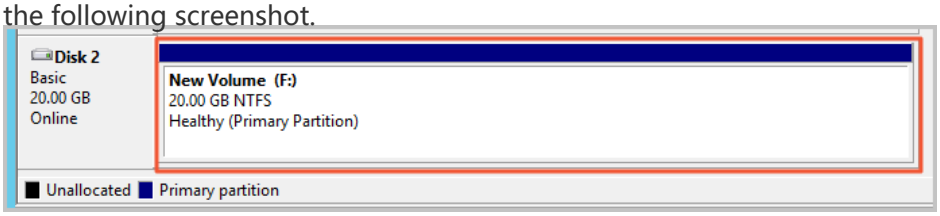
In the **Disk Management** window, right click the **Unallocated** area of **Disk 2** and select **New Simple Volume**.



In the **New Simple Volume Wizard**, do the followings:

i. Click **Next**.
ii. **Specify Volume Size**: Specify the size of the simple volume to create. If you need only one primary partition, use the default value, and then click **Next**.
iii. **Assign Drive Letter or Path**: Select a drive letter (in this example, F). Click **Next**.
iv. **Format Partition**: Select the formatting settings, including file system, allocation unit size, and volume label, and confirm whether to **Perform a quick format** and **Enable file and folder compression**. Use the default settings here. Click **Next**.
v. Start creating a new simple volume. When a new simple volume has been created, click **Finish** to close **New Simple Volume Wizard**.

After the partition formatting is completed, the status of **Disk 2** in **Disk Management** is as shown in the following screenshot.



You can view the new drive **New Volume (F:)** in **This PC**, which means the data disk is ready for use.