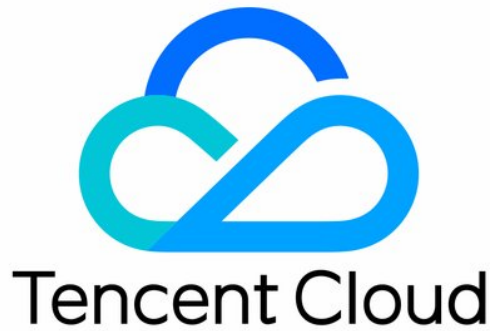


# TencentDB for Redis®

## Product Introduction



## Copyright Notice

©2013–2024 Tencent Cloud. All rights reserved.

The complete copyright of this document, including all text, data, images, and other content, is solely and exclusively owned by Tencent Cloud Computing (Beijing) Co., Ltd. ("Tencent Cloud"); Without prior explicit written permission from Tencent Cloud, no entity shall reproduce, modify, use, plagiarize, or disseminate the entire or partial content of this document in any form. Such actions constitute an infringement of Tencent Cloud's copyright, and Tencent Cloud will take legal measures to pursue liability under the applicable laws.

## Trademark Notice



This trademark and its related service trademarks are owned by Tencent Cloud Computing (Beijing) Co., Ltd. and its affiliated companies ("Tencent Cloud"). The trademarks of third parties mentioned in this document are the property of their respective owners under the applicable laws. Without the written permission of Tencent Cloud and the relevant trademark rights owners, no entity shall use, reproduce, modify, disseminate, or copy the trademarks as mentioned above in any way. Any such actions will constitute an infringement of Tencent Cloud's and the relevant owners' trademark rights, and Tencent Cloud will take legal measures to pursue liability under the applicable laws.

## Service Notice

This document provides an overview of the as-is details of Tencent Cloud's products and services in their entirety or part. The descriptions of certain products and services may be subject to adjustments from time to time.

The commercial contract concluded by you and Tencent Cloud will provide the specific types of Tencent Cloud products and services you purchase and the service standards. Unless otherwise agreed upon by both parties, Tencent Cloud does not make any explicit or implied commitments or warranties regarding the content of this document.

## Contact Us

We are committed to providing personalized pre-sales consultation and technical after-sale support. Don't hesitate to contact us at 4009100100 or 95716 for any inquiries or concerns.

# Contents

## Product Introduction

Overview

Features

Storage Engine

Product Series

Memory Edition (Standard Architecture)

Memory Edition (Cluster Architecture)

Performance

Read/Write Separation

Multi-AZ Deployment

Regions and AZs

# Product Introduction

## Overview

Last updated: 2024-11-01 15:53:51

TencentDB for Redis®, a cache database provided by Tencent Cloud, is compatible with Redis protocol and features high availability, reliability, and elasticity. It supports Redis 2.8, 4.0, 5.0, and 6.2 protocols, offering both Standard and Cluster Editions. With a maximum storage capacity of 8 TB and support for millions of concurrent requests, it meets the diverse needs of businesses in caching, storage, and computing scenarios.

[Watch video](#)

## Relevant Concepts

TencentDB for Redis® generally involves the following concepts:

- **Instance:** A database environment running independently in Tencent Cloud. One database instance can contain multiple user-created databases.
- **Virtual Private Cloud:** A custom virtual network space, logically isolated from other resources.
- **Security Group:** A security group manages access to a Redis instance. It specifies which IP addresses can access the instance through which ports using which protocols.
- **Regions and Availability Zones:** The physical locations of Redis instances and other resources.
- **Tencent Cloud Console:** A web-based user interface.
- **Project:** A feature developed to help users better manage their cloud products, primarily organized by projects. It enables project management by assigning various cloud products to different projects.
- **Read/Write Separation:** TencentDB for Redis® supports enabling and disabling read/write separation for business scenarios with more reads but less writes, addressing read requests concentrating on hot data. It supports up to 1-master 5-replica mode, providing up to 5 times the read performance expansion capability.

## Relevant Products

TencentDB for Redis® generally involves the following products:

- You can deploy your computing services by purchasing Cloud Virtual Machine (CVM) instances. For more information, see [CVM](#).

- You can use Tencent Cloud Observability Platform (TCOP) to monitor the running status of your TencentDB for Redis® instances. For more information, see [TCOP](#).
- You can write code to call TencentCloud APIs to access Tencent Cloud products and services. For more information, see the [TencentCloud API](#) documentation.

# Features

Last updated: 2024-11-01 15:54:07

## Features

### Ease of use

- **Open-source compatibility:** TencentDB for Redis® is fully compatible with the open-source Redis protocol, so you can directly use Redis clients to communicate with TencentDB for Redis® instances without any code modifications.
- **Installation-Free Use:** You can directly apply TencentDB for Redis® instances on the purchase page and select the desired system architecture, with no need to install them on your own.
- **Cloud migration:** TencentDB for Redis® supports cold and hot data migration in various self-built database environments such as self-built Tencent Cloud-based, VPN-based, Direct Connect-based, and IDC-based environments.

### Rich specifications

TencentDB for Redis® offers a choice of 0.25 GB–8 TB capacity specifications available in Standard and Cluster Editions.

- Memory Edition (Standard Architecture) supports 1–5 replicas to meet the different requirements for availability and performance of your business in different scenarios. If the number of replicas is greater than or equal to 1, read/write separation can be enabled to extend the read performance through replica nodes.
- TencentDB for Redis® Memory Edition (Cluster Architecture) supports horizontal scaling of 1–128 shards and replica scaling of 1–5 replica, where the scaling and specification adjustment are virtually imperceptible to the business, maximizing the system availability.

### Ultra-High Performance

- **High throughput:** The Standard Edition has a performance of up to 100,000+ QPS for concurrent response, and the Cluster Edition supports up to tens of millions of QPS as the number of shards increases. Their ultra-high performance can perfectly meet the needs in most business scenarios ranging from gaming, mobile apps to advertising and ecommerce.
- **Low latency:** TencentDB for Redis® delivers stable low-latency response, where 99% requests can be responded to within 1 ms or 10 ms under 50% or 90% load respectively.

- **Read expansion:** In hot key scenarios, TencentDB for Redis® allows dynamically increasing replicas to expand the read performance. It supports up to 5 replicas to read up to 500,000 hot keys.
- **Write expansion:** The performance of the Cluster Edition equals to the shard quantity multiplied by the performance per shard and increases linearly as the shard quantity increases between 3 and 128.

## Flexible architecture

TencentDB for Redis® supports multiple system architectures to meet the needs of various business scenarios. You can deploy the appropriate architecture based on your actual business requirements.

- **Standard architecture:** It refers to the edition that supports one or more replicas (nodes other than the master node) and is the most common Redis edition.
- **Cluster architecture:** It is a new edition of Redis built by Tencent Cloud based on Community Edition of Redis Cluster. It uses a distributed architecture to enable elastic scaling and features high flexibility, availability, and performance of tens of millions of QPS.
- **Architecture upgrade:** TencentDB for Redis® allows you to upgrade from standard architecture to cluster architecture if the performance and capacity of standard architecture are insufficient.
- **Read/Write separation:** Automatic read/write separation is implemented at the proxy layer. The proxy writes data only to the master node, while read requests are automatically routed to the configured read-only nodes by load balancing.
- **Multi-AZ deployment:** It refers to the disaster recovery architecture where the master and replica nodes of an instance are deployed in different AZs in the same region to deliver a higher availability.
- **Global replication:** Based on the original master-replica replication scheme, a new log file is added for remote replication to ensure the eventual data consistency for instances in different regions in the replication group. This solves the problem of cross-region data inconsistency and offers region-level disaster recovery capabilities.
- **Automatic Disaster Recovery:** TencentDB for Redis® adopts a master/replica hot backup architecture. In case of failure of the master, the access can be switched to the replica in a matter of seconds. The switch process does not require any operations at your side, reducing the labor and time costs of developing a master/replica system architecture.

## Data Security

- **RDB persistence:** The TencentDB backend service periodically creates snapshots for the data stored in the storage engine according to the backup policy to generate RDB files and

then saves them in disks for data persistence.

- **Network security protection:** TencentDB for Redis® supports configuring security groups in VPCs to implement allowlist-enabled network access control, which ensures the security and reliability of network environments.
- **Account authentication:** TencentDB for Redis® allows you to authorize root accounts, sub-accounts, or across accounts, so you can manage your resources at a fine-grained level and benefit from enterprise-grade security protection.
- **Data isolation:** TencentDB for Redis® supports multiple layers of data isolation between different regions, AZs, networks, and account levels to ensure the data security and integrity.

## Smart Ops

- **System Monitoring:** TencentDB for Redis® offers seamless and unobtrusive system monitoring capabilities. With the Tencent Cloud Observability Platform (TCOP), you can set alarm rules for over 30 automated monitoring metrics, keeping you informed about the operational status of your Redis instances for quick issue resolution.
- **TencentDB for DBbrain:** This intelligent database management tool can monitor and diagnose database instance anomalies in real-time, including slow log analysis, big key and hot key analysis, latency analysis, and more. It generates health reports and provides expert-level optimization suggestions to help you improve database performance promptly.
- **Management APIs:** TencentDB for Redis® offers a comprehensive suite of management APIs, enabling a range of self-service resource management and operational tasks, including instance provisioning, scaling, configuration, and maintenance.

## Comparison with Self-Built Databases

TencentDB for Redis® provides the capabilities of NoSQL databases as a service, which has great strengths in terms of flexibility, ease of use, high availability, fully managed Ops, data security, and data reliability.

Comparative Items	Self-Built Redis	TencentDB for Redis®
Performance	80,000-100,000/shard	80,000-100,000/shard
Cost	For a self-built database, purchasing a single storage server can be expensive, and	Opting for TencentDB instances eliminates the need for hardware and software investments and offers a range of specification choices. You can apply for

	<p>constructing a high-availability master-slave architecture requires three servers, potentially leading to resource redundancy and waste. Furthermore, the need to hire professional database administrators significantly increases labor costs.</p>	<p>the necessary resources based on your actual business needs, maximizing resource utilization. Additionally, you can access free management, backup, cloning, monitoring, and alert features. Consequently, Tencent Cloud presents an advantageous cost-performance ratio with competitively priced resources.</p>
<b>HA</b>	> 60s	< 60s
<b>Availability</b>	<p>You need to fix failures and build master/replica cluster architecture on your own.</p>	<ul style="list-style-type: none"> <li>Hot backup is supported based on two or more servers, with automatic disaster recovery, failover, and imperceptible migration features available, delivering a 99.95% availability.</li> <li>Read preference from replica databases is supported to ensure high concurrent read capability.</li> </ul>
<b>Scaling up/down</b>	<p>Scaling necessitates independently managing tasks like hardware procurement, data center hosting, and application redeployment, all of which contribute to a prolonged process. Additionally, processes like double-writing and data migration can lead to brief service interruptions during transitions, posing a higher risk of data migration errors.</p>	<p>Automated features enable one-click, on-demand scaling and rapid deployment. This ensures continuous operation with no interruptions, no downtime, and no restriction to read-only mode.</p>
<b>Read/Write Separation</b>	<p>Separate development is required.</p>	<p>Automatic read/write separation is supported and can be enabled/disabled based on the business needs.</p>
<b>Isolation</b>	<p>CPU and memory isolation</p>	<p>CPU, memory, traffic, and connection isolation are supported.</p>

<b>Multi-account</b>	Unavailable	Support Multi-account.
<b>Blocklist</b>	Unavailable	High-risk commands can be disabled.
<b>Monitoring</b>	Only a few monitoring metrics are available, which are collected by the server.	There are over 30 monitoring metrics for clusters and nodes available, including exclusive ones for latency, traffic, big key, and hot key. Alarms can be triggered for them, so you can perform Ops tasks proactively to prevent risks.

# Storage Engine

Last updated: 2024-02-27 11:47:18

## Memory Edition Engine

The Memory Edition Engine offers a native Redis experience with extensive scenario support. The Redis Memory Edition supports both standard and cluster deployment architectures, catering to various business requirements of users.

Editions supported by the Memory Edition engine include:

- **Memory Edition (Standard Architecture)**: When the number of replicas is greater than 0, the primary node (Master) and replica node (Slave) data are synchronized in real-time. In case of primary node failure, the system automatically switches to the replica node within seconds, taking over the business operations seamlessly and without any impact. The primary-secondary architecture ensures high availability of system services, offering specifications ranging from 0.25GB to 64GB.
- **Memory Edition (Cluster Architecture)**: Cluster instances employ a distributed architecture, allowing flexible selection of shard quantity, shard capacity, and replica quantity. They provide seamless scaling and descaling services for businesses, offering specifications ranging from 2GB to 8TB and supporting tens of millions of QPS performance.

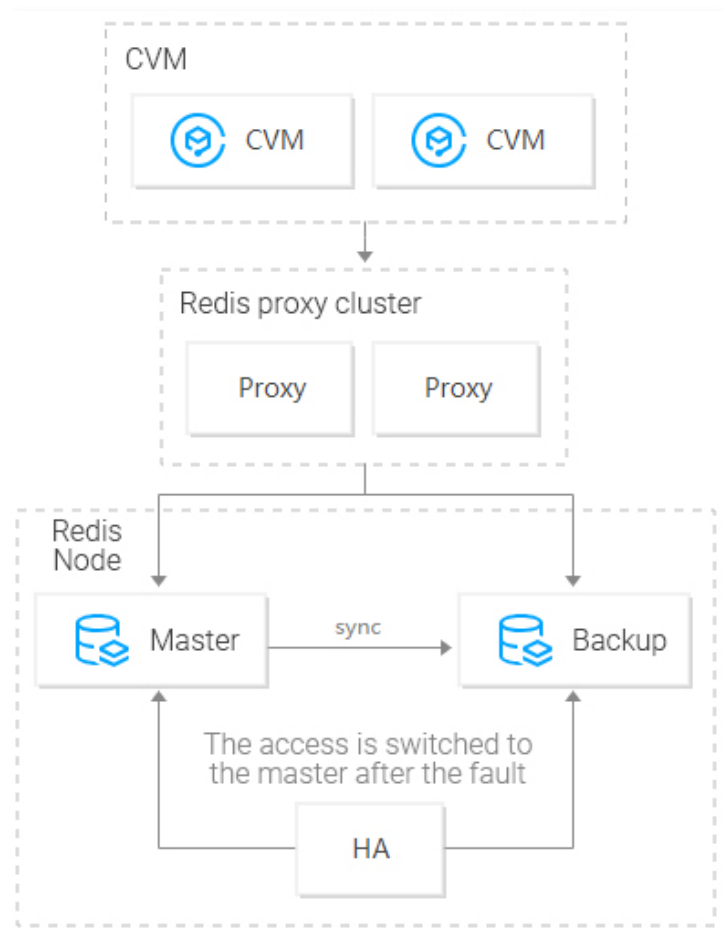
# Product Series

## Memory Edition (Standard Architecture)

Last updated: 2024-11-01 17:02:08

TencentDB for Redis® Memory Edition (Standard Architecture) refers to the version that supports one or more replicas (replicas being the non-master nodes) and is the most commonly used Redis edition, compatible with the protocols and commands of Redis 2.8, 4.0, 5.0, and 6.2, offering data persistence and backup, ideal for scenarios that demand both data reliability and availability. The master node serves daily access, while the replica nodes ensure high availability (HA). In case of a master node failure, the system automatically switches to a replica node, ensuring smooth business operations.

Taking 1 replica as an example, Memory Edition (Standard Architecture) is as follows:



## Replica Description

Memory Edition (Standard Architecture) supports 1-5 replicas to meet the different requirements for availability and performance of your business in different scenarios. All replicas of Memory Edition (Standard Architecture) play a role in supporting system's high availability, so the more replicas, the higher the availability. If the number of replicas is greater than or equal to 1, read/write separation can be enabled to extend the read performance through replica nodes.

#### Glossary:

- **Master node:** A Redis node that provides read and write capabilities.
- **Replica node:** A Redis node that provides high availability or read-only capability. A master node cannot be a replica node.

#### Replica support:

Instance Version	Supported Replica Quantity	Read/Write Separation
2.8 Memory Edition (Standard Architecture)	1	Unavailable
4.0 Memory Edition (Standard Architecture)	1 - 5	This feature is supported.
5.0 Memory Edition (Standard Architecture)	1 - 5	This feature is supported.
6.2 Memory Edition (Standard Architecture)	1 - 5	This feature is supported.

#### Read-Only Replica (Read/Write Separation):

- **Supported editions:** TencentDB for Redis® 4.0 Memory Edition (Standard Architecture) and later instance. When the number of replicas is greater than or equal to 1, automatic read/write separation can be enabled to extend the read performance vertically. Up to 5 replica nodes can be supported.
- **Implementation Principle:** After enabling read-only replicas, write requests are routed to the master node, while read requests are routed to all replica nodes through a load balancing algorithm, and the master node no longer processes read requests. The read/write separation feature is provided by the built-in Proxy component of TencentDB for Redis®.
- **Enabling/Disabling:** The read-only replica feature can be enabled or disabled on the instance creation page of the TencentDB for Redis® Console. It can also be managed via the TencentCloud API.

## Feature

- **Service reliability (1–5 replicas)**

With a dual-server master/slave architecture, the master and slave nodes reside on different physical machines with the master node providing external access. You can perform data CRUD using the Redis command line or client. In case that the master node fails, the proprietary HA system will automatically perform master/slave switchover to ensure smooth operation of the business.

- **Data reliability (1–5 replicas)**

The data persistence feature is enabled by default. Memory Edition (Standard Architecture) supports data backup. You can roll back or clone instances for backup files to effectively cope with data maloperations and other issues.

## Usage Limits

- Memory Edition (Standard Architecture) supports 0.25–64 GB of storage capacity. For higher specifications, use Cluster Edition that supports up to 8 TB of capacity.
- Memory Edition (Standard Architecture) supports up to 100,000 QPS (SET command concurrencies). If you need a higher QPS, you can choose multi-replica read/write separation or use Redis Cluster Edition that supports tens of millions of QPS.

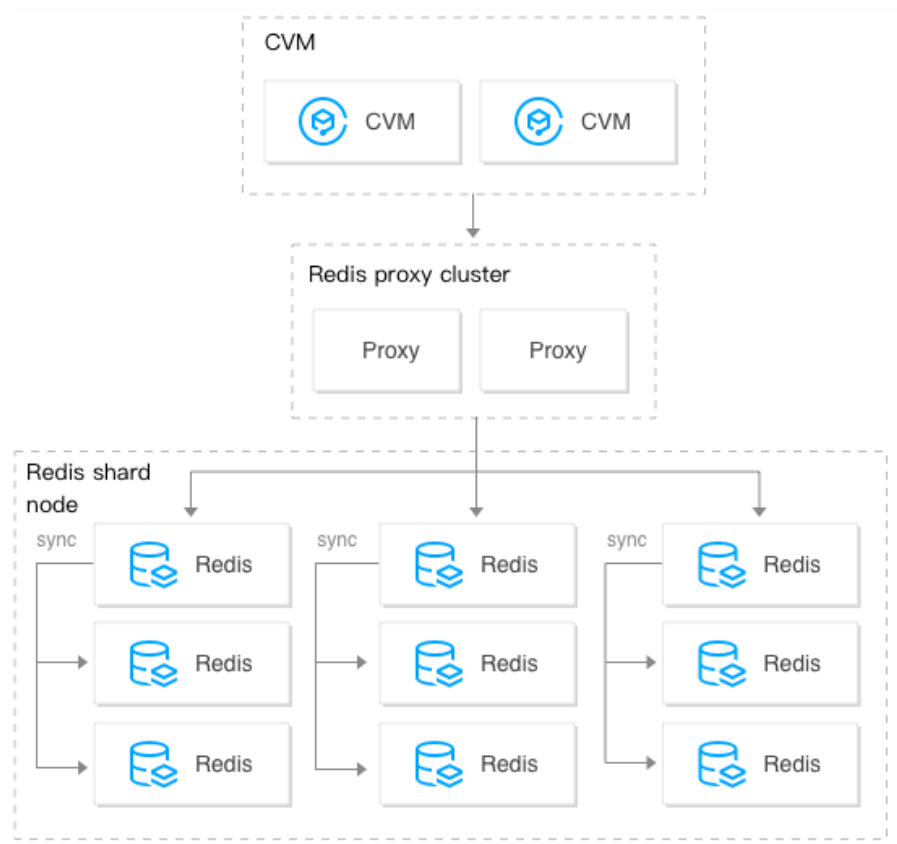
## Notes on command compatibility

For more information on command support, see [Command Compatibility Overview](#).

# Memory Edition (Cluster Architecture)

Last updated: 2024-11-01 17:02:17

TencentDB for Redis® Memory Edition (Cluster Architecture) is a brand-new version built by Tencent Cloud based on the Community Edition of Redis Cluster. It is compatible with Redis 4.0, 5.0, and 6.2 version commands, utilizing a distributed architecture to support elastic scaling of shards and replicas. This edition boasts high flexibility, availability, and performance with up to tens of millions of QPS. Redis Memory Edition (Cluster Architecture) supports horizontal scaling from 1 to 128 shards and replica expansion from 1 to 5 replica sets. The scaling, migration, and resizing processes are virtually imperceptible to the business, ensuring maximum service availability.



## Scenarios

### Master/replica high-availability (HA) scenarios

Memory Edition (Cluster Architecture) allows you to configure a replica set for a single node to achieve high master/replica availability. It features dual-server hot backup and automatic failover to ensure high reliability and availability of the Redis service.

### Read/write separation scenarios

When the number of replica nodes is one or more, automatic read/write separation can be enabled for the TencentDB for Redis® instance to extend the read performance of a single node. Up to five replica sets can be supported, and read access weights across the master and replica nodes can be configured.

### Multi-Shard High-Performance Scenarios

Memory Edition (Cluster Architecture) automatically enables automatic sharding and achieves horizontal scaling of system performance by assigning different keys to multiple nodes.

## Cluster Specifications

- Shard specification (GB): 1, 2, 4, 8, 12, 16, 20, 24, 32, 40, 48, 64.
- Shard quantity: 1, 3, 5, 8, 12, 16, 24, 32, 40, 48, 64, 80, 96, 128.
- Number of replicas: 1, 2, 3, 4, 5.

## Cluster Mode

- In cluster mode, data is automatically sharded. The system provides data load balancing and migration capabilities.
- The cluster mode is compatible with certain commands of the non-cluster mode, mainly reflected in cross-slot data access. For more information, see [Command compatibility](#).

## Replica Description

- When there is only one replica, Redis provides master/replica real-time hot backup for high data reliability and availability (server-level HA is supported in a single AZ). When the HA system detects a node failure, it requests for switching to a replica node and adds a new replica node to the system.
- When the number of replicas is greater than 1, Redis provides master/slave real-time hot backup with the slave nodes being read-only.

## Feature

### Flexibility

Memory Edition (Cluster Architecture) supports horizontal scaling and resizing from a minimum of 1 node to a maximum of 128 nodes, as well as replica expansion and contraction from 1 to 5 replica sets, making it suitable for various application scenarios.

### Availability

The expansion and contraction of shard and replica quantities in Memory Edition (Cluster Architecture) are completely imperceptible to the business, ensuring higher availability.

### Compatibility

Memory Edition (Cluster Architecture) supports native Cluster usage scenarios in the community edition, compatibility with intelligent clients such as Jedis, and compatibility with

Codis usage scenarios.

### **Maintainability**

Memory Edition (Cluster Architecture) maximizes system capabilities by providing shard-level monitoring and management, shard data migration and balancing, as well as advanced features such as large key monitoring and hot key monitoring. It offers fully intelligent monitoring and maintenance of cluster services.

## **Notes on command compatibility**

For more information on command support, see [Command Compatibility Overview](#).

# Performance

Last updated: 2024-02-27 11:55:38

## Specification Series

### Memory Edition

#### ⓘ Note:

- As a trial version, the 256 MB specification on v4.0 or v5.0 is only suitable for product verification in testing environments but not recommended for use in production environments. It is available only in: Guangzhou (Zones 6 and 7), Shanghai (Zones 2, 3, 4 and 5), Beijing (Zones 1, 2, 3, 4, 5, 6, and 7), and Shenzhen Finance (Zones 1, 2, and 3). Other 1 GB and above specifications can be smoothly downgraded to the 256 MB specification.
- v2.8 is unavailable for purchase currently, and v4.0 or later is recommended. To purchase v2.8, [submit a ticket](#).

Features	Standard Architecture			Cluster Architecture	
Redis- Compatible Versions	2.8	4.0、5.0	6.2	4.0、5.0	6.2
Memory Specifications	256MB – 64GB	256MB – 64GB	1GB – 64GB	2GB – 8TB	
Shard Quantity	Unavailable			1、3、5、8、12、16、24、32、40、48、64、80、96、128	
QPS	80,000–100,000	80,000–100,000	80,000–100,000	80,000–100,000 per shard	
Max connections	10,000 by default; up to 40,000	10,000 by default; up to 40,000		10,000/shard by default; up to 40,000/shard	

Traffic Limit	10MB/S – 64MB/S	528MB/s – 608MB/s	288MB/s – 72GB/s	
Multi-DB	This feature is supported.	This feature is supported.	This feature is supported.	This feature is supported.
Mget、Mset	This feature is supported.	This feature is supported.	This feature is supported.	This feature is supported.
lua	This feature is supported.	This feature is supported.	Yes (cross-slot access not supported)	Yes (cross-slot access not supported)
Scale out	Unavailable	Unavailable	This feature is supported.	This feature is supported.
Replica Expansion	Unavailable	This feature is supported.	This feature is supported.	This feature is supported.
Read/Write Separation	Unavailable	This feature is supported.	This feature is supported.	This feature is supported.
GEO	Unavailable	This feature is supported.	This feature is supported.	This feature is supported.
Number of Replicas	1	1 – 5		

## Description of CPU core number

To ensure the stable operations of the system, each node is assigned with 2-core CPU, one of which is used to process the backend tasks.

## CKV edition

SDK	Standard Architecture	Cluster Architecture
-----	-----------------------	----------------------

Redis- Compatibl e Versions	3.2	3.2
Memory Specificati ons	4GB – 384GB	12GB – 48TB
Shard Quantity	–	3 – 128
QPS	80,000–120,000	Tens of millions
Max connectio ns	12000 – 24000	12,000–24,000 per shard
Traffic Limit	16MB/S – 256MB/S	72MB/S – 32GB/S
Multi-DB	This feature is supported.	This feature is supported.
Mget、 Mset	This feature is supported.	This feature is supported.
lua	This feature is supported.	Limited support (to use Lua in the cluster edition, you need to make sure that the keys accessed in the Lua script are in the same slot, and the <code>key</code> field must be included in the command parameters)
Scale out	Unavailable	This feature is supported.
Replica Expansion	Unavailable	Unavailable
Read/Writ e Separatio n	Unavailable	Unavailable
GEO	This feature is supported.	This feature is supported.

## Traffic and Connections

## Memory Edition

Specification (GB)	Max connections	Maximum throughput (MB/s)
0.25	3000	10
1	40000	16
2	40000	24
4	40000	24
8	40000	24
12	40000	32
16	40000	32
20	40000	48
24	40000	48
32	40000	48
40	40000	64
48	40000	64
60	40000	64

## CKV edition

Specification (GB)	Max connections	Maximum throughput (MB/s)
4	10000	24
8	10000	24
16	10000	32
24	10000	32
32	10000	32
48	18000	64
64	18000	64
80	18000	64

96	18000	64
128	24000	128
160	24000	128
192	24000	128
256	24000	256
320	24000	256
384	24000	256

Cluster edition connections = number of connections per shard \* number of shards;

Cluster edition throughput = shard throughput \* number of shards

#### Note:

After scaling, legacy instances capable of up to 9,000 connections will be capable of 10,000 ones.

## Performance Statistics

### Performance references

The time needed to execute Redis commands varies. Businesses use different database commands in their production environments; therefore, the corresponding performance values will also vary. The test results listed here are obtained with specified parameters and are for your reference only. Conduct tests in your actual business environment for more accurate results.

### Single-node Test Performance

Redis Instance Specification	Number of connections	QPS
Memory Edition (Standard Architecture) 8 GB	10000	80,000–100,000
Memory Edition (Cluster Architecture) 8 GB (single-shard)	10000	80,000–100,000
CKV Edition (Standard Architecture) 8 GB	12000	80,000–120,000

## Cluster architecture test performance

Memory Edition (Cluster Architecture) performance = Memory Edition (Standard Architecture) performance \* number of shards

CKV Edition (Cluster Architecture) performance = CKV Edition (Standard Architecture) performance \* number of shards

## Testing Method

### Test environment

Test Client CVM Instances	CVM Cores	CVM Memory	Region	Redis Instance Specification
3	2 cores	8GB	Guangz hou Zone 3	Memory Edition (Standard Architecture) 8 GB
3	2 cores	8GB	Guangz hou Zone 3	CKV Edition (Standard Architecture) 8 GB

### Test Parameters

```
redis-benchmark -h 10.66.187.x -p 6379 -a crs-1znib6aw:chen2016 -t set -c 3500 -d 128 -n 25000000 -r 5000000
redis-benchmark -h 10.66.187.x -p 6379 -a crs-1z5536aw:chen2016 -t set -c 3500 -d 128 -n 25000000 -r 5000000
redis-benchmark -h 10.66.187.x -p 6379 -a crs-090rjlih:1234567 -t set -c 3500 -d 128 -n 25000000 -r 5000000
```

### QPS Calculation

Sum of the QPS values of 3 pressure test clients (tested by redis-benchmark).

# Read/Write Separation

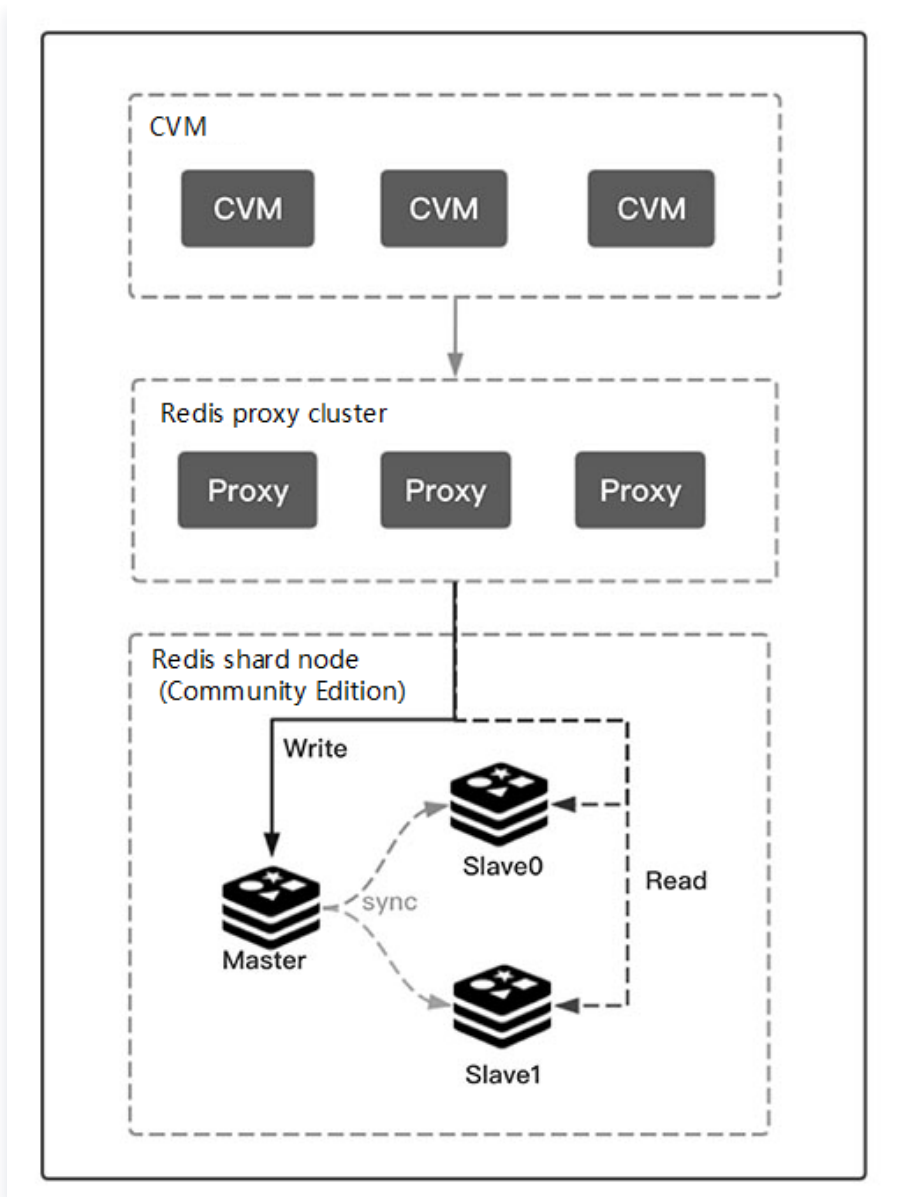
Last updated: 2024-11-01 15:57:06

TencentDB for Redis® supports read-write separation for business scenarios with more reads but less writes, which can well cope with read requests concentrating on frequently read data. It supports up to 1-master 5-replica mode to offer 5x read performance.

## How It Works

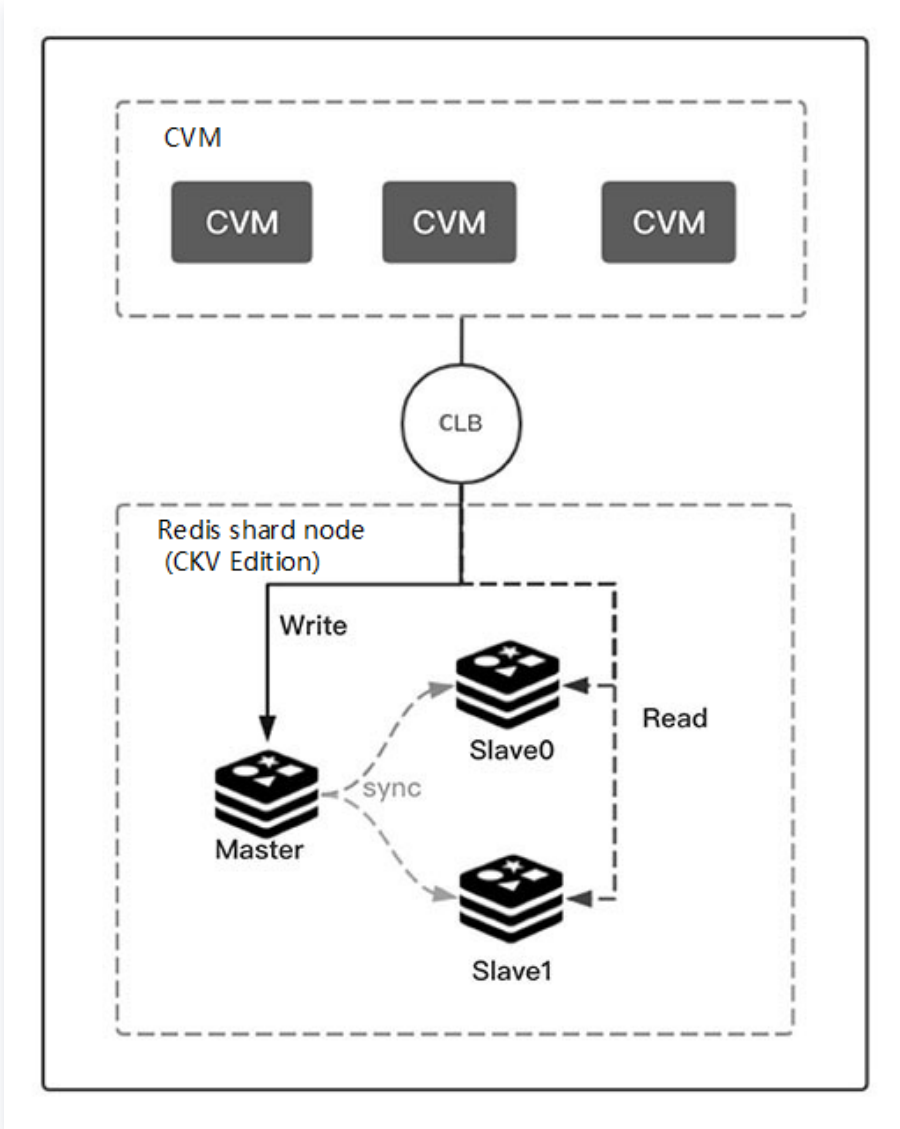
### Memory Edition

- **Read/Write Separation Principle:** Fully compatible with standard and cluster architectures of Redis 4.0 and later, this system implements automatic read/write separation via the proxy layer.
- **Read/Write Separation Weight:** When read/write separation is enabled, the proxy writes data only to the master node and evenly distributes read requests among the replica nodes.



## CKV edition

- **Read/Write Separation Principle:** The CKV version system inherently supports read/write separation architecture. All requests are distributed to nodes in each cluster through the load balancing (CLB) gateway with each node possessing global slot routing information. When read/write separation is enabled, if the key being read falls on the current node, the data is read directly and returned; otherwise, the request is forwarded to the corresponding node based on the routing information. The data is then read and returned to the original node, which in turn sends it back to the client.
- **Read/Write Separation Weight:** Within the CKV version, the CLB distributes requests, leading to an even allocation of read/write weights based on the TCP connection's quadruple elements: source IP, source port, destination IP, and destination port.



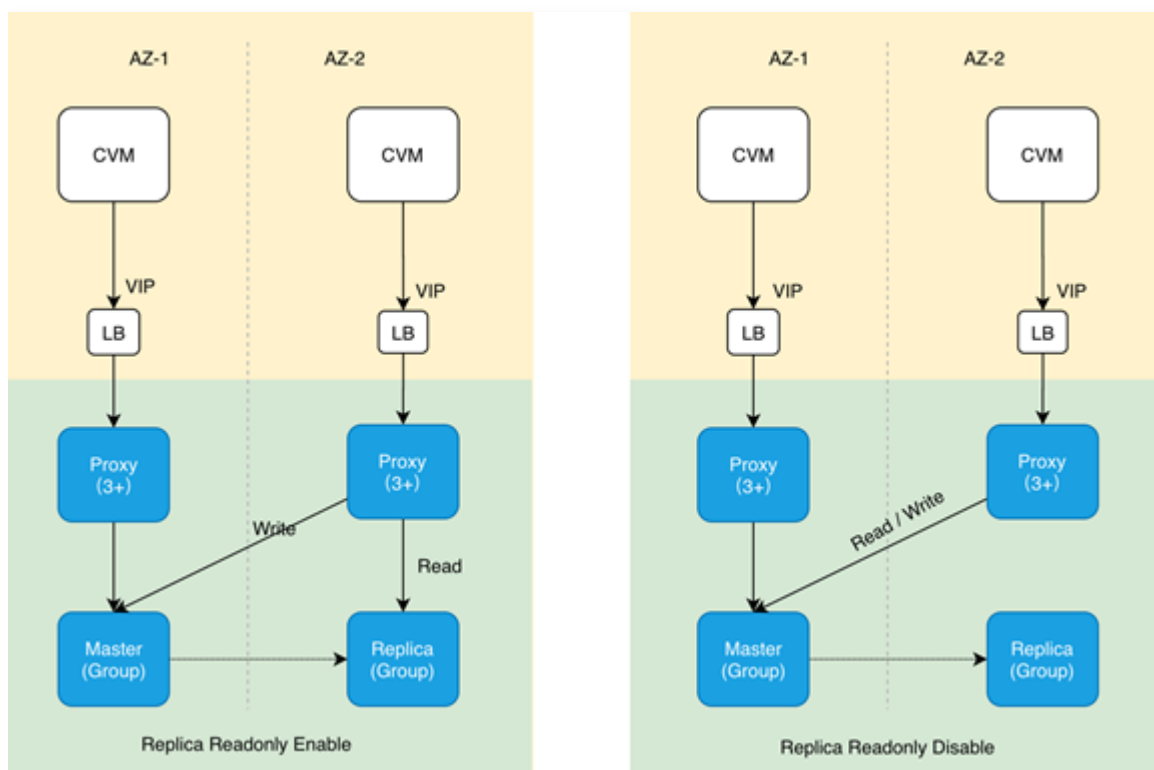
# Multi-AZ Deployment

Last updated: 2024-11-01 15:58:16

TencentDB for Redis® supports deploying replicas across multiple availability zones within the same region, providing higher availability and disaster recovery capabilities compared to single availability zone instances (where the master and replica nodes are in the same availability zone).

- Single-AZ deployed instance: Host- and rack-level disaster recovery.
- Multi-AZ deployed instance: Host-, rack-, and AZ-level disaster recovery.

## Deployment architecture



### Description:

- **LB (Load Balancer):** A TencentDB for Redis® instance in standard architecture or cluster architecture has at least three proxies which need to be accessed through LB.
- **VIP:** In Multi-AZ (Availability Zone) instances, a single VIP is used, accessible across the entire region. Redis High Availability (HA) does not cause any changes to the VIP.
- **Proxy:** Access service that implements request routing and is responsible for distributing user data requests.
  - If read-only replica is enabled, each instance has at least three proxies. For a standard architecture instance, the number of proxies = 3 + (the number of replicas - 1); for a

cluster architecture instance, the number of proxies = Max [the number of shards *the number of replicas*; *shard quantity* \* 1.5 (rounded up)]

- If read-only replica is not enabled, for a standard architecture instance, the number of proxies is 3; for a cluster architecture instance, the number of proxies is shard quantity \* 1.5 (rounded up).
- Master (Group): "Master" refers to the master node of a TencentDB for Redis® instance in standard architecture; "Master Group" refers to the master nodes of all shards of a TencentDB for Redis® instance in cluster architecture.
- Replica (Group): "Replica" refers to the replica nodes of a TencentDB for Redis® instance in standard architecture; "Replica Group" refers to the collection of replica nodes, each of which comes from a different shard of a TencentDB for Redis® instance in the cluster architecture. For a cluster architecture instance, the replicas of a shard are divided into multiple Replica Groups so that these Replica Groups can be deployed in different AZs.
- Master AZ: The master AZ refers to the AZ where the master node resides. Unless manually changed in the console, the master AZ will remain the same. If the master node fails, it may be temporarily switched to a replica AZ, and will be automatically switched back to the master AZ in a few minutes once certain conditions are met. This switching back process won't affect your business unless your business uses block commands, such as `blpop` or `blpush`.

## Failover (HA)

- Fault determination: Both Redis standard and cluster architectures use the native Redis Cluster management mechanism, relying on the Gossip protocol for node status determination among cluster nodes. The timeliness of node fault determination depends on the cluster-node-timeout, which has a default value of 15 s. It is recommended not to change this parameter. For node fault determination, please see [Native Redis Cluster Design](#).
- Promote a replica to master: TencentDB for Redis® adopts a failover mechanism different from that of Redis Cluster, which gives priority to promoting replicas in the master AZ to reduce access delay of the master AZ. The details are as follows:
  - Promote the replica if it has the latest data
  - Promote the replica in the master AZ if all replicas have the same data

## Cross-AZ Access

### Instance with read-only replica disabled

Read/Write separation disabled (that is, replicas can be written to and read from): Write/read requests in a replica AZ are routed by proxy to the master node, and the master node synchronizes with replica nodes to ensure consistent data across all nodes. In this process, only one cross-AZ access happens.

## Instance with read-only replica enabled

Read/Write separation enabled (that is, replicas can only be read from): Write requests are routed by proxy to the master node, but read requests are routed to the replica node in the same AZ as the proxy, so that read requests can get responded by the nearest node.

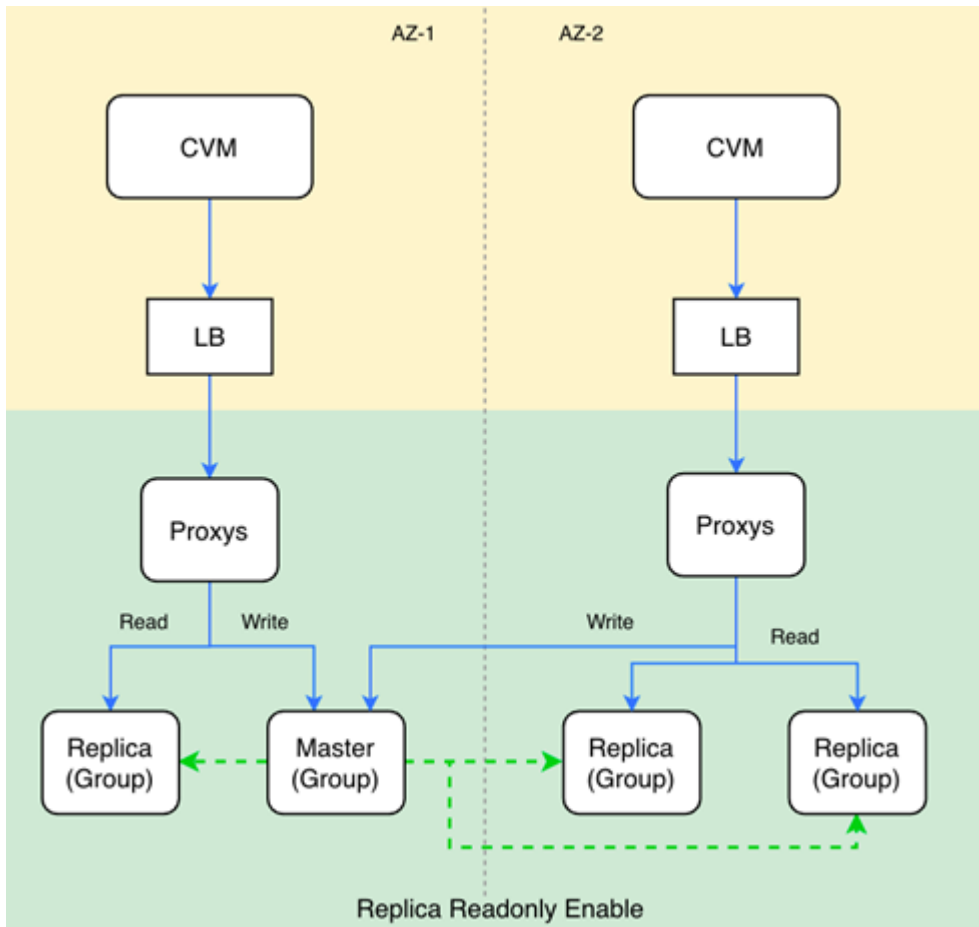
### Note:

There is a 2-5 ms delay during cross-AZ access.

## Deployment recommendations

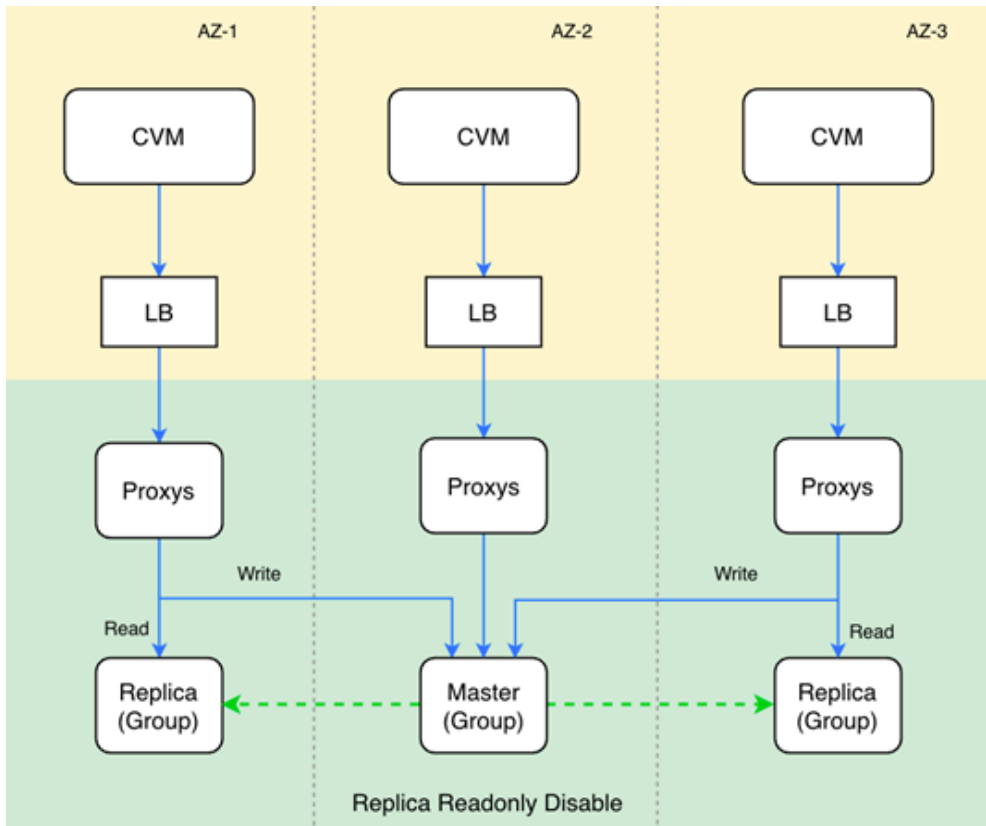
### Two-AZ deployment

Two AZs with 1 master and 3 replicas: one master and one replica in the primary AZ, and two replicas in the secondary AZ. Load balancing is implemented across both AZs, with two nodes in each AZ. This ensures that read requests do not cross AZs in case of a single-node failure, and high availability is maintained even if one AZ fails. This solution is suitable for scenarios with high requirements for availability and access latency.



## Three-AZ deployment

A deployment with one master node and two replica nodes across three availability zones, with one node in each AZ. This architecture ensures high availability across AZs in case of single host or single AZ failure. It is suitable for scenarios with high availability requirements but less sensitivity to latency.



## Related Actions

- TencentDB for Redis® supports configuring and viewing multi-AZ information through the console. For more information, see [Configuring Multi-AZ](#).
- TencentDB for Redis® supports upgrading from single-AZ to multi-AZ deployment through the console. Please see [Upgrading to Multi-AZ Deployment](#).
- TencentDB for Redis® supports enabling and disabling read/write separation through the console. Please see [Enable/Disable Read/Write Separation](#).
- Both single-AZ and multi-AZ deployed TencentDB for Redis® instances can be easily accessed at a VIP address. For more information, please see [Accessing Multi-AZ Instances](#).
- TencentDB for Redis® supports the automatic failover feature to ensure service availability. Please see [Failover](#) for more information.
- TencentDB for Redis® provides the automatic failback feature for instances deployed across AZs. Please see [Auto Failback](#) for more information.
- TencentDB for Redis® provides a manual promotion feature for instances deployed across multiple availability zones. You can deploy the master node to a specified availability zone or replica (group) based on your business requirements. Please see [Manually Promoting the Master Node](#).

- To reduce the access latency of a multi-AZ deployed instance, TencentDB for Redis® allows you to read local nodes only. For more information, please see [Proximity Access](#).

# Regions and AZs

Last updated: 2024-11-20 17:13:01

## Regions

### Feature Overview

A region is the physical location of an IDC. Data centers in different Tencent Cloud regions are isolated to ensure maximal stability and fault tolerance. We recommend that you choose the region closest to your end users to minimize access latency and improve access speed.

### Related features

- The networks of different regions are fully isolated from each other, and Tencent Cloud services in different regions **cannot communicate with each other over the private network by default**.
- Tencent Cloud services in different VPCs can communicate with each other via [Cloud Connect Network](#) which offers faster and more stable communication.

#### ⓘ Note:

**Shenzhen/Shanghai Finance Zones** are custom AZs designed to comply with regulatory requirements in the finance industry, featuring high security and isolation. They currently offer services such as cloud servers, financial databases, Redis storage, and face recognition. Verified financial customers can apply for access to these zones through [online consultation](#). For more information, see [Introduction to Finance Zones](#).

## Availability Zones

### Feature Overview

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.

### Related features

Tencent Cloud services in the same VPC are interconnected via the private network, which means they can communicate using [private IPs](#), even if they are in different AZs of the same

region.

**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.

## List of Regions and AZs

**Note:**

The availability of resources in various regions and AZs is subject to change; they may sell out and become temporarily unavailable, with previously depleted resources potentially being restocked. The resource availability will be assessed and adjusted based on the actual business usage as displayed on the purchase page in the console.

## China

Regions	Availability Zones	Zone ID
South China (Guangzhou) ap-guangzhou	Guangzhou Zone 1 ap-guangzhou-1	100001
	Guangzhou Zone 3 ap-guangzhou-2	100002
	Guangzhou Zone 3 ap-guangzhou-3	100003
	Guangzhou Zone 4 ap-guangzhou-4	100004
	Guangzhou Zone 6 ap-guangzhou-6	100006
	Guangzhou Zone 7 ap-guangzhou-7	100007
South China (Shenzhen Finance) ap-shenzhen-fsi	Shenzhen Finance Zone 1 (available exclusively for financial institutions and enterprises through <a href="#">online consultation</a> application) ap-shenzhen-fsi-1	110001

	Shenzhen Finance Zone 2 (only financial institutions and enterprises can contact us to apply for activation) ap-shenzhen-fsi-2	11000 2
	Shenzhen Finance Zone 3 (only financial institutions and enterprises can contact us to apply for activation) ap-shenzhen-fsi-3	11000 3
East China (Shanghai) ap-shanghai	Shanghai Zone 1 ap-shanghai-1	2000 01
	Shanghai Zone 2 ap-shanghai-2	2000 02
	Shanghai Zone 3 ap-shanghai-3	2000 03
	Shanghai Zone 4 ap-shanghai-4	2000 04
	Shanghai Zone 5 ap-shanghai-5	2000 05
	ap-shanghai-7 ap-shanghai-7	2000 07
	Shanghai Zone 8 ap-shanghai-8	2000 08
East China (Shanghai Finance) ap-shanghai-fsi	Shanghai Finance Zone 1 (only financial institutions and enterprises can contact us to apply for activation) ap-shanghai-fsi-1	7000 01
	Shanghai Finance Zone 2 (only financial institutions and enterprises can contact us to apply for activation) ap-shanghai-fsi-2	7000 02
	Shanghai Finance Zone 3 (only financial institutions and enterprises can contact us to apply for activation) ap-shanghai-fsi-3	7000 03

East China (Nanjing) ap-nanjing	Nanjing Zone 1 ap-nanjing-1	3300 01
	Nanjing Zone 2 ap-nanjing-2	3300 02
	Nanjing Zone 3 ap-nanjing-3	3300 03
ap-beijing-1 ap-beijing	ap-beijing-1 ap-beijing-1	8000 01
	Beijing Zone 2 ap-beijing-2	8000 02
	Beijing Zone 3 ap-beijing-3	8000 03
	Beijing Zone 4 ap-beijing-4	8000 04
	Beijing Zone 5 ap-beijing-5	8000 05
	Beijing Zone 6 ap-beijing-6	8000 06
	Beijing Zone 7 ap-beijing-7	8000 07
North China (Beijing Finance) ap-beijing-fsi	Beijing Finance Zone 1 (only financial institutions and enterprises can contact us to apply for activation) ap-beijing-fsi-1	4600 01
Southwest China (Chengdu) ap-chengdu	Chengdu Zone 1 ap-chengdu-1	1600 01
	Chengdu Zone 2 ap-chengdu-2	1600 02
Southwest China (Chongqing) ap-chongqing	Chongqing Zone 1 ap-chongqing-1	1900 01
Hong Kong, Macao and Taiwan, China (Hong Kong)	Hong Kong Zone 1 (Hong Kong nodes cover services in the China regions of Hong Kong, Macao, and Taiwan)	3000 01

ap-hongkong	ap-hongkong-1	
	Hong Kong Zone 2 (Hong Kong nodes cover services in the China regions of Hong Kong, Macao, and Taiwan) ap-hongkong-2	3000 02
	Hong Kong Zone 3 (Hong Kong nodes cover services in the China regions of Hong Kong, Macao, and Taiwan) ap-hongkong-3	3000 03

## Other Countries and Regions

Regions	Availability Zones	Zonel d
Southeast Asia (Singapore) ap-singapore	Singapore Zone 1 (Nodes in Singapore can cover Southeast Asia) ap-singapore-1	9000 01
	Singapore Zone 2 (Nodes in Singapore can cover Southeast Asia) ap-singapore-2	9000 02
	Singapore Zone 3 (Nodes in Singapore can cover Southeast Asia) ap-singapore-3	9000 03
	Singapore Zone 4 (Nodes in Singapore can cover Southeast Asia) ap-singapore-4	9000 04
Southeast Asia (Jakarta) ap-jakarta	Jakarta Zone 1 (Nodes in Jakarta can cover Southeast Asia) ap-jakarta-1	7200 01
	Jakarta Zone 2 (Jakarta nodes cover services in Southeast Asia) ap-jakarta-2	7200 02
Southeast Asia (Bangkok) ap-bangkok	Bangkok Zone 1 (Bangkok nodes cover services in Southeast Asia) ap-bangkok-1	2300 01
	Bangkok Zone 2 (Bangkok nodes cover services in Southeast Asia)	2300 02

	ap-bangkok-2	
South Asia (Mumbai) ap-mumbai	Mumbai Zone 1 (Mumbai nodes cover services in South Asia) ap-mumbai-1	21000 1
	Mumbai Zone 2 (Mumbai nodes cover services in South Asia) ap-mumbai-2	21000 2
Northeast Asia (Seoul) ap-seoul	Seoul Zone 1 (Seoul nodes cover services in Northeast Asia) ap-seoul-1	1800 01
	Seoul Zone 2 (Seoul nodes cover services in Northeast Asia) ap-seoul-2	1800 02
Northeast Asia (Tokyo) ap-tokyo	Tokyo Zone 1 (Tokyo nodes cover services in Northeast Asia) ap-tokyo-1	2500 01
	Tokyo Zone 2 (Tokyo nodes cover services in Northeast Asia) ap-tokyo-2	2500 02
US West (Silicon Valley) na-siliconvalley	Silicon Valley Zone 1 (Silicon Valley nodes cover services in West US) na-siliconvalley-1	1500 01
	Silicon Valley Zone 2 (Silicon Valley nodes cover services in West US) na-siliconvalley-2	1500 02
US East (Virginia) na-ashburn	Virginia Zone 1 (Virginia nodes cover services in East US) na-ashburn-1	2200 01
	Virginia Zone 2 (Virginia nodes cover services in East US) na-ashburn-2	2200 02
North America (Toronto) na-toronto	Toronto Zone 1 (Toronto nodes cover services in North America) na-toronto-1	4000 01

Europe (Frankfurt) eu-frankfurt	Frankfurt Zone 1 (Frankfurt nodes cover services in Europe) eu-frankfurt-1	17000 1
	Frankfurt Zone 2 (Frankfurt nodes cover services in Europe) eu-frankfurt-2	17000 2
South America (Sao Paulo) sa-saopaulo	São Paulo Zone 1 (São Paulo nodes cover services in South America) sa-saopaulo-1	7400 01

## Selection of Regions and AZs

When you purchase Tencent Cloud services, we recommend that you select the region closest to your end users to minimize access latency and improve download speed.