# GeminiDB DynamoDB-Compatible

# User Guide

| | |
|---|---|
| **Issue** | 01 |
| **Date** | 2025-09-04 |

**HUAWEI CLOUD COMPUTING TECHNOLOGIES CO., LTD.**

## Trademarks and Permissions


and other Huawei trademarks are the property of Huawei Technologies Co., Ltd.
All other trademarks and trade names mentioned in this document are the property of their respective holders.

## Notice

The purchased products, services and features are stipulated by the contract made between Huawei Cloud and the customer. All or part of the products, services and features described in this document may not be within the purchase scope or the usage scope. Unless otherwise specified in the contract, all statements, information, and recommendations in this document are provided "AS IS" without warranties, guarantees or representations of any kind, either express or implied.

The information in this document is subject to change without notice. Every effort has been made in the preparation of this document to ensure accuracy of the contents, but all statements, information, and recommendations in this document do not constitute a warranty of any kind, express or implied.

# Contents



# 1 Service Overview

## 1.1 What Is GeminiDB DynamoDB-Compatible API?

GeminiDB is completely compatible with Amazon DynamoDB, so you can access GeminiDB DynamoDB-Compatible API through SDKs or CLIs and smoothly migrate data from DynamoDB to GeminiDB without refactoring. GeminiDB DynamoDB-Compatible API is secure and reliable. It removes constraints of closed-source software.

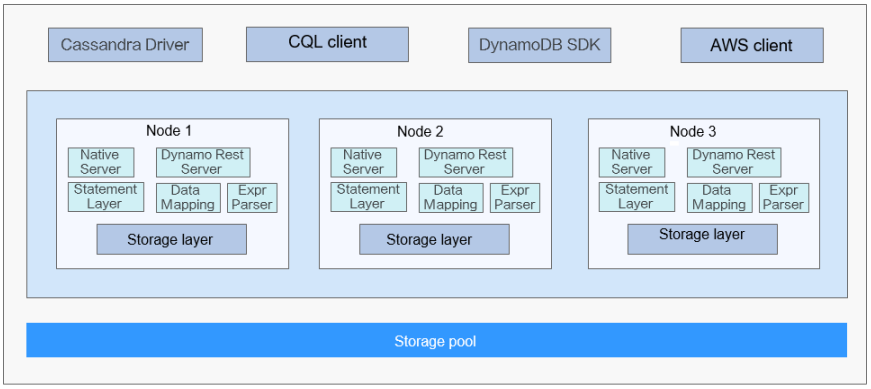
- Secure and reliable
  - VPCs, subnets, security groups, SSL, and fine-grained permissions control ensure database security and user privacy.
  - Cross-region active-active DR is supported. You can deploy an instance across three AZs and quickly back up or restore data.
  - The distributed architecture ensures fault tolerance for *N-1* nodes.
- No pain points of closed-source software
  - Storage can be scaled in seconds without affecting services.
  - Compute nodes can be added in minutes. A jitter may last only a few seconds.

### Architecture

GeminiDB DynamoDB-Compatible API is embedded in kernel services of GeminiDB Cassandra API and communicates with DynamoDB over HTTPS. The following figure shows the overall architecture.

**Figure 1-1** GeminiDB DynamoDB-Compatible API architecture




GeminiDB strictly follows the DynamoDB syntax and data models, so DynamoDB applications can be easily migrated to GeminiDB DynamoDB-Compatible instances. In addition, GeminiDB DynamoDB-Compatible API provides automated management and O&M functions, such as automated backup, fault detection, and fault tolerance.

## Highlights

- Low costs

  Companies generate increasing amounts of data as their businesses grow. Handling frequent read and write requests on DynamoDB becomes costly. GeminiDB provides higher throughput at half the cost of DynamoDB, especially when there is a large amount of data.

- Flexible deployment

  DynamoDB is a fully managed service that needs to follow certain rigid deployment requirements. In contrast, GeminiDB offers flexible deployment options. It can run on both private and public clouds, avoiding platform lock-in. GeminiDB also delivers stable and efficient performance in all deployments.

- Smooth migration

  GeminiDB is fully compatible with DynamoDB syntax and data models, so you can migrate data from DynamoDB to GeminiDB without changing applications. This makes the migration simpler, faster, and less disruptive to your operations.

# 1.2 API Compatibility

This section describes APIs supported by GeminiDB DynamoDB-Compatible API.

## Usage Notes

- You can add, delete, and modify tables of GeminiDB DynamoDB-Compatible instances only through DynamoDB APIs. Tables can only be read on the console.
- Tables created through non-DynamoDB APIs cannot be used on GeminiDB DynamoDB-Compatible instances.

- The current version does not support traffic statistics, traffic control, or transaction functions (TransactWriteItems and TransactGetItems).
- The current version does not support the following on-demand backup APIs: CreateBackup, DescribeBackup, DeleteBackup, ListBackups, and RestoreTableFromBackup.
- Hash or partition keys determine where items are stored in an internal storage device of a GeminiDB DynamoDB-Compatible instance. Amazon DynamoDB uses a dedicated hash function, while GeminiDB DynamoDB-Compatible API uses the Murmur3 algorithm. As a result, the sequence of data returned by a scan operation is different.
- The table or index name cannot contain periods (.).
- Due to different encoding methods, the size of each row of data in a table on a GeminiDB DynamoDB-Compatible instance is not strictly limited to 400 KB.
- All DDL APIs are synchronous APIs.

## Supported Expressions

ConditionExpression of GeminiDB DynamoDB-Compatible API supports the following function syntax:

```
function ::=
    attribute_exists (path)
    | attribute_not_exists (path)
    | attribute_type (path, type)
    | begins_with (path, substr)
    | contains (path, operand)
```

**operand** in **contains(path,operand)** can only be an attribute value (that is, the value placeholder).

UpdateExpression of GeminiDB DynamoDB-Compatible API supports the following syntax:

```
update-expression ::=
     [ SET action [, action] ... ]
     [ REMOVE action [, action] ...]
     [ ADD action [, action] ... ]
     [ DELETE action [, action] ...]

 set-action ::=
     path =value

 value ::=
     operand
     | operand '+' operand
     | operand '-' operand

 operand ::=
     path | function

 function ::=
     if_not_exists (path, value)
     | list_append (list1, list2)

 remove-action ::=
     path

 add-action ::=
     path value

 delete-action ::=
     path value
```

In SET path = operand, **operand** cannot be set to **path**.

In SET path = operand1 '+'|'-' operand2, **operand1** must be set to **path**, and **operand2** must be an attribute

value.
In SET path = if_not_exists (path, value), values of the two **path** parameters must be the same and must be expression attribute values.
All values can only be placeholders, for example, **':placeholder'**.

## Compatible APIs

**Table 1-1** API list

| API |
| --- |
| **CreateTable** |
| **UpdateTable** |
| **DescribeTable** |
| **ListTables** |
| **DeleteTable** |
| **PutItem** |
| **UpdateItem** |
| **GetItem** |
| **DeleteItem** |
| **BatchWriteItem** |
| **BatchGetItem** |
| **Query** |
| **Scan** |
| **UpdateTimeToLive** |
| **DescribeTimeToLive** |

**Table 1-2** CreateTable

| API | Parameter | | Mandatory | Supported | Remarks |
| --- | --- | --- | --- | --- | --- |
| CreateTable | Request parameter | AttributeDefinitions | Yes | Yes | - |
| | | KeySchema | Yes | Yes | - |
| | | TableName | Yes | Yes | Can contain 3 to 48 characters. The regular expression is ([\w-]+). |

| API | Parameter | | Mandatory | Supported | Remarks |
|---|---|---|---|---|---|
| | | BillingMode | No | No | Billing parameter, which is not supported |
| | | GlobalSecondaryIndexes | No | Yes | - |
| | | LocalSecondaryIndexes | No | Yes | - |
| | | ProvisionedThroughput | Yes | Yes | - |
| | | SSESpecification | No | No | Not supported in the current version (planning). |
| | | StreamSpecification | No | Yes | - |
| | | Tags | No | No | Not supported in the current version (planning). |
| | Response parameter | TableDescription | - | Yes | - |

**Table 1-3** UpdateTable

| API | Parameter | | Mandatory | Supported | Remarks |
|---|---|---|---|---|---|
| UpdateTable | Request parameter | AttributeDefinitions | No | Yes | - |
| | | BillingMode | No | No | Billing parameter, which is not supported |
| | | GlobalSecondaryIndexesUpdates | Yes | Yes | CREATE and DELETE are supported, but UPDATE is not. |
| | | ProvisionedThroughput | No | No | Billing parameter, which is not supported |

| API | Parameter | | Mandatory | Supported | Remarks |
| --- | --- | --- | --- | --- | --- |
| | | ReplicaUpdates | No | No | Default strong consistency |
| | | SSESpecification | No | No | Not supported in the current version (planning). |
| | | StreamSpecification | No | Yes | - |
| | | TableName | Yes | Yes | - |
| | Response parameter | TableDescription | - | Yes | - |

**Table 1-4** DescribeTable

| API | Parameter | | Mandatory | Supported | Remarks |
| --- | --- | --- | --- | --- | --- |
| DescribeTable | Request parameter | TableName | Yes | Yes | - |
| | Response parameter | Table | - | Yes | **itemCount** is not supported. |

**Table 1-5** ListTables

| API | Parameter | | Mandatory | Supported | Remarks |
| --- | --- | --- | --- | --- | --- |
| ListTables | Request parameter | ExclusiveStartTableName | No | Yes | - |
| | | Limit | No | Yes | - |
| | Response parameter | LastEvaluatedTableName | - | Yes | - |
| | | TableNames | - | Yes | - |

**Table 1-6** DeleteTable

| API | Parameter | | Mandatory | Supported | Remarks |
|---|---|---|---|---|---|
| DeleteTable | Request parameter | TableName | Yes | Yes | - |
| | Response parameter | TableDescription | - | Yes | - |

**Table 1-7** PutItem

| API | Parameter | | Mandatory | Supported | Remarks |
|---|---|---|---|---|---|
| PutItem | Request parameter | Item | Yes | Yes | - |
| | | TableName | Yes | Yes | - |
| | | ConditionalOperator | No | No | Deprecated parameter. For details, see **ConditionExpression**. |
| | | ConditionExpression | No | Yes | - |
| | | Expected | No | No | Deprecated parameter. For details, see **ConditionExpression**. |
| | | ExpressionAttribute-Names | No | Yes | - |
| | | ExpressionAttribute-Values | No | Yes | - |
| | | ReturnConsumedCapacity | No | No | Billing parameter, which is not supported |
| | | ReturnItemCollec-tionMetrics | No | No | Statistics parameter, which is not supported |
| | | ReturnValues | No | Yes | - |

| API | Parameter | | Mandatory | Supported | Remarks |
| --- | --- | --- | --- | --- | --- |
| | Response parameter | Attributes | - | Yes | - |
| | | ConsumedCapacity | - | No | Billing parameter, which is not supported |
| | | ItemCollectionMetrics | - | No | Statistics parameter, which is not supported |

**Table 1-8** UpdateItem

| API | Parameter | | Mandatory | Supported | Remarks |
| --- | --- | --- | --- | --- | --- |
| UpdateItem | Request parameter | Key | Yes | Yes | - |
| | | TableName | Yes | Yes | - |
| | | AttributeUpdates | No | No | Deprecated parameter. For details, see **UpdateExpression**. |
| | | ConditionalOperator | No | No | Deprecated parameter. For details, see **UpdateExpression**. |
| | | ConditionExpression | No | Yes | - |
| | | Expected | No | No | Deprecated parameter. For details, see **UpdateExpression**. |
| | | ExpressionAttributeNames | No | Yes | - |
| | | ExpressionAttributeValues | No | Yes | - |
| | | ReturnConsumedCapacity | No | No | Billing parameter, which is not supported |

| API | Parameter | | Mandatory | Supported | Remarks |
|---|---|---|---|---|---|
| | | ReturnItemCollectionMetrics | No | No | Statistics parameter, which is not supported |
| | | ReturnValues | No | Yes | - |
| | | UpdateExpression | No | Yes | - |
| | Response parameter | Attributes | - | Yes | - |
| | | ConsumedCapacity | - | No | Billing parameter, which is not supported |
| | | ItemCollectionMetrics | - | No | Statistics parameter, which is not supported |

**Table 1-9** GetItem

| API | Parameter | | Mandatory | Supported | Remarks |
|---|---|---|---|---|---|
| GetItem | Request parameter | Key | Yes | Yes | - |
| | | TableName | Yes | Yes | - |
| | | AttributesToGet | No | No | Deprecated parameter. For details, see **ProjectionExpression**. |
| | | ConsistentRead | No | No | Default strong consistency |
| | | ExpressionAttributeNames | No | Yes | - |
| | | ProjectionExpression | No | Yes | - |
| | | ReturnConsumedCapacity | No | No | Billing parameter, which is not supported |

| API | Parameter | | Mandatory | Supported | Remarks |
| --- | --- | --- | --- | --- | --- |
| | Response parameter | ConsumedCapacity | - | No | Billing parameter, which is not supported |
| | | Item | - | Yes | - |

**Table 1-10** DeleteItem

| API | Parameter | | Mandatory | Supported | Remarks |
| --- | --- | --- | --- | --- | --- |
| DeleteItem | Request parameter | Key | Yes | Yes | - |
| | | TableName | Yes | Yes | - |
| | | ConditionalOperator | No | No | Deprecated parameter. For details, see **ConditionExpression**. |
| | | ConditionExpression | No | Yes | - |
| | | Expected | No | No | Deprecated parameter. For details, see **ConditionExpression**. |
| | | ExpressionAttributeNames | No | Yes | - |
| | | ExpressionAttributeValues | No | Yes | - |
| | | ReturnConsumedCapacity | No | No | Billing parameter, which is not supported |
| | | ReturnItemCollectionMetrics | No | No | Statistics parameter, which is not supported |
| | | ReturnValues | No | Yes | - |
| | Response parameter | Attributes | - | Yes | - |
| | | ConsumedCapacity | - | No | Billing parameter, which is not supported |

| API | Parameter | | Mandatory | Supported | Remarks |
|---|---|---|---|---|---|
| | | ItemCollectionMetrics | - | No | Statistics parameter, which is not supported |

**Table 1-11** BatchWriteItem

| API | Parameter | | Mandatory | Supported | Remarks |
|---|---|---|---|---|---|
| BatchWriteItem | Request parameter | RequestItems | Yes | Yes | - |
| | | ReturnConsumedCapacity | No | No | Billing parameter, which is not supported |
| | | ReturnItemCollectionMetrics | No | No | Statistics parameter, which is not supported |
| | Response parameter | ConsumedCapacity | - | No | Billing parameter, which is not supported |
| | | ItemCollectionMetrics | - | No | Statistics parameter, which is not supported |
| | | UnprocessedItems | - | Yes | - |

**Table 1-12** BatchGetItem

| API | Parameter | | Mandatory | Supported | Remarks |
|---|---|---|---|---|---|
| BatchGetItem | Request parameter | RequestItems | Yes | Yes | None |
| | | ReturnConsumedCapacity | No | No | Billing parameter, which is not supported |
| | Response parameter | ConsumedCapacity | - | No | Statistics parameter, which is not supported |
| | | Responses | - | Yes | - |

| API | Parameter | | Mandatory | Supported | Remarks |
|---|---|---|---|---|---|
| | | UnprocessedKeys | - | Yes | - |

**Table 1-13** Query

| API | Parameter | | Mandatory | Supported | Remarks |
|---|---|---|---|---|---|
| Query | Request parameter | TableName | Yes | Yes | - |
| | | AttributesToGet | No | No | Deprecated parameter. For details, see **ProjectionExpression**. |
| | | ConditionalOperator | No | No | Deprecated parameter. For details, see **ProjectionExpression**. |
| | | ConsistentRead | No | No | Default strong consistency |
| | | ExclusiveStartKey | No | Yes | - |
| | | ExpressionAttributeNames | No | Yes | - |
| | | ExpressionAttributeValues | No | Yes | - |
| | | FilterExpression | No | Yes | - |
| | | IndexName | No | Yes | - |
| | | KeyConditionExpression | No | Yes | - |
| | | KeyConditions | No | No | Deprecated parameter. For details, see **KeyConditionExpression**. |
| | | Limit | No | Yes | - |
| | | ProjectionExpression | No | Yes | - |

| API | Parameter | | Mandatory | Supported | Remarks |
|---|---|---|---|---|---|
| | | QueryFilter | No | No | Deprecated parameter. For details, see **FilterExpression**. |
| | | ReturnConsumedCapacity | No | No | Billing parameter, which is not supported |
| | | ScanIndexForward | No | Yes | - |
| | | Select | No | Yes | COUNT is not supported. |
| | Response parameter | ConsumedCapacity | - | No | Billing parameter, which is not supported |
| | | Count | - | Yes | - |
| | | Items | - | Yes | - |
| | | LastEvaluatedKey | - | Yes | - |
| | | ScannedCount | - | Yes | - |

**Table 1-14** Scan

| API | Parameter | | Mandatory | Supported | Remarks |
|---|---|---|---|---|---|
| Scan | Request parameter | TableName | Yes | Yes | - |
| | | AttributesToGet | No | No | Deprecated parameter. For details, see **ProjectionExpression**. |
| | | ConditionalOperator | No | No | Deprecated parameter. For details, see **ConditionExpression**. |
| | | ConsistentRead | No | No | Default strong consistency |
| | | ExclusiveStartKey | No | Yes | - |

| API | Parameter | | Mandatory | Supported | Remarks |
|---|---|---|---|---|---|
| | | ExpressionAttributeNames | No | Yes | - |
| | | ExpressionAttributeValues | No | Yes | - |
| | | FilterExpression | No | Yes | - |
| | | IndexName | No | Yes | - |
| | | Limit | No | Yes | - |
| | | ProjectionExpression | No | Yes | - |
| | | ReturnConsumedCapacity | No | No | Billing parameter, which is not supported |
| | | ScanFilter | No | No | Deprecated parameter. For details, see **FilterExpression**. |
| | | Segment | No | Yes | - |
| | | Select | No | Yes | COUNT is not supported. |
| | | TotalSegments | No | Yes | - |
| | Response parameter | ConsumedCapacity | No | No | Billing parameter, which is not supported |
| | | Count | - | Yes | - |
| | | Items | - | Yes | - |
| | | LastEvaluatedKey | - | Yes | - |
| | | ScannedCount | - | Yes | - |

**Table 1-15** UpdateTimeToLive

| API | Parameter | | Mandatory | Supported | Remarks |
|---|---|---|---|---|---|
| UpdateTimeToLive | Request parameter | TableName | Yes | Yes | - |
| | | TimeToLiveSpecification | Yes | Yes | - |

| API | Parameter | | Mandatory | Supported | Remarks |
| --- | --- | --- | --- | --- | --- |
| | Response parameter | TimeToLiveSpecification | - | Yes | - |

**Table 1-16** DescribeTimeToLive

| API | Parameter | | Mandatory | Supported | Remarks |
| --- | --- | --- | --- | --- | --- |
| DescribeTimeToLive | Request parameter | TableName | Yes | Yes | - |
| | Response parameter | TimeToLiveSpecification | - | Yes | - |

# 1.3 Instance Specifications

Each instance type comes with various specifications based on memory configurations.

This section describes specifications of a GeminiDB DynamoDB-Compatible instance. The instance specifications depend on the selected CPU model.

**Table 1-17** Specifications of a GeminiDB DynamoDB-Compatible instance with cloud native storage

| Data Node Flavor | vCPUs | Memory (GB) | Min. Storage Space (GB) | Max. Storage Space (GB) |
| --- | --- | --- | --- | --- |
| geminidb.cassandra-geminifs.large.4 | 2 | 8 | 10 | 64000 |
| geminidb.cassandra-geminifs.xlarge.4 | 4 | 16 | 10 | 64000 |
| geminidb.cassandra-geminifs.2xlarge.4 | 8 | 32 | 10 | 64000 |
| geminidb.cassandra-geminifs.4xlarge.4 | 16 | 64 | 10 | 64000 |

| Data Node Flavor | vCPUs | Memory (GB) | Min. Storage Space (GB) | Max. Storage Space (GB) |
|---|---|---|---|---|
| geminidb.cassandra-geminifs.8xlarge.4 | 32 | 128 | 10 | 64000 |

# 1.4 Instance Statuses

The status of an instance indicates the health of the instance. You can view the status of an instance on the console.

**Table 1-18** Instance statuses

| Status | Description |
|---|---|
| Available | The DB instance is available. |
| Abnormal | The instance is abnormal. |
| Creating | The instance is being created. |
| Creation failed | DB instance creation fails. |
| Restarting | The instance is being restarted. |
| Resetting password | The administrator password is being reset. |
| Adding node | Nodes are being added to an instance. |
| Deleting node | Nodes are being deleted from an instance. |
| Scaling storage space | The storage space of an instance is being scaled up. |
| Changing specifications | The vCPUs and memory of an instance are being changed. |
| Uploading backup | The backup file is being uploaded. |
| Backing up | A database backup is being created. |
| Checking restoration | The backup of the instance is being restored to a new instance. |
| Changing to yearly/monthly | The billing mode is being changed from pay-per-use to yearly/monthly. |
| Changing to pay-per-use | The billing mode is being changed from yearly/monthly to pay-per-use. |

| Status | Description |
| --- | --- |
| Creating a DR cluster | A DR instance is being created. |
| Canceling DR relationship | A DR instance is being deleted. |
| Configuring SSL | SSL is being enabled or disabled. |
| Frozen | The instance is frozen because your balance drops to or below zero. |
| Unfreezing | Overdue payments are cleared, and the DB instance is being unfrozen. |
| Checking changes | The yearly/monthly instance is pending check when its billing mode is changed. |

# 1.5 Constraints

The following tables list the constraints designed to ensure stability and security of GeminiDB DynamoDB-Compatible instances.

## Specifications

**Table 1-19** Specifications

| Resource Type | Specifications | Description |
| --- | --- | --- |
| CPU and memory | GeminiDB DynamoDB-Compatible API supports cluster instances. | • For details about specifications of different instance types, see **Instance Specifications**.<br>• You can change the specifications to meet your service requirements by following **Changing vCPUs and Memory**. |
| Storage | The storage space depends on the selected **instance specifications**. | Storage can be scaled up or down. For details, see **Overview**. |

## Quotas

**Table 1-20** Quotas

| Resource Type | Constraint | Description |
| --- | --- | --- |
| Tag | A maximum of 20 tags can be added for each instance. | For more information, see **Tag Management**. |
| Free backup storage | GeminiDB DynamoDB-Compatible instances provide free additional backup storage of the same size as you purchased. | For more information, see **Backup Storage**. |
| Retention period | The default value is 7 days. The value ranges from 1 to 3660 days. | For more information, see **Configuring an Automated Backup Policy**. |

## Naming Rules

**Table 1-21** Naming rules

| Item | Description |
| --- | --- |
| Instance name | • Can contain 4 to 64 characters.<br>• Must start with a letter. Only letters (case-sensitive), digits, hyphens (-), and underscores (_) are allowed. |
| Backup name | • Can contain 4 to 64 characters.<br>• Must start with a letter. Only letters (case sensitive), digits, hyphens (-), and underscores (_) are allowed. |
| Parameter template name | • Can contain 1 to 64 characters.<br>• Only letters (case sensitive), digits, hyphens (-), underscores (_), and periods (.) are allowed. |

## Security

**Table 1-22** Security

| Item | Description |
| --- | --- |
| Password of database administrator **rwuser** | • Can contain 8 to 32 characters.<br>• Can contain at least two types of the following characters: uppercase letters, lowercase letters, digits, and special characters ~!@#%^*-_=+? For more information, see **Resetting the Administrator Password**.<br>• Keep your password secure. The system cannot retrieve it if it is lost. |
| Database port | Database port number.<br>Port 8000 is used by default and cannot be changed for a GeminiDB DynamoDB-Compatible instance. |
| VPC | After a GeminiDB DynamoDB-Compatible instance is created, the VPC where the instance is deployed cannot be changed. |
| Security group | A security group controls access between GeminiDB DynamoDB-Compatible API and other services. Ensure that the security group you selected allows your client to access the instance.<br>If no security group is available, the system creates one for you. |

## Instance Operations

**Table 1-23** Instance operations

| Function | Constraint |
| --- | --- |
| Database access | • If remote access is not enabled, GeminiDB DynamoDB-Compatible instances and their associated ECSs must be in the same VPC subnet.<br>• The security group must allow access from the associated ECS.<br>By default, a GeminiDB DynamoDB-Compatible instance cannot be accessed through an ECS in a different security group. You need to add an inbound rule to the security group.<br>• The default port of a GeminiDB DynamoDB-Compatible instance is 8000.<br>• The database port cannot be set during instance creation and cannot be changed after the instance is created. |
| Instance deployment | The servers where instances are deployed are not directly visible to you. You can only access the instances through IP addresses and database ports. |
| Restarting a GeminiDB DynamoDB-Compatible instance | • GeminiDB DynamoDB-Compatible instances cannot be restarted through commands. They must be restarted on the console.<br>• To avoid downtime, you are advised to restart an instance during off-peak hours. Ensure that your application can be reconnected. |
| Viewing backups of a GeminiDB DynamoDB-Compatible instance | GeminiDB DynamoDB-Compatible instance backups are stored in OBS buckets and are transparent to you. |
| Changing the CPU or memory of a GeminiDB DynamoDB-Compatible instance | • Second-level intermittent disconnection occurs once when the specifications are changed on a single node. Therefore, the entire instance is intermittently disconnected several times. Ensure that the client can be reconnected. You are advised to change the specifications during off-peak hours.<br>• For a node whose specifications are being changed, its computing tasks are handed over to other nodes. Change specifications of nodes during off-peak hours to prevent instance overload. |
| Data restoration | To prevent data loss, you are advised to back up key data before data restoration. |

| Function | Constraint |
| --- | --- |
| Storage | If the instance storage is full, data cannot be written to databases. You are advised to periodically check the storage.<br>The storage can be automatically scaled up in case of a sudden surge in data volumes. Enable autoscaling by following **Automatically Scaling Up Storage Space**. |
| Recycle bin | • You can move unsubscribed yearly/monthly instances and deleted pay-per-use instances to the recycle bin. You can restore an instance that was deleted up to 7 days ago from the recycle bin.<br>• The recycle bin is enabled by default and cannot be disabled. Instances in the recycle bin can be retained for 7 days by default. This function is free of charge.<br>• Currently, you can put a maximum of 100 instances into the recycle bin.<br>• If you delete an instance running out of storage, it will not be moved to the recycle bin. |

# 2 Billing

## 2.1 Billing Overview

In this document, you will learn about how a GeminiDB DynamoDB-Compatible instance is billed, how you can renew subscriptions and manage costs, and what happens if your account goes into arrears.

- **Billing modes**

  There are two billing modes: yearly/monthly and pay-per-use.

  - Yearly/Monthly: You are billed in advance based on the duration you plan to use the instance. Ensure your top-up account has enough funds or a valid payment method pre-configured.
  - Pay-per-use: You are billed for your immediate needs based on how long you have actually used a product.

  For details about the two billing modes, see **Overview**.

  You can also change the billing mode later if it no longer meets your needs. For details, see **Overview**.

- **Billing items**

  You will be billed for GeminiDB DynamoDB-Compatible instance specifications, storage, backup storage, and EIP bandwidth. For details about the billing factors and formulas of each billing item, see **Billing Items**.

  For more information about the billing samples and the billing for each item, see **Billing Examples**.

- **Renewal**

  After a yearly/monthly instance expires, it cannot run properly. To continue using the instance, renew it within the specified period. Otherwise, compute and storage resources will be automatically released and data may be lost.

  You can renew the instance manually or automatically. For details, see **Overview**.

- **Bills**

  You can choose **Billing & Costs** > **Bills** to check the instance transactions and bills. For details, see **Bills**.

- **Arrears**

  If the available account balance is less than amount to be settled and there is no other payment method configured, your account will go into arrears. To continue using your cloud services, top up your account in a timely manner. For details, see **Arrears**.

- **Billing termination**

  If your instance is no longer used, you can unsubscribe from or delete it to avoid unexpected bills. For details, see **Billing Termination**.

- **Cost management**

  Runtime costs are primarily composed of resource and O&M costs. You can allocate, analyze, and optimize instance costs to save more money. For details, see **Cost Management**.

# 2.2 Billing Modes

## 2.2.1 Overview

There are two billing modes: yearly/monthly and pay-per-use.

- Yearly/Monthly is a prepaid billing mode. You pay in advance for a subscription term, and in exchange, you get a discounted rate. The longer the subscription term, the bigger the discount. Yearly/Monthly billing is a good option for long-term, stable services.
- Pay-per-use is a postpaid billing mode. You pay as you go and just pay for what you use. The instance usage is calculated by the second but billed every hour. Pay-per-use billing is a good option for scenarios where there are sudden traffic bursts, such as e-commerce promotions.

**Table 2-1** lists their differences.

**Table 2-1** Billing modes

| **Billing Mode** | Yearly/Monthly | Pay-per-use |
| --- | --- | --- |
| **Payment** | Prepaid<br>Settled based on the subscription period specified in your order. | Postpaid<br>Settled based on how long you have actually used your instance. |
| **Billing Cycle** | Billed based on the subscription period specified in your order. | Billed by the second and settled by the hour. |
| **Billing Item** | Instance specifications (vCPUs and memory), storage, backup storage, and EIP | Instance specifications (vCPUs and memory), storage, backup storage, and EIP |

| | | |
|---|---|---|
| **Billing Mode Change** | A yearly/monthly instance can be changed to pay-per-use. The change is applied only after the yearly/monthly instance expires. For details, see **Changing a Yearly/Monthly Instance to Pay-per-Use**. | A pay-per-use instance can be changed to yearly/monthly. For details, see **Changing a Pay-per-Use Instance to Yearly/Monthly**. |
| **Specification Change** | Supported | Supported |
| **Scenario** | Maintains a long-term commitment and offers significant discounts compared to pay-per-use billing. | Aligns with your scaling needs of compute resources. You can create or delete your instance anytime. |

## 2.2.2 Yearly/Monthly Billing

Yearly/Monthly is ideal for users who are confident in their long-term needs and want to secure a lower price. You will enjoy a discount on a yearly/monthly instance. This section describes the billing rules of a yearly/monthly GeminiDB DynamoDB-Compatible instance.

### Application Scenarios

If you want to ensure resource stability over a certain period of time, yearly/monthly billing is a good choice for the following types of workloads:

- Long-term workloads with stable resource requirements, such as official websites, online malls, and blogs.
- Long-term projects, such as scientific research projects and large-scale events.
- Workloads with predictable traffic bursts, for example, e-commerce promotions or festivals.
- Workloads with high data security requirements.

### Billing Items

You are billed for the following items on a yearly/monthly basis.

**Table 2-2** Items billed on a yearly/monthly basis

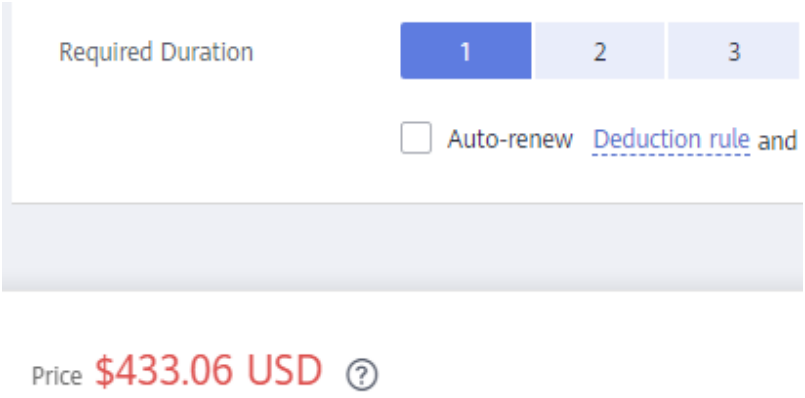
| Billing Item | Description |
|---|---|
| Instance specifications | vCPUs and memory |

| Billing Item | Description |
| --- | --- |
| Storage | If the actual storage usage exceeds your purchased storage, you will be billed for additional storage on a pay-per-use basis. |
| Backup storage | GeminiDB DynamoDB-Compatible API provides free backup storage equal to the amount of storage you purchased.<br>After the free backup storage is used up, additional usage will be priced by the hour based on the backup storage pricing details. If it has been used less than one hour, you will be billed based on the actual duration. |
| (Optional) EIP bandwidth | You will be billed for the bandwidth you use to access GeminiDB DynamoDB-Compatible instances over a public network. Accessing the instances over a private network will not be billed. |

Assume that you want to buy a three-node GeminiDB DynamoDB-Compatible instance with 2 vCPUs, 8 GB of memory, and 100 GB of storage. At the bottom of the page for buying an instance, price details (excluding the backup storage fee) will be displayed.

**Figure 2-1** Example price




You are billed for:

- Selected specifications
- Storage

NOTE

The backup space fee is not included. For details about the backup price, see **Product Pricing Details**.

**Backup Storage Space**

| DB Instance Type | Hourly | Currency |
| --- | --- | --- |
| Cluster | 0.00004 | Price per GB |

## Billing Cycle

The billing cycle of a yearly/monthly GeminiDB DynamoDB-Compatible instance is determined by the duration (UTC+8) you commit to the subscription. The billing starts from when you activated or renewed the subscription and ends at 23:59:59 of the expiry date.

For example, if you bought a GeminiDB DynamoDB-Compatible instance for one month at 15:50:04 on March 8, 2023, the billing cycle is from 15:50:04 on March 8, 2023 to 23:59:59 on April 8, 2023.

## Billing Examples

Assume that you bought a three-node GeminiDB DynamoDB-Compatible instance with 2 vCPUs, 8 GB of memory, 100 GB of storage, 110 GB (100 GB for free) of backup storage for one month at 15:50:04 on March 8, 2023 and renewed the subscription for one more month before it expired. The billing items include instance specifications (vCPUs, memory, and nodes), storage, backup storage, and EIP bandwidth.

- The first billing cycle is from 15:50:04 on March 8, 2023 to 23:59:59 on April 8, 2023.
- The second billing cycle is from 23:59:59 on April 8, 2023 to 23:59:59 on May 8, 2023.
  - From 23:59:59 on April 8, 2023 to 23:59:59 on May 1, 2023, 50 GB of free backup storage was used.
  - From 23:59:59 on May 1, 2023 to 23:59:59 on May 8, 2023, another 10 GB of backup storage was billed for 168 hours.

You need to pay in advance for each billing cycle. Each resource is billed separately.

**Table 2-3** Billing formulas

| Resource | Formula | Unit Price |
| --- | --- | --- |
| Instance specifications (including vCPUs and memory) | Unit price of the instance specifications x Required duration x Number of nodes | See the estimated price of a cluster instance in **GeminiDB Price Calculator**. |

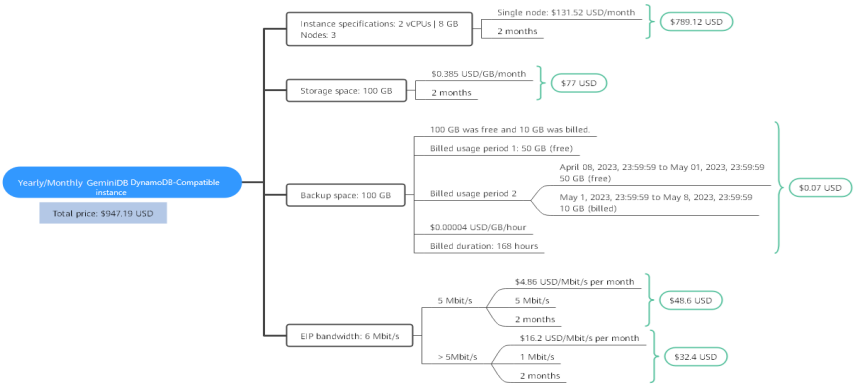
| Resource | Formula | Unit Price |
| --- | --- | --- |
| Storage space | Storage space unit price x Required duration x Storage space (GB) | See the estimated price of a cluster instance with specified storage in **GeminiDB Price Calculator**. |
| Backup storage | Backup storage unit price x Required duration x (Backup storage – Storage) (GB)<br>**NOTE**<br>The billing duration indicates how long the storage exceeding a free quota was used. | See the estimated price of an instance with specified backup storage in **GeminiDB Price Calculator**. |
| EIP bandwidth | Billed by fixed bandwidth | For details, see **EIP Price Calculator**. |

**Figure 2-2** shows how the total price is calculated.

NOTE

The prices in the following figure are for reference only. For the actual prices, see **GeminiDB Price Calculator**.

**Figure 2-2** Pricing example of a yearly/monthly GeminiDB DynamoDB-Compatible instance




## Impact on Billing After Specification Changes

You can change the specifications on the console to meet your service requirements. The system will recalculate the price and either bill or refund you the difference.

- If you increase instance specifications, you need to pay the difference in price.
- If you decrease instance specifications, Huawei Cloud will refund you the difference.

Decreasing instance specifications will affect instance performance. You are not advised to do so. Assume that you bought a yearly/monthly three-node GeminiDB DynamoDB-Compatible instance with 2 vCPUs and 8 GB of memory for one month on April 8, 2023 and increased its specifications to 4 vCPUs and 16 GB of memory on April 18, 2023. The old specifications cost USD587.06/month and the new specifications USD981.62/month. The calculation formula is as follows:

**Price difference = Price of new specifications x Remaining period – Price of old specifications x Remaining period**

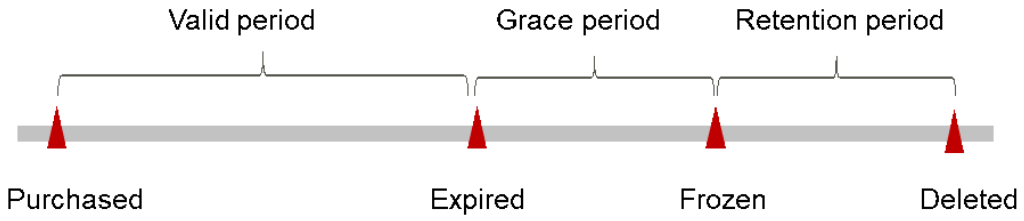
The remaining period in the formula is the number of remaining days in each calendar month divided by the maximum number of days in the calendar month. In this example, the remaining period is 0.6581 (12/30 + 8/31). The fee for increasing specifications is USD259.66 (981.62 x 0.6581 – 587.06 x 0.6581).

For more details, see **Pricing of a Changed Specification**.

## Impact of Expiration

**Figure 2-3** shows the statuses of a yearly/monthly GeminiDB DynamoDB-Compatible instance. After an instance is purchased, it enters the valid period and runs normally during this period. If the instance is not renewed after it expires, before being deleted, it first enters a grace period and then a retention period.

**Figure 2-3** Lifecycle of a yearly/monthly GeminiDB DynamoDB-Compatible instance




### Expiration reminder

Before a yearly/monthly GeminiDB DynamoDB-Compatible instance expires, the system will send an expiration reminder to the creator of the Huawei Cloud account by email, SMS, and in-app message.

- Expiration notifications will be sent 30 days, 15 days, 7 days, 3 days, and 1 day before yearly resources expire.
- Expiration notifications will be sent 15 days, 7 days, 3 days, and 1 day before monthly resources expire.

### Impact

If your yearly/monthly instance is not renewed after it expires, its status changes to **Expired** and it enters a grace period. During the grace period, you can access the instance but cannot:

- Change instance specifications.
- Change the billing mode from yearly/monthly to pay-per-use.

- Unsubscribe from the instance.

If the yearly/monthly instance is not renewed after the grace period ends, its status turns to **Frozen** and it enters a retention period. You cannot perform any operations on the instance while it is in the retention period.

If the yearly/monthly instance is not renewed by the time the retention period ends, it will be released and data cannot be restored.

NOTE

- For details about the renewal, see **Overview**.

## 2.2.3 Pay-per-use Billing

Pay-per-use billing means you pay nothing up front and are not tied into any contract or commitment. This section describes the billing rules of a pay-per-use GeminiDB DynamoDB-Compatible instance.

### Application Scenarios

Pay-per-use billing is good for short-term, bursty, or unpredictable workloads that cannot tolerate any interruptions, such as applications for e-commerce flash sales, temporary testing, and scientific computing.

### Billing Items

You are billed for the following items on a pay-per-use basis.

**Table 2-4** Items billed on a pay-per-use basis

| Billing Item | Description |
| --- | --- |
| Instance specifications | vCPUs and memory |
| Storage | Instance storage space, which is billed hourly on a pay-per-use basis. |
| Backup storage | GeminiDB DynamoDB-Compatible API provides free backup storage equal to the amount of storage you purchased.<br>After the free backup storage is used up, additional usage will be priced by the hour based on the backup storage pricing details. If it has been used less than one hour, you will be billed based on the actual duration. |
| (Optional) EIP bandwidth | You will be billed for the bandwidth when accessing GeminiDB DynamoDB-Compatible instances over a public network. Accessing the instances over a private network will not be billed. |

Assume that you want to buy a three-node GeminiDB DynamoDB-Compatible instance with 2 vCPUs, 8 GB of memory, and 500 GB of storage. At the bottom of

the page for buying an instance, price details (excluding the backup storage fee) will be displayed.

**Figure 2-4** Example price


You are billed for:

- Instance specifications (including vCPUs and memory)
- Selected storage space

NOTE

The backup space fee is not included. For details about the backup price, see **Product Pricing Details**.

**Backup Storage Space**

| DB Instance Type | Hourly | Currency |
| --- | --- | --- |
| Cluster | 0.00004 | Price per GB |

## Billing Cycle

Usage of a pay-per-use instance is calculated by the second and priced by the hour (GMT+8). The billing starts when the instance is created and ends when it is deleted.

NOTE

It takes a certain time to create an instance. The billing starts from the time when the instance is successfully created. You can view the two time points on the **Basic Information** page. You can view the time when the instance is created beside the **Created** field.

For example, if you buy a pay-per-use GeminiDB DynamoDB-Compatible instance at 8:45:30 and delete it at 8:55:00, you are billed for the 600 seconds from 8:45:30 to 8:55:30. The billing items include compute resources (vCPUs and nodes), storage, and backup storage.

## Billing Examples

Assume that you bought a pay-per-use 3-node instance with 2 vCPUs, 8 GB of memory, 100 GB of storage, and 110 GB of backup storage (100 GB for free) at 09:59:30 on April 18 and deleted the instance at 10:45:46 on April 18, 2023. The billing items include compute resources (vCPUs and nodes) and storage.

- Usage of 30 seconds from 9:59:30 to 10:00:00
- Usage of 2,746 seconds from 10:00:00 to 10:45:46
  - The free backup storage is used from 10:00:00 to 10:45:00.

- 10 GB of backup storage is billed for 46 seconds from 10:45:00 to 10:45:46.

You will be billed for each billing cycle. Resources of the GeminiDB DynamoDB-Compatible instance are billed separately. The price per hour is shown in the pricing details. To calculate the price per second, divide it by 3,600.

**Table 2-5** Billing formulas

| Resource | Formula | Unit Price |
| --- | --- | --- |
| Compute resources (including vCPUs and nodes) | Unit price of instance specifications x Required duration | See the estimated price of a cluster instance in **GeminiDB Price Calculator**. |
| Storage | Storage unit price x Required duration | See the estimated price of a cluster instance with specified storage in **GeminiDB Price Calculator**. |
| Backup storage | Backup storage unit price x Required duration x (Backup storage – Storage) (GB)<br>**NOTE**<br>The billing duration indicates how long the storage exceeding a free quota was used. | See the estimated price of a cluster instance with specified backup storage in **GeminiDB Price Calculator**. |
| Public network traffic | Tiered billing by fixed bandwidth<br>• 0 Mbit/s to 5 Mbit/s (included): billed at a fixed unit price per Mbit/s<br>• Greater than 5 Mbit/s: billed at a different price per Mbit/s | For details, see the estimated bandwidth price in **ECS Price Calculator** or **EIP Price Calculator**. |

**Figure 2-5** shows how the total price is calculated.

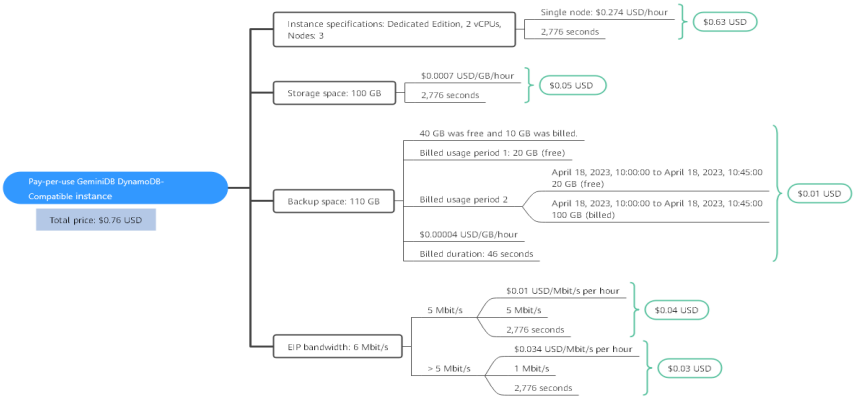
NOTE

The prices in the following figure are for reference only. For the actual prices, see **GeminiDB Price Calculator**.

If the price is not an integer, it is rounded off to the nearest two decimal places. If the rounded price is less than USD0.01, USD0.01 will be displayed.

**Figure 2-5** Billing example of a pay-per-use GeminiDB DynamoDB-Compatible instance




## Impact on Billing After Specification Changes

If you change the specifications of a pay-per-use instance, the original order will become invalid and a new order will be placed. You will be billed based on the new specifications.

If you change instance specifications within a given hour, multiple records will be generated. Different records record the billing for different specifications.

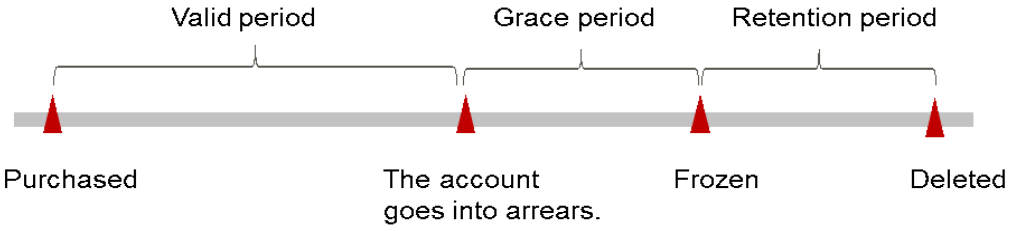
For example, if you buy a pay-per-use instance with 2 vCPUs and 8 GB of memory at 9:00:00 and increase the specifications to 4 vCPUs and 16 GB of memory at 9:30:00, two billing records will be generated between 9:00:00 and 10:00:00:

- 2 vCPUs and 8 GB of memory from 9:00:00 to 9:30:00
- 4 vCPUs and 16 GB of memory from 9:30:00 to 10:00:00

## Impact of Arrears

Figure 2-6 shows the statuses of a pay-per-use GeminiDB DynamoDB-Compatible instance. After you buy a pay-per-use GeminiDB DynamoDB-Compatible instance, it is valid within the billing cycle. If your account falls into arrears because of automatic fee deduction, it will transition into a grace period followed by a retention period.

**Figure 2-6** Lifecycle of a pay-per-use GeminiDB DynamoDB-Compatible instance




**Arrears reminder**

The system will bill you for pay-per-use resources after each billing cycle ends. If your account goes into arrears, the system will send an email, SMS message, or in-app message to the one who created the Huawei Cloud account.

**Impact**

When your account falls into arrears due to automatic fee deduction for the pay-per-use GeminiDB DynamoDB-Compatible instance, your account will be marked as in arrears, and the instance will enter the grace period. After you top up your account, Huawei Cloud will bill you for expenditures generated during the grace period. You can view the charges on the **Billing Center** > **Overview** page.

If you do not pay the arrears within the grace period, your instance enters the retention period and its status changes to **Frozen**. You cannot perform any operations on the instance in the retention period.

If you do not pay the arrears within the retention period, your instance will be released, and data will be lost.

NOTE

- During the retention period, you cannot access or use your instance but the data stored in it can be retained. The retention period for Huawei Cloud International website is 15 days.
- During the grace period, you can access and use only some resources of your instance. The grace period for Huawei Cloud International website is 15 days.
- For details about top-up, see **Topping Up an Account**.

# 2.3 Billing Items

## Billing

You will be billed for GeminiDB DynamoDB-Compatible instance specifications, storage, backup storage, and public traffic. For details, see **Table 2-6**.

NOTE

Billing items marked with an asterisk (*) are mandatory.

**Table 2-6** Billing items of GeminiDB DynamoDB-Compatible instances

| Billing Item | Description | Billing Mode | Formula |
| --- | --- | --- | --- |
| * Specifications | Billed by instance specifications, including vCPUs and memory. Compute and storage capabilities vary by the number of vCPUs and memory size. | Yearly/Monthly and pay-per-use | **Unit price x Required duration**<br>See the estimated price of a cluster instance in **GeminiDB Price Calculator**. |

| Billing Item | Description | Billing Mode | Formula |
| --- | --- | --- | --- |
| * Storage space | Billed based on unified standards. | Yearly/ Monthly and pay-per-use | **Unit price x Storage space x Required duration**<br>See the estimated price of a cluster instance with specified storage in **GeminiDB Price Calculator**. |
| Backup storage | Billed based on unified standards. | Pay-per-use | **Unit price x Billed backup storage x Required duration**<br>See the estimated price of a cluster instance with specified backup storage in **GeminiDB Price Calculator**.<br>**NOTE**<br>The billing duration indicates how long the storage exceeding a free quota was used. |
| (Optional) Cross-region backup | Billed based on unified standards. | Pay-per-use | **Unit price x Storage x Required duration**<br>Unit price of storage: CNY0.0009/GB/hour |
| | Billed based on unified standards. | Billed by storage | **Unit price of cross-region backup traffic x Storage**<br>Unit price of cross-region backup traffic: CNY0.5/GB |

| Billing Item | Description | Billing Mode | Formula |
| --- | --- | --- | --- |
| Public network traffic | An EIP is required if an instance needs to access the Internet.<br>Billing factors: bandwidth, traffic, and IP reservation<br>• EIP for a yearly/monthly instance: billed by bandwidth.<br>• EIP for a pay-per-use instance: billed by bandwidth, traffic, or shared bandwidth. You are also billed for IP reservation. | Yearly/ Monthly and pay-per-use.<br>You can purchase a bandwidth add-on package or a shared traffic package. | **Tiered pricing based on fixed bandwidth**.<br>• 0 Mbit/s to 5 Mbit/s (included): billed at a fixed unit price per Mbit/s.<br>• Greater than 5 Mbit/s: billed at a different price per Mbit/s.<br>For details, see the estimated EIP bandwidth unit price in **ECS Price Calculator** or **EIP Price Calculator**. |

## Billing Examples

Assume that you bought a three-node GeminiDB DynamoDB-Compatible instance with 2 vCPUs, 8 GB of memory, 100 GB of storage, 110 GB (100 GB for free) of backup storage for one month at 15:50:04 on March 8, 2023 and renewed the subscription for one more month before it expired. The billing items include instance specifications (vCPUs, memory, and nodes), storage, backup storage, and EIP bandwidth.

- The first billing cycle is from 15:50:04 on March 8, 2023 to 23:59:59 on April 8, 2023.
- The second billing cycle is from 23:59:59 on April 8, 2023 to 23:59:59 on May 8, 2023.
  - From 23:59:59 on April 8, 2023 to 23:59:59 on May 1, 2023, 50 GB of free backup storage was used.
  - From 23:59:59 on May 1, 2023 to 23:59:59 on May 8, 2023, another 10 GB of backup storage was billed for 168 hours.

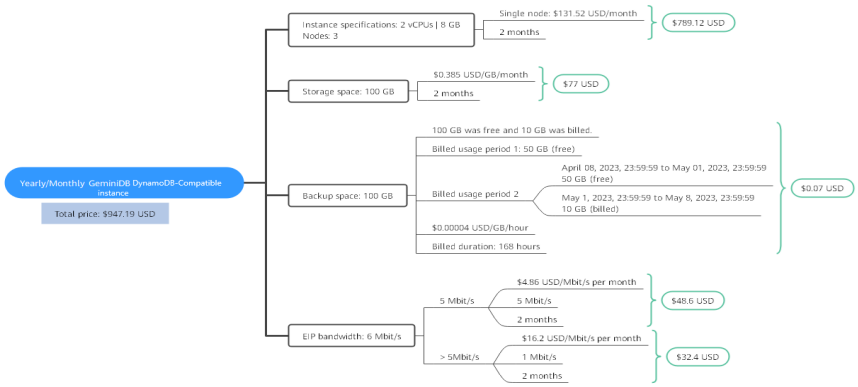
**Figure 2-7** shows how the total price is calculated.

NOTE

The prices in the following figure are for reference only. For the actual prices, see **GeminiDB Price Calculator**.

**Figure 2-7** Pricing example of a yearly/monthly GeminiDB DynamoDB-Compatible instance




For more billing examples of a pay-per-use instance, see **Billing Examples**.

# 2.4 Billing Examples

## Billing Scenario

A user bought a pay-per-use GeminiDB DynamoDB-Compatible instance with the following configurations at 15:30:00 on March 18, 2023:

- 2 vCPUs and 8 GB of memory
- Nodes: 3
- EIP bandwidth: 6 Mbit/s

After a period, the user found that the current instance specifications could not meet service requirements and increased the specifications to 4 vCPUs and 16 GB at 9:00:00 on March 20, 2023. The user changed the billing mode to yearly/monthly at 10:30:00 on March 20, 2023 and paid the instance in advance for one month. So how much will the user be billed for this instance in March and April?

## Billing Analysis

This GeminiDB DynamoDB-Compatible instance is billed in two phases:

- 15:30:00 on March 18, 2023 to 10:30:00 on March 20, 2023: pay-per-use
  - 15:30:00 on March 18, 2023 to 9:00:00 on March 20, 2023
    - Instance specifications: 2 vCPUs and 8 GB of memory
    - Nodes: 3
    - Storage: 100 GB
    - Backup storage: 100 GB
    - EIP bandwidth: 6 Mbit/s
  - 9:00:00 on March 20, 2023 to 10:30:00 on March 20, 2023

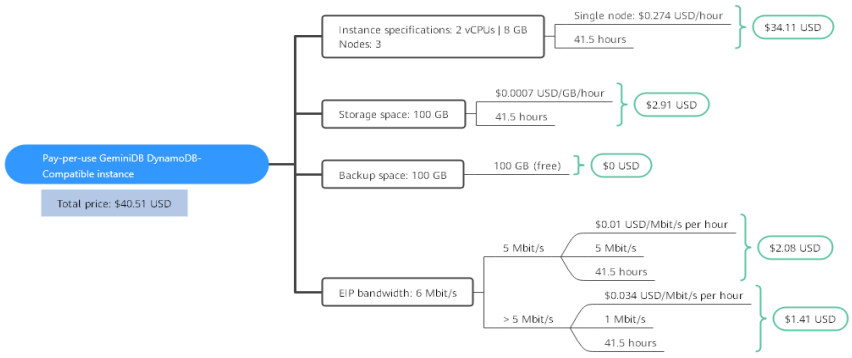
    - Instance specifications: 4 vCPUs and 16 GB of memory
    - Nodes: 3
    - Storage: 200 GB
    - Backup storage: 210 GB (pay-per-use from 10:00:00 to 10:30:00 on March 20, 2023)
    - EIP bandwidth: 6 Mbit/s
- 10:30:00 on Mar 20, 2023 to 23:59:59 on Apr 20, 2023: yearly/monthly
  - Instance specifications: 4 vCPUs and 16 GB of memory
  - Nodes: 3
  - Storage: 200 GB
  - Backup storage: 300 GB (pay-per-use from 23:59:59 on April 10, 2023 to 23:59:59 on April 20, 2023)
  - EIP bandwidth: 6 Mbit/s
  - Billing duration: one month

NOTE

Unit prices in this example are used for reference only, and the calculated prices are only estimates. The actual unit price and cost may vary. For details, see the data released on the Huawei Cloud official website.

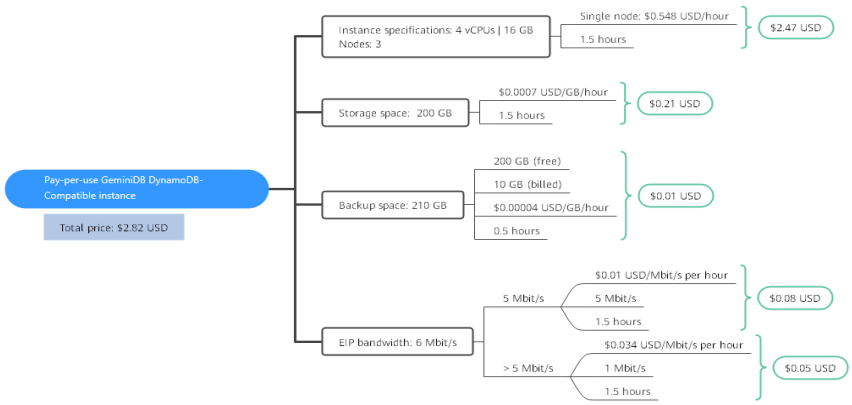
**Pay-per-use**

From 15:30:00 on March 18, 2023 to 09:00:00 on March 20, 2023, an instance with 2 vCPUs and 8 GB of memory was used for 41.5 hours, so the price is calculated as follows.



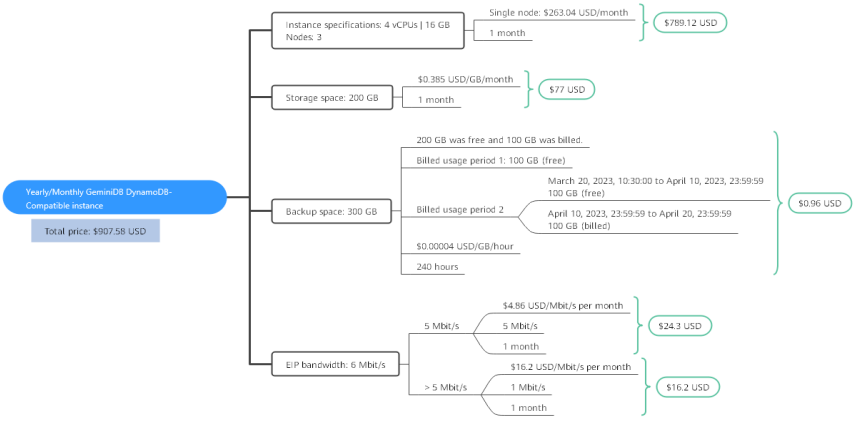

From 9:00:00 to 10:30:00 on March 20, 2023, an instance with 4 vCPUs and 16 GB of memory was used for 1.5 hours, so the price is calculated as follows.




**Yearly/Monthly**

From 10:30:00 on March 20, 2023 to 23:59:59 on April 20, 2023, the yearly/monthly instance was used for one month. The price is calculated as follows.




From March to April, the total price of this instance is USD950.91 (40.51 + 2.82 + 907.58).

# 2.5 Billing Mode Changes

## 2.5.1 Overview

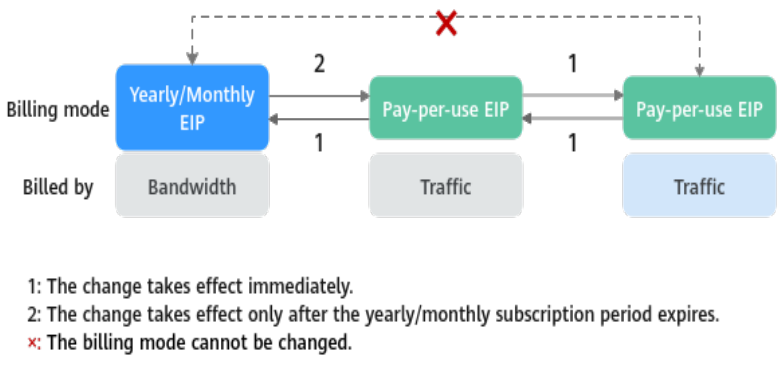
You can change the billing mode if it no longer meets your needs. **Table 2-7** lists billing items that can be changed.

**Table 2-7** Changeable billing items of a GeminiDB DynamoDB-Compatible instance

| Billing Item | Change Description | Reference |
| --- | --- | --- |
| Instance specifications (vCPUs and nodes) | If you change the billing mode of a GeminiDB DynamoDB-Compatible instance, the billing mode of compute resources (vCPUs and nodes) will also be changed.<br>• Change from pay-per-use to yearly/monthly to enjoy lower prices.<br>• Change from yearly/monthly to pay-per-use to save costs and use the instance more flexibly.<br>**NOTE**<br>Such a change is only applied after the yearly/monthly subscription expires. | • **Changing a Pay-per-Use Instance to Yearly/Monthly**<br>• **Changing a Yearly/Monthly Instance to Pay-per-Use** |
| EIP | • A yearly/monthly EIP can be changed to a pay-per-use EIP billed by bandwidth after the yearly/monthly subscription ends.<br>• A pay-per-use EIP billed by bandwidth can be changed to a yearly/monthly EIP.<br>• Pay-per-use EIPs billed by bandwidth can be changed to pay-per-use EIPs billed by traffic, and pay-per-use EIPs billed by traffic can be changed to pay-per-use EIPs billed by bandwidth.<br>For details, see **Figure 2-8**. | • **Changing a Pay-per-Use Instance to Yearly/Monthly**<br>• **Changing a Yearly/Monthly Instance to Pay-per-Use** |

**Figure 2-8** Changing the EIP billing mode



# 2.5.2 Changing a Pay-per-Use Instance to Yearly/Monthly

To commit to a long subscription and save costs, you can change a pay-per-use GeminiDB DynamoDB-Compatible instance to yearly/monthly. Doing so will create an order. After you pay for the order, yearly/monthly billing will be applied immediately.

Suppose you bought a pay-per-use GeminiDB DynamoDB-Compatible instance at 15:29:16 on April 18, 2023 and changed it to yearly/monthly at 16:30:30 on the same day. After you paid for the order, yearly/monthly billing was applied immediately. On the **Billing Center** > **Billing** page, three line items were generated.

- Pay-per-use expenditures for 15:29:16 to 16:00:00 on April 18, 2023
- Pay-per-use expenditures for 16:00:00 to 16:30:30 on April 18, 2023
- A yearly/monthly expenditure generated at 16:30:30 on April 18, 2023

## Constraints

Resources such as EIPs that are used by an instance may not support the change with this instance. For details about their billing mode change rules and handling methods, see **Table 2-8**.

**Table 2-8** EIP billing mode change rules

| Resource | Billing Mode | Billed By | Bandwidth Type | Changed to Yearly/Monthly Billing with the GeminiDB DynamoDB-Compatible Instance | Handling Measure |
|---|---|---|---|---|---|
| EIP | Pay-per-use | Bandwidth | Dedicated | Supported | Change the EIP to yearly/monthly billing on the EIP console.<br>For details, see **Changing EIP Billing Mode**. |

| Resource | Billing Mode | Billed By | Bandwidth Type | Changed to Yearly/ Monthly Billing with the GeminiDB DynamoDB-Compatible Instance | Handling Measure |
|---|---|---|---|---|---|
| EIP | Pay-per-use | Traffic | Dedicated | Not supported | An EIP that is billed by traffic on a pay-per-use basis cannot be directly changed to be billed on a yearly/monthly basis. To change this:<br>1. Change the EIP to be billed by bandwidth on a pay-per-use basis.<br>2. Change the EIP to be billed on a yearly/ monthly basis.<br>For details, see **Changing EIP Billing Mode**. |

## Prerequisites

- The billing mode of the instance is pay-per-use.
- The instance status is **Available**.

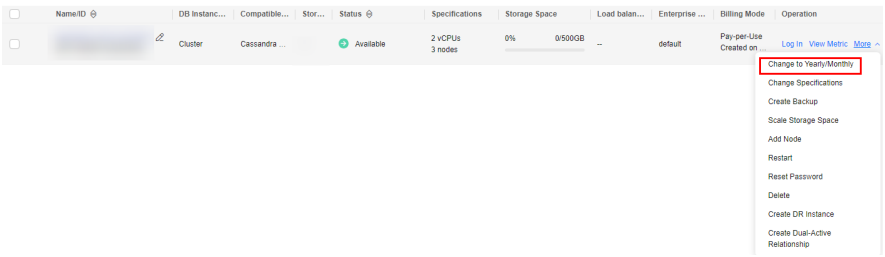
## Procedure

**Step 1** **Log in to the Huawei Cloud console**.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, locate the target instance and click **Change to Yearly/ Monthly** in the **Operation** column.

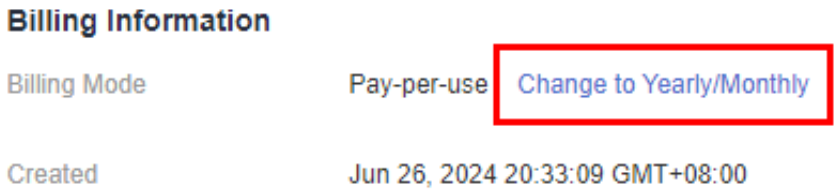
**Figure 2-9** Change to Yearly/Monthly



Alternatively, click the instance name to go to the **Basic Information** page. In the **Billing Information** area, click **Change to Yearly/Monthly** in the **Billing Mode** field.

**Figure 2-10** Change to Yearly/Monthly




NOTE

The billing mode of multiple instances can be changed in batches. Perform the following steps:

1. Select the instances whose billing mode you want to change.
2. Click **Change to Yearly/Monthly** above the instance list.

**Step 4** On the displayed page, specify a subscription duration in month. The minimum duration is one month.

If you do not need to modify your settings, click **Pay** to go to the payment page.

**Step 5** Select a payment method and click **Confirm**.

**Step 6** View the results on the **Instances** page.

In the upper right corner of the instance list, click to refresh the list. The instance status will become **Available** after the change is successful. The billing mode changes to **Yearly/Monthly**.

**----End**

## 2.5.3 Changing a Yearly/Monthly Instance to Pay-per-Use

You can change a yearly/monthly GeminiDB DynamoDB-Compatible instance to pay-per-use to reclaim a portion of your subscription costs while adapting to your usage needs.

Suppose you bought a yearly/monthly GeminiDB DynamoDB-Compatible instance at 15:29:16 on April 18, 2023 and changed it to pay-per-use billing at 16:30:00 on May 18, 2023. On the **Billing Center** > **Billing** page, bill information is generated as follows:

- Yearly/Monthly expenditures for 15:29:16 on April 18 to 23:59:59 on May 18, 2023
- Pay-per-use expenditures for 23:59:59 on May 18, 2023 to the end time of pay-per-use billing. A bill was generated every hour.

NOTE

The pay-per-use billing mode will take effect only after the yearly/monthly subscription has expired. Auto-renewal will not be in effect.

## Constraints

Resources such as EIPs that are used by an instance may not support the change with this instance. For details about their billing mode change rules and handling methods, see **Table 2-9**.

**Table 2-9** EIP billing mode change rules

| Resource | Billing Mode | Billed By | Bandwidth Type | Changed to Pay-per-Use Billing with the GeminiDB DynamoDB-Compatible Instance | Handling Measure |
|---|---|---|---|---|---|
| EIP | Yearly/Monthly | Bandwidth | Dedicated | Not supported | Change the EIP to yearly/monthly billing on the EIP console.<br>For details, see **Changing EIP Billing Mode**. |
| EIP | Yearly/Monthly | Traffic | Dedicated | Not supported | An EIP billed on a yearly/monthly basis cannot be directly changed to be billed by traffic on a pay-per-use basis. To change this:<br>1. Change the EIP to be billed by bandwidth on a pay-per-use basis.<br>2. Change the EIP to be billed by traffic on a pay-per-use basis.<br>For details, see **Changing EIP Billing Mode**. |

## Procedure

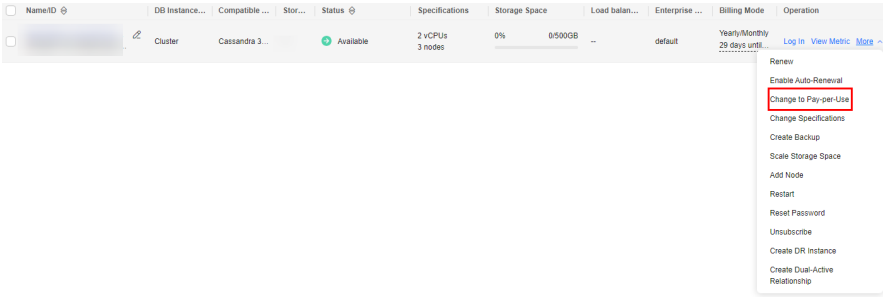
**Step 1** **Log in to the Huawei Cloud console**.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, locate the target instance and choose **More** > **Change to Pay-per-Use** in the **Operation** column.

**Figure 2-11** Change to Pay-per-Use




NOTE

The billing mode of multiple pay-per-use instances can be changed in batches. Perform the following steps:

1. Select the instances whose billing mode you want to change.
2. Click **More** > **Change to Pay-per-Use** in the **Operation** column

**Step 4** On the displayed page, confirm the instance information and click **Change to Pay-per-Use**. The billing mode will change to pay-per-use after the instance expires. After the billing mode is changed, auto-renewal will be disabled.

**Step 5** After you submit the change, check whether a message is displayed in the **Billing Mode** column, indicating that the billing mode will be changed to pay-per-use after the subscription expires.

**Step 6** To cancel the change, choose **Billing** > **Renewal** to enter the Billing Center. On the **Renewals** page, locate the instance and click **More** > **Cancel Change to Pay-per-Use**.

**Step 7** In the displayed dialog box, click **Yes**.

**----End**

# 2.6 Renewing Subscriptions

## 2.6.1 Overview

### When to Renew Subscriptions

Expired yearly/monthly instances cannot run properly. To continue using it, renew your subscription within a specified period. Otherwise, resources such as vCPUs and memory will be automatically released, and data will be lost and cannot be restored.

Only yearly/monthly instance subscriptions can be renewed. If you use pay-per-use instances, just ensure that your account has a valid payment method configured or a top-up account with a sufficient balance.

If you renew the instance before it expires, resources will be retained and you can continue using the instance. For details about statuses of the expired instance, see **Impact of Expiration**.

## How to Renew Subscriptions

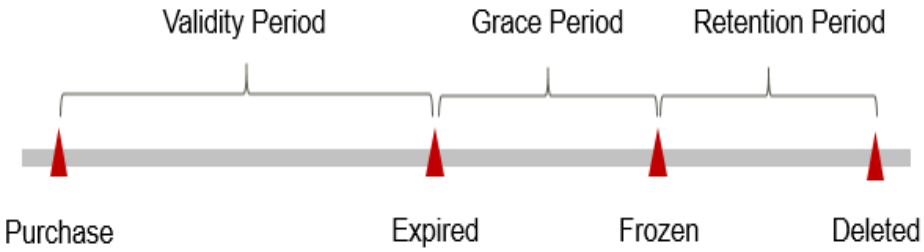
You can renew a yearly/monthly instance manually or automatically.

**Table 2-10** Renewing a yearly/monthly instance

| Method | Description |
| --- | --- |
| **Manually Renewing an Instance** | You can renew a yearly/monthly instance anytime on the console before it is automatically deleted. |
| **Auto-Renewal** | You can enable auto-renewal to automatically renew the instance before it expires. This prevents resources from being deleted in case you forget to renew a subscription. |

You can renew a yearly/monthly instance in different phases of its lifecycle. For details, see **Figure 2-12**.

**Figure 2-12** Selecting a renewal method based on the instance's current phase




- An instance is in the **Provisioned** state after it is provisioned.
- When an instance subscription expires, the status will change from **Provisioned** to **Expired**.
- If an expired instance is not renewed, it enters a grace period. If it is not renewed by the time the grace period expires, the instance will be frozen and enter a retention period.
- If you do not renew the subscription before the retention period expires, your resources will be automatically deleted.

NOTE

- During the retention period, you cannot access or use your instance but the data stored in it can be retained. The retention period for Huawei Cloud International website is 15 days.
- During the grace period, you can access and use only some resources of your instance. The grace period for Huawei Cloud International website is 15 days.

You can enable auto-renewal any time before an instance expires. By default, the system will make the first attempt to charge your account for the renewal at 03:00, seven days before the expiry date. If this attempt fails, it will make another attempt at 03:00 every day until the subscription is renewed or expired. You can change the auto-payment date for renewal as required.

## 2.6.2 Manually Renewing an Instance

You can renew a yearly/monthly instance anytime on the console before it is automatically deleted.

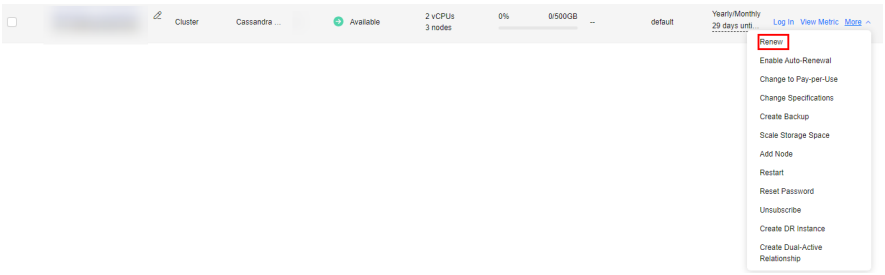
### Renewing an Instance on the Console

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, locate the target instance and choose **More** > **Renew** in the **Operation** column.

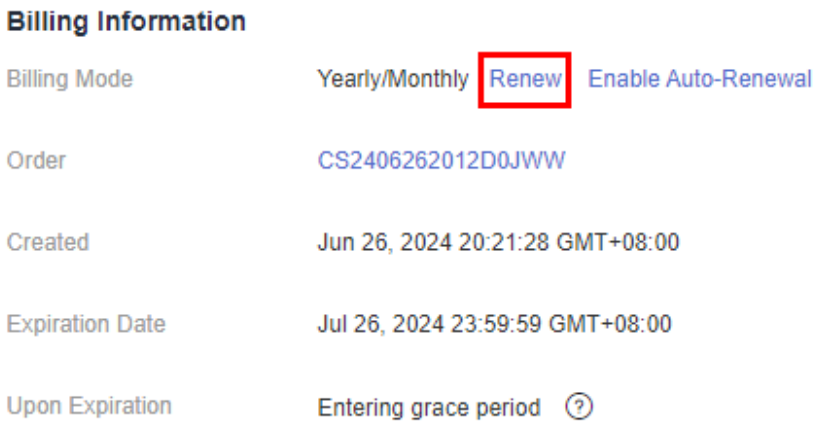
**Figure 2-13** Renewing an instance




Alternatively, click the instance name to go to the **Basic Information** page. In the **Billing Information** area, click **Renew** next to the **Billing Mode** field.

**Figure 2-14** Renewing an instance




**NOTE**

You can also renew multiple instances all at once:

1. Select the yearly/monthly instances to be renewed.
2. Click **Renew** above the instance list.

**Step 4** On the displayed page, renew the instances.

**----End**

## Renewing an Instance in Billing Center

**Step 1** **Log in to the Huawei Cloud console**.

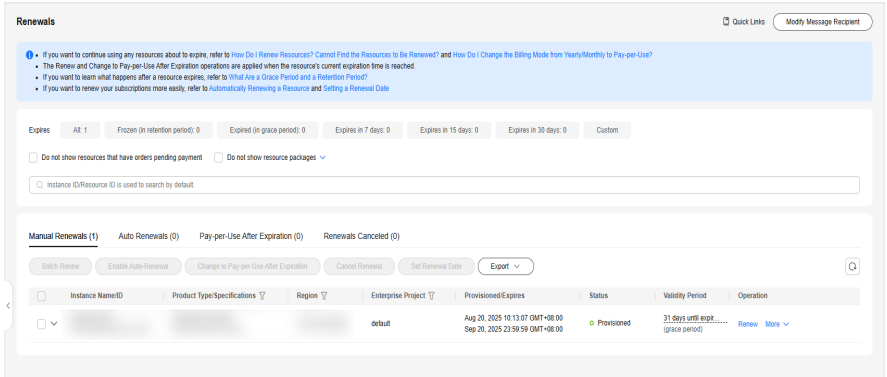
**Step 2** On the top menu bar, choose **Billing** > **Renewal**.

The **Renewals** page is displayed.

**Step 3** Select the search criteria.

On the **Manual Renewals**, **Auto Renewals**, **Pay-per-Use After Expiration**, and **Renewals Canceled** pages, you can view the instances to be renewed.

**Figure 2-15** Renewal management



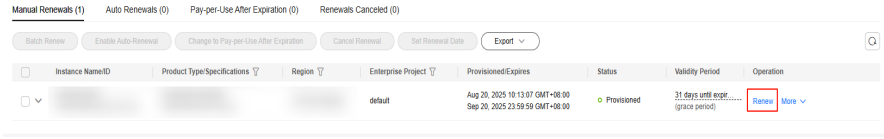

  

You can move all resources to be manually renewed to the **Manual Renewals** tab. For details, see **Restoring to Manual Renewal**.

**Step 4** Manually renew resources.

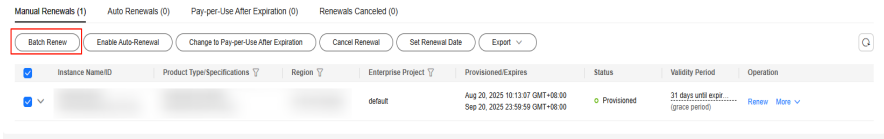
- Individual renewal: Locate an instance that you want to renew and click **Renew** in the **Operation** column.

  **Figure 2-16** Individual renewal

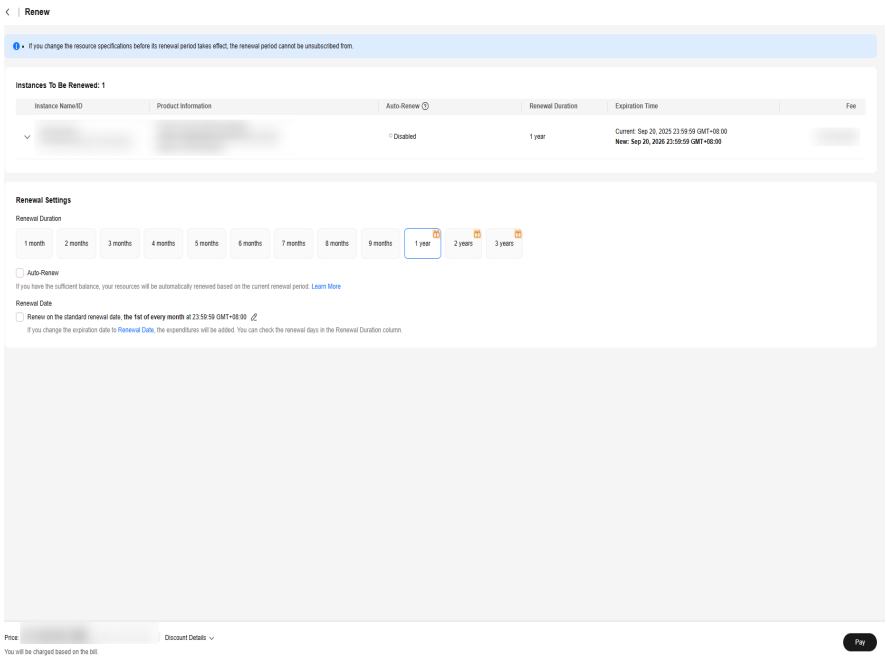
- Batch renewal: Select multiple instances that you want to renew and click **Batch Renew** in the upper left corner.

  **Figure 2-17** Batch renewal

**Step 5** Select a renewal duration and optionally select **Renew on the standard renewal date**. For details, see **Setting the Same Renewal Day for Yearly/Monthly Resources**. Confirm the price and click **Pay**.

**Figure 2-18** Confirming renewal



**Step 6** Select a payment method and make your payment. Once the order is paid for, the renewal is complete.

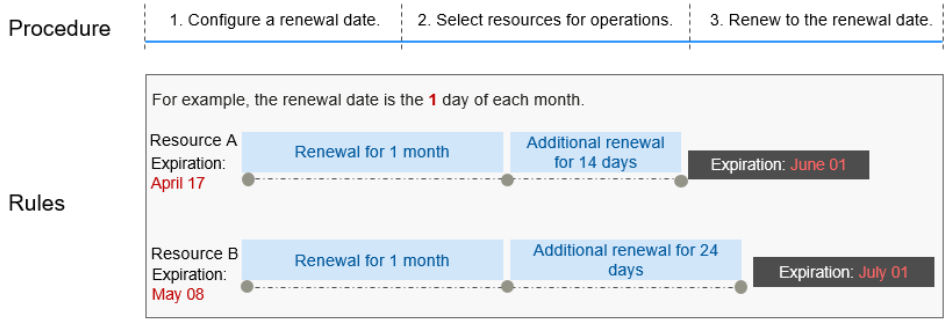
**----End**

### Setting the Same Renewal Day for Yearly/Monthly Resources

If the instances have different expiry dates, you can set the same renewal day, for example, the first day of each month, to make it easier to manage renewals.

In **Figure 2-19**, a user sets the same renewal day for two resources that will expire at different dates.

**Figure 2-19** Setting the same renewal day for resources with different expiry dates




For more details, see **Setting a Renewal Date**.

## 2.6.3 Auto-Renewal

Auto-renewal can prevent instances from being automatically deleted if you forget to manually renew them. The auto-renewal rules are as follows:

- The first auto-renewal date is based on when an instance expires and the billing cycle.
- The auto-renewal period is subject to the renewal duration you select.
  - Your monthly subscription will be renewed each month.
  - Your yearly subscription will be renewed each year.
- You can enable auto-renewal anytime before an instance expires. By default, the system will make the first attempt to renew your account at 03:00 seven days before the expiry date. If this attempt fails, it will make another attempt at 03:00 every day until the subscription is renewed or expired.
- After auto-renewal is enabled, you can still renew the instance manually if you want to. After a manual renewal is complete, auto-renewal is still valid, and the renewal fee will be deducted from your account seven days before the new expiry date.
- By default, the renewal fee is deducted from your account seven days before the new expiry date. You can change this auto-renewal payment date as required.

For more information about auto-renewal rules, see **Auto-Renewal Rules**.

## Prerequisites

Your yearly/monthly instance is not expired.

## Enabling Auto-Renewal When Buying an Instance

You can enable auto-renewal on the page for buying an instance. For details, see **Buying an Instance**.

**Figure 2-20** Enabling auto-renewal


## Enabling Auto-Renewal on the Renewals Page

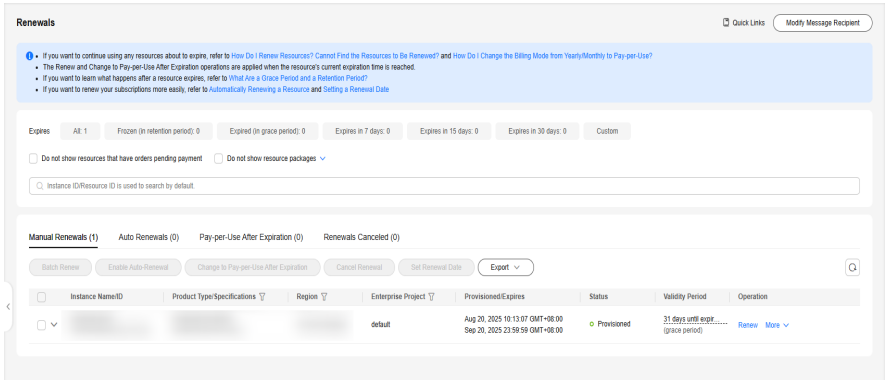
**Step 1** **Log in to the Huawei Cloud console**.

**Step 2** On the top navigation bar, choose **Billing** > **Renewal**.

**Step 3** Select the search criteria.

- On the **Auto Renewals** page, you can view the resources that auto-renewal has been enabled for.
- You can enable auto-renewal for resources on the **Manual Renewals**, **Pay-per-Use After Expiration**, and **Renewals Canceled** pages.

**Figure 2-21** Renewal management

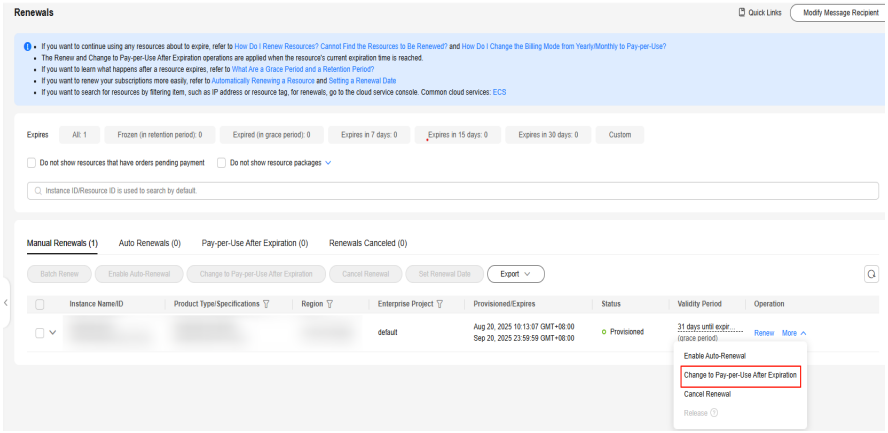



**Step 4** Enable auto-renewal for yearly/monthly resources.

- Enabling auto-renewal for a single instance: Locate the instance that you want to enable auto-renewal for and choose **More** > **Enable Auto-Renew** in the **Operation** column.

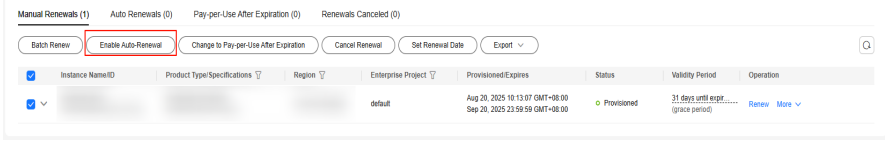
  

**Figure 2-22** Enabling auto-renewal for an instance



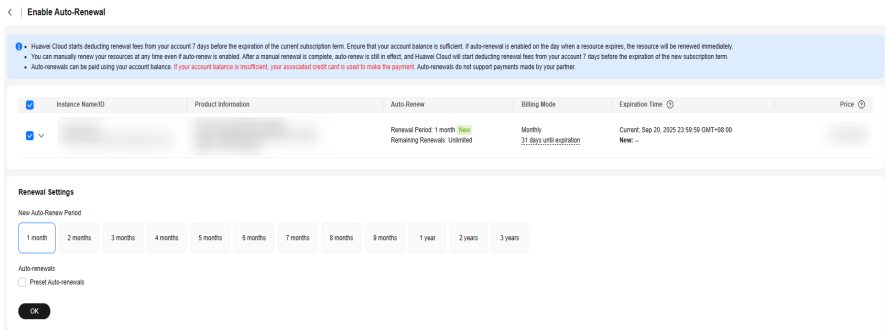

- Enabling auto-renewal for multiple instances at a time: Select the instances that you want to enable auto-renewal for and click **Enable Auto-Renew** above the list.

**Figure 2-23** Enabling auto-renewal for multiple instances




**Step 5** Select a renewal period, specify the auto-renewal times, and click **Pay**.

**Figure 2-24** Enabling auto-renewal




**----End**

# 2.7 Bills

You can view the resource usage and bills for different billing cycles on the **Bills** page in the Billing Center.

## Bill Generation

A bill is generated after a yearly/monthly instance is paid.

The usage of pay-per-use instances is reported to the billing system at a fixed interval. Pay-per-use instances can be settled by the hour, day, or month based on the usage type. For details, see **Bill Run for Pay-per-Use Resources**. Pay-per-use GeminiDB DynamoDB-Compatible instances are settled by the hour.

The fee deduction time of a pay-per-use instance may be later than the settlement period. For example, if an instance is deleted at 08:30, the fees generated from 08:00 to 09:00 are usually deducted at about 10:00. In Billing Center, choose **Billing** > **Transactions and Detailed Bills** > **Transaction Bills**. **Expenditure Time** in the bill indicates the time when the pay-per-use resource is used.

## Viewing Bills of a Specific Resource

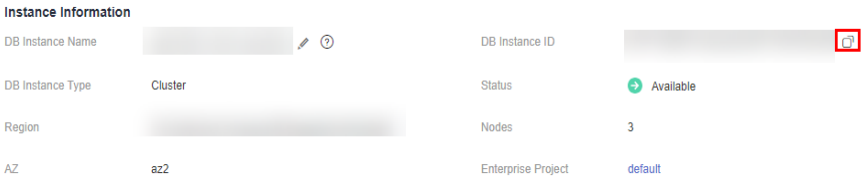
[Method 1: Use the instance ID to search for a bill.]

**Step 1** **Log in to the Huawei Cloud console**. Choose **Databases** > **GeminiDB**.

**Step 2** On the **Instances** page, locate the instance whose bill you want to view and click its name.

**Step 3** Click the icon shown in the figure below to copy the instance ID.

**Figure 2-25** Copying the instance ID



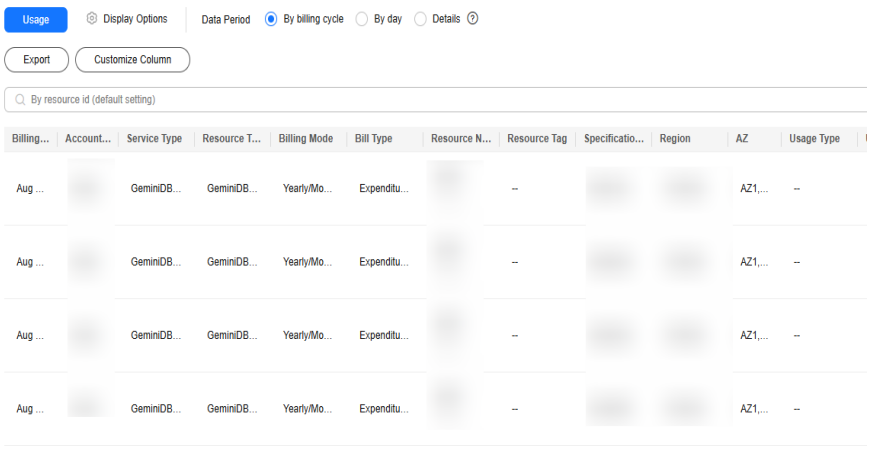

**Step 4** On the top menu bar, choose **Billing** > **Bills**.

The **Bills** page is displayed.

**Step 5** In the navigation pane, choose **Billing** > **Expenditure Details**. Select **Resource ID** as the filter, enter the resource ID, and click to search for the resource bills.

**Figure 2-26** Querying resource bills




By default, the bill details are displayed by usage and billing cycle. You can choose other display options. For details, see **Bills**.

**----End**

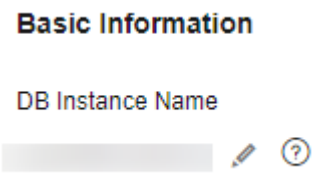
[Method 2: Use the resource name to search for a bill.]

**Step 1** **Log in to the Huawei Cloud console**. Choose **Databases** > **GeminiDB**.

**Step 2** On the **Instances** page, locate the instance whose bill you want to view and click its name.

**Step 3** On the **Basic Information** > **Instance Information** page, obtain the instance name.

**Figure 2-27** Obtaining an instance name



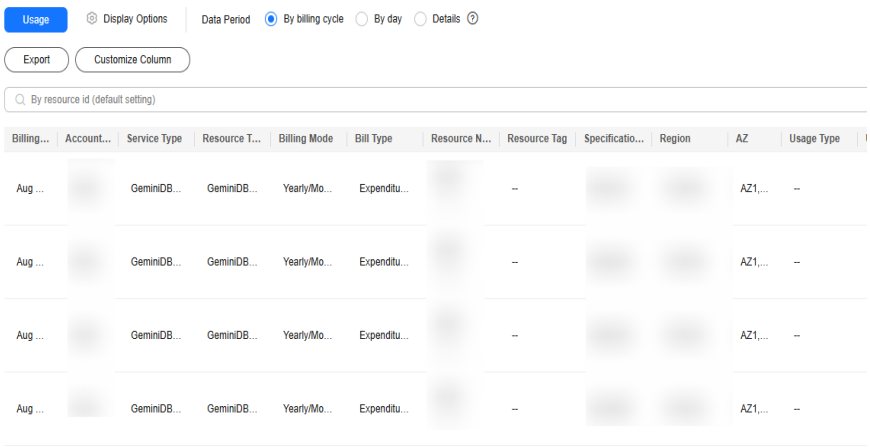

**Step 4** On the top menu bar, choose **Billing** > **Bills**.

The **Bills** page is displayed.

**Step 5** In the navigation pane, choose **Billing** > **Expenditure Details**. Select **Resource Name** as the filter, enter the resource name, and click to search for the resource bills.

**Figure 2-28** Querying resource bills




By default, the bill details are displayed by usage and billing cycle. You can choose other display options. For details, see **Bills**.

**----End**

## Scenario Example: Checking the Consistency of the Actual Usage and Billed Usage

Assume that you bought a pay-per-use GeminiDB DynamoDB-Compatible instance at 10:09:06 on April 8, 2023, and deleted it at 12:09:06 on April 8, 2023.

- GeminiDB DynamoDB-Compatible instance transaction records

  Pricing is listed on a per-hour basis, and bills are calculated down to the second. You can check the transaction records against the actual usage. Resources are billed separately in the transaction records.

  **Table 2-11** GeminiDB DynamoDB-Compatible instance transaction records

| | |
|---|---|
| **Product Type** | GeminiDB DynamoDB-Compatible API |
| **Resource Type** | GeminiDB DynamoDB-Compatible instance storage |
| **Billing Mode** | Pay-per-use |
| **Expenditure Time** | For the period of time from 10:09:06 to 12:09:06 on April 08, 2023, 6 transaction records would be generated for the resource usage in the following periods:<br>• 10:09:06 – 11:00:00<br>• 11:00:00 – 12:00:00<br>• 12:00:00 – 12:09:06 |

| | |
|---|---|
| **List Price** | List price on the official website = Usage x Unit price x Capacity<br>The instance was used for 3,054 seconds in the first period. You can check its unit price in **GeminiDB Price Calculator**. The list price in the first period is USD0.02375333 = (3054/3600) × 0.0007 × 40. You can also calculate the list price in the other periods. |
| **Discounted Amount** | You can enjoy discounts on cloud services, such as business, partner-authorized, and promotional discounts. The discounts are calculated based on the list price. |
| **Truncated Amount** | The billing of Huawei Cloud is calculated to the 8th decimal place (USD). However, the amount due is truncated to the 2nd decimal place. The third and later decimal places are referred to as the truncated amounts.<br>For example, in the first billing cycle, the truncated amount is USD0.00375333. |
| **Amount Due** | Amount due = List price – Discount amount – Truncated amount<br>Take the first period as an example. If the discount amount is 0, the amount due is $0.02 USD (0.02375333 – 0 – 0.00375333). |

- GeminiDB DynamoDB-Compatible instance bill details

  Bill details can be displayed in multiple ways. By default, the bill details are displayed by usage and billing cycle. **Table 2-12** lists the bill details, which can be used to check the actual usage.

  **Table 2-12** GeminiDB DynamoDB-Compatible instance bill details

| | |
|---|---|
| **Product Type** | GeminiDB DynamoDB-Compatible API |
| **Resource Type** | GeminiDB DynamoDB-Compatible instance storage |
| **Billing Mode** | Pay-per-use |
| **Resource Name/ID** | Name and ID of a GeminiDB DynamoDB-Compatible instance<br>Example: nosql-b388 and 21e8811a64bf4de88bc2e2556da17983in12 |
| **Specifications** | GeminiDB DynamoDB-Compatible instance storage |
| **Usage Type** | Duration |

| | |
|---|---|
| **Unit Price** | When pay-per-use billing is used, the unit price is only provided if the amount is equal to the usage multiplied by the unit price. No unit price is provided in other pricing modes, for example, tiered pricing.<br>You can check the unit price of a pay-per-use instance in **GeminiDB Price Calculator**. |
| **Unit** | USD/GB/Hour in **GeminiDB Price Calculator** |
| **Usage** | Depends on the unit of the unit price, which is USD/GB/hour. Storage usage is priced by the hour. In this example, the total duration is 2 hours. |
| **Usage Unit** | Hour |
| **List Price** | List price on the official website = Usage x Unit price x Capacity<br>The instance has been used for 2 hours. Its unit price is displayed in **GeminiDB Price Calculator**. The list price is USD0.056 (2 × 0.0007 × 40). |
| **Discounted Amount** | You can enjoy discounts on cloud services, such as business, partner-authorized, and promotional discounts. The discounts are calculated based on the list price. |
| **Amount Due** | Amount that should be paid for used cloud services after discounts are applied. |

# 2.8 Arrears

If the available account balance is less than the amount to be settled, the account will be in arrears. To continue using your instances, you need to top up your account in a timely manner.

## Arrears Reason

If you do not have yearly/monthly instances, your account falls into arrears any time your configured payment method is unable to pay for the used resources on the pay-per-use basis.

## Impact of Arrears

- Yearly/Monthly

  You have prepaid for yearly/monthly GeminiDB DynamoDB-Compatible instances. If your account is in arrears, the existing yearly/monthly instances can still be used. However, you cannot perform operations that involve fees, such as buying another GeminiDB DynamoDB-Compatible instance, upgrading specifications, and renewing the existing instances.

- Pay-per-use

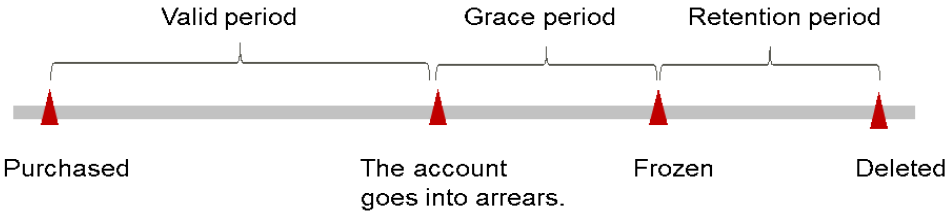
  

When your account falls into arrears due to automatic fee deduction for GeminiDB DynamoDB-Compatible instances, your account will be marked as in arrears, and the pay-per-use instances enter the grace period. After you top up your account, Huawei Cloud will bill you for expenditures generated by the resources during the grace period. You can view the expenditures on the **Overview** page of the Billing Center.

If your account is still in arrears after the grace period ends, the resources enter the retention period and their status turns to **Frozen**. You cannot perform any operations on these resources.

After the retention period ends, the compute resources (vCPUs and memory) and EIPs will be released and cannot be restored.

**Figure 2-29** Lifecycle of a pay-per-use instance




NOTE

The grace period and retention period are both 15 days.

### Avoiding and Handling Arrears

Top up your account in time. For details, see **Topping Up an Account**.

If you no longer need the instances, delete them to avoid additional fees.

To help make sure your account never falls into arrears, you can configure the **Balance Alert** on the **Overview** page of the Billing Center. Then, anytime an expenditure quota drops to below the threshold you specify, Huawei Cloud automatically notifies you by SMS or email.

## 2.9 Billing Termination

### Yearly/Monthly Resources

You are billed upfront for yearly/monthly resources, such as yearly/monthly GeminiDB DynamoDB-Compatible instances. Your access to these instances will automatically cease upon expiration.

- You can unsubscribe from a yearly/monthly resource before it expires. Depending on whether coupons or discounts were used, Huawei Cloud may issue you a refund. For details about unsubscription rules, see **Unsubscriptions**.
- If you have enabled auto-renewal but no longer wish to automatically renew the subscription, disable it before the auto-renewal date (7 days before the expiration date by default) to avoid unexpected expenditures.

## Pay-per-Use Resources

If you no longer need the pay-per-use instances and want to stop being billed for them, simply delete them.

## Searching for Resources from Bills and Stopping Billing

To ensure that all related resources are deleted, you can search the billing records by resource ID, and then delete the resources you identify in this way. The procedure is as follows:

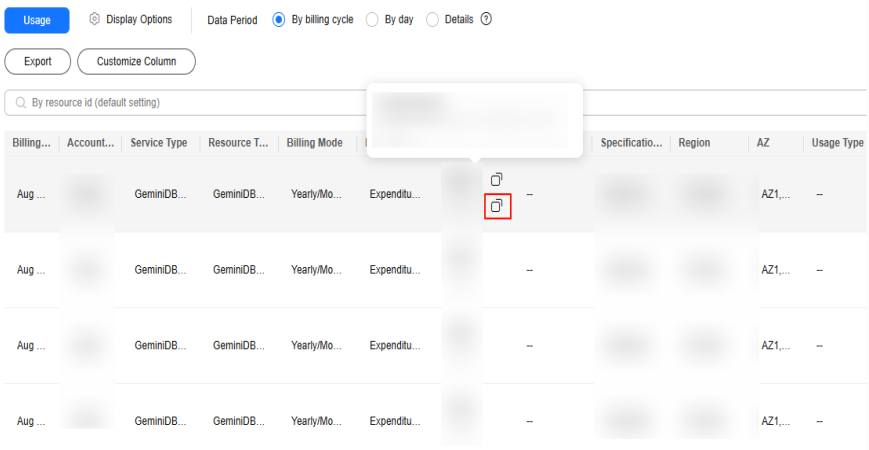
[Method 1: Use the resource ID in the bill to search for the resource.]

**Step 1** Log in to the Huawei Cloud console. On the top menu bar, choose **Billing** > **Bills**.

The **Bills** page is displayed.

**Step 2** In the navigation pane, choose **Billing** > **Expenditure Details**. Click the icon shown in the following figure to copy the resource ID.

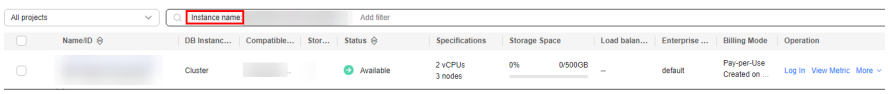
**Figure 2-30** Copying the resource ID




**Step 3** Log in to the Huawei Cloud console. Choose **Databases** > **GeminiDB**.

**Step 4** Select the region where the resource is located. Select **Instance ID**, enter the resource ID copied from **Step 2**, and click to search for the resource.

**Figure 2-31** Searching for resources




**Step 5** Locate the instance you want to delete and click **More** > **Delete** in the **Operation** column. Ensure that the resource is not found in the list.

NOTE

You are billed one hour after the resource usage is calculated, so a bill may still be generated after the pay-per-use resource is deleted. For example, if you delete an instance (which is billed on an hourly basis) at 08:30, the expenditures for that hour from 08:00 to 09:00 are usually not billed until about 10:00.

**----End**

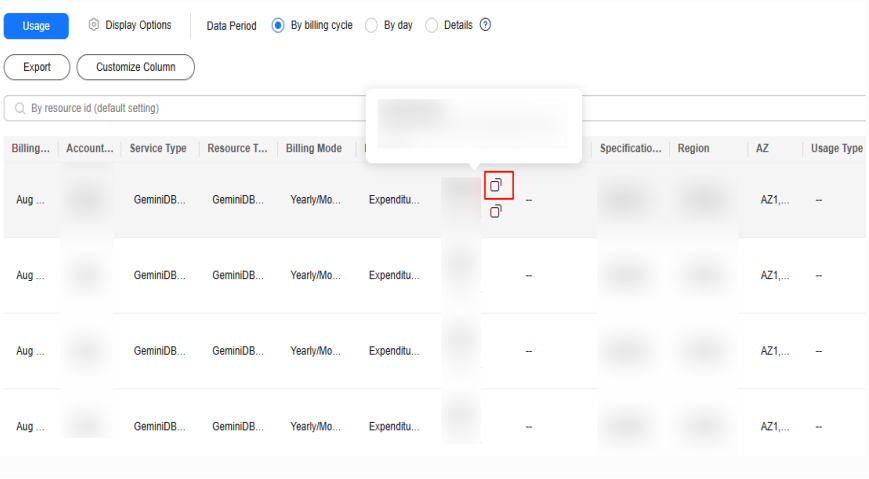
[Method 2: Use the resource name in the bill to search for the resource.]

**Step 1** **Log in to the Huawei Cloud console**. On the top menu bar, choose **Billing** > **Bills**.

The **Bills** page is displayed.

**Step 2** In the navigation pane, choose **Billing** > **Expenditure Details**. Click the icon shown in the following figure to copy the resource name.

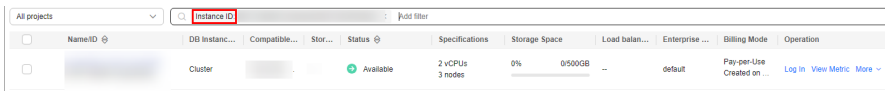
**Figure 2-32** Copying the resource name




**Step 3** **Log in to the Huawei Cloud console**. Choose **Databases** > **GeminiDB**.

**Step 4** Enter the instance name copied from **Step 2** in the search box and click .

**Figure 2-33** Searching for resources




**Step 5** Locate the instance you want to delete and click **More** > **Delete** in the **Operation** column. Ensure that the resource is not found in the list.

NOTE

You are billed one hour after the resource usage is calculated, so a bill may still be generated after the pay-per-use resource is deleted. For example, if you delete an instance (which is billed on an hourly basis) at 08:30, the expenditures for that hour from 08:00 to 09:00 are usually not billed until about 10:00.

**----End**

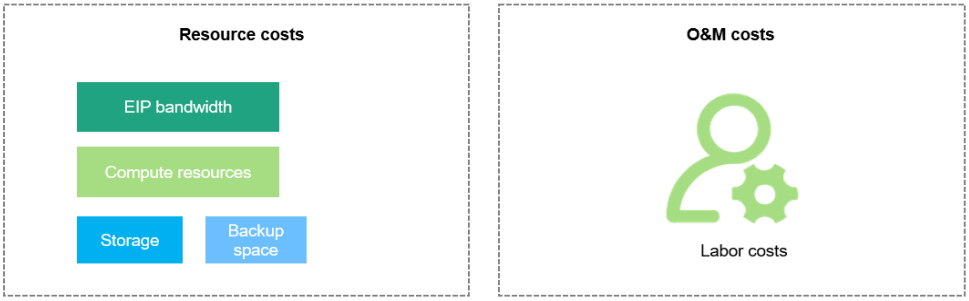
  

# 2.10 Cost Management

## 2.10.1 Costs

The costs of a GeminiDB DynamoDB-Compatible instance include:

- Resource costs: costs of compute and storage resources. For details, see **Billing Modes**.
- O&M costs: labor costs generated when a GeminiDB DynamoDB-Compatible instance is used




## 2.10.2 Cost Allocation

A good cost accountability system is a prerequisite for cost management. It ensures that departments, business teams, and owners are accountable for their respective cloud costs. An enterprise can allocate cloud costs to different teams or projects so as to have a clear picture of their respective costs.

Huawei Cloud **Cost Center** provides various tools for you to group costs in different ways. You can experiment with these tools and find a way that works best for you.

- **By linked account**

  The enterprise master account can manage costs by grouping the costs of its member accounts by linked account. For details, see **Viewing Costs by Linked Account**.

- **By enterprise project**

  Before allocating costs, enable Enterprise Project Management Service (EPS) and plan your enterprise projects based on your organizational structure or service needs. When purchasing cloud resources, select an enterprise project so that the costs of resources will be allocated to the selected enterprise project. For details, see **Viewing Costs by Enterprise Project**.

  **Figure 2-34** Selecting an enterprise project

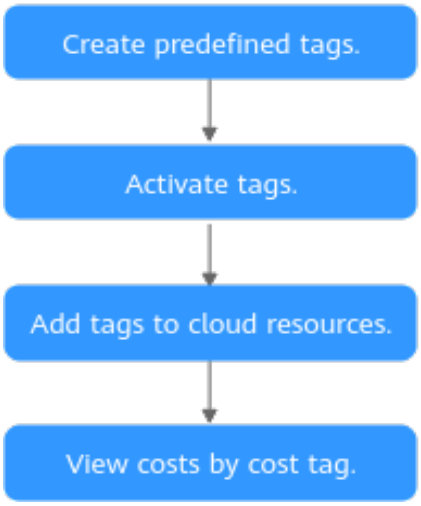
  
  

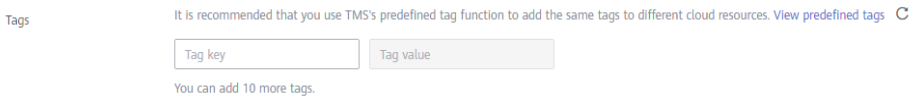
- **By cost tag**

You use tags to sort your Huawei Cloud resources in a variety of different ways, for example, by purpose, owner, or environment. The following is the process of managing costs by predefined tags (recommended).

**Figure 2-35** Adding a tag




For details, see **Viewing Costs by Cost Tag**.

- **By cost category**

  You can use cost categories provided by **Cost Center** to split shared costs. Shared costs are the costs of resources (compute, network, storage, or resource packages) shared across multiple departments or the costs that cannot be directly split by cost tag or enterprise project. These costs are not directly attributable to a singular owner, and they cannot be categorized into a singular cost type. In this case, you can define cost splitting rules to fairly allocate these costs among teams or business units. For more information, see **Allocating Costs By Cost Category**.

## 2.10.3 Cost Analysis

To precisely control and optimize your costs, you need a clear understanding of what parts of your enterprise incurred different costs. **Cost Center** visualizes your original costs and amortized costs using various dimensions and display filters for cost analysis so that you can analyze the trends and drivers of your service usage and costs from a variety of perspectives or within different defined scopes.

You can also use cost anomaly detection provided by **Cost Center** to detect unexpected expenses in a timely manner. In this way, costs can be monitored, analyzed, and traced.

For details, see **Performing Cost Analysis to Explore Costs and Usage** and **Enabling Cost Anomaly Detection to Identify Anomalies**.

## 2.10.4 Cost Optimization

You can identify resources with high costs based on the analysis results in the cost center, determine the causes of high costs, and take optimization measures accordingly.

### Resource optimization

- On Cloud Eye, view metrics of your GeminiDB DynamoDB-Compatible instance, such as CPUs, memory, and disk usage. If the current configurations are too high, you can use lower specifications instead.
- Monitor idle GeminiDB DynamoDB-Compatible instances and delete them in a timely manner.

### Billing mode selection

Different types of services have different requirements on resource usage periods, so the most economical billing mode for one resource may not be the best option for another resource.

- For mature services that tend to be stable for the long term, select yearly/monthly billing.
- For short-term, unpredictable services that experience traffic bursts and cannot afford to be interrupted, select pay-per-use billing.
- Monitor the lifecycle of instances and renew yearly/monthly resources that are about to expire in a timely manner.

# 2.11 Billing FAQs

## 2.11.1 What Are the Differences Between Yearly/Monthly and Pay-per-Use Billing?

Yearly/Monthly is a prepaid billing mode in which resources are billed based on the service duration. This cost-effective mode is ideal when the duration of resource usage is predictable. It is recommended for long-term users.

Pay-per-use is a postpaid mode. You are only billed for how long you have actually used your instance. This mode can be a good option when future requirements are unpredictable. Pay-per-use instances are priced by the hour, but if an instance is used for less than one hour, you will be billed based on the actual duration.

## 2.11.2 Can I Switch Between Yearly/Monthly and Pay-per-use Billing Modes?

Yes.

- For details about how to change a yearly/monthly GeminiDB DynamoDB-Compatible instance to pay-per-use, see **Changing a Yearly/Monthly Instance to Pay-per-Use**.

- For details about how to change a pay-per-use GeminiDB DynamoDB-Compatible instance to yearly/monthly, see **Changing a Pay-per-Use Instance to Yearly/Monthly**.

# 2.11.3 How Do I Renew a Single or Multiple Yearly/Monthly Instances?

You can renew your yearly/monthly GeminiDB DynamoDB-Compatible instance.

## Usage Notes

Pay-per-use instances do not support this function.

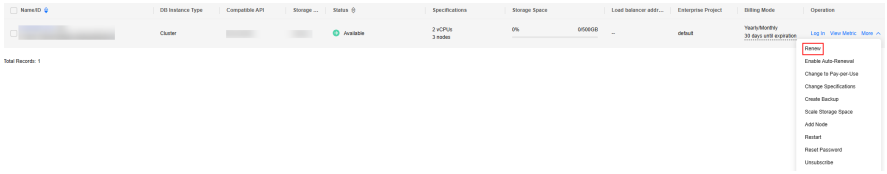
## Renewing a Single Yearly/Monthly Instance

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

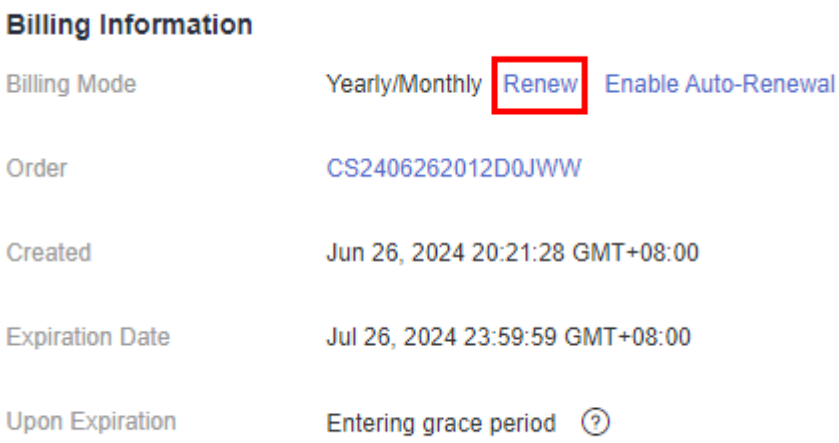
**Step 3** On the **Instances** page, locate the target instance and choose **More** > **Renew** in the **Operation** column.

**Figure 2-36** Renewal



Alternatively, click the instance name to go to the **Basic Information** page. In the **Billing Information** area, click **Renew** next to the **Billing Mode** field.

**Figure 2-37** Renewal




**Step 4** On the displayed page, renew the instance.

**----End**

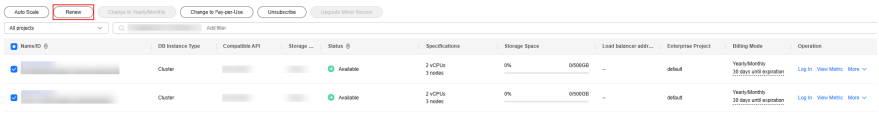
## Renewing Instances in Batches

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, select the instances that you want to renew and click **Renew** above the instance list.

**Figure 2-38** Renewing instances in batches



**Step 4** In the displayed dialog box, click **Yes**.

**----End**

# 2.11.4 How Do I Unsubscribe from a Yearly/Monthly Instance?

If you do not need a yearly/monthly instance any longer, unsubscribe from it.

## Usage Notes

- The unsubscription action cannot be undone. To retain data, create a manual backup before unsubscription. For details, see **Creating a Manual Backup**.
- After an unsubscription request is submitted, resources and data will be deleted and cannot be retrieved. To retain data, back it up before submitting the unsubscription request.

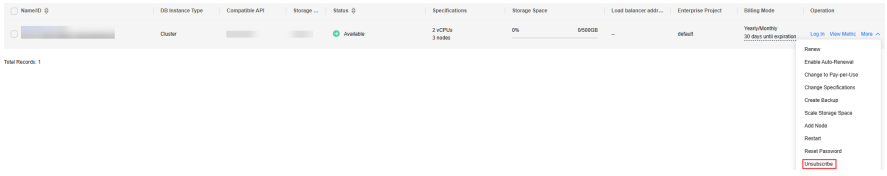
## Unsubscribing from a Single Yearly/Monthly Instance

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, locate the target instance and choose **More** > **Unsubscribe** in the **Operation** column.

**Figure 2-39** Unsubscribe



**Step 4** In the displayed dialog box, click **Yes**.

**Step 5** On the displayed page, confirm the order to be unsubscribed and select a reason. Then, click **Confirm**.

For details, see **Unsubscription Rules**.

**Step 6** In the displayed dialog box, click **Yes**.

NOTE

1. After an unsubscription request is submitted, resources and data will be deleted and cannot be retrieved.
2. To retain data, back it up before submitting the unsubscription request.

**Step 7** View the unsubscription result. After you unsubscribe from the instance order, the instance is no longer displayed in the instance list on the **Instances** page.

**----End**

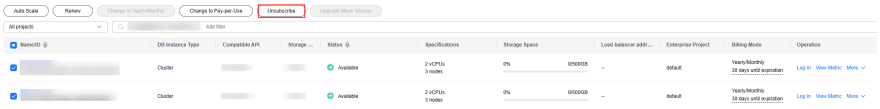
## Batch Unsubscribing from Yearly/Monthly Instances

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** Choose **Instances** in the navigation pane on the left, select the instances you want to unsubscribe from and click **Unsubscribe** above the instance list.

**Figure 2-40** Unsubscribe




**Step 4** In the displayed dialog box, click **Yes**.

**Step 5** On the displayed page, confirm the order to be unsubscribed and select a reason. Then, click **Confirm**.

For details, see **Unsubscription Rules**.

**Step 6** In the displayed dialog box, click **Yes**.

NOTE

1. After an unsubscription request is submitted, resources and data will be deleted and cannot be retrieved.
2. To retain data, back it up before submitting the unsubscription request.

**Step 7** View the unsubscription result. After you unsubscribe from the instance order, the instance is no longer displayed in the instance list on the **Instances** page.

**----End**

# 3 Getting Started with GeminiDB DynamoDB-Compatible API

## 3.1 Getting to Know GeminiDB DynamoDB-Compatible API

This section instructs you to create and connect to a GeminiDB DynamoDB-Compatible instance.

**Connection Methods**

You can use Java and Python to connect to a GeminiDB DynamoDB-Compatible instance. A load balancer address (port: **80**) is recommended.

**Table 3-1** Connection methods

| Method | Scenario | Description |
| --- | --- | --- |
| Connecting to a GeminiDB DynamoDB-Compatible Instance Using Java | An example of connecting to a GeminiDB DynamoDB-Compatible instance using Java | - |

| Method | Scenario | Description |
| --- | --- | --- |
| Connecting to a GeminiDB DynamoDB-Compatible Instance Using Python | An example of connecting to a GeminiDB DynamoDB-Compatible instance using Python | - |

**More Connection Operations**

- See **Connection Methods**.

# 3.2 Buying and Connecting to a GeminiDB DynamoDB-Compatible Instance

This section instructs you to buy a GeminiDB DynamoDB-Compatible instance on the GeminiDB console.

By default, each tenant can create a maximum of 50 GeminiDB DynamoDB-Compatible instances. To request a higher quota, choose **Service Tickets > Create Service Ticket** in the upper right corner of the console and contact customer service personnel.

- **Buying an Instance**
- **Connecting to an Instance Using Java**

  For details about other connection methods, see **Connection Methods**.

## Prerequisites

- You have created a Huawei Cloud account.

## Buying an Instance

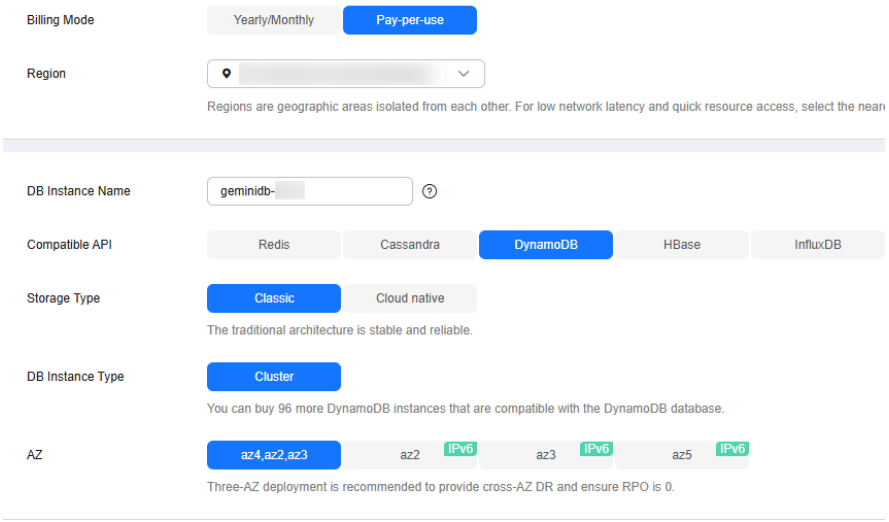
For details, see **Buying a GeminiDB DynamoDB-Compatible Instance**.

1. **Log in to the Huawei Cloud console**.
2. In the service list, choose **Databases** > **GeminiDB**.
3. On the **Instances** page, click **Buy DB Instance**.
4. On the displayed page, select a billing mode, configure instance parameters, and click **Next**.

   The following parameters are for reference only. Select proper specifications as needed. **Table 4-1** lists details about the parameters.

**Figure 3-1** Billing mode and basic information

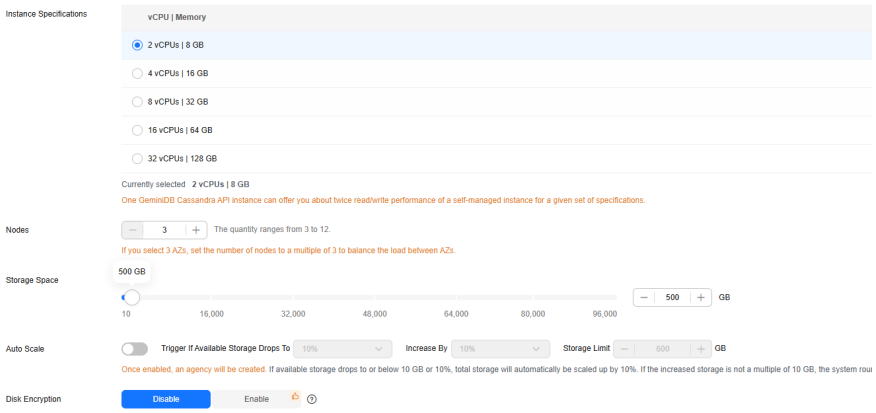



| Parameter | Example Value | Description |
| --- | --- | --- |
| Billing Mode | Pay-per-use | Billing mode of an instance<br>• **Yearly/Monthly**: A prepaid billing mode in which you pay for resources before using it. Bills are settled based on the subscription period. The longer the subscription term, the bigger the discount. This mode is a good option for long-term stable services.<br>• **Pay-per-use** is a postpaid mode. You are billed based on how long you have actually used GeminiDB. Pricing is listed on a per-hour basis, and bills are calculated down to the second. This mode allows you to adjust resource usage easily. You neither need to prepare for resources in advance, nor end up with excessive or insufficient preset resources. |
| Region | Select **CN-Hong Kong**. | Region where a tenant is located<br>**NOTE**<br>To reduce network latency, select a region nearest from which you will access the instance. Instances deployed in different regions cannot communicate with each other over a private network. After you buy an instance, you cannot change its region. |

| Parameter | Example Value | Description |
| --- | --- | --- |
| DB Instance Name | User-defined | The instance name:<br>• Can be the same as an existing instance name.<br>• Can contain 4 to 64 characters and must start with a letter. It is case-sensitive and allows only letters, digits, hyphens (-), and underscores (_). |
| Compatible API | DynamoDB | GeminiDB is compatible with mainstream NoSQL databases, including Redis, DynamoDB, Cassandra, HBase, MongoDB, and InfluxDB. You can select GeminiDB APIs by following **How Do I Select an API?** |
| Storage Type | Classic | • **Classic**: classic architecture with decoupled storage and compute<br>• **Cloud native**: new, more flexible, new-gen version with support for more AZs<br>**NOTE**<br>Cloud native storage is now in OBT. To use it, choose **Service Tickets > Create Service Ticket** in the upper right corner of the console and contact the customer service. |
| DB Instance Type | Cluster | **Cluster**<br>One cluster consists of at least three nodes. A cluster is easy to scale out to meet increasing data growth needs. A cluster is recommended when dealing with stringent availability demands, substantial data volumes, and the need for seamless scalability. |
| AZ | AZ 1, AZ 2, and AZ 3 | Availability zone where the instance is created. An AZ is a part of a region with its own independent power supplies and networks. AZs are physically isolated but can communicate with each other over a private network.<br>Instances can be deployed in a single AZ or three AZs.<br>• To deploy instances in a single AZ, select one AZ.<br>• To deploy instances across AZs for disaster recovery, select three AZs, where the instance nodes will be evenly distributed. |

**Figure 3-2** Specifications and storage




| Parameter | Example Value | Description |
| --- | --- | --- |
| Instance Specifications | 2U8GB | Select appropriate specifications based on the CPU-memory ratio.<br>After an instance is created, you can change its specifications. For details, see **Changing vCPUs and Memory**. |
| Nodes | 3 | Number of nodes that the instance is deployed on.<br>Currently, a maximum of 60 nodes are supported. To add more, choose **Service Tickets > Create Service Ticket** in the upper right corner of the console and contact the customer service.<br>After an instance is created, you can add nodes. For details, see **Manually Adding Instance Nodes**. |
| Storage Space | 500 GB | Instance storage space. The range depends on the instance specifications. For details, see **Instance Specifications**.<br>After an instance is created, you can scale up its storage if necessary. For details, see **Manually Scaling Up Storage Space**. |
| Autoscaling | Toggled off | **Autoscaling** is toggled off by default.<br>You can enable **Auto Scale** after an instance is created. For details, see **Automatically Scaling Up Storage Space**. |

| Parameter | Example Value | Description |
| --- | --- | --- |
| Disk Encryption | Disable | **Disable** is selected by default.<br>If you select **Enable**, your data will be encrypted on disks and stored in ciphertext after you create an instance. When you download encrypted objects, the ciphertext will be decrypted into plain text and then sent to you. Disk encryption can improve data security and may have slight impacts on database writes and reads. |

**Figure 3-3** Network configuration

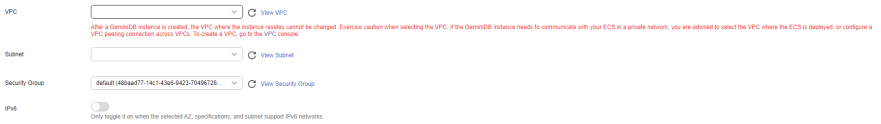
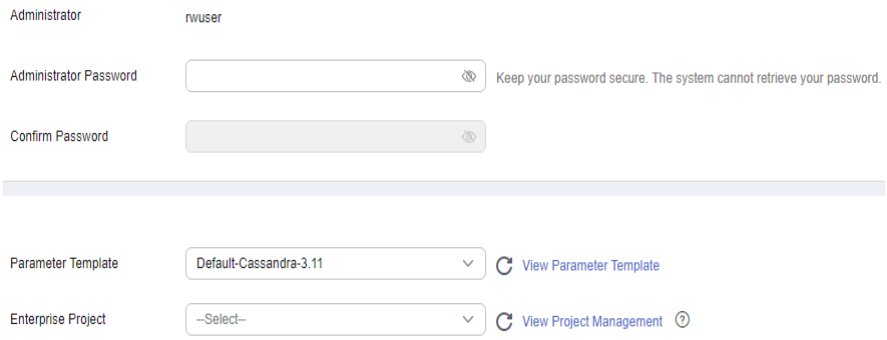



**Figure 3-4** Database configuration




| Parameter | Example Value | Description |
| --- | --- | --- |
| VPC | default_vpc | Virtual private network where your instances are located. A VPC isolates networks for different services. You can select an existing VPC or create a VPC.<br>**NOTE**<br>• After a GeminiDB DynamoDB-Compatible instance is created, the VPC where the instance is deployed cannot be changed.<br>• To connect a GeminiDB DynamoDB-Compatible instance to an ECS over a private network, ensure they are in the same VPC. If they are not, create a **VPC peering connection** between them. |

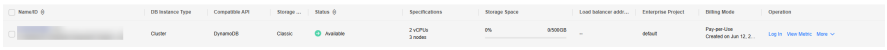
  

| Parameter | Example Value | Description |
| --- | --- | --- |
| Subnet | default_subnet | A subnet provides dedicated network resources that are logically isolated from other networks for security purposes. |
| Security Group | default | A security group controls access between GeminiDB DynamoDB-Compatible API and other services. Ensure that the security group you selected allows your client to access the instance. |
| Administrator Password | Configured based on the password policy | Password of the administrator account. The password:<br>• Can include 8 to 32 characters.<br>• Can include uppercase letters, lowercase letters, digits, and any of the following special characters: ~!@#%^*-_=+?<br>• For security reasons, set a strong password. The system will verify the password strength.<br>Keep your password secure. The system cannot retrieve it if it is lost. |
| Enterprise Project | default | This parameter is provided for enterprise users.<br>An enterprise project groups cloud resources, so you can manage resources and members by project. The default project is **default**.<br>Select an enterprise project from the drop-down list. For more information about enterprise projects, see ***Enterprise Management User Guide***. |

Retain the default values for other parameters.

5. On the displayed page, confirm instance details. To modify the configurations, click **Previous**.
6. If no modification is required, read and agree to the service agreement and click **Submit**.
7. Click **Back to Instance Management** to go to the instance list.
8. On the **Instances** page, view and manage the created instance.
   - It takes about 5 to 9 minutes to create an instance. During the process, the instance status is **Creating**.
   - After the instance is created, its status becomes **Available**.

**Figure 3-5** Available instance



## Connecting to an Instance Using Java

This section describes how to connect to a GeminiDB DynamoDB-Compatible instance using Java.

Prerequisites

- A GeminiDB DynamoDB-Compatible instance has been created.
- For details about how to create an ECS, see **Purchasing an ECS** in *Getting Started with Elastic Cloud Server*.
- JDK has been installed on the ECS.

Obtaining the IP Address of a GeminiDB DynamoDB-Compatible Instance

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

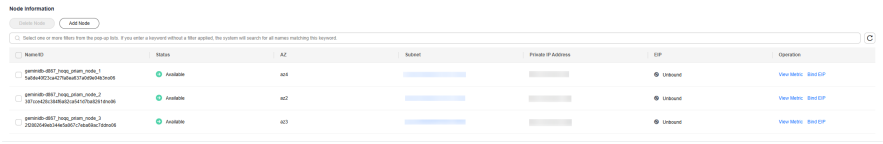
**Step 3** On the **Instances** page, click the target instance.

The port of the GeminiDB DynamoDB-Compatible instance is 8000.

**Method 1**:

In the **Node Information** area on the **Basic Information** page, view the private IP address or EIP of each node.

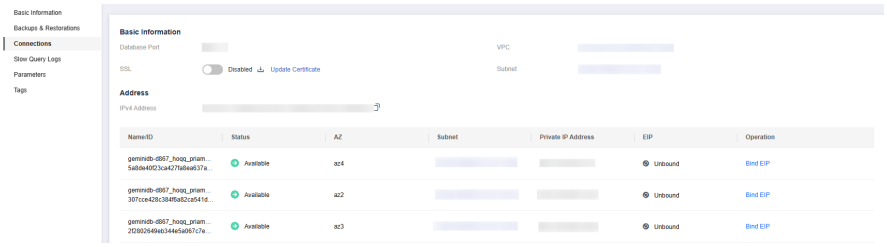
**Figure 3-6** Viewing the IP address



**Method 2**

In the navigation pane, choose **Connections** to view the private IP address and EIP of the instance.

**Figure 3-7** Viewing IP addresses



**----End**

Using a Load Balancing Plug-in to Connect to a GeminiDB DynamoDB-Compatible Instance

**Step 1** To obtain the JAR package and plug-in code, choose **Service Tickets > Create Service Ticket** in the upper right corner of the console and contact the customer service.

**Step 2** Replace the IP address in the following code example with the IP address queried in **Step 3**.

Java code example:

```
package com.huawei.dbs.test;

import com.huawei.dbs.RequestHandler;

import com.amazonaws.SDKGlobalConfiguration;
import com.amazonaws.auth.AWSCredentialsProvider;
import com.amazonaws.auth.AWSStaticCredentialsProvider;
import com.amazonaws.auth.BasicAWSCredentials;
import com.amazonaws.services.dynamodbv2.AmazonDynamoDB;
import com.amazonaws.services.dynamodbv2.AmazonDynamoDBClientBuilder;
import com.amazonaws.services.dynamodbv2.document.DynamoDB;
import com.amazonaws.services.dynamodbv2.document.TableCollection;

import java.net.URI;


public class V1Demo {
    public static AWSCredentialsProvider myCredentials = new AWSStaticCredentialsProvider(
        new BasicAWSCredentials("your_ak", "your_sk"));
    public static String ip = "***.***.***.***";
    public static void main(String[] args) {
        disableCertificateChecks();
        AmazonDynamoDB client = AmazonDynamoDBClientBuilder.standard()
            .withRegion("region-a")
            .withRequestHandlers(new RequestHandler(URI.create("http://" + ip + "8000")))
            .withCredentials(myCredentials)
            .build();
        DynamoDB dynamoDB = new DynamoDB(client);
        TableCollection res = dynamoDB.listTables();
        System.out.println(res);
    }

}
```

**----End**

Connecting to an Instance over HTTPS

**Prerequisites:**

- The instance and ECS must be in the same VPC and subnet.
- The instance security group must allow access from the ECS. For details, see **Setting Security Group Rules for a GeminiDB DynamoDB-Compatible Instance**.
- Enable SSL for the GeminiDB instance by following **Enabling SSL for a GeminiDB DynamoDB-Compatible Instance** and download a certificate by following **Downloading an SSL Certificate**. Take a Java application as an example. The command for importing the certificate is as follows:

  ```
  keytool -importcert -alias hw -file ca.cert -keystore truststore.jks -storepass password
  ```

**CAUTION**

1. For compatibility purposes, you can still use HTTP after SSL is enabled. To meet high security requirements, you can modify a parameter to disable HTTP. For details, choose **Service Tickets > Create Service Ticket** in the upper right corner of the console and contact the customer service.

2. Currently, an EIP cannot be used over HTTPS.

**Procedure**

**Step 1** Add Maven dependencies. Add dependencies related to AWS SDK for Java 2.*x* to the **pom.xml** file.

```
<dependencies>
    <dependency>
        <groupId>software.amazon.awssdk</groupId>
        <artifactId>dynamodb</artifactId>
        <version>2.x.x</version>
    </dependency>
 </dependencies>
```

**Step 2** Connect to the GeminiDB DynamoDB-Compatible instance using DynamoDBClient over HTTPS.

Java code example:

```
import com.amazonaws.auth.AWSCredentialsProvider;
import com.amazonaws.auth.AWSStaticCredentialsProvider;
import com.amazonaws.auth.BasicAWSCredentials;
import com.amazonaws.client.builder.AwsClientBuilder;
import com.amazonaws.services.dynamodbv2.AmazonDynamoDB;
import com.amazonaws.services.dynamodbv2.AmazonDynamoDBClientBuilder;
import com.amazonaws.services.dynamodbv2.model.*;
public class SourceDemo {
// AK/SK credentials can be automatically obtained from environment variable file home/.aws/credentials.
// If there is no such a file, you need to manually specify the AK/SK.
    public static AWSCredentialsProvider myCredentials = new AWSStaticCredentialsProvider(
            new BasicAWSCredentials("your_ak", "your_sk"));
    public static void main(String[] args) {
        AmazonDynamoDB client = AmazonDynamoDBClientBuilder.standard()
                .withEndpointConfiguration(new AwsClientBuilder.EndpointConfiguration("https://
127.0.0.1:8000", "region_a"))
                .withCredentials(myCredentials)
                .build();
        System.out.println(client.listTables());
        CreateTableRequest request = new CreateTableRequest()
                .withTableName("test_001")
                .withProvisionedThroughput(new ProvisionedThroughput(1000L, 1000L))
                .withKeySchema(
                        new KeySchemaElement("id", KeyType.HASH)
                )
                .withAttributeDefinitions(
                        new AttributeDefinition("id", ScalarAttributeType.N)
                );
        System.out.println(client.createTable(request));
    }
}
```

**----End**

GeminiDB DynamoDB-Compatible API is completely compatible with DynamoDB. For details about common operations, see official DynamoDB documents.

### FAQs

Question: What should I do if the DAS console cannot be redirected after I click **Log In** in the **Operation** column in the instance list or click **Log In** on the **Basic Information** page?

Solution: Set your browser to allow pop-ups and try again.

## 3.3 Common Practices

After buying a GeminiDB DynamoDB-Compatible instance, you can perform operations by following its common practices.

**Table 3-2** Common practices

| Practice | | Description |
| --- | --- | --- |
| Instance modifications | **Changing an Instance Name** | Describes how to change instance names to distinguish instances. |
| | **Resetting the Administrator Password** | Describes how to reset the database administrator password. You are advised to change it periodically to prevent risks such as password cracking. |
| | **Changing vCPUs and Memory** | Describes how to change the instance specifications to meet your service requirements. |
| Data backup | **Managing Automated Backups** | Describes how to create automated backups during the backup period of an instance. The system saves the automated backups based on the specified retention period. |
| | **Managing Manual Backups** | Describes how to create manual backups for an instance. These backups can be used to restore data for improved reliability. |
| | **Managing Cross-Region Backups** | Describes how to set a cross-region backup policy for an instance. You can use the backup files in the target region to restore data to a new instance in another region. |
| | **Managing Table-level Backups** | Describes how to create a table-level backup for an instance. If a database or table is deleted maliciously or accidentally, backups can help restore your data. |
| Data restoration | **Restoring a Backup to a New Instance** | Describes how to restore an existing automated or manual backup to a new instance. The restored data is the same as the backup data. |

| Practice | | Description |
| --- | --- | --- |
| | **Restoring a Backup to a Specified Point in Time** | Describes how to use an automated backup to restore instance data to a specified point in time. |
| Log management | **Viewing and Exporting Slow Query Logs** | Describes how to view database-level slow query logs. Any query that takes longer than an execution time threshold (in milliseconds) will be logged. With slow query logs, you can identify and optimize slow statements. |

# 4 Working with GeminiDB DynamoDB-Compatible API

## 4.1 Using IAM to Grant Access to GeminiDB DynamoDB-Compatible API

### 4.1.1 Creating a User and Granting Permissions to Use GeminiDB DynamoDB-Compatible API

This section describes how to use **IAM** to control fine-grained permissions for your GeminiDB resources. With IAM, you can:

- Create IAM users for employees based on your enterprise's organizational structure. Each IAM user will have their own security credentials for accessing GeminiDB resources.
- Grant only the permissions required for users to perform a specific task.
- Entrust a Huawei Cloud account or cloud service to perform efficient O&M on your GeminiDB resources.

If you do not need to create an IAM user, skip this section.

The following describes the procedure for granting permissions (see **Figure 4-1**).

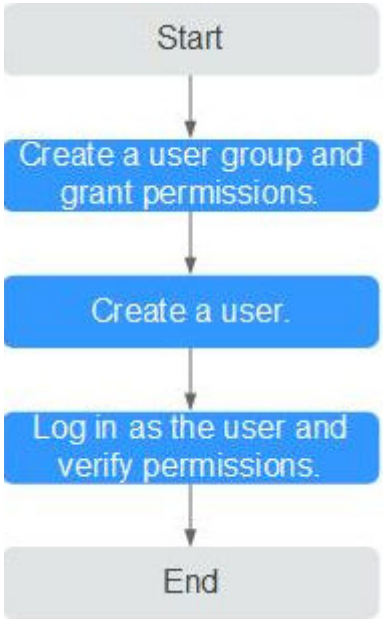
#### Prerequisites

You have learned about permissions that can be granted to the user group. For details, see **GeminiDB Permissions**. For system permissions of other services, see **Permission Policies**.

### Process Flow

**Figure 4-1** Process of granting GeminiDB permissions




1. **Create a user group and assign permissions**.

   Create a user group on the IAM console and assign the **GeminiDB FullAccess** permission to the group.

   NOTE

   To use some interconnected services, you also need to configure permissions of such services.

   For example, when connecting to your instance through DAS, you need to obtain the **GeminiDB FullAccess** and **DAS FullAccess** permissions.

2. **Create an IAM user** and add it to a user group.

   Create a user on the IAM console and add the user to the group created in **1**.

3. **Log in** and verify permissions.

   Log in to the management console using the created user, and verify the user's permissions:

   Choose **Service List** > **GeminiDB** and click **Buy DB Instance**. If you can buy an instance, the required permission policy has taken effect.

## 4.1.2 Custom Policies for GeminiDB DynamoDB-Compatible API

Custom policies can be created to supplement the system-defined policies of GeminiDB. For the actions supported for custom policies, see **Permissions Policies and Supported Actions**.

You can create custom policies in either of the following ways:

- Visual editor: Select cloud services, actions, resources, and request conditions. This does not require knowledge of policy syntax.

- JSON: Edit JSON policies from scratch or based on an existing policy.

For details, see **Creating a Custom Policy**. The following describes examples of common GeminiDB custom policies.

## Example Custom Policy

- Example 1: Allowing users to create GeminiDB instances

```
{
   "Version": "1.1",
   "Statement": [
      {
         "Effect": "Allow",
         "Action": [
            "nosql:instance:create"
         ]
      }
   ]
}
```

- Example 2: Refusing users to delete GeminiDB instances

  A policy with only "Deny" permissions must be used in conjunction with other policies to take effect. If the policies assigned to a user contain both Allow and Deny actions, the Deny actions take precedence over the Allow actions.

  The following method can be used if you need to assign permissions of the **GeminiDB FullAccess** policy to a user but you want to prevent the user from deleting GeminiDB instances. Create a custom policy for denying instance deletion, and attach both policies to the group to which the user belongs. Then, the user can perform all operations on GeminiDB instances except deleting GeminiDB instances. The following is an example of the deny policy:

```
{
   "Version": "1.1",
   "Statement": [
      {
        "Effect": "Deny"
        "Action": [
            "nosql:instance:delete"
         ],
       }
   ]
}
```

- Example 3: Defining permissions for multiple services in a policy

  A custom policy can contain the actions of multiple services that are of the global or project-level type. The following is an example policy containing actions of multiple services:

```
{
    "Version": "1.1",
    "Statement": [
        {
            "Action": [
                "nosql:instance:create",
                "nosql:instance:rename",
                "nosql:instance:delete",
                "vpc:publicIps:list",
                "vpc:publicIps:update"
            ],
            "Effect": "Allow"
        }
    ]
}
```

# 4.2 Buying a GeminiDB DynamoDB-Compatible Instance

This section describes how to buy a GeminiDB DynamoDB-Compatible instance.

By default, each tenant can create a maximum of 50 GeminiDB DynamoDB-Compatible instances. To request a higher quota, choose **Service Tickets > Create Service Ticket** in the upper right corner of the console and contact customer service personnel.

## Prerequisites

- You have created a Huawei Cloud account.

## Procedure

**Step 1** **Log in to the Huawei Cloud console**.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, click **Buy DB Instance**.

**Step 4** On the displayed page, specify a billing mode and instance specifications and click **Next**.

**Figure 4-2** Billing mode and basic information

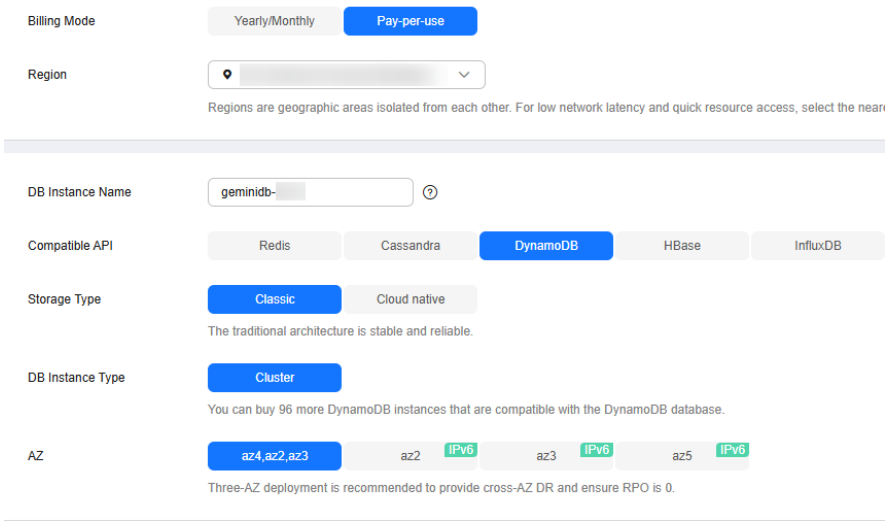


**Table 4-1** Billing parameters

| Parameter | Description |
| --- | --- |
| Billing Mode | Select **Yearly/Monthly** or **Pay-per-use**.<br>• **Yearly/Monthly**<br>– Specify **Required Duration**. The system deducts fees from your account based on the service price.<br>– If you do not need such an instance any longer after it expires, change the billing mode to pay-per-use. For details, see **Changing a Yearly/Monthly Instance to Pay-per-Use**.<br>**NOTE**<br>▪ Yearly/Monthly instances cannot be deleted directly. If such an instance is no longer required, unsubscribe from it. For details, see **How Do I Unsubscribe from a Yearly/Monthly Instance?**.<br>▪ Yearly/Monthly instances with cloud native storage are now in OBT. To use such an instance, choose **Service Tickets > Create Service Ticket** in the upper right corner of the console and contact the customer service.<br>• **Pay-per-use**<br>– If you select this billing mode, you are billed based on how much time the instance is in use.<br>– To use an instance for a long time, change its billing mode to yearly/monthly to reduce costs. For details, see **Changing a Pay-per-Use Instance to Yearly/Monthly**. |

**Table 4-2** Basic information

| Parameter | Description |
| --- | --- |
| Region | Region where a tenant is located<br>**NOTE**<br>To reduce network latency, select a region nearest from which you will access the instance. Instances deployed in different regions cannot communicate with each other over a private network. After you buy an instance, you cannot change its region. |
| DB Instance Name | The instance name:<br>• Can be the same as an existing instance name.<br>• Can contain 4 to 64 characters and must start with a letter. It is case-sensitive and allows only letters, digits, hyphens (-), and underscores (_).<br>After an instance is created, you can change its name. For details, see **Changing an Instance Name**. |

| Parameter | Description |
| --- | --- |
| Compatible API | DynamoDB<br>GeminiDB is compatible with mainstream NoSQL databases, including Redis, DynamoDB, Cassandra, HBase, MongoDB, and InfluxDB. You can select GeminiDB APIs by following **How Do I Select an API?** |
| Storage Type | • **Classic**: classic architecture with decoupled storage and compute<br>• **Cloud native**: more flexible, new-gen version with support for more AZs<br>**NOTE**<br>– The way you use instances with classic or cloud native storage is similar. Cloud native storage supports more AZs. If both classic and cloud native are supported, you can select any of them.<br>– Cloud native storage is now in OBT. To use it, choose **Service Tickets > Create Service Ticket** in the upper right corner of the console and contact the customer service. |
| DB Instance Type | Cluster<br>One cluster consists of at least three nodes. A cluster is easy to scale out to meet increasing data growth needs. A cluster is recommended when dealing with stringent availability demands, substantial data volumes, and the need for seamless scalability. |
| AZ | Availability zone where the instance is created. An AZ is a part of a region with its own independent power supplies and networks. AZs are physically isolated but can communicate with each other over a private network.<br>Instances can be deployed in a single AZ or three AZs.<br>• To deploy instances in a single AZ, select one AZ.<br>• To deploy instances across AZs for disaster recovery, select three AZs, where the instance nodes will be evenly distributed. |

**Figure 4-3** Specifications and storage

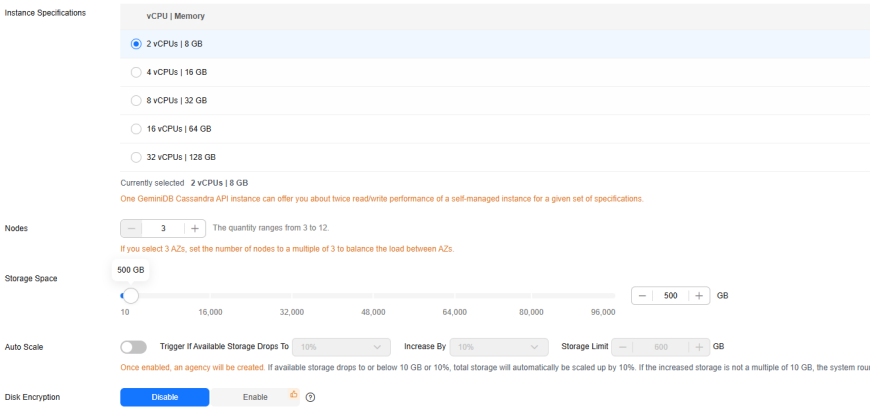


**Table 4-3** Specifications and storage

| Parameter | Description |
| --- | --- |
| Instance Specifications | Decoupled storage and compute and software-hardware synergy deliver twice or more the performance of a self-managed database with the same specifications. When you create an instance, select higher specification and specify as few nodes as possible. For example, if you need 8 vCPUs, 32 GB, and 6 nodes for an open-source instance, then for a GeminiDB DynamoDB-Compatible instance with 8 vCPUs and 32 GB of memory, you only need 3 nodes.<br>Select appropriate specifications based on the CPU-memory ratio.<br>After an instance is created, you can change its specifications. For details, see **Changing vCPUs and Memory**. |
| Nodes | Number of nodes that the instance is deployed on.<br>Currently, a maximum of 60 nodes are supported. To add more, choose **Service Tickets > Create Service Ticket** in the upper right corner of the console and contact the customer service.<br>After an instance is created, you can add nodes. For details, see **Manually Adding Instance Nodes**. |

| Parameter | Description |
| --- | --- |
| Storage Space | Instance storage space. The range depends on the instance specifications. For details, see **Instance Specifications**.<br>To scale up classic storage, you need to add at least 1 GB each time. To scale up cloud native storage, you need to add at least 10 GB each time. The value must be an integer.<br>When configuring storage space of a GeminiDB DynamoDB-Compatible instance, you are advised to enable autoscaling. Then set trigger conditions and limit of autoscaling. After autoscaling is triggered, the system automatically scales up the storage to ensure that the instance has sufficient storage and keeps available. Take care with the following parameters:<br>• **Trigger If Available Storage Drops To**: storage threshold for triggering autoscaling. When the available storage usage drops to a specified threshold or the available storage drops to 10 GB, autoscaling is triggered.<br>• **Increase By**: percentage that your instance storage will be scaled up at. If the increased storage is not a multiple of 10 GB, the system will round it up to the nearest multiple of 10 GB. At least 100 GB is added each time.<br>• **Storage Limit**: maximum amount that the system can automatically scale up an instance's storage to. The value must be no less than the current storage of your instance and cannot exceed the maximum storage supported by your instance.<br>After an instance is created, you can scale up its storage if necessary. For details, see **Manually Scaling Up Storage Space**.<br>**NOTE**<br>• Once **Auto Scale** is enabled, an agency will be created and fees will be automatically deducted.<br>• Autoscaling is available only to users with required permissions. To use it, choose **Service Tickets > Create Service Ticket** in the upper right corner of the console and contact the customer service.<br>• You can enable **Auto Scale** after an instance is created. For details, see **Automatically Scaling Up Storage Space**. |

| Parameter | Description |
| --- | --- |
| Disk Encryption | You can select to enable disk encryption based on service requirements.<br>• **Disable**: Data is not encrypted.<br>• **Enable**: If you select this option, your data will be encrypted on disks and stored in ciphertext after you create an instance. When you download encrypted objects, the ciphertext will be decrypted into plain text and then sent to you. Disk encryption can improve data security and may have slight impacts on database writes and reads.<br>– **Key Name**: Select an existing key or create one.<br>– To use a shared key, ensure that you have created an agency. For details, see **Creating an Agency (by a Delegating Party)**. Select another account from the drop-down list to share the key of the current account. VPC owners can share the keys with one or multiple accounts through **Resource Access Manager (RAM)**. For details, see **Creating a Resource Share**.<br>– Enter a key ID. The key must be in the current region.<br>**NOTE**<br>– This function is now in OBT. To use it, choose **Service Tickets > Create Service Ticket** in the upper right corner of the console and contact the customer service.<br>– An agency will be created after disk encryption is enabled.<br>– After an instance is created, the disk encryption status and the key cannot be changed.<br>– The key cannot be disabled, deleted, or frozen when being used. Otherwise, the database becomes unavailable.<br>– For details about how to create a key, see "Creating a CMK" in *Data Encryption Workshop User Guide*. |

**Figure 4-4** Network configuration

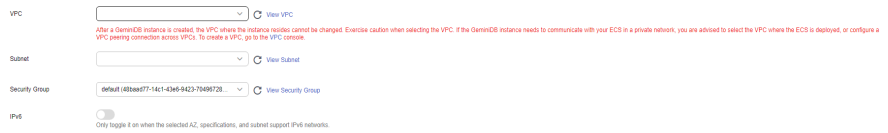


**Table 4-4** Network configuration

| Parameter | Description |
| --- | --- |
| VPC | Virtual private network where your instances are located. A VPC isolates networks for different services. You can select an existing VPC or create a VPC.<br>If there are no VPCs available, the system automatically allocates a VPC to you.<br>For details, see "Creating a VPC" in the *Virtual Private Cloud User Guide*.<br>**NOTE**<br>• After a GeminiDB DynamoDB-Compatible instance is created, the VPC where the instance is deployed cannot be changed.<br>• To connect a GeminiDB DynamoDB-Compatible instance to an ECS over a private network, ensure they are in the same VPC. If they are not, create a **VPC peering connection** between them. |
| Subnet | A subnet where your instance is created. The subnet provides dedicated and isolated networks, improving network security.<br>**NOTE**<br>An IPv6 subnet cannot be associated with your instance. Select an IPv4 subnet. |
| Security Group | A security group controls access between GeminiDB DynamoDB-Compatible API and other services. Ensure that the security group you selected allows your client to access the instance.<br>If no security group is available, the system creates one for you. |
| SSL | A security protocol. Secure Sockets Layer (SSL) certificates set up encrypted connections between clients and servers, preventing data from being tampered with or stolen during transmission.<br>You can enable SSL to improve data security. After an instance is created, connect to the instance through HTTPS. For details, see **Connecting to an Instance over HTTPS**.<br>**NOTE**<br>After SSL is enabled, it cannot be disabled. |

**Figure 4-5** Database configuration

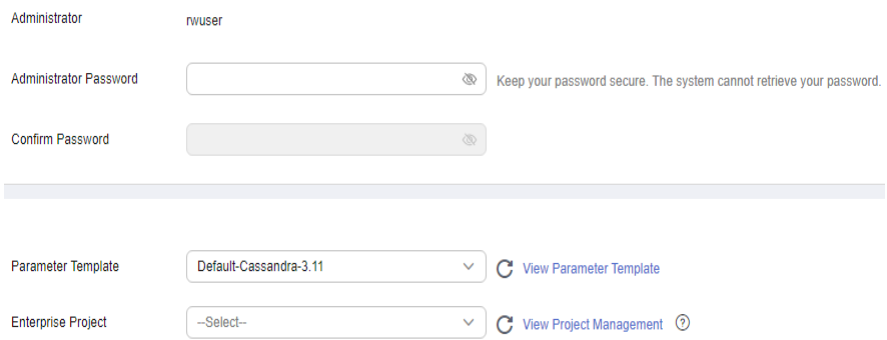


**Table 4-5** Database configuration

| Parameter | Description |
| --- | --- |
| Administrator | Username of the administrator account. The default value is **rwuser**. |
| Administrator Password | Password of the administrator account. The password:<br>• Can include 8 to 32 characters.<br>• Can include uppercase letters, lowercase letters, digits, and any of the following special characters: ~!@#%^*-_=+?<br>• For security reasons, set a strong password. The system will verify the password strength.<br>Keep your password secure. The system cannot retrieve it if it is lost. |
| Confirm Password | This password must be consistent with administrator password. |
| Enterprise Project | This parameter is provided for enterprise users.<br>An enterprise project groups cloud resources, so you can manage resources and members by project. The default project is **default**.<br>Select an enterprise project from the drop-down list. For more information about enterprise projects, see ***Enterprise Management User Guide***. |

**Figure 4-6** Tag configuration

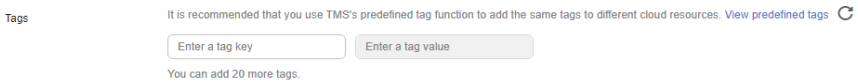


**Table 4-6** Tags

| Parameter | Description |
| --- | --- |
| Tags | (Optional) You can tag your GeminiDB DynamoDB-Compatible instances to identify and manage their resources.<br>A maximum of 20 tags can be added for each instance.<br>If your organization has configured tag policies related to GeminiDB DynamoDB-Compatible API, tag instances based on the tag policies. If a tag does not comply with the policies, an instance may fail to be created. Contact your organization administrator to learn more about tag policies.<br>A tag consists of a tag key and a tag value.<br>• A tag key is mandatory if the instance will be tagged. Each tag key is unique for each instance. It can contain 1 to 128 characters, cannot start with _**sys**_, and cannot start or end with a space. Only letters, digits, spaces, and the following special characters are allowed: -_@.:/+=<br>• A tag value is optional if the instance will be tagged. The value can be empty.<br>The value can contain a maximum of 255 characters. Only letters, digits, spaces, and the following special characters are allowed: _.-:+=@/<br>After an instance is created, you can view its tag details on the **Tags** tab. In addition, you can add, modify, and delete tags of an existing instance. For details, see **Tag Management**. |

**Figure 4-7** Required duration configuration


**Table 4-7** Required duration

| Parameter | Description |
| --- | --- |
| Required Duration | The length of your subscription if you select **Yearly/Monthly** billing. Subscription lengths range from one month to three years. |
| Auto-renew | • This option is not selected by default.<br>• If you select this option, the instance is automatically renewed based on the subscription duration. |

**Step 5** On the displayed page, confirm instance details.

- Yearly/Monthly
  - To modify the configurations, click **Previous**.

  - If no modification is required, read and agree to the service agreement, click **Pay Now**, and complete the payment.
- Pay-per-use
  - To modify the configurations, click **Previous**.
  - If no modification is required, read and agree to the service agreement and click **Submit**.

**Step 6** On the **Instances** page, view and manage the created instance.

- It takes about 5 to 9 minutes to create an instance. During the process, the instance status is **Creating**.
- After the instance is created, its status becomes **Available**.

  You can click  in the upper right corner to refresh the instance status.
- An automated backup policy is enabled by default during instance creation. After the instance is created, a full backup is created.

**Figure 4-8** Instance successfully purchased

**----End**

# 4.3 Instance Connection and Management

## 4.3.1 Connection Methods

Figure 4-9 shows the process of connecting to a GeminiDB DynamoDB-Compatible instance.

**Figure 4-9** Connection Methods

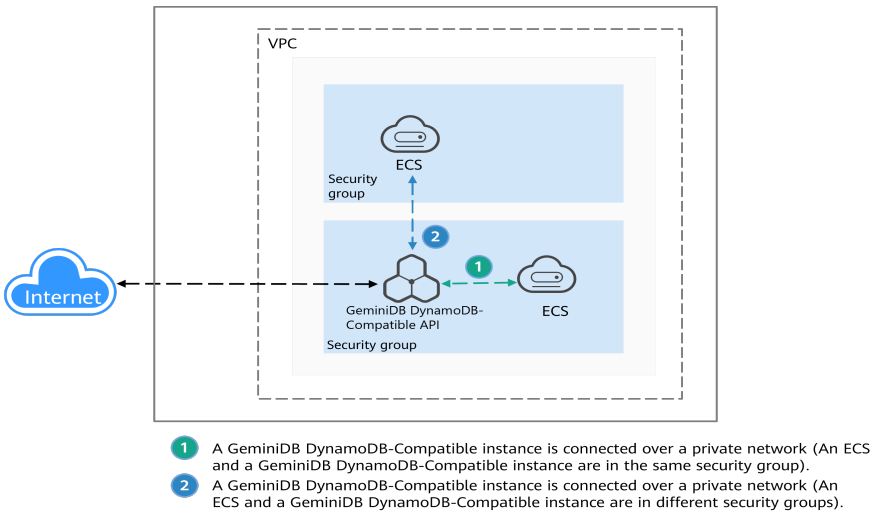


**Table 4-8** Network connection

| Method | Scenario | Default Port | Description |
| --- | --- | --- | --- |
| Load balancer | Suitable for production environments such as ECSs and functions. Requests are evenly distributed to available nodes. | 80 | • Accessible only in a VPC |

**Table 4-9** Code connection

| Method | Scenario | Description |
| --- | --- | --- |
| Connecting to a GeminiDB DynamoDB-Compatible Instance Using Java | An example of connecting to a GeminiDB DynamoDB-Compatible instance using Java | An instance can be connected using Java. |
| Connecting to a GeminiDB DynamoDB-Compatible Instance Using Python | An example of connecting to a GeminiDB DynamoDB-Compatible instance using Python | An instance can be connected using Python. |

# 4.3.2 Connecting to a GeminiDB DynamoDB-Compatible Instance Using Program Code

## 4.3.2.1 Connecting to a GeminiDB DynamoDB-Compatible Instance Using Java

This section describes how to connect to a GeminiDB DynamoDB-Compatible instance using Java.

### Usage Notes

- The target instance and ECS must be in the same VPC and subnet.
- The instance security group must allow access from the ECS.

  **Scenario 1**: If the instance is **associated with the default security group**, you do not need to set security group rules.

  **Scenario 2**: If the instance is **not associated with the default security group**, check whether the security group rules allow access from the ECS.

  - If yes, the ECS can connect to the instance.
  - If no, add an inbound rule to the security group.

    For details, see **Setting Security Group Rules for a GeminiDB DynamoDB-Compatible Instance**.

## Prerequisites

- A GeminiDB DynamoDB-Compatible instance has been created.
- For details about how to create an ECS, see **Purchasing an ECS** in *Getting Started with Elastic Cloud Server*.
- JDK has been installed on the ECS.

## Obtaining the IP Address of a GeminiDB DynamoDB-Compatible Instance

**Step 1** Log in to the Huawei Cloud console.

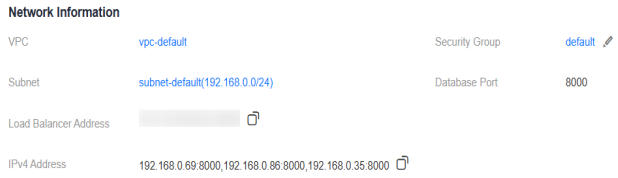
**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, click the instance name. On the displayed **Basic Information** page, view IP addresses of the load balancer and each node.

NOTE

If no load balancer address is available, choose **Service Tickets > Create Service Ticket** in the upper right corner of the console and contact the customer service.

**Figure 4-10** Viewing the IP address




**----End**

Replace the IP address in the following code example with the IP address queried in **Step 3**.

Java code example:

```
package com.huawei.dbs.test;
import com.amazonaws.client.builder.AwsClientBuilder;
import com.amazonaws.auth.AWSCredentialsProvider;
import com.amazonaws.auth.AWSStaticCredentialsProvider;
import com.amazonaws.auth.BasicAWSCredentials;
import com.amazonaws.services.dynamodbv2.AmazonDynamoDB;
import com.amazonaws.services.dynamodbv2.AmazonDynamoDBClientBuilder;
import com.amazonaws.services.dynamodbv2.document.DynamoDB;
import com.amazonaws.services.dynamodbv2.document.TableCollection;
public class ListTableCase {
    public static AWSCredentialsProvider myCredentials = new AWSStaticCredentialsProvider(
            new BasicAWSCredentials("your_ak", "your_sk"));
    public static String ip = "***.***.***.***";
    public static void main(String[] args) {
        AmazonDynamoDB client = AmazonDynamoDBClientBuilder.standard()
```

```
            .withEndpointConfiguration(new AwsClientBuilder.EndpointConfiguration("http://" + ip  ,
"region-01"))
            .withCredentials(myCredentials)
            .build();
        DynamoDB dynamoDB = new DynamoDB(client);
        TableCollection res = dynamoDB.listTables();
        System.out.println(res);
    }
}
```

### 4.3.2.2 Connecting to a GeminiDB DynamoDB-Compatible Instance Using Python

This section describes how to connect to a GeminiDB DynamoDB-Compatible instance using Python.

## Usage Notes

- The target instance and ECS must be in the same VPC and subnet.
- The instance security group must allow access from the ECS.

  **Scenario 1**: If the instance is **associated with the default security group**, you do not need to set security group rules.

  **Scenario 2**: If the instance is **not associated with the default security group**, check whether the security group rules allow access from the ECS.
  - If yes, the ECS can connect to the instance.
  - If no, add an inbound rule to the security group.

    For details, see **Setting Security Group Rules for a GeminiDB DynamoDB-Compatible Instance**.

## Prerequisites

- A GeminiDB DynamoDB-Compatible instance has been created.
- For details about how to create an ECS, see **Purchasing an ECS** in *Getting Started with Elastic Cloud Server*.
- JDK has been installed on the ECS.

## Obtaining the IP Address of a GeminiDB DynamoDB-Compatible Instance

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, click the instance name. On the displayed **Basic Information** page, view IP addresses of the load balancer and each node.

NOTE

If no load balancer address is available, choose **Service Tickets > Create Service Ticket** in the upper right corner of the console and contact the customer service.

**Figure 4-11** Viewing the IP address

**----End**

Replace the IP address in the following code example with the IP address queried in **Step 3**.

Python code example:

```
#!/usr/bin/python
import boto3


url = 'http://***.***.***.***'
dynamodb = boto3.resource('dynamodb',
                          endpoint_url=url,
                          aws_access_key_id='ak',
                          aws_secret_access_key='sk',
                          region_name="region-a")
```

GeminiDB DynamoDB-Compatible API is completely compatible with DynamoDB. For details about common operations, see official DynamoDB documents.

### 4.3.2.3 Connecting to an Instance over HTTPS

This section describes how to connect to a GeminiDB DynamoDB-Compatible instance over HTTPS.

**Prerequisites:**

- The instance and ECS must be in the same VPC and subnet.
- The instance security group must allow access from the ECS. For details, see **Setting Security Group Rules for a GeminiDB DynamoDB-Compatible Instance**.
- Enable SSL for the GeminiDB instance by following **Enabling SSL for a GeminiDB DynamoDB-Compatible Instance** and download a certificate by following **Downloading an SSL Certificate**. Take a Java application as an example. The command for importing the certificate is as follows:
  ```
  keytool -importcert -alias hw -file ca.cert -keystore truststore.jks -storepass password
  ```

**CAUTION**

1. For compatibility purposes, you can still use HTTP after SSL is enabled. To meet high security requirements, you can modify a parameter to disable HTTP. For details, choose **Service Tickets > Create Service Ticket** in the upper right corner of the console and contact the customer service.

2. Currently, an EIP cannot be used over HTTPS.

**Procedure**:

**Step 1** Add Maven dependencies. Add dependencies related to AWS SDK for Java 2.*x* to the **pom.xml** file.

```
<dependencies>
    <dependency>
        <groupId>software.amazon.awssdk</groupId>
        <artifactId>dynamodb</artifactId>
        <version>2.x.x</version>
    </dependency>
 </dependencies>
```

**Step 2** Connect to the GeminiDB DynamoDB-Compatible instance using DynamoDBClient over HTTPS.

Java code example:

```
import com.amazonaws.auth.AWSCredentialsProvider;
import com.amazonaws.auth.AWSStaticCredentialsProvider;
import com.amazonaws.auth.BasicAWSCredentials;
import com.amazonaws.client.builder.AwsClientBuilder;
import com.amazonaws.services.dynamodbv2.AmazonDynamoDB;
import com.amazonaws.services.dynamodbv2.AmazonDynamoDBClientBuilder;
import com.amazonaws.services.dynamodbv2.model.*;
public class SourceDemo {
    public static AWSCredentialsProvider myCredentials = new AWSStaticCredentialsProvider(
            new BasicAWSCredentials("your_ak", "your_sk"));
    public static void main(String[] args) {
        AmazonDynamoDB client = AmazonDynamoDBClientBuilder.standard()
                .withEndpointConfiguration(new AwsClientBuilder.EndpointConfiguration("https://127.0.0.1",
"region_a"))
                .withCredentials(myCredentials)
                .build();
        System.out.println(client.listTables());
        CreateTableRequest request = new CreateTableRequest()
                .withTableName("test_001")
                .withProvisionedThroughput(new ProvisionedThroughput(1000L, 1000L))
                .withKeySchema(
                        new KeySchemaElement("id", KeyType.HASH)
                )
                .withAttributeDefinitions(
                        new AttributeDefinition("id", ScalarAttributeType.N)
                );
        System.out.println(client.createTable(request));
    }
}
```

**----End**

GeminiDB DynamoDB-Compatible API is completely compatible with DynamoDB. For details about common operations, see official DynamoDB documentation.

# 4.3.3 Connection Information Management

## 4.3.3.1 Setting Security Group Rules for a GeminiDB DynamoDB-Compatible Instance

A security group is a collection of access control rules for ECSs and GeminiDB DynamoDB-Compatible instances that have the same security requirements and are mutually trusted in a VPC.

To ensure database security and stability, you need to set a security group and add IP addresses and ports that can access the database before using GeminiDB DynamoDB-Compatible instances.

This section describes how to set security group rules for a GeminiDB DynamoDB-Compatible instance which is connected over a private or public network.

## Usage Notes

- By default, a tenant can create a maximum of 500 security group rules.
- Too many security group rules will increase the first packet latency. You are advised to create a maximum of 50 rules for each security group.

- Currently, each instance can be bound to only one security group.
- Table 4-10 describes the security group rules required for connecting to an instance over a private or public network.

**Table 4-10** Security group rules

| Scenario | Description |
| --- | --- |
| Connecting to an instance over a private network | When connecting to a GeminiDB DynamoDB-Compatible instance over a private network, set security group rules in either of the following ways:<br>• If the ECS and GeminiDB DynamoDB-Compatible instance are in the same security group, they can communicate with each other by default. No security group rule needs to be set.<br>• If they are in different security groups, you need to set security group rules for both of them.<br>– Set an inbound rule for the GeminiDB DynamoDB-Compatible instance by following **Procedure**.<br>– The default security group rule allows all outbound data packets, so you do not need to set a security rule for the ECS. If not all outbound traffic is allowed in the security group, set an outbound rule for the ECS. |
| Connecting to an instance over a public network | Set an inbound rule when connecting to a GeminiDB DynamoDB-Compatible instance over a public network by following **Procedure**. |

## Procedure

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, click the target instance go to the **Basic Information** page.

**Step 4** Set security group rules.

**Method 1**:

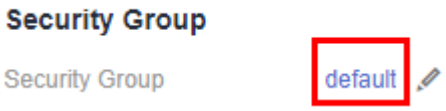
In the **Network Information** area on the **Basic Information** page, click the security group.

**Figure 4-12** Security group

**Method 2**

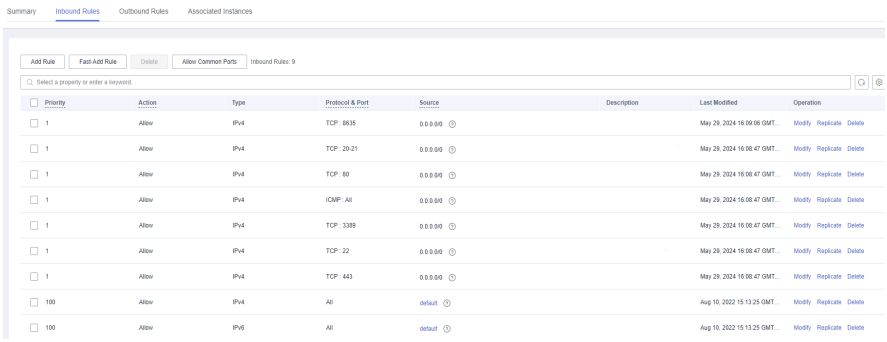
On the **Basic Information** page, choose **Connections** in the navigation pane on the left. In the **Security Group** area on the right, click the name of the security group. The **Security Group** page is displayed.




**Step 5** Add an inbound rule.

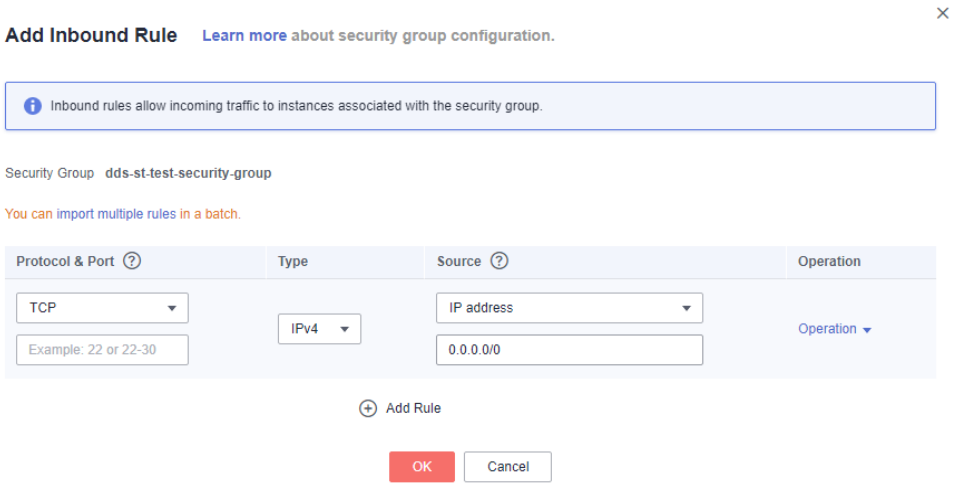
1. Click the **Inbound Rules** tab.

**Figure 4-13** Inbound rule




2. Click **Add Rule**. The **Add Inbound Rule** dialog box is displayed.

**Figure 4-14** Adding a rule




3. Add a security group rule as prompted.

**Table 4-11** Inbound rule settings

| Parameter | Description | Example Value |
| --- | --- | --- |
| Protocol & Port | – **Protocol**: Currently, GeminiDB DynamoDB-Compatible API supports only TCP.<br>– **Port**: The port (1 to 65535) for accessing the ECS. | TCP |
| Type | IP address type. This parameter is available after IPv6 is enabled.<br>– IPv4<br>– IPv6 | IPv4 |
| Source | **Source**: The source can be an IP address, a security group, or an IP address group which allows access from IP addresses or instances in other security groups. For example:<br>– xxx.xxx.xxx.xxx/32 (IPv4 address)<br>– xxx.xxx.xxx.0/24 (subnet)<br>– 0.0.0.0/0 (any IP address)<br>– sg-abc (security group) | 0.0.0.0/0 |
| Description | (Optional) Provides supplementary information about the security group rule.<br>The description can contain a maximum of 255 characters and cannot contain angle brackets (< or >). | - |

**Step 6** Click **OK**.

**----End**

## 4.3.3.2 Binding an EIP to a GeminiDB DynamoDB-Compatible Instance Node

Elastic IP provides independent public IP addresses and bandwidth for public access. After a GeminiDB DynamoDB-Compatible instance is created, you can bind an EIP to the instance to access it over a public network. You can also unbind the EIP from the instance.

## Usage Notes

- To change the EIP that has been bound to a node, unbind it from the node first.

## Binding an EIP

**Step 1** Log in to the Huawei Cloud console.

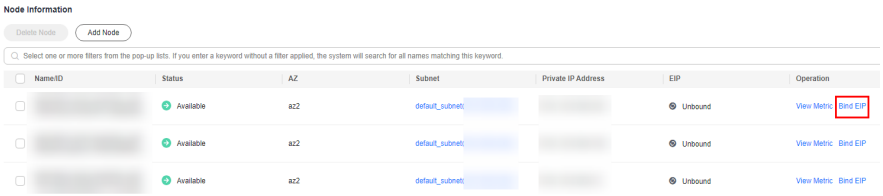
**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, click the target GeminiDB DynamoDB-Compatible instance.

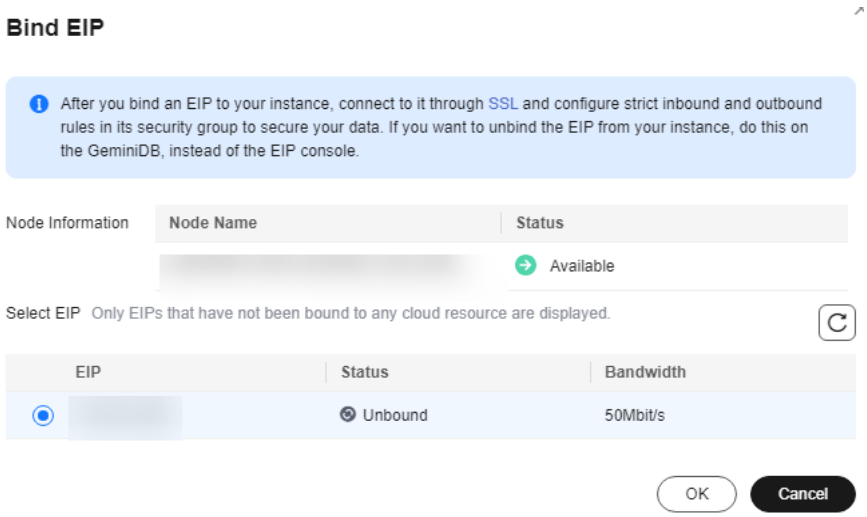
**Step 4** In the **Node Information** area on the **Basic Information** page, locate the target node and click **Bind EIP** in the **Operation** column.

**Figure 4-15** Binding an EIP




**Step 5** In the displayed dialog box, select the required EIP and click **OK**. If no available EIPs are displayed, click **View EIP** and create an EIP.

**Figure 4-16** Selecting an EIP




**Step 6** In the **EIP** column, view the EIP that has been bound.

To unbind the EIP from the DB instance, see **Unbinding an EIP**.

**----End**

## Unbinding an EIP

**Step 1** Log in to the Huawei Cloud console.

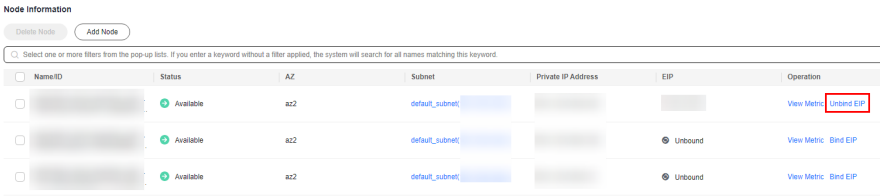
**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** If the node has been bound to an EIP, click the target instance on the **Instances** page.

**Step 4** In the **Node Information** area on the **Basic Information** page, locate the target node and click **Unbind EIP** in the **Operation** column.

**Figure 4-17** Unbinding an EIP




**Step 5** In the displayed dialog box, click **Yes**.

To bind an EIP to the instance again, see **Binding an EIP**.

**----End**

## 4.3.3.3 Viewing the IP Address and Port of a GeminiDB DynamoDB-Compatible Instance

This section describes how to view the IP address and port of a GeminiDB DynamoDB-Compatible instance on the GeminiDB console.

## Procedure

**Step 1** Log in to the Huawei Cloud console.

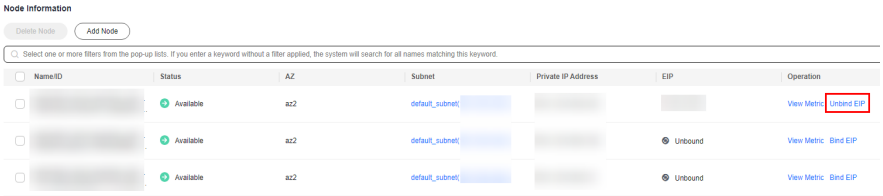
**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, click the target instance.

**Method 1**:

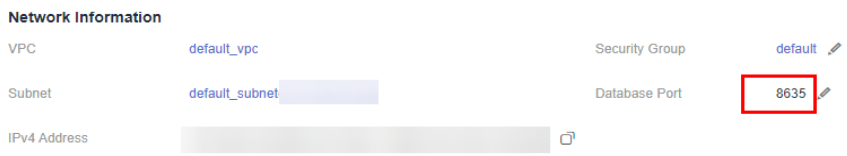
In the **Node Information** area on the **Basic Information** page, obtain the private IP address or EIP of each node.

**Figure 4-18** Viewing the IP address




View the instance port in the **Network Information** area. The default port is **8000**.

**Figure 4-19** Viewing the port



NOTE

Although the displayed port is **8635**, the default port is **8000**.

**Method 2**

In the navigation pane, choose **Connections** to obtain the private IP address, EIP, and port of the instance.

**Figure 4-20** Viewing the IP address and port


**----End**

## 4.3.3.4 Modifying the Security Group of a GeminiDB DynamoDB-Compatible Instance

This section describes how to modify the security group of a GeminiDB DynamoDB-Compatible instance.

## Usage Notes

If you are adding nodes to an instance, the security group cannot be changed.

## Procedure

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, locate the instance whose security group you want to change and click its name.

**Step 4** In the navigation pane, choose **Connections**.

**Step 5** In the **Security Group** area, click next to the security group name and select the security group of the instance.

- Click to submit the modification. This process takes about 1 to 3 minutes.
- Click to cancel the modification.

**Step 6** View the modification result.

**----End**

### 4.3.3.5 Enabling SSL for a GeminiDB DynamoDB-Compatible Instance

Secure Socket Layer (SSL) is an encryption-based Internet security protocol for establishing an encrypted link between a server and a client. It provides privacy, authentication, and integrity to Internet communications.

- Authenticates users and servers, ensuring that data is sent to the correct clients and servers.
- Encrypts data to prevent it from being intercepted during transfer.
- Ensures data integrity during transmission.

After SSL is enabled, you can establish an encrypted connection between your client and the instance you want to access to improve data security.

## Usage Notes

- After you enable or disable SSL, the established connection is interrupted. Restart the instance to apply the change.
- Enabling SSL will prolong network connection response and increase CPU usage. So, evaluate impacts on service performance before enabling SSL.
- GeminiDB DynamoDB-Compatible API supports only TLS 1.3 or later. It does not support TLS 1.0, TLS 1.1, or TLS 1.2.

## Enabling SSL

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, click the target instance go to the **Basic Information** page.

**Step 4** In the **DB Information** area, click  to enable SSL.

**Figure 4-21** Enabling SSL


Alternatively, choose **Connections** in the navigation pane. In the **Basic Information** area, click  in the **SSL** field to enable SSL.

**Figure 4-22** Enabling SSL

After SSL is enabled, you can connect to the GeminiDB DynamoDB-Compatible instance over HTTPS. For details, see **Connecting to an Instance over HTTPS**.

**----End**

## Disabling SSL

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, click the target instance go to the **Basic Information** page.

**Step 4** In the **DB Information** area, click  to disable SSL.

**Figure 4-23** Disabling SSL


Alternatively, choose **Connections** in the navigation pane. In the **Basic Information** area, click  in the **SSL** field to disable SSL.

**Figure 4-24** Disabling SSL


After SSL is disabled, you can connect to the GeminiDB DynamoDB-Compatible instance over HTTP. For details, see **Connecting to a GeminiDB DynamoDB-Compatible Instance Using Java**.

**----End**

## 4.3.3.6 Downloading an SSL Certificate

Secure Sockets Layer (SSL) encrypts connections between clients and servers, preventing data from being tampered with or stolen during transmission.

To improve data security, you can enable SSL when creating an instance. After the instance is created, you can connect to it over SSL. An SSL security certificate is required for connecting to an instance over SSL.

This section describes how to obtain an SSL certificate.

### Procedure

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, click the target instance go to the **Basic Information** page.

**Step 4** In the **DB Information** area, click next to **SSL**.

**Figure 4-25** Downloading an SSL certificate


**----End**

# 4.4 Data Migration

## 4.4.1 Solution Overview

This section describes how to migrate services to a GeminiDB DynamoDB-Compatible instance. If you have any questions about the migration, choose **Service Tickets > Create Service Ticket** in the upper right corner of the console to get technical support.

### Permissions

Ensure that the database port is enabled in the security group of the GeminiDB DynamoDB-Compatible instance.

### Migration Scenarios

**Table 4-12** Migration scenarios

| No. | Source | Destination | Reference |
| --- | --- | --- | --- |
| 1 | DynamoDB | GeminiDB DynamoDB-Compatible API | From DynamoDB to GeminiDB DynamoDB-Compatible API |

# 4.5 Instance Lifecycle Management

# 4.5.1 Restarting a GeminiDB DynamoDB-Compatible Instance

You may need to restart an instance for routine maintenance.

## Usage Notes

- Only instances in states **Available**, **Abnormal**, or **Checking restoration** can be restarted.
- Restarting an instance will interrupt services, so off-peak hours are the best time. Ensure that your application can be reconnected.
- After you restart an instance, all nodes in the instance are also restarted.
- If you enable operation protection, two-factor authentication is required for sensitive operations to secure your account and cloud products. For details about how to enable operation protection, see ***Identity and Access Management User Guide***.

## Procedure

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, locate the target instance and choose **More** > **Restart** in the **Operation** column.

You can also click the instance and click **Restart** in the upper right corner of the displayed page.

**Step 4** If you have enabled operation protection, click **Start Verification** in the **Restart DB Instance** dialog box. On the displayed page, click **Send Code**, enter the verification code, and click **Verify**. The page is closed automatically.

**Step 5** In the displayed dialog box, click **Yes**.

**----End**

# 4.5.2 Exporting Instance Information

## Scenarios

You can export information about all or selected instances to view and analyze instance information.

## Exporting All Instance Information

**Step 1** **Log in to the Huawei Cloud console.**

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, click  in the upper right corner. By default, information about all instances is exported. In the displayed dialog box, you can select the items to be exported and click **Export**.

**Step 4** After the export task is complete, check an XLS file is generated locally.

**----End**

### Exporting Information About Selected Instances

**Step 1** On the **Instances** page, select the target instances or search for required instances by project, compatible API, name, ID, or tag and click in the upper right corner. In the displayed dialog box, select the items to be exported and click **Export**.

**Step 2** After the export task is complete, check an XLS file is generated locally.

**----End**

## 4.5.3 Deleting a Pay-per-Use Instance

You can manually delete a pay-per-use instance on the **Instances** page. Before deleting a yearly/monthly instance, you need to unsubscribe from it. For details, see **How Do I Unsubscribe from a Yearly/Monthly Instance?**.

### Usage Notes

- Instances that an operation is being performed on cannot be deleted. They can be deleted only after the operations are complete.
- If a pay-per-use instance is deleted, its automated backups will also be deleted and you will no longer be billed for them. Manual backups, however, will be retained and generate additional costs.
- After an instance is deleted, all its data and automated backups are automatically deleted as well and cannot be recovered. You are advised to create a backup before deleting an instance. For details, see **Creating a Manual Backup**.
- After you delete an instance, all of its nodes are deleted.
- A deleted instance will be retained in the recycle bin for a period of time after being released, so you can rebuild the instance and restore data from it.

### Procedure

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, locate the target instance and choose **More** > **Delete** in the **Operation** column.

**Step 4** If you have enabled operation protection, click **Start Verification** in the **Delete DB Instance** dialog box. On the displayed page, click **Send Code**, enter the verification code, and click **Verify**. The page is closed automatically.

NOTE

If you enable operation protection, two-factor authentication is required for sensitive operations to secure your account and cloud products. For details about how to enable operation protection, see *Identity and Access Management User Guide*.

**Step 5** In the displayed dialog box, click **Yes**.

Deleted instances are not displayed in the instance list any longer.

**----End**

# 4.5.4 Recycling an Instance

You can move unsubscribed yearly/monthly and deleted pay-per-use GeminiDB DynamoDB-Compatible instances to the recycle bin and rebuild them if necessary.

## Usage Notes

- The recycling bin is enabled by default and cannot be disabled. Instances in the recycle bin are retained for 7 days by default, and this will not incur any charges.
- Currently, you can put a maximum of 100 instances into the recycle bin.
- If you delete an instance of full storage, the deleted instance will not be moved to the recycle bin.
- You can modify the retention period, and the changes only apply to the DB instances deleted after the changes, so exercise caution when performing this operation.
- After an instance is deleted, the most recent automated full backup (if no automated full backup is available one day ago, the latest one is retained) is retained and a full backup is performed. You can select any backup file to restore the instance data.

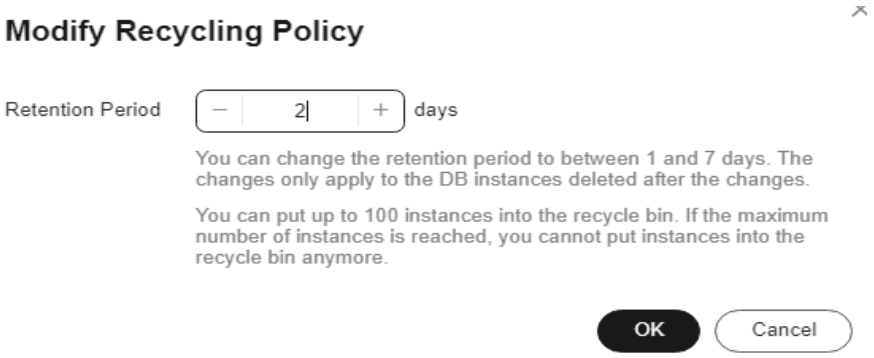
## Modifying the Recycling Policy

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Recycling Bin** page, click **Modify Recycling Policy**. In the displayed dialog box, set the retention period from 1 day to 7 days. Then, click **OK**.

**Figure 4-26** Modifying a recycling policy




**----End**

### Rebuilding an Instance

You can rebuild DB instances from the recycle bin within the retention period to restore data.

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Recycling Bin** page, locate the target instance and click **Rebuild** in the **Operation** column.

**Figure 4-27** Rebuilding an instance

**Step 4** On the displayed page, configure required parameters and submit the task.

**----End**

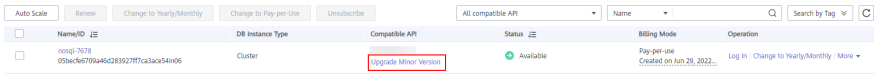
# 4.6 Instance Modifications

## 4.6.1 Upgrading a Minor Version

GeminiDB DynamoDB-Compatible API can be upgraded by installing patches to improve performance, release new features, or fix bugs.

If a new patch is released, you can click **Upgrade Minor Version** to upgrade your instance in the **Compatible API** column on the **Instances** page.

**Figure 4-28** Patch installation



If the kernel version of your instance has potential risks or major defects, has expired, or has been brought offline, the system will notify you by SMS message or email and deliver an upgrade task during maintenance.

### Usage Notes

- Upgrade your instance once there are new patches released.
- If the database version is a risky version, the system prompts you to upgrade the database patch.
- The instance will be restarted and services may be interrupted during the upgrade. The interruption duration depends on services, quantity of nodes, and the amount of service data. Upgrade your instance during off-peak hours.
- When you upgrade a cluster, services may be interrupted a number of times equal to the number of nodes in the cluster plus one. Each interruption will last for no more than 60s and will only affect the services on that node. If your cluster has more than one node, the upgrade duration is as follows:

  600 + (N x 60) ≤ Total upgrade duration (s) ≤ 600 + (N x 120)

For example, if there are 9 nodes in a cluster instance, the upgrade duration is 19 to 28 minutes.

The upgrade duration of most instances is close to 600+ (N x 60). If there are too many tokens on a single node, the upgrade duration may be increased.

- Before you upgrade a DR instance, upgrade the corresponding standby instance first and then the primary instance afterwards.

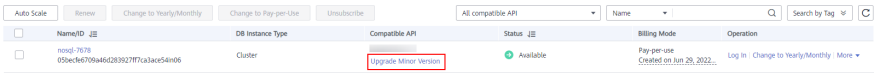
## Procedure

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, locate the instance you want to upgrade and click **Upgrade Minor Version** in the **Compatible API** column.

**Figure 4-29** Patch installation

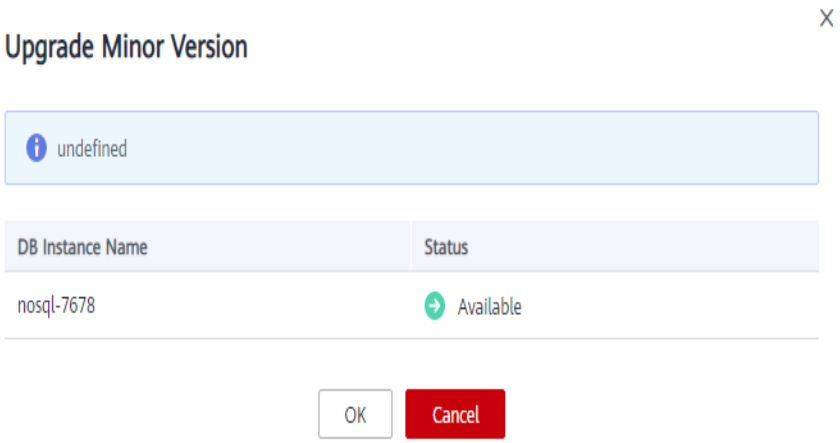



Alternatively, click the instance name to go to the **Basic Information** page. In the **DB Information** area, click **Upgrade Minor Version** in the **Compatible API** field.

**Figure 4-30** Patch installation


**Step 4** In the displayed dialog box, click **OK**.

**Figure 4-31** Confirming information




**Step 5** View the upgrade result on the **Instances** page.

- When the upgrade is ongoing, the instance status is **Upgrading minor version**.

- After the upgrade is complete, the instance status changes **Available**.

**----End**

## 4.6.2 Changing an Instance Name

This section describes how to change GeminiDB DynamoDB-Compatible instance names to distinguish instances.

### Method 1

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, click  next to the target instance name and change it.

- To submit the change, click **OK**.
- To cancel the change, click **Cancel**.

NOTE

The instance name:

- Can be the same as an existing instance name.
- Can include 4 to 64 bytes and must start with a letter. It is case-sensitive and allows only letters, digits, hyphens (-), and underscores (_).

**Step 4** View the results on the **Instances** page.

**----End**

### Method 2

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, click the target instance. The **Basic Information** page is displayed.

**Step 4** In the **Instance Information** area on the **Basic Information** page, click  next to **DB Instance Name** and change the instance name.

- Click  to submit the modification.
- Click  to cancel the modification.

NOTE

The instance name:

- Can be the same as an existing instance name.
- Can include 4 to 64 bytes and must start with a letter. It is case-sensitive and allows only letters, digits, hyphens (-), and underscores (_).

**Step 5** On the **Basic Information** page, check the new name.

**----End**

# 4.6.3 Resetting the Administrator Password

Describes how to reset the database administrator password. You are advised to change it periodically to prevent risks such as password cracking.

## Usage Notes

- If the instance status is **Available**, **Backing up**, **Checking restoration**, **Scaling up** or certain nodes become abnormal, you can reset the administrator password.
- The administrator password takes effect immediately after being reset.
- For two instances with an intra-region DR or cross-region dual-active relationship, make sure that they have the same administrator passwords.
- If you enable operation protection to improve the security of your account and cloud products, two-factor authentication is required for sensitive operations. For details about how to enable operation protection, see ***Identity and Access Management User Guide***.

**CAUTION**

You are advised to change the password during off-peak hours to avoid service interruption.

## Method 1

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, locate the instance whose password you want to reset and choose **More** > **Reset Password** in the **Operation** column.

**Step 4** Enter and confirm the new administrator password and click **OK**.

The password must contain 8 to 32 characters. It can include uppercase letters, lowercase letters, digits, and any of the following special characters: ~!@#%^*-_=+?

**Step 5** If you have enabled operation protection, click **Start Verification** in the displayed dialog box. On the displayed page, click **Send Code**, enter the verification code, and click **Verify**. The page is closed automatically.

**----End**

## Method 2

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, click the instance whose password you want to reset to go to the **Basic Information** page.

**Step 4** In the **DB Information** area, click **Reset Password** in the **Administrator** field.

**Step 5** Enter and confirm the new administrator password and click **OK**.

The password must be 8 to 32 characters in length and contain uppercase letters, lowercase letters, digits, and any of the following special characters: ~!@#%^*-_=+?

**Step 6** If you have enabled operation protection, click **Start Verification** in the displayed dialog box. On the displayed page, click **Send Code**, enter the verification code, and click **Verify**. The page is closed automatically.

**----End**

## 4.6.4 Changing vCPUs and Memory

This section describes how to change your instance vCPUs and memory to suit your service requirements.

### Usage Notes

- You can increase or decrease the vCPUs and memory.
- If one instance has multiple nodes, the change will be performed on the nodes one by one. It takes about 5 to 10 minutes for each node, and the total time required depends on the number of the nodes.
- For a node whose specifications are being changed, its computing tasks are handed over to other nodes. Change specifications of nodes during off-peak hours to prevent the instance from overload.
- Do not perform DDL operations when you change the instance specifications.

  NOTE

  A data definition language (DDL) is a language for defining data structures and database objects. Common examples of DDL statements are CREATE, ALTER, and DROP. Data Definition Language (DDL) is used to create, modify, and delete database objects, such as tables, indexes, views, functions, stored procedures, and triggers.
- vCPU and memory changes are applied on all nodes in sequence. During this process, temporary I/O disruptions or increased latency may occur. You are advised to perform this operation during off-peak hours.
- If you forcibly change the specifications of an instance when the instance is abnormal, services may be affected in seconds.

### Method 1

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, locate the instance whose vCPUs and memory you want to change and click its name.

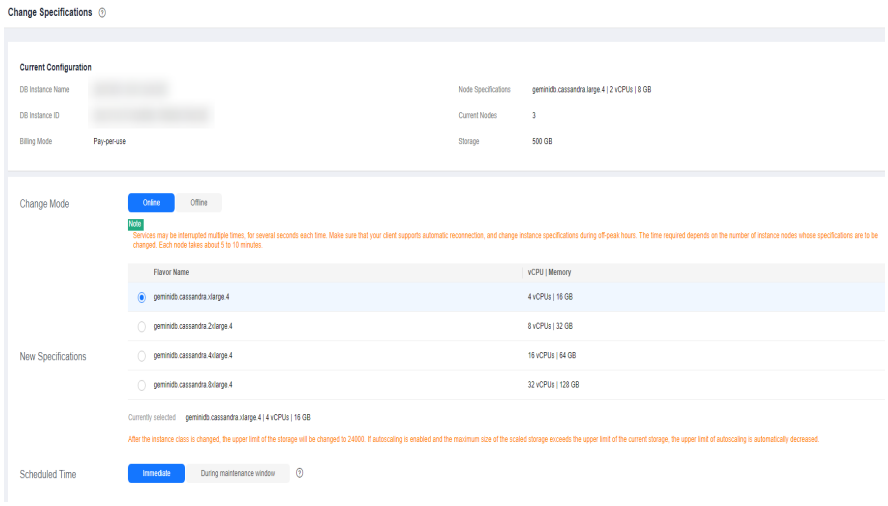
**Step 4** In the in **DB Information** area on the **Basic Information** page, click **Change** next to **Specifications**.

**Figure 4-32** Changing specifications


**Step 5** On the displayed page, select the required specifications and click **Next**.

**Figure 4-33** Changing specifications




**Step 6** On the displayed page, confirm the instance specifications.

- Yearly/Monthly
  - If you need to modify your settings, click **Previous**.
  - If you do not need to modify your settings, click **Submit**. If you are scaling up the instance specifications, go to the payment page, select a payment method, and complete the payment.
- Pay-per-use
  - If you need to modify your settings, click **Previous**.
  - If you do not need to modify your settings, click **Submit**.

**Step 7** View the change results.

In the **DB Information** area on the **Basic Information** page, you can see the new specifications.

**----End**

## Method 2

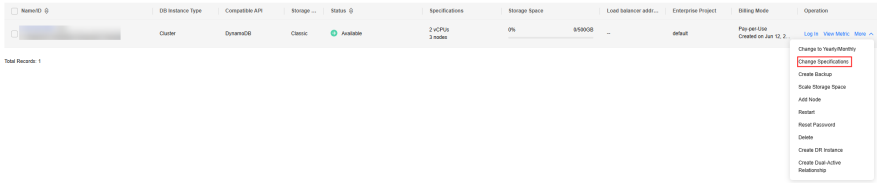
**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

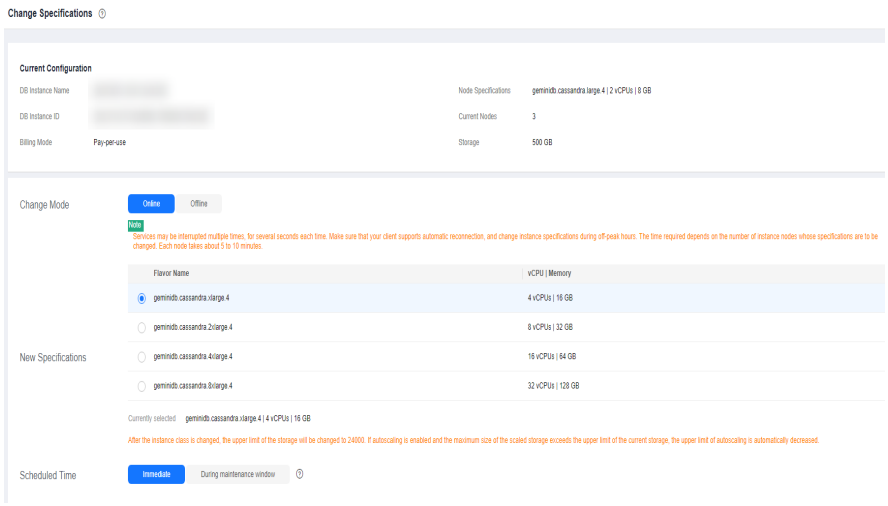
**Step 3** On the **Instances** page, locate the instance whose specifications you want to change and choose **More** > **Change Specifications** in the **Operation** column.

**Figure 4-34** Changing specifications



**Step 4** On the displayed page, select the required specifications and click **Next**.

**Figure 4-35** Changing specifications




**Step 5** On the displayed page, confirm the instance specifications.

- Yearly/Monthly
  - If you need to modify your settings, click **Previous**.
  - If you do not need to modify your settings, click **Submit**. If you are scaling up the instance specifications, go to the payment page, select a payment method, and complete the payment.
- Pay-per-use
  - If you need to modify your settings, click **Previous**.
  - If you do not need to modify your settings, click **Submit**.

**Step 6** View the change results.

In the **DB Information** area on the **Basic Information** page, you can see the new specifications.

**----End**

# 4.6.5 Setting a Maintenance Window

The default maintenance window is 10:00–14:00 (GMT+08:00) but you can change it if needed. To prevent service interruption, set the maintenance window to off-peak hours. Before calling this API:

## Usage Notes

- You can configure a maintenance window only for restarting a DB instance, changing an instance class, or upgrading the minor version of a DB instance.
- The specification change and patch upgrade that have been performed during the maintenance period cannot be performed immediately. The instance can be restarted immediately.
- You can cancel a task to be executed.
- Changing the maintenance window will not affect the timing that has already been scheduled.
- The maintenance window cannot overlap the time window configured for backups. Otherwise, scheduled tasks may fail.
- During the maintenance window, the scheduled task is scanned and executed every 10 minutes. If the task is delivered near the end of the maintenance period, the task may fail to be scanned and the execution is canceled.

## Setting a Maintenance Window

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, click the target instance. The **Basic Information** page is displayed.

**Step 4** In the **DB Information** area, locate **Maintenance Window** and click **Change**.

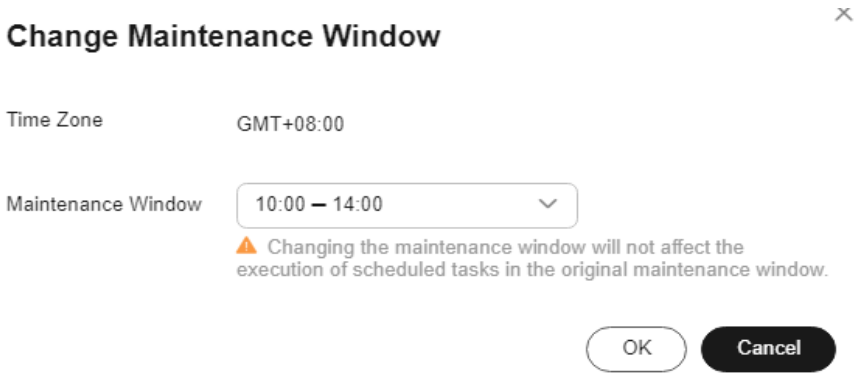
**Figure 4-36** The change button


**Step 5** On the **Change Maintainable Window** page, select the maintenance time period as needed, and then click **OK**.

Supported time periods: 02:00-06:00, 06:00-10:00, 10:00-14:00, 14:00-18:00, 18:00-22:00, and 22:00-02:00

**Figure 4-37** Changing a maintenance window




**Step 6** Check the result.

On the **Basic Information** page, you can view the changed maintenance window.

**----End**

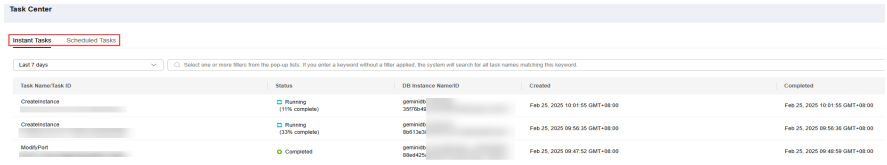
## Querying an Executed Task

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Task Center** page, click the **Instant Tasks** or **Scheduled Tasks** tab to view a task.

**Figure 4-38** Querying a task




**----End**

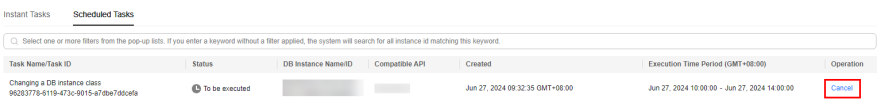
## Canceling a Scheduled Task

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Task Center** page, locate a scheduled task, and click **Cancel** in the **Operation** column.

**Figure 4-39** Canceling a task

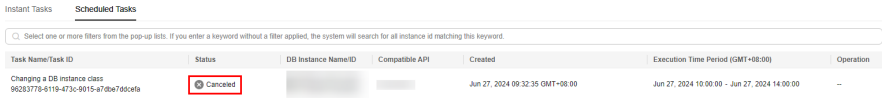



**Step 4** Check the result.

On the **Task Center** page, you can view the result. After the task is cancelled, its status changes to **Cancelled**.

**Figure 4-40** Checking cancelled tasks




**----End**

# 4.6.6 Adding and Deleting Instance Nodes

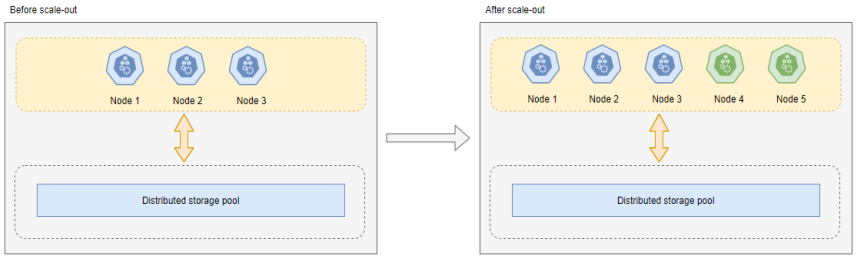
## 4.6.6.1 Overview

After you purchase a GeminiDB DynamoDB-Compatible instance, resource requirements may change along with workload volumes. You can scale your instance nodes in the following ways.

## Manually Adding Instance Nodes

For example, if three nodes have been deployed and two more nodes need to be added, there will be five nodes in total. For details, see **Manually Adding Instance Nodes**.

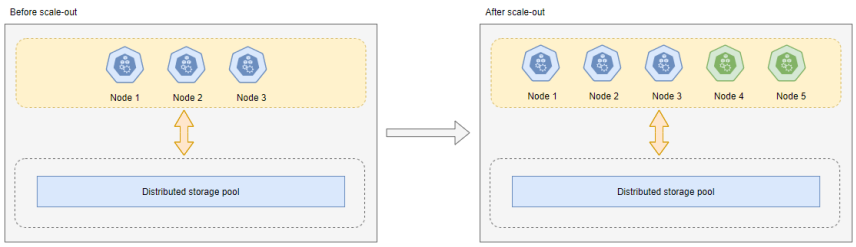
**Figure 4-41** Adding instance nodes




## Automatically Adding Instance Nodes

For example, if three nodes have been deployed and two more nodes need to be added, there will be five nodes in total. For details, see **Automatically Adding Instance Nodes**.

**Figure 4-42** Adding instance nodes

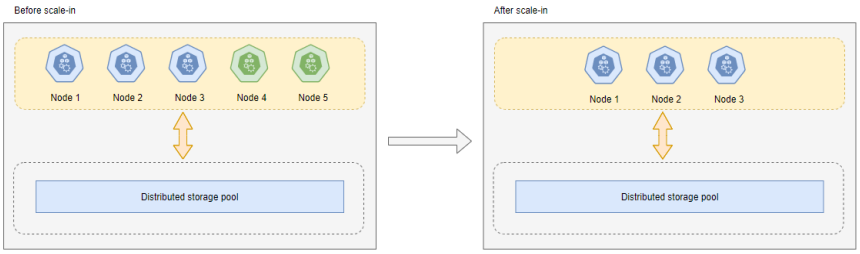


## Manually Deleting Instance Nodes

For example, if five nodes have been deployed and two of them need to be deleted, three nodes will be left. For details, see **Manually Deleting Instance Nodes**.

**Figure 4-43** Deleting instance shards




## 4.6.6.2 Manually Adding Instance Nodes

This section describes how to add nodes to an instance to suit your service requirements.

## Usage Notes

- Adding nodes may lead to the decrease of operations per second (OPS). Perform this operation during off-peak hours.
- You can only add nodes when the instance status is **Available** or **Checking restoration**.
- Instances that one or more nodes are added to cannot be deleted.
- You can also delete nodes by following **Manually Deleting Instance Nodes**.
- Currently, a maximum of 60 nodes are supported. To add more, choose **Service Tickets > Create Service Ticket** in the upper right corner of the console and contact the customer service.

## Method 1

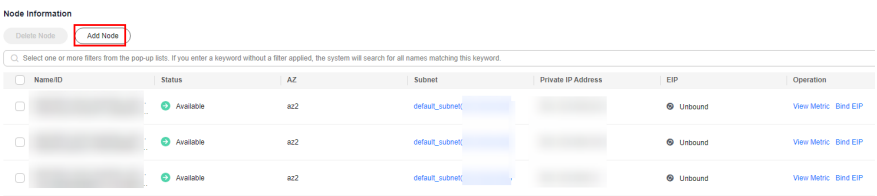
**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, locate the instance that you want to add nodes to and click its name.

**Step 4** In the **Node Information** area on the **Basic Information** page, click **Add Node**.

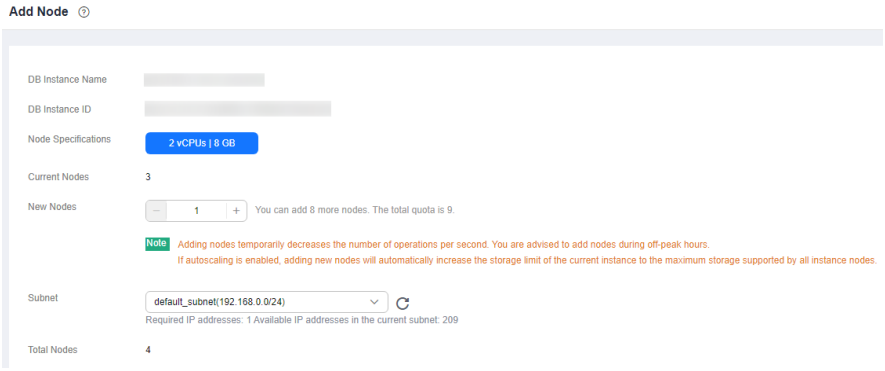
**Figure 4-44** Node information



**Step 5** Specify **Add Nodes** and click **Next**.

**Figure 4-45** Adding nodes




**NOTE**

- New nodes are of the same specifications as existing nodes. Once a new node is added, its specifications cannot be changed.
- New nodes and the instance can be in different subnets of the same VPC.

**Step 6** On the displayed page, confirm the node configurations.

- Yearly/Monthly
  - If you need to modify your settings, click **Previous**.
  - If you do not need to modify your settings, click **Next** and complete the payment.
- Pay-per-use
  - If you need to modify your settings, click **Previous**.
  - If you do not need to modify your settings, click **Submit**.

**Step 7** View the results.

- When new nodes are being added, the instance status is **Adding node**.
- After the nodes are added, the instance status becomes **Available**.
- Click the instance name. In the **Node Information** area on the **Basic Information** page, view information about the new nodes.

**----End**

## Method 2

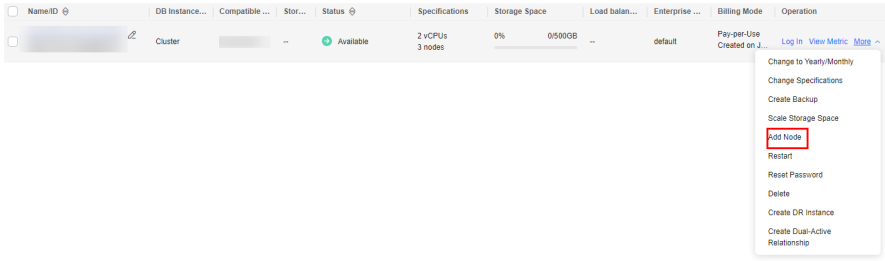
**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, locate the target instance and choose **More** > **Add Node** in the **Operation** column.

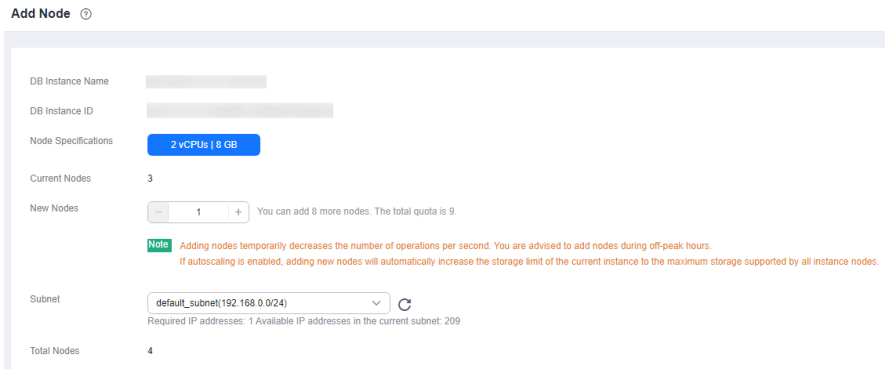
  

**Figure 4-46** Adding nodes




**Step 4** Specify **Add Nodes** and click **Next**.

**Figure 4-47** Adding nodes




**NOTE**

- New nodes are of the same specifications as existing nodes. Once a new node is added, its specifications cannot be changed.
- New nodes and the instance can be in different subnets of the same VPC.

**Step 5** On the displayed page, confirm the node configurations.

- Yearly/Monthly
  - If you need to modify your settings, click **Previous**.
  - If you do not need to modify your settings, click **Next** and complete the payment.
- Pay-per-use
  - If you need to modify your settings, click **Previous**.
  - If you do not need to modify your settings, click **Submit**.

**Step 6** View the results.

- When new nodes are being added, the instance status is **Adding node**.
- After the nodes are added, the instance status becomes **Available**.
- Click the instance name. In the **Node Information** area on the **Basic Information** page, view information about the new nodes.

**----End**

### 4.6.6.3 Automatically Adding Instance Nodes

When an autoscaling threshold is met, GeminiDB DynamoDB-Compatible instance nodes can be automatically added to reduce I/O pressure.

NOTE

- If you enable **Auto Scale** using a Huawei Cloud account, no additional configuration is required.
- If you enable **Auto Scale** as an IAM user first time, you need to obtain the permission to create an agency.

## Configuring Permissions

If you are using an IAM user, perform the following operations to configure GeminiDB and IAM permissions before you enable storage autoscaling:

1. Configure the GeminiDB FullAccess permission.
2. Configure fine-grained permissions for IAM.

   For details about how to configure IAM permissions, see **Creating a Custom Policy**.

   Custom policy in JSON format:

```
{
    "Version":"1.1",
    "Statement":[
        {
            "Effect":"Allow",
            "Action":[
                "iam:permissions:listRolesForAgencyOnProject",
                "iam:permissions:grantRoleToGroupOnProject",
                "iam:agencies:createAgency",
                "iam:agencies:listAgencies",
                "iam:roles:listRoles",
                "iam:roles:createRole"
            ]
        }
    ]
}
```

3. **Create a user group and assign permissions**.

   You can create a user group on the IAM console and grant it custom permissions created in **2** and the security administrator role.

4. **Create an IAM user** and add it to a user group.

   Log in to the IAM console using a Huawei Cloud account or as an IAM user, locate the IAM user that the target instance belongs to, and add it to the user group created in **3**. The IAM user will inherit permissions of the user group.

## Usage Notes

- This function is now in OBT. To use it, choose **Service Tickets > Create Service Ticket** in the upper right corner of the console and contact the customer service.
- If resources in the current region are insufficient, nodes may fail to be added.
- Autoscaling is available only when your account balance is sufficient.
- The instance is in the **Available** status.
- Once autoscaling is enabled, an agency will be created and fees will be automatically deducted.

## Procedure

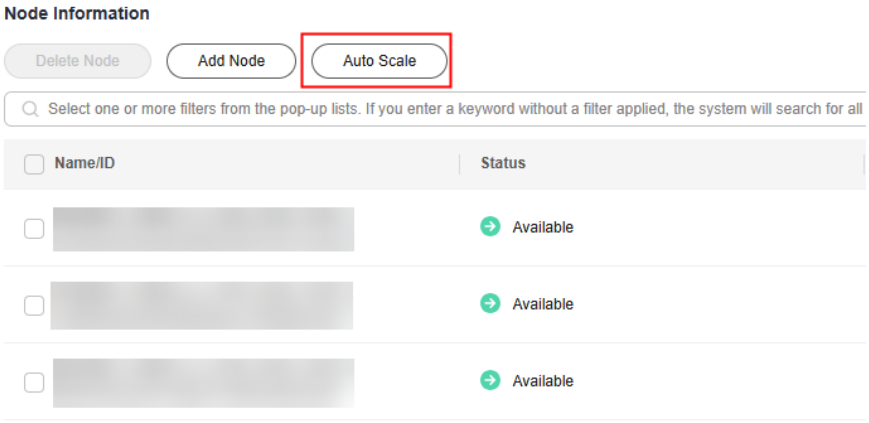
**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, click the target instance. The **Basic Information** page is displayed.

**Step 4** In the **Node Information** area on the **Basic Information** page, click **Auto Scale**.

**Figure 4-48** Auto Scale




**Step 5** Set autoscaling triggers and thresholds listed in **Table 4-13**.

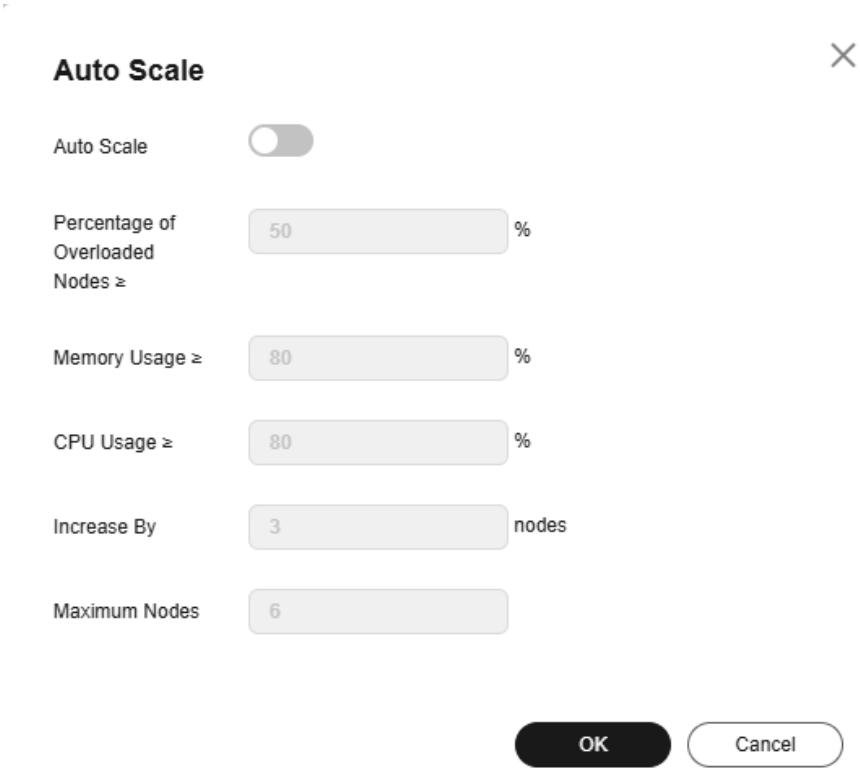


**Table 4-13** Description

| Parameter | Description |
| --- | --- |
| Auto Scale | Whether autoscaling is enabled |
| Percentage of Overloaded Nodes ≥ | Percentage of overloaded nodes that need to be scaled out |
| Memory Usage ≥ | Memory usage of nodes for which autoscaling is triggered |
| CPU Usage ≥ | CPU usage of nodes for which autoscaling is triggered |
| Increase By | Number of nodes to be added each time |
| Maximum Nodes | Maximum number of nodes that can be automatically added |

**----End**

### 4.6.6.4 Manually Deleting Instance Nodes

You can delete nodes that are no longer used to release resources.

## Usage Notes

- Deleted nodes cannot be recovered. Exercise caution when performing this operation.
- Only pay-per-use instances can be deleted.
- Deleting nodes will cause the OPS to decrease for a short period of time. Deleting nodes during off-peak hours.
- If you enable operation protection to improve the security of your account and cloud products, two-factor authentication is required for sensitive operations. For details about how to enable operation protection, see ***Identity and Access Management User Guide***.

## Procedure

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, locate the instance that you want to delete nodes from and click its name.

**Step 4** In the **Node Information** area on the **Basic Information** page, locate the target node and click **Delete** in the **Operation** column.

**Step 5** If you have enabled operation protection, click **Start Verification** in the **Delete Node** dialog box. On the displayed page, click **Send Code**, enter the verification code, and click **Verify**. The page is closed automatically.

**Step 6** In the displayed dialog box, click **Yes**.

- When the node is being deleted, the instance status is **Deleting node**.
- After the node is deleted, the instance status becomes **Available**.

**----End**

# 4.6.7 Scaling Storage Space

## 4.6.7.1 Overview

As more data is added, you may run out of storage. This section describes how to scale up storage space of your instance. As data volumes decrease, you can scale down storage to avoid low database node utilization and resource waste. **Table 4-14** lists the scaling methods supported by GeminiDB DynamoDB-Compatible instances.

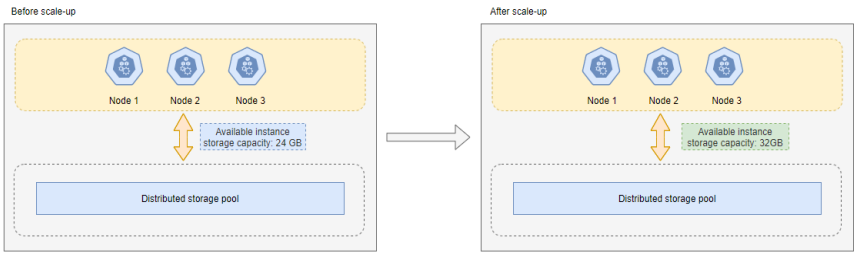
**Table 4-14** Scaling methods

| Method | Description |
| --- | --- |
| **Manually Scaling Up Storage Space** | You can specify how much storage space needs to be added.<br>The added value must be a multiple of 1 (GB). The total storage space cannot exceed the upper limit defined by your instance specifications. |
| **Automatically Scaling Up Storage Space** | If storage usage exceeds the configured threshold, autoscaling will be triggered.<br>The storage is scaled up by a percentage you specify. The added storage space is the current storage space multiplied by the scaling increment. |
| **Manually Scaling Down Storage Space** | You can specify how much storage space needs to be reduced.<br>The storage space to be reduced must be an integer multiple of 1 GB and greater than or equal to 125% of the used storage space. The value is rounded up. |

## Manually Scaling Up Storage Space

For example, if the storage space of a cluster instance is 24 GB and is increased by 8 GB, the storage space will become 32 GB.

**Figure 4-49** Manually scaling up storage space



## Automatically Scaling Up Storage Space

For example, the storage space of a cluster instance is 24 GB before scale-up, the storage usage threshold for triggering autoscaling is set to 80%, and the total storage needs to be automatically scaled up by 10%. If the storage usage is greater than or equal to 80%, the storage space is automatically scaled up by 2.4 GB (24 x 10%), which is rounded up to 3 GB. In this case, the total storage space becomes 27 GB (24 + 3).

**Figure 4-50** Automatically scaling up storage space


## Manually Scaling Down Storage Space

For example, if the storage space of a cluster instance is 32 GB and is decreased by 8 GB, the storage space will become 24 GB.

**Figure 4-51** Manually scaling down storage space


## 4.6.7.2 Manually Scaling Up Storage Space

This section describes how to scale up storage of an instance to suit your service requirements.

## Usage Notes

- **Scaling up storage does not interrupt your services. After storage scale-up is complete, you do not need to restart your instance.**
- If your yearly/monthly instance is running out of storage, additional usage will be billed on a pay-per-use basis. To avoid these extra costs and maintain the benefits of your yearly/monthly subscription, you are advised to scale up storage.

## Procedure

**Step 1** Log in to the Huawei Cloud console.

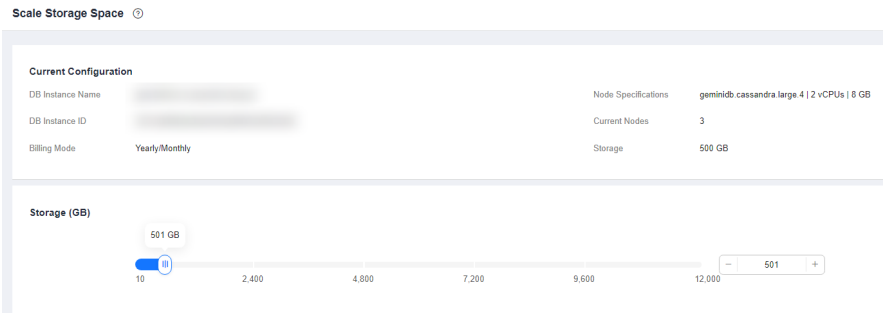
**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, locate the target instance and choose **More** > **Scale Storage Space** in the **Operation** column.

Click the instance name. In the **Storage Space** area on the **Basic Information** page, click **Scale**.

**Step 4** On the displayed page, specify new storage and click **Next**.

**Figure 4-52** Scaling up storage space




- To scale up classic storage, you need to add at least 1 GB each time. The value must be an integer.
- To scale up cloud native storage, you need to add at least 10 GB each time. The value must be an integer multiple of 10.

**Step 5** On the displayed page, confirm the storage space.

- Yearly/Monthly
  - If you need to modify your settings, click **Previous**.
  - If you do not need to modify your settings, click **Next** and complete the payment.
- Pay-per-use
  - If you need to modify your settings, click **Previous**.
  - If you do not need to modify your settings, click **Submit**.

**Step 6** Check the results.

- When the scaling task is ongoing, the instance status is **Scaling storage space**.
- After the scaling process, the instance status becomes **Available**.
- Click the instance name. In the **Storage Space** area on the **Basic Information** page, check the new storage space.

**----End**

## 4.6.7.3 Automatically Scaling Up Storage Space

You can enable storage autoscaling for GeminiDB DynamoDB-Compatible instances. When storage usage reaches the limit, autoscaling is triggered.

You can enable **Auto Scale**:

1. When you create an instance. For details, see **Buying a GeminiDB DynamoDB-Compatible Instance**.
2. After you create an instance

This section describes how to configure **Auto Scale** after an instance is created.

NOTE

- If you enable **Auto Scale** using a Huawei Cloud account, no additional configuration is required.
- If you enable **Auto Scale** as an IAM user first time, you need to obtain the permission to create an agency.

## Configuring Permissions

If you are using an IAM user, perform the following operations to configure GeminiDB and IAM permissions before you enable storage autoscaling:

1. Configure the GeminiDB FullAccess permission.
2. Configure fine-grained permissions for IAM.

   For details about how to configure IAM permissions, see **Creating a Custom Policy**.

   If you use the JSON view to configure a custom policy, the policy content is as follows:

```
{
    "Version":"1.1",
    "Statement":[
        {
            "Effect":"Allow",
            "Action":[
                "iam:permissions:listRolesForAgencyOnProject",
                "iam:permissions:grantRoleToGroupOnProject",
                "iam:agencies:createAgency",
                "iam:agencies:listAgencies",
                "iam:roles:listRoles",
                "iam:roles:createRole"
            ]
        }
    ]
}
```

3. **Create a user group and assign permissions**.

   You can create a user group on the IAM console and grant it custom permissions created in 2 and the security administrator role.

4. **Create an IAM user** and add it to a user group.

   Log in to the IAM console using a Huawei Cloud account or as an IAM user, locate the IAM user that the target instance belongs to, and add it to the user group created in 3. The IAM user will inherit permissions of the user group.

## Usage Notes

- Autoscaling is available only when your account balance is sufficient.
- The instance must be in the **Available** status.
- Once **Auto Scale** is enabled, an agency will be created and fees will be automatically deducted.

## Automatically Scaling Up Storage of a Single Instance

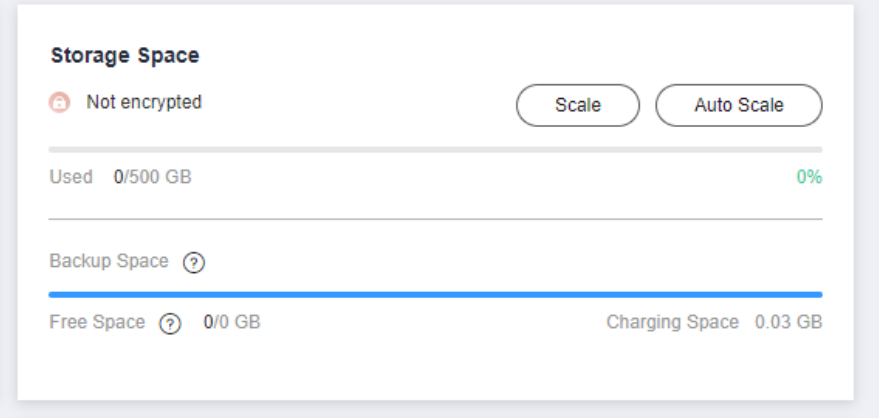
**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, click the target instance. The **Basic Information** page is displayed.

**Step 4** In the **Storage Space** area, click **Auto Scale**.

**Figure 4-53** Auto Scale




**Step 5** Toggle on **Auto Scale** and specify the parameters below.

**Figure 4-54** Configuring autoscaling parameters

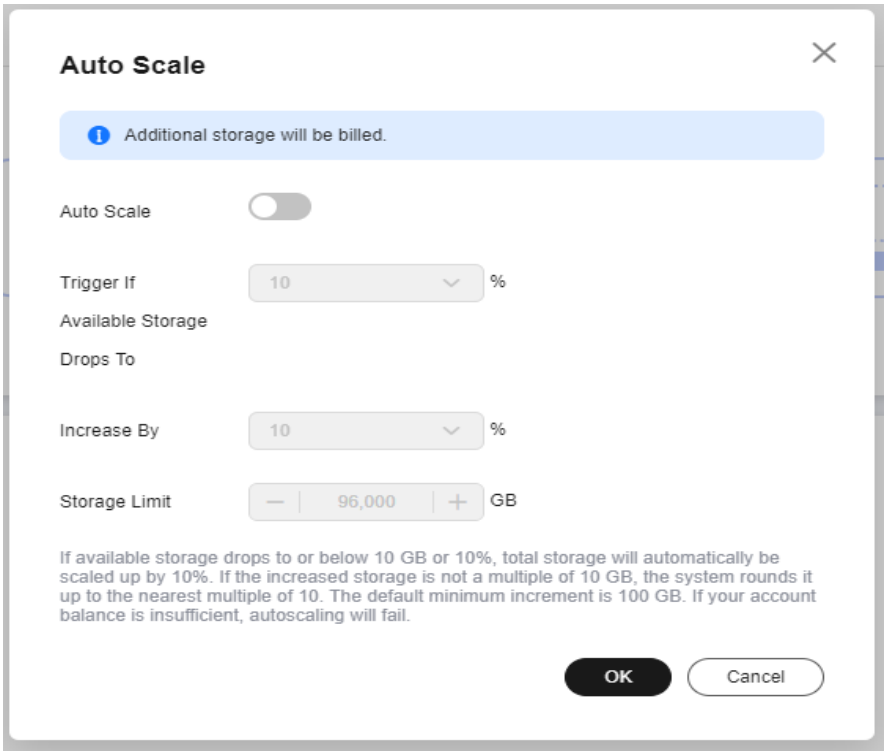


**Table 4-15** Description

| Parameter | Description |
| --- | --- |
| Auto Scale | If you toggle on this switch, autoscaling is enabled. |
| Trigger If Available Storage Drops To | When the available storage usage drops to a specified threshold or the available storage drops to 10 GB, autoscaling is triggered. |
| Increase By | Percentage that your instance storage will be scaled up at. The value can be **10%**, **15%**, or **20%**. If the value is not a multiple of 10, it is rounded up. At least 100 GB is added each time. |
| Storage Limit | Limit of storage (GB) that can be automatically scaled up to.<br>The value must be no less than the storage of your instance and cannot exceed **the maximum storage defined by your instance specifications**. |

**Step 6** Click **OK**.

**----End**

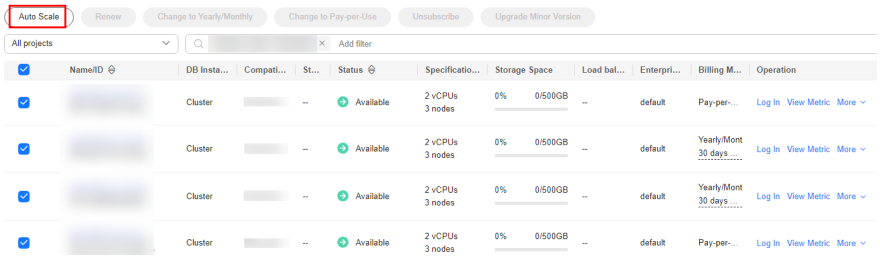
## Automatically Scaling Up Storage of Instances in Batches

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** Select instances and click **Auto Scale**.

**Figure 4-55** Auto Scale




**Step 4** Select an instance, toggle on **Auto Scale**, and specify the parameters below.

**Figure 4-56** Batch Auto Scale

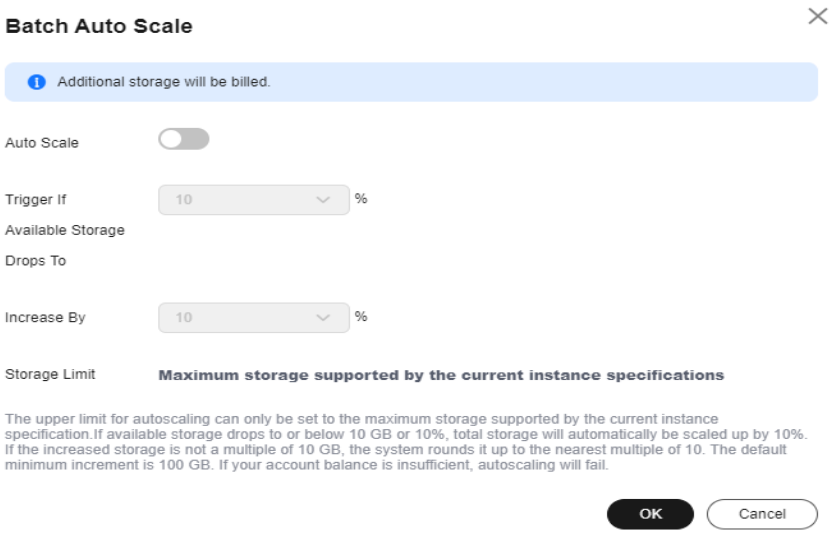



**Table 4-16** Description

| Parameter | Description |
| --- | --- |
| Auto Scale | If you toggle on this switch, autoscaling is enabled. |
| Trigger If Available Storage Drops To | When the available storage usage drops to a specified threshold or the available storage drops to 10 GB, autoscaling is triggered. |
| Increase By | Percentage that your instance storage will be scaled up at. The value can be **10%**, **15%**, or **20%**. If the value is not a multiple of 10, it is rounded up. At least 100 GB is added each time. |
| Storage Limit | The value cannot be specified. By default, the storage is scaled up to the maximum defined by your instance specifications. |

**Step 5** Click **OK**.

**----End**

## 4.6.7.4 Manually Scaling Down Storage Space

As data volumes decrease, you can scale down storage space to avoid low database node utilization and resource waste.

## Usage Notes

- To scale down storage, ensure the new storage space is at least 1.25 times more than the used space and rounded up.

- Scaling down storage does not interrupt your services, and you do not need to restart your instance.
- If your yearly/monthly instance is running out of storage, additional usage will be billed on a pay-per-use basis. To avoid these extra costs and maintain the benefits of your yearly/monthly subscription, you are advised to scale up storage.
- Only classic storage can be scaled in.

## Procedure

**Step 1** Log in to the Huawei Cloud console.

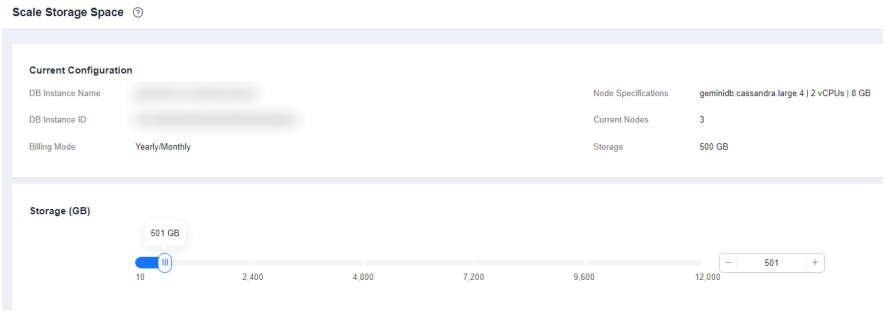
**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, locate the target instance and choose **More** > **Scale Storage Space** in the **Operation** column.

Click the instance name. In the **Storage Space** area on the **Basic Information** page, click **Scale**.

**Step 4** On the displayed page, specify new storage and click **Next**.

**Figure 4-57** Scaling down storage space




Select at least 1 GB each time, and the value must be an integer.

**Step 5** On the displayed page, confirm the storage space.

- Yearly/Monthly
  - To modify your settings, click **Previous** to go back to the page where you specify details.
  - If you do not need to modify your settings, click **Next** and complete the payment.
- Pay-per-use
  - To modify your settings, click **Previous** to go back to the page where you specify details.
  - If you do not need to modify your settings, click **Submit**.

**Step 6** Check the results.

- During the scale-down process, the instance status becomes **Scaling storage space**.
- After the scaling process, the instance status becomes **Available**.

- Click the instance name. In the **Storage Space** area on the **Basic Information** page, check the new storage space.

**----End**

# 4.7 Data Backup

## 4.7.1 Overview

You can back up GeminiDB DynamoDB-Compatible instances to ensure data reliability. After an instance is deleted, the manual backup data is retained. Automated backup data is released together with instances. Backup data cannot be downloaded or exported.

### Usage Notes

Backing up data consumes a few CPUs. Uploading backup files to OBS occupies bandwidth of compute nodes, causing slight latency and jitter.

### Backup Methods

Both automated and manual backups are supported.

- Automated backup

  You can **modify a backup policy** on the GeminiDB console, and the system will automatically back up your instance data based on the time window and backup cycle you configured in the backup policy and will store the data for a length of time you specified.

  Automated backups cannot be manually deleted. You can adjust their retention period by following **Modifying an Automated Backup Policy** . Expired backups will be automatically deleted.

- Manual backup

  A manual backup is a full backup of a DB instance and can be retained until you manually delete it. Manual backup can be triggered at any time to meet your service requirements.

  Regularly backing up your database is recommended. If your database becomes faulty or data is corrupted, you can restore it from backups.

**Table 4-17** Comparison between automated backup and manual backup

| Backup Method | Scenario |
| --- | --- |
| **Managing Automated Backups** | After you set a backup policy, the system automatically backs up your database based on the policy. You can also modify the policy based on service requirements. Either incremental or full backup is supported. |

| Backup Method | Scenario |
| --- | --- |
| Managing Manual Backups | You can enable full backup for your instance based on service requirements. |

Cross-region and table-level backups are supported based on application scenarios.

**Table 4-18** Application scenarios

| Method | Scenario |
| --- | --- |
| Managing Cross-Region Backups | Backups can be stored in the destination region. Then for disaster recovery, you can restore the backups to a new instance in another region. Only an automated full backup is supported. |
| Managing Table-level Backups | If a database or table is deleted maliciously or accidentally, you can use backups to restore data. Manual and automated backups are supported. |

Full and incremental backups are created based on data volumes.

**Table 4-19** Comparison between full and incremental backups

| Backup Type | Full backup | Incremental backup |
| --- | --- | --- |
| Description | All data in an instance is backed up. | Only data that has changed within a certain period is backed up. |
| Enabled by Default | Yes | Yes |

| | | |
|---|---|---|
| **Retention Duration** | • You can specify how many days automated backups can be retained for. If you shorten the retention duration, the new backup policy takes effect for existing backups.<br>• Manual backups are always retained even though a GeminiDB DynamoDB-Compatible instance is deleted. They can only be deleted manually. | Incremental backups will be deleted along with automated full backups. |
| **Feature** | • All data of your instance is backed up in the current point of time.<br>• You can use a full backup to restore all data generated when its backup was created.<br>• Full backups can be created automatically or manually. | • Incremental data in your instance is backed up since the last full backup.<br>• When you use an incremental backup for restoration, the last full backup data and the incremental data generated since then are downloaded.<br>• Incremental backups can be created automatically only. |
| **How to View** | Click an instance name. On the **Backups & Restorations** page, click the **Instance-level Backups** and **Table-level Backups** tabs to view the backup size. | Click an instance name. On the **Backups & Restorations** page, click the **Incremental Backup** tab to view the backup size. |

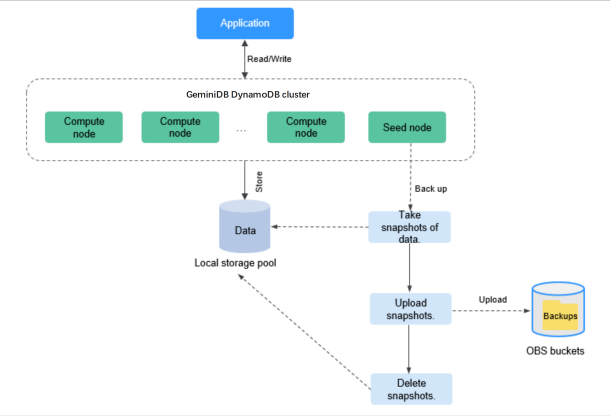
## How Backup Works

A dedicated node (seed node) is used for backup management of GeminiDB DynamoDB-Compatible instances. As shown in the following figure, the seed node backs up data of GeminiDB DynamoDB-Compatible cluster instances. The node

takes snapshots in seconds and then stores them as compressed backups in OBS buckets, without occupying storage of your instance. The CPU usage may increase 5% to 15% because uploading backups consumes CPU resources.

**Figure 4-58** Backup process




## Backup Storage

Storing backup data of GeminiDB DynamoDB-Compatible instances in Object Storage Service (OBS) improves the DR capability and reduces storage usage.

After you buy instance storage, an equivalent amount of backup storage will be provided free of charge. For example, if you buy an instance with 100 GB of storage, you will get additional 100 GB of backup storage at no extra cost. If the backup data size stays within 100 GB, it is stored in OBS at no cost. If it exceeds 100 GB, additional data will be billed based on OBS pricing rules.

# 4.7.2 Managing Automated Backups

You can create automated backups of GeminiDB DynamoDB-Compatible instances to ensure data reliability. If a database or table is deleted, maliciously or accidentally, backups can help recover your data.

## Usage Notes

- Backup files are saved as packages in OBS buckets. Uploading backup files and reading service data both consume bandwidth, so the upload bandwidth of OBS is limited. The upload bandwidth of a single node ranges from 20 MB/s to 70 MB/s.

  You need to specify appropriate nodes based on the backup bandwidth limit to achieve better performance.
- The CPU usage may increase 5% to 15% because uploading backups consumes CPU resources.

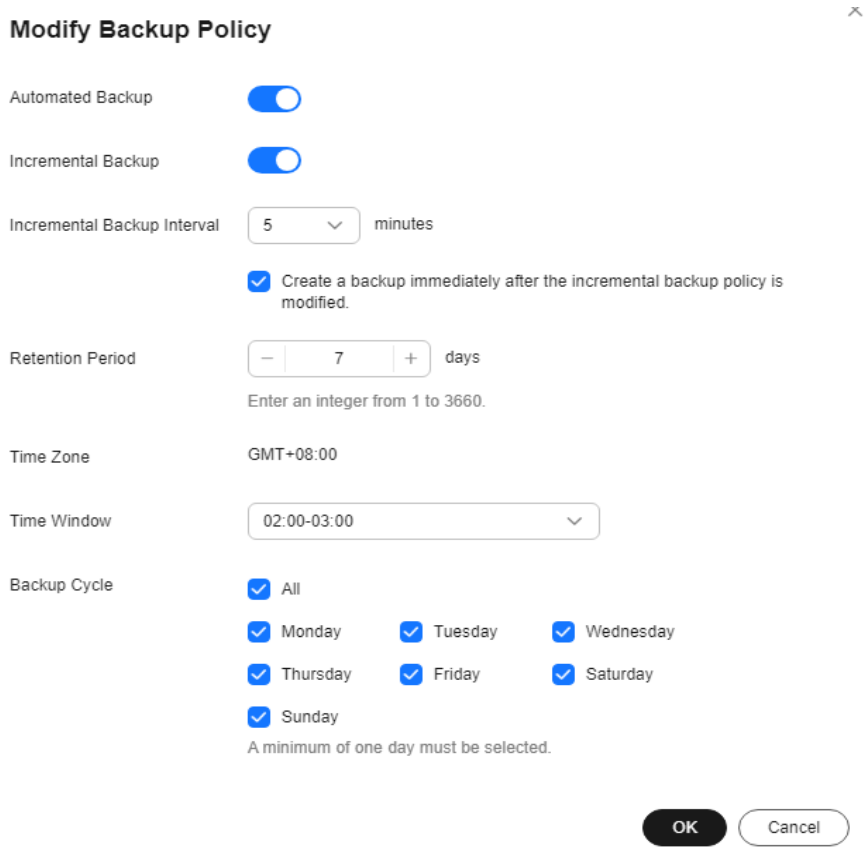
- While backups are uploaded, the memory usage does not increase significantly. Generally, the memory usage is about 300 MB. How much memory usage actually increases depends on the instance data volume. The increased memory mainly caches data during backup upload and read. After the backups are uploaded, the memory usage becomes normal.
- You can manually modify incremental backups of a GeminiDB DynamoDB-Compatible instance.
- To enable the incremental backup function, choose **Service Tickets > Create Service Ticket** in the upper right corner of the console and contact the customer service.
- After the incremental backup function is enabled, differential backup is selected by default. To enable PITR, choose **Service Tickets > Create Service Ticket** in the upper right corner of the console and contact the customer service.

## Configuring an Automated Backup Policy

Automated backups are generated according to a backup policy and saved as packages in OBS buckets to ensure data confidentiality and durability. You are advised to regularly back up your database, in case it becomes faulty or damaged. Backing up data affects the database read and write performance so you are advised to set the automated backup time window to off-peak hours.

When you create an instance, automated backup is enabled by default.

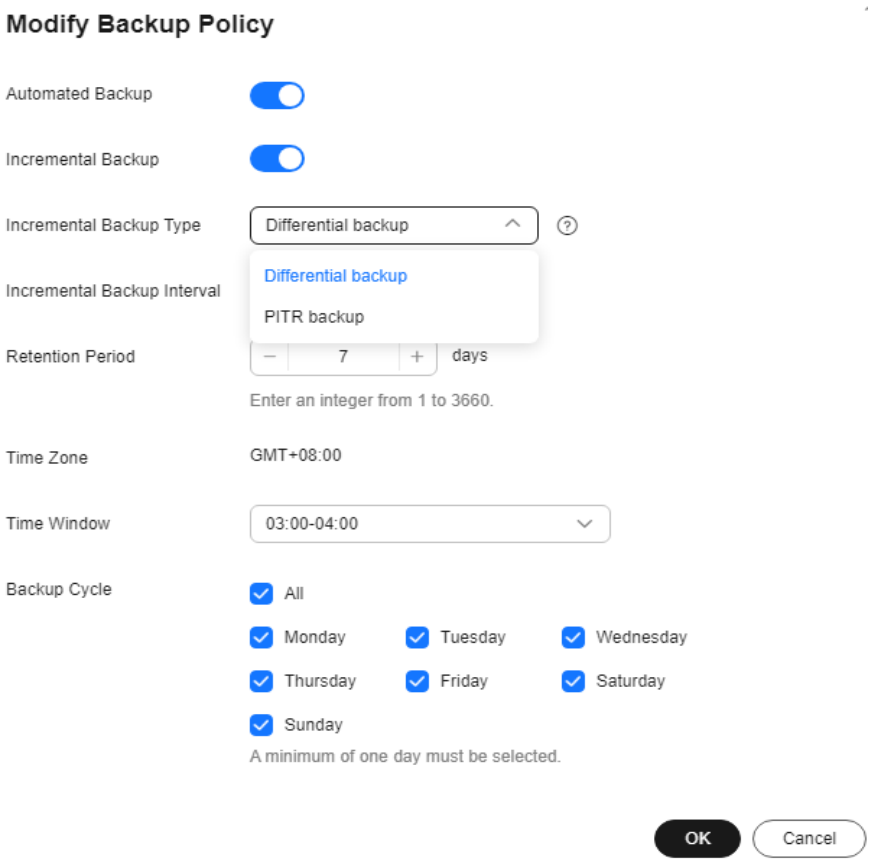
**Figure 4-59** Enabling the automated backup policy



- **Incremental Backup** is enabled by default. You can click  to manually enable or disable it. After it is enabled, the system stores backup data in OBS. Select an incremental backup type. **Differential backup** is selected by default.
  - **Differential backup**: Data can be restored to a specified point in time.
  - **PITR backup**: Data can be restored to any point in time.

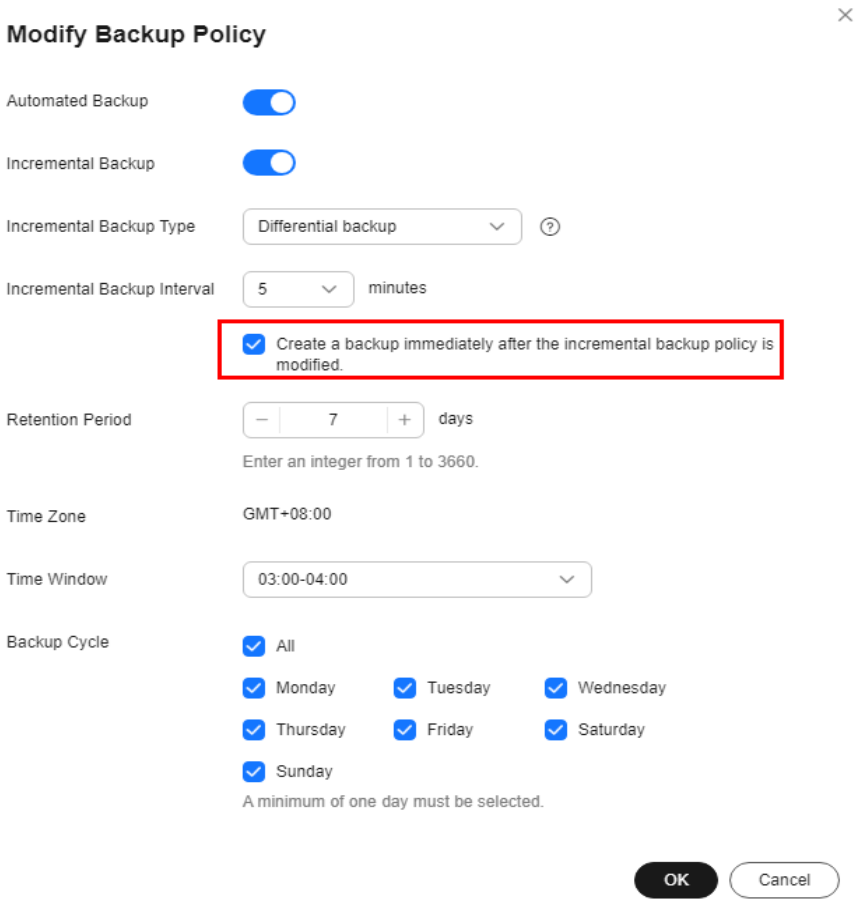
  **Figure 4-60** Selecting an incremental backup type

  Enabling **Incremental Backup** will take effect in the next full backup. You are advised to select **Create a backup immediately after the incremental backup policy is modified**.
  - If you select it, the full backup request is delivered immediately, and the incremental backup takes effect.
  - If you do not select it, the incremental backup will take effect in the next full backup.

**Figure 4-61** Selecting **Create a backup immediately after the incremental backup policy is modified**




- **Incremental Backup Interval**: Incremental backups are generated every 15 minutes.
- **Retention Period**: Automated backup files are saved for seven days by default. The retention period ranges from 1 to 3660 days. Full backups are retained till the retention period ends. However, even if the retention period has ended, the most recent backup will be retained.
  - Extending the retention period improves data reliability. You can extend the retention period as needed.
  - If you shorten the retention period, the new backup policy takes effect for existing backups. Any automated backups (including full and incremental backups) that have expired will be automatically deleted. Manual backups will not be automatically deleted but you can delete them manually.

  NOTE

  - If the retention period is shorter than seven days, the system automatically backs up data daily.
  - The system checks existing automated backups and deletes any backups that exceed the backup retention period you configured.
  - **Time Window**: A one-hour period the backup will be scheduled within 24 hours, such as 00:00–01:00. The backup time is displayed in GMT. After the DST or standard time is switched, the backup time segment changes with the time zone.

If **Retention Period** is set to **2**, full and incremental backups that have been stored for more than two days will be automatically deleted. For instance, a backup generated on Monday will be deleted on Wednesday; or a backup generated on Tuesday will be deleted on Thursday.

**Policy for automatically deleting full backups:**

To ensure data integrity, even after the retention period ends, the most recent backup will be retained:

If **Backup Cycle** was set to **Monday** and **Tuesday** and the **Retention Period** was set to **2**:

- The full backup generated on Monday will be automatically deleted on Thursday. The reasons are as follows:

  The backup generated on Monday expires on Wednesday, but it is the last backup, so it will be retained until a new backup expires. The next backup will be generated on Tuesday and will expire on Thursday. So the full backup generated on Monday will not be automatically deleted until Thursday.

- The full backup generated on Tuesday will be automatically deleted on Wednesday of the following week. The reasons are as follows:

  The backup generated on Tuesday will expire on Thursday, but as it is the last backup, so it will be retained until a new backup expires. The next backup will be generated on the following Monday and will expire on the following Wednesday. So the full backup generated on Tuesday will not be automatically deleted until the following Wednesday.

- **Backup Cycle**: All options are selected by default.
  - **All**: Each day of the week is selected. The system automatically backs up data every day.
  - You can select one or more days in a week. The system automatically backs up data at the specified time.

  NOTE

  A full backup starts within one hour of the time you specify. The amount of time required for the backup depends on the amount of data to be backed up. The more data has to be backed up, the longer it will take.

- After an instance is created, you can set an automated backup policy. The system will back up data based on the automated backup policy.
- If **Automated Backup** is disabled, any automated backups in progress stop immediately.

## Modifying an Automated Backup Policy

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

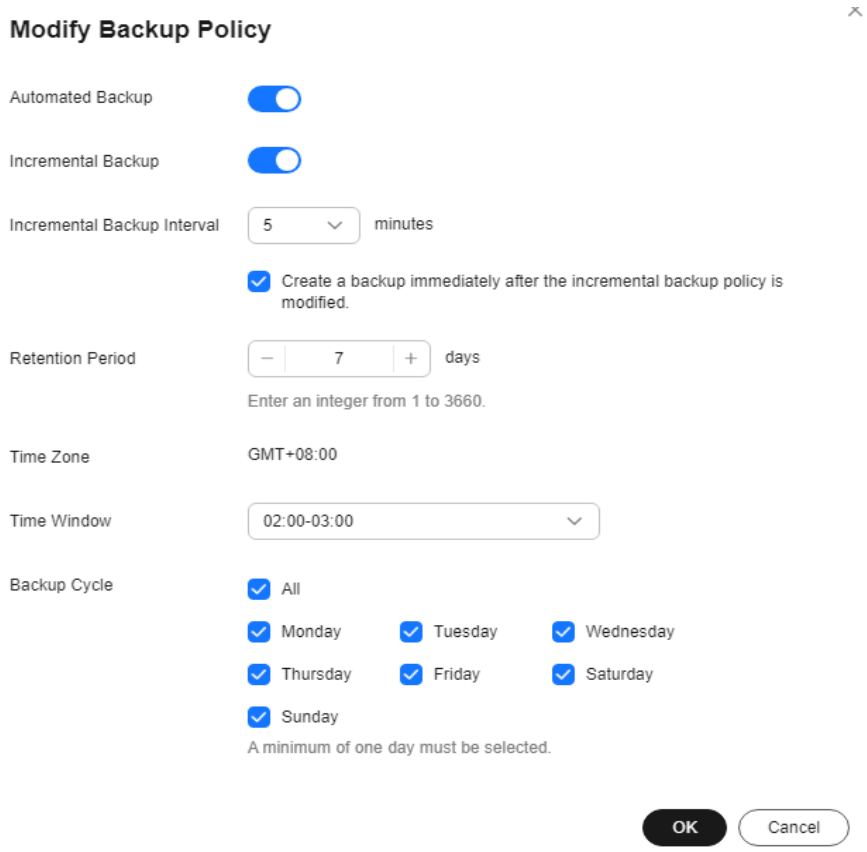
**Step 3** On the **Instances** page, click the instance you want to back up.

**Step 4** Choose **Backups & Restorations** in the navigation pane one the left, and click **Modify Backup Policy**. In the displayed dialog box, configure the backup policy. Click **OK**.

For details about how to set a backup policy, see **Configuring an Automated Backup Policy**.

**Figure 4-62** Modifying the backup policy




**Step 5** Check or manage the generated backups on the **Backups** or **Backups & Restorations** page.

**----End**

## Viewing Incremental Backups

You can view incremental backups and their size of an instance.

- To view the size and records of incremental backups, choose **Service Tickets** > **Create Service Ticket** in the upper right corner of the console and contact the customer service.
- You can view incremental backups and their size only after you enable **Incremental Backup**, or no data is displayed.

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

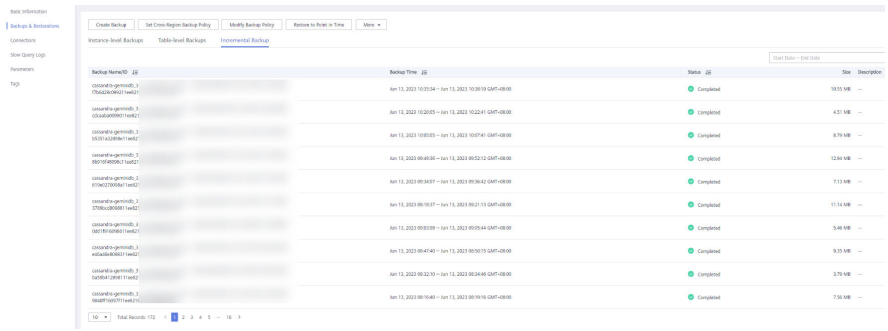
**Step 3** On the **Instances** page, locate the instance that you want to connect to and click its name.

**Step 4** In the navigation pane on the left, choose **Backups & Restorations**.

**Step 5** On the **Backups & Restorations** page, click **Incremental Backup**.

**Figure 4-63** Incremental backup



**Step 6** View incremental backups and their size.

**----End**

## Disabling Incremental Backup

**Step 1** Log in to the Huawei Cloud console.

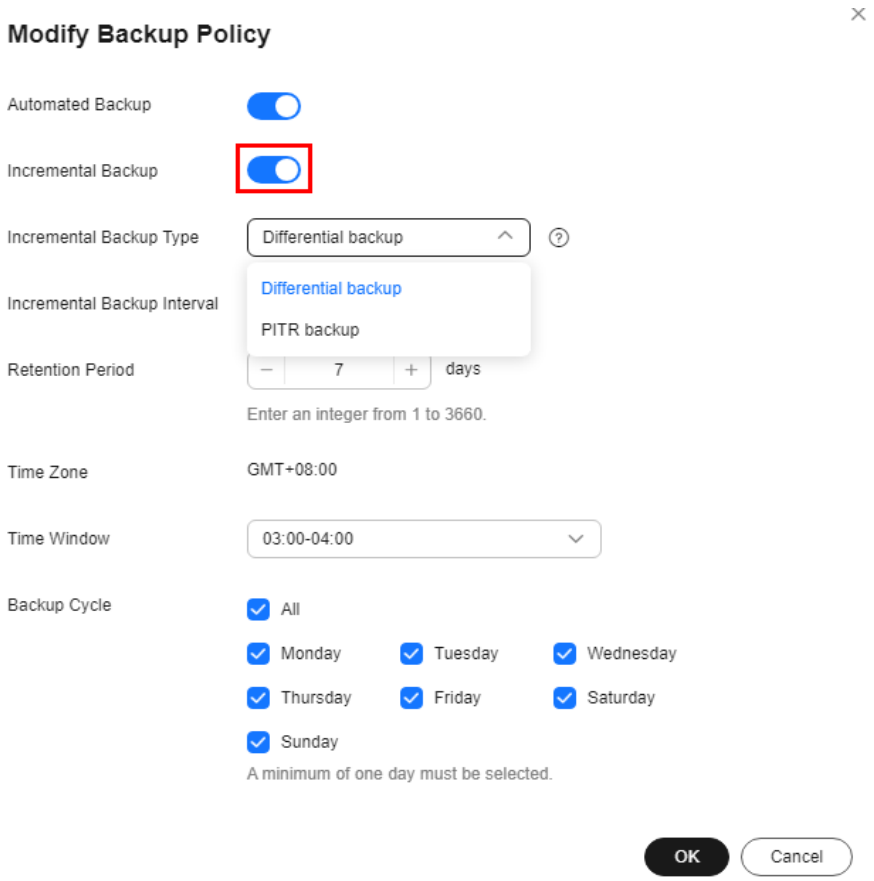
**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, click the instance you want to back up.

**Step 4** In the navigation pane, choose **Backups & Restorations**. On the displayed page, click **Modify Backup Policy** and click  next to **Incremental Backup**.

**Figure 4-64** Disabling Incremental Backup




**----End**

## Disabling Automated Backup

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

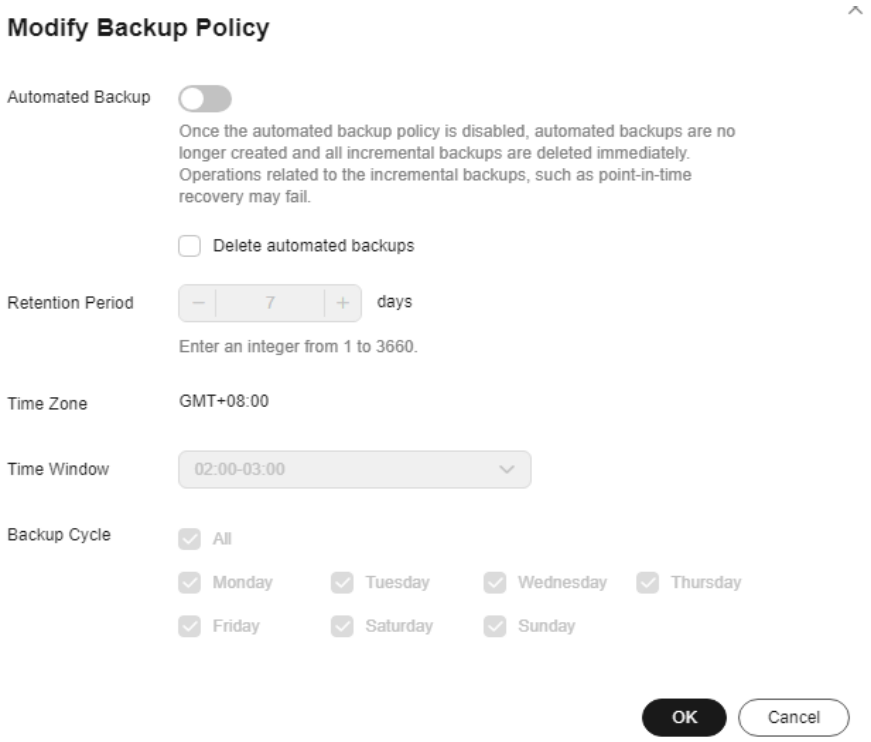
**Step 3** On the **Instances** page, click the instance you want to back up.

**Step 4** Choose **Backups & Restorations** in the navigation pane one the left, and click **Modify Backup Policy**.

**Step 5** In the displayed dialog box, click  and click **OK**.

**Figure 4-65** Disabling backup policies




When your disable **Automated Backup**, specify whether to delete the automated backups:

- If you select **Delete automated backups**, all backup files within the retention period will be deleted. There are no automated backups displayed until you enable automated backup again.
- If you do not select **Delete automated backups**, backup files within the retention period will be retained, but you can still manually delete them later if needed. For details, see **Deleting an Automated Backup**.

If **Automated Backup** is disabled, any automated backups in progress stop immediately.

**----End**

## Deleting an Automated Backup

If automated backup is disabled, you can delete stored automated backups to free up storage space.

If automated backup is enabled, the system will delete automated backups when they expire. You cannot delete them manually.

**CAUTION**

Deleted backups cannot be restored.

- **Method 1**
  a. Log in to the Huawei Cloud console.
  b. In the service list, choose **Databases** > **GeminiDB**.
  c. On the **Instances** page, click the instance you want to back up.
  d. Choose **Backups & Restorations** in the navigation pane on the left, locate the backup you want to delete, and click **Delete** in the **Operation** column.
  e. In the displayed dialog box, confirm the backup details and click **Yes**.
- **Method 2**
  a. Log in to the Huawei Cloud console.
  b. In the service list, choose **Databases** > **GeminiDB**.
  c. On the **Backups** page, locate the backup that you want to delete and click **Delete**.
  d. In the displayed dialog box, confirm the backup details and click **Yes**.

## 4.7.3 Managing Manual Backups

You can create manual backups for GeminiDB DynamoDB-Compatible instances in the **Available** status to ensure data reliability. If a database or table is deleted, maliciously or accidentally, backups can help recover your data.

### Usage Notes

- Manual backups are full backups.
- Backup files are saved as packages in OBS buckets. Uploading backup files and reading service data both consume bandwidth, so the upload bandwidth of OBS is limited. The upload bandwidth of a single node ranges from 20 MB/s to 70 MB/s.

  For better performance, you need to specify appropriate nodes for an instance and take into account the bandwidth for uploading backups.
- The CPU usage may increase 5% to 15% because uploading backups consumes CPU resources.
- While backups are uploaded, the memory usage does not increase significantly. Generally, the memory usage is about 300 MB. How much memory usage actually increases depends on the instance data volume. The increased memory mainly caches data during backup upload and service read. After the backup upload is complete, the memory recovers.
- Manual backups are charged for instances with cloud native storage during OBT.

### Creating a Manual Backup

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** Create a manual backup.

**Method 1**

On the **Instances** page, locate the instance that you want to create a backup for and choose **More** > **Create Backup** in the **Operation** column.

**Method 2**

1. On the **Instances** page, click the instance that you want to create a backup for and click its name.
2. Choose **Backups & Restorations** in the navigation pane on the left, click **Create Backup**.

**Method 3**

In the navigation pane on the left, choose **Backups** and click **Create Backup**.

**Step 4** In the displayed dialog box, enter a backup name and description and click **OK**.

**Figure 4-66** Creating a backup


**Table 4-20** Parameter description

| Parameter | Description |
| --- | --- |
| DB Instance Name | Must be the name of the DB instance to be backed up and cannot be modified. |
| Backup Name | Must be 4 to 64 characters in length and start with a letter. It is case-insensitive and contains only letters, digits, hyphens (-), and underscores (_). |
| Description | Can contain a maximum of 256 characters. Line breaks and special characters >!<"&'= are not allowed. |

**Step 5** View the backup status.

- When the backup is being created, query the backup status on the **Backups** or **Backups & Restorations** page. The backup status is **Backing up**.

- After the backup is created, the backup status is **Completed**.

You can create manual table-level backups by following **Creating and Managing Table-level Backups**.

**----End**

## Deleting a Manual Backup

If you no longer need a manual backup, delete it on the **Backups** or **Backups & Restorations** page.

Deleted backups are not displayed in the backup list.

NOTE

Deleted backups cannot be restored.

Method 1

1. Log in to the Huawei Cloud console.
2. In the service list, choose **Databases** > **GeminiDB**.
3. On the **Instances** page, locate the instance whose backup you want to delete and click its name.
4. Choose **Backups & Restorations** in the navigation pane on the left, locate the backup you want to delete and click **Delete** in the **Operation** column.
5. In the displayed dialog box, confirm the backup details and click **Yes**.

Method 2

1. Log in to the Huawei Cloud console.
2. In the service list, choose **Databases** > **GeminiDB**.
3. On the **Backups** page, locate the backup you want to delete and click **Delete** in the **Operation** column.
4. In the displayed dialog box, confirm the backup details and click **Yes**.

# 4.7.4 Managing Cross-Region Backups

GeminiDB DynamoDB-Compatible instance backups can be stored in the destination region or OBS buckets, so you can use the backups to restore data to a new instance.

After a cross-region backup policy is set for an instance, the system will synchronize backups of the instance to the destination region you specified. You can manage cross-region backup files on the **Backups** page.

## Usage Notes

- To enable the cross-region backup function, choose **Service Tickets > Create Service Ticket** in the upper right corner of the console and contact the customer service.
- Before you configure a cross-region backup policy, make sure to enable automated backup first. Otherwise, the cross-region backup policy cannot take effect. For details, see **Modifying an Automated Backup Policy** .

- Only automated full backups can be created across regions.

## Billing

**Table 4-21** Billing

| Flavor | Billing Item | Unit Price |
| --- | --- | --- |
| geminidb.cassandra.crossreg.backup.space.dfv | Storage space | CNY0.0009/GB/hour |
| geminidb.cassandra.crossreg.backup.flow | Cross-region backup traffic | CNY0.5/GB |

## Setting or Modifying a Cross-Region Backup Policy

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, locate the instance that you want to connect to and click its name.

**Step 4** In the navigation pane on the left, choose **Backups & Restorations**.

**Step 5** On the displayed page, click **Set Cross-Region Backup Policy**.

**Step 6** In the displayed dialog box, set required parameters.

**Figure 4-67** Setting a cross-region backup policy

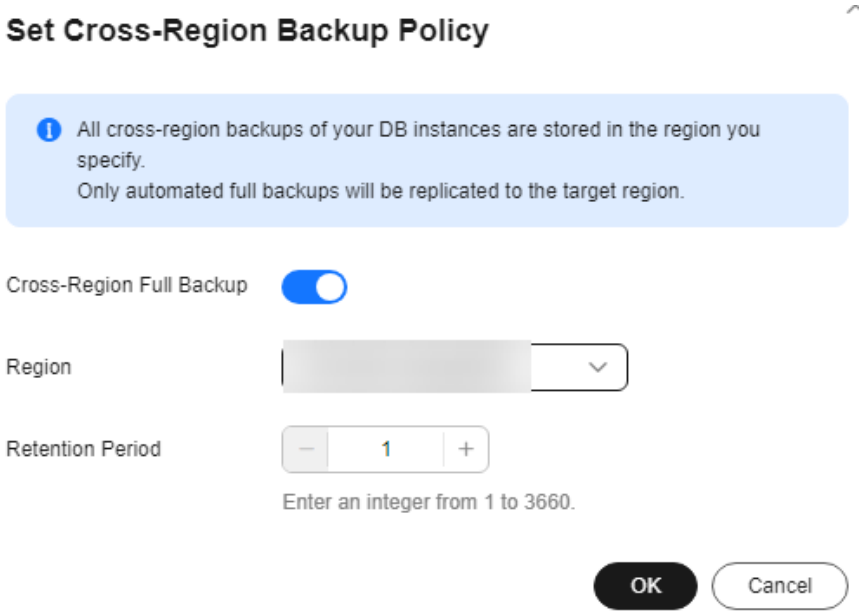


**Table 4-22** Description

| Parameter | Description |
| --- | --- |
| Cross-Region Full Backup | If you enable **Cross-Region Full Backup**, automated full backup files of the instance will be stored in the region you specify. |
| Region | You can select the region for storing backups based on service requirements. |
| Retention Period | Number of days that cross-region backups are kept. The value ranges from **1** to **1825**. You can increase the retention period to improve data reliability. |

NOTE

- Only new backups generated after you set a cross-region backup policy will be stored in the region you specify.
- All cross-region backups of your DB instances are stored in the same region you specify.
- Cross-region backups are synchronized to the destination region you specify only after your instance is backed up locally.
- Only automated full backups are replicated to the destination region.

**Step 7** Click **OK**.

**----End**

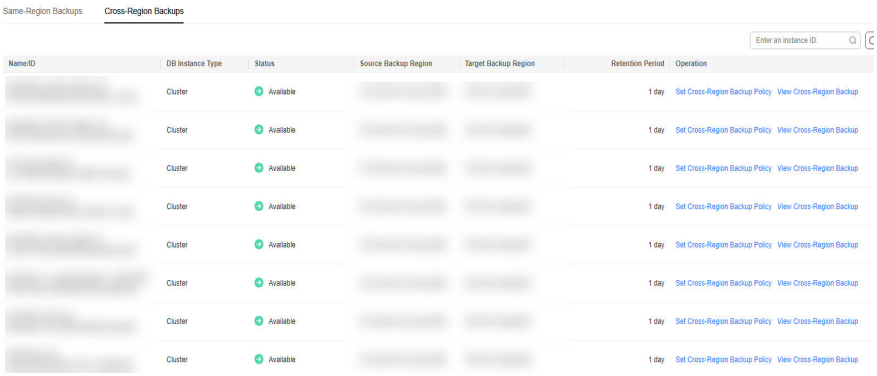
## Managing Cross-Region Backups

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** In the navigation pane, choose **Backups**. Click the **Cross-Region Backups** tab.

**Figure 4-68** Cross-region backups



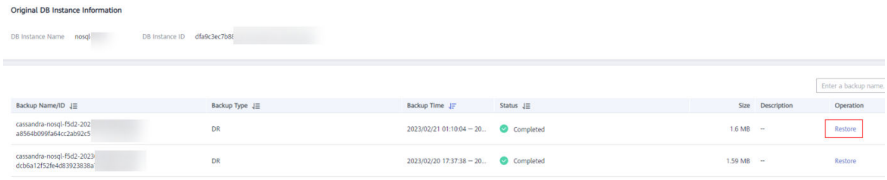

All cross-region backups are displayed by default.

- To modify the cross-region backup policy, click **Set Cross-Region Backup** in the **Operation** column.

- To view all cross-region backups, click **View Cross-Region Backup**. To restore a backup to a new instance, click **Restore** in the **Operation** column. For details, see **Restoring a Backup to a New Instance**.

  **Figure 4-69** Restoring a cross-region backup

  NOTE

  - Cross-region full backup replicates only automated full backups to another region for full restoration. Any of the full backups can be restored to a new instance that has no relationships with the original instance.
  - The new instance uses the same parameter group as the original instance.
  - During the instance restoration, backup files are downloaded from OBS buckets to the data directory of the new instance. The download bandwidth of OBS is 40 MB/s.
  - If the original instance has cross-region backup disabled, the restoration may fail.
  - Cross-region backup is not supported for instances that has disk encryption enabled.

**----End**

## Disabling Cross-Region Backup

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, locate the instance that you want to connect to and click its name.

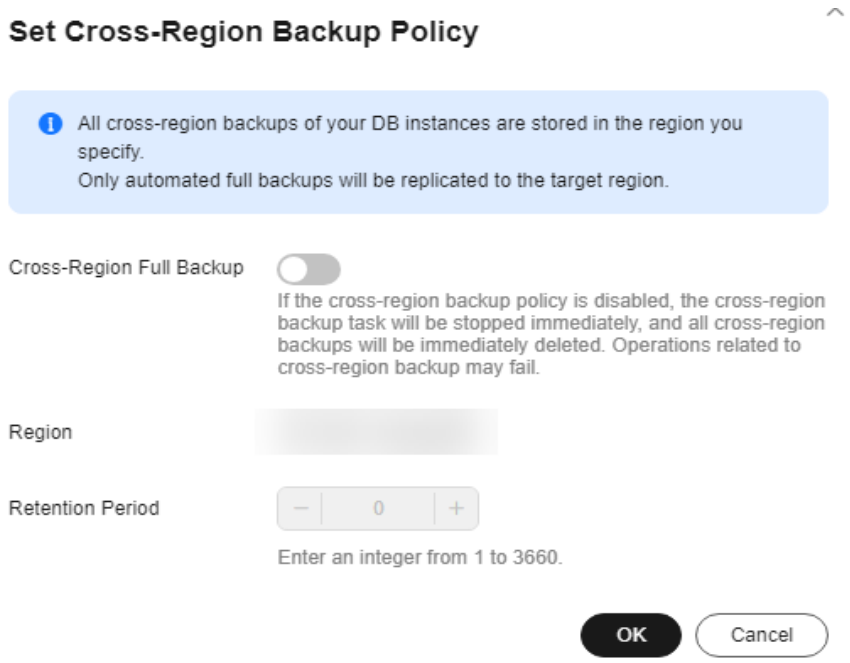
**Step 4** In the navigation pane on the left, choose **Backups & Restorations**.

**Step 5** On the displayed page, click **Set Cross-Region Backup Policy**.

**Step 6** In the displayed dialog box, disable **Cross-Region Full Backup**.

**Figure 4-70** Disabling cross-region backup




NOTE

- After cross-region backup is disabled, the cross-region backup task is stopped and all cross-region backups are deleted immediately. As a result, operations using cross-region backups will fail.
- If an instance with cross-region backup enabled is deleted, its cross-region backups will be retained. The retention period depends on settings of the cross-region backup policy.

**Step 7** Click **OK**.

**----End**

## 4.7.5 Managing Table-level Backups

You can manually create table-level backups for GeminiDB DynamoDB-Compatible instances. If a database or table is deleted maliciously or accidentally, you can use backups to restore data.

### Usage Notes

- To enable the table-level backup, choose **Service Tickets > Create Service Ticket** in the upper right corner of the console and contact the customer service.
- Table-level backups can be created automatically or manually.

### Enabling or Modifying a Table-level Backup Policy

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, click the target instance name to access the **Basic Information** page.

**Step 4** In the navigation pane, choose **Backups & Restorations**.

**Step 5** On the displayed page, click **Configure Table-level Backup Policy**.

**Step 6** On the displayed page, configure related parameters.

**Figure 4-71** Setting a table-level backup policy

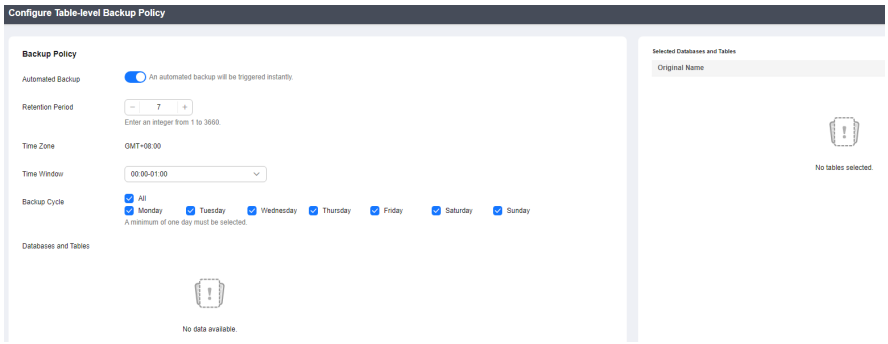



**Table 4-23** Parameters

| Parameter | Description |
| --- | --- |
| Automated Backup | After this function is enabled, database table data is backed up based on the backup policy. |
| Retention Period | Automated backup files are stored for 7 days by default. The retention period ranges from 1 to 732 days.<br>• If the retention period is shorter than seven days, the system automatically backs up data daily.<br>• The system automatically checks existing backup files and deletes files that exceed the retention period you set. |
| Time Window | A one-hour period the backup will be scheduled within 24 hours, such as 01:00-02:00 or 12:00-13:00. The backup time is in GMT format. If the DST or standard time is switched, the time window changes with the time zone. |
| Backup Cycle | • **All**: Each day of the week is selected. This option is selected by default. The system automatically backs up data every day.<br>• You can select one or more days in a week. The system automatically backs up data on the specified days. |
| Databases and Tables | Select tables to be backed up. |

**NOTE**

A full backup starts within one hour of the time you specify. How long the backup takes depends on the data volume.

**Step 7** Click **OK**.

**----End**

## Creating and Managing Table-level Backups

Creating a Manual Table-level Backup

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, locate the instance that you want to create a table-level backup for and click its name.

**Step 4** In the navigation pane on the left, choose **Backups & Restorations**.

**Step 5** On the displayed page, click **Create Table-level Backup**.

**Step 6** On the displayed page, configure related parameters.

**Figure 4-72** Creating a table-level backup

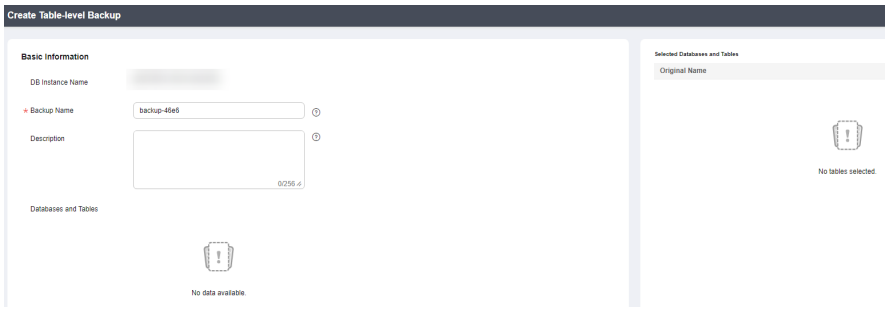



**Table 4-24** Parameters

| Parameter | Description |
| --- | --- |
| Backup Name | Can contain 4 to 64 characters and must start with a letter. The name is case-sensitive and can contain only letters, digits, hyphens (-), or underscores (_). |
| Description | The description can contain a maximum of 256 characters and cannot contain line breaks or special characters >! <''&'= |
| Databases and Tables | You can select the databases and specify tables therein that you want to back up. |

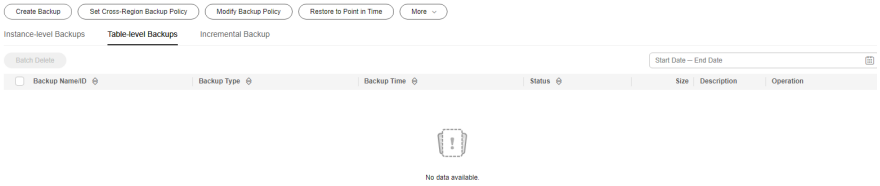
**Step 7** Click **OK**.

**Step 8** Choose **Backups and Restorations** > **Table-level Backup** and manage the created backup.

Alternatively, click **Backups** in the navigation pane on the left, choose **Intra-region Backups** > **Table-level Backups**, and manage the created backup.

**Figure 4-73** Managing the created table-level backup




- Click **View Tables** to view tables contained in the backup file.
- Click **Restore** in the **Operation** column to restore the backup to a new instance. For details, see **Restoring a Backup to a New Instance**.
- Click **Delete** in the **Operation** column to delete the created backup.

  NOTE

  Deleted backups cannot be restored.

**----End**

## Disabling a Table-level Backup Policy

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

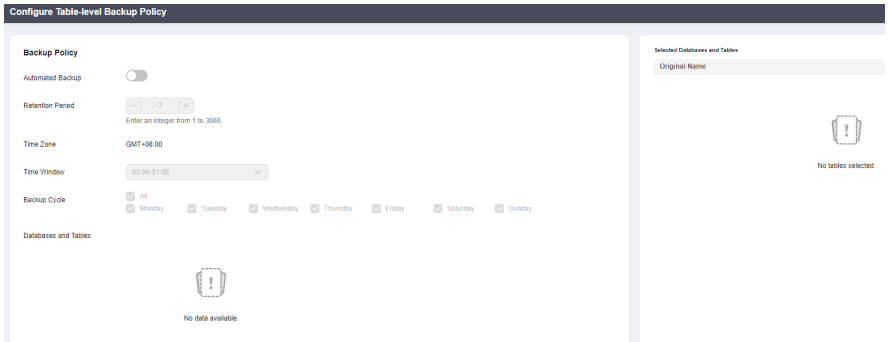
**Step 3** On the **Instances** page, click the target instance name to access the **Basic Information** page.

**Step 4** In the navigation pane, choose **Backups & Restorations**.

**Step 5** On the displayed page, click **Configure Table-level Backup Policy**.

**Step 6** On the displayed page, configure related parameters.

**Figure 4-74** Disabling a table-level backup policy




NOTE

After the table-level backup policy is disabled, any table-level backup task in progress stops immediately, and all table-level backups of the instance are retained. The retention duration depends on **Retention Period** specified when you enabled the table-level backup policy.

**Step 7** Click **OK**.

**----End**

# 4.8 Data Restoration

## 4.8.1 Restoration Methods

Data of GeminiDB DynamoDB-Compatible instances can be restored using different methods.

**Table 4-25** Restoration methods

| Method | Scenario |
| --- | --- |
| **Rebuilding an Instance** | If an instance is deleted by mistake, you can rebuild it within a retention period in the recycle bin. |
| **Restoring a Backup to a New Instance** | You can restore an existing backup file to a new instance. |
| **Restoring a Backup to a Specified Point in Time** | You can use an automated backup to restore an instance to a specified point in time. |

## 4.8.2 Restoring a Backup to a New Instance

You can use existing backups to restore data to a new GeminiDB DynamoDB-Compatible instance.

### Usage Notes

- The new instances must have at least as many nodes as the original instance.
- The new instance must have at least as much storage as the original instance.
- Incremental backup and PITR are not supported.
- Restoration to the current instance is not supported.
- You can scale in the memory, but the memory decrease cannot become less than the actual memory used during the backup.
- The restored instance uses the same parameter group as the original instance.
- During the instance restoration, backups are downloaded from OBS buckets to the data directory of the restored instance. The download bandwidth of OBS is 40 MB/s.

### Procedure

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

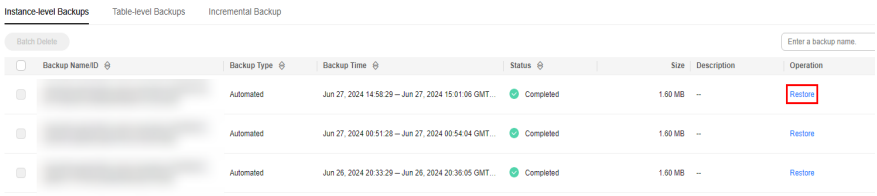
  

**Step 3** Restore an instance from the backup.

Method 1

1. On the **Instances** page, locate the instance whose backup you want to restore and click its name.
2. Choose **Backups & Restorations** in the navigation pane on the left, locate the backup that you want to restore and click **Restore** in the **Operation** column.

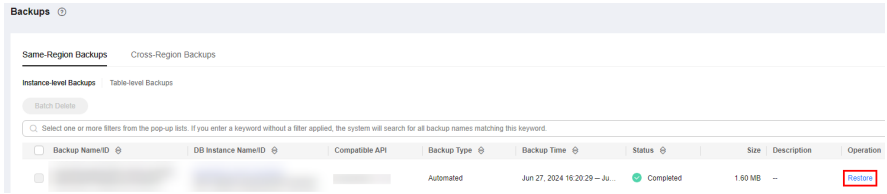
**Figure 4-75** Backups and restorations




Method 2

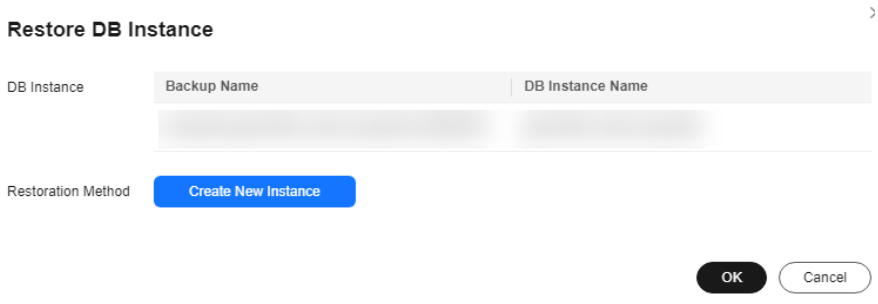
On the **Backups** page, locate the backup that you want to restore and click **Restore** in the **Operation** column.

**Figure 4-76** Backup management




**Step 4** In the displayed dialog box, confirm the current instance details and restoration method and click **OK**.

**Figure 4-77** Restoring data to a new instance




- The default API type and DB engine version are the same as those of the original instance and cannot be changed.
- GeminiDB automatically calculates the minimum storage space required for restoration based on the size of the selected backup file. The storage capacity depends on the instance specifications, and must be an integer.

- The administrator password needs to be reset.
- To modify other parameters, see the description of buying instances of other DB engines in *Getting Started*.

**Step 5** View the results.

A new instance is created using the backup data. The status of the new instance changes from **Creating** to **Available**.

After the restoration, the system will perform a full backup.

The new instance is independent from the original one.

**----End**

# 4.8.3 Restoring a Backup to a Specified Point in Time

You can use existing automated backups to restore GeminiDB DynamoDB-Compatible instance data to a specified point in time.

The most recent full backup will be downloaded from OBS for restoration. After the restoration is complete, incremental backups will be replayed to the specified point in time. The time required depends on the amount of data to be restored.

## Usage Notes

- Data can only be restored to a new instance.
- After automated backup is enabled, the system performs an incremental backup based on the preset incremental backup interval. The incremental backup is stored in OBS.
- Keep your account balance above zero so that backup data can be restored to a new instance.
- Data can be restored to a specified time point only after the automated backup policy is enabled.
- During the instance restoration, backup files are downloaded from OBS buckets to the data directory of the restored instance. The download bandwidth of OBS is 40 MB/s.

## Procedure

**Step 1** Log in to the Huawei Cloud console.

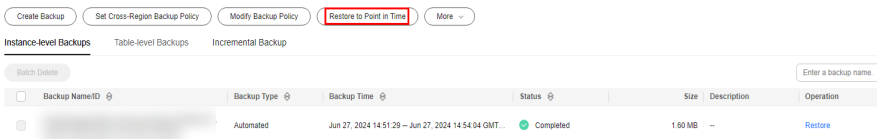
**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, click the target GeminiDB DynamoDB-Compatible instance.

**Step 4** In the navigation pane on the left, choose **Backups & Restorations**.

**Step 5** On the **Backups & Restorations** page, click **Restore to Point in Time**.

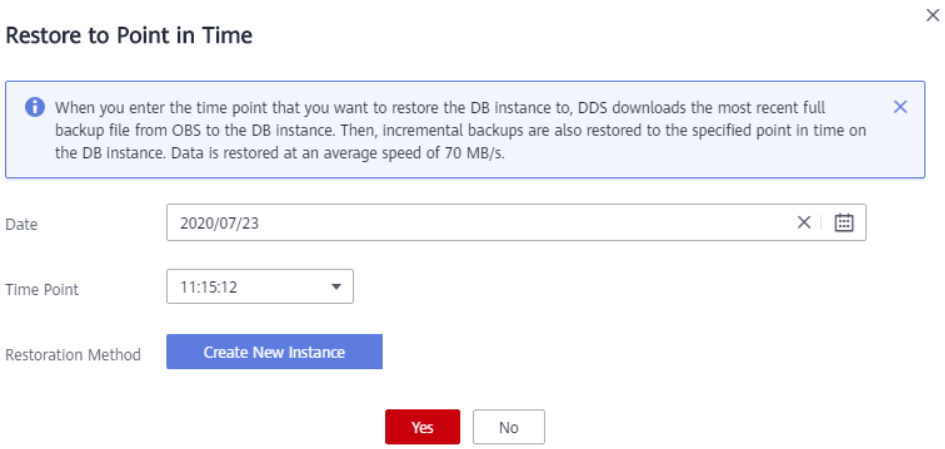
**Figure 4-78** Restoring data to a point in time



**Step 6** Select a restoration date and a time point and click **OK**.

**Figure 4-79** Restore to Point in Time




**Step 7** On the **Create New Instance** page, create an instance of the same specifications as the instance to be restored. The new instance is independent from the original one.

- The new instance should be deployed in a different AZ to ensure that your applications will not be affected by SPOFs.
- The API compatibility, instance type, version, and CPU type are the same as those of the original instance and cannot be changed.
- Other settings are the same as those of the original instance by default but can be modified. For details, see **Buying a GeminiDB DynamoDB-Compatible Instance**.

**----End**

# 4.9 Logs and Audit

## 4.9.1 Viewing and Exporting Slow Query Logs

This section describes how to view database-level slow query logs. Any query that takes longer than an execution time threshold (in milliseconds) will be logged. You can identify the SQL statements that take a long time to execute and tune them based on slow query logs.

### Viewing and Exporting Log Details

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, click the target instance.

**Step 4** In the navigation pane on the left, choose **Slow Query Logs**.

**Step 5** On the **Slow Query Logs** page, set search criteria and click **Search** to view log information.

- Select **All nodes** and view slow query logs of all nodes. Alternatively, select a specific node to view its slow query logs.
- View slow query logs of a specific node in different time ranges.
- View slow query logs of the following types of SQL statements:
  - SELECT

**Step 6** On the **Log Details** page, click in the upper right corner of the log list to export log details.

- You can view the exported CSV file to your local PC.
- Up to 2,000 logs can be exported at a time.

**----End**

# 4.9.2 CTS

## 4.9.2.1 Key Operations Supported by CTS

With CTS, you can record operations on GeminiDB DynamoDB-Compatible instances for later queries, audit, and backtracking.

**Table 4-26** Key operations on GeminiDB DynamoDB-Compatible instances

| Operation | Resource Type | Trace Name |
| --- | --- | --- |
| Creating an instance | instance | NoSQLCreateInstance |
| Deleting an instance | instance | NoSQLDeleteInstance |
| Adding nodes | instance | NoSQLEnlargeInstance |
| Deleting nodes | instance | NoSQLReduceInstance |
| Restarting an instance | instance | NoSQLRestartInstance |
| Restoring data to a new instance | instance | NoSQLRestoreNewInstance |
| Scaling up storage space of an instance | instance | NoSQLExtendInstanceVolume |
| Resetting the password of an instance | instance | NoSQLResetPassword |
| Modifying the name of an instance | instance | NoSQLRenameInstance |
| Changing specifications | instance | NoSQLResizeInstance |
| Binding an EIP | instance | NoSQLBindEIP |
| Unbinding an EIP | instance | NoSQLUnBindEIP |

| Operation | Resource Type | Trace Name |
| --- | --- | --- |
| Freezing an instance | instance | NoSQLFreezeInstance |
| Unfreezing an instance | instance | NoSQLUnfreezeInstance |
| Creating a backup | backup | NoSQLCreateBackup |
| Deleting a backup | backup | NoSQLDeleteBackup |
| Modifying the backup policy of an instance | backup | NoSQLSetBackupPolicy |
| Adding a tag for an instance | tag | NoSQLAddTags |
| Modifying an instance tag | tag | NoSQLModifyInstanceTag |
| Deleting an instance tag | tag | NoSQLDeleteInstanceTag |
| Creating a parameter template | parameterGroup | NoSQLCreateConfigurations |
| Modifying a parameter template | parameterGroup | NoSQLUpdateConfigurations |
| Modifying instance parameters | parameterGroup | NoSQLUpdateInstanceConfigurations |
| Replicating a parameter template | parameterGroup | NoSQLCopyConfigurations |
| Resetting a parameter template | parameterGroup | NoSQLResetConfigurations |
| Applying a parameter template | parameterGroup | NoSQLApplyConfigurations |
| Deleting a parameter template | parameterGroup | NoSQLDeleteConfigurations |
| Deleting the node that fails to be added | instance | NoSQLDeleteEnlargeFailNode |
| Enabling SSL | instance | NoSQLSwitchSSL |
| Changing the security group of an instance | instance | NoSQLModifySecurityGroup |
| Configuring autoscaling | instance | NoSQLModifyAutoEnlargePolicy |
| Creating a dual-active relationship | instance | NoSQLBuildBiactiveInstance |
| Exporting parameter template information for an instance | instance | NoSQLSaveConfigurations |

| Operation | Resource Type | Trace Name |
| --- | --- | --- |
| Modifying the recycling policy | instance | NoSQLModifyRecyclePolicy |

### 4.9.2.2 Querying Traces

After CTS is enabled, CTS starts recording operations on cloud resources. The CTS console stores the last seven days of operation records.

This section describes how to query the last seven days of operation records on the CTS console.

### Procedure

**Step 1** Log in to the Huawei Cloud console.

**Step 2** Click in the upper left corner and select a region and project.

**Step 3** Click **Service List**. Under **Management & Governance**, click **Cloud Trace Service**.

**Step 4** In the navigation pane on the left, click **Trace List**.

**Step 5** Specify filter criteria to search for the required traces. The following four filter criteria are available:

- **Trace Source**, **Resource Type**, and **Search By**

  Select filters from the drop-down list.

  When you select **Trace name** for **Search By**, you need to select a specific trace name.

  When you select **Resource ID** for **Search By**, you also need to select or enter a specific resource ID.

  When you select **Resource name** for **Search By**, you also need to select or enter a specific resource name.
- **Operator**: Select a specific operator (a user other than the tenant).
- **Trace Status**: Select **All trace statuses**, **Normal**, **Warning**, or **Incident**.
- **Start Date** and **End Date**: You can specify a time range to query traces.

**Step 6** Locate the target trace and click to view its details.

**Step 7** Click **View Trace** in the **Operation** column. In the displayed dialog box, the trace structure details are displayed.

**----End**

# 4.10 Viewing Metrics and Configuring Alarms

# 4.10.1 Supported Metrics

This section describes GeminiDB DynamoDB-Compatible API metrics reported to Cloud Eye as well as their namespaces and dimensions. You can use APIs provided by Cloud Eye to query the metrics and alarms.

## Namespace

SYS.NoSQL

## Metrics

NOTE

You can view metrics on instance nodes by following **Viewing Metrics**.

**Table 4-27** GeminiDB DynamoDB-Compatible API metrics

| Metric ID | Name | Description | Value Range | Unit | Number System | Monitored Object | Monitoring Period (Raw Data) |
|---|---|---|---|---|---|---|---|
| nosql005_disk_usage | Storage Space Usage | Storage usage of the current instance. | 0–100 | % | N/A | GeminiDB Dynamo DB-Compatible instance | 1 minute |
| nosql006_disk_total_size | Total Storage Space | Total storage space of the current instance. | ≥ 0 | GB | 1024(IEC) | GeminiDB Dynamo DB-Compatible instance | 1 minute |
| nosql007_disk_used_size | Storage Space Usage | Storage space usage of the current instance. | ≥ 0 | GB | 1024(IEC) | GeminiDB Dynamo DB-Compatible instance | 1 minute |

| Metric ID | Name | Description | Value Range | Unit | Number System | Monitored Object | Monitoring Period (Raw Data) |
|---|---|---|---|---|---|---|---|
| nosql009_dfv_write_delay | Storage Write Latency | Average delay of writing data to the storage layer in a specified period | ≥ 0 | ms | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| nosql010_dfv_read_delay | Storage Read Latency | Average latency of reading data from the storage layer in a specified period | ≥ 0 | ms | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra001_cpu_usage | CPU Usage | CPU usage of an instance | 0–100 | % | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra002_mem_usage | Memory Usage | Memory usage of the instance | 0–100 | % | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra003_bytes_out | Network Output Throughput | Outgoing traffic in bytes per second | ≥ 0 Bytes/s | Bytes/s | 1024(IEC) | GeminiDB DynamoDB-Compatible instance node | 1 minute |

| Metric ID | Name | Description | Value Range | Unit | Number System | Monitored Object | Monitoring Period (Raw Data) |
|---|---|---|---|---|---|---|---|
| cassandra004_bytes_in | Network Input Throughput | Incoming traffic in bytes per second | ≥ 0 | Bytes/s | 1024(IEC) | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra014_connections | Active Node Connections | Active connections to the current GeminiDB DynamoDB-Compatible instance node | ≥ 0 | Counts | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra015_read_latency | Average Read Latency | Average amount of time consumed by read requests | ≥ 0 | ms | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra016_write_latency | Average Write Latency | Average amount of time consumed by write requests | ≥ 0 | ms | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra037_pending_write | Suspended Write Tasks | Number of write tasks waiting in the queue | ≥ 0 | Counts | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |

| Metric ID | Name | Description | Value Range | Unit | Number System | Monitored Object | Monitoring Period (Raw Data) |
|---|---|---|---|---|---|---|---|
| cassandra038_pending_read | Suspended Read Tasks | Number of read tasks waiting in the queue. | ≥ 0 | Counts | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra044_range_slice_latency | Scan Duration | Average time consumed by scan operations | ≥ 0 | ms | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra049_dropped_mutation | Dropped Writes | Average number of dropped writes | ≥ 0 | Counts | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra052_dropped_read | Dropped Reads | Average number of dropped reads | ≥ 0 | Counts | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra092_load_info | Data Volume on a Node | Data volume on a node | ≥ 0 | Byte | 1024(IEC) | GeminiDB DynamoDB-Compatible instance node | 1 minute |

| Metric ID | Name | Description | Value Range | Unit | Number System | Monitored Object | Monitoring Period (Raw Data) |
|---|---|---|---|---|---|---|---|
| cassandra093_write_count_latency | Accumulated Write Requests | Number of write requests initiated by a node | ≥ 0 | Counts | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra094_write_1min_rate | Average Write Rate in the Last Minute | Average write rate in the last minute | ≥ 0 | Counts/s | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra095_write_p75_latency | p75 Write Latency | p75 write latency | ≥ 0 | ms | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra096_write_p95_latency | p95 Write Latency | p95 write latency | ≥ 0 | ms | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra097_write_p99_latency | p99 Write Latency | p99 write latency | ≥ 0 | ms | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |

| Metric ID | Name | Description | Value Range | Unit | Number System | Monitored Object | Monitoring Period (Raw Data) |
|---|---|---|---|---|---|---|---|
| cassandra098_read_count_latency | Accumulated Read Requests | Number of read requests initiated by a node | ≥ 0 | Counts | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra099_read_1min_rate | Average Read Rate in the Last Minute | Average read rate in the last minute | ≥ 0 | Counts/s | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra100_read_p75_latency | p75 Read Latency | p75 read latency | ≥ 0 | ms | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra101_read_p95_latency | p95 Read Latency | p95 read latency | ≥ 0 | ms | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra102_read_p99_latency | p99 Read Latency | p99 read latency | ≥ 0 | ms | N/A | instance node | 1 minute |

| Metric ID | Name | Description | Value Range | Unit | Number System | Monitored Object | Monitoring Period (Raw Data) |
|---|---|---|---|---|---|---|---|
| cassandra103_range_slice_count_latency | Accumulated Range Read Requests | Number of range read requests | ≥ 0 | Counts | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra104_range_slice_1min_rate | Average Range Read Rate in the Last Minute | Average range read rate in the last minute | ≥ 0 | Counts/s | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra105_range_slice_p75_latency | p75 Range Read Latency | p75 range read latency | ≥ 0 | ms | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra106_range_slice_p95_latency | p95 Range Read Latency | p95 range read latency | ≥ 0 | ms | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra107_range_slice_p99_latency | p99 Range Read Latency | p99 range read latency | ≥ 0 | ms | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |

| Metric ID | Name | Description | Value Range | Unit | Number System | Monitored Object | Monitoring Period (Raw Data) |
|---|---|---|---|---|---|---|---|
| cassandra163_write_p999_latency | p999 Write Latency | p999 write latency | ≥ 0 | ms | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra164_read_p999_latency | p999 Read Latency | p999 read latency | ≥ 0 | ms | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra165_large_partition_num | Big Keys | Number of big keys on the current node | ≥ 0 | Counts | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra166_write_max_latency | Maximum Write Latency | Maximum write latency | ≥ 0 | ms | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra167_read_max_latency | Maximum Read Latency | Maximum read latency | ≥ 0 | ms | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |

| Metric ID | Name | Description | Value Range | Unit | Number System | Monitored Object | Monitoring Period (Raw Data) |
|---|---|---|---|---|---|---|---|
| cassandra168_imbalance_table_num | Tables with Uneven Data Distribution | Number of tables whose data is not evenly distributed | ≥ 0 | Counts | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra169_modify_request_size_mean | Average Write Request Size | Average write request size | ≥ 0 | Bytes | 1024(IEC) | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra170_query_response_size_mean | Average Query Response Size | Average size of query requests | ≥ 0 | Bytes | 1024(IEC) | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra173_limit_diff_count_mean | Mean of limit Value and Returned Rows | Mean of limit difference | ≥ 0 | - | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra174_tombstone_query_rate | Tombstone Query Requests per Second | Rate of tombstone query requests | ≥ 0 | Counts/s | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |

| Metric ID | Name | Description | Value Range | Unit | Number System | Monitored Object | Monitoring Period (Raw Data) |
|---|---|---|---|---|---|---|---|
| cassandra175_single_delete_rate | Row Delete Requests per Second | Rate at which a single row is deleted | ≥ 0 | Counts/s | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra176_range_delete_rate | Range Delete Requests per Second | Range deletion rate | ≥ 0 | Counts/s | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra177_large_row_count | Large Rows per Second | Number of large rows | ≥ 0 | Counts/s | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |
| cassandra174_tombstone_query_rate | Maximum sync delay | Maximum synchronization latency between the primary and standby instances | ≥ 0 | ms | N/A | GeminiDB DynamoDB-Compatible instance node | 1 minute |

## Dimensions

| Key | Value |
|---|---|
| cassandra_cluster_id | instance cluster ID |

| Key | Value |
| --- | --- |
| cassandra_node_id | instance node ID |

## 4.10.2 Configuring Alarm Rules

Setting alarm rules allows you to customize objects to be monitored and notification policies so that you can closely monitor your instances.

Alarm rules include the alarm rule name, instance, metric, threshold, monitoring interval, and whether to send notifications. This section describes how to set alarm rules.

NOTE

For more information about alarm rules, see *Cloud Eye User Guide*.

### Procedure

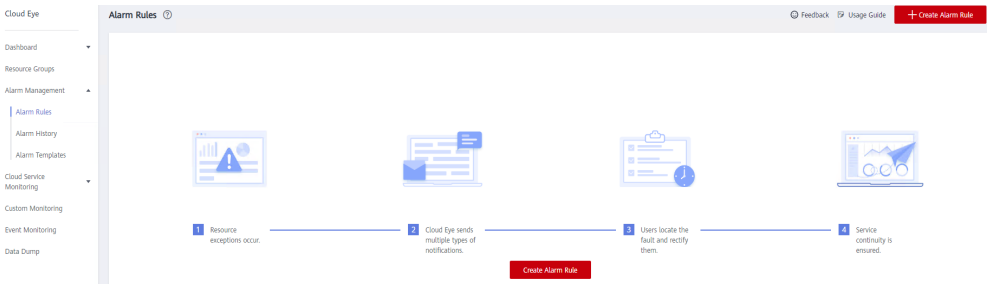
**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Management & Governance** > **Cloud Eye**.

**Step 3** In the navigation pane on the left, choose **Alarm Management** > **Alarm Rules**.

**Step 4** On the **Alarm Rules** page, click **Create Alarm Rule**.

**Figure 4-80** Creating an alarm rule




**Step 5** Set alarm parameters.

1. Configure basic alarm information.

**Figure 4-81** Configuring basic information for an alarm rule

**Table 4-28** Basic alarm rule information

| Parameter | Description | Example Value |
| --- | --- | --- |
| Name | Name of the rule. The system generates a random name and you can modify it. | alarm-cag2 |
| Description | (Optional) Alarm rule description. | - |

2. Select objects to be monitored and specify the monitoring scope.

**Figure 4-82** Configuring objects to be monitored

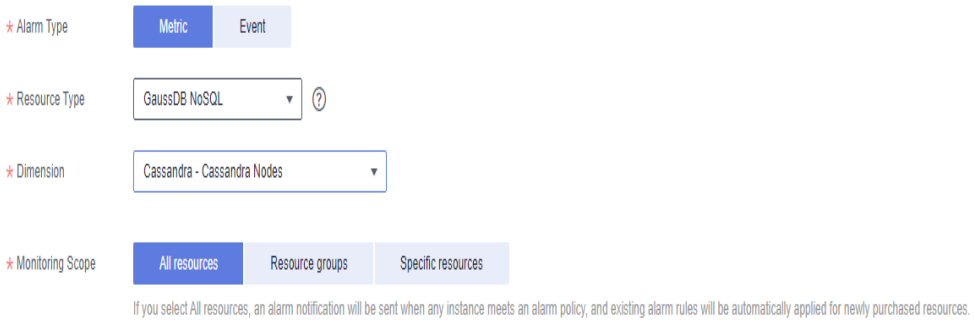



**Table 4-29** Parameter description

| Parameter | Description | Example Value |
| --- | --- | --- |
| Alarm Type | Alarm type that the alarm rule is created for. The value can be **Metric** or **Event**. | Metric |
| Resource Type | Type of the resource the alarm rule is created for.<br>Select **GeminiDB**. | - |
| Dimension | Metric dimension of the alarm rule.<br>Select **Cassandra - Cassandra Nodes**. | - |

| Parameter | Description | Example Value |
| --- | --- | --- |
| Monitoring Scope | Monitoring scope the alarm rule applies to.<br>**NOTE**<br>– If you select **All resources**, an alarm notification will be sent when any instance meets an alarm policy, and existing alarm rules will be automatically applied for newly purchased resources.<br>– If you select **Resource groups** and any resource in the group meets the alarm policy, an alarm notification will be sent.<br>– If you select **Specific resources**, click **Select Specified Resources**, select one or more resources, and click **OK**. | All resources |
| Group | This parameter is mandatory when **Monitoring Scope** is set to **Resource groups**. | - |

3. Configure an alarm policy.

**Figure 4-83** Configuring an alarm policy

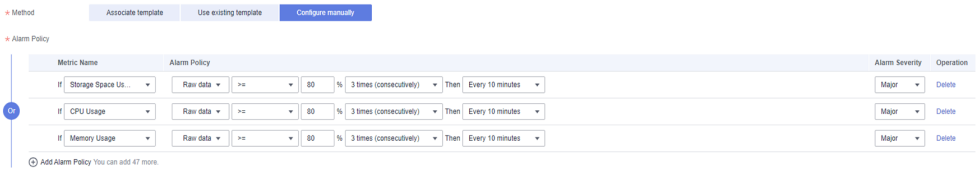



**Table 4-30** Parameter description

| Parameter | Description | Example Value |
| --- | --- | --- |
| Method | Select **Associate template**, **Use existing template**, or **Configure manually**.<br>**NOTE**<br>If you set **Monitoring Scope** to **Specific resources**, you can set **Method** to **Use existing template**. | Configure manually |
| Template | Select the template to be used.<br>This parameter is available only when you set **Method** to **Use existing template**. | - |

| Parameter | Description | Example Value |
| --- | --- | --- |
| Alarm Policy | Policy for triggering an alarm. You can configure the threshold, consecutive periods, alarm interval, and alarm severity based on service requirements.<br>– **Metric Name**: metric that an alarm rule is created for.<br>The following metrics are recommended:<br>**Storage Space Usage**:<br>Storage usage of a GeminiDB DynamoDB-Compatible instance. If the storage usage is greater than 80%, scale up the storage in a timely manner by following **Manually Scaling Up Storage Space**.<br>**CPU Usage** and **Memory Usage**:<br>Compute resource usage of each GeminiDB DynamoDB-Compatible instance node. If the CPU usage or memory usage is greater than 80%, you are advised to **add nodes** or **increase node specifications**.<br>For more metrics, see **Supported Metrics**.<br>– **Alarm Severity**: specifies the severity of the alarm. Valid values are **Critical**, **Major**, **Minor**, and **Informational**.<br>**NOTE**<br>A maximum of 50 alarm policies can be added to an alarm rule. If any one of these alarm policies is met, an alarm is triggered. | Take the CPU usage as an example. The alarm policy configured in **Figure 4-83** indicates that a major alarm notification will be sent to users every 10 minutes if the original CPU usage reaches 80% or above for three consecutive periods. |

4. Configure alarm notification information.

**Figure 4-84** Configuring alarm notification information

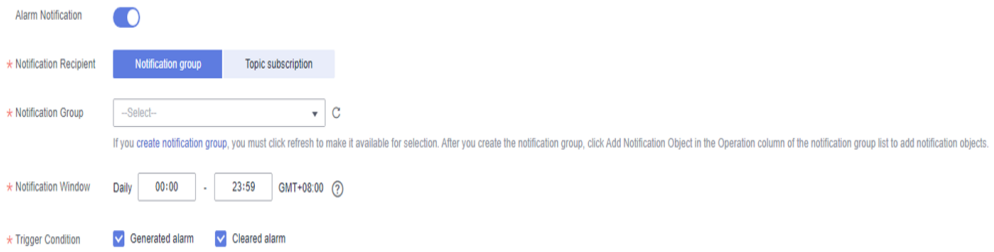


**Table 4-31** Parameter description

| Parameter | Description | Example Value |
| --- | --- | --- |
| Alarm Notification | Whether to notify users when alarms are triggered. Notifications can be sent by email, text message, or HTTP/HTTPS message.<br>Enabling alarm notification is recommended. When the metric data reaches the threshold set in the alarm rule, Cloud Eye immediately notifies you through SMN that an exception has occurred. | Enabled **Alarm Notification**. |
| Notification Recipient | Select **Notification group** or **Topic subscription**. | - |
| Notification Group | Notification group the alarm notification is to be sent to. | - |
| Notification Object | Specifies the object that receives alarm notifications. You can select the account contact or a topic.<br>– Account contact is the mobile phone number and email address provided for registration.<br>– **Topic** is used to publish messages and subscribe to notifications. If the required topic is unavailable, create one first and add subscriptions to it. For details, see **Creating a Topic** and **Adding Subscriptions**. | - |
| Notification Window | Cloud Eye sends notifications only within the notification window specified in the alarm rule.<br>For example, if **Notification Window** is set to **00:00-8:00**, Cloud Eye sends notifications only within 00:00-08:00. | - |

| Parameter | Description | Example Value |
| --- | --- | --- |
| Trigger Condition | Condition for triggering an alarm notification. You can select **Generated alarm** (when an alarm is generated), **Cleared alarm** (when an alarm is cleared), or both. | - |

5. Configure advanced settings.

**Figure 4-85** Advanced settings

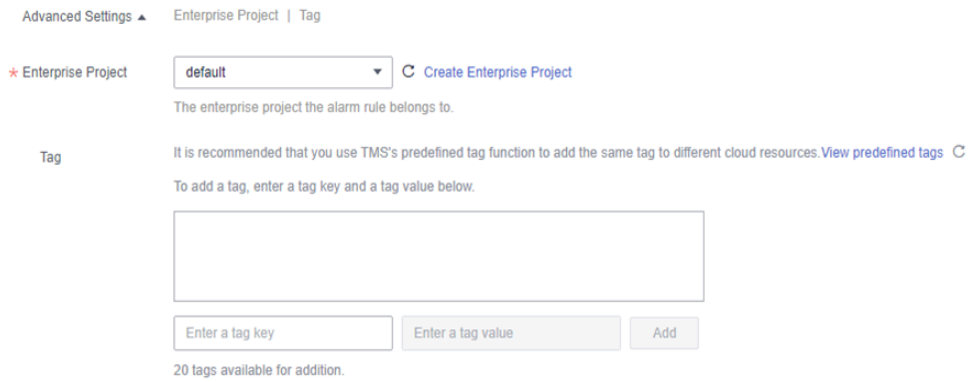



**Table 4-32** Parameter description

| Parameter | Description | Example Value |
| --- | --- | --- |
| Enterprise Project | Enterprise project that the alarm rule belongs to. Only users with the enterprise project permissions can view and manage the alarm rule. For details about how to create an enterprise project, see **Creating an Enterprise Project**. | default |
| Tag | A tag is a key-value pair. Tags identify cloud resources so that you can easily categorize and search for your resources. You are advised to create predefined tags on TMS. For details about how to create predefined tags, see **Creating Predefined Tags**.<br>– A key can contain a maximum of 128 characters, and a value can contain a maximum of 255 characters.<br>– A maximum of 20 tags can be added. | - |

**Step 6** After the configuration is complete, click **Create**.

When the metric data reaches the threshold set in the alarm rule, Cloud Eye immediately notifies you through SMN that an exception has occurred.

**----End**

## 4.10.3 Viewing Metrics

Cloud Eye monitors the status of GeminiDB DynamoDB-Compatible instances. You can check GeminiDB DynamoDB-Compatible API metrics on the console.

Monitored data requires a period of time for transmission and display. The status of the monitored object displayed on the Cloud Eye page is the status obtained 5 to 10 minutes before. You can view the monitored data of a newly created DB instance 5 to 10 minutes later.

### Usage Notes

- The DB instance is running properly.

  Cloud Eye does not display the metrics of a faulty or deleted DB instance. You can view the monitoring information only after the instance is restarted or recovered.
- The DB instance has been properly running for at least 10 minutes.

  The monitoring data and graphics are available for a new DB instance after the instance runs for at least 10 minutes.

### Procedure

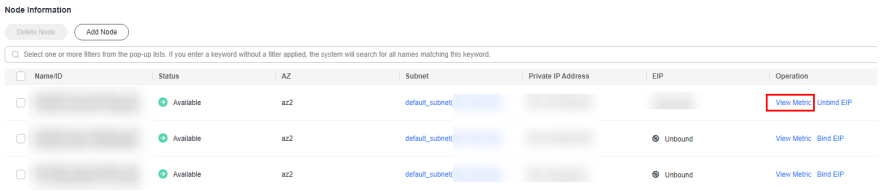
**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instance** page, click the instance whose metrics you want to view and click its name.

**Step 4** In the **Node Information** area on the **Basic Information** page, click **View Metric** in the **Operation** column.

**Figure 4-86** Viewing metrics




**Step 5** In the monitoring area, select a time range to view monitoring data.

You can view the monitoring data in the last 1, 3, or 12 hours.

To view the graph in a longer time range, click .

**----End**

# 4.10.4 Event Monitoring

## 4.10.4.1 Introduction

Events and alarms will be reported on the **Event Monitoring** page. You can create alarm rules for both system and custom events. When a specific event occurs, Cloud Eye generates and sends an alarm for you.

Key operations on GeminiDB DynamoDB-Compatible instances are monitored and recorded by Cloud Eye as events. Events include operations performed by specific users on specific resources, such as changing instance names and specifications.

Event monitoring provides an API for reporting custom events, which helps you collect and report abnormal events or important change events generated by services to Cloud Eye.

Event monitoring is enabled by default. You can view monitoring details of system and custom events. For details about system events, see **Events Supported by Event Monitoring**.

## 4.10.4.2 Viewing Event Monitoring Data

### Scenarios

Events and alarms will be reported on the **Event Monitoring** page. You can create alarm rules for both system and custom events. When a specific event occurs, Cloud Eye generates and sends an alarm for you.

Event monitoring is enabled by default. You can view monitoring details about system events and custom events.

This topic describes how to view the event monitoring data.

### Procedure

**Step 1** **Log in to the Huawei Cloud console.**

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, locate the instance whose event monitoring data you want to view. In the **Node Information** area on the **Basic Information** page, click **View Metric** in the **Operation** column.

**Step 4** Click < in the upper part to return to the Cloud Eye console.

**Step 5** In the navigation pane on the left, choose **Event Monitoring**.

On the displayed **Event Monitoring** page, all system events generated in the last 24 hours are displayed by default.

You can also click **1h**, **3h**, **12h**, **1d**, **7d**, or **30d** to view events generated in different time periods.

**Step 6** Locate an event and click **View Event** in the **Operation** column to view its details.

**----End**

### 4.10.4.3 Creating an Alarm Rule for Event Monitoring

## Scenarios

This topic describes how to create an alarm rule for event monitoring.

## Procedure

**Step 1** Log in to the Huawei Cloud console.

**Step 2** Click ☰ in the upper left corner of the page. Choose **Management & Governance** > **Cloud Eye**.

**Step 3** In the navigation pane on the left, choose **Event Monitoring**.

**Step 4** On the event list page, click **Create Alarm Rule** in the upper right corner.

**Step 5** On the **Create Alarm Rule** page, configure the parameters.

**Table 4-33** Parameter description

| Parameter | Description |
| --- | --- |
| Name | Specifies the name of the alarm rule. The system generates a random name, but you can change it if needed. |
| Description | (Optional) Provides supplementary information about the alarm rule. |
| Enterprise Project | You can select an existing enterprise project or click **Create Enterprise Project** to create one. |
| Alarm Type | Specifies the alarm type corresponding to the alarm rule. |
| Event Type | Specifies the event type of the metric corresponding to the alarm rule. |
| Event Source | Specifies the service the event is generated for.<br>Select GeminiDB. |
| Monitoring Scope | Specifies the monitoring scope for event monitoring. |
| Method | Specifies the event creation method. |
| Alarm Policy | **Event Name** indicates the instantaneous operations users performed on system resources, such as login and logout.<br>For details about events supported by Event Monitoring, see **Events Supported by Event Monitoring**.<br>You can select a trigger mode and alarm severity as needed. |

Click  to enable **Alarm Notification**. The validity period is 24 hours by default. If required topics are not displayed in the drop-down list, click **Create an SMN topic**.

**Table 4-34** Alarm notification parameters

| Parameter | Description |
| --- | --- |
| Alarm Notification | Whether to notify users when alarms are triggered. Notifications can be sent by email, text message, or HTTP/HTTPS message. |
| Notification Object | Object an alarm notification is to be sent to. You can select the account contact or a topic.<br>• Account contact is the mobile phone number and email address provided for registration.<br>• Topic is used to publish messages and subscribe to notifications. If the required topic is unavailable, create one first and add subscriptions to it.<br>For details, see **Creating a Topic** and **Adding Subscriptions**. |
| Validity Period | Notification window which Cloud Eye only sends notifications within.<br>If you set **Validity Period** to **08:00-20:00**, Cloud Eye sends notifications only within 08:00-20:00. |
| Trigger Condition | Condition for triggering the alarm notification. |

**Step 6** After the configuration is complete, click **Create**.

**----End**

### 4.10.4.4 Events Supported by Event Monitoring

**Table 4-35** Events supported by GeminiDB Event Monitoring

| Event Source | Event Name | Event ID | Event Severity | Description | Solution | Impact |
|---|---|---|---|---|---|---|
| NoSQL | Instance creation failure | NoSQLCreateInstanceFailed | Major | The instance quota or underlying resources are insufficient. | Release unnecessary instances and try again. You can also choose **Service Tickets > Create Service Ticket** in the upper right corner of the console and submit a service ticket to adjust the quota. | Instances fail to be created. |
| | Specifications change failure | NoSQLResizeInstanceFailed | Major | The underlying resources are insufficient. | Choose **Service Tickets > Create Service Ticket** in the upper right corner of the console. Submit a service ticket to the O&M personnel to coordinate resources in the background and change the specifications again. | Services are interrupted. |

| Event Source | Event Name | Event ID | Event Severity | Description | Solution | Impact |
|---|---|---|---|---|---|---|
| | Node adding failure | NoSQLAddNodesFailed | Major | The underlying resources are insufficient. | Choose **Service Tickets > Create Service Ticket** in the upper right corner of the console. Submit a service ticket to O&M personnel to coordinate resources in the background, delete nodes that failed to be added, and add the nodes again. | None |
| | Node deletion failure | NoSQLDeleteNodesFailed | Major | Releasing underlying resources failed. | Delete the node again. | None |
| | Storage space scale-up failure | NoSQLScaleUpStorageFailed | Major | The underlying resources are insufficient. | Choose **Service Tickets > Create Service Ticket** in the upper right corner of the console. Submit a service ticket to O&M personnel to coordinate resources in the background and scale up storage again. | Services may be interrupted. |

| Event Source | Event Name | Event ID | Event Severity | Description | Solution | Impact |
|---|---|---|---|---|---|---|
| | Password resetting failure | NoSQLResetPasswordFailed | Major | Resetting the password times out. | Reset the password again. | None |
| | Parameter template change failure | NoSQLUpdateInstanceParamGroupFailed | Major | Changing a parameter template times out. | Change the parameter template again. | None |
| | Backup policy configuration failure | NoSQLSetBackupPolicyFailed | Major | The database connection is abnormal. | Configure the backup policy again. | None |
| | Manual backup creation failure | NoSQLCreateManualBackupFailed | Major | The backup files fail to be exported or uploaded. | Choose **Service Tickets > Create Service Ticket** in the upper right corner of the console and submit a service ticket to O&M personnel. | Data cannot be backed up. |
| | Automated backup creation failure | NoSQLCreateAutomatedBackupFailed | Major | The backup files fail to be exported or uploaded. | Choose **Service Tickets > Create Service Ticket** in the upper right corner of the console and submit a service ticket to O&M personnel. | Data cannot be backed up. |

| Event Source | Event Name | Event ID | Event Severity | Description | Solution | Impact |
|---|---|---|---|---|---|---|
| | Instance status abnormal | NoSQLFaultyDBInstance | Major | This event is a key alarm event and is reported when an instance is faulty due to a disaster or a server failure. | Choose **Service Tickets > Create Service Ticket** in the upper right corner of the console and submit a service ticket to O&M personnel. | The database service may be unavailable. |
| | Instance status recovery | NoSQLDBInstanceRecovered | Major | If a disaster occurs, NoSQL provides an HA tool to automatically or manually rectify the fault. After the fault is rectified, this event is reported. | No further action is required. | None |
| | Node status abnormal | NoSQLFaultyDBNode | Major | This event is a key alarm event and is reported when a database node is faulty due to a disaster or a server failure. | Check whether the database service is functional. Choose **Service Tickets > Create Service Ticket** in the upper right corner of the console and submit a service ticket to O&M personnel. | The database service may be unavailable. |

| Event Source | Event Name | Event ID | Event Severity | Description | Solution | Impact |
|---|---|---|---|---|---|---|
| | Node status recovery | NoSQLDBNodeRecovered | Major | If a disaster occurs, NoSQL provides an HA tool to automatically or manually rectify the fault. After the fault is rectified, this event is reported. | No further action is required. | None |
| | Primary/standby switchover or failover | NoSQLPrimaryStandbySwitched | Major | This event is reported when a primary/standby switchover or a failover is triggered. | No further action is required. | None |
| | Occurrence of hotspot partitioning keys | HotKeyOccurs | Major | Hotspot data is stored in one partition because the primary key is improper. Improper application design causes frequent read and write operations on a key. | 1. Choose a proper partition key.<br>2. Add service cache so that service applications read hotspot data from the cache first. | The service request success rate is affected, and the cluster performance and stability deteriorates. |

| Event Source | Event Name | Event ID | Event Severity | Description | Solution | Impact |
|---|---|---|---|---|---|---|
| | BigKey occurrence | BigKeyOccurs | Major | The primary key design is improper. There are too many records or too much data in a single partition, causing load imbalance on nodes. | 1. Choose a proper partition key.<br>2. Add a new partition key for hashing data. | As more and more data is stored in the partition, cluster stability deteriorates. |
| | Insufficient storage space | NoSQLRiskyDataDiskUsage | Major | The storage space is insufficient. | Scale up storage space. For details, see section "Scaling Up Storage Space" in the user guide of GeminiDB. | The instance is set to read-only and data cannot be written to the instance. |
| | Data disk expanded and being writable | NoSQLDataDiskUsageRecovered | Major | The data disk has been expanded and becomes writable. | No further action is required. | None |

| Event Source | Event Name | Event ID | Event Severity | Description | Solution | Impact |
|---|---|---|---|---|---|---|
| | Index creation failure | NoSQL CreateIndexFailed | Major | The service load exceeds what the instance specifications can take. In this case, creating indexes consumes more instance resources. As a result, the response is slow or even frame freezing occurs, and the creation times out. | 1. Select matched instance specifications based on the service loads.<br>Create indexes during off-peak hours.<br>Create indexes in the background.<br>Select indexes as required. | The index fails to be created or is incomplete. Delete the index and create a new one. |
| | Write speed decrease | NoSQL Stalling Occurs | Major | The write speed is close to the maximum write speed allowed by the cluster scale and instance specifications. As a result, the database flow control mechanism is triggered, and requests may fail. | 1. Adjust the cluster scale or node specifications based on the maximum write rate of services.<br>2. Measure the maximum write request rate of services and distribute the peak write rate of services. | The success rate of service requests is affected. |

| Event Source | Event Name | Event ID | Event Severity | Description | Solution | Impact |
|---|---|---|---|---|---|---|
| | Data write stopped | NoSQLStoppingOccurs | Major | The data write is too fast, reaching the maximum write capability allowed by the cluster scale and instance specifications. As a result, the database flow control mechanism is triggered, and requests may fail. | 1. Adjust the cluster scale or node specifications based on the maximum write rate of services.<br>2. Measure the maximum write request rate of services and distribute the peak write rate of services. | The success rate of service requests is affected. |
| | Database restart failure | NoSQLRestartDBFailed | Major | The instance status is abnormal. | Choose **Service Tickets > Create Service Ticket** in the upper right corner of the console and submit a service ticket to O&M personnel. | The instance status may be abnormal. |

| Event Source | Event Name | Event ID | Event Severity | Description | Solution | Impact |
|---|---|---|---|---|---|---|
| | Restoration to new instance failure | NoSQLRestoreToNewInstanceFailed | Major | The underlying resources are insufficient. | Choose **Service Tickets > Create Service Ticket** in the upper right corner of the console. Submit a service ticket to O&M personnel to coordinate resources in the background and add nodes again. | Data cannot be restored to a new instance. |
| | Restoration to existing instance failure | NoSQLRestoreToExistInstanceFailed | Major | The backup file fails to be downloaded or restored. | Choose **Service Tickets > Create Service Ticket** in the upper right corner of the console and submit a service ticket to O&M personnel. | The current instance may be unavailable. |
| | Backup file deletion failure | NoSQLDeleteBackupFailed | Major | The backup files fail to be deleted from OBS. | Delete the backup files again. | None |

| Event Source | Event Name | Event ID | Event Severity | Description | Solution | Impact |
|---|---|---|---|---|---|---|
| | Failure to display slow query logs in plaintext | NoSQLSwitchSlowlogPlainTextFailed | Major | The API does not support this function. | Refer to *GeminiDB User Guide* to ensure that the API supports slow query logs in plaintext. Choose **Service Tickets > Create Service Ticket** in the upper right corner of the console and submit a service ticket to O&M personnel. | None |
| | EIP binding failure | NoSQLBindEipFailed | Major | The node status is abnormal, an EIP has been bound to the node, or the EIP to be bound is invalid. | Check whether the node is normal and whether the EIP is valid. | The instance cannot be accessed from a public network. |
| | EIP unbinding failure | NoSQLUnbindEipFailed | Major | The node status is abnormal or the EIP has been unbound from the node. | Check whether the node and EIP status are normal. | None |

| Event Source | Event Name | Event ID | Event Severity | Description | Solution | Impact |
|---|---|---|---|---|---|---|
| | Parameter modification failure | NoSQLModifyParameterFailed | Major | The parameter value is invalid. | Check whether the parameter value is valid. Choose **Service Tickets > Create Service Ticket** in the upper right corner of the console and submit a service ticket to O&M personnel. | None |
| | Parameter template application failure | NoSQLApplyParameterGroupFailed | Major | The instance status is abnormal. So, the parameter template cannot be applied. | Choose **Service Tickets > Create Service Ticket** in the upper right corner of the console and submit a service ticket to O&M personnel. | None |

| Event Source | Event Name | Event ID | Event Severity | Description | Solution | Impact |
|---|---|---|---|---|---|---|
| | Enabling or disabling SSL failure | NoSQLSwitchSSLFailed | Major | Enabling or disabling SSL times out. | Try again or choose **Service Tickets > Create Service Ticket** in the upper right corner of the console and submit a service ticket. Retain the SSL connection mode configured before the event occurred. | The SSL connection mode cannot be changed. |
| | Too much data in a single row | LargeRowOccurs | Major | If there is too much data in a single row, queries may time out, causing faults like OOM error. | 1. Limit the write length of each column and row so that the key and value length of each row does not exceed the preset threshold.<br>2. Check whether there are abnormal writes or coding, causing large rows. | If there are too many records in a single row, cluster stability will deteriorate as the data volume increases. |

# 4.11 Enterprise Project

## 4.11.1 Overview

An enterprise project facilitates project-level management and grouping of cloud resources and users. The default project is **default**.

You can also customize enterprise projects to meet your service requirements. For details, see **Enterprise Management User Guide**.

## 4.11.2 Quota Management

GeminiDB DynamoDB-Compatible API provides a quota function that allows you to manage resources by controlling the number of resources in each enterprise project to ensure that resources can be used and managed properly.

This section describes how to query used resources in each enterprise project and its resource quotas.

This function is now in OBT. To use it, choose **Service Tickets > Create Service Ticket** in the upper right corner of the console and contact the customer service.

### Viewing Resource Quotas in Each Enterprise Project

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** In the navigation pane on the left, choose **My Quotas** to view quota details of the current enterprise project.

**Figure 4-87** Quota management

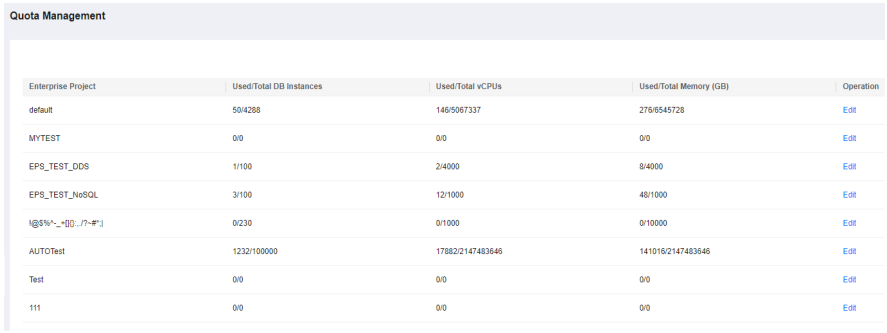



**Table 4-36** Parameter description

| Parameter | Description |
| --- | --- |
| Enterprise Project | Enterprise project that an instance belongs to. |
| Used/Total DB Instances | Number of used instances in the current enterprise project |
| Used/Total vCPUs | vCPUs of all instances in the current enterprise project |

| Parameter | Description |
| --- | --- |
| Used/Total Memory (GB) | Memory of all instances in the current enterprise project |

NOTE

If there are no resources in an enterprise project, the default quota is 0. Before creating an instance, you need to set quotas first by following **Modifying Resource Quotas of an Enterprise Project**.

**----End**

## Modifying Resource Quotas of an Enterprise Project

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** In the navigation pane on the left, choose **My Quotas**. In the quota list, select the enterprise project you want to set quotas for and click **Modify** in the **Operation** column.

**Figure 4-88** Managing quotas

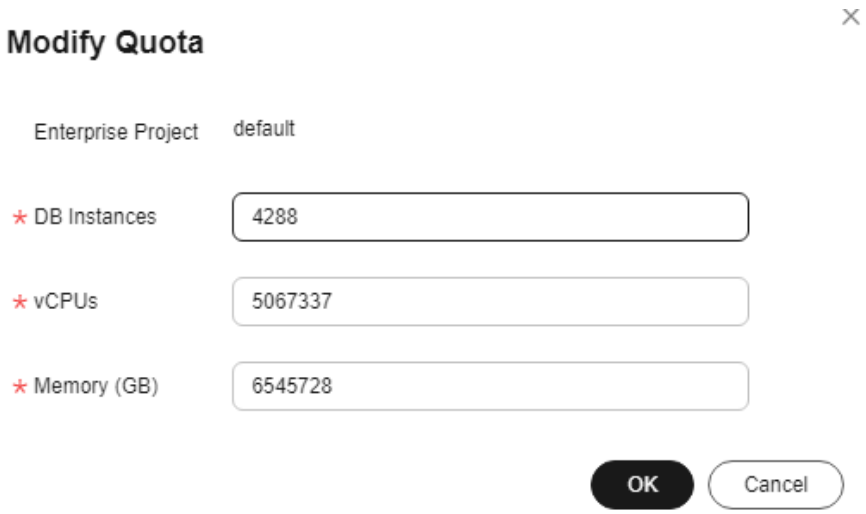



**Table 4-37** Quota management

| Parameter | Value Range |
| --- | --- |
| DB Instances | 0–5,000 |
| vCPUs | 0–8,000,000 |
| Memory (GB) | 0–16,000,000 |

**----End**

# 4.12 Tag Management

Tag Management Service (TMS) enables you to manage resources using tags on the management console. TMS works with other cloud services to manage tags. TMS manages tags globally while other cloud services manage their own tags.

You can tag your GeminiDB DynamoDB-Compatible instance to easily identify and manage its resources. An instance can be tagged when or after it is created.

After an instance is tagged, you can search for the tag key or value to quickly query the instance details.

## Usage Notes

- You are advised to set predefined tags on the TMS console.
- A tag consists of a key and value. You can add only one value for each key. For details about the naming rules of tag keys and tag values, see **Table 4-38**.
- A maximum of 20 tags can be added for each instance.
- The tag name must comply with the naming rules described in **Table 4-38**.

**Table 4-38** Naming rules

| Parameter | Requirement | Example Value |
| --- | --- | --- |
| Tag key | • Cannot be left blank.<br>• Must be unique for each instance.<br>• Can contain a maximum of 128 characters.<br>• Cannot start with _**sys**_ and cannot start or end with a space. Only letters, digits, spaces, and the following special characters are allowed: -_@.:/+= | Organization |
| Tag value | • Can be left blank.<br>• Can contain a maximum of 255 characters.<br>• Only letters, digits, spaces, and the following special characters are allowed: -_@.:/+= | nosql_01 |

## Adding a Tag

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, click the target instance. The **Basic Information** page is displayed.

**Step 4** In the navigation pane on the left, click **Tags**.

**Step 5** On the **Tags** page, click **Add Tag**. In the displayed dialog box, enter a tag key and value, and click **OK**.

**Step 6** View and manage tags on the **Tags** page.

**----End**

## Editing a Tag

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, click the target instance. The **Basic Information** page is displayed.

**Step 4** In the navigation pane on the left, click **Tags**.

**Step 5** On the **Tags** page, locate the tag that you want to edit and click **Edit** in the **Operation** column. In the displayed dialog box, change the tag value and click **OK**.

Only the tag value can be edited.

**Step 6** View and manage tags on the **Tags** page.

**----End**

## Deleting a Tag

**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, click the target instance. The **Basic Information** page is displayed.

**Step 4** In the navigation pane on the left, click **Tags**.

**Step 5** On the **Tags** page, locate the tag that you want to delete and click **Delete** in the **Operation** column. In the displayed dialog box, click **Yes**.

**Step 6** Check whether the deleted tag is displayed on the **Tags** page.

**----End**

## Searching by tag

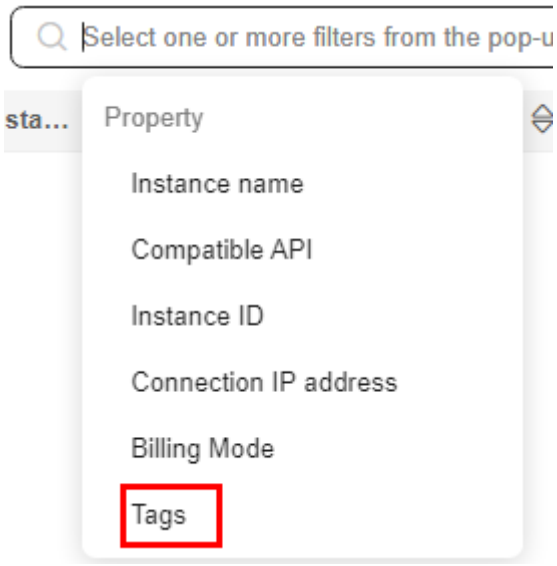
**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, select **Tags** in the search box.

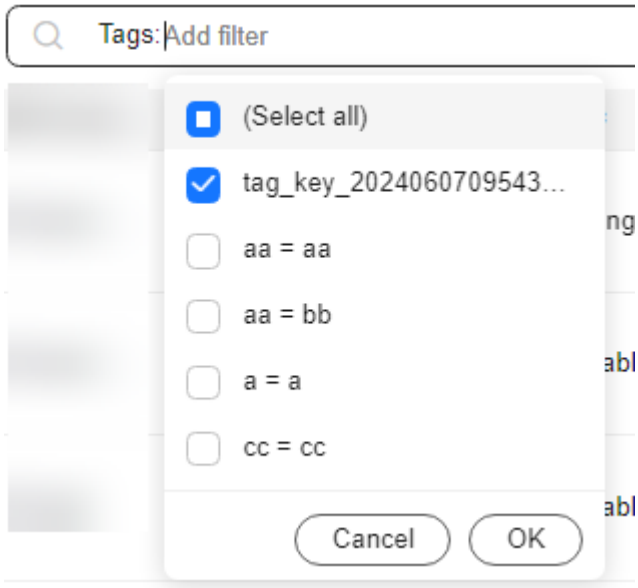
  

**Figure 4-89** Selecting tags




**Step 4** Select the tag to be queried and click **OK** to query information about instances associated with the tag.

**Figure 4-90** Searching by tag




**----End**

# 4.13 User Resource Quotas

## Scenarios

Quotas are enforced for service resources on the platform to prevent unforeseen spikes in resource usage. Quotas limit the number or amount of resources

available to users, for example, the maximum number of GeminiDB instances that you can create.

If a quota cannot meet your needs, apply for a higher quota.

## Checking Quotas

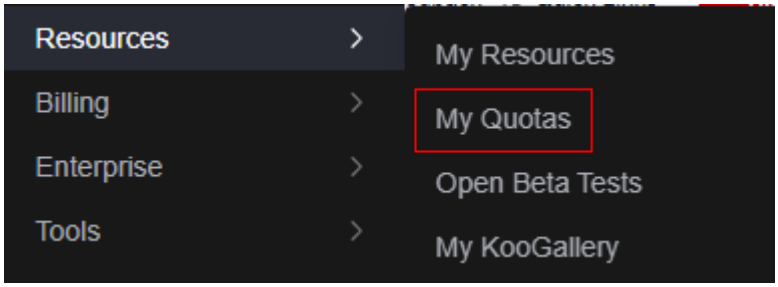
**Step 1** Log in to the Huawei Cloud console.

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** Click in the upper left corner and select a region and project.

**Step 4** In the upper right corner, choose **Resources** > **My Quotas**.

**Figure 4-91** My quotas




**Step 5** On the displayed page, check the used and total quotas of each type of GeminiDB instance resources.

**----End**

## Increasing Quotas

**Step 1** Log in to the Huawei Cloud console.

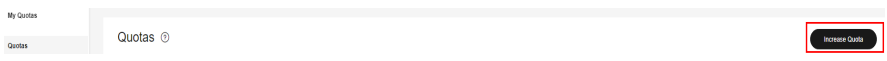
**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** Click in the upper left corner and select a region and project.

**Step 4** In the upper right corner, choose **Resources** > **My Quotas**.

**Step 5** In the upper right corner of the page, click **Increase Quota**.

**Figure 4-92** Increasing quotas




**Step 6** On the **Create Service Ticket** page, configure parameters.

In the **Problem Description** area, describe why you need the adjustment.

**Step 7** After all mandatory parameters are configured, read and agree to the agreement and click **Submit**.

**----End**

# 5 Best Practices

## 5.1 Buying and Connecting to a GeminiDB DynamoDB-Compatible Instance

This section describes how to buy and connect to a GeminiDB DynamoDB-Compatible instance running on a Linux ECS.

- **Buying a GeminiDB DynamoDB-Compatible Instance**
- **Buying an ECS**
- **Connecting to the GeminiDB DynamoDB-Compatible Instance**
- **(Optional) Connecting to the Instance over HTTPS**

### Buying a GeminiDB DynamoDB-Compatible instance

**Step 1** **Log in to the Huawei Cloud console.**

**Step 2** In the service list, choose **Databases** > **GeminiDB**.

**Step 3** On the **Instances** page, click **Buy DB Instance**.

**Step 4** Click **Buy DB Instance**, select a billing mode, and configure instance parameters. Then, click **Next** and complete subsequent operations.

**Figure 5-1** Basic information

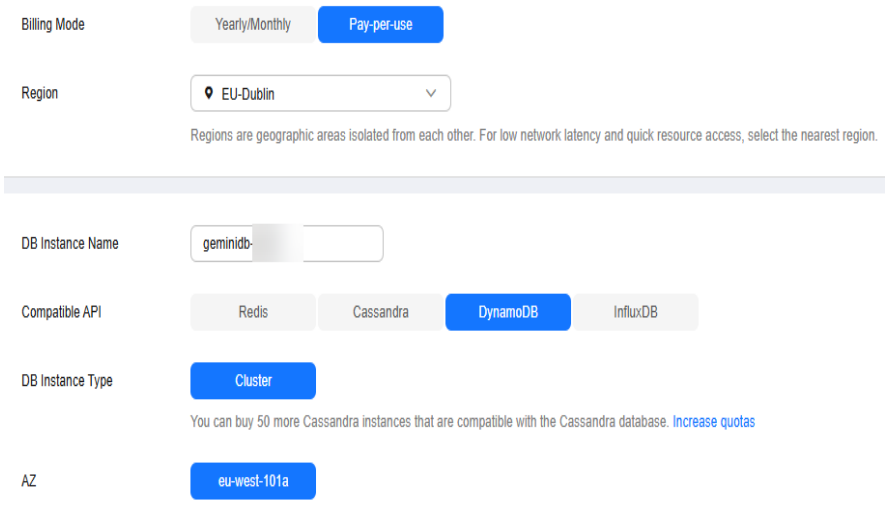
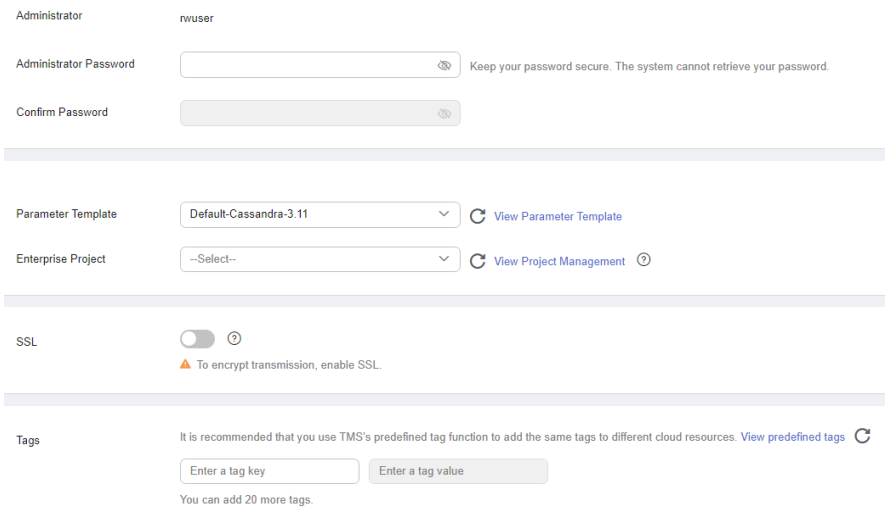
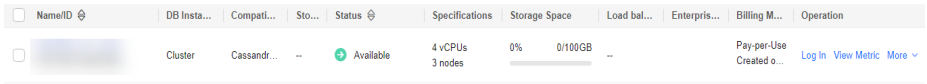



**Figure 5-2** Setting a password




**Step 5** View the purchased GeminiDB DynamoDB-Compatible instance.

**Figure 5-3** Available instance




**----End**

## Buying an ECS

**Step 1** **Log in to the Huawei Cloud console.**

**Step 2** In the service list, choose **Compute** > **Elastic Cloud Server**. On the Elastic Cloud Server console, click **Buy ECS**.

**Step 3** Configure basic settings and click **Next: Configure Network**. The ECS and GeminiDB DynamoDB-Compatible instance must be deployed in the same region, AZ, VPC, and security group.

**Figure 5-4** Basic settings

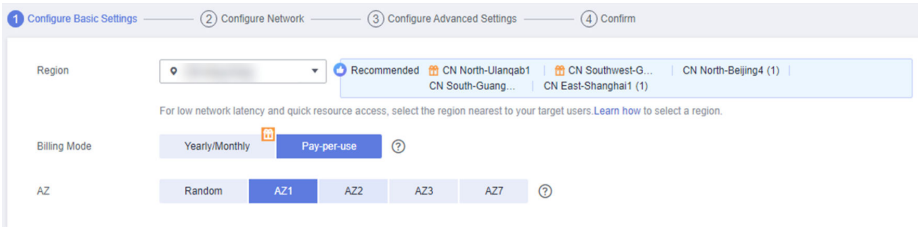



**Figure 5-5** Selecting specifications

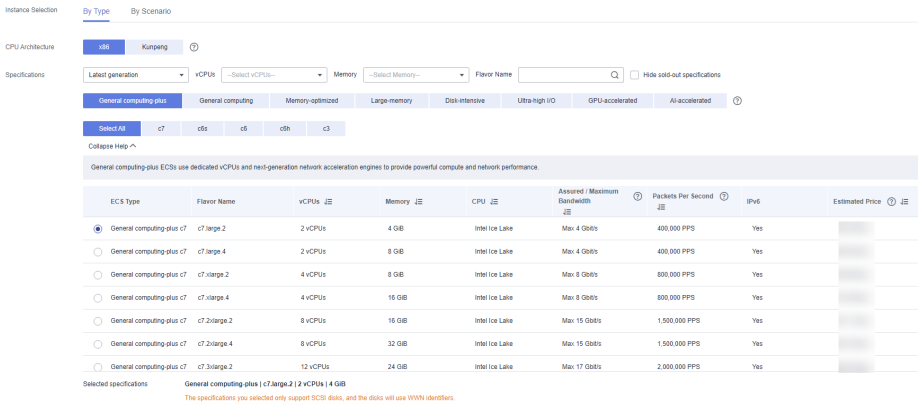
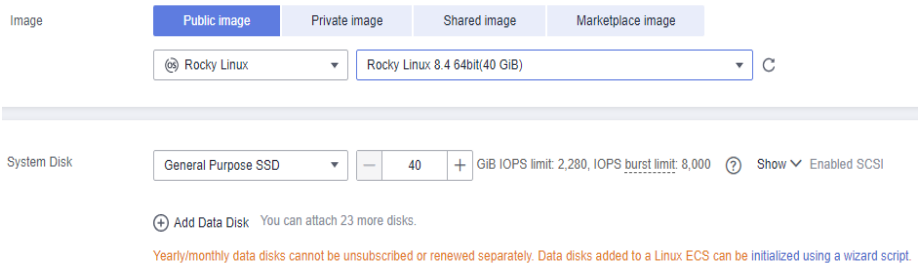



**Figure 5-6** Selecting an image




**Step 4** Configure the network and click **Next: Configure Advanced Settings**. The ECS and the GeminiDB DynamoDB-Compatible instance must be in the same VPC and security group.

- If security group rules allow access from the ECS, you can connect to the instance using the ECS.
- If the security group rules do not allow access from the ECS, add an inbound rule to the security group.

**Figure 5-7** Network settings

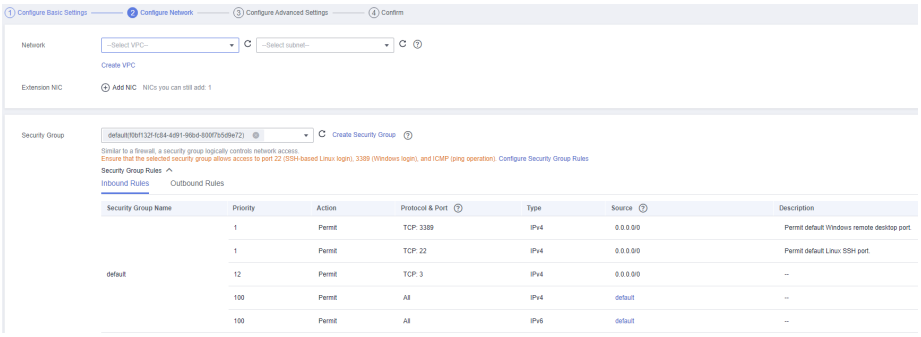
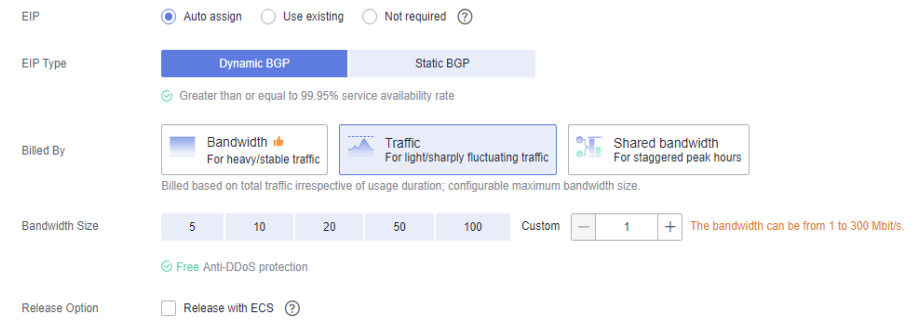
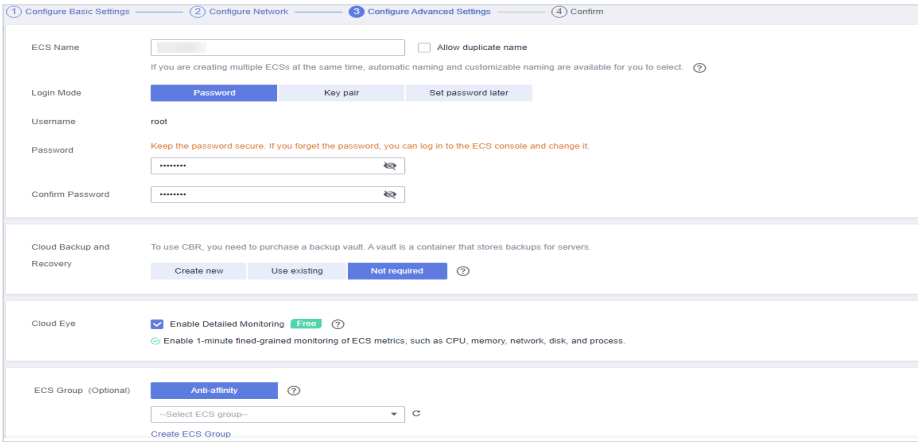



**Figure 5-8** Selecting an EIP




**Step 5** Configure a password for the ECS and click **Next: Confirm**.

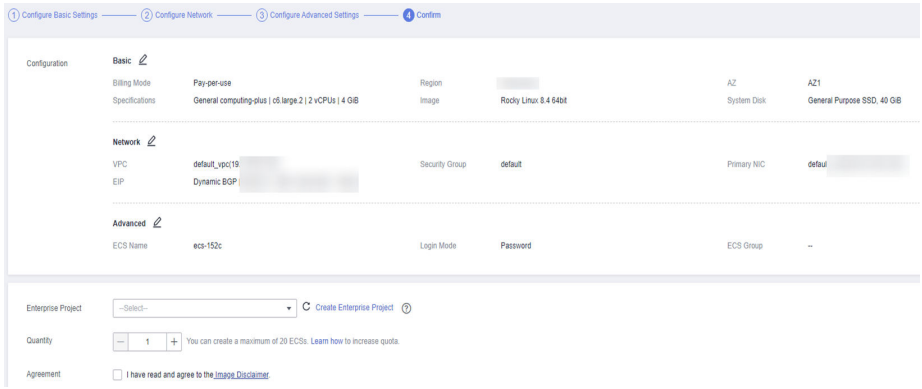
**Figure 5-9** Advanced settings




**Step 6** Confirm the configurations and click **Submit**.

**Figure 5-10** Confirming the configurations



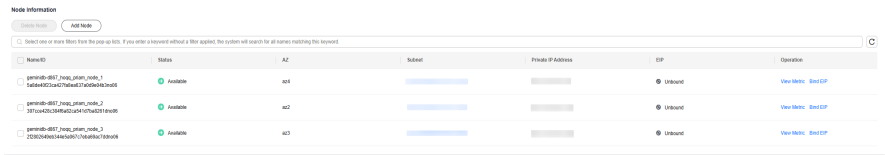

**Step 7** View the purchased ECS.

**----End**

## Connecting to the GeminiDB DynamoDB-Compatible Instance

**Step 1** In the navigation pane, choose **Connections** to obtain the private IP address and EIP of the instance.

**Figure 5-11** Viewing the IP address




**Step 2** Replace the IP address in the following code example with the IP address queried in **Step 1**.

Java code example:

```
package com.huawei.dbs.test;

import com.huawei.dbs.RequestHandler;

import com.amazonaws.SDKGlobalConfiguration;
import com.amazonaws.auth.AWSCredentialsProvider;
import com.amazonaws.auth.AWSStaticCredentialsProvider;
import com.amazonaws.auth.BasicAWSCredentials;
import com.amazonaws.services.dynamodbv2.AmazonDynamoDB;
import com.amazonaws.services.dynamodbv2.AmazonDynamoDBClientBuilder;
import com.amazonaws.services.dynamodbv2.document.DynamoDB;
import com.amazonaws.services.dynamodbv2.document.TableCollection;

import java.net.URI;


public class V1Demo {
    public static AWSCredentialsProvider myCredentials = new AWSStaticCredentialsProvider(
        new BasicAWSCredentials("your_ak", "your_sk"));
    public static String ip = "***.***.***.***";
    public static void main(String[] args) {
        disableCertificateChecks();
        AmazonDynamoDB client = AmazonDynamoDBClientBuilder.standard()
            .withRegion("region-a")
            .withRequestHandlers(new RequestHandler(URI.create("http://" + ip + "8000")))
```

```
            .withCredentials(myCredentials)
            .build();
        DynamoDB dynamoDB = new DynamoDB(client);
        TableCollection res = dynamoDB.listTables();
        System.out.println(res);
    }

}
```

**----End**

## (Optional) Connecting to the Instance over HTTPS

Prerequisites:

- The instance and ECS must be in the same VPC and subnet.
- The ECS must be accessible to the security group to which the instance belongs. For details, see **Setting Security Group Rules for a GeminiDB DynamoDB-Compatible Instance**.
- Enable SSL for the GeminiDB instance by following **Enabling SSL for a GeminiDB DynamoDB-Compatible Instance** and download a certificate by following **Downloading an SSL Certificate**. Take a Java application as an example. The command for importing the certificate is as follows:

  ```
  keytool -importcert -alias hw -file ca.cert -keystore truststore.jks -storepass password
  ```

**CAUTION**

1. For compatibility purposes, you can still use HTTP after SSL is enabled. To meet high security requirements, you can modify a parameter to disable HTTP. For details, choose **Service Tickets > Create Service Ticket** in the upper right corner of the console and contact the customer service.

2. Currently, an EIP cannot be used over HTTPS.

Add Maven dependencies. Add dependencies related to AWS SDK for Java 2.*x* to the **pom.xml** file.

```
<dependencies>
    <dependency>
        <groupId>software.amazon.awssdk</groupId>
        <artifactId>dynamodb</artifactId>
        <version>2.x.x</version>
    </dependency>
 </dependencies>
```

Connect to the GeminiDB DynamoDB-Compatible instance using DynamoDBClient over HTTPS.

Java code example:

```
import com.amazonaws.auth.AWSCredentialsProvider;
import com.amazonaws.auth.AWSStaticCredentialsProvider;
import com.amazonaws.auth.BasicAWSCredentials;
import com.amazonaws.client.builder.AwsClientBuilder;
import com.amazonaws.services.dynamodbv2.AmazonDynamoDB;
import com.amazonaws.services.dynamodbv2.AmazonDynamoDBClientBuilder;
import com.amazonaws.services.dynamodbv2.model.*;
public class SourceDemo {
// AK/SK credentials can be automatically obtained from environment variable file home/.aws/credentials.
// If there is no such a file, you need to manually specify the AK/SK.
    public static AWSCredentialsProvider myCredentials = new AWSStaticCredentialsProvider(
```

```
            new BasicAWSCredentials("your_ak", "your_sk"));
    public static void main(String[] args) {
        AmazonDynamoDB client = AmazonDynamoDBClientBuilder.standard()
                .withEndpointConfiguration(new AwsClientBuilder.EndpointConfiguration("https://
127.0.0.1:8000", "region_a"))
                .withCredentials(myCredentials)
                .build();
        System.out.println(client.listTables());
        CreateTableRequest request = new CreateTableRequest()
                .withTableName("test_001")
                .withProvisionedThroughput(new ProvisionedThroughput(1000L, 1000L))
                .withKeySchema(
                        new KeySchemaElement("id", KeyType.HASH)
                )
                .withAttributeDefinitions(
                        new AttributeDefinition("id", ScalarAttributeType.N)
                );
        System.out.println(client.createTable(request));
    }
}
```

GeminiDB DynamoDB-Compatible API is completely compatible with DynamoDB. For details about common operations, see official DynamoDB documents.

# 5.2 Security Best Practices for GeminiDB DynamoDB-Compatible API

Huawei Cloud and you share the responsibility for security. Huawei Cloud is responsible for the security of cloud services to provide a secure cloud. As a tenant, you should properly use the security capabilities provided by cloud services to protect data, and securely use the cloud. For details, see **Shared Responsibilities**.

This section provides best practices for enhancing security of GeminiDB DynamoDB-Compatible API. You can continuously evaluate the security status of your GeminiDB DynamoDB-Compatible instances and combine different security capabilities provided by GeminiDB DynamoDB-Compatible API. By doing this, data in GeminiDB DynamoDB-Compatible instances can be protected from being disclosed or tampered with.

Consider the following aspects for your security configurations:

- **Enabling SSL and Setting the CCM PCA or SSL Certificate**
- **Using a Unique CCM Certificate for Each Instance**
- **Replacing the Instance Certificate Before It Expires**

## Enabling SSL and Setting the CCM PCA or SSL Certificate

Since April 2017, GeminiDB has offered a new root certificate that has a 20-year validation period. The new certificate takes effect after DB instances are rebooted. Replace the old certificate before it expires to improve system security.

## Using a Unique CCM Certificate for Each Instance

Configuring a unique SSL certificate for each instance can significantly improve the system security and help organizations better cope with potential threats.

### Replacing the Instance Certificate Before It Expires

You can configure notification policies for the instance certificate before it expires. Renew or replace the certificate in a timely manner to avoid affecting instance functions.

## 5.3 Suggestions on Alarm Rules for GeminiDB DynamoDB-Compatible API Metrics

After setting alarm rules on the Cloud Eye console, for example, specifying monitored objects and notification policies, you can stay on top of your instance status. For details, see **Configuring Alarm Rules**.

This section describes recommended alarm rules of GeminiDB DynamoDB-Compatible API.

**Table 5-1** Suggestions on alarm rules for GeminiDB DynamoDB-Compatible API metrics

| Metric ID | Metric Name | Description | Dimension | Threshold (Raw Value) in Best Practices | Alarm Severity in Best Practices | Alarm Handling Suggestion |
|---|---|---|---|---|---|---|
| nosql005_disk_usage | Storage Space Usage | Storage usage<br>Unit: % | Instance | > 80% for 3 consecutive periods | Major | • Evaluate how much storage needs to be added based on data growth. For details, see **Manually Scaling Up Storage Space**.<br>• Enable autoscaling. For details, see **Automatically Scaling Up Storage Space**. |

| Metric ID | Metric Name | Description | Dimension | Threshold (Raw Value) in Best Practices | Alarm Severity in Best Practices | Alarm Handling Suggestion |
|---|---|---|---|---|---|---|
| cassandra001_cpu_usage | CPU Usage | CPU usage of an instance<br>Unit: % | Node | > 80% for 3 consecutive periods | Major | Upgrade CPU specifications. For details, see **Changing vCPUs and Memory**. |
| cassandra002_mem_usage | Memory Usage | Memory usage of an instance<br>Unit: % | Node | > 80% for 3 consecutive periods | Major | Upgrade memory specifications. For details, see **Changing vCPUs and Memory**. |
| cassandra015_read_latency | Average Read Latency | Average latency of database reads<br>Unit: ms | Node | > 900 ms for 3 consecutive periods | Major | Check whether the service traffic increases sharply and whether the database is normal. For details, see **Viewing Metrics**. |
| cassandra016_write_latency | Average Write Latency | Average latency of database writes<br>Unit: ms | Node | > 900 ms for 3 consecutive periods | Major | Check whether the service traffic increases sharply and whether the database is normal. For details, see **Viewing Metrics**. |

| Metric ID | Metric Name | Description | Dimension | Threshold (Raw Value) in Best Practices | Alarm Severity in Best Practices | Alarm Handling Suggestion |
|---|---|---|---|---|---|---|
| cassandra037_pending_write | Suspended Write Tasks | Write tasks waiting in the queue<br>Unit: count | Node | > 3,000 for 3 consecutive periods | Major | Check whether the service traffic increases sharply and whether the database is normal. For details, see **Viewing Metrics**. |
| cassandra038_pending_read | Suspended Read Tasks | Read tasks in the waiting status<br>Unit: count | Node | > 3,000 for 3 consecutive periods | Major | Check whether the service traffic increases sharply and whether the database is normal. For details, see **Viewing Metrics**. |

# 6 Performance White Paper

## 6.1 Performance Test Methods

This section describes performance tests of GeminiDB DynamoDB-Compatible instances, including environment, procedure, and model tests.

### Test Environment

- Region: CN North-Beijing4
- AZ: AZ 1
- Elastic Cloud Server (ECS): h3.4xlarge.2 with 16 vCPUs, 32 GB of memory, and CentOS 7.5 64-bit image
- Stress test data model: 1 KB per row
- Average CPU load during the stress test: 50
- Instance specifications: all specifications described in **Table 6-1**

**Table 6-1** Instance specifications

| No. | Specifications |
| --- | --- |
| Cluster 1 | 4 vCPUs \| 16 GB |
| Cluster 2 | 8 vCPUs \| 32 GB |
| Cluster 3 | 16 vCPUs \| 64 GB |
| Cluster 4 | 32 vCPUs \| 128 GB |

### Test Tool

YCSB is an open-source tool for testing database performance. In this test, YCSB 0.17.0 is used. For details about the load balancing policy, see the load balancing plug-in.

For details on how to use this tool, see **YCSB**.

## Test Metrics

Operations per Second (OPS): operations executed by a database per second

## Test Procedure

1. Configure the **workload** file.

   Set **readproportion**, **insertproportion**, **updateproportion**, **scanproportion**, and **readmodifywriteproportion** in the **workload** file based on **Table 6-2**.

   Set **recordcount** in the **workload** file based on **Table 6-3**.

2. Use workload-insert-mostly as an example. Run the following command to prepare test data:

   ```
   nohup ./bin/ycsb load dynamodb -s -P workloads/workload-insert-mostly -P dynamodb.properties -p operationcount=400000000 -p recordcount=400000000 -threads ${thread} -p maxexecutiontime=${maxexecutiontime} -s 1> data_load.log 2>&1 &
   ```

3. Use workload-insert-mostly as an example. Run the following command to test performance:

   ```
   nohup ./bin/ycsb run dynamodb -s -P workloads/workload-insert-mostly -P dynamodb.properties -p operationcount=9000000 -p recordcount=9000000 -threads ${thread} -s 1> workload-insert-mostly_run.log 2>&1 &
   ```

## Test Models

- Workload model

  **Table 6-2** Workload models

  | Workload Model | Test Models |
  | --- | --- |
  | workload-read-mostly | 100% read |
  | workload-write-mostly | 100% insert |
  | workload-read-write-combination | 50% update, 50% read |

- Data model

  fieldlength = 100, fieldcount = 10

- Preset data volume

  Different preset data volumes were used to test performance of instances of each type of specifications.

  The following table describes the preset data volumes.

  **Table 6-3** Preset data volumes

  | No. | Specifications | Preset Data Volume |
  | --- | --- | --- |
  | Cluster 1 | 4 vCPUs \| 16 GB | 200 GB |
  | Cluster 2 | 8 vCPUs \| 32 GB | 500 GB |
  | Cluster 3 | 16 vCPUs \| 64 GB | 1,000 GB |

| No. | Specifications | Preset Data Volume |
| --- | --- | --- |
| Cluster 4 | 32 vCPUs \| 128 GB | 2,000 GB |

# 6.2 Performance Test Data

OPS of instances with different specifications can be tested using different service models with the same preset data volume. For details, see the numbers in bold in **Table 6-4**.

**Table 6-4** Test data

| Node Specifications | | 4 vCPUs \| 16 GB | 8 vCPUs \| 32 GB | 16 vCPUs \| 64 GB | 32 vCPUs \| 128 GB |
| --- | --- | --- | --- | --- | --- |
| Preset Data Volume | | 200 GB | 500 GB | 1,000 GB | 2,000 GB |
| Test Model No. | workload-read-mostly | **10690** | **23872** | **59169** | **104149** |
| | workload-write-mostly | **13987** | **35183** | **71469** | **130171** |
| | workload-read-write-combination | **11694** | **30528** | **66517** | **119160** |

NOTE

- Operations per Second (OPS): operations executed by a database per second
- Test Model No.: test model sequence number. **Table 6-5** lists the test models.

**Table 6-5** Test models

| Test Model No. | Test Model |
| --- | --- |
| More read requests than write requests workload-read-mostly | 95% read, 5% update |
| Balanced read and write requests workload-read-write-combination | 50% update, 50% read |
| Balanced read and rewrite requests workload-read-modify-write | 50% read, 50% readmodifywrite |
| Read, update, and write requests workload-mixed-operational-analytical | 65% read, 25% update, 10% insert |
| More write requests than read requests workload-insert-mostly | 90% insert, 10% read |

# 7 FAQs

## 7.1 Product Consulting

### 7.1.1 What Are the Precautions for Using GeminiDB DynamoDB-Compatible API?

1. DB instance operating systems (OSs) are invisible to you. Your applications can access a database only through an IP address and a port.
2. Backup files stored in OBS and system containers used by GeminiDB DynamoDB-Compatible API are invisible to you. They are visible only to the background management system.
3. Precautions after purchasing DB instances:

   After purchasing DB instances, you do not need to perform basic database O&M operations, such as applying HA and security patches, but you should still note:

   a. The CPU, input/output operations per second (IOPS), and space are sufficient for the DB instances.
   b. The DB instance has performance problems and whether optimization is required.

### 7.1.2 What Is GeminiDB DynamoDB-Compatible Instance Availability?

The formula for calculating the instance availability is as follows:

DB instance availability = (1 – Failure duration/Total service duration) × 100%

The failure duration refers to the total duration of faults that occur during the running of an instance after you buy the instance. The total service duration refers to the total running time of the instance.

## 7.2 Billing

### 7.2.1 What Are the Differences Between Yearly/Monthly and Pay-per-Use Billing Modes?

Yearly/Monthly is a prepaid mode. If your future usage is predictable, this billing mode is generally less expensive than pay-per-use. Longer subscriptions offer larger discounts.

Pay-per-use is a postpaid mode. You are only billed for how long you have actually used your instance. This mode can be a good option when future requirements are unpredictable. Pay-per-use instances are priced by the hour, but if an instance is used for less than one hour, you will be billed based on the actual duration.

### 7.2.2 Can I Switch Between Yearly/Monthly and Pay-per-Use Billing Modes?

Yes.

- To change a yearly/monthly instance to pay-per-use, see **Changing a Yearly/Monthly Instance to Pay-per-Use**.
- To change a pay-per-use instance to yearly/monthly, see **Changing a Pay-per-Use Instance to Yearly/Monthly**.

## 7.3 Instructions for Use

### 7.3.1 How Can I Use TTL?

Time to live (TTL) of GeminiDB DynamoDB-Compatible API is used to periodically delete unnecessary data rows. You can specify an expiration timestamp for each row in a table. GeminiDB DynamoDB-Compatible API automatically deletes the rows within a period of time after they are expired.

#### Configuring TTL

You can call the UpdateTimeToLive API to enable or disable table-level TTL and set a column name for table-level TTL. The column stores the expiration timestamps.

```
// Enables TTL for a specified table and specifies TTL timestamps for a column.
UpdateTimeToLiveRequest request1 = new UpdateTimeToLiveRequest()
        .withTableName("$tableName")
        .withTimeToLiveSpecification(new TimeToLiveSpecification()
            .withAttributeName("$ttlAttr")
            .withEnabled(true)
        );
// Updates values in the TTL column through APIs such as UpdateItem.
UpdateItemRequest request2 = new UpdateItemRequest()
        .withTableName("$tableName")
        .withKey(keyMap)
        .withUpdateExpression("SET $ttlAttr = :$ttlVal")
        .withExpressionAttributeValues(valueMap);
// Disables TTL for a specified table.
UpdateTimeToLiveRequest request1 = new UpdateTimeToLiveRequest()
        .withTableName("$tableName")
        .withTimeToLiveSpecification(new TimeToLiveSpecification()
            .withEnabled(false)
        );
```

You can call the DescribeTimeToLive API to query table-level TTL configurations, including whether TTL is enabled and the TTL column name.

```
// Queries TTL configurations of a specified table.
DescribeTimeToLiveRequest request1 = new DescribeTimeToLiveRequest().withTableName("$tableName");
```

**NOTE**

- The TTL column must store the Number (*N*) data type.
- TTL column values are converted to Unix timestamps (in seconds) in the GMT+0 time zone and compared with the current time to determine the expiration status.
- TTL runs in the background as a scheduled task. Therefore, the actual deletion time is different from the configured expiration timestamp. The offset is related to the task execution period.
- All user tables with TTL enabled are scanned in the background. You can also disable TTL for a table.
- TTL of GeminiDB DynamoDB-Compatible API is compatible with DynamoDB Streams. Rows that exceed the TTL value are inserted into Dynamo Streams records at the time when they are actually deleted.

# 7.4 Database Connection

## 7.4.1 How Can I Create and Connect to an ECS?

1. To create an ECS, see *Elastic Cloud Server User Guide*.
   - The ECS is used to connect to a GeminiDB instance DynamoDB-Compatible instance. They must be in the same VPC and subnet.
   - Configure the security group rules to allow the ECS to access to the instance.
2. To connect to an ECS, see "Logging in to an ECS" *Getting Started with Elastic Cloud Server User Guide*.

## 7.4.2 Can I Change the VPC After Buying a GeminiDB DynamoDB-Compatible Instance?

After a GeminiDB DynamoDB-Compatible instance is created, its VPC cannot be changed.

However, you can change a VPC by restoring the full backup of your instance to the VPC you want to use. For details, see **Restoring a Backup to a New Instance**.

# 7.5 Backup and Restoration

## 7.5.1 How Long Can GeminiDB DynamoDB-Compatible Instance Backups Be Stored?

The validity period of automated backups of a GeminiDB DynamoDB-Compatible instance is the same as the backup retention period you set. There is no limit for the manual backup retention period. You can delete manually backup files as needed.

# 7.6 Regions and AZs

## 7.6.1 What Is AZ and How Can I Select an AZ?

### AZ

An AZ is a part of a physical region with its own independent power supply and network. An AZ is generally an independent physical equipment room, ensuring independence of the AZ.

Each region contains multiple AZs. If one AZ becomes faulty, the other AZs in the same region can continue to provide services.

AZs in the same region can communicate with each other over an intranet.

### Selecting an AZ

You can deploy your instances in different AZs for high availability. If one of an AZ becomes faulty, databases in other AZs will not be affected. When selecting AZs:

- If only one AZ is available in a region, there is only one AZ in the region.
- The AZ of a purchased DB instance cannot be changed.
- The AZs in one region can communicate with each other over an intranet.

For more information, see **Regions and AZs**.

## 7.6.2 Can Different AZs Communicate with Each Other?

An AZ is a part of a physical region with its own independent power supply and network. An AZ is generally an independent physical equipment room, ensuring independence of the AZ.

Each region contains multiple AZs. If one AZ becomes faulty, the other AZs in the same region can continue to provide services normally.

By default, different AZs in the same VPC can communicate with each other through an internal network.

For more information, see **Regions and AZs**.

## 7.6.3 Can I Change the Region After Buying a GeminiDB DynamoDB-Compatible Instance?

No. After an instance is created, its region cannot be changed.

# 7.7 Instance Freezing, Release, Deletion, and Unsubscription

## Why Is My GeminiDB DynamoDB-Compatible Instance Released?

If your subscriptions have expired but not been renewed, or you are in arrears due to insufficient balance, your instances enter a grace period. If you do not renew your subscription or top up your account before the grace period ends, your instance will enter a retention period, during which the resources will be suspended. If you still do not renew them or top up your account after the retention period ends, your instances will be released and your data stored will be deleted. For details, see **Resource Suspension and Release**.

## Why Is My GeminiDB DynamoDB-Compatible Instance Frozen?

Your instances may be frozen for a variety of reasons. The most common reason is that you are in arrears.

## Can I Back Up Data After My Instance Is Frozen?

No. If your instance is frozen due to arrears, you need to unfreeze it first.

## How Do I Unfreeze My Instance?

Frozen due to arrears: You can renew your resources or top up your account. You can renew, release, or delete instances frozen due to arrears. You cannot unsubscribe from expired yearly/monthly GeminiDB DynamoDB-Compatible instances.

## What Happens After My Resources Are Frozen, Unfrozen, or Released?

- After an instance is frozen:
  - It cannot be accessed, and your services will be interrupted. For example, if your instance is frozen, it cannot be connected.
  - If they are yearly/monthly resources, no changes can be made to them.
  - It can be unsubscribed from or deleted manually.
- After it is unfrozen, you can connect to it again.
- When resources are released, the instance is deleted. Before the deletion, the system determines whether to **move the instance to the recycle bin** based on the recycling policy you specified.

## How Do I Renew My Instance?

After a yearly/monthly GeminiDB DynamoDB-Compatible instance expires, you can renew it by following **Renewal Management**.

## Can I Restore Released or Unsubscribed Resources?

If a deleted instance is moved to the recycle bin, you can restore it by following **Recycling an Instance**. If the recycling policy is not enabled, you cannot restore it.

When you unsubscribe from an instance, confirm the instance information carefully. If you have unsubscribed from an instance by mistake, purchase a new one.

## How Do I Delete a GeminiDB DynamoDB-Compatible Instance?

- To delete a pay-per-use instance, see **Deleting a Pay-per-Use Instance**.
- To delete a yearly/monthly instance, see **How Do I Unsubscribe from a Yearly/Monthly Instance?**.