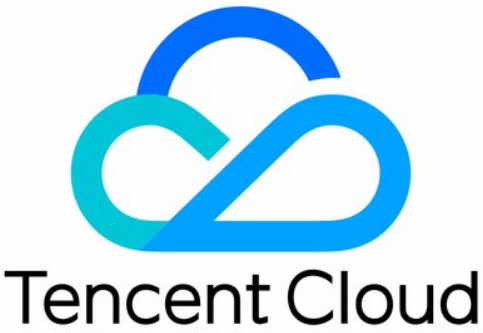
Cloud Virtual Machine

Product Introduction

Product Documentation



Contents

Product Introduction
CVM Overview

Last updated：2024-01-08 09:12:28

CVM Overview

Tencent Cloud Virtual Machine (CVM) is a scalable cloud computing service that frees you from estimation of resource usage and upfront investment. With Tencent Cloud CVM, you can start CVMs and deploy applications immediately.
You can customize all resources of a CVM instance, including CPU, memory, disk, network, and security policies. You can also easily adjust the resources in response to any change in demand.

Using CVM instances

You can configure and manage CVM instances in the following ways:
Console: a web-based UI for configuring and managing CVM instances.
API: Tencent Cloud also provides APIs for configuring and managing CVM instances. For more information, see API Category.
SDK: you can use SDK or Tencent Cloud CLI to call CVM APIs.

Key Concepts

Before using Tencent Cloud CVM, you should familiarize yourself with the following concepts:

Concept	Description
Instance	A virtual computing resource containing basic computing components such as CPU, memory, OS, network, and disks. Tencent Cloud provides various configurations of CPU, MEM, storage, and networking capacity for CVM instances. For more information, please see Instance Types.
Image	A pre-configured template containing an operating system and applications that CVM instances run on. Tencent Cloud CVM provides pre-configured images for Windows, Linux, etc.
Cloud Block Storage	A distributed and persistent block storage device provided by Tencent Cloud that can serve as the system disk or an expandable data disk of an instance.

Virtual Private Cloud	A logically isolated virtual network space in Tencent Cloud.
IP address	Tencent Cloud provides Private IP and Public IP addresses. Private IP addresses are for the interconnection of CVM instances within the same LAN, while public IP addresses are for public-facing services.
Elastic IP	Static public network IP addresses designed especially for dynamic networks to meet the demands for fast troubleshooting.
Security group	A virtual firewall that features stateful data packet filtering. It is used to configure the network access control of CVMs. Security groups are an important measure for network security and isolation.

Purchasing and Customizing CVM Instances

If you have specific needs that our standard CVM specifications cannot meet, use the following guides to learn how to obtain custom configurations:
Customizing Windows CVM Configurations
Customizing Linux CVM Configurations

CVM Prices

CVM supports pay-as-you-go. For more information, see Price of CVM Instance.
For pricing information on CVM instances and other resources, refer to Product Pricing.

Relevant Products

Auto Scaling lets you scale server clusters using pre-defined criteria such as time or load. For more information, refer to the Auto Scaling documentation.
Cloud Load Balancer allows you to automatically distribute client traffic among multiple CVM instances. For more information, refer to the Cloud Load Balancer documentation.
Tencent Kubernetes Engine allows you to manage the lifecycle of applications in a CVM. For more information, refer to the Tencent Kubernetes Engine documentation.
Cloud Monitor keeps track of your CVM instances and their system disks. For more information, refer to the Basic Cloud Monitor documentation.
You can deploy a relational database on the cloud or use TencentDB. For more information, refer to the TencentDB for MySQL documentation.

Advantages

Last updated：2025-06-11 15:16:40

Comprehensive Service Offerings

Tencent Cloud CVM offers a wide range of options for your data storage needs.
Region and availability zones across the globe: Tencent Cloud offers nodes in South China, East China, North China, and Southwest China in Mainland China. Overseas coverage includes Southeast Asia, Asia Pacific, North America, Western U.S., and Europe. Deploy applications in regions close to the location of your clients reduces latency.
Model Specifications
Standard: suitable for small and medium-sized web applications and databases.
MEM optimized: suitable for applications that require memory-intensive operations such as searches and computing.
High IO: suitable for low latency, I/O-intensive applications.
Computing: suitable for large game servers, advertisement service engines, high performance computing, and other compute-intensive applications.
Big data: suitable for throughput-intensive applications such as Hadoop distributed computing, massive log processing, distributed file systems, and large data warehouses.
Heterogeneous: suitable for high-performance applications such as deep learning, scientific computing, video encoding/decoding, and graphics workstations.
Batch-based: suitable for compute-intensive applications that use super large computing nodes frequently in a short time, such as rendering, gene analysis, and crystal pharmacy.

Flexible Configuration

We are committed to building the most flexible cloud service management platform in the industry with the following capabilities:
Hardware: **instant upgrading/downgrading** of hardware specifications for CVMs based on CBS.
Storage: **instant storage expansion** for CVMs based on CBS.
Bandwidth: **instant upgrading/downgrading** of bandwidth for CVMs.
OS: switch between Windows and Linux at any time for CVMs in Mainland China. Coming soon for other areas.
EIP: for CVMs in various network environments.
Image: public images (with multiple Linux and Windows variants) and custom images (image created by users using the image creation feature). Cross-region adjustment and image duplication are supported.

Custom network architecture: Virtual Private Cloud (VPC) provides independent network space, custom IP address ranges, IP addresses, and routing policies. It offers port-level access control to support full logical isolation for networks.

Reliability

We are committed to building the most reliable cloud service in the industry.
CVM reliability: we guarantee 99.975% service availability for CVMs with 99.9999999% data reliability. Features including imperceptible migration, data snapshots, and automatic alerts are supported to ensure the security of your servers.
Cloud disk policy: multiple copies of user data eliminates single point of failure and ensure data reliability. You can rest assured when moving your data to Tencent Cloud and do not have to worry about data loss.
Stable network architecture: our sophisticated network virtualization technology and ENI binding technology deliver high network availability. Operations in T3+ IDCs ensure the reliability of the operating environment and frees you from having to worry about network availability.

Fast

We are committed to providing fast and convenient services for both user operations and CVM performance.
Convenient and rapid operations: deploy one, hundreds, or even thousands of server instances in minutes, and purchase, configure, manage, or expand your services with just one click.
Top-speed public network: our BGP public network with more than 20 lines covers almost all ISPs. No matter which ISP your customers are using, they can always enjoy the same top-speed bandwidth and failover in seconds.
Top-speed private network: the private networks of Tencent Cloud's data centers in the same region are interconnected, all with megabyte or gigabyte connections, ensuring high communication quality within private networks.

Security

Tencent Cloud provides multiple solutions to ensure the security of CVMs. In addition, the backup and rollback mechanisms can protect data security.
Various methods for remote login to the CVM: Various login methods are provided, including login with key/password or VNC login.
Diverse security services: DDoS protection, DNS hijacking detection, intrusion detection, vulnerability scanning, web trojan detection, login protection, and other security services are provided to safeguard your server.

Free cloud monitoring: multiple types of real-time alerts are supported.
Custom access control: customize CVM and network access policies through security groups and network ACL, which serve as firewalls for different scenarios.
Tencent Cloud security highlights:
Comprehensive security protection
Comprehensive security services are provided for CVMs, including security inspections (vulnerability scanning, trojan detection, backdoor detection, port inspection, etc.) and security protection (DDoS protection, intrusion detection, access control) to ensure data security and user privacy.
Real-time alerts and scheduled analysis
24/7 security services so we can identify vulnerabilities and notify you in real time.
Free and easily accessible security protection
You do not need to buy expensive security devices for your data anymore. Our cloud security service comes free with your CVM. On top of that, one-click activation without deployment makes it easy to use.
Professional team, reliable service
The team responsible for Tencent Cloud security has years of experience in providing professional and trustworthy security services to users.

Ease of use

A wide range of officially verified application software and operation and maintenance tools make operation and maintenance more convenient, so that you no longer have to worry about choosing the right tools.
Tencent Cloud CVM provides a web-based user interface, the Console, that allows you to perform operations such as system boot, configuration, and OS reinstallation on CVM instances just like on physical machines. If you have signed up for a Tencent Cloud account, you can log in to the CVM console to operate on your CVM.
Tencent Cloud CVM also offers plenty of APIs which you can utilize to implement the automated operation and maintenance system that tightly integrates with your existing O&M structure and provides the ideal solution for your service needs. These APIs use HTTP/HTTPS requests. For more information about CVM API operations, refer to our API Documentation.

Cost-effective

Tencent Cloud lets you save on IT investment, simplify O&M operations.
You only pay for what you need. No up-front investment, no infrastructure, no preparation. Tencent Cloud is the perfect cloud platform for you.

Regions and Availability Zones

Last updated：2024-09-04 16:39:39

Regions

Introduction

A region is the physical location of an IDC. In Tencent Cloud, regions are fully isolated from each other, ensuring cross-region stability and fault tolerance. When purchasing Tencent Cloud services, we recommend selecting the region closest to your end users to minimize access latency and improve download speed.
You can view the following table or use the DescribeRegions API to get a complete list of regions.

Characteristics

The networks of different regions are fully isolated. Tencent Cloud services in different regions **cannot communicate via a private network by default**.
Tencent Cloud services across regions can communicate with each other through public IPs as instructed in Internet Access, while those in different VPCs can communicate with each other through Cloud Connect Network that is faster and steadier.
Cloud Load Balancer (CLB) currently supports intra-region traffic forwarding by default. If you enable the cross-region binding feature, a CLB instance can be bound to CVM instances in another region.

Availability Zones

Introduction

An availability zone (AZ) is a physical IDC of Tencent Cloud with independent power supply and network in the same region. It can ensure business stability, as failures (except for major disasters or power failures) in one AZ are isolated without affecting other AZs in the same region. By starting an instance in an independent AZ, users can protect their applications from being affected by a single point of failure.
You can view the following table or use the DescribeZones API to get a complete list of availability zones.

Characteristics

Tencent Cloud products that are in the same region, different availability zones, and the same VPC are interconnected through the private network. They can be accessed directly through private IP addresses.
Note:
Private network interconnection refers to the interconnection of resources under the same account. Resources under different accounts are completely isolated on the private network.

Mainland China

Region	AZ
South China (Guangzhou) ap-guangzhou	Guangzhou Zone 1 (sold out) ap-guangzhou-1
	Guangzhou Zone 2 (sold out) ap-guangzhou-2
	Guangzhou Zone 3 (sold out) ap-guangzhou-3
	Guangzhou Zone 4(sold out) ap-guangzhou-4
	Guangzhou Zone 6 ap-guangzhou-6
	Guangzhou Zone 7 ap-guangzhou-7
East China (Shanghai) ap-shanghai	Shanghai Zone 1 (sold out) ap-shanghai-1
	Shanghai Zone 2 ap-shanghai-2
	Shanghai Zone 3(sold out) ap-shanghai-3
	Shanghai Zone 4(sold out) ap-shanghai-4
	Shanghai Zone 5 ap-shanghai-5
	Shanghai Zone 8 ap-shanghai-8
East China (Nanjing) ap-nanjing	Nanjing Zone 1 ap-nanjing-1
	Nanjing Zone 2(sold out) ap-nanjing-2
	Nanjing Zone 3

	ap-nanjing-3
North China (Beijing) ap-beijing	Beijing Zone 1 (sold out) ap-beijing-1
	Beijing Zone 2(sold out) ap-beijing-2
	Beijing Zone 3 ap-beijing-3
	Beijing Zone 4(sold out) ap-beijing-4
	Beijing Zone 5(sold out) ap-beijing-5
	Beijing Zone 6 ap-beijing-6
	Beijing Zone 7 ap-beijing-7
	Beijing Zone 8 ap-beijing-8
Southwest China (Chengdu) ap-chengdu	Chengdu Zone 1 ap-chengdu-1
	Chengdu Zone 2(sold out) ap-chengdu-2
Southwest China (Chongqing) ap-chongqing	Chongqing Zone 1 ap-chongqing-1
Hong Kong/Macao/Taiwan (Hong Kong, China) ap-hongkong	Hong Kong Zone 1 (Nodes in Hong Kong, China can cover Hong Kong/Macao/Taiwan regions) ap-hongkong-1
	Hong Kong Zone 2 (Nodes in Hong Kong, China can cover services in Hong Kong/Macao/Taiwan regions) ap-hongkong-2
	Hong Kong Zone 3 (Nodes in Hong Kong, China can cover Hong Kong/Macao/Taiwan regions) ap-hongkong-3

Note:
The product is in beta test for Jinan, Hangzhou, Fuzhou, Wuhan, Changsha, and Shijiazhuang regions. To try it out, contact the sales rep for application.

Other countries and regions

Region	AZ
Southeast Asia (Singapore) ap-singapore	Singapore Zone 1 (Nodes in Singapore can cover Southeast Asia) ap-singapore-1
	Singapore Zone 2 (Nodes in Singapore can cover Southeast Asia) ap-singapore-2
	Singapore Zone 3 (Nodes in Singapore can cover Southeast Asia) ap-singapore-3
	Singapore Zone 4 (Nodes in Singapore can cover Southeast Asia) ap-singapore-4
Southeast Asia (Jakarta) ap-jakarta	Jakarta Zone 1 (Nodes in Jakarta can cover Southeast Asia) ap-jakarta-1
	Jakarta Zone 2 (Nodes in Jakarta can cover Southeast Asia) ap-jakarta-2
Northeast Asia (Seoul) ap-seoul	Seoul Zone 1 (Nodes in Seoul can cover Northeast Asia) ap-seoul-1
	Seoul Zone 2 (Nodes in Seoul can cover Northeast Asia) ap-seoul-2
Northeast Asia (Tokyo) ap-tokyo	Tokyo Zone 1 (Nodes in Tokyo can cover Northeast Asia) ap-tokyo-1
	Tokyo Zone 2 (Tokyo nodes can cover services in Northeast Asia) ap-tokyo-2
Southeast Asia (Bangkok) ap-bangkok	Bangkok Zone 1 (Nodes in Bangkok can cover Southeast Asia) ap-bangkok-1
	Bangkok Zone 2 (Nodes in Bangkok can cover Southeast Asia) ap-bangkok-2
South America (São	Sao Paulo Zone 1 (Sao Paulo nodes cover services in South America)

Paulo) sa-saopaulo	sa-saopaulo-1
West US (Silicon Valley) na-siliconvalley	Silicon Valley Zone 1 (Nodes in Silicon Valley can cover Western US) na-siliconvalley-1
	Silicon Valley Zone 2 (Nodes in Silicon Valley can cover Western US) na-siliconvalley-2
East US (Virginia) na-ashburn	Virginia Zone 1 (Nodes in Virginia can cover Eastern US) na-ashburn-1
	Virginia Zone 2 (Nodes in Virginia can cover Eastern US) na-ashburn-2
Europe (Frankfurt) eu-frankfurt	Frankfurt Zone 1 (Nodes in Frankfurt can cover Europe) eu-frankfurt-1
	Frankfurt Zone 2 (Nodes in Frankfurt can cover Europe) eu-frankfurt-2

Regions and AZ Selection

When selecting a region and AZ, take the following into consideration:

Your location, the location of your users, and the region of the CVM instances.

We recommend that you choose the region closest to your end users when purchasing CVM instances to minimize access latency and improve access speed.

Other Tencent Cloud services you use.

When you select other Tencent Cloud services, we recommend that you try to locate them all in the same region and AZ to allow them to communicate with each other through the private network, reducing access latency and increasing access speed.

High availability and disaster recovery.

Even if you have just one VPC, we still recommend that you deploy your businesses in different availability zones to prevent a single point of failure and enable cross-AZ disaster recovery.

There may be network latency among different AZs. We recommend you assess your business requirements and find the optimal balance between high availability and low latency.

If you need access to CVM instances in other countries or regions, we recommend you select a CVM in those other countries or regions. If you use a CVM instance in China to access servers in other countries and regions, you may encounter much higher network latency.

Resource Availability

The following table describes which Tencent Cloud resources are global, which are regional, and which are specific to availability zones.

Resource	Resource ID Format -8-Digit String of Numbers and Letters	Type	Description
User Account	No limit	Globally unique	Users can use the same account to access Tencent Cloud resources around the world.
SSH Keys	skey-xxxxxxxx	Global	Users can use an SSH key to bind a CVM in any region under the account.
CVM Instances	ins-xxxxxxxx	Instances are AZ-specific.	Users can only create a CVM instance in a specific AZ.
Custom Images	img-xxxxxxxx	Available in multiple availability zones of a region	Custom images created for the instance are available to all availability zones of the same region. Use Copy Image to copy a custom image if you need to use it in other regions.
Elastic IPs	eip-xxxxxxxx	Available in multiple availability zones of a region	EIPs are region-specific and can only be associated with instances in the same region.
Security Group	sg-xxxxxxxx	Available in multiple availability zones of a region	Security groups can only be associated with instances in the same region. Tencent Cloud automatically creates three default security groups for users.
Cloud Block Storage	disk-xxxxxxxx	Instances are AZ-specific.	Users can only create a Cloud Block Storage disk

			in a specific AZ and attach it to instances in the same availability zone.
Snapshots	snap-xxxxxxxx	Available in multiple availability zones of a region	A snapshot created from a cloud disk can be used for other purposes (such as creating cloud disks) in this region.
Cloud Load Balancer	clb-xxxxxxxx	Available in multiple availability zones of a region	Cloud Load Balancer can be bound with CVMs in different availability zones of a single region for traffic forwarding.
VPC	vpc-xxxxxxxx	Available in multiple availability zones of a region	A VPC in one region can have resources created in different availability zones of the region.
Subnets	subnet-xxxxxxxx	Instances are AZ-specific.	Users cannot create subnets across availability zones.
Route Tables	rtb-xxxxxxxx	Available in multiple availability zones of a region	When creating a route table, users need to specify a VPC. Therefore, route tables are regional as well.

Relevant Operations

Migrating an instance to another availability zone

Once launched, an instance cannot be migrated to another availability zone. However, you can create a custom image of the CVM instance and use the image to launch or update an instance in a different availability zone.

1. Create a custom image for the current instance. For more information, see Create Custom Images.
2. If the instance is on a VPC network environment and you want to retain its current private IP address after the migration, first delete the subnet in the current availability zone and then create a subnet in the new availability zone with the same IP address range. Note that a subnet can be deleted only when it contains no available instances. Therefore, all the instances in the current subnet should be migrated to the new subnet.

3. Create a new instance in the new availability zone by using the custom image you have just created. You can choose the same type and configuration as the original instance, or choose new settings. For more information, see Creating Instances via CVM Purchase Page.
4. If an elastic IP is associated with the original instance, you need to dissociate it from the old instance and associate it with the new instance. For more information, see Elastic IP.
5. (Optional) If the original instance is pay-as-you-go, then you can choose to terminate the original instance. For more information, see Terminate Instances.

Copying images to other regions

Operations such as launching and viewing instances are region-specific. If the image of the instance that you need to launch does not exist in the region, the image needs to be copied to the desired region. For more information, see Copying Images.

Instance
Instance Overview

Last updated：2024-01-08 09:12:28

Overview of a CVM Instance

An **instance** is a Cloud Virtual Machine (CVM). It contains basic computing components such as the CPU, the memory, the operating system, the network, and the disk.
CVM instances provide elastic computing services in the cloud in a secure and reliable way to meet computing requirements. As business demands change, computing resources can be scaled in real time to lower your software and hardware costs and simplify IT OPS work.
Each instance type offers different computing and storage capabilities, making them suitable for different use cases. You can choose the computing capacity, storage, and network access method of the instance based on the scope of the service you need to provide. For more information about instance types and use cases, see Instance Types. After you launch an instance, you can use it as you would any traditional computer. You will also have complete control over your instances.

Instance Image

An **Image** is a template that contains software configurations (operating systems, pre-installed programs, etc.) required for launching CVM instances. You can use an image to launch an instance or multiple instances repeatedly. In other words, an image is the “installed disk” of the CVM.
Tencent Cloud provides the following types of images:
Public image: available to all users and suitable for major operating systems.
Custom image: only available to the creator and the users with whom the image is shared. A custom image is created from running instances or imported from external sources.
Shared image: shared by other users. They can only be used to create instances.
For more information about images, see Overview or Image Types Overview.

Instance Storage

Similar to a normal CVM, instances can be stored in the **system disk** and the **data disk**:
System disk: similar to the C drive in the Windows system. The system disk contains a full copy of the image used to launch an instance and the running environment for the instance. A system disk larger than the used image is required

when an instance is launched.
Data disk: similar to the D and E drives in the Windows system. The data disk saves user data and supports flexible expansion, mounting, and unmounting.
Both the system disk and the data disk support different types of storage provided by Tencent Cloud. For more information, please see Storage Overview.

Instance Security

Tencent Cloud provides the following instance security protection methods:
Policy: when different accounts need to control the same set of cloud resources, you can use a policy to control their access to cloud resources.
Security groups: you can use a security group to control the access of trusted addresses to instances.
Login control: log in to your Linux instances using SSH keys as much as possible. If you log in to your Linux instances using login passwords, change your passwords periodically.

Families and Models

Last updated：2025-03-21 16:40:49

When creating a Tencent Cloud CVM, the instance type you specify determines the instance's host hardware configuration. Each instance type provides different computing, memory and storage features. You can choose the instance type that suits your application scale. Tencent Cloud provides a range of instance families with varying combinations of CPU, memory, storage, heterogeneous hardware, and network bandwidth. This gives you the flexibility in selecting the appropriate mix of resources for your applications.
Tencent Cloud CVM is pay-as-you-go. Instances can be activated and terminated at any time, and you only pay for resources used. This flexible billing mode lets you meet fluctuating demands while keeping costs low. In addition, usage is billed in seconds to further maximize your savings.

Instance Type

Tencent Cloud instance families are categorized as follows:

Type	Sub-type	Description
Standard Instance Family	Standard S8 Standard SA5 Standard SA4 Standard S6 Standard SA3 Standard SR1 Standard S5 Standard Storage Optimized S5se Standard SA2 Standard S4 Standard Network Optimized SN3ne Standard S3 Standard SA1 Standard Network Optimized S2ne Standard S2 Standard S1	Standard instances provide a balance of compute, memory, and network resources to accommodate most applications.
Memory Optimized	MEM-optimized M8 MEM-optimized MA5 MEM-optimized MA4	Memory optimized instances feature large memory and are suitable for applications that require extensive memory

Instance Family	MEM-optimized MA3 Memory Optimized M6 Security-enhanced Memory Optimized M6ce Memory Optimized M5 Memory Optimized MA2 Memory Optimized M4 Memory Optimized M3 Memory Optimized M2 Memory Optimized M1	operations, searches, and computations, such as high-performance databases and distributed memory caching.
High I/O Instance Family	High I/O ITA5 High I/O IT5 High I/O IT3	High I/O instances feature high random IOPS, high throughput and low latency. They are suitable for I/O-intensive applications that require high disk read/write performance and low latency, such as high-performance databases.
Big Data Instance Family	Big Data D3 Big Data D2	This instance family is equipped with massive storage resources, features high throughput, and is suitable for throughput-intensive applications such as Hadoop distributed computing, massive log processing, distributed file systems, and large data warehouses.
Computing Instance Family	Compute Optimized C6 Compute Optimized C5 Compute Optimized C4 Compute Network-optimized CN3 Compute C3 Compute C2	This family comes with a turbo frequency up to 3.8 GHz, and provides the highest single-core computing performance. It is suitable for compute-intensive applications such as batch processing, high performance computing, and dedicated game servers.
BeFast Instance Family	BeFast 1	BeFast 1 instances provide cost-effective, balanced and stable computing, memory and network resources. Instances of this family are dynamically managed through the Tencent Cloud's intelligent scheduling capability, meeting the needs for application resources in most scenarios.
Heterogeneous Computing Instance Family	-	This family is equipped with heterogeneous hardware such as GPU and FPGA to deliver real-time, fast parallel computing and floating-point computing capabilities. It is suitable for high-performance applications such as deep learning, scientific computing, video encoding/decoding, and graphics workstations.

Note:
Some of the families are discontinued.

Glossary

Send/Receive Packets: the maximum number of data packets that the instance can process per second for both sending and receiving, without differentiating private or public network traffic.
Private Network Bandwidth Capacity: the maximum number of bits that the instance can transfer per second over private network.
Multi-queue Technology: multiple send/receive packet queues are supported on a virtual ENI, with the sending and receiving of the data packets of each queue capable of being simultaneously processed by different CPUs.
Number of Queues: the quantity of send/receive packet queues supported by each virtual ENI (N number of queues indicates that N number of receiving queues and N number of sending queues can be simultaneously supported).

Restrictions on Instances

The total number of instances that can be enabled in one region is limited. For more information, please see CVM Instance Purchase Limit.
Restrictions on system and data disks mounted on an instance: to ensure premium disk I/O performance, Tencent Cloud sets limits on the size and type of data disks purchased with an instance. For more information, see the supported disks of each instance family. You can also purchase separate cloud disks if you have higher disk requirements.
Note that the private network bandwidth capacity of an instance specification is the maximum private network bandwidth limit of the corresponding instance. If the CVM private network traffic exceeds this limit, random packet loss may happen within the private network for your instances.
The availability of instance specifications may vary from region to region. Some configurations may be sold out. Please see the purchase page for the latest information.
The send/receive packets performance data noted in this document is the results of a one-time network forwarding test. See Network Performance Test for the testing method. Separate testing is needed to estimate the performance for your business.
Here is a complete list of instance families.

Standard Instance Family

Standard instances provide a balance of compute, memory, and network resources to accommodate most applications.

Standard S8

The Standard S8 instance is the latest generation of standard instances. Based on a newly optimized virtualization platform, it provides balanced, stable computing, memory, and network resources, making it the optimal choice for a wide range of applications.
The Standard S8 instance uses the new Intel Emerald Rapids processor, the latest DDR5 memory, and with default network optimization. The maximum private network packet receive and send processing capability reaches 45 million pps, with the highest private network bandwidth supporting up to 120 Gbps.

Instance Features

The new generation Tencent Cloud self-developed Star Lake two-way server is equipped with an Intel Emerald Rapids processor.
Uses the Intel Emerald Rapids processor with Turbo Boost up to 3.0 GHz.
Offers a variety of processor to memory ratios, including 1:2 and 1:4.
Supports up to 120 GB private network bandwidth, 45 million PPS, with exceptional network packet receive and send processing capability, meeting the demands for high private network transfer requirements.
The network performance of instances corresponds to their specifications. Higher specifications result in stronger network forwarding performance and higher private network bandwidth limits.
Supports the configuration to disable or enable Hyper-Threading.
Supports burst bandwidth.

Note:

Burst bandwidth: The size of private network bandwidth is related to instance specifications, each having its own maximum private network bandwidth limit. In scenarios where network traffic surges within a short duration, Tencent Cloud servers offer burst bandwidth processing capabilities. For cloud servers that meet burst conditions, exceeding of the private network bandwidth limit for a short period is allowed, catering to customer business needs.
Supports Jumbo Frames

Note:

Jumbo frames: Tencent Cloud supports sending Ethernet frames with an 8500-byte payload. Using Jumbo frames can fully maximize network performance, making it suitable for storage scenarios and other high traffic situations, thereby enhancing transmission efficiency.

Use Cases

The Standard S8 instance can be applied to the following scenarios:
Various types and sizes of enterprise-level applications.
Medium and small database systems, caching, and search clusters.

Compute clusters, memory-dependent data processing.
Scenarios with high demanding of network packet receive and send processing, such as video danmaku, live streaming, games, etc.

Instance Requirements

S8 instances can be used as monthly subscription instances as well as pay-as-you-go billing instances.
S8 instances can only be launched in a Virtual Private Cloud.
It is recommended to pair with the TencentOS Server operating system for optimal application performance of the instance.
See the specifications below to see the sizes of S8 instances available for purchase. Ensure the size of the S8 instance you choose meets the minimum CPU and memory requirements of your operating system and application.
An Instance can support up to 120 Gbps of network bandwidth, depending on the kernel version and runtime environment of the instance operating system. When PPS exceeds 10 million and the bandwidth exceeds 50 Gbps, the kernel protocol stack can cause significant network performance loss. In such cases, the bandwidth values tested by netperf may not meet expectations. The DPDK method can be used to overcome the differences in the cloud server's kernel protocol stack, getting the instance's real network performance. For testing methods, see high-throughput network performance testing.

Specification	vCPU	Memory (GB)	Network Packet Receive and Send (pps) (Outbound + Inbound)	Number of Connections	Number of Queues	Standard Bandwidth/Burst Bandwidth (Gbps) (Outbound + Inbound)	
S8.MEDIUM8	2	8	400,000	250,000	2	1.5/10	
S8.LARGE8	4	8	800,000	250,000	4	2/10	
S8.LARGE16	4	16	800,000	250,000	4	2/10	
S8.2XLARGE16	8	16	1,600,000	500,000	8	4/10	
S8.2XLARGE32	8	32	1,600,000	500,000	8	4/10	
S8.4XLARGE32	16	32	3,200,000	1,100,000	16	9/10	
S8.4XLARGE64	16	64	3,200,000	1,100,000	16	9/10	
S8.8XLARGE64	32	64	6,400,000	2,200,000	32	17/25	
S8.8XLARGE128	32	128	6,400,000	2,200,000	32	17/25	

S8.14XLARGE256	56	256	11,200,000	4,000,000	48	30	
S8.16XLARGE256	64	256	12,800,000	4,500,000	48	34	
S8.28XLARGE512	112	512	22,500,000	8,000,000	48	60	
S8.56XLARGE1024	224	1024	45,000,000	16,000,000	48	120	

Standard SA5

The Standard SA5 instance is the latest generation of standard instances. Based on a newly optimized virtualization platform, it provides balanced, stable computing, memory, and network resources, making it the optimal choice for a wide range of applications.

The Standard SA5 instance uses the new AMD EPYC™ Bergamo processor, the latest DDR5 memory, and with default network optimization. The maximum private network packet receive and send processing capability reaches 45 million pps.

Instance Features

The new generation Tencent Cloud self-developed Star Lake two-way server is equipped with an AMD EPYC™ Bergamo processor with Turbo Boost up to 3.1 GHz.

Offers a variety of processor to memory ratios, including 1:2 and 1:4.

Supports up to 45 million PPS, with exceptional network packet receive and send processing capability, meeting the demands for high private network transfer requirements.

The network performance of instances corresponds to their specifications. Higher specifications result in stronger network forwarding performance and higher private network bandwidth limits.

Supports the configuration to disable or enable Hyper-Threading.

Supports burst bandwidth.

Note:

Burst bandwidth: The size of private network bandwidth is related to instance specifications, each having its own maximum private network bandwidth limit. In scenarios where network traffic surges within a short duration, Tencent Cloud servers offer burst bandwidth processing capabilities. For cloud servers that meet burst conditions, exceeding of the private network bandwidth limit for a short period is allowed, catering to customer business needs.

Supports Jumbo Frames

Note:

Jumbo frames: Tencent Cloud supports sending Ethernet frames with an 8500-byte payload. Using Jumbo frames can fully maximize network performance, making it suitable for storage scenarios and other high traffic situations, thereby enhancing transmission efficiency.

Use Cases

The Standard SA5 instance can be applied to the following scenarios:

Various types and sizes of enterprise-level applications.
Medium and small database systems, caching, and search clusters.
Compute clusters, memory-dependent data processing.
Scenarios with high demanding of network packet receive and send processing, such as video danmaku, live streaming, games, etc.

Instance Requirements

SA5 instances can be used as monthly subscription instances as well as pay-as-you-go billing instances.
SA5 instances can only be launched in the Virtual Private Cloud.
It is recommended to pair with the TencentOS Server operating system for optimal application performance of the instance.
See the specifications below to see the sizes of SA5 instances available for purchase. Ensure the size of the SA5 instance you choose meets the minimum CPU and memory requirements of your operating system and application. In many use cases, operating systems with GUIs that consume a lot of memory and CPU resources (such as Windows) may need larger instance sizes. As your workload's memory and CPU needs increase over time, you can scale to higher configurations or choose other types of instances.
Depends on the support from the instance's operating system kernel version and runtime environment. When PPS exceeds 10 million and the bandwidth exceeds 50 Gbps, the kernel protocol stack can cause significant network performance loss. In such cases, the bandwidth values tested by netperf may not meet expectations. The DPDK method can be used to overcome the differences in the cloud server's kernel protocol stack, getting the instance's real network performance. For testing methods, see high-throughput network performance testing.

Specification	vCPU	Memory (GB)	Network Packet Receive and Send (pps) (Outbound + Inbound)	Number of Connections	Number of Queues	Standard Bandwidth/Burst Bandwidth (Gbps) (Outbound + Inbound)
SA5.MEDIUM2	2	2	250,000	250,000	2	1.5/10
SA5.MEDIUM4	2	4	250,000	250,000	2	1.5/10
SA5.LARGE8	4	8	300,000	250,000	4	1.5/10
SA5.LARGE16	4	16	300,000	250,000	4	1.5/10
SA5.2XLARGE16	8	16	700,000	250,000	8	3/10
SA5.2XLARGE32	8	32	700,000	250,000	8	3/10

SA5.4XLARGE32	16	32	1,400,000	500,000	16	5/10
SA5.4XLARGE64	16	64	1,400,000	500,000	16	5/10
SA5.8XLARGE64	32	64	2,800,000	1,000,000	32	10/25
SA5.8XLARGE128	32	128	2,800,000	1,000,000	32	10/25
SA5.12XLARGE96	48	96	4,200,000	1,500,000	48	15/25
SA5.12XLARGE192	48	192	4,200,000	1,500,000	48	15/25
SA5.16XLARGE256	64	256	5,600,000	2,000,000	48	20/25
SA5.16XLARGE288	64	288	5,600,000	2,000,000	48	20/25
SA5.32XLARGE576	128	576	11,200,000	4,000,000	48	40/-
SA5.64XLARGE1152	256	1152	22,500,000	8,000,000	48	80/-
SA5.128XLARGE2304	512	2304	45,000,000	16,000,000	48	160/-

Standard SA4

The Standard SA4 instance is based on a newly optimized virtualization platform, it provides balanced, stable computing, memory, and network resources, making it the optimal choice for a wide range of applications.
The Standard SA4 instance uses the new AMD EPYC™ Genoa processor, the latest DDR5 memory, and with default network optimization. The maximum private network packet receive and send processing capability reaches 45 million pps, with the highest private network bandwidth supporting up to 100 Gbps.

Instance Features

The new generation Tencent Cloud self-developed Star Lake two-way server is equipped with the AMD EPYC™ Genoa processor.
Uses the AMD EPYC™ Genoa processor with Turbo Boost up to 3.7GHz.
Offers a variety of processor to memory ratios, including 1:2 and 1:4
Supports up to 100 GB private network bandwidth, 45 million PPS, with exceptional network packet receive and send processing capability, meeting the demands for high private network transfer requirements.
The network performance of instances corresponds to their specifications. Higher specifications result in stronger network forwarding performance and higher private network bandwidth limits.
Supports the configuration to disable or enable Hyper-Threading.

Use Cases

The Standard SA4 instance can be applied to the following scenarios:
Various types and sizes of enterprise-level applications.
Medium and small database systems, caching, and search clusters.
Compute clusters, memory-dependent data processing.
Scenarios with high demanding of network packet receive and send processing, such as video danmaku, live streaming, games, etc.

Instance Requirements

SA4 instances can be used as monthly subscription instances as well as pay-as-you-go billing instances.
SA4 instances can only be launched in a Virtual Private Cloud.
It is recommended to pair with the TencentOS Server operating system for optimal application performance of the instance.
Refer to the specifications below to see the sizes of SA4 instances available for purchase. Ensure the size of the SA4 instance you choose meets the minimum CPU and memory requirements of your operating system and application. In many use cases, operating systems with GUIs that consume a lot of memory and CPU resources (such as Windows) may need larger instance sizes. As your workload's memory and CPU needs increase over time, you can scale to higher configurations or choose other types of instances.
An Instance can support up to 100 Gbps of network bandwidth, depending on the kernel version and runtime environment of the instance operating system. When PPS exceeds 10 million and the bandwidth exceeds 50 Gbps, the kernel protocol stack can cause significant network performance loss. In such cases, the bandwidth values tested by netperf may not meet expectations. The DPDK method can be used to overcome the differences in the cloud server's kernel protocol stack, getting the instance's real network performance. For testing methods, see high-throughput network performance testing.

Specification	vCPU	Memory (GB)	Network Packet Receive and Send (pps) (Outbound + Inbound)	Number of Connections	Number of Queues	Private Network Bandwidth Capacity (Gbps) (Outbound + Inbound)	Rema
SA4.2XLARGE16	8	16	900,000	300,000	8	2	-
SA4.2XLARGE32	8	32	900,000	300,000	8	2	-
SA4.4XLARGE32	16	32	1,800,000	600,000	16	4	-
SA4.4XLARGE64	16	64	1,800,000	600,000	16	4	-
SA4.8XLARGE64	32	64	3,700,000	1,300,000	32	8	-
SA4.8XLARGE128	32	128	3,700,000	1,300,000	32	8	-

SA4.16XLARGE128	64	128	7,500,000	2,600,000	48	17	-
SA4.16XLARGE256	64	256	7,500,000	2,600,000	48	17	-
SA4.24XLARGE192	96	192	11,200,000	4,000,000	48	25	-
SA4.48XLARGE384	192	384	22,500,000	8,000,000	48	50	-
SA4.96XLARGE768	384	768	45,000,000	16,000,000	48	100	-

Standard S6

As the latest generation of standard instances, S6 instances are powered by an all-new optimized virtualization platform, and provide balanced and stable compute, memory, and network resources. These instances are ideal for many applications.

S6 instances come with the new Intel® Xeon® processors with the latest DDR4 memory. They are network optimized by default, offering a private network throughput of up to 19 million pps and a private network bandwidth of up to 100 Gbps.

Note:

This instance type is now only available to beta users. Contact your account manager to purchase it.

Features

A new-generation Star Lake two-way server developed by Tencent Cloud, with the 3rd generation Intel® Xeon® scalable processor.

Intel® Xeon® Ice Lake processor, with a CPU clock rate of 2.7 GHz and a turbo frequency of 3.3 GHz.

Supported CPU/RAM ratios: 1:2 and 1:4.

Supports up to 100 Gbps of private network bandwidth, providing ultra-high packet throughput to meet high private network transmission requirements.

The network performance of an instance depends on its specification. A higher specification means a greater network forwarding performance and a higher private network bandwidth cap.

Supports hyper-threading.

Use Cases

Standard S6 instances are applicable to the following scenarios:

Enterprise applications of different types and sizes.

Small and medium-sized database systems, caches, and search clusters.

Computing clusters, memory-intensive data processing.

Scenarios that require sending and receiving massive network packets, such as video on-screen comments, live video broadcasting, and gaming.

Requirements

Supports pay-as-you-go.
Supports only VPC.
See below for available S6 instance specifications. Make sure that the instance specification you choose meets the minimum CPU and memory requirements of your operating system and applications. In many cases, GUI operating systems (such as Windows) that consume extensive memory and CPU resources may need higher specifications. As the needs of your workload for memory and CPU increase, you can choose higher configurations or other instance types.

Specification	vCPU	Memory (GB)	Throughput (pps) (received and sent packets)	Number of connections	Number of queues	Private network bandwidth (Gbps) (outbound and inbound)	CPU clock rate
S6.SMALL1	1	1	260,000	260,000	1	1.6	2.7GHz
S6.SMALL2	1	2	260,000	260,000	1	1.6	2.7GHz
S6.SMALL4	1	4	260,000	260,000	1	1.6	2.7GHz
S6.MEDIUM4	2	4	300,000	260,000	2	2	2.7GHz
S6.MEDIUM8	2	8	300,000	260,000	2	2	2.7GHz
S6.LARGE8	4	8	600,000	600,000	4	4	2.7GHz
S6.LARGE16	4	16	600,000	600,000	4	4	2.7GHz
S6.2XLARGE16	8	16	1,200,000	1,000,000	8	7	2.7GHz
S6.2XLARGE32	8	32	1,200,000	1,000,000	8	7	2.7GHz
S6.4XLARGE32	16	32	2,600,000	2,000,000	16	13	2.7GHz
S6.4XLARGE64	16	64	2,600,000	2,000,000	16	13	2.7GHz
S6.8XLARGE64	32	64	6,000,000	4,000,000	32	26	2.7GHz
S6.8XLARGE128	32	128	6,000,000	4,000,000	32	26	2.7GHz
S6.12XLARGE96	48	96	7,000,000	6,000,000	48	39	2.7GHz
S6.12XLARGE192	48	192	7,000,000	6,000,000	48	39	2.7GHz
S6.16XLARGE266	64	266	10,000,000	8,000,000	48	62	2.7GHz

S6.31MEDIUM216	62	216	10,000,000	6,000,000	48	60	2.7GHz
S6.16XLARGE216	64	216	10,000,000	8,000,000	48	50	2.7GHz
S6.32XLARGE432	128	432	19,000,000	16,000,000	48	100	2.7GHz

Standard SA3

SA3 instances are optimized with the latest generation of AMD EPYCTM processor. They deliver ultra-high cost performance with a variety of specifications, and provide a balance of compute, storage, and network resources, making them ideal for many applications.

SA3 instances come with the new AMD EPYCTM Milan processors with the latest DDR4 memory. They are network optimized by default, offering a private network throughput up to 19 million pps and a private network bandwidth up to 100 Gbps.

Note:

This instance type is now only available to beta users. Contact your account manager to purchase it.

Features

Provides high performance with high reliability, security, and stability based on the Star Lake servers developed by Tencent Cloud.

Comes with 2.55 GHz AMD EPYCTM Milan processor, with a turbo frequency of 3.5GHz and the latest generation of 8-channel DDR4 for stable memory computing performance.

Higher specification, SA3.58XLARGE940, offering 232 vCPUs and 940 GB memory.

A wide range of CPU/RAM ratios, such as 1:2 and 1:4.

Supports up to 100 Gbps of private network bandwidth for ultra-high packet throughput to meet extremely high private network transmission requirements.

The network performance of an instance depends on its specification. A higher specification means a greater network forwarding performance and a higher private network bandwidth cap.

Use Cases

Standard SA3 instances are applicable to the following scenarios:

Enterprise applications of different types and sizes.

Search and other computing clusters.

Video encoding and decoding, video rendering and other applications sensitive to the single-core performance.

Requirements

Supports pay-as-you-go.

Supports only VPC.

See below for available SA3 instance specifications. Make sure that the instance specification you choose meets the minimum CPU and memory requirements of your operating system and applications. In many cases, GUI operating

systems (such as Windows) that consume extensive memory and CPU resources may need higher specifications. As the needs of your workload for memory and CPU increase, you can choose higher configurations or other instance types.

With the support of appropriate operating system kernel version and runtime environment, the instance's network bandwidth can reach up to 100 Gbps. If the network throughput is larger than 10 million pps and the network bandwidth is higher than 50 Gbps, the kernel protocol stack consumes a lot of network resources. In this case, the resulted bandwidth of netperf test may not meet the expectation. To obtain the actual network performance of the instance, use the DPDK method to shield the difference caused by the kernel protocol stack.

Specification	vCPU	Memory (GB)	Throughput (pps) (received and sent packets)	Number of connections	Number of queues	Private network bandwidth (Gbps) (outbound and inbound)	CPU clock rate
SA3.SMALL1	1	1	250,000	250,000	1	1.5	2.55GH
SA3.SMALL2	1	2	250,000	250,000	1	1.5	2.55GH
SA3.SMALL4	1	4	250,000	250,000	1	1.5	2.55GH
SA3.MEDIUM4	2	4	300,000	250,000	2	1.5	2.55GH
SA3.MEDIUM8	2	8	300,000	250,000	2	1.5	2.55GH
SA3.LARGE8	4	8	500,000	250,000	4	2	2.55GH
SA3.LARGE16	4	16	500,000	250,000	4	2	2.55GH
SA3.2XLARGE16	8	32	800,000	500,000	8	4	2.55GH
SA3.2XLARGE32	8	32	800,000	500,000	8	4	2.55GH
SA3.4XLARGE32	16	32	1,500,000	1,100,000	16	7	2.55GH
SA3.4XLARGE64	16	64	1,500,000	1,100,000	16	7	2.55GH
SA3.8XLARGE64	32	64	2,500,000	2,200,000	32	14	2.55GH
SA3.8XLARGE128	32	128	2,500,000	2,200,000	32	14	2.55GH
SA3.12XLARGE96	48	96	4,000,000	3,300,000	48	21	2.55GH
SA3.12XLARGE192	48	192	4,000,000	3,300,000	48	21	2.55GH

SA3.16XLARGE128	64	128	5,200,000	4,400,000	48	28	2.55GI
SA3.16XLARGE256	64	256	5,200,000	4,400,000	48	28	2.55GI
SA3.20XLARGE160	80	160	6,500,000	5,500,000	48	35	2.55GI
SA3.20XLARGE320	80	320	6,500,000	5,500,000	48	35	2.55GI
SA3.24XLARGE192	96	192	7,800,000	6,600,000	48	42	2.55GI
SA3.24XLARGE384	96	384	7,800,000	6,600,000	48	42	2.55GI
SA3.29XLARGE216	116	216	9,500,000	8,000,000	48	50	2.55GI
SA3.29XLARGE470	116	470	9,500,000	8,000,000	48	50	2.55GI
SA3.32XLARGE256	128	256	10,000,000	8,800,000	48	56	2.55GI
SA3.32XLARGE512	128	512	10,000,000	8,800,000	48	56	2.55GI
SA3.40XLARGE320	160	320	13,000,000	11,000,000	48	69	2.55GI
SA3.40XLARGE640	160	640	13,000,000	11,000,000	48	69	2.55GI
SA3.58XLARGE432	232	432	19,000,000	16,000,000	48	100	2.55GI
SA3.58XLARGE940	232	940	19,000,000	16,000,000	48	100	2.55GI

Standard SR1

Standard SR1 instances are the latest generation of ARM standard instances. Based on an all-new optimized virtualization platform, this family provides a balance of stable computing, memory, and network resources. This is a premium choice for many applications.

Features

Use the Ampere® Altra® processor with a clock rate of 2.8 GHz and a turbo frequency of 3.0 GHz.
A wide range of processor to memory ratio, such as 1:2 and 1:4.
Support premium, SSD and enhance SSD cloud diks.
Support up to 25 Gbps of private network bandwidth for ultra-high packet throughput to meet the needs of extremely high private network transmission demand.
The network performance of an instance depends on its specification. A higher specification means a greater network forwarding performance and a higher private network bandwidth cap.

Use Cases

Standard SR1 instances are applicable to the following scenarios:

Enterprise applications of different types and sizes.
ARM simulation testing.
Embedded development and testing.
CPU-based machine learning and reasoning.

Requirements

Support purchase configuration. See the instance specifications below to make sure that the size of the SR1 instance you selected meets the minimum CPU memory requirement for your operating system and application. As your workload demands more memory and CPU over time, you can expand to a higher configuration or choose other types of instances.

Specification	vCPU	Memory (GB)	Throughput (pps) (received and sent packets)	Number of queues	Private network bandwidth (Gbps) (outbound and inbound bandwidth)	Clock rate	Notes
SR1.SMALL1	1	1	250,000	1	1.5	2.8GHz	-
SR1.SMALL2	1	2	250,000	1	1.5	2.8GHz	-
SR1.MEDIUM4	2	4	250,000	2	1.5	2.8GHz	-
SR1.MEDIUM8	2	8	250,000	2	1.5	2.8GHz	-
SR1.LARGE8	4	8	250,000	2	1.5	2.8GHz	-
SR1.LARGE16	4	16	250,000	2	1.5	2.8GHz	-
SR1.2XLARGE16	8	16	500,000	2	3	2.8GHz	-
SR1.2XLARGE32	8	32	500,000	2	3	2.8GHz	-
SR1.4XLARGE32	16	32	1,100,000	4	6	2.8GHz	-
SR1.4XLARGE64	16	64	1,100,000	4	6	2.8GHz	-
SR1.8XLARGE64	32	64	2,200,000	8	12	2.8GHz	-
SR1.8XLARGE128	32	128	2,200,000	8	12	2.8GHz	-
SR1.16XLARGE128	64	128	4,500,000	16	16	2.8GHz	-

Standard S5

Standard S5 instances are the latest generation of standard instances. Based on an all-new optimized virtualization platform, this family provides a balance of stable computing, memory, and network resources. This is a premium choice for many applications.

Features

Intel® Xeon® Cascade Lake or Intel® Xeon® Cooper Lake processor, with a CPU clock rate of 2.5 GHz and turbo frequency of 3.1 GHz.
New Intel Advanced Vector Extension (AVX-512) instruction set.
Supported CPU/RAM ratios: 1:2 and 1:4.
Supports up to 28 Gbps of private network bandwidth for ultra-high packet throughput to meet the needs of extremely high private network transmission demand.
The network performance of an instance depends on its specification. A higher specification means a greater network forwarding performance and a higher private network bandwidth cap.
Supports the configuration to disable or enable Hyper-Threading.

Application Scenarios

Standard S5 instances are applicable to the following scenarios:
Enterprise applications of different types and sizes
Small and medium-sized database systems, caches, and search clusters.
Computing clusters, memory-intensive data processing.
Scenarios that require sending and receiving massive network packets, such as video on-screen comments, live video broadcasting, and gaming.

Requirements

Supports pay-as-you-go.
Supports only VPC.
Configuration purchase is available for S5 instances. Please see the instance specifications below. Make sure that the size of the S5 instance you choose meet the minimum CPU memory requirements of your operating system and applications. In many cases, GUI operating systems (such as Windows) that consume extensive memory and CPU resources may need larger instances. As the needs of your workload for memory and CPU increase, you can expand to higher configurations or choose other instance types.

Specification	vCPU	Memory (GB)	Throughput (pps) (received and sent packets)	Number of connections	Number of queues	Private network bandwidth (Gbps) (outbound and inbound)	CPU clock rate

S5.SMALL1	1	1	250,000	250,000	1	1.5	2.5 GHz
S5.SMALL2	1	2	250,000	250,000	1	1.5	2.5 GHz
S5.SMALL4	1	4	250,000	250,000	1	1.5	2.5 GHz
S5.MEDIUM4	2	4	300,000	250,000	2	1.5	2.5 GHz
S5.MEDIUM8	2	8	300,000	250,000	2	1.5	2.5 GHz
S5.LARGE8	4	8	500,000	250,000	2	1.5	2.5 GHz
S5.LARGE16	4	16	500,000	250,000	2	1.5	2.5 GHz
S5.2XLARGE16	8	16	800,000	250,000	2	3.0	2.5 GHz
S5.2XLARGE32	8	32	800,000	250,000	2	3.0	2.5 GHz
S5.4XLARGE32	16	32	1,500,000	300,000	4	6.0	2.5 GHz
S5.4XLARGE64	16	64	1,500,000	300,000	4	6.0	2.5 GHz
S5.6XLARGE48	24	48	2,000,000	400,000	6	9.0	2.5 GHz
S5.6XLARGE96	24	96	2,000,000	400,000	6	9.0	2.5 GHz
S5.8XLARGE64	32	64	2,500,000	600,000	8	12	2.5 GHz
S5.8XLARGE128	32	128	2,500,000	600,000	8	12	2.5 GHz
S5.12XLARGE96	48	96	4,000,000	900,000	12	17.0	2.5 GHz

S5.12XLARGE192	48	192	4,000,000	900,000	12	17.0	2.5 GHz
S5.16XLARGE256	64	256	5,000,000	1,200,000	16	23.0	2.5 GHz

Standard Storage Optimized S5se

Based on an all-new optimized virtualization platform, S5se instances improve the overall cloud disk performance. They are powered by new Xeon® Cascade Lake processors with the latest DDR4 memory to provide the excellent cloud disk performance.

Features

Supports Premium Cloud Storage and SSD as the system disk.
Supports Tremendous SSD as the data disk.
Provides cloud disk storage performance of up to 1,200,000 IOPS, and read/write bandwidth of 4 GB/s.
2.5 GHz Intel® Xeon® Cascade Lake processor, with a turbo frequency 3.1 GHz and the latest generation 6-channel DDR4 for stable memory computing performance.
New Intel Advanced Vector Extension (AVX-512) instruction set.
Supported CPU/RAM ratios: 1:2 and 1:4.
Supports up to 25 Gbps of private network bandwidth for ultra-high packet throughput to meet extremely high private network transmission requirements.
The network performance of an instance depends on its specification. A higher specification means a greater network forwarding performance and a higher private network bandwidth cap.
Supports the configuration to disable or enable Hyper-Threading.

Application Scenarios

Standard Storage Optimized S5se instances are applicable to the following scenarios:
Enterprise applications of different types and sizes.
IO-intensive applications including large databases, NoSQL databases, audio/video processing services, and Elasticsearch cluster.

Requirements

Supports pay-as-you-go.
Supports only VPC.
See below for available S5se specifications:

Specification	vCPU	Memory (GB)	Throughput (pps) (received	Number of connections	Number of queues	Private network bandwidth (Gbps)	CPU clock rate

			and sent packets)			(outbound and inbound)	
S5se.LARGE16	4	16	500,000	250,000	2	1.5	2.5GH
S5se.2XLARGE16	8	16	800,000	250,000	2	3.0	2.5GH
S5se.2XLARGE32	8	32	800,000	250,000	2	3.0	2.5GH
S5se.4XLARGE32	16	32	1,500,000	300,000	4	6.0	2.5GH
S5se.4XLARGE64	16	64	1,500,000	300,000	4	6.0	2.5GH
S5se.8XLARGE64	32	64	2,500,000	600,000	8	12.0	2.5GH
S5se.8XLARGE128	32	128	2,500,000	600,000	8	12.0	2.5GH
S5se.17XLARGE316	68	316	5,000,000	1,200,000	16	25.0	2.5GH

Standard SA2

SA2 instances are optimized with the latest generation AMD EPYC™ processor, deliver ultra-high cost performance with a variety of specifications, and provide a balance of compute, memory, and network resources, making them ideal for many applications.
SA2 instances come with the new AMD EPYC™ ROME processors with the latest DDR4 memory. They are network optimized by default, offering a private network throughput up to 7,500,000 pps and support up to 25 Gbps of private network bandwidth.

Features

Provides high performance with high reliability, security, and stability based on the Star Lake servers developed by Tencent Cloud.
2.6 GHz AMD EPYC™ ROME processor, with a turbo frequency 3.3 GHz and the latest generation 8-channel DDR4 for stable memory computing performance.
Higher specification, SA2.45XLARGE464, offering 180 vCPUs and 464 GB memory.
A wide range of CPU/RAM ratios, such as 1:2 and 1:4.
Supports up to 25 Gbps of private network bandwidth for ultra-high packet throughput to meet extremely high private network transmission requirements.
The network performance of an instance depends on its specification. A higher specification means a greater network forwarding performance and a higher private network bandwidth cap.
Supports the configuration to disable or enable Hyper-Threading.

Application Scenarios

Standard SA2 instances are applicable to the following scenarios:
Enterprise applications of different types and sizes.
Search and other computing clusters.
Video encoding and decoding, video rendering and other applications sensitive to the single-core performance.

Requirements

Supports pay-as-you-go.
Supports only VPC.
See below for available SA2 instance specifications. Make sure that the instance specification you choose meet the minimum CPU memory requirements of your operating system and applications. In many cases, GUI operating systems (such as Windows) that consume extensive memory and CPU resources may need higher specifications. As the needs of your workload for memory and CPU increase, you can choose higher configurations or other instance types.

Specification	vCPU	Memory (GB)	Throughput (pps) (received and sent packets)	Number of queues	Private network bandwidth (Gbps) (outbound and inbound)	CPU clock rate	Notes
SA2.SMALL1	1	1	250,000	1	1.5	2.6 GHz	-
SA2.SMALL2	1	2	250,000	1	1.5	2.6 GHz	-
SA2.SMALL4	1	4	250,000	1	1.5	2.6 GHz	-
SA2.MEDIUM4	2	4	300,000	2	1.5	2.6 GHz	-
SA2.MEDIUM8	2	8	300,000	2	1.5	2.6 GHz	-
SA2.LARGE8	4	8	500,000	2	1.5	2.6 GHz	-
SA2.LARGE16	4	16	500,000	2	1.5	2.6 GHz	-
SA2.2XLARGE16	8	16	700,000	2	1.5	2.6 GHz	-

SA2.2XLARGE32	8	32	700,000	2	1.5	2.6 GHz	-
SA2.4XLARGE32	16	32	1,000,000	4	3.0	2.6GHz	-
SA2.4XLARGE64	16	64	1,000,000	4	3.0	2.6GHz	-
SA2.8XLARGE64	32	64	1,400,000	8	5.0	2.6GHz	-
SA2.12XLARGE96	48	96	2,100,000	12	7.0	2.6GHz	-
SA2.16XLARGE128	64	128	2,800,000	16	9.0	2.6GHz	-
SA2.20XLARGE160	80	160	3,500,000	16	12.0	2.6GHz	-
SA2.22XLARGE224	90	224	3,750,000	16	13.0	2.6GHz	-
SA2.24XLARGE192	96	192	4,200,000	16	14.0	2.6GHz	-
SA2.32XLARGE256	128	256	5,600,000	32	18.0	2.6GHz	-
SA2.40XLARGE320	160	320	7,100,000	32	23.0	2.6GHz	-
SA2.45XLARGE464	180	464	7,500,000	32	25.0	2.6GHz	-

Standard S4

S4 instances provide a balance of compute, memory, and network resources, making them ideal for many applications.
S4 instances come with new Xeon® Skylake processors with the latest DDR4 memory. They are network optimized by default, offering a private network throughput up to 6,000,000 pps, and support up to 25 Gbps of private network bandwidth.

Features

2.4 GHz Intel® Xeon® Skylake 6148 processor, offering stable computing performance.
New Intel Advanced Vector Extension (AVX-512) instruction set.
The latest generation of 6-channel DDR4 memory with a bandwidth of 2,666 MT/s.
Higher specification, S4.18XLARGE228, offering 72 vCPUs and 228 GB memory.
Supported CPU/RAM ratios: 1:2 and 1:4.
Supports up to 25 Gbps of private network bandwidth for ultra-high packet throughput to meet extremely high private network transmission requirements.
The network performance of an instance depends on its specification. A higher specification means a greater network forwarding performance and a higher private network bandwidth cap.

Application Scenarios

Standard S4 instances are applicable to the following scenarios:
Enterprise applications of different types and sizes.
Small and medium-sized database systems, caches, and search clusters.
Computing clusters, memory-intensive data processing.
Scenarios that require massive send/receive network packets, such as on-screen comments, live video broadcasting, and gaming.

Requirements

Supports pay-as-you-go.
Supports only VPC.
See below for available S4 instance specifications. Make sure that the instance specification you choose meet the minimum CPU memory requirements of your operating system and applications. In many cases, GUI operating systems (such as Windows) that consume extensive memory and CPU resources may need higher specifications. As the needs of your workload for memory and CPU increase, you can choose higher configurations or other instance types.

Specification	vCPU	Memory(GB)	Throughput (pps) (received and sent packets)	Number of queues	Private network bandwidth (Gbps) (outbound and inbound)	CPU clock rate	Notes
S4.SMALL1	1	1	250,000	1	1.5	2.4GHz	-
S4.SMALL2	1	2	250,000	1	1.5	2.4GHz	-
S4.SMALL4	1	4	250,000	1	1.5	2.4GHz	-
S4.MEDIUM4	2	4	300,000	2	1.5	2.4GHz	-
S4.MEDIUM8	2	8	300,000	2	1.5	2.4GHz	-
S4.LARGE8	4	8	500,000	2	1.5	2.4GHz	-
S4.LARGE16	4	16	500,000	2	1.5	2.4GHz	-
S4.2XLARGE16	8	16	800,000	2	3.0	2.4 GHz	-
S4.2XLARGE32	8	32	800,000	2	3.0	2.4 GHz	-

S4.4XLARGE32	16	32	1,500,000	4	6.0	2.4 GHz	-
S4.4XLARGE64	16	64	1,500,000	4	6.0	2.4 GHz	-
S4.6XLARGE48	24	48	2,000,000	6	8.0	2.4 GHz	-
S4.6XLARGE96	24	96	2,000,000	6	8.0	2.4 GHz	-
S4.8XLARGE64	32	64	2,500,000	8	11.0	2.4 GHz	-
S4.8XLARGE128	32	128	2,500,000	8	11.0	2.4 GHz	-
S4.12XLARGE96	48	96	4,000,000	12	16.0	2.4 GHz	-
S4.12XLARGE192	48	192	4,000,000	12	16.0	2.4 GHz	-
S4.16XLARGE256	64	256	5,000,000	16	22.0	2.4 GHz	-
S4.18XLARGE288	72	288	6,000,000	16	24.0	2.4 GHz	-

Standard Network Optimized SN3ne

Standard SN3ne instances are a relatively new generation of network optimized instances. This family provides a balance of computing, memory, and network resources, with outstanding network throughput, and it is a good choice for many applications.

Standard Network Optimized SN3ne instances are equipped with the latest Xeon® Skylake processors with the latest DDR4 memory, default network optimization, and private network throughput up to 6,000,000 pps, with performance is nearly 8 times greater than the Standard S3 family. It can support up to 25 Gbps of private network bandwidth, with performance 2.5 times greater than the Standard S3 family.

Features

2.5 GHz Intel Xeon® Skylake 6133 processors with stable computing performance.

The latest generation of 6-channel DDR4 memory with a memory bandwidth of 2,666 MT/s.

Larger instance size, SN3ne.18XLARGE228, offering 72 vCPUs and 228 GB of memory.

Supported CPU/RAM ratios: 1:2 and 1:4.

Supports up to 25 Gbps of private network bandwidth for ultra-high packet throughput to meet the needs of extremely high private network transmission demand.
The network performance of an instance depends on its specification. A higher specification means a greater network forwarding performance and a higher private network bandwidth cap.
Support all types of cloud disks.

Application Scenarios

Standard Network Optimized SN3ne instances are applicable to the following scenarios:
Enterprise applications of different types and sizes.
Small and medium-sized database systems, caches, and search clusters.
Computing clusters, memory-intensive data processing.
Scenarios that require sending and receiving massive network packets, such as video on-screen comments, live video broadcasting, and gaming.

Requirements

SN3ne instances are pay-as-you-go.
SN3ne instances can only be launched in a VPC.
Configuration purchase is available for SN3ne instances. Please see the instance specifications below. Make sure that the size of the SN3ne instance you choose meet the minimum CPU memory requirements of your operating system and applications. In many cases, GUI operating systems (such as Windows) that consume extensive memory and CPU resources may need larger instances. As the needs of your workload for memory and CPU increase, you can expand to higher configurations or choose other instance types.

Specification	vCPU	Memory (GB)	Throughput (pps) (received and sent packets)	Number of queues	Private network bandwidth (Gbps) (outbound and inbound)	CPU clock rate	Notes
SN3ne.SMALL2	1	2	250,000	1	1.5	2.5 GHz	-
SN3ne.MEDIUM4	2	4	300,000	2	1.5	2.5 GHz	-
SN3ne.LARGE8	4	8	500,000	2	1.5	2.5 GHz	-
SN3ne.LARGE16	4	16	500,000	2	1.5	2.5 GHz	-

SN3ne.2XLARGE16	8	16	800,000	2	3.0	2.5 GHz	-
SN3ne.2XLARGE32	8	32	800,000	2	3.0	2.5 GHz	-
SN3ne.3XLARGE24	12	24	1,000,000	3	4.0	2.5 GHz	-
SN3ne.4XLARGE32	16	32	1,500,000	4	6.0	2.5 GHz	-
SN3ne.4XLARGE64	16	64	1,500,000	4	6.0	2.5 GHz	-
SN3ne.6XLARGE48	24	48	2,000,000	6	8.0	2.5 GHz	-
SN3ne.6XLARGE96	24	96	2,000,000	6	8.0	2.5 GHz	-
SN3ne.8XLARGE64	32	64	2,500,000	8	11.0	2.5 GHz	-
SN3ne.8XLARGE128	32	128	2,500,000	8	11.0	2.5 GHz	-
SN3ne.12XLARGE96	48	96	4,000,000	12	16.0	2.5 GHz	-
SN3ne.12XLARGE192	48	192	4,000,000	12	16.0	2.5 GHz	-
SN3ne.16XLARGE128	64	128	5,000,000	16	22.0	2.5 GHz	-
SN3ne.16XLARGE256	64	256	5,000,000	16	22.0	2.5 GHz	-
SN3ne.18XLARGE288	72	288	6,000,000	16	24.0	2.5 GHz	-

Standard S3

Standard S3 instances are a relatively new generation of standard instances. This family provides a balance of computing, memory, and network resources, and it is a good choice for many applications.

Standard S3 instances are equipped with the latest Xeon® Skylake processors with the latest DDR4 memory and support for up to 10 Gbps of private network bandwidth.

Features

2.5 GHz Intel Xeon® Skylake 6133 processors with stable computing performance.
The latest generation of 6-channel DDR4 memory with a memory bandwidth of 2,666 MT/s.
Larger instance size, S3.20XLARGE320, offering 80 vCPUs and 320 GB of memory.
Supported CPU/RAM ratios: 1:2 and 1:4.
The network performance of an instance depends on its specification. A higher specification means a greater network forwarding performance and a higher private network bandwidth cap.
Support all types of cloud disks.

Application Scenarios

Standard S3 instances are applicable to the following scenarios:
Enterprise applications of different types and sizes.
Small and medium-sized database systems, caches, and search clusters.
Computing clusters, memory-intensive data processing.

Requirements

Supports pay-as-you-go.
Supports both classic network and VPC.
See below for available S3 instance specifications. Make sure that the instance specification you choose meets the minimum CPU and memory requirements of your operating system and applications. In many cases, GUI operating systems (such as Windows) that consume extensive memory and CPU resources may need higher specifications. As the needs of your workload for memory and CPU increase, you can choose higher configurations or other instance types.

Specification	vCPU	Memory (GB)	Throughput (pps) (received and sent packets)	Number of queues	Private network bandwidth (Gbps) (outbound and inbound)	CPU clock rate	Notes
S3.SMALL1	1	1	200,000	1	1.5	2.5 GHz	-
S3.SMALL2	1	2	200,000	1	1.5	2.5 GHz	-
S3.SMALL4	1	4	200,000	1	1.5	2.5	-

						GHz	
S3.MEDIUM8	2	8	250,000	2	1.5	2.5 GHz	-
S3.LARGE8	4	8	450,000	4	1.5	2.5 GHz	-
S3.LARGE16	4	16	450,000	4	1.5	2.5 GHz	-
S3.2XLARGE16	8	16	850,000	8	1.5	2.5 GHz	-
S3.2XLARGE32	8	32	850,000	8	1.5	2.5 GHz	-
S3.3XLARGE24	12	24	850,000	12	1.5	2.5 GHz	-
S3.3XLARGE48	12	48	850,000	12	1.5	2.5 GHz	-
S3.4XLARGE32	16	32	850,000	16	2.0	2.5 GHz	-
S3.4XLARGE64	16	64	850,000	16	2.0	2.5 GHz	-
S3.6XLARGE48	24	48	850,000	16	3.0	2.5 GHz	-
S3.6XLARGE96	24	96	850,000	16	3.0	2.5 GHz	-
S3.8XLARGE64	32	64	850,000	16	4.0	2.5 GHz	-
S3.8XLARGE128	32	128	850,000	16	4.0	2.5 GHz	-
S3.12XLARGE96	48	96	850,000	16	6.0	2.5 GHz	-
S3.12XLARGE192	48	192	850,000	16	6.0	2.5 GHz	-
S3.16XLARGE256	64	256	850,000	16	8.0	2.5	-

						GHz	
S3.20XLARGE320	80	320	850,000	16	10.0	2.5 GHz	-

Standard SA1

A Standard SA1 instance uses the AMD EPYC™ processor. This ultra-high cost performance instance family supports a variety of specifications, ensuring cost efficiency. Standard SA1 instances provides a balance of computing, memory, and network resources, and it is a good choice for many applications.

Features

2.0 GHz AMD EPYC™ 7551 processors with a Turbo Boost clock of 2.55 GHz, offering stable computing performance.
The latest generation of 8-channel DDR4 memory with a memory bandwidth of 2,666 MT/s.
Supported CPU/RAM ratios: 1:2 and 1:4.
The network performance of an instance depends on its specification. A higher specification means a greater network forwarding performance and a higher private network bandwidth cap.
Support all types of cloud disks.

Application Scenarios

Standard SA1 instances are applicable to the following scenarios:
Enterprise applications of different types and sizes.
Small and medium-sized database systems, caches, and search clusters.
Computing clusters, memory-intensive data processing.

Requirements

Supports pay-as-you-go.
Supports both classic network and VPC.
See below for available SA1 instance specifications. As the needs of your workload for memory and CPU increase, you can choose higher configurations or other instance types.

Specification	vCPU	Memory (GB)	Throughput (pps) (received and sent packets)	Number of queues	Private network bandwidth (Gbps) (outbound and inbound)	CPU clock rate	Notes
SA1.SMALL1	1	1	-	1	1.5	2.0 GHz	-

SA1.SMALL2	1	2	-	1	1.5	2.0 GHz	-
SA1.SMALL4	1	4	-	1	1.5	2.0 GHz	-
SA1.MEDIUM4	2	4	-	2	1.5	2.0 GHz	-
SA1.MEDIUM8	2	8	-	2	1.5	2.0 GHz	-
SA1.LARGE8	4	8	-	4	1.5	2.0 GHz	-
SA1.LARGE16	4	16	-	4	1.5	2.0 GHz	-
SA1.2XLARGE16	8	16	-	8	1.5	2.0 GHz	-
SA1.2XLARGE32	8	32	-	8	1.5	2.0 GHz	-
SA1.4XLARGE32	16	32	-	16	1.5	2.0 GHz	-
SA1.4XLARGE64	16	64	-	16	1.5	2.0 GHz	-

Note:

In the "SA1 Instances Specification List", "-" in the Send/Receive Packets column indicates that, for this specification, the network packet rate and server CPU performance are not guaranteed. If you have strict requirement on the performance, please choose models with performance guarantee.

Standard Network Optimized S2ne

Standard Network Optimized S2ne instances are the best choice for applications that require sending and receiving massive network packets. They support sending and receiving up to millions of network packets per second. This family is recommended for businesses with high network PPS requirements, such as large game servers, videos, and live streaming.

Note:

This instance type is currently only available to beta users. Contact your account manager if you wish to purchase these instance type.

Features

Intel Xeon E5-2680 Broadwell (v4) processors with base frequency of 2.4 GHz and DDR4 memory, offering stable computing performance.
Up to 48 cores and 192 GB are available for sale.
Supported CPU/RAM ratios: 1:2 and 1:4.
The network performance of an instance depends on its specification. A higher specification means a greater network forwarding performance and a higher private network bandwidth cap.
Support all types of cloud disks.

Application Scenarios

Scenarios that require sending and receiving massive network packets, such as game services, video services, and financial analysis.
Enterprise applications of different types and sizes.

Requirements

Supports pay-as-you-go.
Supports only VPC.
See below for available S2ne specifications:

Specification	vCPU	Memory (GB)	Throughput (pps) (received and sent packets)	Number of queues	Private network bandwidth (Gbps) (outbound and inbound)	CPU clock rate	Notes
S2ne.SMALL2	1	2	120,000	1	1.5	2.4 GHz	-
S2ne.MEDIUM4	2	4	150,000	2	1.5	2.4 GHz	-
S2ne.LARGE8	4	8	300,000	2	1.5	2.4 GHz	-
S2ne.LARGE16	4	16	300,000	2	1.5	2.4 GHz	-
S2ne.2XLARGE16	8	16	600,000	2	2.0	2.4 GHz	-
S2ne.2XLARGE32	8	32	600,000	2	2.0	2.4 GHz	-

S2ne.3XLARGE24	12	24	900,000	3	2.5	2.4 GHz	-
S2ne.3XLARGE48	12	48	900,000	3	2.5	2.4 GHz	-
S2ne.4XLARGE32	16	32	1,200,000	4	3.5	2.4 GHz	-
S2ne.4XLARGE64	16	64	1,200,000	4	3.5	2.4 GHz	-
S2ne.6XLARGE48	24	48	1,800,000	6	5.0	2.4 GHz	-
S2ne.6XLARGE96	24	96	1,800,000	6	5.0	2.4 GHz	-
S2ne.8XLARGE64	32	64	2,400,000	8	6.5	2.4 GHz	-
S2ne.8XLARGE128	32	128	2,400,000	8	6.5	2.4 GHz	-
S2ne.12XLARGE192	48	192	3,600,000	12	9.5	2.4 GHz	-

Standard S2

Standard S2 instances are a relatively new generation of instances. This family provides a balance of computing, memory, and network resources, and it is a good choice for many applications.
Standard S2 instances are equipped with Intel® Xeon® Broadwell processors with DDR4 memory.

Features

Intel Xeon E5-2680 Broadwell (v4) processors with base frequency of 2.4 GHz and DDR4 memory.
Up to 56 cores and 224 GB are available for sale.
Supported CPU/RAM ratios: 1:2 and 1:4.
Balance of computing, memory, and network resources.

Application Scenarios

This family is used for small and mid-size databases, data processing tasks that require additional memory and cache fleets, and for running backend servers for SAP, Microsoft SharePoint, cluster computing and other enterprise applications.

Requirements

S2 instances are pay-as-you-go, and can also be used as production instances of standard host HS20 in CDHs.
S2 instances can be launched in classic network and VPCs.
See below for available S2 instance specifications. Make sure that the instance specification you choose meets the minimum CPU and memory requirements of your operating system and applications. In many cases, GUI operating systems (such as Windows) that consume extensive memory and CPU resources may need higher specifications. As the needs of your workload for memory and CPU increase, you can choose higher configurations or other instance types.

Specification	vCPU	Memory (GB)	Throughput (pps) (received and sent packets)	Number of queues	Private network bandwidth (Gbps) (outbound and inbound)	CPU clock rate	Notes
S2.SMALL1	1	1	200,000	1	1.5	2.4 GHz	-
S2.SMALL2	1	2	200,000	1	1.5	2.4 GHz	-
S2.SMALL4	1	4	200,000	1	1.5	2.4 GHz	-
S2.MEDIUM2	2	2	250,000	2	1.5	2.4 GHz	-
S2.MEDIUM4	2	4	250,000	2	1.5	2.4 GHz	-
S2.MEDIUM8	2	8	250,000	2	1.5	2.4 GHz	-
S2.LARGE8	4	8	450,000	4	1.5	2.4 GHz	-
S2.LARGE16	4	16	450,000	4	1.5	2.4 GHz	-
S2.2XLARGE16	8	16	500,000	8	1.5	2.4 GHz	-
S2.2XLARGE32	8	32	500,000	8	1.5	2.4 GHz	-

S2.3XLARGE24	12	24	500,000	8	2.5	2.4 GHz	-
S2.3XLARGE48	12	48	500,000	8	2.5	2.4 GHz	-
S2.4XLARGE32	16	32	500,000	8	3.0	2.4 GHz	-
S2.4XLARGE64	16	64	500,000	8	3.0	2.4 GHz	-
S2.6XLARGE48	24	48	700,000	8	4.5	2.4 GHz	-
S2.6XLARGE96	24	96	700,000	8	4.5	2.4 GHz	-
S2.8XLARGE64	32	64	700,000	8	6.0	2.4 GHz	-
S2.8XLARGE128	32	128	700,000	8	6.0	2.4 GHz	-
S2.14XLARGE224	56	224	700,000	8	10.0	2.4 GHz	-

Standard S1

Standard S1 instances offer a range of CPUs from low to high core count. They deliver significant savings with flexible configuration options to meet diverse needs. They also offer a range of data disk options, including local disks and SSD cloud disks. (Options may vary depending on hardware specifications).

Features

Standard S1 instances have the following features:
Offer a range of CPUs from low to high core count for flexible configuration options.
Intel Xeon CPUs and DDR3 memories.
Support storage options of local disks and SSD cloud disks.
Balance of computing, memory, and network resources.

Application Scenarios

Standard S1 instances are suitable for large, medium, and small-sized applications and databases.

Requirements

S1 instances are pay-as-you-go, and can also be used as production instances of standard host in CDHs.
S1 instances can be launched in classic network and VPCs.
"-" indicates that, for this specification, the network packet rate and server CPU performance are not guaranteed. If you have strict requirement on the performance, please choose models with performance guarantee.
See below for available S1 instance specifications. Make sure that the instance specification you choose meets the minimum CPU and memory requirements of your operating system and applications. In many cases, GUI operating systems (such as Windows) that consume extensive memory and CPU resources may need higher specifications. As the needs of your workload for memory and CPU increase, you can choose higher configurations or other instance types.

Specification	vCPU	Memory (GB)	Throughput (pps) (received and sent packets)	Number of queues	Private network bandwidth (Gbps) (outbound and inbound)	CPU clock rate	Notes
S1.SMALL1	1	1	-	1	1.5	-	-
S1.SMALL2	1	2	-	1	1.5	-	-
S1.SMALL4	1	4	-	1	1.5	-	-
S1.MEDIUM2	2	2	-	2	1.5	-	-
S1.MEDIUM4	2	4	-	2	1.5	-	-
S1.MEDIUM8	2	8	-	2	1.5	-	-
S1.MEDIUM12	2	12	-	2	1.5	-	-
S1.LARGE4	4	4	-	4	1.5	-	-
S1.LARGE8	4	8	-	4	1.5	-	-
S1.LARGE16	4	16	-	4	1.5	-	-
S1.2XLARGE8	8	8	-	8	2.0	-	-
S1.2XLARGE16	8	16	-	8	2.0	-	-
S1.2XLARGE32	8	32	-	8	2.0	-	-
S1.3XLARGE24	12	24	-	8	2.5	-	-
S1.3XLARGE48	12	48	-	8	2.5	-	-

S1.4XLARGE16	16	16	-	8	3.5	-	-
S1.4XLARGE32	16	32	-	8	3.5	-	-
S1.4XLARGE48	16	48	-	8	3.5	-	-
S1.4XLARGE64	16	64	-	8	3.5	-	-
S1.6XLARGE48	24	48	-	8	5.0	-	-
S1.8XLARGE64	32	64	-	8	7.0	-	-
S1.12XLARGE96	48	96	-	8	10.0	-	-

Note:
In the S1 instance list, "-" indicates that, for this specification, the send/receive packets and CPU clock rate are not guaranteed. If you have a strict performance requirement, please choose models with performance guarantee.

Memory Optimized Instance Family

This family features a large memory. It is well suited for memory-intensive operations, searches, and applications, such as high-performance databases and distributed in-memory caching.

MEM-optimized M8

The MEM-optimized M8 instance is the latest generation of memory-optimized instances. Based on a newly optimized virtualization platform, it is designed to deliver fast and stable performance for workloads processing large datasets in memory. It is the optimal choice for high memory computing applications.

Instance Features

The new generation of Tencent Cloud self-developed Star Lake two-way server is equipped with the Intel Emerald Rapids processor.
Offers a variety of processor to memory ratios, such as 1:8.
Purchases the same-size memory at the lowest price.
The network performance of instances corresponds to their specifications. Higher specifications result in stronger network forwarding performance and higher private network bandwidth limits.
Supports the configuration to disable or enable Hyper-Threading.
Supports burst bandwidth.

Note:
Burst bandwidth: The size of private network bandwidth is related to instance specifications, each having its own maximum private network bandwidth limit. In scenarios where network traffic surges within a short duration, Tencent

Cloud servers offer burst bandwidth processing capabilities. For cloud servers that meet burst conditions, exceeding of the private network bandwidth limit for a short period is allowed, catering to customer business needs.
Supports Jumbo Frames

Note:

Jumbo frames: Tencent Cloud supports sending Ethernet frames with an 8500-byte payload. Using Jumbo frames can fully maximize network performance, making it suitable for storage scenarios and other high traffic situations, thereby enhancing transmission efficiency.

Use Cases

The MEM-optimized M8 instance is highly suitable for the following scenarios:
High-performance databases, distributed memory caching, and other applications that require extensive memory operations, searches, and computations.
Users who build their own Hadoop clusters or use Redis for gene computing purposes.
Scenarios with high demanding of network packet receive and send processing, such as video danmaku, live streaming, games, etc.

Instance Requirements

M8 instances can be used as monthly subscription instances as well as pay-as-you-go billing instances.
M8 instances can only be launched in a Virtual Private Cloud.
It is recommended to pair with the TencentOS Server operating system for optimal application performance of the instance.
Refer to the specifications below to see the sizes of M8 instances available for purchase. Ensure the size of the M8 instance you choose meets the minimum CPU and memory requirements of your operating system and application. In many use cases, operating systems with GUIs that consume a lot of memory and CPU resources (such as Windows) may need larger instance sizes. As your workload's memory and CPU needs increase over time, you can scale to higher configurations or choose other types of instances.
Depends on the support from the instance's operating system kernel version and runtime environment. When PPS exceeds 10 million and the bandwidth exceeds 50 Gbps, the kernel protocol stack can cause significant network performance loss. In such cases, the bandwidth values tested by netperf may not meet expectations. The DPDK method can be used to overcome the differences in the cloud server's kernel protocol stack, getting the instance's real network performance. For testing methods, see high-throughput network performance testing.

Specification	vCPU	Memory (GB)	Network Packet Receive and Send (pps) (Outbound + Inbound)	Number of Connections	Number of Queues	Standard Bandwidth/Burst Bandwidth (Gbps) (Outbound + Inbound)	C B S S B ((

M8.MEDIUM16	2	16	400,000	250,000	2	1.5/10	2
M8.LARGE32	4	32	800,000	250,000	4	2/10	2
M8.2XLARGE64	8	64	1,400,000	500,000	8	4/10	4
M8.4XLARGE128	16	128	3,200,000	1,100,000	16	9/10	6
M8.8XLARGE256	32	256	6,400,000	2,200,000	32	17/25	1
M8.16XLARGE512	64	512	12,800,000	4,500,000	48	34	2

MEM-optimized MA5

The MEM-optimized MA5 instance is the latest generation of memory-optimized instances. Based on a newly optimized virtualization platform, it is designed to deliver fast and stable performance for workloads processing large datasets in memory. It is the optimal choice for high memory computing applications.

Instance Features

The new generation of Tencent Cloud self-developed Star Lake two-way server is equipped with the AMD EPYC™ Bergamo processor.
Offers a variety of processor to memory ratios, such as 1:8.
Purchases the same-size memory at the lowest price.
The network performance of instances corresponds to their specifications. Higher specifications result in stronger network forwarding performance and higher private network bandwidth limits.
Supports the configuration to disable or enable Hyper-Threading.
Supports burst bandwidth.

Note:

Burst bandwidth: The size of private network bandwidth is related to instance specifications, each having its own maximum private network bandwidth limit. In scenarios where network traffic surges within a short duration, Tencent Cloud servers offer burst bandwidth processing capabilities. For cloud servers that meet burst conditions, exceeding of the private network bandwidth limit for a short period is allowed, catering to customer business needs.
Supports Jumbo Frames

Note:

Jumbo frames: Tencent Cloud supports sending Ethernet frames with an 8500-byte payload. Using Jumbo frames can fully maximize network performance, making it suitable for storage scenarios and other high traffic situations, thereby enhancing transmission efficiency.

Use Cases

The MEM-optimized MA5 instance is highly suitable for the following scenarios:

High-performance databases, distributed memory caching, and other applications that require extensive memory operations, searches, and computations.
Users who build their own Hadoop clusters or use Redis for gene computing purposes.
Scenarios with high demanding of network packet receive and send processing, such as video danmaku, live streaming, games, etc.

Instance Requirements

MA5 instances can be used as monthly subscription instances as well as pay-as-you-go billing instances.
MA5 instances can only be launched in a Virtual Private Cloud.
It is recommended to pair with the TencentOS Server operating system for optimal application performance of the instance.
Refer to the specifications below to see the sizes of MA5 instances available for purchase. Ensure the size of the MA5 instance you choose meets the minimum CPU and memory requirements of your operating system and application. In many use cases, operating systems with GUIs that consume a lot of memory and CPU resources (such as Windows) may need larger instance sizes. As your workload's memory and CPU needs increase over time, you can scale to higher configurations or choose other types of instances.
Dependent on the support of the operating system kernel version and the runtime environment.

Specification	vCPU	Memory (GB)	Network Packet Receive and Send (pps) (Outbound + Inbound)	Number of Connections	Number of Queues	Standard Bandwidth/Burst Bandwidth (Gbps) (Outbound + Inbound)	
MA5.LARGE32	4	32	300,000	250,000	4	1.5/10	
MA5.2XLARGE64	8	64	700,000	250,000	8	3/10	
MA5.4XLARGE128	16	128	1,400,000	500,000	16	5/10	
MA5.8XLARGE256	32	256	2,800,000	1,000,000	32	10/25	
MA5.16XLARGE512	64	512	5,600,000	2,000,000	48	20/25	

MEM-optimized MA4

The MEM-optimized MA4 instance is the latest generation of memory-optimized instances. Based on a newly optimized virtualization platform, it is designed to deliver fast and stable performance for workloads processing large datasets in memory. It is the optimal choice for high memory computing applications.

Instance Features

The new generation of Tencent Cloud self-developed Star Lake two-way server is equipped with the AMD EPYC™ Genoa processor.
Offers a variety of processor to memory ratios, such as 1:8.
Purchases the same-size memory at the lowest price.
The network performance of instances corresponds to their specifications. Higher specifications result in stronger network forwarding performance and higher private network bandwidth limits.
Supports the configuration to disable or enable Hyper-Threading.

Use Cases

The MEM-optimized MA4 instance is highly suitable for the following scenarios:
High-performance databases, distributed memory caching, and other applications that require extensive memory operations, searches, and computations.
Users who build their own Hadoop clusters or use Redis for gene computing purposes.
Scenarios with high demanding of network packet receive and send processing, such as video danmaku, live streaming, games, etc.

Instance Requirements

MA4 instances can be used as monthly subscription instances as well as pay-as-you-go billing instances.
MA4 instances can only be launched in a Virtual Private Cloud.
It is recommended to pair with the TencentOS Server operating system for optimal application performance of the instance.
Refer to the specifications below to see the sizes of MA4 instances available for purchase. Ensure the size of the MA4 instance you choose meets the minimum CPU and memory requirements of your operating system and application. In many use cases, operating systems with GUIs that consume a lot of memory and CPU resources (such as Windows) may need larger instance sizes. As your workload's memory and CPU needs increase over time, you can scale to higher configurations or choose other types of instances.
Depends on the support from the instance's operating system kernel version and runtime environment. When PPS exceeds 10 million and the bandwidth exceeds 50 Gbps, the kernel protocol stack can cause significant network performance loss. In such cases, the bandwidth values tested by netperf may not meet expectations. The DPDK method can be used to overcome the differences in the cloud server's kernel protocol stack, getting the instance's real network performance. For testing methods, see high-throughput network performance testing.

Specification	vCPU	Memory (GB)	Throughput (pps) (received and sent packets)	Number of connections	Number of queues	Private network bandwidth (Gbps) (outbound and inbound)	Cloud Block Storag Stand Bandv (Gbps

MA4.LARGE32	4	32	400,000	250,000	4	1.5	2
MA4.2XLARGE64	8	64	900,000	300,000	8	2	2.5
MA4.4XLARGE128	16	128	1,800,000	600,000	16	4	3
MA4.8XLARGE256	32	256	3,700,000	1,300,000	32	8	4
MA4.16XLARGE512	64	512	7,500,000	2,600,000	48	17	6

MEM-optimized MA3

The MEM-optimized MA3 instance is the latest generation of memory-optimized instances. Based on a newly optimized virtualization platform, it is designed to deliver fast and stable performance for workloads processing large datasets in memory. It is the optimal choice for high memory computing applications.

Instance Features

The new generation of Tencent Cloud self-developed Star Lake two-way server is equipped with the AMD EPYC™ Milan processor.
Offers a variety of processor to memory ratios, such as 1:8.
Purchases the same-size memory at the lowest price.
The network performance of instances corresponds to their specifications. Higher specifications result in stronger network forwarding performance and higher private network bandwidth limits.
Supports the configuration to disable or enable Hyper-Threading.

Use Cases

The MEM-optimized MA3 instance is highly suitable for the following scenarios:
High-performance databases, distributed memory caching, and other applications that require extensive memory operations, searches, and computations.
Users who build their own Hadoop clusters or use Redis for gene computing purposes.
Scenarios with high demanding of network packet receive and send processing, such as video danmaku, live streaming, games, etc.

Instance Requirements

MA3 instances can be used as monthly subscription instances as well as pay-as-you-go billing instances.
MA3 instances can only be launched in a Virtual Private Cloud.
It is recommended to pair with the TencentOS Server operating system for optimal application performance of the instance.
Refer to the specifications below to see the sizes of MA3 instances available for purchase. Ensure the size of the MA3 instance you choose meets the minimum CPU and memory requirements of your operating system and application. In many use cases, operating systems with GUIs that consume a lot of memory and CPU resources (such as Windows)

may need larger instance sizes. As your workload's memory and CPU needs increase over time, you can scale to higher configurations or choose other types of instances.

Specification	vCPU	Memory (GB)	Throughput (pps) (received and sent packets)	Number of queues	Private network bandwidth (Gbps) (outbound and inbound)	CPU clock rate	Notes
MA3.SMALL8	1	8	250,000	1	1.5	2.55GHz	-
MA3.MEDIUM16	2	16	300,000	2	1.5	2.55GHz	-
MA3.LARGE32	4	32	500,000	4	2	2.55GHz	-
MA3.2XLARGE64	8	64	800,000	8	4	2.55GHz	-
MA3.4XLARGE128	16	128	1,500,000	16	7	2.55GHz	-
MA3.8XLARGE256	32	256	2,500,000	32	14	2.55GHz	-

Memory Optimized M6

As the latest generation of memory optimized instances, M6 instances are powered by an all-new optimized virtualization platform and designed to deliver fast and stable performance for workloads that process large data sets in memory. They are ideal for applications that require high in-memory computing performance.

Features

A new-generation Star Lake two-way server developed by Tencent Cloud, with the 3rd generation Intel® Xeon® scalable processor.
Intel® Xeon® Ice Lake processor, with a CPU clock rate of 2.7 GHz and a turbo frequency of 3.3 GHz.
Intel® Optane™ persistent memory for much better cloud database performance and significantly improved memory capacity and write performance.
CPU/RAM ratio 1:8.
Lowest price per GB of memory among instance types.
The network performance of an instance depends on its specification. A higher specification means a greater network forwarding performance and a higher private network bandwidth cap.

Use Cases

These instances are ideal for:

Applications that require memory-intensive operations, searches, and computations, such as high-performance databases and distributed in-memory caching.
External Hadoop clusters or Redis for fields such as computational genomics.
Scenarios that require sending and receiving massive network packets, such as video on-screen comments, live video broadcasting, and gaming.

Requirements

Supports pay-as-you-go.
Supports only VPC.
See below for available M6 specifications:

Specification	vCPU	Memory (GB)	Throughput (pps) (received and sent packets)	Number of connections	Number of queues	Private network bandwidth (Gbps) (outbound and inbound)	CPU clock rate
M6.SMALL8	1	8	250,000	250,000	1	1.5	2.7 GHz
M6.MEDIUM16	2	16	300,000	250,000	2	2	2.7 GHz
M6.LARGE32	4	32	600,000	500,000	4	4	2.7 GHz
M6.2XLARGE64	8	64	1,200,000	1,000,000	8	7	2.7 GHz
M6.4XLARGE128	16	128	2,500,000	2,000,000	16	13	2.7 GHz
M6.8XLARGE256	32	256	5,000,000	4,000,000	32	26	2.7 GHz
M6.16XLARGE512	64	512	10,000,000	8,000,000	48	52	2.7 GHz
M6.31MEDIUM470	62	470	10,000,000	8,000,000	48	50	2.7 GHz
M6.31XLARGE940	124	940	19,000,000	16,000,000	48	100	2.7 GHz

Security-enhanced Memory Optimized M6ce

As the latest generation of memory optimized instances, M6ce instances are powered by an all-new optimized virtualization platform and designed to deliver fast and stable performance for workloads that process large data sets in memory. They are ideal for applications that require high in-memory computing performance.

Note:

This instance type is now only available to beta users. Contact your account manager to purchase it.

Features

A new-generation Star Lake two-way server developed by Tencent Cloud, with the 3rd generation Intel® Xeon® scalable processor.
Intel® Xeon® Ice Lake processor, with a CPU clock rate of 2.7 GHz and a turbo frequency of 3.3 GHz.
Intel® SGX encrypted computing, with a single instance containing up to 428 GB encrypted memory, ensuring the confidentiality and security of critical data and codes.
The processor-to-memory ratio is 1: 8. The encrypted memory accounts for about 50% of the memory.
Lowest price per GB of memory among instance types.
The network performance of an instance depends on its specification. A higher specification means a greater network forwarding performance and a higher private network bandwidth cap.

Use Cases

These instances are ideal for:
Data sharing and computing, protecting confidential data shared between different users or vendors.
Blockchain applications, enhancing the privacy and security of transactions and key storage, etc.
Scenarios with high security and reliability requirements, such as finance, government agencies, medical care, etc.
Confidential computing scenarios and data encryption applications.

Requirements

Supports pay-as-you-go.
Supports only VPC.
See below for available M6ce specifications:

Note:

If you use the key bound to hardware to encrypt data (such as SGX Sealing) in Intel SGX Enclave, you cannot decrypt the corresponding data once the host changes in the following scenarios. It is recommended that you back up data in the application layer to ensure the application reliability. Scenarios such as host failure, adjustment of instance types, no charges when shutdown, modification of instance placement group, etc., may lead to host changes.
For more information on how to use security-enhanced memory optimized M6ce instances, see Building an SGX Confidential Computing Environment.

Specification	vCPU	Memory (GB)	With	Throughput (pps)	Number of connections	Number of	Privat netwo

			encrypted memory (GB)	(received and sent packets)		queues	bandv (Gbps (outbc and inbou
M6ce.MEDIUM16	2	16	8	300,000	250,000	2	2
M6ce.LARGE32	4	32	16	600,000	500,000	4	4
M6ce.2XLARGE64	8	64	32	1,200,000	1,000,000	8	7
M6ce.4XLARGE128	16	128	64	2,500,000	2,000,000	16	13
M6ce.8XLARGE256	32	256	128	5,000,000	4,000,000	32	26
M6ce.31MEDIUM428	62	428	214	10,000,000	8,000,000	48	50
M6ce.31XLARGE856	124	856	428	19,000,000	8,000,000	48	100

Memory Optimized M5

As the latest generation of memory optimized instances, M5 instances are powered by an all-new optimized virtualization platform and designed to deliver fast and stable performance for workloads that process large data sets in memory. They are ideal for applications that require high in-memory computing performance.

Features

Intel® Xeon® Cascade Lake or Intel® Xeon® Cooper Lake processor, with a CPU clock rate of 2.5 GHz and turbo frequency of 3.1 GHz.
New Intel Advanced Vector Extension (AVX-512) instruction set.
CPU/RAM ratio 1:8.
Lowest price per GB of memory among instance types.
The network performance of an instance depends on its specification. A higher specification means a greater network forwarding performance and a higher private network bandwidth cap.
Supports the configuration to disable or enable Hyper-Threading.

Application Scenarios

These instances are ideal for:
Applications that require memory-intensive operations, searches, and computations, such as high-performance databases and distributed in-memory caching.
Self-built Hadoop clusters or Redis for fields such as computational genomics.
Scenarios that require massive send/receive packets, such as on-screen comments, live video broadcasting, and gaming.

Requirements

Supports pay-as-you-go.

Supports only VPC.

See below for available M5 specifications:

Specification	vCPU	Memory (GB)	Throughput (pps) (received and sent packets)	Number of connections	Number of queues	Private network bandwidth (Gbps) (outbound and inbound)	CPU clock rate
M5.SMALL8	1	8	250,000	250,000	1	1.5	2.5GHz
M5.MEDIUM16	2	16	300,000	250,000	2	1.5	2.5GHz
M5.LARGE32	4	32	500,000	250,000	2	1.5	2.5GHz
M5.2XLARGE64	8	64	800,000	250,000	2	3.0	2.5GHz
M5.3XLARGE96	12	96	1,000,000	250,000	3	5.0	2.5GHz
M5.4XLARGE128	16	128	1,500,000	300,000	4	6.0	2.5GHz
M5.8XLARGE256	32	256	2,500,000	600,000	8	12.0	2.5GHz
M5.16XLARGE512	64	512	5,000,000	1,200,000	16	23.0	2.5GHz

Memory Optimized MA2

As the latest generation of memory optimized instances, MA2 instances are powered by an all-new optimized virtualization platform and designed to deliver fast and stable performance for workloads that process large data sets in memory. They are ideal for applications that require high in-memory computing performance.

Features

Provides high performance with high reliability, security, and stability based on the Star Lake servers developed by Tencent Cloud.

Comes with AMD EPYC™ ROME processor, with a CPU clock rate of 2.6 GHz and turbo frequency of 3.3 GHz.

CPU/RAM ratio 1:8.

Lowest price per GB of memory among instance types.

The network performance of an instance depends on its specification. A higher specification means a greater network forwarding performance and a higher private network bandwidth cap.

Use Cases

Memory Optimized MA2 instances are applicable to the following scenarios:
Applications that require extensive memory operations, searches, and computing, such as high-performance databases and distributed memory caching.
External Hadoop clusters or Redis for fields such as computational genomics.
Scenarios that require sending and receiving massive network packets, such as video on-screen comments, live video broadcasting, and gaming.

Requirements

Supports pay-as-you-go.
Supports only VPC.
See below for available MA2 specifications:

Specification	vCPU	Memory (GB)	Throughput (pps) (received and sent packets)	Number of connections	Number of queues	Private network bandwidth (Gbps) (outbound and inbound)	CPU clock rate
MA2.SMALL8	1	8	250,000	250,000	1	1.5	2.6GH
MA2.LARGE32	4	32	500,000	250,000	4	1.5	2.6GH
MA2.2XLARGE64	8	64	700,000	250,000	8	1.5	2.6GH
MA2. 2XLARGE384	8	384	700,000	250,000	8	1.5	2.6GH
MA2.4 XLARGE128	16	128	1,000,000	300,000	16	3	2.6GH
MA2.8 XLARGE256	32	256	1,400,000	700,000	32	5	2.6GH
MA2.12XLARGE384	48	384	2,100,000	1,000,000	48	7	2.6GH

Memory Optimized M4

M4 instances are designed to deliver fast performance for workloads that process large data sets in memory. They are ideal for applications that require high in-memory computing performance.
As a new generation of memory optimized instances, M4 instances come with new Xeon® Skylake processors with the latest DDR4 memory. They are network optimized by default, offering a private network throughput up to 6,000,000 pps, and support for ultra-high private network bandwidth.

Application Scenarios

These instances are an ideal choice for:
Applications that require extensive memory operations, searches, and computations, such as high-performance databases and distributed memory caching.
Users who build their own Hadoop clusters or Redis to use in fields such as computational genomics.
Scenarios that require sending and receiving massive network packets, such as video on-screen comments, live video broadcasting, and gaming.

Features

2.4 GHz Intel® Xeon® Skylake 6148 processor, offering stable computing performance.
New Intel Advanced Vector Extension (AVX-512) instruction set.
The latest generation of 6-channel DDR4 memory with a bandwidth of 2,666 MT/s.
Higher specification, M4.16XLARGE512, offering 64 vCPUs and 512 GB memory.
Supported CPU/RAM ratios: 1:8 and 1:12.
Priced lower than any other memory-optimized instances with the same memory.
The network performance of an instance depends on its specification. A higher specification means a greater network forwarding performance and a higher private network bandwidth cap.

Requirements

Supports pay-as-you-go.
Supports only VPC.
See below for available M4 specifications:

Specification	vCPU	Memory (GB)	Throughput (pps) (received and sent packets)	Number of queues	Private network bandwidth (Gbps) (outbound and inbound)	CPU clock rate	Notes
M4.SMALL8	1	8	250,000	1	1.5	2.4GHz	-
M4.MEDIUM16	2	16	300,000	2	1.5	2.4GHz	-
M4.LARGE32	4	32	500,000	2	1.5	2.4GHz	-
M4.2XLARGE64	8	64	800,000	2	3.0	2.4GHz	-
M4.3XLARGE96	12	96	1,000,000	3	4.0	2.4GHz	-
M4.3XLARGE144	12	144	1,000,000	3	4.0	2.4GHz	-
M4.4XLARGE128	16	128	1,500,000	4	6.0	2.4GHz	-

M4.4XLARGE192	16	192	1,500,000	4	6.0	2.4GHz	-
M4.8XLARGE256	32	256	2,500,000	8	11.0	2.4GHz	-
M4.8XLARGE384	32	384	2,500,000	8	11.0	2.4GHz	-
M4.16XLARGE512	64	512	5,000,000	16	22.0	2.4GHz	-
M4.18XLARGE648	72	648	6,000,000	16	24.0	2.4GHz	-

Memory Optimized M3

M3 instances are designed to deliver fast performance for workloads that process large data sets in memory. They are ideal for applications that require high in-memory computing performance.
M3 instances are powered by new Intel® Xeon® Skylake processors with the latest DDR4 memory, and support up to 10 Gbps of private network bandwidth.

Application Scenarios

These instances are an ideal choice for:
Applications that require extensive memory operations, searches, and computations, such as high-performance databases and distributed memory caching.
Users who build their own Hadoop clusters or Redis to use in fields such as computational genomics.

Features

2.5 GHz Intel® Xeon® Skylake 6133 processor, offering stable computing performance.
The latest generation of 6-channel DDR4 memory with a bandwidth of 2,666 MT/sec.
Higher specification, M3.16XLARGE512, offering 64 vCPUs and 512 GB memory.
Supported CPU/RAM ratios: 1:8 and 1:12.
Priced lower than any other memory-optimized instances with the same memory.
The network performance of an instance depends on its specification. A higher specification means a greater network forwarding performance and a higher private network bandwidth cap.

Requirements

Supports pay-as-you-go.
Supports both classic network and VPC.
See below for available M3 specifications:

Specification	vCPU	Memory (GB)	Throughput (pps) (received and sent packets)	Number of queues	Private network bandwidth (Gbps) (outbound	CPU clock rate	Notes

					and inbound)		
M3.SMALL8	1	8	300,000	1	1.5	2.5 GHz	-
M3.MEDIUM16	2	16	400,000	2	1.5	2.5 GHz	-
M3.LARGE32	4	32	800,000	4	1.5	2.5 GHz	-
M3.2XLARGE64	8	64	1,200,000	8	1.5	2.5GHz	-
M3.3XLARGE96	12	96	1,200,000	12	1.5	2.5 GHz	-
M3.3XLARGE144	12	144	1,200,000	12	1.5	2.5 GHz	-
M3.4XLARGE128	16	128	1,200,000	16	2.0	2.5 GHz	-
M3.4XLARGE192	16	192	1,200,000	16	2.0	2.5 GHz	-
M3.8XLARGE256	32	256	1,200,000	16	4.0	2.5 GHz	-
M3.8XLARGE384	32	384	1,200,000	16	4.0	2.5 GHz	-
M3.16XLARGE512	64	512	1,200,000	16	8.0	2.5 GHz	-

Memory Optimized M2

M2 instances are designed to deliver fast performance for workloads that process large data sets in memory. They feature a large memory, making them ideal for applications that require high in-memory computing performance.

Application Scenarios

These instances are ideal for:

Applications that require extensive memory operations, searches, and computations, such as high-performance databases and distributed memory caching.

Users who build their own Hadoop clusters or Redis to use in fields such as computational genomics.

Features

2.4 GHz Intel® Xeon® E5-2680v4 processor with DDR4 memory.
Memory of up to 448 GB is available. The M2.14XLARGE448 model offers 56 vCPUs and 448 GB memory.
CPU/RAM ratio 1:8.
Priced lower than any other memory-optimized instances with the same memory.

Requirements

M2 instances are pay-as-you-go. They can also be used as production instances of memory optimized host HM20 in CDHs.
Supports only VPC.
See below for available M2 specifications:

Specification	vCPU	Memory (GB)	Throughput (pps) (received and sent packets)	Number of queues	Private network bandwidth (Gbps) (outbound and inbound)	CPU clock rate	Notes
M2.SMALL8	1	8	200,000	1	1.5	2.4 GHz	-
M2.MEDIUM16	2	16	250,000	2	1.5	2.4 GHz	-
M2.LARGE32	4	32	450,000	4	1.5	2.4 GHz	-
M2.2XLARGE64	8	64	500,000	8	1.5	2.4 GHz	-
M2.3XLARGE96	12	96	500,000	8	2.5	2.4 GHz	-
M2.4XLARGE128	16	128	500,000	8	3.0	2.4 GHz	-
M2.6XLARGE192	24	192	700,000	8	4.5	2.4GHz	-
M2.8XLARGE256	32	256	700,000	8	6.0	2.4GHz	-
M2.12XLARGE384	48	384	700,000	8	9.0	2.4GHz	-
M2.14XLARGE448	56	448	700,000	8	10.0	2.4GHz	-

Memory Optimized M1

M1 instances come with approximately 1:8 CPU to memory ratio. They are well suited for applications that require memory-intensive operations, searches, and computations, such as high-performance databases and distributed in-memory caching.

Features

2.3 GHz Intel Xeon® E5-2670 v3 processor with DDR3 memory, providing more instance sizes with stronger computing capacities.
Golden CPU to memory ratio for memory-intensive applications, satisfying the needs for large-scale business deployment.

Application Scenarios

These instances are ideal for:
Applications that require memory-intensive operations, searches, and computations, such as high-performance databases and distributed in-memory caching.
Self-built Hadoop clusters or Redis for fields such as computational genomics.

Requirements

Supports pay-as-you-go.
Supports both classic network and VPC.
In the M1 instance list, "-" indicates that, for this specification, the send/receive packets and CPU clock rate are not guaranteed. If you have a strict performance requirement, please choose models with performance guarantee.
See below for available M1 specifications:

Specification	vCPU	Memory (GB)	Throughput (pps) (received and sent packets)	Number of queues	Private network bandwidth (Gbps) (outbound and inbound)	CPU clock rate	Notes
M1.SMALL8	1	8	-	1	1.5	2.3GHz	-
M1.MEDIUM16	2	16	-	2	1.5	2.3GHz	-
M1.LARGE32	4	32	-	4	1.5	2.3GHz	-
M1.2XLARGE64	8	64	-	8	2.0	2.3GHz	-
M1.3XLARGE96	12	96	-	8	2.5	2.3GHz	-
M1.4XLARGE128	16	128	-	8	3.5	2.3GHz	-

M1.6XLARGE192	24	192	-	8	5.0	2.3GHz	-
M1.8XLARGE256	32	256	-	8	7.0	2.3GHz	-
M1.12XLARGE368	48	368	-	8	10.0	2.3GHz	-

Note:
In the M1 instance list, "-" indicates that, for this specification, the send/receive packets and CPU clock rate are not guaranteed. If you have a strict performance requirement, please choose models with performance guarantee.

High I/O Family

This family features high random IOPS, high throughput and low latency, making it suitable for I/O-intensive applications that require high disk read/write performance and low latency, such as high-performance databases.
Note:
ITA5、IT3 and IT5 instances use local disk as the data disk, which may lose data (e.g., when the host crashes). If your application cannot guarantee data reliability, we recommend you choose an instance that can use cloud disks as the data disk.

High I/O ITA5

Note:
This instance specification family is in Beta Testing, and we welcome your suggestions and feedback.

High I/O Instance Family

The High I/O ITA5 Instance is the latest generation of high I/O instances designed specifically for I/O-intensive workloads. Based on NVMe SSD instance storage, it offers low latency, ultra-high IOPS, and high throughput storage resources at a lower cost. Suitable for high-performance relational databases, Elasticsearch, and other I/O-intensive services.

Instance Characteristics

The new generation of Tencent Cloud's self-developed StarSea dual-socket server, equipped with the AMD EPYC™ Bergamo processor.
Equipped with AMD EPYC™ Bergamo Processor, Turbo Boost up to 3.1GHz.
Instance network performance is matched to specifications; higher specifications enable stronger network forwarding performance and higher internal network bandwidth limits.
Utilizing NVMe SSD for instance storage, it provides low latency and ultra-high IOPS.

Application scenario

High-performance databases, NoSQL databases (e.g., MongoDB), clustered databases.
Online Transaction Processing (OLTP) systems, Elasticsearch search, and other I/O-intensive applications requiring low latency.

Instance Requirements

If the local hard disk is damaged, we support online disk replacement operations.
If a cloud server instance has already experienced downtime, we will inform you and proceed with maintenance operations.
An ITA5 instance not having monitoring components installed means the platform cannot monitor the instance as thoroughly. If the instance fails, normal notifications cannot be sent, Please activate cloud monitoring during purchasing.
ITA5 instances can be used as annual and monthly subscription instances and on-demand billing instances.
Only in Virtual Private Cloud can ITA5 instances be launched.
ITA5 instances do not support adjusting configuration or the 'shutdown not charged' feature.
ITA5 instances support purchasing configuration; please see the instance specifications below.
The instance supports up to 160G of network bandwidth, depending on the instance kernel version and running environment support. When pps exceeds 10 million and bandwidth is greater than 50Gbps, the kernel protocol stack causes significant network performance loss. In such cases, the bandwidth measured by netperf testing may not meet expectations. The DPDK method can be used to shield the differences of the cloud service kernel protocol stack, accessing the true network performance of the instance. For testing methods, please refer to High-throughput network performance testing.

Specification	vCPU	Memory (GB)	Network packet transmission (pps) (Outbound + Inbound)	Number of connections	Queue count	Private Network Standard/Burst Bandwidth (Gbps) (Outbound + Inbound)
ITA5.4XLARGE64	16	64	1.4 million	500,000	16	5/10
ITA5.8XLARGE128	32	128	2.8 million	1 million	32	10/25

ITA5.16XLARGE256	64	256	5.6 million	2 million	48	20/25
ITA5.32XLARGE512	128	512	11.2 million	4 million	48	40/-
ITA5.64XLARGE1024	256	1024	22.5 million	8 million	48	80/-
ITA5.128XLARGE2304	512	2304	45 million	16 million	48	160/-

High I/O IT5

IT5 instances, based on NVMe SSD storage, are designed for I/O-intensive workloads. They come with storage resources featuring low latency, ultra-high IOPS and high throughput. They offer an IOPS of up to 2.05 million and a throughput of up to 11 GB/s, allowing for an ultra-high IOPS at a low cost. This instance type is suitable for I/O-intensive applications including high-performance relational databases, Elasticsearch, etc.

Application Scenarios

High-performance databases, NoSQL databases (e.g. MongoDB), and clustered databases.
I/O intensive applications that require low latency, such as online transaction processing (OLTP) systems, and Elasticsearch.

Features

2.5 GHz Intel® Xeon® Cascade Lake processor, offering stable computing performance.
The latest generation of 6-channel DDR4 memory.
Supports up to 23 Gbps of private network bandwidth to meet extremely high private network transmission requirements.
NVMe SSD is used for instance storage, offering low latency and ultra-high IOPS.
Up to 650,000 random read IOPS (block size of 4 KB), and up to 2.8 GB/s sequential read throughput (block size of 128 KB) for a single disk.
Up to 2.05 million random read IOPS (block size of 4 KB), and up to 11 GB/s sequential read throughput (block size of 128 KB) for a CVM.

Requirements

If the local disk is damaged, you will need to shut down the CVM instance for us to replace the local disk.
If the CVM instance crashes, we will notify you and perform fixes.
If Agent has not been installed, IT5 instances cannot be monitored. You will not receive a notification when the CVM instance fails, which is very risky. For more information about how to install Agent, see Installing CVM Agents.
Supports pay-as-you-go.
Supports only VPC.
The configuration of IT5 instances cannot be changed.
See below for available IT5 specifications:

Specification	vCPU	Memory (GB)	Throughput (pps) (received and sent packets)	Number of queues	Private network bandwidth (Gbps) (outbound and inbound)	CPU clock rate	Notes
IT5.4XLARGE64	16	64	1,500,000	4	6	2.5GHz	1 x 3570 GB local NVMe SSD
IT5.8XLARGE128	32	128	2,500,000	8	12	2.5GHz	2 x 3570 GB local NVMe SSDs
IT5.16XLARGE256	64	256	5,000,000	16	23	2.5GHz	4 x 3570 GB local NVMe SSDs

High I/O IT3

IT3 instances, based on NVMe SSD storage, are designed for I/O-intensive workloads. They come with storage resources featuring low latency, ultra-high IOPS and high throughput. They offer an IOPS of up to 1.8 million and a throughput of up to 11 GB/s, allowing for an ultra-high IOPS at a low cost. This instance type is suitable for I/O-intensive applications including high-performance relational databases, Elasticsearch, etc.
IT3 instances feature inexpensive and intensive storage. They can provide each CPU core with higher IOPS and larger private network bandwidth.

Application Scenarios

High-performance databases, NoSQL databases (e.g. MongoDB), and clustered databases.
I/O intensive applications that require low latency, such as online transaction processing (OLTP) systems, and Elasticsearch.

Features

2.5 GHz Intel® Xeon® Skylake 6133 processor, offering stable computing performance.
The latest generation of 6-channel DDR4 memory.
Supports up to 23 Gbps of private network bandwidth to meet extremely high private network transmission requirements.
NVMe SSD is used for instance storage, offering low latency and ultra-high IOPS.
Up to 620,000 random read IOPS (block size of 4 KB), and up to 2.9 GB/s sequential read throughput (block size of 128 KB) for a single disk.
Up to 1.8 million random read IOPS (block size of 4 KB), and up to 11 GB/s sequential read throughput (block size of 128 KB) for a CVM.

Requirements

If the local disk is damaged, you will need to shut down the CVM instance for us to replace the local disk.
If the CVM instance crashes, we will notify you and perform fixes.
If Agent has not been installed, IT3 instances cannot be monitored. You will not receive a notification when the CVM instance fails, which is very risky. For more information about how to install Agent, see Installing CVM Agents.
Supports pay-as-you-go.
Supports only VPC.
The configuration of IT3 instances cannot be changed.
See below for available IT3 specifications:

Specification	vCPU	Memory (GB)	Throughput (pps) (received and sent packets)	Number of queues	Private network bandwidth (Gbps) (outbound	CPU clock rate	Notes

					and inbound)		
IT3.4XLARGE64	16	64	1,500,000	4	6	2.5GHz	1 × 3720 GB local NVMe SSD
IT3.8XLARGE128	32	128	2,500,000	8	12	2.5GHz	2 × 3720 GB local NVMe SSDs
IT3.16XLARGE256	64	256	5,000,000	16	23	2.5GHz	4 × 3720 GB local NVMe SSDs

Big Data Family

The big data family is equipped with massive storage resources, features high throughput, and is suitable for throughput-intensive applications such as Hadoop distributed computing, massive log processing, distributed file systems, and large data warehouses.

Note:

D3 and D2 instances use local disk as the data disk, which may lose data (e.g., when the host crashes). If your application cannot guarantee data reliability, we recommend you choose an instance that can use cloud disks as the data disk.

Big Data D3

As the latest generation of big data instances, D3 instances are equipped with massive storage resources, and carry up to 94 TB SATA HDD local storage, making them suitable for throughput-intensive services, such as Hadoop distributed computing and parallel data processing.

Application Scenarios

Distributed computing services such as Hadoop MapReduce/HDFS/Hive/HBase.

Business scenarios such as Elasticsearch, log processing, and large data warehouse.
Internet, finance, and industries that require big data computing and storage analysis, as well as applications that require massive data storage and computing.

Features

2.5 GHz Intel® Xeon® Cascade Lake processor with DDR4 memory.
Instances carry up to 24 local disks of 4 TB and up to 94 TB of HDD local storage.
At least 190 MB/s sequential read throughput performance for a single disk.
Millisecond read/write latency.
CPU/RAM ratio 1:4, tailored for the big data scenarios.

Requirements

If the local disk is damaged, you will need to shut down the CVM instance for us to replace the local disk.
We will notify you and perform fixes if the CVM instance crashes.
If Agent has not been installed, D3 instances cannot be monitored. You will not receive a notification when the CVM instance fails, which is very risky. For more information about how to install Agent, see Installing CVM Agents.
Supports pay-as-you-go.
Supports only VPC.
The configuration of D3 instances cannot be changed.
See below for available D3 specifications:

Specification	vCPU	Memory (GB)	Throughput (pps) (received and sent packets)	Number of connections	Number of queues	Private network bandwidth (Gbps) (outbound and inbound)	CPU clock rate
D3.2XLARGE32	8	32	800,000	250,000	2	4.0	2.5GHz
D3.4XLARGE64	16	64	1,500,000	300,000	4	7.0	2.5 GHz

D3.8XLARGE128	32	128	2,500,000	600,000	8	14.0	2.5 GHz
D3.16XLARGE256	64	256	5,000,000	1,200,000	12	27.0	2.5GHz
D3.21XLARGE320	84	320	6,000,000	1,600,000	16	32.0	2.5 GHz

Big Data D2

D2 instances are equipped with massive storage resources, and carry up to 144 TB SATA HDD local storage, making them suitable for throughput-intensive services, such as Hadoop distributed computing and parallel data processing.

Application Scenarios

Distributed computing services such as Hadoop MapReduce/HDFS/Hive/HBase
Business scenarios such as Elasticsearch, log processing, and large data warehouse.
Internet, finance, and industries that require big data computing and storage analysis, as well as applications that require massive data storage and computing.

Features

2.4 GHz Intel® Xeon® Skylake 6148 processor with DDR4 memory.
Instances carry up to 12 local disks of 12 TB and up to 144 TB of HDD local storage
For a single disk, the sequential read throughput is 220+ MB/s and sequential write throughput is 220+ MB/s (128 KB block size and depth of 32)
For the CVM, the throughput can reach up to 2.8 GB/s (block size of 128 KB and queue depth of 32).
Read/write latency is as low as 2-5 ms
Higher specification, D2.19XLARGE320, offering 76 vCPUs and 320 GB of memory.
Local storage with a unit price as low as 1/10 of S2, making its total cost close to that of a self-built Hadoop clusters in IDC.

CPU/RAM ratio 1:4, tailored for the big data scenarios.

Requirements

If the local disk is damaged, you will need to shut down the CVM instance for us to replace the local disk.

If the CVM instance crashes, we will notify you and perform fixes.

Supports pay-as-you-go.

Supports only VPC.

The configuration of D2 instances cannot be changed.

See below for available D2 specifications:

Specification	vCPU	Memory (GB)	Throughput (pps) (received and sent packets)	Number of queues	Private network bandwidth (Gbps) (outbound and inbound)	CPU clock rate	Notes
D2.2XLARGE32	8	32	800,000	2	3.0	2.4GHz	1 x 11176 GB local SATA HDD
D2.4XLARGE64	16	64	1,500,000	4	6.0	2.4GHz	2 x 11176 GB local SATA HDDs
D2.6XLARGE96	24	96	2,000,000	6	8.0	2.4GHz	3 x 11176 GB local SATA HDDs
D2.8XLARGE128	32	128	2,500,000	8	11.0	2.4GHz	4 x 11176 GB local SATA HDDs

D2.16XLARGE256	64	256	5,000,000	16	22.0	2.4GHz	8 x 11176 GB local SATA HDDs
D2.19XLARGE320	76	320	6,000,000	16	25.0	2.4GHz	12 x 11176 GB local SATA HDDs

Computing Instance Family

This family provides the highest single-core computing performance with a turbo frequency up to 3.8 GHz. It is suitable for compute-intensive applications such as batch processing, high performance computing, and dedicated game servers.

Compute Optimized C6

As the latest generation of computing instances, C6 instances are powered by an all-new virtualization platform and provide stable and reliable high-performance computing services for greater bandwidth and lower latency. These instances have the most powerful processors and the best cost performance among CVMs. They are ideal for applications that require high computing performance and high-concurrence reads/writes.
C6 instances come with new Intel® Xeon® processors, and support up to 100 Gbps of private network bandwidth.
Note:
This instance type is now only available to beta users. Contact your account manager to purchase it.

Features

A new-generation Star Lake two-way server developed by Tencent Cloud, with the 3rd generation Intel® Xeon® scalable processor.
Intel® Xeon® Ice Lake processor, with a CPU clock rate of 3.2 GHz and a turbo frequency of 3.5 GHz.
Supported CPU/RAM ratios: 1:2 and 1:4.
Supports up to 100 Gbps of private network bandwidth, providing ultra-high packet throughput to meet high private network transmission requirements.
Supports all kinds of cloud disks.
Supports the configuration to disable or enable Hyper-Threading.

Use Cases

Compute Optimized C6 instances are applicable to the following scenarios:

Batch processing workloads and high performance computing (HPC).

High-traffic Web frontend server.

Other compute-intensive services such as massively multiplayer online (MMO) game servers.

Requirements

Supports pay-as-you-go.

Supports only VPC.

See below for available C6 specifications:

Specification	vCPU	Memory (GB)	Throughput (pps) (received and sent packets)	Number of connections	Number of queues	Private network bandwidth (Gbps) (outbound and inbound)	CPU clock rate
C6.LARGE8	4	8	800,000	600,000	4	5	3.2 GHz
C6.LARGE16	4	16	800,000	600,000	4	5	3.2 GHz
C6.2XLARGE16	8	16	1,600,000	1,300,000	8	9	3.2 GHz
C6.2XLARGE32	8	32	1,600,000	1,300,000	8	9	3.2 GHz
C6.4XLARGE32	16	32	3,300,000	2,700,000	16	18	3.2 GHz
C6.4XLARGE64	16	64	3,300,000	2,700,000	16	18	3.2 GHz
C6.8XLARGE128	32	128	6,600,000	5,500,000	32	35	3.2 GHz
C6.23MEDIUM216	46	216	9,500,000	8,000,000	46	50	3.2 GHz
C6.16XLARGE256	64	256	13,000,000	11,000,000	48	70	3.2 GHz

C6.20XLARGE320	80	320	16,000,000	13,900,000	48	87	3.2 GHz
C6.23XLARGE432	92	432	19,000,000	16,000,000	48	100	3.2 GHz

Compute Optimized C5

As the latest generation of computing instances, C5 instances are powered by an all-new virtualization platform to provide stable and reliable high-performance computing services for greater bandwidth and lower latency. These instances have the most powerful processors and the best cost performance among CVMs. They are ideal for applications that require high computing performance and high-concurrence reads/writes.
C5 instances come with the latest Xeon® Cooper Lake processors.
Note:
This instance type is now only available to beta users. Contact your account manager to purchase it.

Features

Intel® Xeon® Cooper Lake processor, with a base frequency of 3.4 GHz and turbo frequency of 3.8 GHz.
New Intel Brain Floating Point 16-bit (blfoat16) instruction set.
Supported CPU/RAM ratios: 1:2 and 1:4.
The network performance of an instance depends on its specification. A higher specification means a greater network forwarding performance and a higher private network bandwidth cap.
Supports all kinds of cloud disks.
Supports the configuration to disable or enable Hyper-Threading.

Application Scenarios

These instances are ideal for:
Batch processing workloads and high performance computing (HPC).
High-traffic Web frontend server.
Other compute-intensive services such as massively multiplayer online (MMO) game servers.

Requirements

Supports pay-as-you-go.
Supports both classic network and VPC.
See below for available C5 specifications:

Specification	vCPU	Memory (GB)	Throughput (pps) (received and sent packets)	Number of connections	Number of queues	Private network bandwidth (Gbps) (outbound	CPU clock rate

						and inbound)	
C5.LARGE8	4	8	500,000	250,000	4	1.5	3.4GHz
C5.LARGE16	4	16	500,000	250,000	4	1.5	3.4GHz
C5.2XLARGE16	8	16	800,000	250,000	8	3.0	3.4GHz
C5.2XLARGE32	8	32	800,000	250,000	8	3.0	3.4GHz
C5.4XLARGE32	16	32	1,500,000	300,000	8	6.0	3.4GHz
C5.4XLARGE64	16	64	1,500,000	300,000	8	6.0	3.4GHz
C5.8XLARGE64	32	64	2,500,000	600,000	16	12.0	3.4GHz
C5.8XLARGE128	32	128	2,500,000	600,000	16	12.0	3.4GHz
C5.12XLARGE96	48	96	4,000,000	900,000	24	18.0	3.4GHz
C5.12XLARGE192	48	192	4,000,000	900,000	24	18.0	3.4GHz
C5.13XLARGE184	52	184	4,000,000	1,000,000	32	19.0	3.4GHz
C5.16XLARGE256	64	256	5,000,000	1,200,000	32	24.0	3.4GHz
C5.26XLARGE368	104	368	6,000,000	2,000,000	32	36.0	3.4GHz

Compute Optimized C4

C4 instances have up to 25 Gbps of private network bandwidth for greater bandwidth and lower latency. These instances have the most powerful processors and the best cost performance among CVMs. They are ideal for applications that require high computing performance and high-concurrence reads/writes.
C4 instances are equipped with the latest Xeon® Cascade Lake processors, and support up to 25 Gbps of private network bandwidth.

Use Cases

These instances are an ideal choice for:
Batch processing workloads and high performance computing (HPC).
High-traffic Web frontend server.
Other compute-intensive services such as massively multiplayer online (MMO) game servers.

Features

3.2 GHz Intel® Xeon® Cascade Lake processor, with a turbo frequency 3.7 GHz.

New Intel Advanced Vector Extension (AVX-512) instruction set.
Supports up to 25 Gbps of private network bandwidth to meet extremely high private network transmission requirements.
Supported CPU/RAM ratios: 1:2 and 1:4.
The network performance of an instance depends on its specification. A higher specification means a greater network forwarding performance and a higher private network bandwidth cap.
Supports all kinds of cloud disks.
Supports the configuration to disable or enable Hyper-Threading.

Requirements

Supports pay-as-you-go.
Supports both classic network and VPC.
See below for available C4 specifications:

Specifications	vCPU	MEM (GB)	Throughput (pps) (received and sent packets)	Number of queues	Private network bandwidth (Gbps) (outbound and inbound)	CPU clock rate	Notes
C4.LARGE8	4	8	600,000	4	2.0	3.2GHz	-
C4.LARGE16	4	16	600,000	4	2.0	3.2GHz	-
C4.2XLARGE16	8	16	1,000,000	8	4.0	3.2GHz	-
C4.2XLARGE32	8	32	1,000,000	8	4.0	3.2GHz	-
C4.4XLARGE64	16	64	1,000,000	16	7.0	3.2GHz	-
C4.8XLARGE174	32	174	1,000,000	16	13.0	3.2GHz	-
C4.16XLARGE348	64	348	1,000,000	16	25.0	3.2GHz	-

Compute Network-optimized CN3

CN3 instances have up to 25 Gbps of private network bandwidth for greater bandwidth and lower latency. These instances have the most powerful processors and the best cost performance among CVMs. They are ideal for applications that require high computing performance and high-concurrence reads/writes.
Compute CN3 instances are equipped with the new Xeon® Skylake processors and support up to 25 Gbps of private network bandwidth, which is 2.5 times greater than Compute C3.

Application Scenarios

These instances are an ideal choice for:

Batch processing workloads and high performance computing (HPC)

High-traffic Web frontend server

Other compute-intensive services such as massively multiplayer online (MMO) game servers

Features

3.2 GHz Intel Xeon® Skylake 6146 processors

New Intel Advanced Vector Extension (AVX-512) instruction set

Supports up to 25 Gbps of private network bandwidth to meet extremely high private network transmission requirements

The latest generation of 6-channel DDR4 memory with a memory bandwidth of 2,666 MT/s.

Supported CPU/RAM ratios: 1:2 and 1:4

The network performance of an instance depends on its specification. A higher specification means a greater network forwarding performance and a higher private network bandwidth cap.

Support all types of cloud disks

Requirements

Supports pay-as-you-go.

Supports both classic network and VPC.

See below for available CN3 specifications:

Specification	vCPU	Memory (GB)	Throughput (pps) (received and sent packets)	Number of queues	Private network bandwidth (Gbps) (outbound and inbound)	CPU clock rate	Notes
CN3.LARGE8	4	8	600,000	4	3.0	3.2 GHz	-
CN3.LARGE16	4	16	600,000	4	3.0	3.2 GHz	-
CN3.2XLARGE16	8	16	1,000,000	8	5.0	3.2 GHz	-
CN3.2XLARGE32	8	32	1,000,000	8	5.0	3.2 GHz	-
CN3.4XLARGE32	16	32	1,000,000	16	9.0	3.2	-

						GHz	
CN3.4XLARGE64	16	64	1,000,000	16	9.0	3.2 GHz	-
CN3.8XLARGE64	32	64	1,000,000	16	17.0	3.2 GHz	-
CN3.8XLARGE128	32	128	1,000,000	16	17.0	3.2 GHz	-
CN3.12XLARGE160	48	160	1,000,000	16	25.0	3.2 GHz	-

Compute C3

C3 instances have the most powerful processors and the best cost performance among CVMs. They are ideal for applications that require high computing performance and high-concurrence reads/writes.
Compute C3 instances are equipped with the latest Xeon® Skylake processors with the latest DDR4 memory and support for higher private network bandwidth.

Application Scenarios

These instances are an ideal choice for:
Batch processing workloads and high performance computing (HPC).
High-traffic Web frontend server.
Other compute-intensive services such as massively multiplayer online (MMO) game servers.

Features

3.2 GHz Intel Xeon® Skylake 6146 processors.
New Intel Advanced Vector Extension (AVX-512) instruction set.
The latest generation of 6-channel DDR4 memory with a memory bandwidth of 2,666 MT/s.
Supported CPU/RAM ratios: 1:2 and 1:4.
The network performance of an instance depends on its specification. A higher specification means a greater network forwarding performance and a higher private network bandwidth cap.
Support all types of cloud disks.

Requirements

Supports pay-as-you-go.
Supports both classic network and VPC.
See below for available C3 specifications:

Specification	vCPU	Memory	Throughput	Number	Private	CPU	Notes

		(GB)	(pps) (received and sent packets)	of queues	network bandwidth (Gbps) (outbound and inbound)	clock rate	
C3.LARGE8	4	8	600,000	2	2.5	3.2 GHz	-
C3.LARGE16	4	16	600,000	2	2.5	3.2 GHz	-
C3.2XLARGE16	8	16	1,000,000	2	3.0	3.2 GHz	-
C3.2XLARGE32	8	32	1,000,000	2	3.0	3.2 GHz	-
C3.4XLARGE32	16	32	1,000,000	4	4.5	3.2 GHz	-
C3.4XLARGE64	16	64	1,000,000	4	4.5	3.2 GHz	-
C3.8XLARGE64	32	64	1,000,000	8	8.0	3.2 GHz	-
C3.8XLARGE128	32	128	1,000,000	8	8.0	3.2 GHz	-

Compute C2

C2 instances are optimized with high-performance processors and deliver high cost performance. They are ideal for applications that require high computing performance and high-concurrence reads/writes.

Application Scenarios

These instances are an ideal choice for:
Batch processing workloads.
High-traffic Web servers, massively multiplayer online (MMO) game servers.
High-performance computing (HPC) and other compute-intensive applications.

Features

3.2 GHz Intel Xeon® E5-2667v4 processors with up to 3.6 GHz of turbo frequency; DDR4 memory.
Supported CPU/RAM ratios: 1:2 and 1:4.

The network performance of an instance depends on its specification. A higher specification means a greater network forwarding performance and a higher private network bandwidth cap.
Support all types of cloud disks.

Requirements

C2 instances are pay-as-you-go, and can also be used as production instances of computing host HC20 in CDHs.
Supports both classic network and VPC.
See below for available C2 specifications:

Specification	vCPU	Memory (GB)	Throughput (pps) (received and sent packets)	Number of queues	Private network bandwidth (Gbps) (outbound and inbound)	CPU clock rate	Notes
C2.LARGE8	4	8	550,000	4	2.5	3.2 GHz	-
C2.LARGE16	4	16	550,000	4	2.5	3.2 GHz	-
C2.LARGE32	4	32	550,000	4	2.5	3.2 GHz	-
C2.2XLARGE16	8	16	550,000	8	3.5	3.2 GHz	-
C2.2XLARGE32	8	32	550,000	8	3.5	3.2 GHz	-
C2.4XLARGE32	16	32	850,000	8	6.0	3.2 GHz	-
C2.4XLARGE64	16	64	850,000	8	6.0	3.2 GHz	-
C2.8XLARGE96	32	96	850,000	8	10.0	3.2 GHz	-

BeFast Instance Family

BeFast 1

BeFast 1 instances are of high performance-cost ratio and customized for small and medium-sized businesses. They can provide cost-effective, balanced and stable computing, memory and network resources. This instance family uses Tencent Cloud's intelligent scheduling algorithm for dynamic resource management, continuously providing stable computing power and ensuring resource supply for applications. It supports processor-to-memory ratios of 1:1, 1:2, and 1:4.
BeFast 1 instances offer the same user experience, computing power assurance, resource supply, and service support as standard instances. The underlying layer is adaptive to multiple generations of servers. The benchmark vCPU computing power is on par with that of the 5th-generation enterprise-level instances (such as S5, SA2), and the maximum vCPU computing power is on par with that of the latest generation of enterprise-level instances. They are suitable for business scenarios compatible with different CPU models.

Application Scenarios

BeFast 1 instances are quite suitable for the following scenarios:
Small and medium-sized enterprise-level applications.
Small and medium-sized database systems, caches, search clusters.
Computing clusters, and memory-dependent data processing.

Features

Cost-effective, with the lowest unit price among all other families of instances in specification comparison.
CPU clock rate is not less than 2.25 GHz. CPU/RAM ratio supports 1:1/1:2/1:4.
The network performance of instances corresponds to their specifications. Higher specifications result in stronger network forwarding performance and higher private network bandwidth limits.
Note:
Instances of this family are randomly deployed on different processor platforms. During the lifecycle of an instance, it may be migrated to another processor platform. Since there are clear performance differences among businesses across different platforms, it is recommended that you purchase standard instances if you have strong needs for business consistency, .

Requirements

BeFast 1 instances support monthly subscription and pay-as-you-go.
BeFast 1 instances can only be launched in VPCs.
It is recommended to pair with the TencentOS Server operating system for optimal application performance of the instance.
See the specifications below to learn the sizes of BeFast 1 instances available for purchase. Ensure the size of the BeFast 1 instance you choose meets the minimum CPU and memory requirements of your operating system and applications. In many cases, GUI operating system (such as Windows) that consumes extensive memory and CPU resources may need higher specifications. As the needs of your workload for memory and CPU increase, you can choose higher specifications or other instance types.

Specification	vCPU	Memory (GB)	Throughput (pps) (received and sent packets)	Number of Connections	Number of Queues	Private network bandwidth (Gbps) (outbound and inbound)	Notes
BF1.MEDIUM2	2	2	250,000 or more	250,000 or more	2	1.5 or higher	-
BF1.MEDIUM4	2	4	250,000 or more	250,000 or more	2	1.5 or higher	-
BF1.MEDIUM8	2	8	250,000 or more	250,000 or more	2	1.5 or higher	-
BF1.LARGE4	4	4	300,000 or more	250,000 or more	2	1.5 or higher	-
BF1.LARGE8	4	8	300,000 or more	250,000 or more	2	1.5 or higher	-
BF1.LARGE16	4	16	300,000 or more	250,000 or more	2	1.5 or higher	-
BF1.2XLARGE8	8	8	600,000 or more	250,000 or more	2	1.5 or higher	-
BF1.2XLARGE16	8	16	600,000 or more	250,000 or more	2	1.5 or higher	-
BF1.2XLARGE32	8	32	600,000 or more	250,000 or more	2	1.5 or higher	-
BF1.4XLARGE16	16	16	1.3 million or more	500,000 or more	4	3 or higher	-
BF1.4XLARGE32	16	32	1.3 million or more	500,000 or more	4	3 or higher	-
BF1.4XLARGE64	16	64	1.3 million or more	500,000 or more	4	3 or higher	-
BF1.8XLARG64	32	64	2.6 million or more	1 million or more	8	5 or higher	-
BF1.8XLARG128	32	128	2.6 million	1 million or	8	5 or	-

			or more	more		higher	
BF1.12XLARG96	48	96	3.9 million or more	1.5 million or more	12	7 or higher	-
BF1.12XLARG192	48	192	3.9 million or more	1.5 million or more	12	7 or higher	-
BF1.16XLARG128	64	128	5.2 million or more	2 million or more	16	9 or higher	-
BF1.16XLARG256	64	256	5.2 million or more	2 million or more	16	9 or higher	-

Heterogeneous Computing Instance Family

This family is equipped with heterogeneous hardware such as GPU and FPGA to deliver real-time, fast parallel computing and floating-point computing capabilities. It is suitable for high-performance applications such as deep learning, scientific computing, video encoding/decoding, and graphics workstations.
The AMD GPU instances use AMD FirePro S7150 to provide outstanding graphics processing capability. It is the best choice for applications such as remote desktop, 3D rendering, and cloud gaming.
NVIDIA GPU instances use NVIDIA Tesla GPUs, including the previous generation M40, the current mainstream choice P4/P40, and the latest generation V100, to provide excellent general-purpose computing capability, making this the top choice for applications such as deep learning training/reasoning and scientific computing. NVIDIA GPU Instances >

Life Cycle

Last updated：2024-01-08 09:12:28

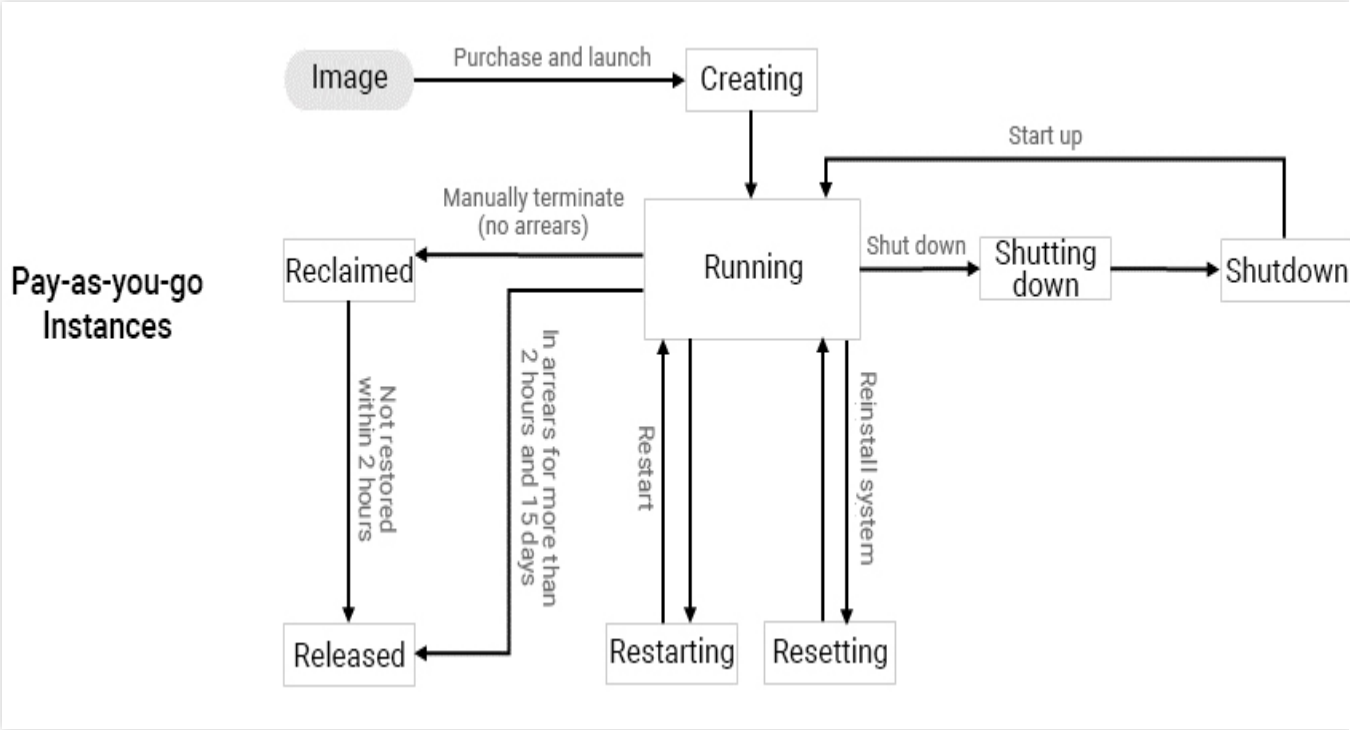
The lifecycle of a Tencent Cloud CVM instance refers to all statuses from the launch of the instance to its release. Properly managing the CVM instance in different statuses can ensure that applications running on the instance provide services economically and efficiently.

Instance Status

An instance has the following statuses:

Status Name	Status Attributes	Description
Creating	Intermediate status	The instance has been created but is not running yet.
Running	Steady status	The instance is running normally, and you can run your services on this instance.
Restarting	Intermediate status	A restart operation has been performed for the instance via the console or APIs, but the instance is not running yet. If this status lasts for a long time, there may be an exception.
Reinstalling	Intermediate status	The instance's system has been reinstalled or its disk has been reconfigured via the console or APIs, but the instance is not running yet.
Shutting Down	Intermediate status	A shutdown operation has been performed for the instance via the console or APIs, but the instance has not been shut down yet. If this status lasts for a long time, there may be an exception. We do not recommend forced shutdown.
Shutdown	Steady status	The instance has been shut down normally and cannot provide external services. Some instance attributes can only be modified when the instance is in shutdown status.
Terminating	Intermediate status	The instance has expired for 7 days or the user has performed the termination operation, but it has not completed yet.
Reclaimed	Steady status	A pay-as-you-go instance has been manually terminated for less than 2 hours and put into the recycle bin. In this status, the instance does not provide external services.
Released	Steady status	The release operation has been completed. The original instance no longer exists, cannot provide services, and its data is completely cleared.

Instance Status Transition:



Launching an instance

After an instance is launched, it goes to the "Creating" status. For an instance in this status, its hardware specifications will be configured according to the specified instance type, and the system will launch the instance using the image specified at launch.

An instance goes to "Running" status after it is created. An instance in the running status can be connected to and accessed normally.

For more information about instance startup, see Creating an Instance, Logging In to a Windows Instance, and Logging In to a Linux instance.

Restarting an instance

We recommend you restart an instance via the Tencent Cloud Console or Tencent Cloud APIs instead of running the OS restart command in the instance.

After the restart operation is performed, the instance will enter the restarting status.

Restarting an instance is like restarting a computer. After restarted, the instance will retain its public IP address, private IP address and all data on its disk.

Depending on its configuration, restarting the instance normally takes dozens of seconds to several minutes.

For more information about how to restart an instance, see Restarting Instances.

Shutting down an instance

You can shut down the instance in the console or by using APIs.
Shutting down an instance is like shutting down a computer.
A shutdown instance no longer provides services, but the billing of the instance continues.
A shutdown instance is still visible in the console.
You need to shutdown an instance for some operations, such as adjusting hardware configurations and resetting the password.
The shutdown operation does not change the CVM's public IP, private IP, or any data on its disks.
For more information about shutting down an instance, see Shutting Down Instances.

Terminating and releasing an instance

If you no longer need an instance, you can terminate and release it in the console or by using APIs.
Manual termination: Pay-as-you-go instances are released after being moved to the recycle bin for two hours.
Auto termination due to expiry or overdue payment: Pay-as-you-go instances will be automatically terminated when the account balance drops below 0 for 2 hours and 15 days. Billing will continue for the first 2 hours, then the instance will shut down and no longer be billed. The overdue pay-as-you-go instance will not enter the recycle bin and can be viewed on the instance list. You can continue to use the instance if you renew it within the specified time.
When an instance is terminated, its system disk and data disks specified at purchase will be released. However, cloud disks attached to the instance will not be affected.
For more information on terminating an instance, see Terminating Instances.

Spot Instance

Last updated：2024-06-25 15:51:01

Overview

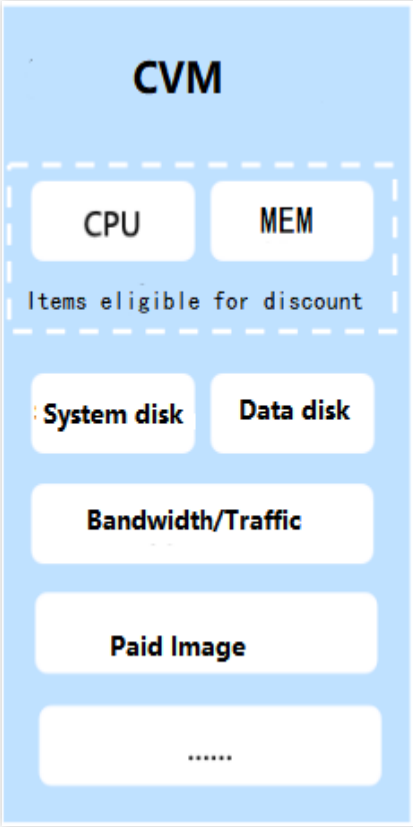
Spot instances represent a billing mode for Cloud Virtual Machine (CVM) services, where prices fluctuate in real time based on market supply and demand dynamics. Their most distinctive features include discounted sales and a system interruption mechanism. This means you can purchase instances at a certain discount, but the system may automatically reclaim these discounted instances. Once you purchase a spot instance, its use is virtually indistinguishable from that of a pay-as-you-go CVM instance, encompassing console operations, remote login, service deployment, and VPC association, among others. Compared to the pay-as-you-go billing mode, opting for spot instances can result in savings of up to 97% on instance costs.
Reference: FAQ > About Instance > Spot Instances
Reference: Spot Instances

Important Considerations

System interruption due to insufficient supply: Currently, spot instances will not be interrupted due to the price reason, but may be interrupted if spot instances are in short supply. If the supply is insufficient, Tencent Cloud will randomly repossess allocated spot instances without retaining the instance data.
Available in almost all regions: Spot instances are available in almost all Tencent Cloud regions. The instance types supported by spot instances are the same as those supported by pay-as-you-go instances. For the latest regions and instance types supported, see Spot Instances.

Features

1. Cost-effectiveness



Guangzhou Zone 3 Take S2.2XLARGE16 as an example Unit: USD/hour	Pay-as-you-go Instances	Spot Instances (Take maximal discou as an example)
CPU (8-core) MEM (16 GB)	0.38	0.076
System Disk (50 GB HDD disk)	0.01	0.01
Bandwidth (1 Mbps, bill-by-bandwidth)	0.06	0.06
Total Fee	0.45	0.146

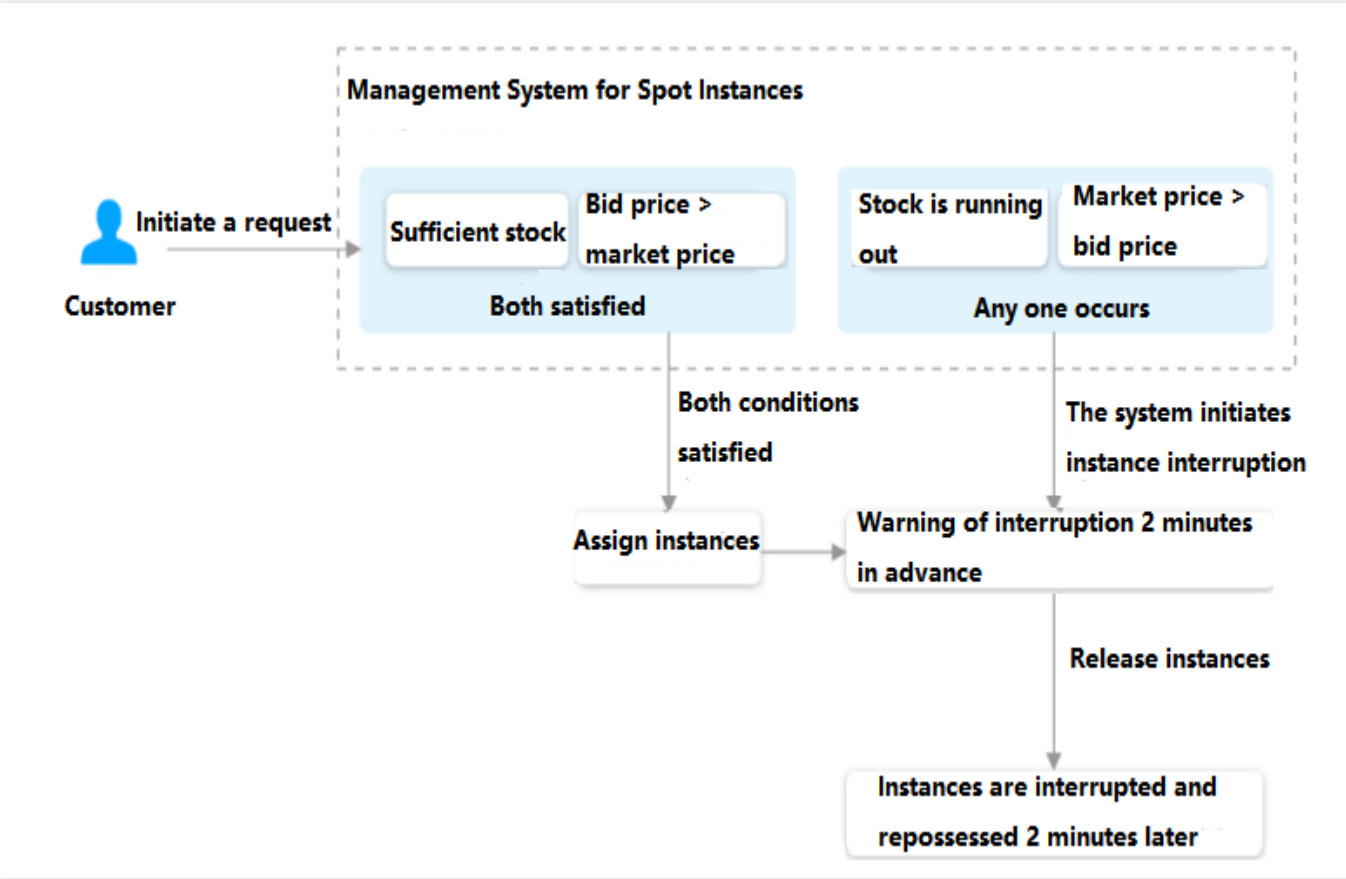
Spot instances are sold at a discount of up to 97% off the prices of pay-as-you-go instances.

Discount range: Spot instances are sold at a discount of up to 97% off the price of pay-as-you-go instances of the same specification.

Exclusions: The discount only applies to CVM CPU and memory. Other items including system disks, data disks, bandwidth, and paid images are not eligible for the discount.

Price fluctuations: When market supply and demand dynamics shift, prices will experience fluctuations.

2. System interruption mechanism



Unlike pay-as-you-go instances which can only be released by users, spot instances may be interrupted by the system due to price or resource availability reasons.

System interruption due to insufficient supply: Currently, spot instances will not be interrupted due to the price reason, but may be interrupted if spot instances are in short supply. If the supply is insufficient, Tencent Cloud will randomly repossess allocated spot instances without retaining the instance data.

Non-applicable Scenarios

As spot instances may be interrupted, their lifecycle is not under your control. Therefore, it is not recommended to run services with high stability requirement on a spot instance. For example:

Database services

Online and website services without load balancers

Core control nodes in a distributed architecture

Prolonged big data computing job lasting over 10 hours

Applicable Scenarios and Industries

Applicable scenarios

Big data computing
Online and website services with load balancers
Web crawler service
Other computing scenarios with fine granularity or support for checkpoint restart

Applicable industries

Gene sequencing and analysis
Drug crystal form analysis
Video transcoding and rendering
Financial and transaction data analysis
Image and multimedia processing
Science calculations, such as in geography and hydromechanics.

Limits and Restrictions

Quota limits: Currently, the spot instances for each account per availability zone can contain up to 200 vCPU cores in total. To increase the quota, please submit a ticket.
Purchase limits: Vouchers are not available to spot instances.
Operation restriction 1: You cannot upgrade and degrade the configuration of spot instances.
Operation restriction 3: **No charges when shut down** is not supported for spot instances.
Operation restriction 4: System reinstallation is not supported for spot instances.
Operation restriction 5: You cannot expand the system disks and data disks of spot instances.

Best Practices

1. Splitting tasks

Split a prolonged task into fine-grained subtasks for lower possibility of interruption.
Use a big data suite such as EMR that comes with a splitting mechanism.

2. Using load balancers to ensure the stability of online and website services

Use load balancers, such as CLB, at the access layer.
Use a combination of some pay-as-you-go instances and many spot instances for backend resources.
Monitor the interruptions of spot instances and remove instances that are about to be interrupted from the CLB.

3. Using a computing scheduling mode that supports checkpoint restart

Store intermediate computing results on permanent storage products such as COS, CFS, and NAS.
Be aware of the instance metadata to monitor which instances are about to be interrupted and save the computing results within the retention period of 2 minutes.
Resume the last computing when a spot instance is started again.

Storage
Storage Overview

Last updated：2024-01-08 09:12:28

Tencent Cloud provides a wide range of flexible and cost-effective data storage devices for CVM instances. The performance and price varies by the type of storage device, which is suitable for different scenarios.

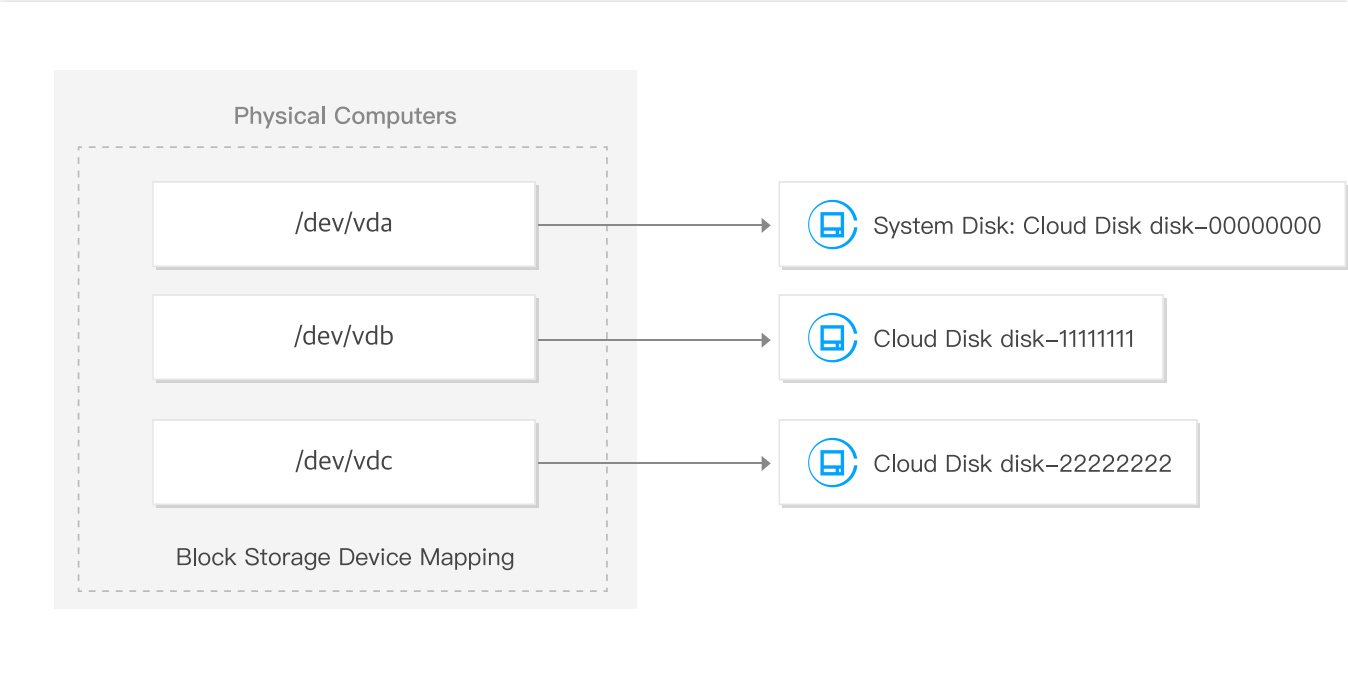
Storage Devices

Storage devices can be divided into the following categories by dimension:

Dimension	Category	Description
Storage Medium	HDD disk	Uses hard disk drive (HHD) as the storage medium. It features a low price and fast read/write speeds.
	SSD disk	Uses solid state drive (SSD) as the storage medium. It features excellent IOPS and read/write speeds, with an IOPS and throughput of up to 20 times and 16 times higher than those of the HDD disk. It is more expensive than the HDD disk.
Use Cases	System disk	Used to store the collection of systems that control and schedule the operation of CVM. It uses images.
	Data disk	Used to store all user data.
Architecture	Cloud Block Storage	Cloud Block Storage (CBS) is an elastic, highly available, highly reliable, low-cost, and customizable network block device that can be used as a standalone scalable disk for CVM. It provides data storage at a data block level and employs a 3-copy distributed mechanism to ensure data reliability. **You can adjust the hardware, disk, and network of a CVM with CBS.**
	Cloud Object Storage	Cloud Object Storage (COS) is a data storage device on the Internet. It allows data retrieval from any location on the CVM instance or the Internet, reducing storage costs. It is unsuitable for low-latency and high-IO scenarios.

Block storage device mapping

Each instance has a system disk to ensure basic data operations, and may mount more data disks. An instance uses block storage device-mapping to map the storage devices to locations it can identify.
CBS puts data into blocks in bytes and allows random access. Tencent Cloud supports two types of CBS devices: local disk and cloud disk.



This figure shows how CBS maps the block storage device to the CVM: it maps `/dev/vda` to the system disk and maps the two data disks to `/dev/vdb` and `/dev/vdc`, respectively.
The CVM instance can automatically create block storage device mapping for local disks and cloud disks mounted to itself.

Cloud Block Storage

Last updated：2024-01-08 09:12:28

Tencent Cloud Block Storage (CBS) provides a persistent block storage service for CVM instances.
CBS automatically stores data in multiple redundant copies in an availability zone to eliminate the risk of single points of data failure, providing up to 99.9999999% data reliability.
CBS offers cloud disks of multiple types and specifications to achieve stable and low-latency storage performance.
Cloud disks can be attached to and detached from CVM instances in the same availability zone. It takes only a few minutes to adjust the disk capacity.

Typical Use Cases

When your CVM is running out of disk space, you can purchase one or more cloud disks and attach them to the CVM. You don’t need to purchase extra storage capacity while purchasing a CVM. You can purchase cloud disks later when it’s necessary.
When you need to transfer data from one CVM to another, just detach the cloud data disk from the source CVM and attach it to the target CVM.
Use multiple cloud disks to form a Logical Volume Manager (LVM) to go beyond the physical limit of a single cloud disk.
Use multiple cloud disks to form a Redundant Array of Independent Disks (RAID) configuration to go beyond the I/O performance limit of a single cloud disk.

Lifecycle

The lifecycle of a **non-elastic cloud disk** is the same as that of the CVM. It is purchased with the CVM and used as a system disk. It cannot be attached to or detached from CVMs.
The lifecycle of an **elastic cloud disk** is independent from CVM instances. You can attach multiple cloud disks to a CVM instance as data disks, detach them, and then reattach them to another instance.

Types

Four types of cloud disks are provided, including **Premium Cloud Storage**, **SSD**, **Enhanced SSD**, and **Tremendous SSD**. Each type has unique performance and characteristics, and the price varies. You can choose the cloud disk type that best suits your application requirements. For more information, see Cloud Disk Types and Price Overview.

Relevant Operations

For information on CVM instance and cloud disk configurations, see Creating Cloud Disks and Attaching Cloud Disks. For information about best practices for capacity expansion, detachment, termination, and other operations, see the CBS documentation.

Local Storage

Last updated：2024-01-08 09:12:28

Overview

A local disk is a storage device on the same physical server as the CVM instance. It features high read/write I/O and low latency.
The local disk is a local storage device on the same physical server as the CVM instance. It is a reserved storage space on the physical server (currently only available to high I/O and big data CVM instances). The reliability of data stored on a local disk depends on that of the physical server. There may be a single point of failure.
Note:
In case of a hardware failure on the physical server of the CVM, local disk data lose may occur. We recommend data redundancy at the application layer to ensure reliability. If your application does not support this, consider using Cloud Block Storage to improve data reliability.
For a CVM using a local disk as the system disk, hardware (CPU, memory, and storage) upgrade is not supported. You can only adjust the bandwidth of the instance.

Use Cases

IO-intensive applications: For large relational databases, NoSQL, ElasticSearch, and other I/O-intensive applications that are more sensitive to latency, you can use the NVME SSD local disk that comes with high I/O CVM instances, but note that it carries the risk of a single point of failure.
Big data applications: For big data applications such as EMR that are less sensitive to latency and feature data redundancy at the upper layer to tolerate a single point of failure, you can use the SATA HDD disk that comes with big data CVM instances.

Local Disk Lifecycle

The lifecycle of a local disk is the same as that of the CVM instance it is attached. Therefore, local disks launch and terminate with CVM instances.

Local Disk Types

Local disks are local storage devices on the same physical server as the CVM instance. There are two types of local disks by media: SATA HDD and NVME SSD.

Disk Type	CVM
Local SATA HDD	Big data CVM
Local NVME SSD	High I/O CVM

Purchasing a Local Disk

A local disk can only be purchased together with a CVM instance. For more information on purchasing a CVM instance, see Creating Instances.

Cloud Object Storage (COS)

Last updated：2024-01-08 09:12:28

Cloud Object Storage (COS) provided by Tencent Cloud is a distributed storage service designed to handle a massive amount of files. You can store all your files in one place and manage them using an easy-to-use interface at any time.

COS stores data redundantly across multiple regions and allows multiple clients or application threads to perform read/write operations on the data at the same time. The data on COS can be retrieved in a similar way to data on Internet domain names and be accessed by using an HTTP URL address.

For more information on COS, refer to COS Product Overview.

Image
Overview

Last updated：2024-01-08 09:12:28

Images

A Tencent Cloud image provides all the information needed to launch a CVM instance. With an image, you can launch instances with similar configurations easily. Generally speaking, an image is the "installation disk" of a CVM instance.

Image Types

Tencent Cloud provides the following types of images:
Public image: Available to all users and cover major operating systems.
Custom image: Only available to the creator and the users with whom the image is shared. A custom image can be created from a running instance or imported from external sources.
Shared image: Shared by other users. You can use shared images to create instances, but cannot modify the shared images.
For more information, see Image Types.

Image Billing

For billing details of images, see Billing Overview.

Deployment with Images vs Manual Deployment

Mode/Item	Deployment with Images	Manual Deployment
Deployment duration	3 to 5 minutes	1 to 2 days
Deployment process	Quickly create a suitable CVM based on mature marketplace solutions or the already used solutions.	Select the appropriate operating system, database, application software and plugins to create a CVM, and installation and debugging is required.

Security	Public images and custom images have been tested and audited by Tencent Cloud. The sources of shared images need to be identified by users.	It depends on the development and deployment personnel.
Applicable scenarios	Public images: Genuine operating systems, which contain the initialized add-ons provided by Tencent Cloud.Custom images: Quickly create the same software environment as an existing CVM, or perform environment backup.Shared images: Quickly create the same software environment as other users' CVMs.	Configure a CVM completely by yourself, without basic settings provided.

Use Cases

Deploy a specific software environment

You can use shared images and custom images to quickly set up an environment with the specific software.

Deploy software environment in batch

By creating an image of a CVM instance, you can create multiple CVM instances with the same environment by choosing the image as the OS.

Back up a server runtime environment

You can create an image of a CVM instance to back up the running environment. If the CVM instance ceases to run properly because its software environment is damaged, this image can be used to recover the environment.

Custom Image Lifecycle

After creating or importing a custom image, you can use it to launch new instances. The diagram below shows the lifecycle of a custom image.

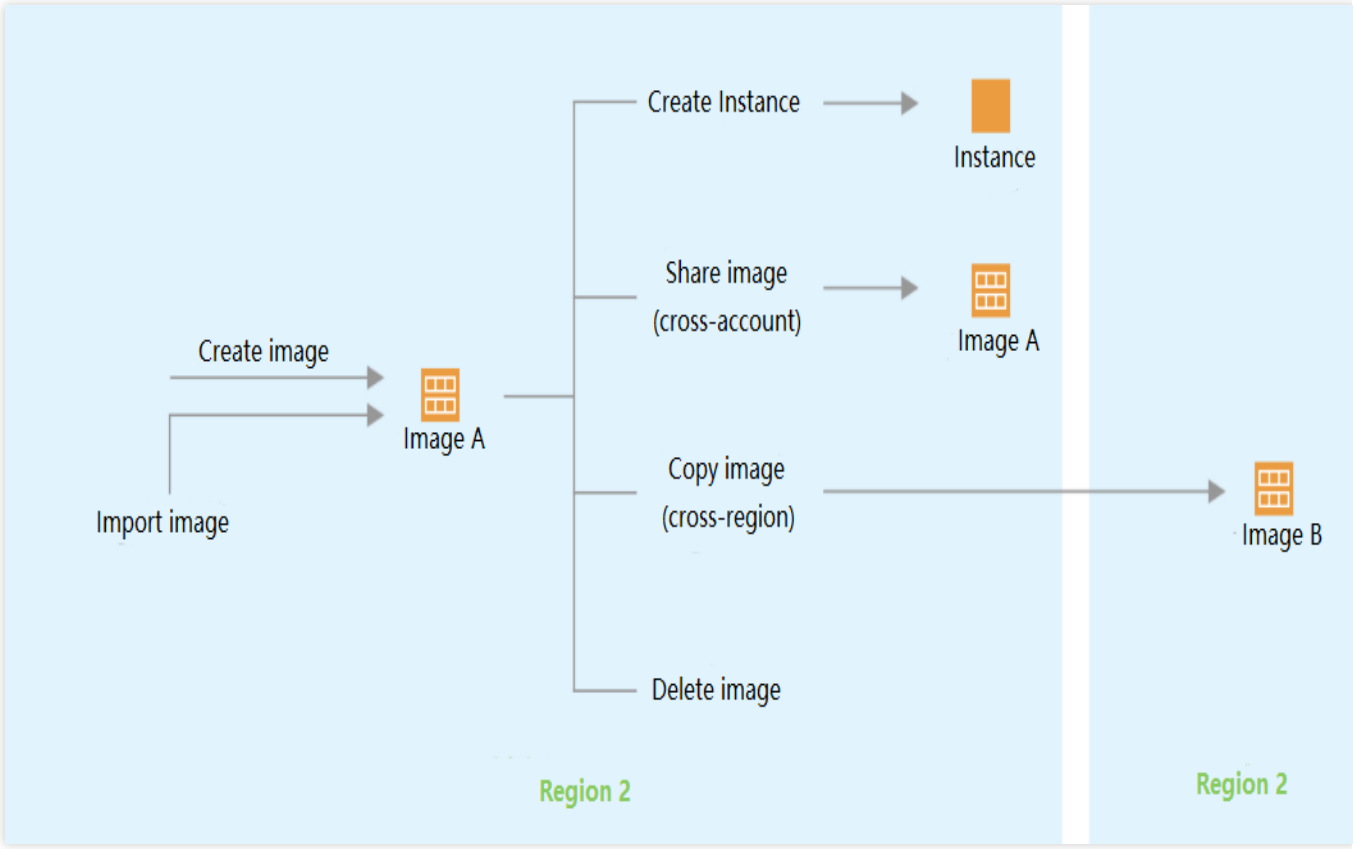

Create Instance
Instance
Share image
(cross-account)
Image A
Create image
Image A
Import image
Copy image
(cross-region)
Image B
Delete image
Region 2
Region 2

Image Types

Last updated：2024-05-16 10:23:34

You can select an image based on the following attributes:
Location (see Regions and Availability Zones)
Operating system
Architecture (32-bit or 64-bit)
Tencent Cloud provides three types of images, namely public images, custom images, and shared images.

Public Images

Public images are images provided by Tencent Cloud. Each image contains an operating system and initialization components provided by Tencent Cloud, and is available to all users.
Features:
Operating system: you can select a preferred operating system, such as Linux or Windows. Operating systems are updated regularly.
Software: public images are integrated with software packages provided by Tencent Cloud, and support multiple versions of software with full permissions, such as Java, MySQL, SQL Server, Python, Ruby, and Tomcat.
Security: all provided operating systems are licensed. All images are created by Tencent Cloud's security and OPS team, and are strictly tested. The Tencent Cloud security components are also available.
Limit: none.
Fees: all public images are free of charge, except for Windows images in certain regions outside the Chinese mainland where a license fee may be charged.
Technical support:
Tencent Linux is provided and maintained by Tencent Cloud.
For issues involving third-party images, contact the open-source community or the operating system vendor. Tencent Cloud will provide technical assistance during the troubleshooting process if needed.

Custom Images

Custom images are made by users using the image creation feature, or imported using the image import feature. Custom images are only available to creators and the people they share them with.
Features:
Use case: an image created from a CVM instance with applications deployed can be used to quickly create more instances that contain the same configuration.

Features: you can create, copy, share and terminate images.
Limitation: each region only supports a maximum of 500 custom images.
Fees: image creation may incur fees. For the actual cost, refer to the price shown when you create the image. Cross-region replication of custom image is free of charge.
For more information on custom images, see Creating Custom Images, Copying Images, Sharing Custom Images, Cancelling Image Sharing, and Importing Images.

Shared Images

Shared images are custom images shared with you by another Tencent Cloud user using the image sharing feature. These images are displayed in the same region as the original images.
Features:
Use case: help other users quickly create a CVM instance.
Features: the shared image can only be used to create CVM instances. Renaming, copying, or re-sharing the image with others are not supported.
Security: shared images are not checked by Tencent Cloud, which may pose security risks. We strongly recommend against using images from unknown sources.
Limitations: each custom image can be shared with a maximum of 500 Tencent Cloud users. Custom images can only be shared with accounts in the same region as the source account.
For more information on shared images, see Sharing Custom Images and Cancelling Image Sharing.

Public Image
Official Maintenance End Plans of Operating Systems

Last updated：2024-03-28 10:32:42

The public image maintenance plan of Tencent Cloud is consistent with the official maintenance plans of operating system platforms. This document provides links to the image version maintenance plans of these platforms.

Note:

When an OS is discontinued, any software maintenance and support, including bug fixes and feature updates, are unavailable. Please update your operating system in time or try another image tag.

Tencent Cloud Home-grown OS

TencentOS Server

Version	End of maintenance
TencentOS Server 2	2032-12-31
TencentOS Server 3	2032-12-31

Tencent Cloud will keep providing maintenance services for TencentOS Server 2 and TencentOS Server 3 till December 31, 2032.

Third-Party Operating Systems

Windows

Version	Mainstream Update End Date	Extension Update End Date
Windows Server 2022 Datacenter	2026-10-13	2031-10-14
Windows Server 2019 Datacenter	2024-01-09	2029-01-09
Windows Server 2016 Datacenter	2022-01-11	2027-01-12
Windows Server 2012 R2 Datacenter	2018-10-09	2023-10-10

For more information about official maintenance plans, visit the Microsoft Windows Server official website.

OpenCloudOS

Version	Maintenance End Date
OpenCloudOS 8	2029-12

CentOS

Version	Whole Update End Date	Maintenance Update End Date
CentOS Stream 9	2027-05-31	2027-05-31
CentOS Stream 8	2024-05-31	2024-05-31
CentOS 8	2021-12-31	2021-12-31
CentOS 7	2020-08-06	2024-06-30
CentOS 6	2017-05-10	2020-11-30

For more information about official maintenance plans, visit the CentOS official website.

Note:

CentOS plans to officially discontinue support for CentOS Linux. Tencent Cloud provides you with two alternative solutions: OpenCloudOS and TencentOS Server, both compatible with CentOS.

Ubuntu

Version	Standard Support End Date	Extension Update End Date
Ubuntu 22.04 LTS	2027-04	2032-04
Ubuntu 20.04 LTS	2025-04	2030-04
Ubuntu 18.04 LTS	2023-04	2028-04
Ubuntu 16.04 LTS	2021-04	2026-04

For more information about official maintenance plans, visit the Ubuntu official website.

Debian

Version	Support End Date	EOL-Long-Term Support End Date	EOL-Extension Long-Term Support End Date

Debian 11	2024-07	2026-06	None
Debian 10	2022-07	2024-06	2029-06-30
Debian 9	2020-07-06	2022-06-30	2027-06-30

For more information about official maintenance plans, visit the Debian official website.

Red Hat Enterprise Linux

Version	Fully Support End Date	First-Stage Maintenance Support End Date	Second-Stage Maintenance Support End Date	Extended Lifecycle Support (ELS) End Date
Red Hat Enterprise Linux 8	2024-05-31	N/A	2029-05-31	2031-05-31
Red Hat Enterprise Linux 7	2019-08-06	2020-08-06	2024-06-30	2026-06-30

For more information, see Red Hat Enterprise Linux Life Cycle.

Note:

You can select an instance model that has been certified for Red Hat Enterprise Linux to use the Red Hat Enterprise Linux image. For supported image tags and instance models, see FAQs about Red Hat Enterprise Linux Image.

AlmaLinux

Version	Maintenance End Date
AlmaLinux 8.5	2031-11

For more information about official maintenance plans, visit the AlmaLinux official website.

CoreOS

Version	Maintenance End Date
CoreOS Container Linux	2020-05-26

FreeBSD

Version	Maintenance Update End Date
FreeBSD 13.1	2023-06-30
FreeBSD 13.0	2022-08-31
FreeBSD 12.3	2023-03-31
FreeBSD 12.2	2022-03-31

For more information about official maintenance plans, visit the FreeBSD official website.

Rocky Linux

Version	Maintenance End Date
Rocky Linux 9.0	2032-05-31
Rocky Linux 8.6	2029-05-31
Rocky Linux 8.5	2029-05-31

For more information about official maintenance plans, visit the Rocky Linux official website.

OpenSUSE

Version	Maintenance Update End Date
OpenSUSE Leap 15.4	2023-11
OpenSUSE Leap 15.3	2022-11
OpenSUSE Leap 15.2	2022-01-04
OpenSUSE Leap 15.1	2021-02-02

For more information about official maintenance plans, visit the OpenSUSE official website.

TencentOS Server
TencentOS Server Introduction

Last updated：2025-01-03 16:53:32

TencentOS Server is a Linux distribution developed by Tencent Cloud for cloud scenarios. It offers specific features and performance optimizations to deliver a higher performance and a more secure and reliable runtime environment for applications in CVM instances. Independently developed and designed based on the Linux kernel on top of Tencent's technical experience accumulated in the field of operating systems over the last decade, TencentOS Server has been verified and improved by Tencent's many internal businesses for many years. It is used as the OS of more than 99% of internal servers in all Tencent businesses. Meanwhile, as Tencent has the widest variety of business ecosystems in China ranging from social networking, gaming, and payment to AI and security, TencentOS Server's core capabilities such as stability, security, compatibility, and performance have been fully proven for a long time. Compared with community OS distributions, TencentOS Server is comprehensively enhanced and optimized in terms of stability, performance, and container infrastructure. It can provide enterprises with stable and highly available services to meet their demanding workload requirements, making it a better enterprise-grade OS solution. Tencent Cloud strives to make TencentOS Server the best OS in the cloud.

Use Cases

TencentOS Server is suitable for most models, including Standard, Compute, MEM Optimized, and High IO models. It also supports CPM 2.0 and high-performance computing clusters.

Note:

If you need to use TencentOS Server to run a GPU instance, please install the corresponding GPU driver.

TencentOS Server Strengths

Safety and Stability Verified by Tens of Millions of Nodes

Tencent has a complete operation management system for TencentOS Server. It has designed an automatic downtime analysis system that can monitor anomalies in real time, proactively collect OS exception information, analyze and identify anomalies, and correct faults in real time through hot patches.

After more than 10 years of continuous improvement, TencentOS Server has been verified by over 10 million nodes running massive business operations. It has achieved enterprise-level stability with deployments reaching tens of millions.

According to Tencent's actual operation data, TencentOS Server's downtime rate is 70% lower compared to other Linux distributions, delivering an overall availability of over 99.999%.

Comprehensive Optimization for a Higher Performance

The deeply optimized high-performance TencentOS Server has been improved for various software programs in the system, improving typical business performance by more than 50%. When used with CVM and bare metal scenarios, it has higher data access efficiency, better computing, network, and storage performance, greater data security and isolation than the native community version. It also has superior high bandwidth and high packet sending and receiving capabilities, making it suitable for applications that require extensive network interactions.

User Mode Compatible With CentOS

Based on the self-developed kernel, TencentOS Server versions 2 and 3 have user-mode components that are homogeneous and compatible (sourced from the same upstream open-source software repository as CentOS, rebuilt and packaged to form a distribution that maintains 100% compatibility with the northbound software ecosystem). Applications originally running in the CentOS environment can run directly on TencentOS Server without any additional adaptation.
Meanwhile, TencentOS Server also provides a complete migration tool suite to assist in the **in-place migration** from CentOS, minimizing migration risks and costs. These tool suites have been extensively tested and applied.

Self-Developed Features for Cloud-native Scenarios

TencentOS Server combines Tencent's continuous efforts in the cloud-native field to form a full stack open-source cloud-native solution. As the foundational layer of the entire technology stack, it provides core support for the stable and efficient operation of the cloud-native full stack.
At the kernel level, TencentOS Server is designed and integrated with a complete set of original cloud-native features, forming a cloud-native kernel, which include: Tencent Cloud-native Scheduler (TCNs), Cloud-native Resource Utilization Enhancement (RUE), Quality Monitor (service quality monitor), Cloud-native SLI (specialized metrics from a container's perspective), Mbuf (normalized kernel key monitoring), Cgroupfs (container resource view isolation). A typical feature description is as follows:
Cloud-native Resource Utilization Enhancement (RUE) provides perfect and comprehensive isolation capabilities for multi-priority hybrid deployment scenarios.
Cloud-native SLI provides more precise and professional monitoring systems and metrics for cloud-native services.
Quality Monitor and Mbuf provide more timely and routine Trace capabilities for cloud-native services.

Cgroupfs provides an independent resource view for cloud-native services.
Full support for eBPF core features. eBPF capabilities are aligned with the upstream 5.18 kernel, comprehensively supporting new technology architectures and features for cloud-native scenarios, such as full support for Cilium's key features.
Memory Hierarchical Uninstall (Wujing) reduces overall memory consumption through adaptive business profiling and dynamic memory reclamation technology without affecting business operations.
Nettrace is a network packet lifecycle tracking system for complex cloud-native network scenarios. It can improve network maintenance capabilities in cloud-native scenarios.

Security Compliance With Zero-downtime Fix

TencentOS Server uses a self-developed Vulnerability Management Platform with third-party intelligence platforms and YunDing Laboratory as information inputs. Upon acquiring vulnerability intelligence information, it quickly performs analysis and reproduction, conducts targeted repair verification, and ultimately delivers hotfixes to users. Meanwhile, the Tencent Operating System team has established a clear SLA reporting mechanism for different levels of security vulnerabilities, shortening repair duration and avoiding unnecessary downtime for users.
For details, please see TencentOS Server Security Advisories.

Professional Maintenance Without Worries

In terms of the technical service system, Tencent has a relatively comprehensive technical service system with higher SLA service standards (7 x 24 by default) and rich, long-term technical service experience for large-scale server scenarios. For the technical service team, Tencent uses a self-developed kernel and has been independently developing the OS since 2010. It has accumulated a large number of kernel talents, continuous contributions to the upstream community, and strong technical influence.

TencentOS Server Image Version

Currently, Tencent Cloud offers three TencentOS Server images within their maintenance period for users to select from:

Image Version	Description
TencentOS Server 4	TencentOS Server 4 is Tencent's new generation commercial distribution version of the server operating system. The kernel and user-mode software of the upstream distribution version that TencentOS Server 4 depends on are independently evolved based on the upstream community, with independent

	selection and maintenance, no longer relying on any third-party distribution. This version realizes independent compilation and maintenance of 3000+ (BaseOS, AppStream repository) user-mode software and 8000+ (EPOL) additional software packages. It features high performance, security, and support for multiple hardware platforms, providing reliable basic environments and service capabilities for cloud-based products and businesses.
TencentOS Server 3	It is compatible with the CentOS 8 user mode and uses the tkernel4 version deeply optimized based on the community 5.4 LTS kernel.
TencentOS Server 2.4	It is compatible with the CentOS 7 user mode and uses the tkernel4 version deeply optimized based on the community 5.4 LTS kernel.

Using TencentOS Server

Use in the Cloud

You can select a public image and choose the corresponding version of OpenCloudOS when creating an instance or reinstalling the operating system for an existing one. For operation details, please see Creating Instances via CVM Purchase Page and Reinstalling System.

Services and Updates

Tencent Cloud provides maintenance and updates for each major version of TencentOS Server for over 13 years. Existing servers can be upgraded via yum to promptly complete vulnerability repairs.

Note:

If you need more information about TencentOS Server, you can consult Tencent Cloud Assistant through the mini-program.

TencentOS Server Image Update Log

Last updated：2025-03-04 14:24:16

Note

The image update history is organized by release time.

Images are released region by region. If an image is not of the latest version in the update history when you create a CVM instance, the image may have not yet been released to the region.

If an image in the image update history cannot be found in the console, the image may have not been fully open.

Image Update History

The image architectures involved in this document are all x86. The update dates and features are as follows:

TencentOS Server 3

2025-01-20

Upgraded additional user-mode software packages.

2024-12-16

Upgraded additional user-mode software packages.

Updated the kernel to v5.4.119-19.0009.56.

Incorporate configurations into pip.conf

2024-11-28

Upgraded additional user-mode software packages.

Updated the kernel to v5.4.119-19.0009.54.

2024-10-08

Upgraded additional user-mode software packages.

Downgrade the release package to tencentos-release-3.1-15.tl3.x86_64

2024-08-11

Upgraded additional user-mode software packages.

Switched yum source.

2024-06-07

Upgraded additional user-mode software packages.

Updated the kernel to v5.4.119-19.0009.44.

2024-05-22

Optimized the default crashkernel configuration of the image.

Upgraded additional user-mode software packages.

2024-03-28

Updated the kernel to v5.4.119-19.0009.40.

Upgraded other user-mode software packages.

2024-01-31

Updated user-mode software.

Updated the kernel to v0009.37.

Optimized /etc/motd display content.

No longer pre-installs MySQL related software packages by default.

Added ntp*.tencent.com in chrony.conf.

2023-09-05

Upgraded the kernel to 5.4.119-19.0009.28.

Upgraded other user mode software.

2022-11-07

Updated the user mode software.

Updated the kernel to 5.4.119-19-0009.11.

Enabled disk_setup in the cloud_init configuration file.

Modified the default open files configuration of the image.

2022-10-13

Updated the user mode software.

Updated the kernel to 5.4.119-19-0009.6.

Added the pre-installation of the kernel-debuginfo package.

Removed the pre-installation of the docker-ce package.

2022-07-21

Updated the kernel to 5.4.119-19.0009.3.

Updated other user mode software.

2022-04-19

Upgraded the Broadcom network card driver to 1.10.2-218.1.182.18.

2022-02-23

Installed java-8-konajdk-8.0.8-1.1.312.x86_64.

Updated polkit to 0.112-26.tl.1.

2022-01-28

Updated the kernel to 5.4.119-19-0009.1.

Fixed CVE-2022-0185.

2022-01-05

Updated the kernel to 5.4.119-19-0009.

Updated the default yum source domain name.

Installed docker-ce by default.

Enabled the rngd service by default.

2021-09-28

Updated the kernel to 5.4.119-19-0008.

Fixed several security vulnerabilities.

2021-06-23

Updated the kernel to 5.4.119-19-0007.

Fixed several security vulnerabilities.

2021-03-19

Updated the kernel to 5.4.32-19-0003.

Fixed several security vulnerabilities.

TencentOS Server 2.4

2024-12-23

Upgraded additional user-mode software packages.

Updated the kernel to v5.4.119-19.0009.56.

Incorporate configurations into pip.conf

2024-11-28

Upgraded additional user-mode software packages.

Updated the kernel to v5.4.119-19.0009.54.

2024-10-08

Upgraded other user-mode software packages.

2024-08-27

No longer pre-installs MLNX drivers by default

Pre-installs the openssl-devel package

Updates other user-space software packages

Fixes the logic for configuring boot entries with grubby

2024-06-07

Updated the kernel to v5.4.119-19.0009.44.

Upgraded additional user-mode software packages.

2024-05-22

Upgraded additional user-mode software packages.

Optimized the default crashkernel configuration of the image.

2024-04-03

Upgraded other user-mode software packages.

Updated the kernel to v5.4.119-19.0009.40.

Fixed an issue where the kernel-debuginfo version did not meet expectations.

2023-12-01

Updated the user mode software.

Upgraded the kernel to 5.4.119-19.0009.37.

Upgraded the MLNX driver to 5.4-3.6.8.1.

2023-09-05

Upgraded the kernel to 5.4.119-19.0009.28.

Upgraded other user mode software.

2022-11-07

Updated the user mode software.

Updated the kernel to 5.4.119-19-0009.11.

Enabled disk_setup in the cloud_init configuration file.

Modified the default open files configuration of the image.

2022-10-13

Updated the user mode software.

Updated the kernel to 5.4.119-19-0009.6.

Added default installation of the scl, devtoolset-8, and devtoolset-9 components.

2022-05-12

Updated the rdma driver to 5.4-3.1.0.0LTS.

Upgraded the kernel to 5.4.119-19.0009.3.

Updated the user mode software.

Upgraded the Broadcom network card driver to 1.10.2-218.1.182.18.

2022-02-23

Installed java-8-konajdk-8.0.8-1.1.312.x86_64.

Updated polkit to 0.112-26.tl.1.

2022-01-28

Updated the kernel to 5.4.119-19-0009.1.

Fixed CVE-2022-0185.

2022-01-11

Updated the kernel to 5.4.119-19-0009.

Fixed several security vulnerabilities.

2021-10-26

Updated the kernel to 5.4.119-19-0008.

Fixed several security vulnerabilities.

2021-07-16

Updated the kernel to 5.4.119-19-0007.

Fixed several security vulnerabilities.

2021-03-03

Launched TencentOS Server 2.4 (TK4) to the public cloud.

TencentOS Server 2.4（TK3）

2023-06

Updates were stopped. The image was taken offline in June 2023.

2022-08-31

Updated other user mode software.

Updated the settings of crashkernel= in the startup parameters.

2022-07-28

Updated the kernel to 4.14.105-19.0024.

Updated other user mode software.

2022-04-19

Upgraded the Broadcom network card driver to 1.10.2-218.1.182.18.

2022-02-23

Installed java-8-konajdk-8.0.8-1.1.312.x86_64.

2021-11-30

Enabled the rngd service by default.

Updated the user mode software to the latest version.

2021-10-27

Updated the kernel to 4.14.105-19-0022.

Updated the user mode software to the latest version.

2021-07-21

Updated the kernel to 4.14.105-19-0020.

Updated the user mode software to the latest version.

2021-04-14

Updated the kernel to 4.14.105-19-0018.

Updated the user mode software to the latest version.

2020-09-24

Updated the kernel to 4.14.105-19-0017.

Updated the user mode software to the latest version.

2020-06-11

Updated the kernel to 4.14.105-19-0016.

Updated the user mode software to the latest version.

2019-12-23

Updated the kernel to 4.14.105-19-0015.

Updated the user mode software to the latest version.

2019-08-23

Updated the kernel to 4.14.105-19-0014.

Updated the user mode software to the latest version.

2019-06-28

Launched TencentOS Server 2.4 to the public cloud.

TencentOS Server 4 Release Notes

Last updated：2025-01-21 18:01:27

2024

Update Summary	Update Date
Incorporate into the file located at /etc/rc.local -> /etc/rc.d/rc.local, and configure it with executable permissions. Install glibc-langpack-en and glibc-langpack-zh to address compatibility issues with language configurations following the subdivision of the glibc-common package.	2024-12-23
Updated the kernel to 6.6.47-12, incorporating the Hygon patch to rectify the issue preventing the Metasequoia model from booting properly. Upgraded additional user-mode software packages.	2024-11-28
Public version first conversion test rootfs uses XFS file system	2024-09-03

TencentOS Server Features

Last updated：2024-11-22 17:16:57

In cloud-native scenarios, Namespace and Cgroup provide basic support for resource isolation, but the overall isolation capability of the container is still incomplete. Particularly, some resource statistics in the /proc and /sys file systems are not containerized, causing some commonly used commands (such as free and top) that run inside the container not to accurately display the usage of container resources.
To address this issue, the TencentOS kernel introduces a cgroupfs solution, aiming to improve the display of container resources. cgroupfs offers a virtual file system, including the /proc and /sys files needed for services from the container perspective, with the directory structure consistent with the global procfs and sysfs to ensure compatibility with user tools. When these files are actually read, cgroupfs dynamically generates the corresponding view of container information through the context of a reader process.

Mounting cgroupfs File System

1. Run the following commands to mount a cgroupfs file system.

```
mount -t cgroupfs cgroupfs /cgroupfs/
```

The screenshot is shown as follows:

```
[root@localhost ~]# tree /cgroupfs/
/cgroupfs/

0 directories, 0 files
[root@localhost ~]# mount -t cgroupfs cgroupfs /cgroupfs/
[root@localhost ~]# tree -L 2 /cgroupfs/
/cgroupfs/
├── proc
│   ├── cpuinfo
│   ├── diskstats
│   ├── loadavg
│   ├── meminfo
│   ├── stat
│   ├── uptime
│   └── vmstat
└── sys
    └── devices

3 directories, 7 files
```

2. Mount the files from `cgroupfs` to a specified container by using the option of `-v`. The command for enabling the docker is as follows:

```
docker run -itd --cpus 2 --cpuset-cpus 0,1,2,4 -v
/cgroupfs/sys/devices/system/cpu/:/sys/devices/system/cpu -v
/cgroupfs/proc/cpuinfo:/proc/cpuinfo -v
/cgroupfs/proc/stat:/proc/stat -v
/cgroupfs/proc/meminfo:/proc/meminfo <image-id> /bin/bash
```

3. Then the information of cpuinfo, stat, and meminfo from the container perspective only can be viewed in the container.

Isolating pagecache

When the process file page is excessively used, a large amount of memory will be occupied, reducing the available memory for other services. This will make the memory allocation of a machine frequently fall into an insufficient path, which will easily trigger an "out of memory (OOM)" error. In the cloud-native scenarios, we extend the page cache of the kernel to achieve a page cache limit for a container, thereby limiting the page cache of a certain container.

pagecache Limit of the Complete Machine:

```
sysctl -w vm.memory_qos=1
```

```
sysctl -w vm.pagecache_limit_global=1
```

```
echo x > /proc/sys/vm/pagecache_limit_ratio
#(0 < x < 100) enables the page cache limit. When it is not 0, for example, 30,
it means the page cache is limited to 30% of the total system memory.
```

The process of dirty page reclaim is relatively time-consuming. `pagecache_limit_ignore_dirty` is used to determine whether to ignore dirty pages when the memory occupation of `page cache` is calculated. Its location is as follows:

`/proc/sys/vm/pagecache_limit_ignore_dirty.`

The default value is 1, which means to ignore dirty pages.

Set Reclaim Method for page cache:

```
/proc/sys/vm/pagecache_limit_async
```

1 means asynchronous reclaim of page cache. TencentOS will create a kernel thread of [kpclimitd], which is responsible for the reclaim of the page cache.

0 means synchronous reclaim, in which no dedicated reclaim thread will be created, and the reclaim will directly occur in the context of the process that triggers the page cache limit. The default value is 0.

pagecache Limit of Container:

In addition to supporting the pagecache limit at the global system level, TencentOS also supports the pagecache limit at the container level. The usage is as follows:

1. Enable the memory QOS:

```
sysctl -w vm.memory_qos=1
```

2. Disable the global pagecache limit:

```
sysctl -w vm.pagecache_limit_global=0
```

3. Enter the memcg where the container is located, and set the memory limit:

```
echo value > memory.limit_in_bytes
```

4. Set the maximum usage of pagecache as a percentage such as 10% of the current memory limit:

```
echo 10 > memory.pagecache.max_ratio
```

5. Set the reclaim ratio for pagecache after excess:

```
echo 5 > memory.pagecache.reclaim_ratio
```

Unified Read-Write Throttling

The original IO throttling scheme of the Linux kernel separates the read and write speeds, requiring the administrator to divide the IO bandwidth into separate read and write limits based on the service model and implement them separately. This leads to bandwidth waste. For example, if the configuration is 50 MB/s for reading and 50 MB/s for writing, but the actual IO bandwidth is 20 MB/s for reading and 50 MB/s for writing, then the bandwidth of 30 MB/s for reading will be wasted. To address this issue, TencentOS introduced a unified read-write throttling scheme. This scheme provides a unified read-write throttling configuration interface for customers in the user mode and dynamically allocates the read-write ratio based on the service traffic in the kernel mode. The usage is as follows:
Enter cgroup of the corresponding blkio service, and use the following configuration: echo MAJ:MIN VAL > FILE.
Wherein MAJ:MIN represents the device number, and FILE and VAL are as shown in the table below:

FILE	VAL
blkio.throttle.readwrite_bps_device	Total Throttle of Read-Write Bps
blkio.throttle.readwrite_iops_device	Total Throttle of Read-Write iops
blkio.throttle.readwrite_dynamic_ratio	Dynamic Prediction of Read-Write Ratio: 0: Disabled. A fixed ratio of (Read: Write - 3:1) is used. 1-5: The dynamic prediction scheme is enabled.

Throttling Support of buffer io

In the native kernel of Linux, cgroup v1 has certain defects in throttling the buffer IO. Buffer IO write-back is usually an asynchronous process. When the kernel proceeds with asynchronous dirty page flushing, it is unable to determine which blkio cgroup the IO should be submitted to, thereby unable to apply the corresponding blkio throttling policy. Based on this, TencentOS has further improved the buffer IO throttling feature under cgroup v1, making it consistent with the buffer IO throttling feature based on cgroup v2. For cgroup v1, we provide a user mode interface, allowing the binding of the mem_cgroup that the page cache belongs to with the corresponding blkio cgroup, thereby enabling the kernel to throttle the buffer IO based on the binding information.
1. To enable the throttling feature of buffer IO, two features of kernel.io_qos and kernel.io_cgv1_buff_wb need to be enabled.

```
sysctl -w kernel.io_qos=1 # Enable IO QoS feature
sysctl -w kernel.io_cgv1_buff_wb=1 # Enable Buffer IO feature (enabled by
default)
```

2. To implement buffer IO throttling for containers, it is necessary to explicitly bind the corresponding memcg of the container with cgroup of blkcg. The operation is as follows:

```
echo /sys/fs/cgroup/blkio/A > /sys/fs/cgroup/memory/A/memory.bind_blkio
```

After binding, the throttling mechanism on the buffer IO resources provided by blkio cgroup for the container can be viewed through the following interfaces:
blkio.throttle.read_bps_device
blkio.throttle.write_bps_device
blkio.throttle.write_iops_device
blkio.throttle.read_iops_device
blkio.throttle.readwrite_bps_device
blkio.throttle.readwrite_iops_device

Asynchronous fork

When a large memory service executes the fork system call to create a subprocess, the duration of the fork calling process will be relatively long, causing the service to be in the kernel mode for an extended period and unable to process the service requests. Therefore, it is particularly necessary to optimize the fork time of the kernel for this scenario.
In Linux, during the default process of the processing fork in the kernel, the parent process needs to copy numerous process metadata to the subprocess, with the most time-consuming part being the copying of page tables, which usually occupies more than 97% of the fork calling time. The design philosophy of asynchronous fork is to transfer the work of copying page tables from the parent process to the subprocess, thereby shortening the time for the parent process to call the fork system call and fall into the kernel mode, making the application return to the user mode as soon as possible and handle the service requests, thereby resolving the performance jitter issue caused by fork.
This enabling and disabling of the feature is controlled by `cgroup` , with the basic usage as follows:

```
echo 1 > <cgroup directory>/memory.async_fork # Enable the asynchronous fork
feature for the current cgroup directory
echo 0 > <cgroup directory>/memory.async_fork # Disable the asynchronous fork
feature for the current cgroup directory
```

The default value of this interface is 0, that is, disabled by default.

FAQs about TencentOS Server

Last updated：2025-01-03 16:55:14

What Is TencentOS Server?

TencentOS Server is an enterprise-level Linux server operating system launched by Tencent Cloud. TencentOS Server features self-developed kernel core technology, supports the mainstream hardware platforms, and offers high performance and high reliability to meet enterprise-level requirements.

How Long is the Lifecycle of TencentOS Server?

Below is the lifecycle of each TencentOS Server edition. Before the lifecycle ends, bugfixes and security patch updates will be provided continuously.
TencentOS Server 2.4 distribution: The regular lifecycle will last until December 31, 2029, with maintenance extended until December 31, 2032.
TencentOS Server 3 distribution: The regular lifecycle will last until December 31, 2029, with maintenance extended until December 31, 2032.
TencentOS Server 4 distribution: The regular lifecycle will last until April 30, 2033, with maintenance extended until April 30, 2036.

What are the Characteristics of TencentOS Server?

TencentOS Server has the following characteristics:
Deeply customized out-of-the-box service with no complex configuration required.
High security and compliance with support for hotfixes and zero-downtime repair.
Fully open-source, and long-term support from a powerful operations support team.
Comprehensively optimized high-performance OS specifically designed for cloud scenarios.

Compared to Community Editions of OS Such As Rocky Linux and AlmaLinux, What are the Advantages of TencentOS Server?

As an enterprise-level OS, TencentOS Server surpasses Community Editions of OS in terms of maintenance lifecycle, stability, security, performance, and the integration or support for cloud-native features and virtualization technology. Currently, TencentOS Server 2 and 3 editions are independently evolved based on the upstream LTS 5.4 version, whereas the mainstream Community Editions of OS on the market are mostly based on 3.x or 4.x version kernels. Compared to these version kernels, the 5.x series has made further improvements and enhancements in areas such as SELinux, Kernel Security Reinforcement, Efficient Memory Management, Scheduler/File System/Network Stack Optimization, and improvement of KVM/Xen Virtualization Platform. These improvements are particularly important for the high availability, data security, performance, and efficiency of cloud computing services.
Community OS is not as stable as RHEL or traditional CentOS versions that have undergone long-term testing and certification, therefore, it is not suitable for production environment applications and users with high requirements for

long-term stable operation and compatibility of OS.

How Do I Use TencentOS Server in Tencent Cloud?

Tencent Cloud currently offers three versions of TencentOS Server, You can select an image on the CVM purchase page to start using it.

Which CVM Instance Types Does TencentOS Server Support?

TencentOS Server supports all CVM instance types. You can select an image on the CVM purchase page to start using it.

How Do I Install and Upgrade Software Programs When TencentOS Server is Used?

On TencentOS Server distribution, you can run the `yum` command, as well as the `t` command that comes with TencentOS Server to manage software packages. In addition, TencentOS Server 3 supports running the `dnf` command to manage software packages.

Can I Install and Use TencentOS Server on My Local Server?

Yes. You can download the TencentOS Server distribution ISO file from Tencent Cloud's software source.
TencentOS Server 2.4
TencentOS Server 3
TencentOS Server 4
You can install and use TencentOS Server on your local server or on a Virtual Machine such as VirtualBox.

Can I View the TencentOS Server Source Code?

TencentOS Server is fully open-source. You can go to Tencent Cloud Software Source to get the source package, or use the `yumdownloader --source kernel` command in the system to obtain it.

Does TencentOS Server Support 32-bit Applications and Libraries?

No. Currently, TencentOS Server only supports the installation of some 32-bit software packages via yum.

How Does TencentOS Server Guarantee the System Security?

TencentOS Server has binary compatibility with CentOS 7 and CentOS 8. Tencent Cloud guarantees the security of TencentOS Server in the following ways:
Tencent's proprietary vulnerability scanning tool and other standard vulnerability scanning and security check tools in the industry are used for regular security scanning.
TencentOS Server cooperates with the Tencent Security team to implement system security scanning and enhancement.
CVE patches from mainstream communities are regularly assessed, and user-mode software packages are updated regularly to fix security vulnerabilities.

CWP is used to regularly check the system security, trigger user security alarms, and provide repair solutions.

OpenCloudOS
OpenCloudOS Overview

Last updated：2024-01-08 09:12:28

OpenCloudOS, an open source operating system community jointly initiated by Tencent and its partners. It integrates the advantages of vendors in software and open source ecosystem and Tencent's over 10 years of technical accumulation in the operating system and kernel fields. It supports most platforms, and is completely neutral, fully open, secure, stable and high-performance.
OpenCloudOS 8.6 is the latest official release of the OpenCloudOS community. Its basic libraries and user-mode components are fully compatible with CentOS 8. It is optimized and enhanced at the kernel level. It has been used on more than 10 million nodes. The result shows that its stability is 70% better than CentOS 8 and the performance is improved by 50% in specific scenarios.

Use Cases

OpenCloudOS works with the vast majority of production instances on the cloud, including Cloud Virtual Machine (CVM) and Cloud Bare Metal (CBM).

Note:

OpenCloudOS does not come with a preinstalled GPU driver. You need to install it on your own if necessary.

OpenCloudOS Image Tag

Tencent Cloud supports OpenCloudOS 8.6 image. The image is fully compatible with the CentOS 8 user mode and comes with the OpenCLoudOS kernel that is built based on community 5.4 LTS.

OpenCloudOS Kernel

The OpenCloudOS kernel is built based on community 5.4 LTS. It includes the latest key features of the community, as well as tailored optimizations for different business scenarios.

Using OpenCloudOS

When you create an instance or reinstall the operating system of an existing instance, select **Public images** and choose an OpenCloudOS version as needed. For more information, see Creating Instances via CVM Purchase Page and Reinstalling System.

Downloading OpenCloudOS

Download address: OpenCloudOS 8.6.

Update History

For update logs, see OpenCloudOS Image Update Log.

Services and Updates

The OpenCloudOS community provides up to 10 years of maintenance and updates for each major version of OpenCloudOS (such as OpenCloudOS 8), including the latest kernel features, vulnerability fixes, and bug fixes.

OpenCloudOS V8 Image Update Log

Last updated：2024-10-14 11:31:53

Note

The image update history is organized by release time.

Images are released region by region. If an image is not of the latest version in the update history when you create a CVM instance, the image may have not yet been released to the region.

If an image in the image update history cannot be found in the console, the image may have not been fully open. In this case, you can submit a ticket to obtain more information about the image.

2024

Updated Feature	Update Date
Upgraded the user-mode software packages.	2024-09-11
Updated the user mode to 8.10. Added virtio_scsi in /etc/dracut.conf.	2024-08-16
Upgraded the kernel to 5.4.119-20.0009.32. Upgraded the user-mode software packages.	2024-06-26
Optimized the crashkernel value configured in /etc/default/grub. Upgraded the user-mode software packages.	2024-05-07

2023

Updated Feature	**Update Date**
Rectified the i8042.noaux parameter issue. Rectified the issue of an empty keymap. Rectified the issue of certain repository URLs lacking mirrors.tencentyun.com.	2023-07-15
Upgraded the kernel to kernel-5.4.119-20.0009.20.oc8. Upgraded the user mode software. Upgraded the dracut-related software to rectify the startup exceptions of the kdump service.	2023-03-15

Launched OpenCloudOS 8.6 to the public cloud.	

2022

Updated Feature	Update Date
Disabled the firewalld\\sssd\\rngd service. Uninstalled the microcode_ctl/nss-softokn/avahi package. Set keymap. Set timezone. Set the cloudinit.target dependency for the boot of kdump. Configured mirrors.tencentyun.com as the first URL in repo. Modified the /etc/rc.d/rc.local file permission to 755. Fixed permission errors of some directories in /var/lib/.	2022-09-16
Updated the kernel version to 5.4.119-19.0010. Updated other user mode software. Updated the image timestamp.	2022-07-27
Launched OpenCloudOS 8.5 to the public cloud.	2022-03-04

OpenCloudOS V9 Image Update Log

Last updated：2024-10-14 11:30:34

Note

The image update history is organized by release time.

Images are released region by region. If an image is not of the latest version in the update history when you create a CVM instance, the image may have not yet been released to the region.

If an image in the image update history cannot be found in the console, the image may have not been fully open. In this case, you can submit a ticket to obtain more information about the image.

Image Update History

2024

Updated Feature	Update Date
Upgraded the user-mode software packages.	2024-09-10
Upgraded the kernel to 6.6.34-9. Upgraded openssh to openssh-9.3p2-12.oc9.x86_64.rpm, fixing CVE-2024-6387. Upgraded the user-mode software packages.	2024-07-09
Changed the rootfs file system to XFS. Upgraded the user-mode software packages.	2024-06-27
Upgraded the kernel to 6.6.6-2401.0.1.oc9.4. Upgraded other user-mode software packages. Fixed the issue of two duplicate URLs in the yum configuration. Optimized the crashkernel value configured in /etc/default/grub.	2024-05-07

2023

Updated Feature	Update Date
Configured dracut.conf to adapt to specific device models.	2023-08-04
Disabled irqbalance.service. Configured selinux = disable.	2023-04-19

Set locale to en_US.UTF-8.	
Rectified cloud-init. Rectified locale.	2023-03-27
Launched OpenCloudOS 9.0 to the public cloud.	2023-03-17

Snapshot

Last updated：2024-01-08 09:12:28

Overview

Real-time replica of online data

Snapshots are fully usable copies of cloud disks. When a problem occurs to a cloud disk where a snapshot has been created, you can use the snapshot to quickly restore the cloud disk to normal status. We recommend you to create a snapshot for the cloud disk before making any major changes to your businesses, so that data can be quickly restored if the business changes failed.

Persistent backup at critical milestones

Snapshots can be used as persistent backups of business data to keep the business data at milestones.

Quick business deployment

Snapshots allow you to quickly clone multiple cloud disks for quick server deployment.

Use Cases

Snapshots provide convenient and efficient data protection service, which can be used in the following business scenarios:

Daily data backup

You can use snapshots to regularly back up important business data to avoid data loss caused by incorrect operations, attacks, and viruses.

Quick data recovery

You can create snapshots before performing major operations, such as changing operating systems, upgrading applications, or migrating business data. If any problem occurs, you can use snapshots to restore the business data.

Application of multiple replicas of production data

You can create snapshots for the production data to provide near real-time data for applications such as data mining, report query, and development testing.

Quick environment deployment

You can create snapshots for a CVM instance to create a custom image, and use the custom image to create CVM instances for batch deployment with the same environment and less configuration time.

Billing

For more information about snapshot prices, see Billing Overview and Price Overview.

Quota Limits

For more information about snapshot quota limits, see Use Limits.

Snapshot Types

Manual snapshot

You can manually create a snapshot for a cloud disk at a certain point in time. This snapshot can be used to create more cloud disks with identical data, or to restore the cloud disk to the point in time when the snapshot was created. For more information, see Creating Snapshots.

Scheduled snapshot

For businesses that are updated continuously, you can use scheduled snapshots to provide continuous data backups. To achieve continuous backups of cloud disk data over a certain time period, you only need to configure a backup policy and associate it with cloud disks, which significantly enhances data security. For more information, see Scheduled Snapshot.

Note:

During snapshot creation, application data saved in the memory may not be persistently stored. As such, snapshots may not capture the latest and most complete cloud disk data. For more information, refer to Notes to ensure snapshot data consistency.

Case Review

Case 1: No manual snapshots were created before performing high-risk operations, resulting in data loss

Customer A has never created a snapshot for the cloud disk. In May 2019, an operator performed a fio test on the cloud disk. The file system was corrupted. The data was damaged and could not be recovered.

Analysis: If customer A has created a snapshot for the cloud disk before testing, the snapshot can be used to roll back data and resume business immediately after the data damage.

Case 2: No scheduled snapshots were created for important data disks, resulting in data loss

Customer B created snapshots for multiple cloud disks, except those purchased after January 2019 for cost reasons. In June 2019, a cloud disk without snapshot protection had an unrecoverable data loss due to an accidental deletion of system-layer file data.

Analysis: If customer B has created scheduled snapshots for this cloud disk, the data can be recovered to the point in

time when the last snapshot was created, thereby minimizing loss. After the incident, customer B created a snapshot for that cloud disk to enhance data protection.

Case 3: Using the scheduled snapshot to roll back data and restore business after an incorrect operation

Customer C created snapshots for all cloud disks. In May 2019, a startup exception occurred due to an incorrect operation.

Analysis: Customer C promptly restored data using the scheduled snapshot that was created two days ago, and the business remains stable.

These cases all involve data loss due to incorrect operations, but the results are different. By comparison, we can find that:

In situations where **a snapshot hasn't been created**, data is rarely recoverable when a server or cloud disk exception occurs, resulting in major loss.

In situations where **a snapshot has been created**, data can be recovered when a server or cloud disk exception occurs, minimizing loss.

We recommend regularly creating snapshots for businesses based on business types, enhancing data security and achieving low-cost, high-efficiency disaster recovery.

Others

For other questions, refer to Snapshot FAQs.

Network and Security Overview

Last updated：2024-01-08 09:12:28

Tencent Cloud offers network infrastructure and security services to ensure your business stays secure, efficient, and flexible.

Encrypted Login

Tencent Cloud provides two encrypted login methods using Login Password and SSH Key. You can use either one of these to connect to your CVM instance. Windows CVM does not support login using SSH keys.

Network access

Tencent Cloud products can communicate with one another using Internet Access or Private Network Access.
Internet Access: CMV instances and other Tencent Cloud products use Internet access to provide public-facing services. They are assigned public IP addresses to communicate with other computers on the Internet.
Private Network Access: Tencent Cloud assigns private IP addresses to resources in the same region so they can communicate with one another on the same LAN for free.

Network Environment

Tencent Cloud provides two types of Network Environments: basic network and private networks (VPC).
Basic network is a shared networking resource pool for all users. We recommend basic network for beginners.
VPCs are custom network spaces that are logically isolated. You can launch CVM instances on a predefined and customized IP range to isolate your resources from those of other users. We recommend VPCs for users who are comfortable with network administration.

Security Groups

A security group is a stateful virtual firewall capable of filtering. As an important means of network security isolation provided by Tencent Cloud, it can be used to set network access controls for one or more TencentDB instances.
The following is a list of typical security group use cases:

Create multiple security groups and populate them with different rules.
Associate one or more security groups with each of your CVM instances. The system uses these security groups to control traffic to your instances and the resources your instances can access.
Configure your security groups so that only specific IP addresses or security groups can access your instances.

Elastic Public IPs

Elastic Public IPs, or EIPs, are static IP addresses specially designed for cloud computing.
We recommend using EIPs in the following cases:
An instance can go down for unforeseeable reasons, and the failover instance needs to use the same IP address to provide uninterrupted service.
An instance does not have a public IP address but still needs a static IP address.

ENIs

An Elastic Network Interface (ENI) is an elastic network interface that can be bound to a CVM instance on a VPC to provide network connection. ENIs can be freely migrated among instances. This is essential when configuring and managing networks and implementing high availability deployments.

Placement Group

Last updated：2024-03-21 09:19:57

A placement group is a policy for distributing instances on the underlying hardware. The instances you create in the placement group feature disaster recovery and high availability when launched. Tencent Cloud CVM provides an instance placement policy, which can force the instances to be dispersed with a certain policy and reduce the impact of the underlying hardware/software failure on CVM services. You can use the placement group to deploy CVM instances on different physical servers to ensure service high availability and the underlying disaster recovery capability. When you create instances in a placement group, we will launch them in a specified region according to the deployment policy you configured in advance. If you did not configure a placement group for the instances, we will try to launch them on different physical servers to ensure service availability.

Spread Placement Group

Currently, spread placement group is supported. A spread placement group is a group of instances that are placed on different underlying hardwares and have high availability. We recommend that you use a spread placement group for applications of important instances that need to be placed separately, such as master/slave databases and high-availability clusters. By launching instances in a spread placement group, you can minimize the simultaneous failure of instances with the same underlying hardware.
A spread placement group is region specific and can be deployed across multiple availability zones. There is a limit on the number of instances that can be placed in a group. For more information, please see the Console.
Note:
When you launch instances in a spread placement group, the request will fail if you have insufficient hardware. You can wait for a while and try again.

Spread placement group rules and limits

Before using a spread placement group, note the following rules:
Placement groups cannot be merged.
An instance can only be added to one placement group.
Spread placement group can be placed on physical machines, switches, or racks.
Placement groups support setting affinity levels across three tiers of physical machines, switches, and racks, with the affinity range from 1 to 10. This represents the maximum number of instances that can be distributed at the corresponding tier. While adhering to a relatively stringent dispersion policy, this approach also allows for a degree of flexibility.

The maximum number of instances supported by a spread placement group on a physical machine, switch, or rack is different. For more information, see the official website.
If you specify and use the disaster recovery group policy, it will be strictly followed. Please note that if there is not enough hardware to distribute instances, the creation of some instances will fail.
Dedicated CVM instances cannot be added to a spread placement group.

Operation Guide

For more information on operations, see Spread Placement Group.

Network Environment

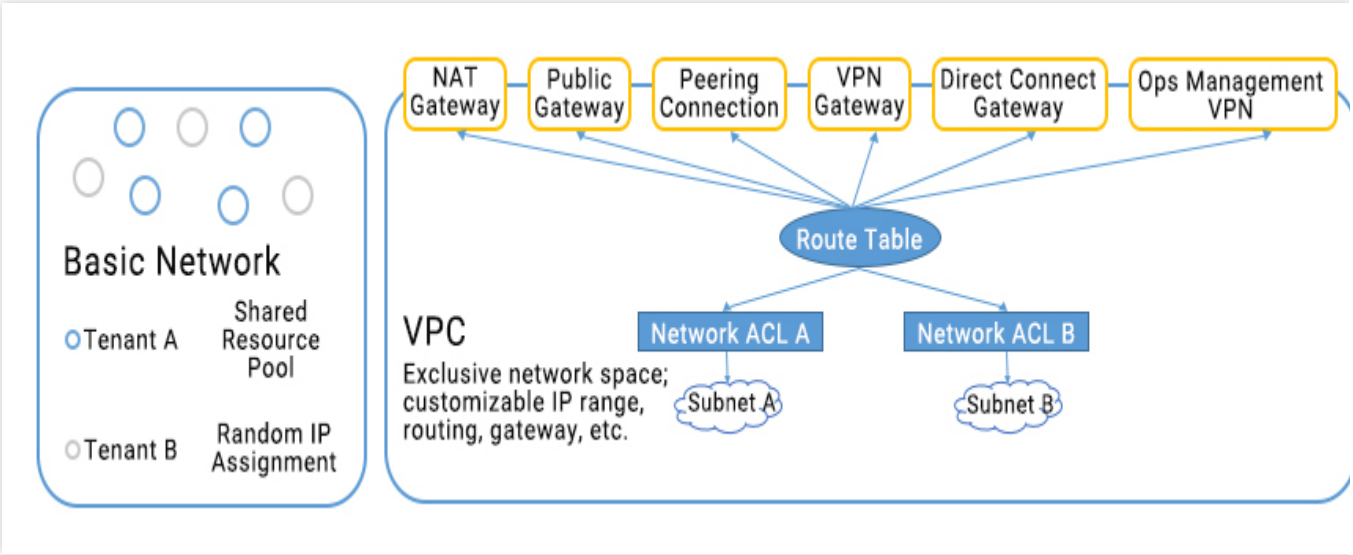
Last updated：2024-01-08 09:12:28

Tencent Cloud provides two network environments, Virtual Private Cloud (VPC) and classic network.
For Tencent Cloud accounts created after June 13, 2017, only VPC is available. We recommend you use a VPC for the following reasons:
Complete features: VPC covers all classic network features, while providing more flexible network services such as custom IP range, routing, Direct Connect, VPN, NAT, etc.
Smooth migration: There is no completely smooth migration solution in the industry (the need to shut down, change private IP, etc). If you need to use a VPC for business development later, the migration process may affect your business.

VPC and Classic Network

VPC

With Tencent Cloud's Virtual Private Cloud (VPC), you can customize a logical isolated virtual network on the cloud. Even in the same region, different VPCs cannot communicate with each other by default. Similar to the traditional network that you run in your data center, your Tencent Cloud resources, including CVM, Cloud Load Balancer, and TencentDB for MySQL, are hosted on Tencent Cloud's VPC to give you full control over the environment. For more information on configuration and application scenarios, see VPC. The VPC helps you build more complex network architecture and is suitable for those who are familiar with network management.



Classic network

Classic network is the public network resource pool for all Tencent Cloud users. All your Tencent Cloud resources will be centrally managed by Tencent Cloud.

Feature differences

Feature	Classic Network	VPC
Tenant association	Tenant association	Logical isolated network based on GRE encapsulation
Network customization	Not Supported	Supported
Routing customization	Not Supported	Supported
Custom IP	Not Supported	Supported
Interconnection rules	Interconnection is allowed for the same tenant in the same region	Cross-region and cross-account interconnection are supported
Security control	Security groups	Security groups and Network ACL

Sharing Between Classic Network and VPC

Some Tencent Cloud resources and features support both the classic network and VPC, and can be shared or accessed via the two different networks.

Resource	Description
Image	An image can be used to launch a CVM instance in any network environment
Elastic Public IP	Elastic Public IPs can bind to a CVM instance under any network environment
Instances	Instances in the classic network and VPC can communicate with each other through the Public IP or Classiclink
SSH Key	SSH key supports loading a CVM instance under any network environment
Security Groups	Security groups support binding CVM instances under any network environment

Note:

A Cloud Load Balancer can not work on both the classic network and VPC, even if the VPC and the classic network are interconnected.

Migrating from Classic Network to VPC

Please see Switch to VPC to migrate instances in the classic network to VPC.

Private Network Access

Last updated：2024-01-08 09:12:28

Private network services are LAN services, where cloud services can access each other through internal linkages. Tencent Cloud services can access each other through internet access or the private network of Tencent Cloud. Tencent Cloud data centers are interconnected with underlying networks of megabytes/gigabytes. They enable communications via private networks with large bandwidth and low latency, which are free of charge if in the same region to help you build a network architecture flexibly.

Private IP Address

Overview

Private IPs are addresses that cannot be accessed through the Internet. Each instance has a default network interface (i.e., eth0) for assigning private IPs. Private IPs can be automatically assigned by Tencent Cloud, or you can specify on your own (only for VPC environment).

Note:

If you change the private IP within the operating system, the private network may be interrupted.

Attributes

User-level isolation: The private networks of different users are isolated from each other. By default, cloud services of different accounts can not communicate over the private network.

Regional isolation: The private networks in different regions are isolated from each other. By default, cloud services under the same account in different regions can not communicate over the private network.

Application Scenarios

Private IP can be used for the communication between CLBs and CVM instances, and between CVM instances and other Tencent Cloud services (such as TencentDB).

IP Allocation

Each CVM instance is assigned a default private IP address upon the launch. The private IP varies by the network environment:

Classic network: private IP address is automatically assigned by Tencent Cloud and cannot be changed.

VPC: Tencent Cloud VPC CIDR currently allows you to use one of the following IP ranges, and the maximum:

10.0.0.0 - **10.255**.255.255 （minimum masks are /12 and /28）

172.16.0.0 - **172.31**.255.255 （minimum masks are /12 and /28）

192.168.0.0 - **192.168**.255.255 （minimum masks are /16 and /28）

Private Network DNS

DNS Server Address

Private network DNS service is used for domain name resolution. If DNS is configured incorrectly, domain name cannot be accessed.

Tencent Cloud provides reliable private network DNS servers in different regions. Specific configurations are shown below:

Network Environment	Region	Private Network DNS Server
Classic Network	Guangzhou	Guangzhou Zone 1: 10.112.65.31 10.112.65.32
		Guangzhou Zone 2: 10.112.65.31 10.112.65.32
		Guangzhou Zone 3: 10.59.218.193 10.59.218.194
		Guangzhou Zone 4: 100.121.190.140 100.121.190.141
	Shanghai	10.236.158.114 10.236.158.106
	Beijing	10.53.216.182 10.53.216.198
	North America	10.116.19.188 10.116.19.185
	Hong Kong, China	10.243.28.52 10.164.55.3
	Singapore	100.78.90.19 100.78.90.8
	Guangzhou Open Zone	10.59.218.18 10.112.65.51
	Chengdu	100.88.222.14

		100.88.222.16
	Silicon Valley	100.102.22.21 100.102.22.30
	Frankfurt	100.120.52.60 100.120.52.61
	Seoul	10.165.180.53 10.165.180.62
VPC	All Regions	183.60.83.19 183.60.82.98

Operation Guide

You can view or modify the private IP address of the instance. For detailed instructions, see:

Getting Private IP Addresses and Setting DNS

Modifying Private IP Addresses

Internet Access

Last updated：2024-01-08 09:12:28

In order to host public-facing services on your CVM instance, which involves transmitting data over the Internet, you need a public IP address. Tencent Cloud provides Internet access via high-speed connected networks of Tencent Cloud IDC. The domestic multi-line BGP network covers more than 20 network operators and the public network egress can implement cross-region switchover in a matter of seconds. This ensures that your users can enjoy a high-speed and secure network no matter which networks they use.

Public IP address

Overview: public IP addresses are non-reserved addresses on the Internet. A CVM with a public IP address can access other computers on the Internet and be accessed by other computers.
Obtaining a public IP address: set your bandwidth to more than 0 Mbps when creating your CVM instance. The system automatically picks a public IP address from the public IP address pool and assigns it to your instance. This address can be changed later. For details, refer to Change Instance Public IP.
Configuring your CVM instance: you can log in to your CVM instance from a public network to configure it. For more information, refer to Log in to Linux Instances and Log in to Windows Instances.
Translating your public IP address: you can map your public IP address to the private IP address of your CVM instance using Network Address Translation (NAT).
Retrieving public IP address information: all Tencent Cloud public network interfaces are managed by Tencent Gateway (TGW), which means all public network interfaces of all CVM instances are maintained by TGW. This process is imperceptible to CVM instances. Therefore, when you run the command `ifconfig` on your Linux CVM instance or `ipconfig` on your Windows CVM instance, you get private IP address information. For information regarding your public IP addresses, log in to your CVM console and check instance detail pages.
Billing: using public IP addresses to offer services may incur fees. For details, refer to Public Network Billing Methods.

Public IP Address Release

You cannot actively associate or release a public IP address that is already associated with an instance.
In the following cases, a public IP address associated with an instance is released or reassigned:
When you terminate an instance. Tencent Cloud releases a public IP address when you terminate an on-demand CVM instance.

When you associate or disassociate an Elastic Public IP with an instance. When you associate an EIP with your instance, the existing public IP address is released back into the public IP pool. When you disassociate an EIP from your instance, Tencent Cloud assigns a new public IP address to your instance. You will not be able to use your old public IP address.
If you need a static public IP address, use an Elastic Public IP (EIP).

Operation Guide

For detailed instructions on how to obtain and change your public IP address, refer to:
Obtaining a Public IP Address of the Instance
Changing Your Public IP Address

Elastic IP (EIP)

Last updated：2024-04-10 14:28:59

Overview

Elastic IP (EIP) is a regional static IP designed for dynamic cloud computing. With EIP, you can quickly map an address to another instance (or NAT gateway instance) in your account for failure isolation.
The lifecycle of a public IP is bound with a CVM, which means that the public IP is released when the associated CVM is terminated. However, the lifecycle of an EIP is independent of the CVM. An EIP is always available under your account unless you release it. If you want to retain a public IP after the termination of the associated CVM, convert it to an EIP.

IP Address Type

Tencent Cloud supports various types of EIPs, such as general BGP IP, premium BGP IP, accelerated IP, and static single-line IP.
General BGP IP: The domestic multi-line BGP network covers more than twenty ISPs (including the three major ISPs, CERNET, and China Broadnet). The BGP public network outbound supports switchover across regions within seconds, providing your users with high-speed and secure networks.
Premium BGP IP: Dedicated lines can avoid the use of international ISP services. The latency is lower, which effectively improves the quality of overseas services for users in Chinese Mainland.
Accelerated IP: Anycast is used for acceleration to ensure more stable and reliable public network access with a low latency.
Static single-line IP: Users can access the public network using services of a single ISP, featuring low cost and convenient scheduling.
Anti-DDoS EIP: This type of GBP IP provides Tbps-level cloud-native DDoS protection capability and should be used together with Anti-DDoS Pro for Enterprise. After the IP address is bound to business resources and Anti-DDoS Pro resources, users can enjoy the anti-DDoS capability.

IP Resource Pool

If certain adjacent IP addresses need to be reserved for your services or IP addresses in a specific network segment should be allocated, you can submit a ticket for consultation. A dedicated IP resource pool will be assigned for you. Dedicated resource pools are supported for general and premium BGP IPs and static single-line IPs currently.

For the fees, please consult your business manager.

Public IP vs EIP

Both public IPs and EIPs are used of internet access.

Public IP: It is assigned automatically upon the creation of CVM and cannot be unbound from the CVM.

EIP: It's purchased independently. It can be bound with/unbound from cloud resources (CVM, NAT gateway, ENI, and HAVIP) at any time.

Note:

Only general BGP IP lines are applicable for the current common public IP addresses.

For more details about EIPs, see IP Addresses.

Item	Public IP	EIP
Public network access	✓	✓
Independent lifecycle	×	✓
Bind/unbind anytime	×	✓
Adjust bandwidth1	✓	✓
Idle fee	×	✓

Note:

In the Public IP console, you can only adjust the bandwidth of EIPs. To adjust the bandwidth of public IPs, see Adjusting Network Configuration.

EIPs can be decoupled from the lifecycle of the cloud resource and operate independently. For example, if you need to keep a certain public IP that is strongly associated with your business, you can change it to an EIP and keep it in your account.

Rules and Limits

Use Rules

An EIP is applicable to instances in the basic network and VPC, and NAT gateway instances in VPC.

When you bind an EIP with a CVM instance, the original public IP is released at the same time.

When a CVM/NAT gateway instance is terminated, it will be disassociated from its EIP.

For details on EIP billing rules, see Elastic IP Billing.

For the step-by-step operations of EIPs, see Elastic IP.

Quota

Resource	Quota
EIPs/region	20
Daily purchase (times)	Regional EIP quota * 2
Public IPs assigned due to disassociation of EIPs/day	10

Note:

By default, the EIP quota cannot be adjusted. You can free some quota by converging IPs in NAT gateway and Cloud Load Balancer.

If you do need to adjust the quota, submit a ticket or contact your sales rep.

Limits on public IPs bound to CVM

Note:

For CVM instances purchased before 00:00, September 18, 2019, the number of public IPs can be bound to each instance is equal to the number of supported private IPs.

CPU cores	Public IPs + EIPs
1-5	2
6-11	3
12-17	4
18-23	5
24-29	6
30-35	7
36-41	8
42-47	9
≥ 48	10

Elastic Network Interface (ENI)

Last updated：2024-01-08 09:12:28

Elastic Network Interface (ENI) is an elastic network interface bound to CVMs in a VPC, which can be migrated among multiple CVMs. ENI is very useful for configuring management networks and establishing highly reliable network solutions.
ENI has VPC, availability zone and subnet attributes. You can only bind it to CVMs under the same availability zone. A CVM can be bound with multiple ENIs, and the maximum number allowed varies by CVM specifications.

Concepts

Primary ENI and secondary ENI: The ENI created when creating the CVM in a VPC is the primary ENI, and those created by users are secondary ENIs. The primary ENI does not support binding and unbinding, but secondary ENIs do.
Primary private IP: The primary private IP of an ENI is randomly assigned by the system or customized by users when the ENI is created. You can modify the primary private IP of the primary ENI, but not those of secondary ENIs.
Secondary private IP: The secondary private IPs bound to the ENI, other than the primary IP, are configured by users when they create or modify the ENI. Users can also bind/unbind these IPs.
EIP: Bound with private IPs of an ENI one by one.
Security group: An ENI can be bound with one or more security groups.
MAC address: An ENI has a globally unique MAC address.

Applications

Isolation among private network, public network and management network:
Isolation among private network, public network and management network for secure data transmission is required for the network deployment of key businesses. Different routing policies and security group policies can guarantee data security and network isolation. Like a physical server, CVM can be bound with ENIs residing in different subnets to achieve isolation among three networks.
Highly reliable application deployment:
The high availability of key components in the system architecture is ensured through multi-server hot backup. Tencent Cloud provides ENIs and private IPs that can be flexibly bound and unbound. You can configure Keepalived disaster recovery to achieve highly available deployment of key components.

Limits

Based on CPU and memory configurations, the number of ENIs bound to a CVM differs from the number of private IPs bound to an ENI. The quotes are as shown below:

Caution:

The number of IPs that can be bound to a single ENI indicates the upper limit of IPs bound only. The EIP use limits, rather than the upper limit, determines the actual EIP quota.

Number of ENIs bound to a CVM instance

Number of private IPs bound to a single ENI on CVM instances

Model	Instance Type	Number of ENIs								
		CPU: 1 core	CPU: 2 cores	CPU: 4 cores	CPU: 6 cores	CPU: 8 cores	CPU: 10 cores	CPU: 12 cores	CPU: 14 cores	CPU 16 core
Standard	Standard S5	2	4	4	-	6	-	-	-	8
	Standard Storage Optimized S5se	-	-	4	-	6	-	-	-	8
	Standard SA2	2	4	4	-	6	-	-	-	8
	Standard S4	2	4	4	-	6	-	-	-	8
	Standard Network-optimized SN3ne	2	4	4	-	6	-	8	-	8
	Standard S3	2	4	4	-	6	-	8	-	8
	Standard SA1	2	2	4	-	6	-	-	-	8
	Standard S2	2	4	4	-	6	-	8	-	8
	Standard S1	2	4	4	-	6	-	8	-	8

High IO	High IO IT5	-	-	-	-	-	-	-	-	8
	High IO IT3	-	-	-	-	-	-	-	-	8
MEM Optimized	Memory Optimized M5	2	4	4	-	6	-	8	-	8
	Memory Optimized M4	2	4	4	-	6	-	8	-	8
	Memory Optimized M3	2	4	4	-	6	-	8	-	8
	Memory Optimized M2	2	4	4	-	6	-	8	-	8
	Memory Optimized M1	2	4	4	-	6	-	8	-	8
Compute	Compute Optimized C4	-	-	4	-	6	-	-	-	8
	Compute Network-optimized CN3	-	-	4	-	6	-	-	-	8
	Compute C3	-	-	4	-	6	-	-	-	8
	Compute C2	-	-	4	-	6	-	-	-	8
GPU-based	GPU Compute GN2	-	-	-	-	-	-	-	-	-
	GPU Compute GN6	-	-	-	-	-	-	-	-	-

	GPU Compute GN6S	-	-	4	-	6	-	-	-	-
	GPU Compute GN7	-	-	4	-	6	-	-	-	-
	GPU Compute GN8	-	-	-	4	-	-	-	8	-
	GPU Compute GN10X	-	-	-	-	6	-	-	-	-
	GPU Compute GN10Xp	-	-	-	-	-	6	-	-	-
FPGA-based	FPGA Accelerated FX4	-	-	-	-	-	6	-	-	-
Big Data	Big Data D3	-	-	-	-	6	-	-	-	8
	Big Data D2	-	-	-	-	6	-	-	-	8
	Big Data D1	-	-	-	-	6	-	-	-	-
Cloud Physical Machine 2.0		Not supported								

Model	Instance Type	Number of private IPs bound to a single ENI							
		CPU: 1 core	CPU: 2 cores	CPU: 4 cores	CPU: 6 cores	CPU: 8 cores	CPU: 10 cores	CPU: 12 cores	CPU: 14 cores
Standard	Standard S5	6	10	10	-	20	-	-	-

	Standard Storage Optimized S5se	-	-	20	-	20	-	-	-
	Standard SA2	6	10	10	-	20	-	-	-
	Standard S4	6	10	10	-	20	-	-	-
	Standard Network-optimized SN3ne	6	10	10	-	20	-	30	-
	Standard S3	6	10	10	-	20	-	30	-
	Standard SA1	1 GB memory: 2>1 GB memory: 6	10	8 GB memory: 1016 GB memory: 20	-	20	-	-	-
	Standard S2	6	10	10	-	20	-	30	-
	Standard S1	6	10	10	-	20	-	30	-
High IO	High IO IT5	-	-	-	-	-	-	-	-
	High IO IT3	-	-	-	-	-	-	-	-
MEM Optimized	Memory Optimized M5	6	10	10	-	20	-	30	-
	Memory Optimized M4	6	10	10	-	20	-	30	-
	Memory Optimized M3	6	10	10	-	20	-	30	-

	Memory Optimized M2	6	10	10	-	20	-	30	-
	Memory Optimized M1	6	10	10	-	20	-	30	-
Compute	Compute Optimized C4	-	-	10	-	20	-	-	-
	Compute Network-optimized CN3	-	-	10	-	20	-	-	-
	Compute C3	-	-	10	-	20	-	-	-
	Compute C2	-	-	10	-	20	-	-	-
GPU-based	GPU Compute GN2	-	-	-	-	-	-	-	-
	GPU Compute GN6	-	-	-	-	-	-	-	-
	GPU Compute GN6S	-	-	10	-	20	-	-	-
	GPU Compute GN7	-	-	10	-	20	-	-	-
	GPU Compute GN8	-	-	-	10	-	-	-	30
	GPU Compute GN10X	-	-	-	-	20	-	-	-
	GPU	-	-	-	-	-	20	-	-

	Compute GN10Xp								
FPGA-based	FPGA Accelerated FX4	-	-	-	-	-	20	-	-
Big Data	Big Data D3	-	-	-	-	20	-	-	-
	Big Data D2	-	-	-	-	20	-	-	-
	Big Data D1	-	-	-	-	20	-	-	-
Cloud Physical Machine 2.0		Not supported							

API Overview

The APIs related to the ENI and CVM are shown in the following table. For more information, see API Category.

Feature	Action ID	Description
Create ENI	CreateNetworkInterface	Create an ENI
Assign Private IP for ENI	AssignPrivateIpAddresses	Assign a private IP for an ENI
Bind ENI to CVM	AttachNetworkInterface	Bind an ENI to a CVM

Login Password

Last updated：2024-01-08 09:12:28

The password is a unique login credential for the CVM instance. To ensure instance security, Tencent Cloud provides the following two encrypted login methods:

Password login

SSH key login

When creating the CVM, you can refer to the following documents to choose different encrypted login methods based on your CVM operating system.

Custom Configuration for Windows CVM

Custom Configuration for Linux CVM

Anyone with an instance login password can log into the CVM instance remotely through a public network address allowed by the security group. Therefore, we recommend you use a secure password and change it periodically.

For more password-related operations, please see Manage Login Password.

SSH Keys

Last updated：2024-01-08 09:12:28

To ensure the security and reliability of instances, Tencent Cloud provides users with two encrypted login methods, password and SSH key pair. This document describes how to log in with an SSH key pair.
You can select SSH key pair as the encrypted CVM login method when Customizing Linux CVM Configurations.

SSH Key Overview

We recommend **using SSH key pair** to log in to a Linux instance. A key pair contains a public key and a private key. It is generated with RSA 2048-bit encryption method.
Public key: When the SSH key pair is generated, Tencent Cloud only stores the public key, and save it in the `~/.ssh/authorized_keys` file.
Private key: You need to download the private key and keep it secretly. The private key can only be downloaded once. Tencent Cloud will not keep your private key. Anyone with your private key can access your instance. Be sure to keep it safe.
You can use the key pair to connect to CVM securely. This method is more secure than logging in with a password. You only need to specify the key pair when creating a Linux instance, or bind key pair to an existing instance, so that you can use the private key to log in to the instance without entering a password.

Features and Advantages

Compared to traditional password authentication methods, SSH key pair login has the following advantages:
SSH key pair login is more complex and difficult to brute-force.
SSH key pair login is easier to use. You can log in to instances remotely with a few simple configuration steps on the console and your local client, and do not need to enter a password when you log in again.

Use Limits

SSH key pair login is only available for Linux instances.
Each Tencent Cloud account can have up to 100 SSH key pairs.
Tencent Cloud will not retain your private key. You need to download the private key after creating an SSH key, and keep it safe.
To ensure data security, you need to shut the instance down before loading the key.
To improve CVM security, you cannot use the password login method after binding a key pair to the instance.

Use Cases

To learn how to create, bind, unbind, or delete a key, see Managing SSH Key Pairs.

To learn about how to log in to CVM instances remotely using an SSH key pair, see

Logging in to Linux Instances via Remote Login Tools

Logging in to Linux Instance via SSH Key

Cloud Workload Protection Platform

Last updated：2024-01-08 09:12:28

Overview

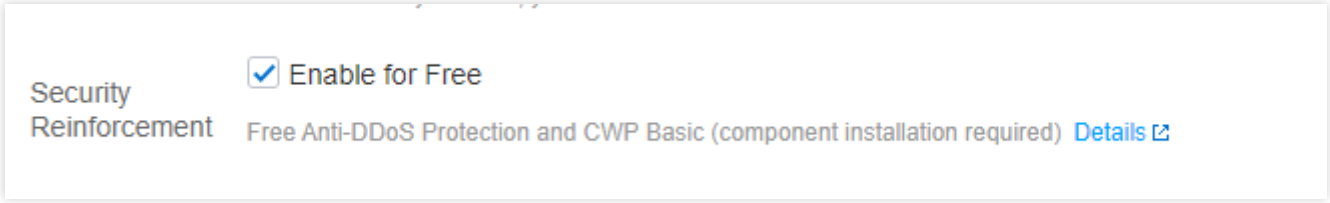
Based on Tencent Security's massive amounts of threat data, Tencent Cloud Workload Protection Platform（CWPP) leverages machine learning to provide a wide variety of security services ranging from intrusion detection to vulnerability alerting. It offers various security features such as brute force attack prevention, unusual login location reminding, trojan protection, and high-risk vulnerability detection, helping build a security protection system to cope with major network security risks faced by servers and prevent data leakage.
CWPP is available in Basic and Pro editions. When creating a CVM instance, you can choose to activate CWPP Basic by default.
Note:
For more information on the features of CWPP Basic and Pro editions and their differences, see Features in Different Editions.

Billing Mode

CCWPP Basic is free of charge.

Installing CWPP Basic

You can install CWPP Basic in one of the following methods based on the actual conditions:
Automatic installation during CVM instance creation
Manual installation for existing CVM instance
When creating a CVM instance, you can choose to activate CWPP Basic by default. On the CVM instance purchase page, select **Enable for Free** for **Security Reinforcement** to automatically install CWPP as shown below:



You can use one of the following methods to install CWPP for an existing instance based on its operating system:
Windows CVM environment

Linux CVM environment

After successful installation, you can view the security status of the CVM instance on the in the CWPP console.

See Also

Features in Different Editions

Security Overview

Anti-DDoS Basic

Last updated：2024-01-08 09:12:28

Overview

Anti-DDoS Basic provides CVM, CLB and other resources with free basic protection capability to meet your daily security protection needs. Tencent Cloud will dynamically adjust the blocking threshold according to your security reputation score. Anti-DDoS Basic is enabled by default to monitor network traffic in real time, cleanse attack traffic once detected, and start protection for Public IPs on Tencent Cloud within seconds.

Billing Mode

Anti-DDoS Basic service is free of charge. For details on other DDoS protection schemes, see Comparison of Anti-DDoS Protection Schemes.

Enabling Anti-DDoS Basic

Tencent Cloud CVM can enable Anti-DDoS Basic for free by default when creating an instance. You can click "Security Reinforcement" on the CVM instance purchase page and check "Free Enablement" to enable Anti-DDoS Basic.
After Anti-DDoS Basic is enabled, you can view the protection configuration of CVM on the Overview Page of CVM Console or Anti-DDoS Basic page of DDOS protection console.

OPS and Monitoring
Health Check

Last updated：2024-01-08 09:12:28

Overview

Health check can check the status of performance, fees, network, and disk utilization of a CVM instance. With this feature, you can discover and solve instance problems promptly.

Use Cases

We recommend you use health check in the following two scenarios:
Troubleshooting: If any failure or problem occurs during instance operations, you can run a health check to locate and troubleshoot it.
All-around instance check: During daily Ops, health check can help you keep up to date with the overall instance running status.

Check Item Description

The health check items are as detailed below:

Local network

Check Item	Description	RiskLevel	Solution
Network latency	An HTTP request is sent to check whether the network latency of the instance is high based on the following criteria: If the latency is above 600 ms, the network quality is considered poor. If no response is received within 5s, the request is considered timed out. If all requests time out, the network is considered disconnected.	Exception	Check your local network and solve the problem.
Network jitter	The average of the differences between the latency values of adjacent requests is the network jitter value. If the network	-	

	jitter/latency is below or equal to 0.15, the network is stable; otherwise, the network is jittering.		
Upstream bandwidth	Data packets are uploaded to the instance to calculate its upstream bandwidth.	-	
Downstream bandwidth	Data packets are downloaded from the instance to calculate its downstream bandwidth.	-	

Security group

Check Item	Description	RiskLevel	Solution
Common ports	Checks whether requests to common ports such as ports 22 and 3389 used by the TCP protocol for the inbound traffic are blocked in the security group.	Warning	If requests to port 22 used by the TCP protocol are blocked in an ingress rule in the security group, SSH login may be abnormal. You can open the required ports as instructed in Security Group Use Cases.

Billing

Check Item	Description	RiskLevel	Solution
Cloud disk status	Checks whether cloud disks associated with the instance have expired and whether they can be read/written.	Exception	A cloud disk associated with the instance has expired. Go to the CBS console to renew it as soon as possible.
	Checks whether cloud disks associated with the pay-as-you-go instance are unavailable due to expiration.	Warning	Auto-Renewal is not enabled for the cloud disks attached to the instance. Go to the CBS console to configure auto-renewal for the cloud disk.

Instance storage

Check Item	Description	Risk Level	Solution
High cloud disk	Checks whether the I/O	Warning	A cloud disk associated with the

latency	performance metric svctm is abnormal.		instance has a high latency. We recommend you pay attention to the cloud disk usage.
Cloud disk I/O	Checks for cloud disk I/O hang	Warning	A cloud disk associated with the instance has an I/O hang. We recommend you pay attention to the cloud disk usage.
System disk inode utilization	Checks whether the inode utilization of the cloud disk has reached 100%.	Warning	Pay attention to the cloud disk usage and troubleshoot as instructed in Kernel and I/O Issues.
System disk read-only	Checks whether the cloud disk is read-only.	Exception	
System disk space utilization	Checks whether the utilization of the cloud disk has reached 100%.	Warning	
Partition I/O utilization	Checks whether the io_util of the cloud disk has reached 100%.	Warning	

Instance status

Check Item	Description	RiskLevel	Solution
Instance shutdown	Checks whether the instance is shut down.	Warning	The instance is shut down. Go to the CVM console to start it up.
Instance restart history	Checks whether the instance has been restarted in the last 12 hours.	Warning	The instance has been restarted in the last 12 hours. Pay attention to the instance running status.
Instance kernel crash	Checks whether a hung task has occurred in the instance in the last 12 hours.	Exception	A hung task, panic, or soft deadlock has occurred in the instance in the last 12 hours. Pay attention to the instance running status and troubleshoot as instructed in Kernel and I/O Issues.
	Checks whether a panic has occurred in the instance in the last 12 hours.	Exception	

	Checks whether a soft deadlock occurred in the instance in the last 12 hours.	Exception	

Instance performance

Check Item	Description	RiskLevel	Solution
CPU utilization	Checks whether the instance has experienced a high CPU utilization in the last 12 hours.	Warning	Check your CPU utilization and adjust the configuration according. Troubleshoot it as instructed below: Windows instance: Login Failure Due To High CPU and Memory Usage (Windows) Linux instance: Login Failure Due To High CPU and Memory Usage (Linux)
Memory utilization	Checks whether the instance has experienced a high memory utilization in the last 12 hours.	Warning	
Basic CPU utilization	Checks whether the instance has experienced a high CPU utilization in the last 12 hours.	Warning	

Instance network

Check Item	Description	RiskLevel	Solution
Public EIP connection	Checks whether the EIP is isolated due to overdue payments.	Exception	An EIP may be disconnected from the public network due to overdue payments. You need to go to the Billing Center to make the overdue payment.
Existence of EIP	Checks whether the instance has an EIP.	Warning	The instance has no EIPs. If you want to use an EIP to access the public network, go to the EIP console to bind one.
EIP blocked	Checks whether the EIP is blocked due to DDoS attacks	Exception	An instance EIP is blocked due to DDoS attacks.
Public network bandwidth utilization	Checks whether the instance has	Warning	Check the network usage and troubleshoot as instructed in Login

	experienced a high public network inbound bandwidth utilization in the last 12 hours.		Failure Due to High Bandwidth Utilization.
	Checks whether the instance has experienced a high public network outbound bandwidth utilization in the last 12 hours.	Warning	
Private network bandwidth utilization	Checks whether the instance has experienced a high private network inbound bandwidth utilization in the last 12 hours.	Warning	
	Checks whether the instance has experienced a high private network outbound bandwidth utilization in the last 12 hours.	Warning	
Packet loss	Checks whether the instance has experienced TCP packet loss due to triggering of traffic throttling in the last 12 hours.	Warning	Check the traffic to your application. For more information, see Network Packet Loss.
	Checks whether the instance has experienced UDP packet loss due to triggering of traffic throttling in the last 12 hours.	Warning	
	Checks whether the instance has experienced packet loss due to a soft interrupt in the last 12 hours.	Warning	
Kernel network	Checks whether the	Warning	

conditions	instance has experienced a full UDP send buffer in the last 12 hours.		
	Checks whether the instance has experienced a full UDP receive buffer in the last 12 hours.	Warning	
	Checks whether the instance has experienced a full TCP complete connection queue in the last 12 hours.	Warning	
	Checks whether the instance has experienced TCP request overflow in the last 12 hours.	Warning	
Connection utilization	Checks whether the number of connections of the instance has reached the upper limit in the last 12 hours.	Warning	

Related Operations

You can generate an instance detection result report or view historical reports as instructed in Checking Instance Status.

Monitoring & Alarming

Last updated：2024-01-08 09:12:28

Monitoring and Alarming help ensure high reliability, availability and performance of CVMs. When you create a CVM, Cloud Monitor will be activated for free by default, through which you can analyze and implement alarms and get CVM monitoring metrics.
This document describes the CVM monitoring and alarm features. For more information, see the Cloud Monitor documentation .

Overview

CVM Monitoring and Alarming displays complete monitoring data of different CVM key metrics in real time. With this feature, you can have a comprehensive understanding of the resource usage, performance, and running status of the CVM instance. You can also configure custom alarm thresholds and notification rules.

Basic Features

You can access the following CVM monitoring and alarms features in the Cloud Monitor console:

Component	Capability	Main Features
Monitor Overview	Displays the overall monitoring information for Tencent Cloud services	Provides the overall information and alarms for Tencent Cloud services
Alarm Policy	Displays the custom alarm policy list	Supports alarm configuration for CVM
Cloud Virtual Machine	Displays the specific CVM monitoring information	Allows you to view the CVM monitoring data
Dashboard	Displays the custom monitoring dashboard	Graphically displays monitoring data to facilitate the dynamic metric analysis
Custom Monitoring	Displays the custom monitoring metrics	Allows you to view the predefined custom monitoring metrics and reported data
Traffic Monitoring	Displays the traffic monitoring	Allows you to view your overall bandwidth usage

For more information about Cloud Monitor, see Product Overview.

Overview

Daily management: log in to the Cloud Monitor console and view the running status of each monitored product.
Troubleshooting exceptions: Cloud Monitor will send you alarm notifications promptly when the monitoring data reaches the alarm threshold so that you can troubleshoot the issue.
Timely expansion: you can configure alarm policies for bandwidth, number of connections, disk utilization and other monitoring items to check the overall status of your Tencent Cloud services. You will receive alarm notifications when business surges and can expand your CVMs accordingly.

Monitoring Items

To monitor instance performance benchmarks, you should monitor at least the following items. You can go to the CVM console to obtain related monitoring information on the instance details page.

Monitoring Item	Monitoring Metric	Description
CPU usage	cpu_usage	CPU usage ratio. The data is collected and reported by the internal monitoring component of the server, making the data more accurate.
Memory utilization	mem_usage	The ratio of the actual amount of memory used by the user to the total amount of memory, excluding the memory occupied by buffer and system cache.
Private network outbound bandwidth	lan_outtraffic	Average outbound traffic per second of private ENI.
Private network inbound bandwidth	lan_intraffic	Average inbound traffic per second of private ENI.
Public network outbound bandwidth	wan_outtraffic	Average outbound traffic per second over the public network. The minimum granularity for bandwidth statistics is 10 seconds (bandwidth calculation method: total traffic in 10 seconds divided by 10 seconds).
Public network inbound bandwidth	wan_intraffic	Average inbound traffic per second of the public network.
Disk usage	disk_usage	Disk usage.
Disk I/O wait time	disk_io_await	Average wait time per disk I/O operation.

Monitoring Data

Monitoring interval: Cloud Monitor provides monitoring data at different statistical granularities, including 1 minute, 5 minutes, 1 hour, and 1 day. CVM supports the monitoring granularity of 1 minute, meaning data is collected every 1 minute. The default interval is 5 minutes.
Data storage: monitoring data at a 1-minute, 5-minute, and 1-hour granularity will be retained for 31 days; monitoring data at a 1-day granularity will be retained for half a year.
Alarm display: data is displayed in easy-to-read charts. The Cloud Monitor console displays the monitoring data of all products to give you a comprehensive overview of their running status.
Alarm settings: you can set limits for monitoring metrics. When the condition is met, alarm notification will be sent to the recipient group promptly. For more information, see Configuring Alarm Policies.
Dashboard settings: you can create a monitoring metric dashboard to dynamically analyze abnormal metrics and view the metric change in real time for prompt resource expansion. For more information, see Creating a Dashboard.

Access Management
Access Control Overview

Last updated：2024-01-08 09:12:28

If you use multiple Tencent Cloud services such as Cloud Virtual Machine (CVM), VPC, and TencentDB, it is likely they are managed by different users who all your account key. This presents the following challenges:
Your key is shared by multiple users, leading to high risk of compromise.
You cannot restrict user access to different resources, which poses a security risk due to potential misoperations.
You can use sub-accounts to solve these issues. Sub-accounts allow you to assign each user a different account so they can perform their duties without the risk of interfering with others. However, sub-accounts by default do not have the permission to access CVM instances and related resources. You need to create appropriate policies for each sub-account to grant them access.
Tencent Cloud provides a web service called Cloud Access Management (CAM) to help customers securely manage access to their resources under their Tencent Cloud accounts. CAM allows you to create, manage, or terminate users and user groups, and provides identity management and policy management to control who is allowed to access and use your Tencent Cloud resources.
You can associate CAM policies with a user or a user group to grant or deny them permissions to use specific resources to perform specific tasks. For more information on CAM policy, see Policy Syntax. For more information on how to use CAM policies, see Policies.
If you do not need to manage multi-user resource access, you can skip this section. Doing so does not affect your understanding of the rest of the article.

Getting started

A CAM policy must allow or deny one or more CVM operations, as well as the resources these actions target. A policy can also include the conditions governing the use of resources.
Some of the CVM APIs do not support resource-level permissions, which means that you must specify all resources, rather than specific resources, when performing such API operations.

Task	Link
Learn more about the basic policy structure	Policy Syntax
Define operations in the policy	CVM Operations
Define resources in the policy	CVM Resource Path
Limit the policy with conditions	CVM Condition Keys
Resource-level permissions supported by CVM	Resource-level Permissions Supported by CVM

Console samples	Console Samples

Authorizable Resource Type

Last updated：2024-01-08 09:12:28

Resource-level permission refers to the ability to specify which resources users are allowed to operate on. Cloud Virtual Machine(CVM) has partial support for resource-level permission. This means that for certain CVMs, you can control when users are allowed to operate on them, and what specific resources users are allowed to use. For example, you authorize users to perform operations on specific CVMs in Guangzhou.

The types of resources can be authorized in Cloud Access Management (CAM) are as follows:

Resource Type	Resource Description Method in Authorization Policy
CVM Instance	`qcs::cvm:$region::instance/*`
CVM Key	`qcs::cvm:$region::keypair/*`
CVM Image	`qcs::cvm:$region:$account:image/*`

CVM Instance, CVM Key and CVM Image introduce CVM API operations that currently support resource-level permission, as well as resources and condition keys supported by these CVM API operations. **When configuring the resource path,** you need to change variable parameters such as `$ region` , `$ account` into your actual parameter information. You can also use wildcard * in the path. For more information, see Operation Examples.

Note:

CVM API operations not listed in the table do not support resource-level permission. You can still authorize a user to perform these operations, but you must specify * as the resource element in the policy statement.

CVM Instance

API Operation	Resource Path
DescribeInstanceInternetBandwidthConfigs	`qcs::cvm:$region:$account:instance/*` `qcs::cvm:$region:$account:instance/$instanceId`
ModifyInstanceInternetChargeType	`qcs::cvm:$region:$account:instance/*` `qcs::cvm:$region:$account:instance/$instanceId`
ModifyInstancesAttribute	`qcs::cvm:$region:$account:instance/*` `qcs::cvm:$region:$account:instance/$instanceId`
ModifyInstancesProject	`qcs::cvm:$region:$account:instance/*` `qcs::cvm:$region:$account:instance/$instanceId`

ModifyInstancesRenewFlag	qcs::cvm:$region:$account:instance/* qcs::cvm:$region:$account:instance/$instanceId
RebootInstances	qcs::cvm:$region:$account:instance/* qcs::cvm:$region:$account:instance/$instanceId
RenewInstances	qcs::cvm:$region:$account:instance/* qcs::cvm:$region:$account:instance/$instanceId
ResetInstance	qcs::cvm:$region:$account:instance/* qcs::cvm:$region:$account:instance/$instanceId qcs::cvm:$region:$account:image/* qcs::cvm:$region:$account:image/$imageId qcs::cvm:$region:$account:keypair/* qcs::cvm:$region:$account:keypair/$keyId qcs::cvm:$region:$account:systemdisk/*
ResetInstancesInternetMaxBandwidth	qcs::cvm:$region:$account:instance/* qcs::cvm:$region:$account:instance/$instanceId
ResetInstancesPassword	qcs::cvm:$region:$account:instance/* qcs::cvm:$region:$account:instance/$instanceId
ResetInstancesType	qcs::cvm:$region:$account:instance/* qcs::cvm:$region:$account:instance/$instanceId
ResizeInstanceDisks	qcs::cvm:$region:$account:instance/* qcs::cvm:$region:$account:instance/$instanceId
RunInstances	qcs::cvm:$region:$account:instance/* qcs::cvm:$region:$account:image/* qcs::cvm:$region:$account:image/$imageId qcs::cvm:$region:$account:keypair/* qcs::cvm:$region:$account:keypair/$keyId qcs::cvm:$region:$account:sg/* qcs::cvm:$region:$account:sg/$sgId qcs::vpc:$region:$account:subnet/*

	qcs::vpc:$region:$account:subnet/$subnetId qcs::cvm:$region:$account:systemdisk/* qcs::cvm:$region:$account:datadisk/* qcs::vpc:$region:$account:vpc/* qcs::vpc:$region:$account:vpc/$vpcId
StartInstances	qcs::cvm:$region:$account:instance/* qcs::cvm:$region:$account:instance/$instanceId
StopInstances	qcs::cvm:$region:$account:instance/* qcs::cvm:$region:$account:instance/$instanceId
TerminateInstances	qcs::cvm:$region:$account:instance/* qcs::cvm:$region:$account:instance/$instanceId

CVM Key

API Operation	Resource Path	Conditio Key
AssociateInstancesKeyPairs	qcs::cvm:$region:$account:instance/* qcs::cvm:$region:$account:instance/$instanceId qcs::cvm:$region:$account:keypair/* qcs::cvm:$region:$account:keypair/$keyId	-
CreateKeyPair	qcs::cvm:$region:$account:keypair/*	-
DeleteKeyPairs	qcs::cvm:$region:$account:keypair/* qcs::cvm:$region:$account:keypair/$keyId	-
DescribeKeyPairs	qcs::cvm:$region:$account:keypair/*	-
DescribeKeyPairsAttribute	qcs::cvm:$region:$account:keypair/* qcs::cvm:$region:$account:keypair/$keyId	-
DisassociateInstancesKeyPairs	qcs::cvm:$region:$account:instance/* qcs::cvm:$region:$account:instance/$instanceId qcs::cvm:$region:$account:keypair/* qcs::cvm:$region:$account:keypair/$keyId	-
ImportKeyPair	qcs::cvm:$region:$account:keypair/*	-
ModifyKeyPairAttribute	qcs::cvm:$region:$account:keypair/*	-

	qcs::cvm:$region:$account:keypair/$keyId	

CVM Image

API Operation	Resource Path	Conditi Key
CreateImage	qcs::cvm:$region:$account:instance/* qcs::cvm:$region:$account:instance/$instanceId qcs::cvm:$region:$account:image/*	cvm:req
DeleteImages	qcs::cvm:$region:$account:image/* qcs::cvm:$region:$account:image/$imageId	cvm:req
DescribeImages	qcs::cvm:$region:$account:image/*	cvm:req
DescribeImagesAttribute	qcs::cvm:$region:$account:image/* qcs::cvm:$region:$account:image/$imageId	cvm:req
DescribeImageSharePermission	qcs::cvm:$region:$account:image/*	cvm:req
ModifyImageAttribute	qcs::cvm:$region:$account:image/* qcs::cvm:$region:$account:image/$imageId	cvm:req
ModifyImageSharePermission	qcs::cvm:$region:$account:image/* qcs::cvm:$region:$account:image/$imageId	cvm:req
SyncImages	qcs::cvm:$region:$account:image/* qcs::cvm:$region:$account:image/$imageId	cvm:req

Authorization Policy Syntax

Last updated：2024-01-08 09:12:28

Policy Syntax

CAM policy:

```
{
        "version":"2.0",
        "statement":
        [
           {
              "effect":"effect",
              "action":["action"],
              "resource":["resource"],
               "condition": {"key":{"value"}}
           }
       ]
}
```

version is required. Currently, only "2.0" is supported.
statement describes the details of one or more permissions. This element contains a permission or permission set consisting of other elements such as effect, action, resource, and condition. One policy has only one statement.
1.1 **effect** describes whether the result produced by the statement is "allowed" (allow) or "denied" (deny). This element is required.
1.2 **Action** describes the allowed or denied actions. An action can be an API (described using the prefix "name") or a feature set (a set of specific APIs, described using the prefix "permit"). This element is required.
1.3 **resource** describes the authorization details. A resource is described in a six-piece format. Detailed resource definitions vary by product. For more information on how to specify a resource, see the documentation for the relevant product. This element is required.
1.4 **condition** describes the condition for the policy to take effect. A condition consists of operator, action key, and action value. A condition value may contain information such as time and IP address. Some services allow you to specify additional values in a condition. This element is optional.

CVM Operations

A CAM policy allows you to perform API operations in any Tencent Cloud service that supports CAM. For CVM, use the prefix `name/cvm:` with any API, such as `name/cvm:RunInstances` or `name/cvm:ResetInstancesPassword` .
To specify multiple actions in a single statement, separate them with commas, as shown below:

```
"action":["name/cvm:action1","name/cvm:action2"]
```

You can also specify multiple actions using a wildcard. For example, you can specify all APIs whose names begin with "Describe", as shown below:

```
"action":["name/cvm:Describe*"]
```

To specify all CVM operations, use the wildcard "*" as follows:

```
"action": ["name/cvm:*"]
```

CVM Resource Path

Each CAM policy defines its own resources.

The general format of resource paths is as follows:

```
qcs:project_id:service_type:region:account:resource
```

project_id: project information, which is only used for compatibility purposes and can be left blank.

service_type: abbreviation of a product, such as CVM.

region: region of the resource, such as bj.

account: the root account of the resource owner, such as uin/164256472.

resource: detailed resource information of each product, such as instance/instance_id1 or instance/*.

For example, you can specify a specific instance (i-15931881scv4) in the statement as follows:

```
"resource":[ "qcs::cvm:bj:uin/164256472:instance/i-15931881scv4"]
```

You can also use the wildcard "*" to specify all instances that belong to a specific account as shown below:

```
"resource":[ "qcs::redis:bj:uin/164256472:instance/*"]
```

If you want to specify all resources or if any API operation does not support resource-level permissions, you can use wildcard "*" in `resource` as shown below:

```
"resource":["*"]
```

To specify multiple resources in one instruction, separate them with commas. In the following example, two resources are specified:

```
"resource":["resource1","resource2"]
```

The following table describes CVM resources and the corresponding resource description methods.

In the following table, names with the prefix $ are placeholders.

$project is the ID of the project.

$region is the region of the resource.
$account is the ID of the account.

Resource	Syntax
Instance	qcs::cvm:$region:$account:instance/$instanceId
Key	qcs::cvm:$region:$account:keypair/$keyId
VPC	qcs::vpc:$region:$account:vpc/$vpcId
Subnet	qcs::vpc:$region:$account:subnet/$subnetId
Image	qcs::cvm:$region:$account:image/*
CBS	qcs::cvm:$region:$account:volume/$diskid
Security group	qcs::cvm:$region:$account:sg/$sgId
EIP	qcs::cvm:$region:$account:eip/*

CVM Condition Keys

You can use conditions to specify the conditions under which policies take effect. Each condition consists of one or more key pairs. These are not case-sensitive.
If you specify multiple conditions or multiple keys in one condition, they are connected with the logical operator "AND".
If you specify a key with multiple values in one condition, they are connected with the logical operator "OR".
The following table describes CVM condition keys for specific services.

Condition key	Reference type	Key pair
cvm:instance_type	String	cvm:instance_type= `instance_type` `instance_type` is the model of the CVM instance, such as S1.SMALL1.
cvm:image_type	String	cvm:image_type= `image_type` `image_type` is the type of the image, such as IMAGE_PUBLIC.
vpc:region	String	vpc:region= `region` `region` is the region of the CVM instance, such as ap-guangzhou.
cvm:disk_size	Integer	cvm:disk_size= `disk_size` `disk_size` is the size of the disk, such as 500.

cvm:disk_type	String	cvm_disk_type= `disk_type` `disk_type` is the type of the disk, such as CLOUD_BASIC.
cvm:region	String	cvm:region= `region` `region` is the region of the CVM instance, such as ap-guangzhou.

APIs Supporting CAM

Last updated：2024-01-08 09:12:28

Basic information

Product Name	Abbreviation	Authorization Granularity
Cloud Virtual Machine	CVM	Resource-level

Note:
Three authorization granularity levels are supported by the Tencent Cloud product: service level, operation level, and resource level.
Service level: It defines whether a user has the permission to access the service as a whole. A user can have either full access or no access to the service. The Tencent Cloud product at the service level doesn't support authorization of specific APIs.
Operation level: It defines whether a user has the permission to call a specific API of the service. For example, granting an account read-only access to the CVM service is an authorization at the operation level.
Resource level: It is the finest authorization granularity which defines whether a user has the permission to access specific resources. For example, granting an account read/write access to a specific CVM instance is an authorization at the resource level. The Tencent Cloud product that supports resource-level API authorization supports resource-level authorization.

API authorization granularity

Resource-level API: It supports the authorization of a specific resource.
Operation-level API: It doesn't support the authorization of a specific resource.
For authentication through a resource-level API, the Tencent Cloud product will deliver the six-segment format of the specific resource to CAM for authentication, thereby supporting authorization and authentication of the specific resource.
For authentication through an operation-level API, the Tencent Cloud product will not deliver the six-segment format of the specific resource to CAM for authentication. Instead, it will deliver the `*` of any resource. To be more specific, if a resource is specified in the policy syntax during authorization but not delivered through the API for authentication, CAM will identify the API as out of the authorization scope, that is, having no permission.