ALIBABA CLOUD

Alibaba Cloud

云原生关系型数据库PolarDB O引擎

Management Guide

Document Version: 20220704

Alibaba Cloud

Legal disclaimer

Alibaba Cloud reminds you to carefully read and fully understand the terms and conditions of this legal disclaimer before you read or use this document. If you have read or used this document, it shall be deemed as your total acceptance of this legal disclaimer.

1. You shall download and obtain this document from the Alibaba Cloud website or other Alibaba Cloud-authorized channels, and use this document for your own legal business activities only. The content of this document is considered confidential information of Alibaba Cloud. You shall strictly abide by the confidentiality obligations. No part of this document shall be disclosed or provided to any third party for use without the prior written consent of Alibaba Cloud.

2. No part of this document shall be excerpted, translated, reproduced, transmitted, or disseminated by any organization, company or individual in any form or by any means without the prior written consent of Alibaba Cloud.

3. The content of this document may be changed because of product version upgrade, adjustment, or other reasons. Alibaba Cloud reserves the right to modify the content of this document without notice and an updated version of this document will be released through Alibaba Cloud-authorized channels from time to time. You should pay attention to the version changes of this document as they occur and download and obtain the most up-to-date version of this document from Alibaba Cloud-authorized channels.

4. This document serves only as a reference guide for your use of Alibaba Cloud products and services. Alibaba Cloud provides this document based on the "status quo", "being defective", and "existing functions" of its products and services. Alibaba Cloud makes every effort to provide relevant operational guidance based on existing technologies. However, Alibaba Cloud hereby makes a clear statement that it in no way guarantees the accuracy, integrity, applicability, and reliability of the content of this document, either explicitly or implicitly. Alibaba Cloud shall not take legal responsibility for any errors or lost profits incurred by any organization, company, or individual arising from download, use, or trust in this document. Alibaba Cloud shall not, under any circumstances, take responsibility for any indirect, consequential, punitive, contingent, special, or punitive damages, including lost profits arising from the use or trust in this document (even if Alibaba Cloud has been notified of the possibility of such a loss).

5. By law, all the contents in Alibaba Cloud documents, including but not limited to pictures, architecture design, page layout, and text description, are intellectual property of Alibaba Cloud and/or its affiliates. This intellectual property includes, but is not limited to, trademark rights, patent rights, copyrights, and trade secrets. No part of this document shall be used, modified, reproduced, publicly transmitted, changed, disseminated, distributed, or published without the prior written consent of Alibaba Cloud and/or its affiliates. The names owned by Alibaba Cloud shall not be used, published, or reproduced for marketing, advertising, promotion, or other purposes without the prior written consent of Alibaba Cloud. The names owned by Alibaba Cloud include, but are not limited to, "Alibaba Cloud", "Aliyun", "HiChina", and other brands of Alibaba Cloud and/or its affiliates, which appear separately or in combination, as well as the auxiliary signs and patterns of the preceding brands, or anything similar to the company names, trade names, trademarks, product or service names, domain names, patterns, logos, marks, signs, or special descriptions that third parties identify as Alibaba Cloud and/or its affiliates.

6. Please directly contact Alibaba Cloud for any errors of this document.

Document conventions

Style	Description	Example
Danger	A danger notice indicates a situation that will cause major system changes, faults, physical injuries, and other adverse results.	**Danger:** Resetting will result in the loss of user configuration data.
Warning	A warning notice indicates a situation that may cause major system changes, faults, physical injuries, and other adverse results.	**Warning:** Restarting will cause business interruption. About 10 minutes are required to restart an instance.
Notice	A caution notice indicates warning information, supplementary instructions, and other content that the user must understand.	**Notice:** If the weight is set to 0, the server no longer receives new requests.
Note	A note indicates supplemental instructions, best practices, tips, and other content.	**Note:** You can use Ctrl + A to select all files.
>	Closing angle brackets are used to indicate a multi-level menu cascade.	Click **Settings**> **Network**> **Set network type**.
Bold	Bold formatting is used for buttons , menus, page names, and other UI elements.	Click **OK**.
`Courier font`	Courier font is used for commands	Run the `cd /d C:/window` command to enter the Windows system folder.
Italic	Italic formatting is used for parameters and variables.	`bae log list --instanceid` *Instance_ID*
[] or [a|b]	This format is used for an optional value, where only one item can be selected.	`ipconfig [-all|-t]`
{} or {a|b}	This format is used for a required value, where only one item can be selected.	`switch {active|stand}`

Table of Contents

1.Overview 07
2.Comparison with Oracle on database management 08
3.Configure a whitelist for a cluster 12
4.Connect to PolarDB 16
4.1. View or apply for an endpoint 16
4.2. Modify or delete an endpoint 18
4.3. Connect to a PolarDB cluster 20
4.4. Private domain names 26
5.Cluster endpoint 28
5.1. Overview 28
5.2. Features 29
5.2.1. Read/write splitting 29
5.2.2. Consistency levels 31
5.2.3. Transaction splitting 33
5.3. Create and modify a custom cluster endpoint 34
5.4. Customize a routing rule for requests that contain a speci... 37
5.5. FAQ 39
6.Cluster management 41
6.1. Create a cluster 41
6.2. Perform a temporary upgrade 44
6.3. Change the specifications of a PolarDB cluster 46
6.4. Add or remove a read-only node 49
6.5. Set a maintenance window 51
6.6. Restart nodes 52
6.7. Release a cluster 53
6.8. Cluster lock feature 54

6.9. Automatic failover and manual failover 56
6.10. Upgrade the minor version 57
6.11. Deploy a cluster across zones and change the primary zo... 58
7.Account management 62
7.1. Account overview 62
7.2. Register and log on to an Alibaba Cloud account 63
7.3. Create and authorize a RAM user 63
7.4. Create database accounts 64
7.5. Manage database accounts 66
8.DBLink 68
8.1. Overview 68
8.2. Create a database link from PolarDB for Oracle to PolarDB.. 68
8.3. Create a database link from PolarDB for Oracle to Postgre... 70
8.4. Use a database link to query data across databases 72
8.5. Delete a database link 73
9.Database 75
10.Backup and restoration 77
10.1. Overview 77
10.2. Billing 79
10.3. Backup methods 80
10.3.1. Configure a backup policy 80
10.3.2. Backup method 1: Automatic backup 83
10.3.3. Backup method 2: Manual backup 84
10.4. Restoration methods 85
10.4.1. Restoration method 1: Restore data to a specific point... 85
10.4.2. Restoration method 2: Restore data from a backup se... 87
10.5. FAQ 89
11.Cluster Recycle 92

11.1. Pricing 92
11.2. Restore a released cluster 92
11.3. Delete a released cluster 95
12.Data Security and Encryption 96
12.1. Configure SSL encryption 96
12.2. Configure TDE 100
13.Diagnostics and optimization 103
13.1. SQL Explorer 103
13.2. Performance monitoring 106
13.3. Create an alert rule 107
13.4. Manage alert rules 109
13.5. Performance insight 109
14.Configuration parameters 111
14.1. polar_comp_redwood_date 111
14.2. polar_comp_redwood_raw_names 111
14.3. polar_comp_redwood_strings 112
14.4. polar_comp_stmt_level_tx 113
14.5. polar_create_table_with_full_replica_identity 114
14.6. Custom parameters 115
14.7. Configure cluster parameters 117
15.Version Management 119
16.SQL firewalls 121
17.More operations 125
17.1. Clone a cluster 125
17.2. View the database storage usage 127
17.3. View or cancel a scheduled task 127

1.Overview

is a new-generation that is developed by Alibaba Group. This service decouples computing from storage and uses integrated software and hardware. PolarDB is a secure and reliable database service that provides auto scaling, high performance, and mass storage. PolarDB is fully compatible with MySQL 5.6, MySQL 5.7, MySQL 8.0, and PostgreSQL 11. PolarDB is highly compatible with Oracle.

uses an architecture that decouples computing from storage. All compute nodes share one set of data. PolarDB allows you to upgrade or downgrade specifications within minutes, and supports disaster recovery within seconds. PolarDB ensures that global data is consistent and offers data backup and disaster recovery for free. provides the benefits of both commercial databases and open source cloud databases. The benefits of commercial databases include stability, reliability, high performance, and scalability. The benefits of open source cloud databases include ease of use, openness, and self iteration.

Terms

- Cluster

  A PolarDB cluster of the contains one primary node and a maximum of 15 read-only nodes. A minimum of one read-only node is required to provide high availability in active-active mode. If a cluster ID starts with `pc`, the cluster is a cluster.

- Node

  A node is a database service process that exclusively occupies physical memory. If a node ID starts with `pi`, the node is a instance.

- Database

  A database is a logical unit that is created on a node. You can create multiple databases on a node. The name of each database on the node must be unique.

- Region and zone

  A region is a geographic area where data centers can be deployed. A zone is a geographic area in a region. This area has an independent power supply and network. For more information, see Global infrastructure of Alibaba Cloud.

Console

Alibaba Cloud provides an easy-to-use web console to help you manage various Alibaba Cloud services, including the cloud database service . In the console, you can create, connect to, and configure databases.

Click the following link to log on to the console: PolarDB console

2.Comparison with Oracle on database management

This topic describes the differences between and native Oracle on database management from several aspects.

Databases

- By default, one Oracle instance of the versions earlier than Oracle Database 12c has only one database. The versions later than Oracle Database 12c provide the multitenancy concept. Each container database (CDB) can include multiple pluggable databases (PDBs).
- An cluster corresponds to an Oracle instance. You can create multiple databases for each database cluster.

Users

- Similarities:

  Both and Oracle have the user concept. Users are owners of database objects and have access to databases.

- Differences:

  An Oracle user can log on to a database only after this user is granted the CREATE SESSION privilege. By default, users are granted the LOGIN privilege to log on to a database.

The following syntax that is used to create a user is available:

- Oracle syntax:

```
CREATE USER user
   IDENTIFIED { BY password
              | EXTERNALLY [ AS 'certificate_DN' ]
              | GLOBALLY [ AS '[ directory_DN ]' ]
              }
   [ DEFAULT TABLESPACE tablespace
   | TEMPORARY TABLESPACE
        { tablespace | tablespace_group_name }
   | QUOTA size_clause
           | UNLIMITED
           }
           ON tablespace
     [ QUOTA size_clause
             | UNLIMITED
             }
             ON tablespace
     ]...
   | PROFILE profile
   | PASSWORD EXPIRE
   | ACCOUNT { LOCK | UNLOCK }
     [ DEFAULT TABLESPACE tablespace
     | TEMPORARY TABLESPACE
          { tablespace | tablespace_group_name }
     | QUOTA size_clause
             | UNLIMITED
             }
             ON tablespace
       [ QUOTA size_clause
               | UNLIMITED
               }
               ON tablespace
       ]...
     | PROFILE profile
     | PASSWORD EXPIRE
     | ACCOUNT { LOCK | UNLOCK }
     ]...
  ] ;
```

- syntax:

```
CREATE USER|ROLE name [[WITH] option […]] [IDENTIFIED BY password]
where option can be the following compatible clauses:
 PROFILE profile_name
 | ACCOUNT {LOCK|UNLOCK}
 | PASSWORD EXPIRE [AT 'timestamp']
or option can be the following non-compatible clauses:
 | LOCK TIME 'timestamp'
```

is also compatible with the syntax of CREATE USER of PostgreSQL. For more information, see SQL Commands.

Roles

- In Oracle, a role is a group of privileges and cannot be regarded as an owner of database objects. This role cannot be granted privileges of other roles and does not have access to databases.
- In , a role is equivalent to a user. This role can be regarded as an owner of database objects, can be granted privileges of other roles or users, and can have access to databases. By default, a role does not have the LOGIN privilege compared with a user. A role can access databases only after the role is granted the LOGIN privilege. A user has the LOGIN privilege.

The following syntax that is used to create a role is available:

- Oracle syntax:

```
CREATE ROLE role
   [ NOT IDENTIFIED
   | IDENTIFIED { BY password
                | USING [ schema. ] package
                | EXTERNALLY
                | GLOBALLY
                }
   ] ;
```

- syntax:

  The syntax of CREATE ROLE is consistent with that of CREATE USER.

Schemas

- Similarities:

  A schema is a logical concept that represents a collection of database objects, such as tables, indexes, and views. These objects are also called schema objects.

- Differences:

Oracle	
You cannot separately create a schema.	You can execute the CREATE SCHEMA statement to create a schema.
When you create a database user, the system automatically creates a schema that has the same name as the username.	Each database has a default schema that is named PUBLIC. You can use `SET SEARCH_PATH TO 'xxx ';` to modify the current default schema.

Privileges

The privileges of are similar to those of Oracle. The privileges are divided into system privileges and object privileges.

- System privileges
  - Oracle

    System privileges allow you to perform specific actions, such as CREATE USER, CREATE TABLE, and CREATE TABLESPACE.

    System privileges also include some administrative rights:

    - SYSDBA and SYSOPER: have the privileges of almost all the database objects. You are authorized to perform some standard database operations, such as starting and shutting down databases, creating server parameter files (SPFILEs) of a database, and changing database archived logs.
    - SYSBACKUP: performs backup and restoration operations.
    - SYSDG: performs the Data Guard operations.
    - SYSKM: manages transparent data encryption (TDE) wallets.
    - SYSRAC: performs the operations on Oracle Real Application Clusters (RACs).

  - supports multiple system privileges, such as LOGIN, POLAR_SUPERUSER, CREATEDB, and CREATEROLE. When you execute the CREATE ROLE or CREATE USER statement, you can specify whether the user has the corresponding privileges.

- Object privileges

  Object privileges are the privileges to perform operations on specified objects. Database objects include tables, views, sequences, large objects, schemas, functions, and procedural language. Object privileges include SELECT, INSERT, UPDATE, DELETE, ALTER, INDEX, REFERENCES, and EXECUTE. The object privilege varies based on the object type.

  - Oracle

    Users are granted privileges on all the objects for a schema.

- Only the object owner and the superuser are authorized to modify or delete objects.

  > **Note** A superuser is a user who has the POLAR_SUPERUSER privilege.

Monitoring and O&M

- Oracle

  For more information, see Documentation at the Oracle official website.

- Metric monitoring and log monitoring are supported.
  - Metric monitoring: includes performance monitoring, alerts, and performance insights. For more information, see Performance monitoring and Performance insight.
  - Log monitoring: includes slow query logs and SQL Explorer. For more information, see SQL Explorer.

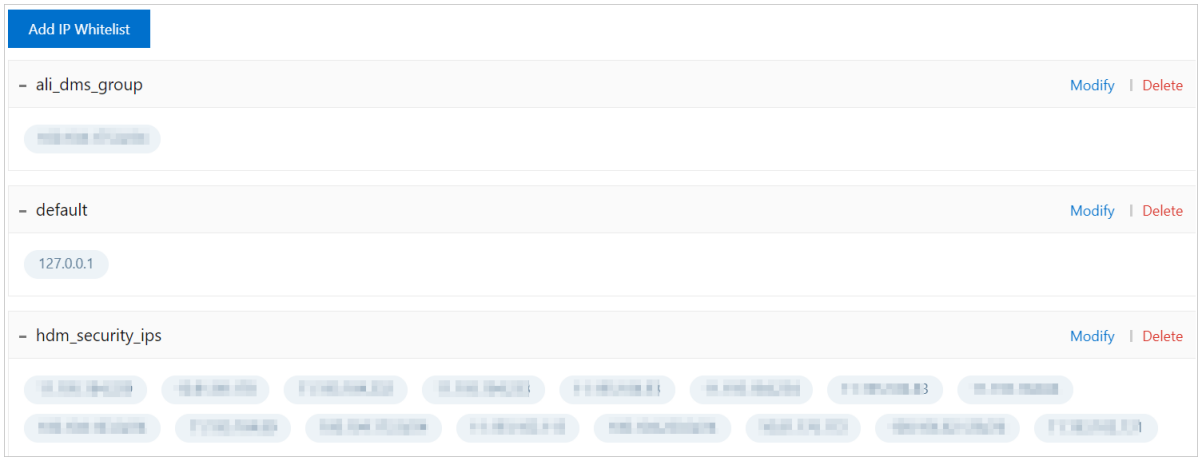
3.Configure a whitelist for a cluster

After you create a cluster, you must add IP addresses to a whitelist for the cluster and create an initial account to access and manage the cluster.

Notes

- cannot automatically obtain the private IP addresses of ECS instances in virtual private clouds (VPCs). If you want to use the private IP address of an ECS instance to access a cluster, you must manually add the private IP address to the IP whitelist of the cluster.
- The **ali_dms_group** (for Data Management), **hdm_security_ips** (for Database Autonomy Service), and **dtspolardb** (for Data Transmission Service) whitelists are automatically created when you use the relevant services. To ensure that the services can be used as normal, do not modify or delete these IP whitelists.

> **Notice** Do not add your service IP addresses to these IP whitelists. Otherwise, your service IP addresses may be overwritten when the related services are updated. Consequently, service interruption may occur.

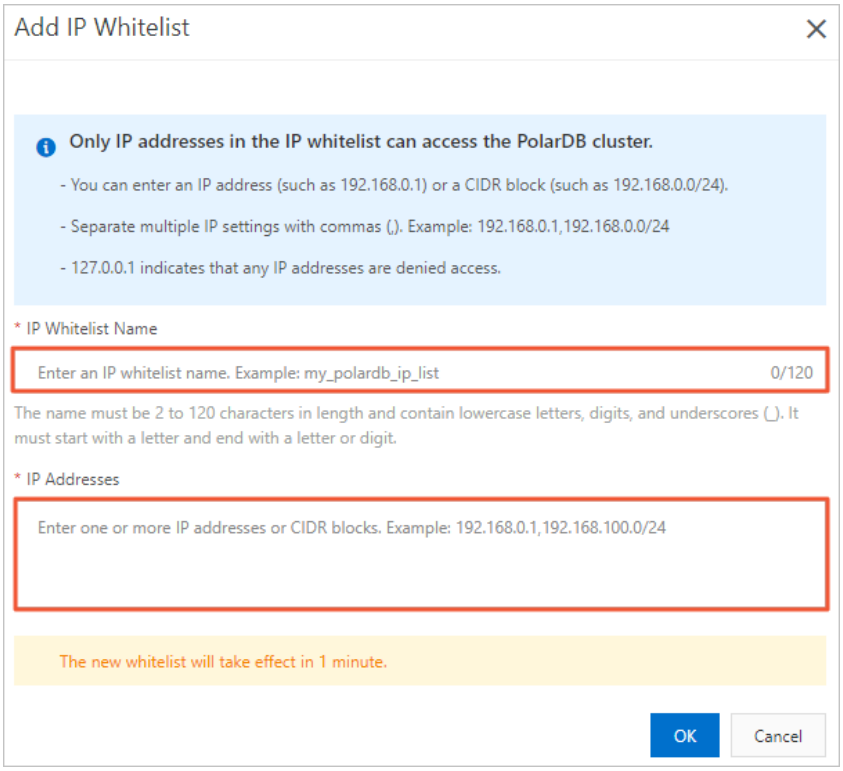


Configure a whitelist

1.
2.
3.
4. In the left-side navigation pane, choose **Settings and Management > Whitelists**.
5. On the **Whitelists** page, you can click **Add IP Whitelist** to add an IP whitelist or click **Modify** to modify an existing IP whitelist.

   - Add an IP whitelist
     a. Click **Add IP Whitelist**.

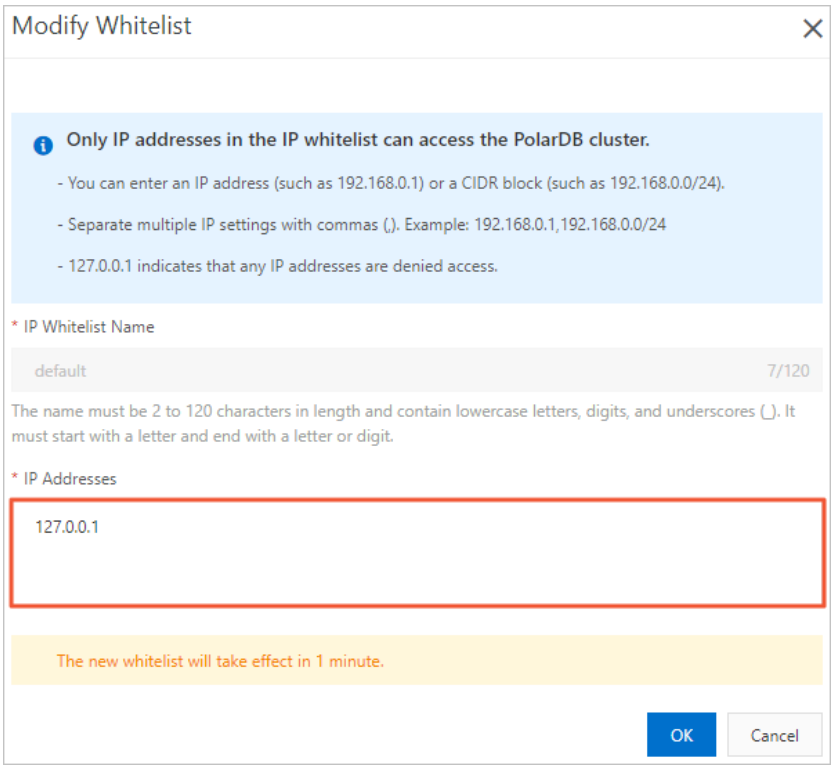
b. In the **Add IP Whitelist** panel, specify the name of the IP whitelist and enter the IP addresses that are allowed to access the cluster.



> **Note** The name of the IP whitelist must meet the following requirements:
> - The name can contain lowercase letters, digits, and underscores (_).
> - The name must start with a letter and end with a letter or digit.
> - The name must be 2 to 120 characters in length.

- Modify an IP whitelist
  a. On the right side of an IP whitelist name, click **Modify**.

b. In the **Modify Whitelist** panel, enter the IP addresses that are allowed to access the cluster.



> **Note**
> - A `default` IP whitelist that contains only the IP address `127.0.0.1` is automatically created for each cluster. This IP whitelist blocks all IP addresses.
> - If you set an IP whitelist to a percent sign (`%`) or `0.0.0.0/0`, all IP addresses are allowed to access the cluster. We recommend that you do not use this configuration unless necessary because it compromises database security.

6. Click **OK**.

> **Note** You can create at most 50 IP whitelists and add at most 1,000 IP addresses or CIDR blocks to the 50 IP whitelists.

What to do next

After you configure whitelists and create database accounts, you can connect to the cluster and manage the databases.

- Create database accounts
- Connect to a PolarDB cluster

FAQ

- How can I allow a server to access only a specified node in a cluster?

  You can use the custom cluster endpoint feature. This feature allows a server to access only a specified node in a cluster.

- What is the maximum number of IP addresses in all the IP whitelists?

  You can add a maximum of 1,000 entries to the IP whitelists. Each entry can be an IP address or a CIDR block.

- After I add the IP address of an Elastic Compute Service (ECS) instance to the IP whitelist of my cluster, why am I unable to connect the ECS instance to the cluster?

You can perform the following steps for troubleshooting:

i. Check whether the IP whitelist is configured in a correct way. If you connect the ECS instance to the cluster by using an internal endpoint, you must add the private IP address of the ECS instance to the whitelist. If you connect the ECS instance to the cluster by using a public endpoint, you must add the public IP address of the ECS instance to the whitelist.

ii. Check whether the ECS instance and the cluster run in the same type of network. If the ECS instance runs in the classic network,you can migrate the ECS instance to the virtual private cloud (VPC) where the PolarDB cluster is deployed. For more information, see Classic network-to-VPC migration.

> **Note** If you want to connect the ECS instance to other internal resources that are located in the classic network, do not migrate the ECS instance to the VPC. The ECS instance cannot connect to the classic network after you migrate the ECS instance to the VPC.

You can also usethe ClassicLink feature to connect the classic network to the VPC.

iii. Check whether the ECS instance and the PolarDB cluster run in the same VPC. If the instance and cluster do not run in the same VPC, you must purchase a new PolarDB cluster,or activate Cloud Enterprise Network to connect the two VPCs for data access.

- Why am I unable to access the cluster by using a public endpoint?

  If you cannot access the cluster by using the public endpoint, perform the following steps for troubleshooting:

  - If you connect to the cluster from an ECS instance through a public endpoint, make sure that you have added the public IP address of the ECS instance to an IP whitelist.
  - Set the IP address in the IP whitelist to 0.0.0.0/0 and try again. If you can connect to the cluster, the public IP address that was specified in the IP whitelist is invalid. You must check the public endpoint. For more information, see View endpoints and ports.

Related API operations

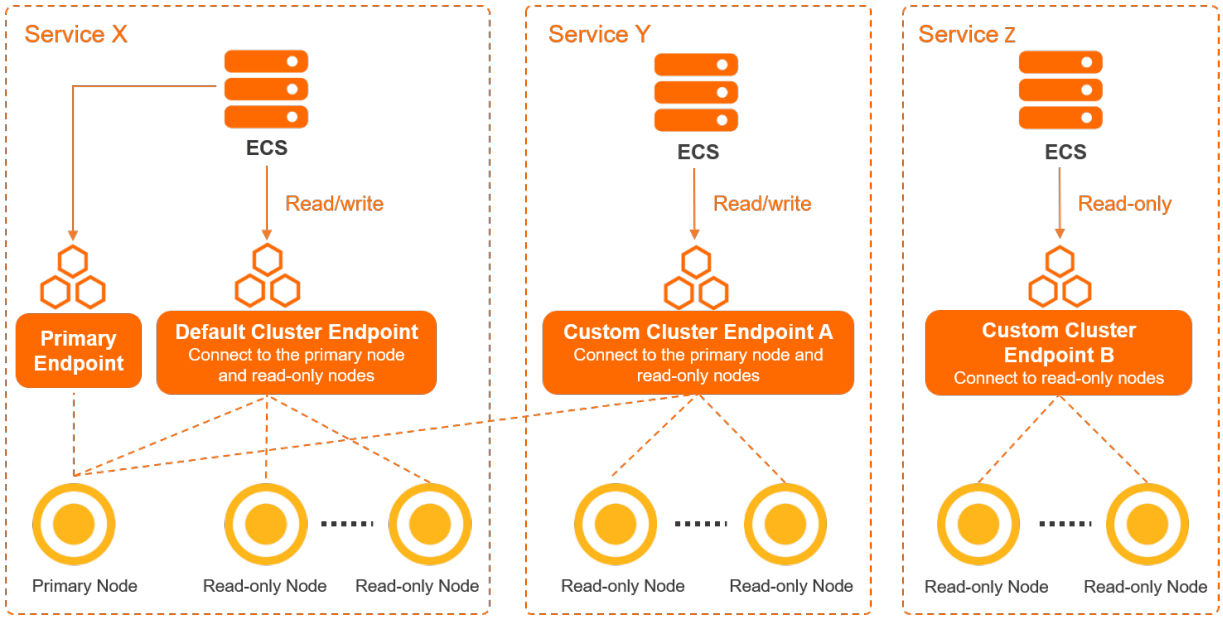
API	Description
DescribeDBClusterAccessWhitelist	Queries the IP addresses that are allowed to access a specified database cluster.
ModifyDBClusterAccessWhitelist	Modifies the IP addresses that are allowed to access a specified database cluster.

4.Connect to PolarDB

4.1. View or apply for an endpoint

To connect to a cluster, enter an endpoint of the cluster. supports cluster endpoints and primary endpoints. For each type of endpoints, you can apply for an internal endpoint or a public endpoint to connect to the cluster. This topic describes how to view or apply for an endpoint in the PolarDB console.

Cluster endpoint and primary endpoint



Endpoint	Description	Network type
Cluster endpoint (recommended)	• An application can connect to multiple nodes by connecting to one cluster endpoint. • Cluster endpoints support read/write splitting. Write requests are forwarded to the primary node. Read requests are forwarded to the primary node or read-only nodes based on the load on each node. **Note** A cluster provides a default cluster endpoint. You can create one or more custom cluster endpoints based on your needs. When you create a custom cluster endpoint, you can specify the nodes that can be connected by using the endpoint and set the read/write mode. For more information, see Create and modify a custom cluster endpoint.	• Internal network • Internet
Primary endpoint	• A primary endpoint allows you to connect to the primary node of the cluster. The endpoint can be used for read and write operations. • If the primary node fails, the primary endpoint is resolved to the new primary node.	

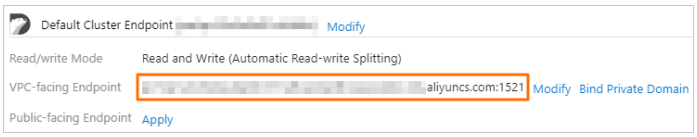
Internal endpoint and public endpoint

Network type	Description	Scenario
Internal network	• A cluster achieves optimal performance when the cluster is connected through an internal endpoint. • When you create a cluster, a default internal endpoint is generated. You can modify the endpoint but cannot delete it. For more information, see Modify an endpoint.	Examples: • If your Elastic Compute Service (ECS) instance runs in the same virtual private cloud (VPC) as the cluster, your ECS instance can connect to the cluster by using the internal endpoint. • You can connect to your cluster by using Data Management (DMS).
Internet	• You can apply for or delete a public endpoint. For more information, see Apply for an endpoint and Delete an endpoint. • The public endpoint enables connections over the Internet. A cluster cannot achieve optimal performance when the cluster is connected through a public endpoint.	Example: You can connect to your PolarDB cluster through a public endpoint to maintain databases.

View endpoints and ports

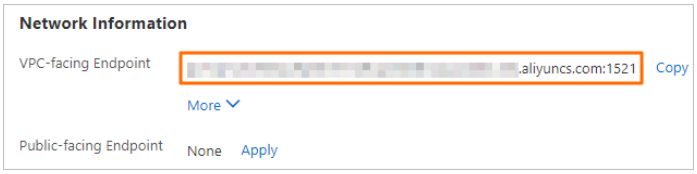
1.
2.
3.
4. In the **Endpoints** section of the **Overview** page, view the endpoint and port information by using one of the following methods:
   - Method 1

     In the upper-right corner of the **Endpoints** section, click the icon to switch views to view the endpoint and port information.

   - Method 2

     Click **Modify** on the right of the cluster endpoint. In the dialog box that appears, view **Network Information** that includes the endpoint and the port number.

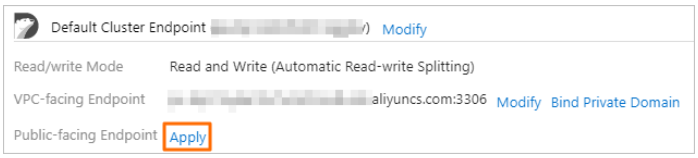
     

> **Note**
>
> - If you use a domain name to connect to a database, you can click **Bind Private Domain** to bind the domain name to an internal endpoint. This allows you to retain the original database domain name after the database is migrated to the cloud. You can bind private domain names only to **VPC-facing Endpoint** endpoints. For more information, see Private domain names.
> - By default, a cluster includes port 1521 in its endpoint. You cannot modify the port.

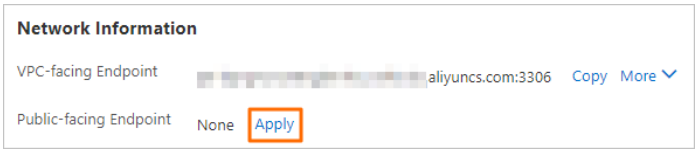
Apply for an endpoint

1.

2.
3.
4. In the **Endpoints** section of the **Overview** page, view the endpoint and port information by using one of the following methods:
5. Click **Apply**.
   - Method 1
     a. In the upper-right corner of the **Endpoints** section, click the icon to switch views.
     b. Click **Apply**.



   - Method 2
     a. Click **Modify** on the right of the cluster endpoint.
     b. In the **Network Information** section of the dialog box that appears, click **Apply**.



> **Note**
> - You can apply only for **Public-facing Endpoint** endpoints.
> - When you create a cluster, a default **VPC-facing Endpoint** endpoint is generated. You do not need to apply for this endpoint.

6. In the dialog box that appears, specify a prefix for the endpoint and click **OK**.

> **Note** The prefix of the endpoint must meet the following requirements:
> - The prefix must be 6 to 30 characters in length, and can contain lowercase letters, digits, and hyphens (-).
> - The prefix must start with a letter and end with a digit or a letter.

What to do next

Connect to a PolarDB cluster

Related API operations

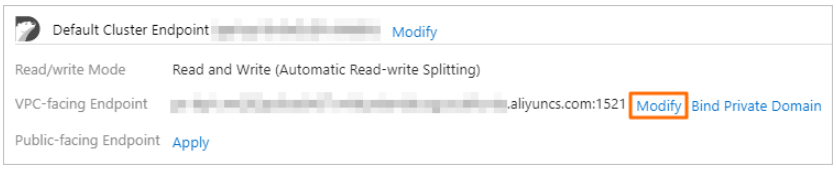
API	Description
DescribeDBClusterEndpoints	Queries the endpoint of a specified PolarDB cluster.
CreateDBEndpointAddress	Creates a public endpoint for a specified PolarDB cluster.
ModifyDBEndpointAddress	Modifies the default endpoint of a specified PolarDB cluster.
DeleteDBEndpointAddress	Deletes a cluster endpoint of a specified PolarDB cluster.

4.2. Modify or delete an endpoint

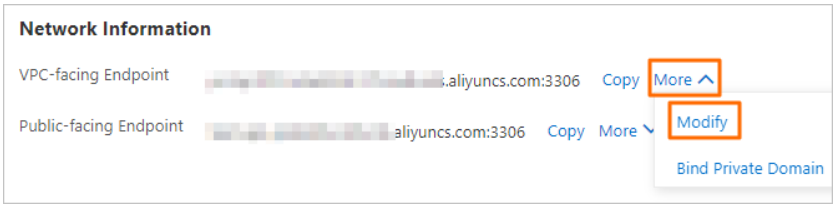
To connect to a cluster, enter an endpoint of the cluster. supports cluster endpoints and primary endpoints. For each type of endpoints, you can apply for an internal endpoint or a public endpoint to connect to the cluster. This topic describes how to modify or delete an endpoint in the console.

Modify an endpoint

1.
2.
3.
4. In the **Endpoints** section of the **Overview** page, view and modify the endpoint by using one of the following methods:
   - Method 1
     a. In the upper-right corner of the **Endpoints** section, click the icon to switch views.
     b. Find the endpoint that you want to modify and click **Modify**.



   - Method 2
     a. Click **Modify** on the right of the cluster endpoint.
     b. In the **Network Information** section of the dialog box that appears, choose **More > Modify**.



5. In the dialog box that appears, modify the prefix of the **Public-facing Endpoint** or **VPC-facing Endpoint** endpoint.

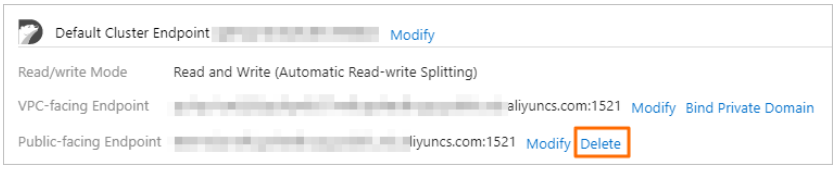
> **Notice**
> - The prefix of the endpoint must meet the following requirements:
>   - The prefix must be 6 to 30 characters in length, and can contain lowercase letters, digits, and hyphens (-).
>   - The prefix must start with a letter and end with a digit or a letter.
> - If Secure Sockets Layer (SSL) is enabled for the endpoint, the cluster is restarted after you modify the endpoint.
> - If SSL is enabled for the endpoint, the total length of the new endpoint cannot exceed 64 characters.

6. Click **OK**.

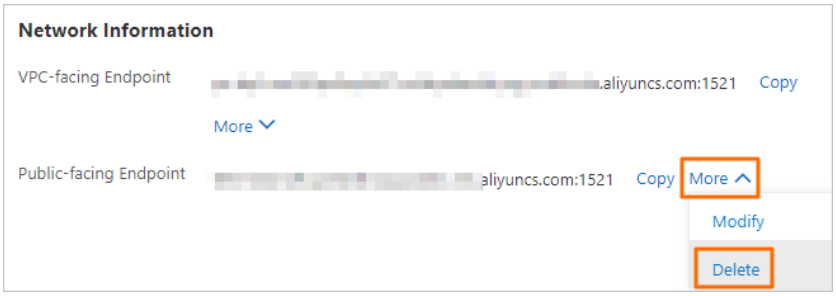
Delete an endpoint

> **Warning**
> - Before you delete an endpoint, make sure that your application is connected to the cluster through a new endpoint.
> - The deleted endpoint cannot be restored. You can click **Apply** next to the required endpoint type in the console to apply for a new endpoint. For more information, see Apply for an endpoint.

1.
2.
3.
4. In the **Endpoints** section of the **Overview** page, delete an endpoint by using one of the following methods:
   - Method 1
     a. In the upper-right corner of the **Endpoints** section, click the icon to switch views.
     b. Find the endpoint that you want to delete and click **Delete**.



   - Method 2
     a. Click **Modify** on the right of the cluster endpoint.
     b. In the **Network Information** section of the dialog box that appears, choose **More > Delete**.



> **Note** You can delete only **Public-facing Endpoint** endpoints.

5. Click **OK**.

Related API operations

API	Description
DescribeDBClusterEndpoints	Queries the endpoint of a specified PolarDB cluster.
CreateDBEndpointAddress	Creates a public endpoint for a specified PolarDB cluster.
ModifyDBEndpointAddress	Modifies the default endpoint of a specified PolarDB cluster.
DeleteDBEndpointAddress	Deletes a cluster endpoint of a specified PolarDB cluster.

4.3. Connect to a PolarDB cluster

This topic describes how to use Data Management (DMS) or a client to connect to a cluster.

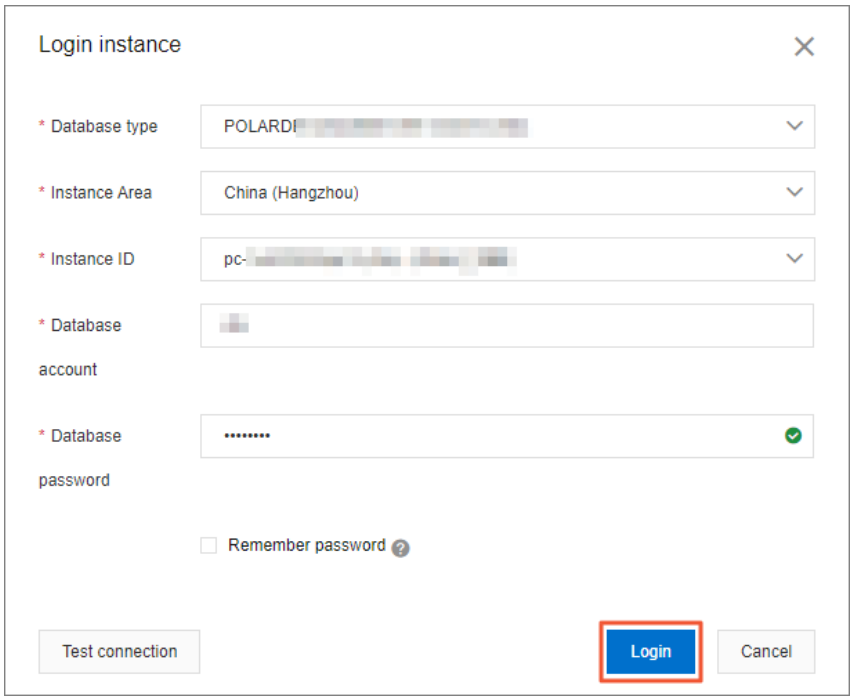
Prerequisites

- A privileged account or a standard account for a PolarDB cluster is created. For more information, see Create database accounts.
- The IP address of the host that you want to connect to the cluster is added to the whitelist. For more information, see Configure a whitelist for a cluster.

Use DMS to connect to a cluster

DMS provides an integrated data management solution. DMS supports data management, schema management, access control, BI charts, trend analysis, data tracking, performance optimization, and server management. DMS allows you to manage relational databases such as MySQL, SQL Server, and PostgreSQL, as well as NoSQL databases such as MongoDB and Redis. DMS also allows you to manage Linux servers.

1.
2.
3.
4. In the upper-right corner of the **Overview** page, click **Log On to Database**.
5. In the dialog box that appears, enter the **database account** and **password** that you create in the cluster.



6. Click **Login**.

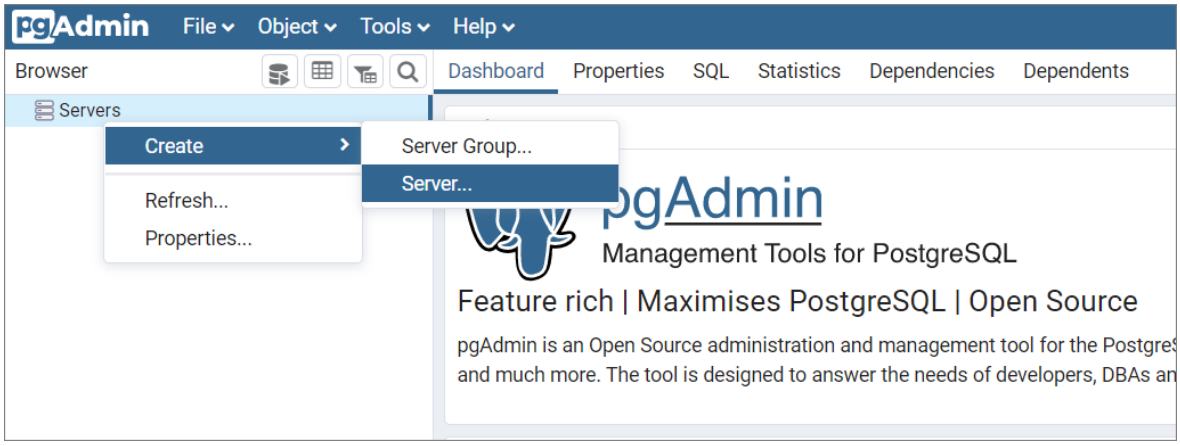
> **Note** If you are using DMS to connect to the cluster for the first time, you are prompted to set the whitelist. Click **Configure Whitelist** to complete the authorization.

7. After you log on to DMS, refresh the page. In the left-side navigation pane, click **Logged in instance**.
8. Find and double-click the name of the database that you want to manage. Then, you can manage the database.

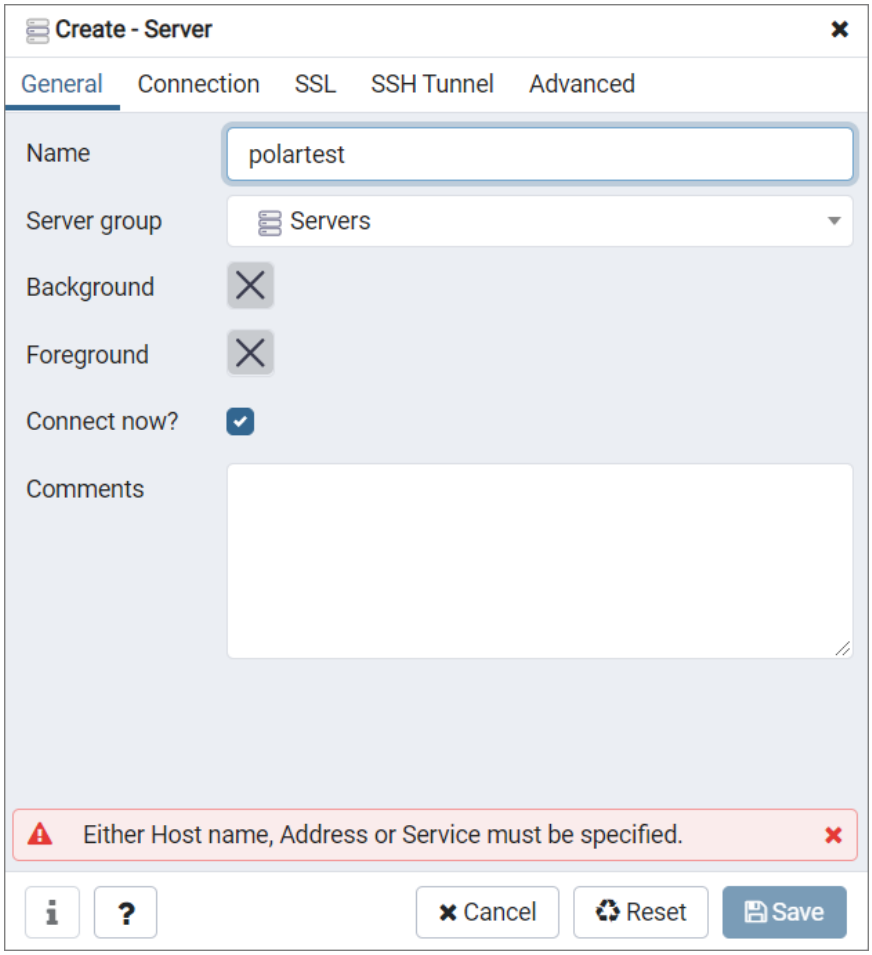
Use a client to connect to a cluster

You can also use your pgAdmin 4 client to connect to a cluster.

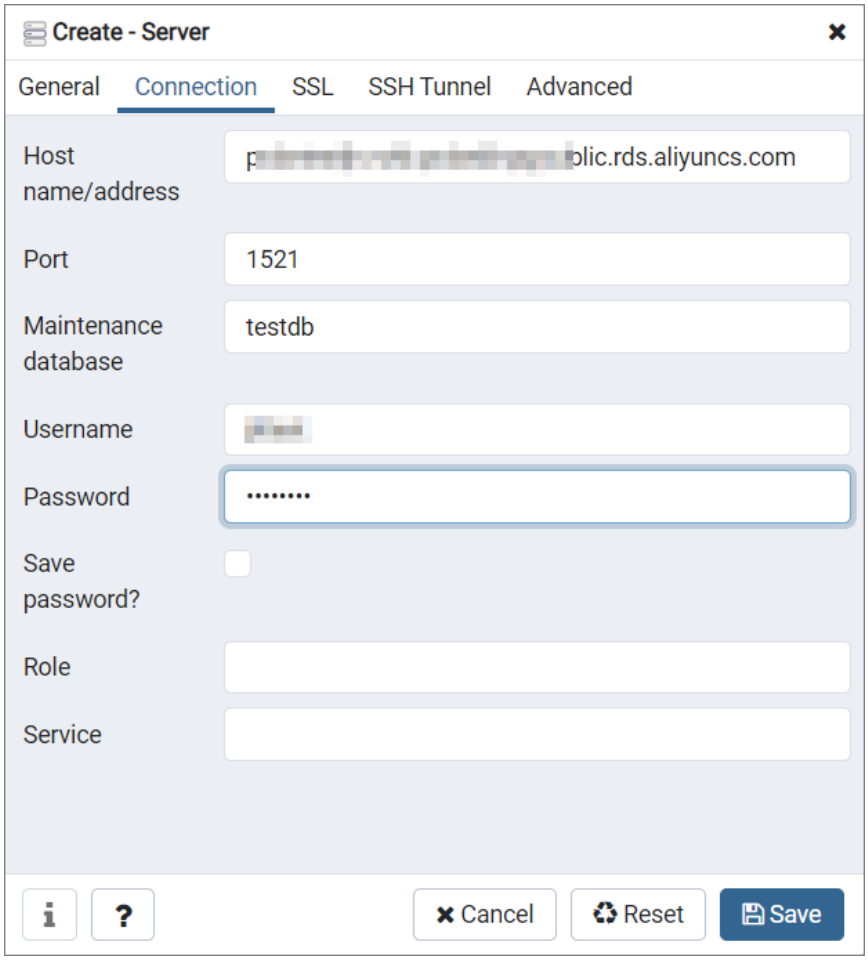
1. Launch the pgAdmin 4 client.
2. Right-click **Servers** and choose **Create > Server**, as shown in the following figure.



3. On the **General** tab of the **Create - Server** dialog box, enter the **name** of the server, as shown in the following figure.



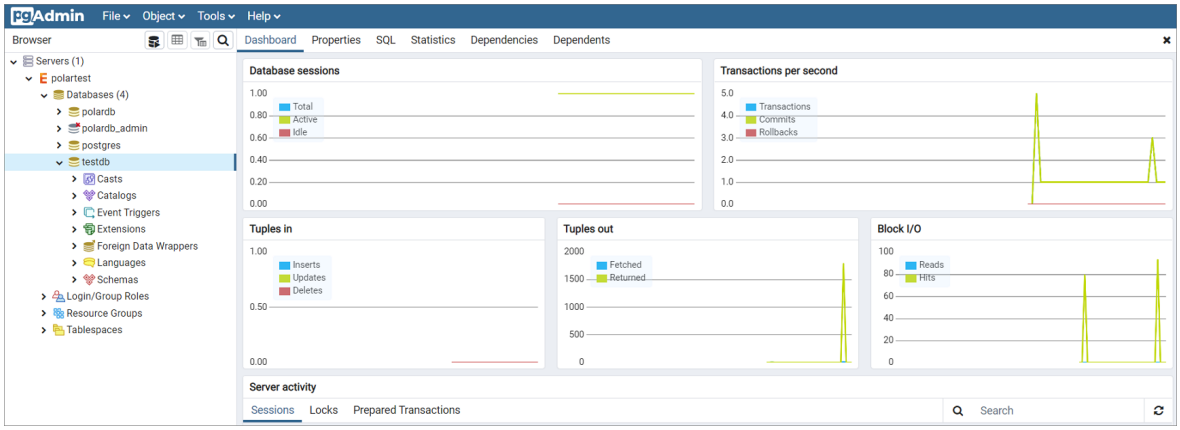
4. Click the **Connection** tab and specify the information about the cluster that you want to connect to. The following table describes the parameters.



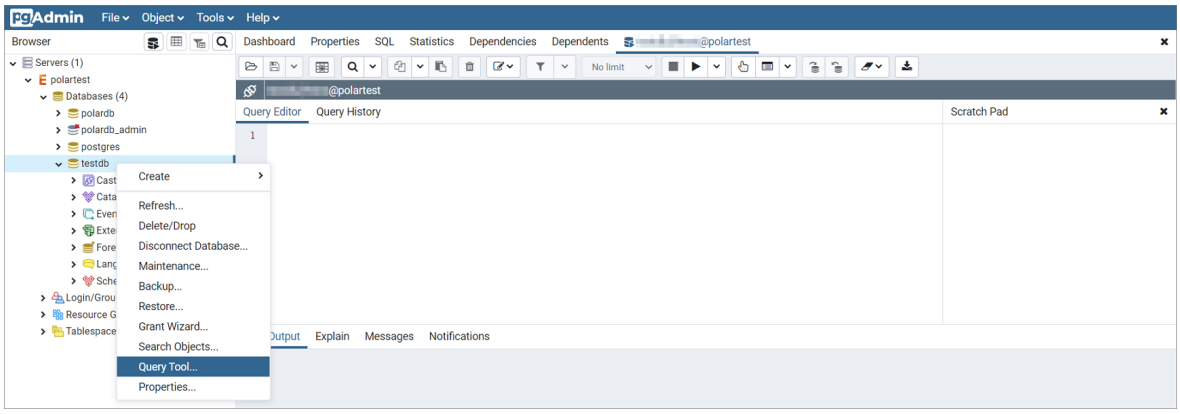
Parameter description

Parameter	Description
Host name/address	The endpoint of the cluster. To view the endpoint and port information about the cluster, perform the following steps: i. Log on to the PolarDB console. ii. In the upper-left corner of the console, select the region where the cluster is deployed. iii. On the Clusters page, click the ID of the cluster that you want to manage. iv. In the **Endpoints** section, view the endpoints of the cluster.
Port	The port for the cluster. Default value: 1521.
Maintenance database	The maintenance database. Default value: **postgres**.
Username	The account of the cluster. For more information about how to create an account, see Create database accounts.
Password	The password of the account for the cluster.

5. Click **Save**.
6. If the connection information is valid, a page that is similar to the following page appears after you click the database name. This indicates that the database is connected.



7. Right-click the name of the database that you want to manage and click **Query Tool...**. On the page that appears, you can add, delete, update, and query data in the database.



Use psql to connect to a PolarDB cluster

In addition to the preceding methods of connecting to your cluster, you can also download and install a PostgreSQL client and use psql to connect to your cluster.

> **Note**
>
> - The method of using psql to connect to your cluster for Windows operating systems is the same as that for Linux operating systems.
> - For more information about how to use psql, see psql.

1. Enter the following command in the psql command line interface (CLI) and press Enter:

```
psql -U <username> -h <host> -p <port> <dbname>
```

Parameter	Description
`username`	The account of the cluster. For more information about how to create an account, see Create database accounts.
`host`	The endpoint of the cluster. To view the endpoint and port information about the cluster, perform the following steps: i. Log on to the PolarDB console. ii. In the upper-left corner of the console, select the region where the cluster is deployed. iii. On the Clusters page, click the ID of the cluster that you want to manage. iv. In the **Endpoints** section, view the endpoints of the cluster.

Parameter	Description
`port`	The port for the cluster. Default value: 1521.
`dbname`	The name of the maintenance database that you want to manage. For more information about how to create a database, see Create a database.

The following code provides an example:

```
psql -U testuser -h hostname -p 1521 testdb
```

2. Enter the password of the specified `username` and press Enter. Then, the connection is established.

What to do next

Oracle compatibility

Troubleshooting

- The IP address whitelist is invalid.
  - The default whitelist contains only the IP address 127.0.0.1. 127.0.0.1 indicates that no IP address is allowed to access the cluster. Therefore, you must add the IP addresses of the clients that you use to access your cluster to the whitelist. For more information, see Configure a whitelist for a cluster.
  - The entry in the whitelist is set to 0.0.0.0. The valid format is 0.0.0.0/0.

    Notice 0.0.0.0/0 indicates that all IP addresses are allowed to access the cluster. Proceed with caution.

  - The public IP addresses that you add to the whitelist are invalid. For example, the public IP addresses may be dynamic IP addresses, or the tools or websites used to check the public IP addresses provide invalid IP addresses.
- The internal or public endpoint is incorrectly used.

  If you use an internal endpoint to establish a connection over the Internet or use a public endpoint to establish a connection over an internal network, the connection fails.

  Use the required endpoint. If you want to connect to the cluster over an internal network, use an internal endpoint of the cluster. If you want to connect to the cluster over the Internet, use a public endpoint of the cluster.
- A Domain Name System (DNS) server fails to resolve the endpoint of your cluster.
  - The endpoint that you enter to connect to the PolarDB cluster is invalid. In this case, troubleshoot the issue:
    - The endpoint of the PolarDB cluster is invalid. In this case, view the valid endpoints in the PolarDB console. For more information, see View endpoints and ports.
    - The endpoint that you enter is a public endpoint. However, the public endpoint is manually deleted.
    - Some applications have limits on the length of endpoints and the endpoint you enter is truncated.
  - If the endpoint of the PolarDB cluster is valid, change the IP address of the DNS server to that of the Alibaba Cloud DNS server.

Network type	IP address of the Alibaba Cloud DNS server
Internal network (classic network)	10.143.22.116 10.143.22.118
Internal network (virtual private cloud)	100.100.2.136 100.100.2.138
Internet	223.5.5.5 223.6.6.6

4.4. Private domain names

Assume that you use domain names to connect to databases and you want to retain the original domain names of the databases after the databases are migrated to the cloud. In this case, you can bind the private domain names by using the private domain name feature.

Scenarios

You can bind a private domain name to each VPC-facing endpoint of . Private domain names take effect in only the VPC that you specify in the current region. Private domain names have a higher priority for resolution than the domain names that take effect in the globe.

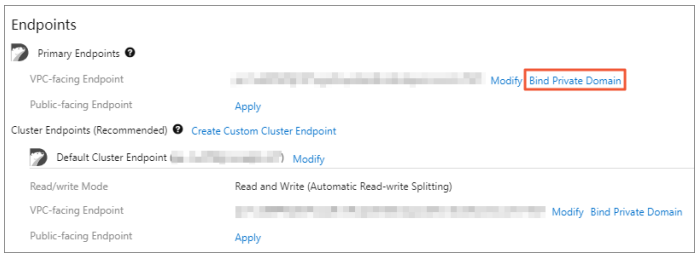
For example, the original domain name of a database is developer.aliyundoc.com, and the database is migrated to the cluster. The endpoint of the cluster is image.developer.aliyundoc.com. To allow the original domain name to remain unchanged, you can create a private domain name to bind developer.aliyundoc.com that is a CNAME record to image.developer.aliyundoc.com. After the domain name is bound to the endpoint, you can access the cluster by visiting developer.aliyundoc.com in the specified VPC, as shown in the following figure.

Billing description

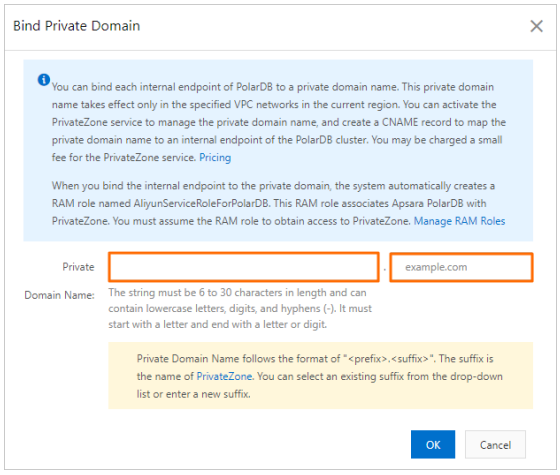
The private domain name feature of is realized by mapping the private domain names that are managed by PrivateZone to the VPC-facing endpoints of . PrivateZone charges a small amount of fee. For more information about pricing, see Pricing.

Bind a private domain name

1.
2.
3.
4. In the upper-right corner of the **Endpoints** section on the **Overview** page, click the icon to switch the view.
5. On the right side of the VPC-facing endpoint, click **Bind Private Domain**.



6. In the **Bind Private Domain** dialog box, enter the prefix and the suffix of the private domain name.



The format of private domain names is `<prefix>.<suffix>`. The following table describes the format of the private domain names.

Configuration	Description
Prefix of a private domain name	The prefix of the private domain name must be 6 to 30 characters in length and can contain at least one of the following types of characters: lowercase letters, digits, and hyphens (-). The prefix must start with a letter and end with a digit or a letter.
Suffix of the private domain name (zone)	You can select an existing zone from the drop-down list or enter a new zone. For more information about zones, see PrivateZone. **Note** ◦ If the VPC where your cluster resides is not in the configured zone, the system automatically binds the VPC to the zone. ◦ You canview and manage zones in the PrivateZone console.

Note When you bind a private domain name, the system automatically creates an AliyunServiceRoleForPolarDB role. For more information, see RAM role linked to Apsara PolarDB.

7. Click **OK**.
8. In the **Bind Private Domain** dialog box, confirm the information about the domain name again and click **OK**.

Related API operations

API	Description
ModifyDBEndpointAddress	Modifies the endpoints of a cluster, including the primary endpoint, default cluster endpoint, custom cluster endpoint, and private domain name.

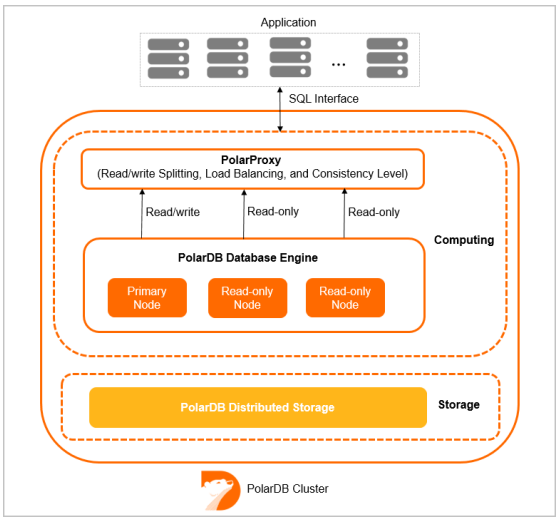
5.Cluster endpoint

5.1. Overview

PolarProxy is a proxy that is developed for . This topic describes the features of PolarProxy.

PolarProxy is deployed between your database system of and your applications. PolarProxy receives requests from your applications and then routes the requests to the primary and read-only instances of your database system. PolarProxy is easy to use and maintain, and provides high availability and high performance. In addition, PolarProxy provides advanced features such as automatic read/write splitting and transaction splitting.

PolarDB architecture and PolarProxy



Standard Edition has the following characteristics:

- A PolarDB cluster consists of one primary node and one or more read-only nodes.
- By default, PolarDB provides two types of endpoints: primary endpoints and cluster endpoints.

> **Note** The requests from the applications that are connected to cluster endpoints must pass through PolarProxy. Cluster endpoints are classified into read-only cluster endpoints and read/write cluster endpoints. Read-only cluster endpoints allow PolarDB clusters to distribute read requests to read-only nodes based on the number of connections. For information about read/write cluster endpoints, see Read/write splitting.

Read/write splitting

Read/writing splitting is automatically performed in the clusters of the . After an application is connected to a cluster endpoint of a PolarDB cluster, the write requests from the application are forwarded to the primary node of the cluster and read requests are forwarded to the primary and read-only nodes based on the load of each node. The number of pending requests on a node indicates the load of the node. For more information, see Read/write splitting.

Transaction splitting

supports transaction splitting. Transaction splitting ensures that data is consistent in a session and allows PolarDB to send read requests to read-only nodes to reduce the load of the primary node. For more information, see Transaction splitting.

Related API operations

Operation	Description
CreateDBEndpointAddress	Creates a public endpoint for a specified cluster.

Operation	Description
CreateDBClusterEndpoint	Creates a custom cluster endpoint for a specified cluster.
DescribeDBClusterEndpoints	Queries the information about the endpoints of a specified cluster.
ModifyDBClusterEndpoint	Modifies the configuration of a cluster endpoint for a specified cluster.
ModifyDBEndpointAddress	Modifies the endpoints such as custom cluster endpoints of a specified cluster.
DeleteDBEndpointAddress	Deletes a cluster endpoint of a specified cluster. This operation cannot be used to delete private custom cluster endpoints.
DeleteDBClusterEndpoint	Deletes a custom cluster endpoint of a specified cluster.

5.2. Features

5.2.1. Read/write splitting

clusters enable the read/write splitting feature. This feature enables PolarDB clusters to distribute read and write requests from applications by using cluster endpoints. The built-in proxy of a PolarDB cluster forwards write requests to the primary node. The proxy forwards read requests to the primary node or read-only nodes. The destination node depends on the loads on nodes. The number of requests that are not processed on a node indicates the loads on the node.

Benefits

- Easy maintenance based on a unified endpoint

  If you do not use cluster endpoints whose read/write mode is Read and Write (Automatic Read-write Splitting), you must configure the endpoints of the primary node and each read-only node in your application. Otherwise, you cannot send write requests to the primary node and read requests to read-only nodes. If you connect your application to cluster endpoints whose read/write mode is Read and Write (Automatic Read-write Splitting), the cluster endpoints can automatically forward read and write requests to the relevant nodes. This reduces maintenance costs. You need only to add read-only nodes to improve the processing capabilities of clusters, and do not need to modify your applications.

- Session-level read consistency

  When a client connects to the backend by using the cluster endpoint, the built-in proxy for read/write splitting automatically establishes a connection to the primary node and each read-only node. In the same session, the built-in proxy first selects an appropriate node based on the data synchronization progress of each database node. Then, the proxy forwards read and write requests to the nodes whose data is up-to-date and correct. This balances read and write requests among the nodes.

- Even distribution of the PREPARE statements

  The PREPARE statements that contain write operations and the related EXECUTE statements are sent to only the primary node. The PREPARE statements that contain read-only operations are broadcast to all the nodes, and the related EXECUTE statements are routed based on the loads on these nodes. This achieves load balancing for query requests.

- Support for native high security links, and improved performance

  You can build your own proxy on the cloud to achieve read/write splitting. However, an excessive latency may occur because data is parsed and forwarded by multiple components before the data arrives at a database. However, PolarDB uses a built-in proxy that works as a cluster component for read/write splitting. The built-in proxy provides a lower latency and higher data processing speed than external components.

- Node health checks to enhance database availability

The read/write splitting module of PolarDB performs health checks on the primary node and read-only nodes of a cluster. If a node fails or its latency exceeds a specified threshold, PolarDB stops distributing read requests to this node, and distributes write and read requests to other healthy nodes. This ensures that applications can access the cluster even if a single read-only node fails. After the node recovers, PolarDB automatically adds the node into the list of nodes that are available to receive requests.

Limits

- PolarDB does not support the following statements or features:
  - Connect to a cluster through the replication-mode method. If you need to set up dual-node clusters based on a primary/secondary replication architecture, use the endpoint of the primary node.
  - Use the name of the temporary table to declare the %ROWTYPE attribute.

    ```
    create temp table fullname (first text, last text);
    select '(Joe,von Blow)'::fullname, '(Joe,d''Blow)'::fullname;
    ```

  - Create temporary resources by using functions.
    - If you create a temporary table by using functions and execute an SQL statement to query the temporary table, an error message may be returned. The error message indicates that the table does not exist.
    - If your function contains the PREPARE statement, an error message may be returned when you execute the EXECUTE statement. The error message indicates that the PREPARE statement name does not exist.
- Routing-related restrictions:
  - Requests in the transaction are routed to the primary node, and load balancing is resumed after the transaction terminates.
  - All statements that use functions except aggregate functions such as COUNT() and SUM() are routed to the primary node.

Create or modify a cluster endpoint

- For more information about how to create a custom cluster endpoint, see Create and modify a custom cluster endpoint.
- For more information about how to modify a cluster endpoint, see FAQ.

Configure transaction splitting (advanced setting)

For more information, see Transaction splitting.

Specify a consistency level (advanced setting)

For more information, see Consistency levels.

FAQ

- Why am I unable to immediately retrieve a record after I insert the record?

  This is because in a read/write splitting architecture, a replication delay may occur during data replication between the primary node and read-only nodes. However, PolarDB supports session consistency. This allows you to query the updates within a session.

- Why do read-only nodes have no loads?

  By default, all the requests in transactions are sent to the primary node. If you use sysbench for stress testing, you can use --oltp-skip-trx=on in sysbench 0.5 or use --skip-trx=on in sysbench 1.0 to skip transactions. If the loads on read-only nodes are excessively low due to a large number of transactions in your business, you can enable the transaction splitting feature in the console. For more information, see Configure transaction splitting (advanced setting).

- Why does a node receive more requests than other nodes?

  Requests are distributed to each node based on loads. The node that has low loads receives more requests.

- Does PolarDB support zero-delay read access to data?

No, PolarDB does not support zero-delay read access to data. When the primary node and read-only nodes process normal loads, a delay of several milliseconds occurs. If the read/write mode of a cluster endpoint is Read and Write (Automatic Read-write Splitting), data cannot be read with a zero delay after the data is written. If you require zero-delay read access to data, connect your applications to the primary endpoint to send all the read and write requests to the primary node. The primary endpoint is always connected to your PolarDB primary node.

- Are new read-only nodes automatically available to receive read requests for read/write splitting?

  If a session that supports read/write splitting is created after you add a read-only node, the read requests are forwarded to the read-only node. If a session that supports read/write splitting is created before you add a read-only node, the read requests are not forwarded to the read-only node. To ensure that the read requests are forwarded to the read-only node, close the session and establish the session again. For example, you can restart your application to establish the session.

5.2.2. Consistency levels

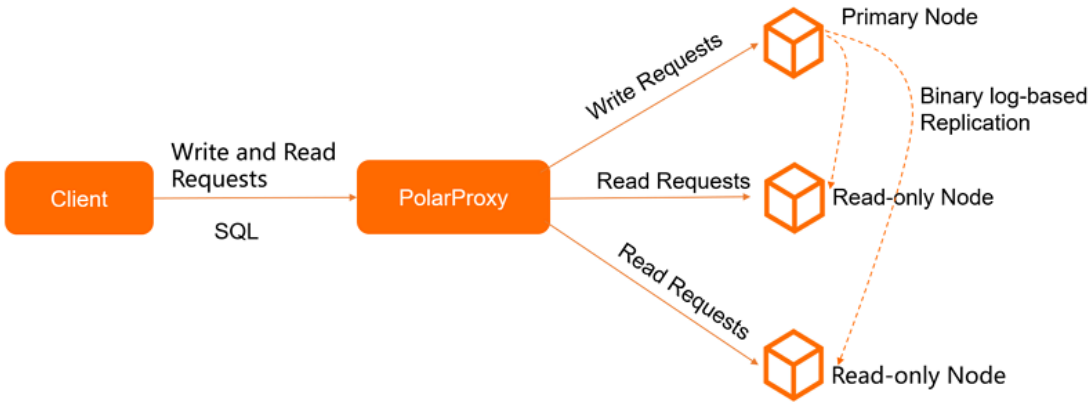
provides two consistency levels to meet the requirements of different business scenarios: eventual consistency and session consistency. Consistency refers to the consistency feature in atomicity, consistency, isolation, durability (ACID).

Issues and solutions

Data replication is a simple method used to replicate data from the primary node to read-only nodes in . You only need to asynchronously transfer the write-ahead logs (WALs) of the primary node to read-only nodes. Data replication enables read-only nodes to process queries. This reduces loads on the primary node and ensures high availability. If you use read-only nodes to process read requests, the following issues may occur:

- 1. Typically, the primary database and secondary databases provide different endpoints. When you access different databases, you must modify the code in your applications to change the endpoint that is used to connect to databases.
- 2. Data is asynchronously replicated. Data may not be immediately synchronized to read replicas after a client commits data modifications. Data in read-only nodes may not be up-to-date. In this case, data is not consistent.

To resolve the first issue, uses PolarProxy as a proxy to perform read/write splitting operations. In most cases, the proxy establishes connections from applications to and parses each SQL statement. Write requests such as UPDATE, DELETE, INSERT, and CREATE operations are forwarded to the primary database by the proxy. The SELECT operations are forwarded to secondary databases.



The read/write splitting feature cannot fix the data inconsistency issue that is caused by a replication delay. If the loads on databases are heavy, the replication delay increases. For example, when you execute DDL statements to add columns to a large table or insert a large amount of data. In this case, you cannot retrieve the most recent data from read-only nodes.

synchronizes data between the primary node and read-only nodes by performing asynchronous physical replication. After the data on the primary node is updated, the updates are synchronized to read-only nodes. The replication delay varies based on the write loads on the primary node. The replication delay is just a few milliseconds. The asynchronous replication ensures eventual consistency among the primary and read-only nodes. provides the following consistency levels to meet your different consistency requirements:

- Eventual consistency

- Session consistency

> **Note** For more information about how to change the consistency level, see Configure PolarProxy.

Eventual consistency

- Description

  runs in a read/write splitting architecture. Traditional read/write splitting ensures only eventual consistency. The retrieved results from different nodes may be different due to a primary/secondary replication delay. For example, if you repeatedly execute the following statements within a session, the result returned by each SELECT statement may be different. The actual query result depends on the replication delay.

```
INSERT INTO t1(id, price) VALUES(111, 96);
UPDATE t1 SET price = 100 WHERE id=111;
SELECT price FROM t1;
```

- Applicable scenarios

  To reduce loads on the primary node and send as a number of read requests as possible to read-only nodes, we recommend that you select eventual consistency.

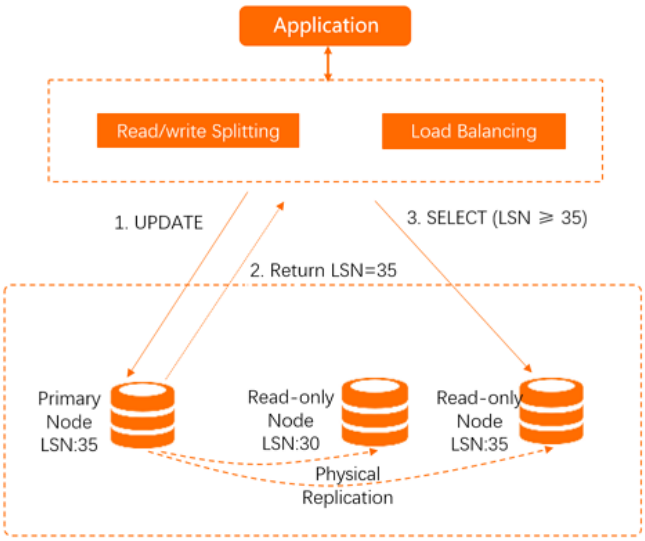
Session consistency

- Description

  In most cases, requests are split to eliminate data inconsistencies caused by eventual consistency. The requests that require high consistency are sent to the primary node. The requests that require eventual consistency are sent to read-only nodes by using the read/write splitting feature. This increases the loads on the primary node, reduces read/write splitting performance, and makes application development more complex.

  To solve the issue, supports session consistency. Session consistency is known as causal consistency. Session consistency ensures that the data that is updated before read requests are sent in a session can be obtained. This ensures that data is monotonic.

  uses PolarProxy to perform read/write splitting operations. PolarProxy tracks redo logs that are applied on each node and records each log sequence number (LSN). When the data in the primary node is updated, records the LSN of the new update as a session LSN. When a new read request is received, compares the session LSN with the LSN on each node and forwards the request to a node in which the LSN is greater than or equal to the session LSN. This ensures session consistency. performs physical replication in an efficient manner.



  To ensure efficient synchronization, data is being replicated to other read-only nodes when the read-only node returns the result to the client. This way, data is updated on read-only nodes before subsequent read requests are received. In most scenarios, a large number of read requests and a small number of write requests exist. Therefore, this mechanism can ensure session consistency, read/write splitting, and load balancing based on the verification result.

- Applicable scenarios

  A higher consistency level of a cluster indicates heavier loads on the primary database and lower cluster performance. We recommend that you use session consistency. This consistency level meets the requirements of most application scenarios and minimizes the impact on cluster performance.

Best practices for consistency levels

- A higher consistency level of a cluster indicates lower cluster performance. We recommend that you use session consistency. This consistency level meets the consistency requirements in most of the application scenarios and minimizes the impact on cluster performance.
- If you require high data consistency between different sessions, you can select one of the following solutions:

  Use hints to forcibly send specific queries to the primary node.

```
/*FORCE_MASTER*/ select * from user;
```

> **Note**
>
> - If you want to execute the preceding statement that contains the hint on the official command line of MySQL, add the -c parameter in the statement. Otherwise, the hint becomes invalid because the official command line of MySQL filters out the hint. For more information, see mysql Client Options.
> - Hints are assigned the highest priority for routing and are not limited by consistency levels or transaction splitting. Before you use hints, evaluate the impacts on your business.
> - Hints cannot contain statements that change environment variables. For example, if you execute the `/*FORCE_SLAVE*/ set names utf8;` statement, errors may occur.

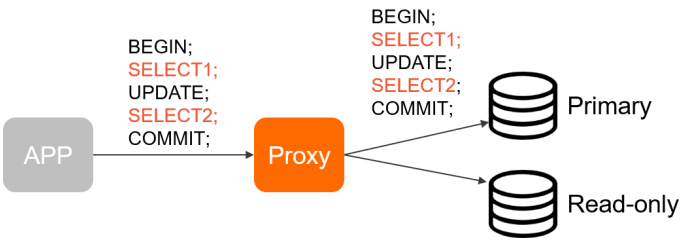
5.2.3. Transaction splitting

supports transaction splitting. Transaction splitting ensures that data is consistent in a session and allows PolarDB to send read requests to read-only nodes to reduce the loads on the primary node.

Context

If you use a cluster endpoint in read/write mode, the proxy forwards read and write requests to the primary node and read-only nodes. To ensure read/write consistency of transactions in a session, the proxy sends all transaction requests in the session to the primary node.

For example, some database client drivers, such as the Java Database Connectivity (JDBC) driver, encapsulate all requests in transactions by default. In this case, all requests from applications are sent to the primary node. This results in heavy loads on the primary node and low loads on read-only nodes, as shown in the following figure.

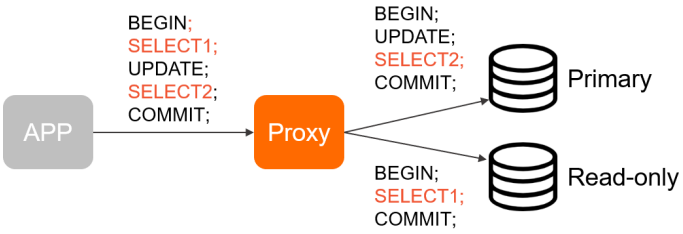


To fix the preceding issue, PolarDB provides the transaction splitting feature that can be used to ensure read/write consistency. This feature reduces the loads on the primary node by sending read requests in transactions to read-only nodes.

Description

Basic transaction splitting

To reduce the loads on the primary node, PolarProxy sends read requests that are received before the first write request in a transaction is sent to read-only nodes. Uncommitted data in transactions cannot be queried from read-only nodes. To ensure data consistency in transactions, all read and write requests that are received after the first write request are still forwarded to the primary node. For more information about how to enable basic transaction splitting, see Create and modify a custom cluster endpoint.



Benefits

This feature allows you to transfer the read loads from the primary node to read-only nodes without modifying application code or configurations. This makes the primary node more stable.

Note

- Only transactions that use the Read Committed isolation level can be split.
- If you enable basic transaction splitting and the consistency level is not set to **eventual consistency**, the proxy sends the read requests that are received before the first write request in a transaction to read-only nodes only after read-only nodes synchronize all data from the primary node. Otherwise, the proxy still sends read requests to the primary node. For more information about consistency levels, see Consistency levels.

Related API operations

API	Description
ModifyDBClusterEndpoint	Modifies the attributes of a cluster endpoint. For example, you can modify the following attributes: read/write mode, consistency level, transaction splitting, and offload read requests from the primary node. You can also specify whether to associate the specified cluster endpoint with newly added nodes.

5.3. Create and modify a custom cluster endpoint

This topic describes how to create and modify a custom cluster endpoint. You can enable or disable features such as read/write splitting, transaction splitting, and consistency level when you create or modify a custom cluster endpoint.

Create a custom cluster endpoint

When you create a custom cluster endpoint, you can enable or disable read/write splitting, transaction splitting, and consistency level.

Procedure

1.
2.
3.
4. In the **Endpoints** section, click **Create Custom Cluster Endpoint**.
5. In the **Create Custom Cluster Endpoint** dialog box, set the following parameters.

   Parameters for creating a custom cluster endpoint

Parameter		Description
Network Information		By default, provides a public endpoint for each cluster. For more information about how to modify the public endpoint or apply for a VPC endpoint,see Modify an endpoint and Apply for an endpoint.
Cluster Settings	**Read/write Mode**	Specifies the read/write mode for the custom cluster endpoint. You can select **Read Only** or **Read and Write (Automatic Read-write Splitting)**. **Note** You can change the read/write mode of a custom cluster endpoint after the custom cluster endpoint is added. After you change the read/write mode, the new mode takes effect only on newly created connections. The existing connections still work in the original mode.
	Endpoint Name	The name of the cluster endpoint.
Node Settings	**Unselected Nodes** and **Selected Nodes**	Select the nodes that you want to add to process read requests from the **Unselected Nodes** section on the left and click the icon. Then, move the nodes to the **Selected Nodes** section on the right. **Note** ◦ The **Unselected Nodes** list shows the primary node and all read-only nodes. ◦ The types of nodes that you select do not affect the read/write mode. If you set the read/write mode to **Read and Write (Automatic Read-write Splitting)**, write requests are sent only to the primary node regardless of whether the primary node is selected. ◦ allows you to create a cluster endpoint that is associated with only the primary node. However, when the read/write mode is set to **Read Only**, you are not allowed to create a cluster endpoint that is associated only with the primary node.
	Automatically Associate New Nodes	Specifies whether to automatically associate a newly added node with the cluster endpoint.
SLB Settings	**Load Balancing Policy**	Specifies the load balancing policy that is used to distribute read requests to multiple read-only nodes when read/write splitting is enabled. The default value is **Load-based Automatic Scheduling** and cannot be changed.
	Primary Node Accepts Read Requests	After you set Primary Node Accepts Read Requests to No, SQL query requests are sent only to read-only nodes. This reduces the loads on the primary node and ensures the service stability of the primary node. For more information, see Read/write splitting. **Note** This parameter is available only if the read/write mode is set to **Read and Write (Automatic Read-write Splitting)**.
	Transaction Splitting	Specifies whether to enable the transaction splitting feature. For more information, see Transaction splitting. **Note** This parameter is available only if you set the read/write mode to **Read and Write (Automatic Read-write Splitting)**.

Parameter		Description
Consistency Settings	**Consistency Level**	◦ If you set Read/write Mode to **Read and Write (Automatic Read-write Splitting)**, the following consistency levels are available: **Eventual Consistency (Weak)** and **Session Consistency (Medium)**. For more information, see Consistency levels. ◦ If you set the read/write mode to **Read Only**, the default consistency level is **Eventual Consistency (Weak)** and cannot be changed. **Note** Changes to the consistency level immediately take effect on all connections.

6. Click **OK**.

Modify a custom cluster endpoint

When you modify a custom cluster endpoint, you can enable or disable read/write splitting, transaction splitting, and consistency level.

Procedure

1.
2.
3.
4. In the **Endpoints** section, find the custom cluster endpoint that you want to modify, and choose **Modify** next to the cluster endpoint.
5. In the **Configure Nodes** dialog box, you can set relevant parameters. For more information about the parameters, see Parameters for creating a custom cluster endpoint.
6. Click **OK**.

Delete a custom cluster endpoint

Note

- You can delete only custom cluster endpoints. The default cluster endpoint cannot be deleted.
- The deleted custom cluster endpoint cannot be recovered. You must change the endpoint for connecting the client to the cluster at the earliest opportunity.

1.
2.
3.
4. In the **Endpoints** section, find the cluster endpoint that you want to modify, and choose **Settings > Delete** on the right side of the cluster endpoint.
5. In the message that appears, click **OK**.

Related API operations

Operation	Description
CreateDBClusterEndpoint	Creates a custom cluster endpoint for a specified PolarDB cluster.
DescribeDBClusterEndpoints	Queries the cluster endpoints of a PolarDB cluster.
DeleteDBClusterEndpoint	Deletes a custom cluster endpoint of a PolarDB cluster.

5.4. Customize a routing rule for requests that contain a specified function or a specified table

PolarProxy allows you to customize a routing rule for requests that contain a specified function or a specified table. By default, read requests that contain functions in your PolarDB cluster are routed to the read-only nodes and read requests that contain tables in your PolarDB cluster are routed to the primary node. You can customize a routing rule to route read requests that contain a specified function in your PolarDB cluster to the read-only nodes. You can also customize a routing rule to route read requests that contain a specified table to the primary node.

Prerequisites

A privileged account is used to connect to the primary endpoint of your PolarDB cluster.

Create the polar_proxy_utils plug-in

Submit a ticket to create the polar_proxy_utils plug-in.

Customize a routing rule

To customize a routing rule, execute the following statement:

```
polar_add_proxy_routing_strategy(_name, _type, rw_mode);
```

Note The previous statement contains the following parameters:

- _name: the name of the table or function for which you want to customize a routing rule.
- _type: specifies whether the object that is specified by the _name parameter is a table or a function. The value t specifies a table. The value f specifies a function.
- rw_mode: specifies whether to route requests to the primary node or the read-only nodes. The value w specifies the primary node. The value r specifies the read-only nodes.

For example, execute the `polar_add_proxy_routing_strategy('lol', 't', 'w');` statement.

```
postgres=#
postgres=#
postgres=#
postgres=# select polar_list_proxy_routing_strategy();
 polar_list_proxy_routing_strategy
-----------------------------------
(0 rows)

postgres=# select polar_add_proxy_routing_strategy('lol', 't', 'w');
 polar_add_proxy_routing_strategy
----------------------------------

(1 row)

postgres=# select polar_list_proxy_routing_strategy();
       polar_list_proxy_routing_strategy
------------------------------------------------
 (lol,table,write,"2020-03-24 04:12:56.245526")
(1 row)
```

- Before this statement is executed, `select * from lol` operations are routed to the read-only nodes.
- After this statement is executed, `select * from lol` operations are routed to the primary node.

Query routing rules

To query routing rules, execute the following statement:

```
select polar_list_proxy_routing_strategy();
```

```
postgres=# select polar_list_proxy_routing_strategy();
        polar_list_proxy_routing_strategy
-------------------------------------------------
 (lol,table,write,"2020-03-24 04:12:56.245526")
 (abc,function,read,"2020-03-24 04:15:23.854227")
(2 rows)
```

Delete a routing rule

To delete a routing rule, execute the following statement:

```
select polar_delete_proxy_routing_strategy(_name, _type);
```

Note The previous statement contains the following parameters:

- _name: the name of the table or function that is associated with the routing rule to be deleted.
- _type: specifies whether the object specified by the _name parameter is a table or a function. The value t specifies a table. The value f specifies a function.

For example, run the `select polar_delete_proxy_routing_strategy('lol', 't');` statement.

```
postgres=#
postgres=# select polar_list_proxy_routing_strategy();
        polar_list_proxy_routing_strategy
-------------------------------------------------
 (lol,table,write,"2020-03-24 04:12:56.245526")
 (abc,function,read,"2020-03-24 04:15:23.854227")
(2 rows)

postgres=# select polar_delete_proxy_routing_strategy('lol', 't');
 polar_delete_proxy_routing_strategy
-------------------------------------

(1 row)

postgres=# select polar_list_proxy_routing_strategy();
        polar_list_proxy_routing_strategy
-------------------------------------------------
 (abc,function,read,"2020-03-24 04:15:23.854227")
(1 row)
```

- Before this statement is executed, `select * from lol` operations are routed to the primary node.
- After this statement is executed, `select * from lol` operations are routed to the read-only nodes.

Delete all routing rules

To delete all routing rules, execute the following statement:

```
select polar_truncate_proxy_routing_strategy();
```

```
postgres=# select polar_list_proxy_routing_strategy();
        polar_list_proxy_routing_strategy
-------------------------------------------------
 (abc,function,read,"2020-03-24 04:15:23.854227")
 (lol,table,write,"2020-03-24 04:20:17.186641")
(2 rows)

postgres=# select polar_truncate_proxy_routing_strategy();
 polar_truncate_proxy_routing_strategy
---------------------------------------

(1 row)

postgres=# select polar_list_proxy_routing_strategy();
 polar_list_proxy_routing_strategy
-----------------------------------
(0 rows)
```

5.5. FAQ

This topic provides answers to frequently asked questions (FAQ) about PolarProxy provided by .

Read/write splitting

- Why am I unable to retrieve a record immediately after I insert the record?

  In a read/write splitting architecture, a delay occurs when data is being replicated among the primary node and read-only nodes. supports session consistency to ensure that you can query updates within a session. You can retrieve the inserted record after the replication is complete. For more information, see the "Session consistency" section in Consistency levels.

- Can data be read immediately after it is written into PolarDB?

  No, data cannot be read immediately after the data is written into PolarDB. A delay of a few milliseconds occurs when you read data by using an endpoint for which read/write splitting is enabled, even if the loads on the primary node and read-only nodes of an cluster are not heavy. To eliminate this delay, you can use the primary endpoint to connect to the cluster. This way, read and write requests are sent to the primary node. For more information about how to view the primary endpoint, see View the endpoint and port.

- Why do low loads exist on read-only nodes when the loads on the primary node are high?

  By default, requests in transactions are routed only to the primary node. To balance loads across the primary and read-only nodes, you can use the following solutions:

  - When you perform stress tests by using Sysbench, specify `--oltp-skip-trx=on` in the code if the version of Sysbench is 0.5 or `--skip-trx=on` if the version of Sysbench is 1.0. This way, you do not need to execute the BEGIN and COMMIT statements.
  - If a large number of transactions cause heavy loads on the primary node, you can enable the transaction splitting feature to reduce the loads on the primary node. For more information, see Transaction splitting.

- Why does a specific node receive more requests than other nodes?

  Requests are distributed to each node based on loads. The nodes on which lighter loads exist receive more requests.

- Does a new read-only node automatically receive read requests?

  This depends on whether a session that supports read/write splitting is created after you add a read-only node. If yes, requests are automatically forwarded to the read-only node. If no, read requests are not forwarded to the read-only node. In this case, you can close a connection and then reconnect to your cluster. This way, read requests sent over the connection are forwarded to the read-only node. For example, you can restart your application to establish a new connection.

Cluster endpoints

- What is the maximum number of single-node cluster endpoints that I can create in a cluster?

You can create up to three custom cluster endpoints for a cluster. The custom cluster endpoints can be single-node cluster endpoints. For more information about how to create a single-node cluster endpoint, see Create and modify a custom cluster endpoint.

> **Warning** If you create a single-node cluster endpoint for a read-only node and the read-only node fails to run as expected, the single-node cluster endpoint may be unavailable for up to 1 hour. We recommend that you do not create single-node cluster endpoints in your production environment.

- If a single-node cluster endpoint is created for a read-only node, can the read-only node be used as the new primary node after a failover?

  The read-only node for which a single-node cluster endpoint is created cannot be automatically used as the new primary node after a failover. However, you can manually promote the read-only node as the new primary node. For more information, see Automatic failover and manual failover.

- What is the maximum number of cluster endpoints that can be created for a cluster?

  A cluster can contain a maximum of four cluster endpoints. One cluster endpoint is the default cluster endpoint and the other endpoints are custom cluster endpoints.

- Can I modify a cluster endpoint?

  Yes, you can modify the default cluster endpoint and custom cluster endpoints. For more information, see Modify a custom cluster endpoint.

- Can I delete a cluster endpoint?

  Yes, you can delete only custom cluster endpoints. You cannot delete the default cluster endpoint. For more information, see Delete a custom cluster endpoint.

6.Cluster management

6.1. Create a cluster

This topic describes how to create a cluster by using the console.

Prerequisites

An Alibaba Cloud account is created and is used to log on to the Alibaba Cloud Management Console. For more information, see Register and log on to an Alibaba Cloud account.

Context

A cluster consists of one primary node and a maximum of 15 read-only nodes. To ensure high availability, at least one read-only node is required to implement the active-active architecture. A node is a virtual database server. You can create and manage multiple databases on a node.

> **Note**
> - supports only Virtual Private Cloud (VPC). Each VPC is an isolated network on Alibaba Cloud and is more secure than the classic network.
> - To optimize the performance of , clusters must be deployed within the same internal network as other Alibaba Cloud services. We recommend that you deploy clusters and Elastic Compute Service (ECS) instances in the same VPC to ensure the optimal performance of . If your ECS instance is deployed in the classic network, you must migrate the ECS instance to a VPC.

Procedure

1.
2. In the upper-left corner of the page, click **Create Cluster**.
3. Select **Subscription** or **Pay-As-You-Go**.

> **Note**
> - **Subscription**: If you select this billing method when you create the cluster, you must pay for compute nodes (a primary node and a read-only node) in advance. In addition, you are charged for the consumed storage resources on an hourly basis. The charge of storage resources is deducted from your account on an hourly basis. The **Subscription** billing method is more cost-effective than the Pay-As-You-Go billing method if you want to use the new cluster for a long period of time. You are offered larger discounts for longer subscription periods.
> - **Pay-As-You-Go**:If you select this billing method when you create the cluster, you do not need to pay in advance. You are charged for compute nodes and the consumed storage resources on an hourly basis. These charges are deducted from your account balance on an hourly basis. We recommend that you select the **Pay-As-You-Go** billing method for the short-term use. You can reduce costs by releasing the cluster based on your business requirements.

4. Specify the parameters described in the following table.

Parameter	Description
Region	The region where the cluster is deployed. The region cannot be changed after the cluster is created. **Note** Make sure that the cluster is created in the same region as the Elastic Compute Service (ECS) instance to which you want to connect. Otherwise, the cluster and the ECS instance can communicate only over the Internet. As a result, the performance of the cluster may be compromised.

Parameter	Description
Creation Method	The method used to create the cluster. ○ **Create Primary Cluster**: creates a cluster. ○ **Restore from Recycle**: creates a cluster by restoring a backup of a deleted cluster from the recycle bin. ■ **Source Version**: the version of the deleted cluster that you want to restore. ■ **Deleted Clusters**: the name of the deleted cluster that you want to restore. ■ **Backup History**: the backup that you want to restore. **Note** You can select other options to create databases of other engines.
Primary Availability Zone	The primary zone where the cluster is deployed. ○ Each zone is an independent geographical location in a region. All of the zones in a region provide the same level of service performance. ○ You can choose to create your cluster in the same zone as an ECS instance or in a different zone from the zone of the instance. ○ You must specify only the primary zone. The system automatically selects a secondary zone.
Network Type	This parameter can be set only to **VPC**. You do not need to specify this parameter.
VPC **VSwitch**	Make sure that the cluster is created in the same VPC as the ECS instance to which you want to connect. Otherwise, the cluster and the ECS instance cannot communicate over the internal network to achieve optimal performance. ○ If you have an existing VPC that meets your network requirements, select the VPC. For example, if you have an existing ECS instance and the VPC to which the ECS instance belongs meets your network requirements, select this VPC. ○ Otherwise, use the default VPC and the default vSwitch. ■ Default VPC: ■ Only one VPC is specified as the default VPC in the region that you select. ■ The default VPC uses a 16-bit subnet mask. For example, the CIDR block 172.31.0.0/16 provides up to 65,536 private IP addresses. ■ The default VPC does not count towards the quota of VPCs that you can create on Alibaba Cloud. ■ Default vSwitch: ■ Only one vSwitch is specified as the default vSwitch in the zone that you select. ■ The default VPC uses a 20-bit subnet mask. For example, the CIDR block 172.16.0.0/20 provides up to 4,096 private IP addresses. ■ The default vSwitch does not count towards the quota of vSwitches that you can create in a VPC. ○ If the default VPC and vSwitch cannot meet your requirements, you can create your own VPC and vSwitch. For more information, see Create and manage a VPC.

Parameter	Description
Compatibility	◦ **MySQL 8.0**: fully compatible with MySQL 8.0. MySQL 8.0 supports parallel queries. In specific scenarios, the database performance increases by 10 times. For more information, see Parallel query. ◦ **MySQL 5.7**: fully compatible with MySQL 5.7. ◦ **MySQL 5.6**: fully compatible with MySQL 5.6. ◦ **PostgreSQL 11**: fully compatible with PostgreSQL 11. ◦ **Compatible with Oracle**: highly compatible with Oracle. For more information, see Oracle compatibility. **Note** **PostgreSQL 11** and **Compatible with Oracle** are not supported in the following regions: China (Qingdao), US (Virginia), UK (London), and Australia (Sydney).
Edition	By default, this parameter is set to .
Node Specification	Select node specifications based on your requirements. All nodes in the PolarDB cluster are dedicated nodes with stable and reliable performance. For more information about compute node specifications, see Specifications and pricing.
Nodes	If the source cluster edition is , the system creates a primary node and a read-only node that have the same specifications. In this case, you do not need to specify this parameter. **Note** If the primary node fails, the system uses the read-only node as the primary node and creates another read-only node. For more information about read-only nodes, see Architecture.
Storage Cost	You do not need to specify this parameter. The system charges you on an hourly basis based on the amount of storage that is consumed by your data. For more information, see Specifications and pricing. **Note** You do not need to specify the storage capacity when you create a cluster. The system automatically scales storage resources based on data volume.
Enable TDE	Specify whether to enable Transparent Data Encryption (TDE). After TDE is enabled, encrypts the data files of your cluster. You do not need to modify the code to allow access to your cluster. However, TDE reduces the performance of your cluster by 5% to 10%. **Note** TDE cannot be disabled after it is enabled.
Cluster Name	Enter the name of the cluster. The name must meet the following requirements: ◦ The name cannot start with `http://` or `https://`. ◦ The name must be 2 to 256 characters in length. If this parameter is left empty, the system automatically generates a cluster name. You can change the cluster name after the cluster is created.

Parameter	Description
Resource Group	Select a resource group from available resource groups. For more information, see Create a resource group. **Note** A resource group is a group of resources that belong to an Alibaba Cloud account. Resource groups allow you to manage these resources in a centralized manner. A resource belongs to only one resource group. For more information, see Use RAM to create and authorize resource groups.

5. Specify the **Number** parameter and click **Buy Now**.

   Note You can create a maximum of 50 clusters at a time. This allows you to create multiple clusters in specific scenarios. For example, you can deploy multiple game servers at a time.

6. On the **Confirm Order** page, confirm your order information. Read and accept the terms of service, and then click **Buy Now**.

   After you complete the payment, it requires 10 to 15 minutes to create the cluster. Then, the newly created cluster is displayed on the **Clusters** page.

   Note
   - If nodes in your cluster are in the **Creating** state, the cluster is being created and unavailable. The cluster is available only when it is in the **Running** state.
   - Make sure that you have selected the region where the cluster is deployed. Otherwise, you cannot view the cluster.
   - We recommend that you purchase storage plans if you want to store a large volume of data. Storage plans are more cost-effective than pay-as-you-go storage. Larger storage plans provide more storage for lower costs. For more information, see .

What to do next

Configure a whitelist for a cluster

Related API operations

API	Description
CreateDBCluster	Creates a PolarDB cluster.
DescribeDBClusters	Queries PolarDB clusters.
DescribeDBClusterAttribute	Queries the detailed information about a specified cluster.
DescribeAutoRenewAttribute	Queries the auto-renewal of a specified subscription cluster.
ModifyAutoRenewAttribute	Configures auto-renewal for a specified subscription cluster.

6.2. Perform a temporary upgrade

For a cluster you subscribed to, you can temporarily upgrade the specifications of the cluster to meet the business peak requirements in the specified validity period of the temporary upgrade.

Prerequisites

- The cluster is subscription-based.
- No pending orders of renewal or specification changes are available for the cluster.

- No pending orders of temporary upgrades are available for the cluser.
- The product edition must be . Both and do not support this feature. For more information about the three editions, see Overview.

Background information

A temporary upgrade allows you to temporarily upgrade the specifications of a PolarDB cluster to improve overall performance. When the specified restoration time is reached, the specifications of the cluster are automatically restored to the status before the temporary upgrade.

> **Note** PolarDB clusters do not support temporary downgrades. For more information about downgrades, see Manually upgrade or downgrade a PolarDB cluster.

Considerations

- Transient connection errors may occur during the restoration process. Therefore, you must ensure that your application implements an automatic reconnection mechanism.
- The restoration time must be at least one day earlier than the expiration time of the cluster. For example, if a cluster expires on January 10, the restoration time must be on January 9 at latest.
- Regular methods of changing specifications are not supported during the temporary upgrade. For more information, see Manually upgrade or downgrade a PolarDB cluster.
- The minimum validity period for a temporary upgrade is one hour. We recommend that you set the validity period to up to 14 days.
- After you perform a temporary upgrade on a cluster, the cluster performance may not meet the business requirements or you may need to extend the restoration time. In this case, you can perform a maximum of one more temporary upgrade before the **restoration time** is reached. The restoration time that you specify for the second temporary upgrade cannot be earlier than the previous one.

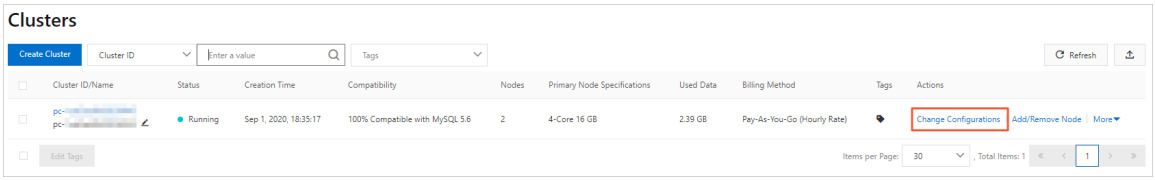
Pricing

The price of a temporary upgrade is 1.5 times the price difference between the original and new specifications. Assume that the validity period of the temporary upgrade is N days. You can calculate the price based on the following formula:

Price = (Monthly subscription fee for the new specifications - Monthly subscription fee for the original specifications)/30 × 1.5 × N

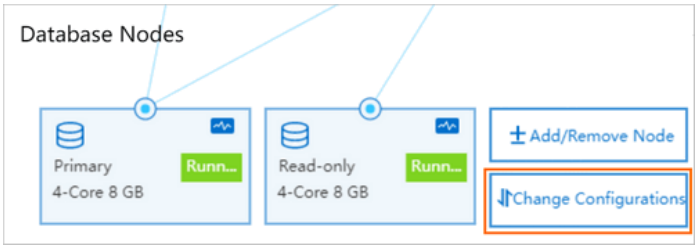
Procedure

1.
2.
3. On the **Clusters** page, find the cluster for which you want to perform a temporary upgrade.
4. Go to the **Change Configurations (Subscription)** dialog box by using the following two methods:
   - Find the cluster and click **Change Configurations** in the **Actions** column.



   - a. Find the cluster, and click the ID of the cluster to go to the **Overview** page.

b. In the **Database Nodes** section, click **Change Configurations**.



5. In the **Change Configurations (Subscription)** dialog box, select **Temporary Upgrade** and click **OK**.

> **Note** **Temporary Upgrade** is available for only subscription clusters.

6. On the page that appears, specify the following parameters.

Parameter	Description
Node	The specifications that you want to select for the upgraded node.
Restore Time	The time when the temporary upgrade expires. When the time is reached, the cluster is restored to the original specifications. **Note** ◦ After you perform a temporary upgrade on a cluster, the cluster performance may not meet the business requirements. In this case, you can perform a maximum of one more temporary upgrade before the restoration time is reached. The **restoration time** that you specify for the second temporary upgrade cannot be earlier than the previous one. ◦ The minimum validity period for a temporary upgrade is one hour. We recommend that you set the validity period to up to 14 days. It is because the restoration time cannot be changed after it is specified. ◦ The restoration time must be at least one day earlier than the expiration time of the cluster.

7. Read the service agreement, select the check box, click **Buy Now**, and then pay for the order.
8. On the **Purchase** page, confirm the unpaid order and click **Purchase**.

6.3. Change the specifications of a PolarDB cluster

This topic describes how to upgrade or downgrade the specifications of a PolarDB cluster. It takes only 5 to 10 minutes for the new specifications of each node to take effect.

Prerequisites

You can change cluster specifications only when the cluster does not have pending specification changes.

Context

This topic describes how to change the specifications of a cluster to meet your business requirements. PolarDB supports capacity scaling in three dimensions:

- Vertically scale the computing capacity: Upgrade or downgrade the specifications of the cluster. This topic describes the details of vertically scaling.
- Horizontal scaling of computing capabilities: Add or remove read-only nodes. For more information, see Add or remove a read-only node.

- Horizontal scaling of storage capacity: The storage capacity is provisioned in a serverless model. The storage capacity is automatically scaled based on the amount of data.

Configuration change fees

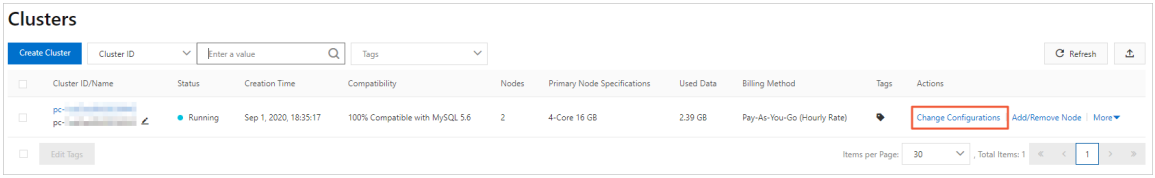
For more information, see Configuration change fees.

Precautions

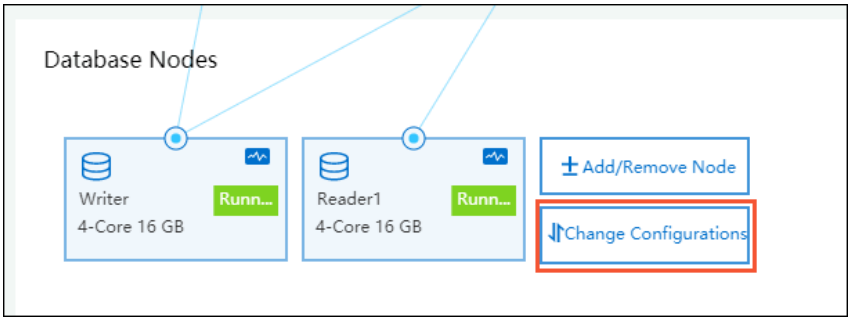
- You can upgrade or downgrade only clusters. You cannot upgrade or downgrade a single node in a cluster.
- When you upgrade or downgrade a cluster, data stored in the cluster is not affected.
- During the upgrade or downgrade process, your applications are temporarily disconnected from each endpoint for no more than 30 seconds. We recommend that you upgrade or downgrade your cluster during off-peak hours and make sure that your applications are configured with the automatic reconnection mechanism.

Procedure

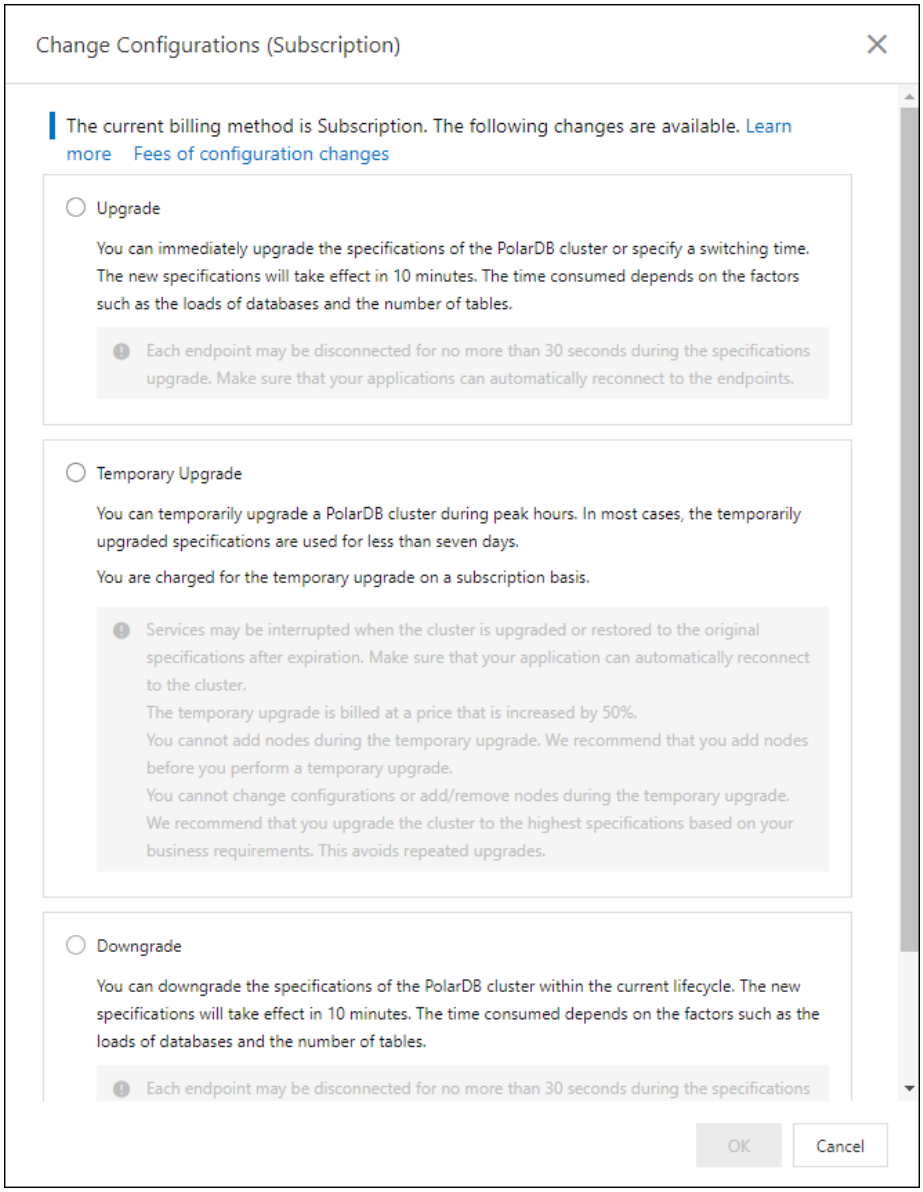
1. Log on to the PolarDB console.
2. In the upper-left corner of the console, select the region in which the cluster that you want to manage is deployed.
3. Go to the **Change Configurations** page by using one of the following two methods:
   - Find the cluster whose specifications you want to change, and click **Change Configurations** in the **Actions** column.



   - Click the ID of the cluster whose specifications you want to change. Then, in the Database Nodes section of the **Overview** page, click **Change Configurations**.



4. Select **Upgrade** or **Downgrade** and click **OK**.



> **Note** Only subscription clusters support **Temporary Upgrade**. For more information, see Perform a temporary upgrade.

5. Select the desired specifications and switching time.

> **Note** All nodes in a cluster must use the same specifications.

6. Read and agree to the terms of the service. Then, click **Buy Now** and complete the payment.

> **Note** The new specifications take effect within 10 minutes.

Related API operations

Operation	Description
ModifyDBNodeClass	Changes the node specifications of a cluster.

6.4. Add or remove a read-only node

After you create PolarDB clusters, you can manually add or remove read-only nodes to or from the clusters.

Context

A PolarDB cluster can contain a maximum of 15 read-only nodes. Each cluster must have at least one read-only node to ensure high availability. All the nodes in a cluster have the same specifications.

Billing methods of new nodes

You are charged for the nodes that are added to an existing cluster based on the following billing methods:

- If nodes are added to a **subscription** cluster, you are charged for the added nodes based on the **subscription** billing method.
- If nodes are added to a **pay-as-you-go** cluster, you are charged for the added nodes on a **pay-as-you-go** basis.

> **Note**
> - You can release read-only nodes whose billing methods are subscription and pay-as-you-go. After you release the nodes, the system refunds fees for the remaining subscription period or stops billing. For more information, see Configuration change fees.
> - Only the node specifications of the added nodes are charged. For more information, see Billable items. The storage fee is charged based on the amount of the used storage space, regardless of the number of nodes.

Considerations

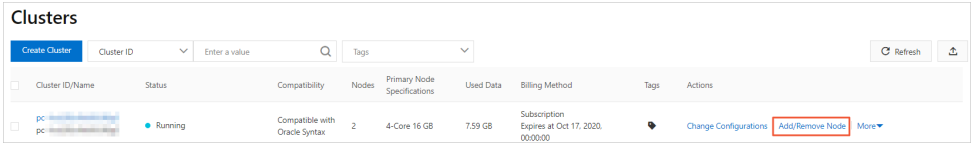
- You can add or remove read-only nodes only when the cluster does not have pending configuration changes.
- To avoid misoperations, you can add or remove only one read-only node at a time. To add or remove multiple read-only nodes, you must repeat the operation for multiple times.
- It takes about 5 minutes to add or remove a node.

Add a read-only node

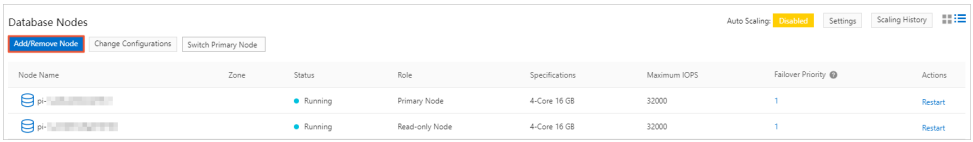
> **Note** A read/write splitting connection that is created after a read-only node is added forwards requests to the read-only node. A read/write splitting connection that is created before a read-only node is added does not forward requests to the read-only node. You must close the current connection and establish the connection again. For example, you can restart the application to establish the connection.

1.
2.
3. Open the **Add/Remove Node** dialog box by using one of the following methods:
   - Open the **Add/Remove Node** dialog box on the **Clusters** page.

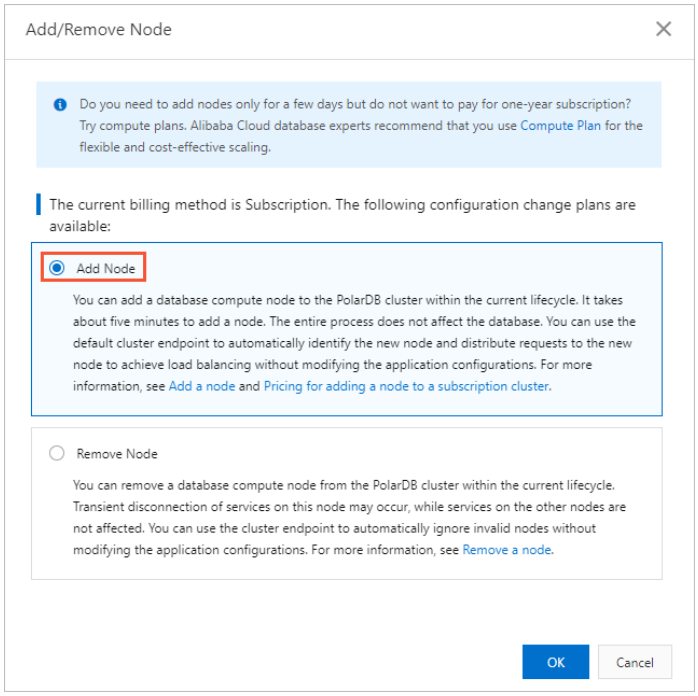
     Find the cluster that you want to manage and click **Add/Remove Node** in the **Actions** column.

   - Open the **Add/Remove Node** dialog box on the **Overview** page of the cluster.
     a. Find the cluster that you want to manage and click the cluster ID. The **Overview** page appears.
     b. In the **Database Nodes** section, click the icon to change the display mode.
     c. Click **Add/Remove Node**.



4. Select **Add Node** and click **OK**.

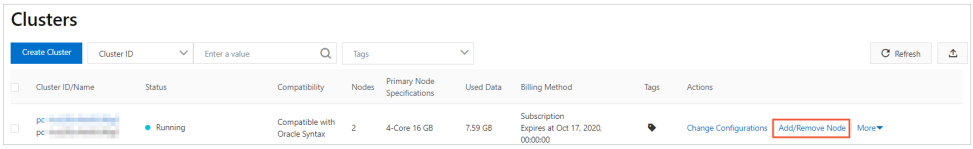


5. Click **+ Add a read-only node** to add a read-only node. Read and agree to the service agreement. To agree to the service agreement, select the check box. Then, click **Buy Now** and complete the payment.

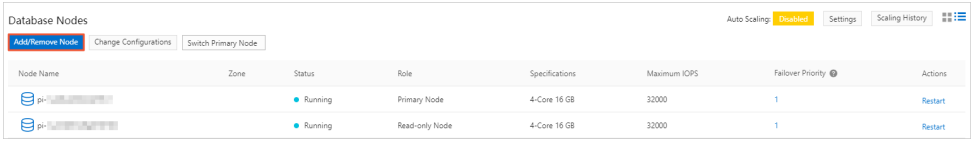
Remove a read-only node

1.
2.
3. Open the **Add/Remove Node** dialog box by using one of the following methods:
   - Open the **Add/Remove Node** dialog box on the **Clusters** page.

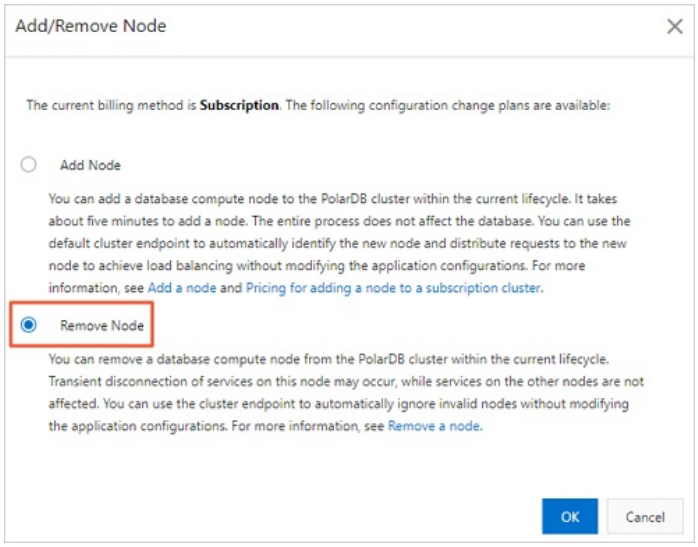
     Find the cluster that you want to manage and click **Add/Remove Node** in the **Actions** column.



   - Open the **Add/Remove Node** dialog box on the **Overview** page of the cluster.
     a. Find the cluster that you want to manage and click the cluster ID. The **Overview** page appears.
     b. In the **Database Nodes** section, click the icon to change the display mode.
     c. Click **Add/Remove Node**.



4. Select **Remove Node** and click **OK**.



5. Click the
   
   icon to the left of a node name to remove the node.

   > **Note** Only clusters support the concurrent removal of multiple read-only nodes. However, you must keep at least one read-only node in the cluster to ensure high availability.

6. Read and accept the terms of service, and then click **Buy Now**.

   > **Note** After a node is removed, the system refunds fees for the remaining subscription periods or stops billing. For more information, see Configuration change fees.

Related API operations

API	Description
CreateDBNodes	Adds a read-only node to a cluster.
ModifyDBNodeClass	Changes the node specifications of a cluster.
RestartDBNode	Restarts a specified node in a cluster.
DeleteDBNodes	Deletes read-only nodes from a cluster.

6.5. Set a maintenance window

This topic describes how to set a maintenance window for a cluster so that your business is not affected during the maintenance process.

Context

To ensure the stability of clusters, the backend system performs maintenance operations on the clusters from time to time. We recommend that you select a maintenance window within the off-peak hours of your business to minimize the impact on the business during the maintenance process.

Considerations

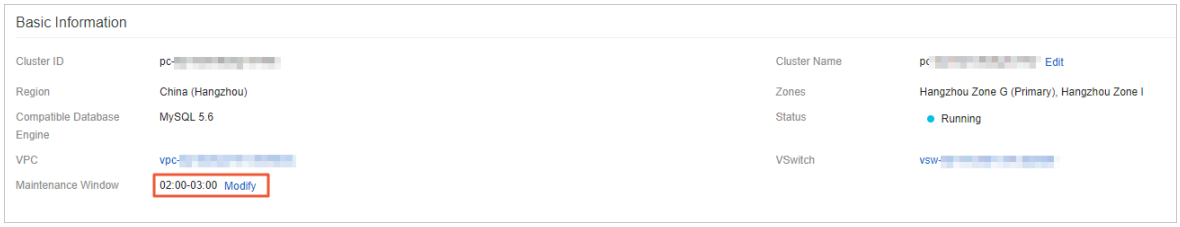
- Before the maintenance is performed on a cluster, sends SMS messages and emails to contacts listed in your Alibaba Cloud account.
- To ensure the stability of a cluster during the maintenance process, the cluster enters the Under Maintenance state

before the specified maintenance window starts. When the cluster is in the Under Maintenance state, you can access data in the databases of the cluster. However, features that are related to configuration changes become unavailable in the console except for the account management, database management, and whitelisting features. For example, you cannot upgrade, downgrade, or restart the cluster. Query features such as performance monitoring are still available.

- Within the maintenance window of a cluster, the cluster may experience one or two transient disconnections. Make sure that the application has an automatic reconnection mechanism. The cluster recovers to the normal state immediately after the disconnection.

Procedure

1.
2.
3.
4. On the **Overview** page, click **Modify** next to **Maintenance Window**.



5. In the Modify Maintenance Window dialog box, select a maintenance window, and click **OK**.

> **Note**
> - To ensure the stability of clusters, the backend system performs maintenance operations on the clusters from time to time. We recommend that you select a maintenance window within the off-peak hours of your business to minimize the impact on the business during the maintenance process.
> - Within the maintenance window of a cluster, the cluster may experience one or two transient disconnections. Make sure that the application has an automatic reconnection mechanism.

Related API operations

API operation	Description
CreateDBCluster	Creates a cluster.
ModifyDBClusterMaintainTime	Modifies the maintenance window for a cluster.

6.6. Restart nodes

allows you to restart nodes. When the number of database connections reaches the upper limit or the database performance is compromised, you can manually restart nodes.

Usage notes

- A read/write splitting connection that is established after a read-only node is restarted forwards requests to the read-only node. If a read/write splitting connection is established before a read-only node is restarted, the connection does not forward requests to the read-only node. You can restart your application to close the read/write splitting connection and establish the connection again.
- During the restart, services may be interrupted for up to 1 minute. We recommend that you perform this operation during off-peak hours and make sure that your application is configured to automatically reconnect to the database service.
- The time required to restart a node depends on the data volume. Several hours may be required to restart a node. Proceed with caution.

Procedure

1.
2.
3.
4. In the upper-right corner of the **Database Nodes** section for the **Overview** page, click the icon to switch the display mode.
5. Find the node that you want to restart, and click **Restart** in the **Actions** column.

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

Related API operations

API	Description
RestartDBNode	Restarts a node of a cluster.

6.7. Release a cluster

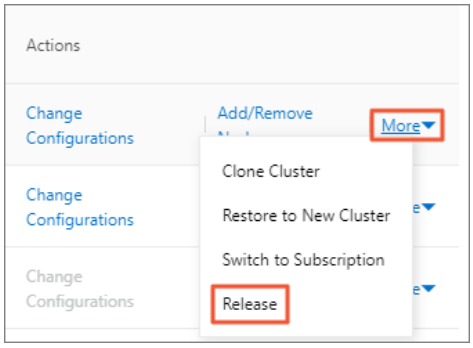
You can manually release pay-as-you-go clusters based on your business requirements. The pay-as-you-go clusters are charged on an hourly basis. This topic describes how to manually release clusters.

Considerations

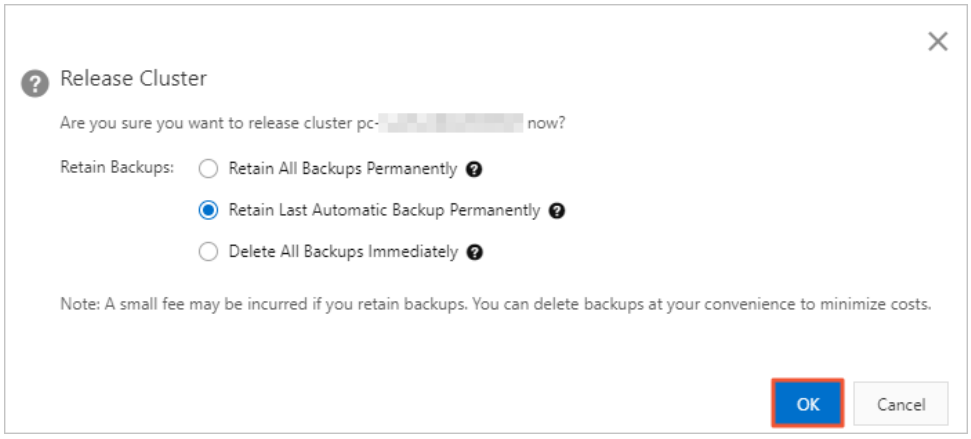
- You cannot manually release **subscription** clusters. Subscription clusters are automatically released when they expire.
- You can manually release the cluster whose **Status** is only **Running**.
- You can use this feature to release clusters. If this feature is used, all the nodes of the clusters are released. For more information about how to release a single read-only node, see Add or remove a read-only node.

Procedure

1.
2.
3. On the **Clusters** page, find the cluster that you want to release and choose **More > Release** in the **Actions** column.



4. In the **Release Cluster** dialog box, select a backup retention policy and click **OK**.



Retain backups	Description
Permanently Retain All Backups	Retains all the backups for a cluster when you delete the cluster.
Permanently Retain Last Automatic Backup	Retains the last backup for a cluster when you delete the cluster.
Immediately Delete All Backups	Deletes all the backups for a cluster when you delete the cluster. **Warning** If you select this policy, the deleted clusters cannot be restored.

Note

- If you select the **Permanently Retain All Backups** or **Permanently Retain Last Automatic Backup** policy, the system runs an automatic backup task to retain all the data about a cluster when you delete the cluster.
- After you delete a cluster, level-1 backups are automatically transferred to level-2 backups. You can go to the **Cluster Recycle** page to view retained backups. For more information, see Restore a released cluster.

Related API operations

API	Description
DescribeDBClusters	Queries PolarDB clusters.
DeleteDBCluster	Deletes a specified PolarDB cluster.

6.8. Cluster lock feature

You can enable the cluster lock feature for your pay-as-you-go clusters to prevent potential irreversible consequences arising from accidental manual release of the clusters. This topic describes how to enable or disable the cluster lock feature.

Prerequisites

The billing method of the cluster is pay-as-you-go.

Precautions

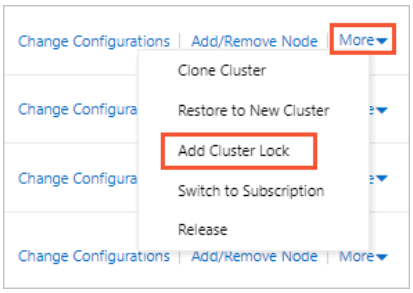
- The billing method of clusters with the cluster lock feature enabled cannot be changed to subscription.

- The cluster lock feature cannot prevent the automatic release of clusters in normal cases such as the following ones:
  - A payment in your account is overdue for more than eight days.
  - The cluster does not comply with the applicable security compliance policies.

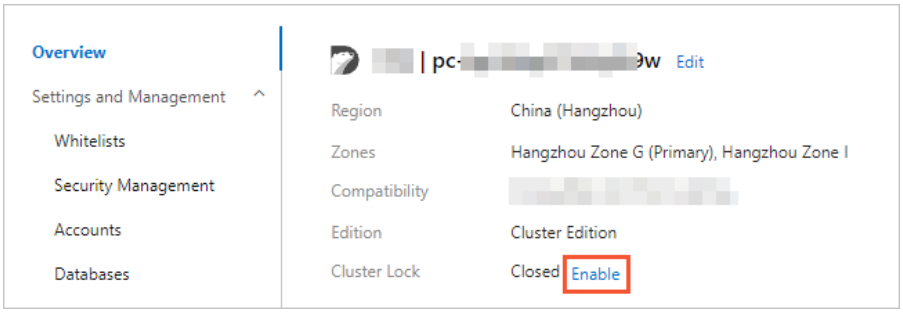
Enable the cluster lock feature

1.
2.
3. You can use one of the following methods to enable the cluster lock feature:
   - Method 1:

     On the **Clusters** page, find the cluster and choose **More > Add Cluster Lock** in the **Actions** column.

   - Method 2:
     a. On the **Clusters** page, click the cluster.
     b. On the **Overview** page, click **Enable** next to **Cluster Lock**.

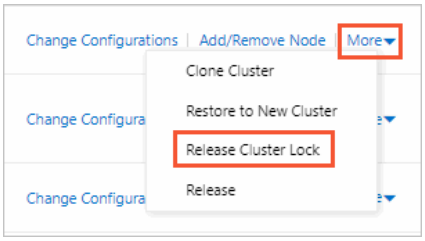
     

4. In the message that appears, click **OK**.

Disable the cluster lock feature

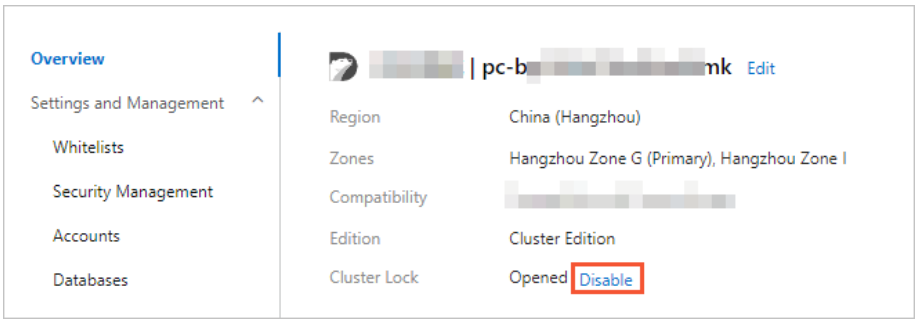
1.
2.
3. You can use one of the following methods to disable the cluster lock feature:
   - Method 1:

     On the **Clusters** page, find the cluster and choose **More > Release Cluster Lock** in the **Actions** column.

   - Method 2:
     a. On the **Clusters** page, click the cluster.

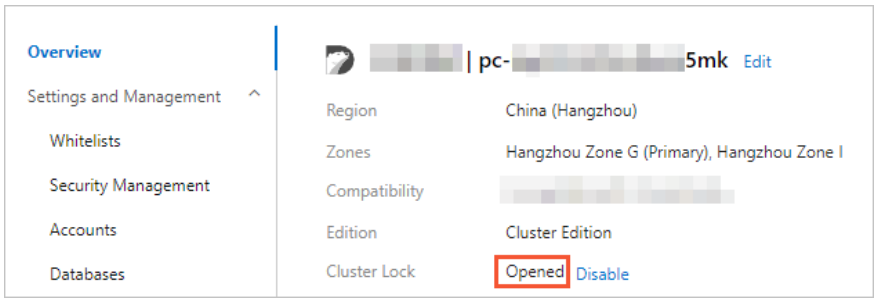
b. On the **Overview** page, click **Disable** next to **Cluster Lock**.



4. In the message that appears, click **OK**.

View the status of the cluster lock feature

1.
2.
3. On the **Clusters** page, click the cluster.
4. On the **Overview** page, view the status of **Cluster Lock**.



Related API operations

Operation	Description
	Enables or disables the cluster lock feature.

6.9. Automatic failover and manual failover

When a system failure occurs, a cluster can automatically fail over services from the primary node to a read-only node. You can specify a read-only node as the new primary node to fail over services from the primary node to the read-only node.

Precautions

During an automatic failover or a manual failover, transient disconnections may occur. Each transient disconnection lasts 20s to 30s. Make sure that your applications can be automatically reconnected to the cluster.

Automatic failover

clusters use an active-active architecture that ensures high availability. If the primary node that supports reads and writes is faulty, services are automatically failed over to the read-only node that is elected by the system as the new primary node.

A failover priority is assigned by the system to each node in a cluster. During a failover, a node is elected as the primary node based on the probability that is determined by this priority. The probability of being elected as the primary node is the same for the nodes that are assigned the same failover priority.

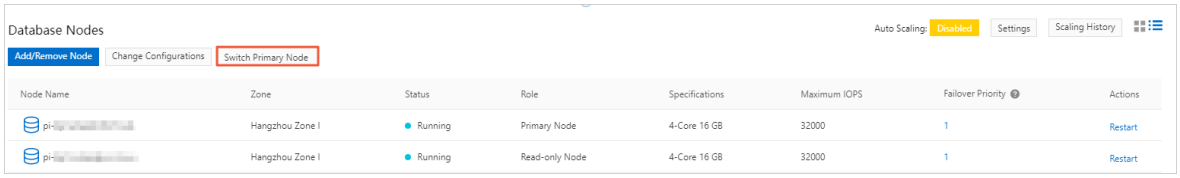
The system performs the following steps to promote a read-only node to the primary node:

1. Find all the read-only nodes that can be promoted to the primary node.

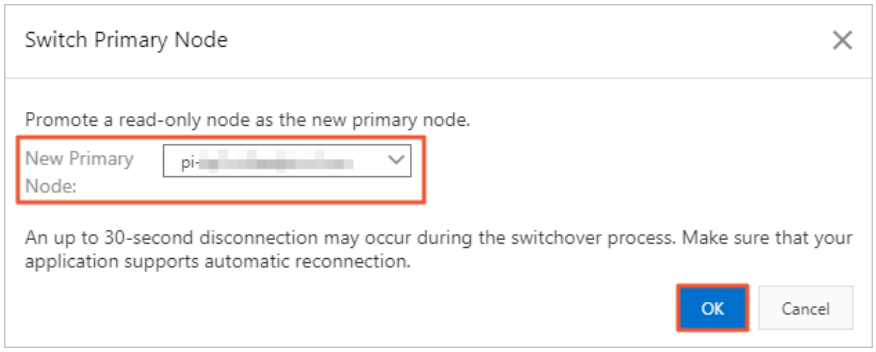
2. Select one or more read-only nodes that have the highest failover priority.
3. If the first node fails to be promoted to the primary node due to network or replication errors, the system attempts to promote the next available node. The system continues this process until the failover is successful.

Manual failover

1. Log on to the PolarDB console.
2.
3.
4. In the upper-right corner of the **Database Nodes** section on the **Overview** page, click the icon to switch the display mode.
5. Click **Switch Primary Node**.



6. In the dialog box that appears, select a new primary node from the **New Primary Node** drop-down list and click **OK**.



Related API operations

API	Description
FailoverDBCluster	Performs a manual failover by promoting a read-only node to a new primary node in a cluster.

6.10. Upgrade the minor version

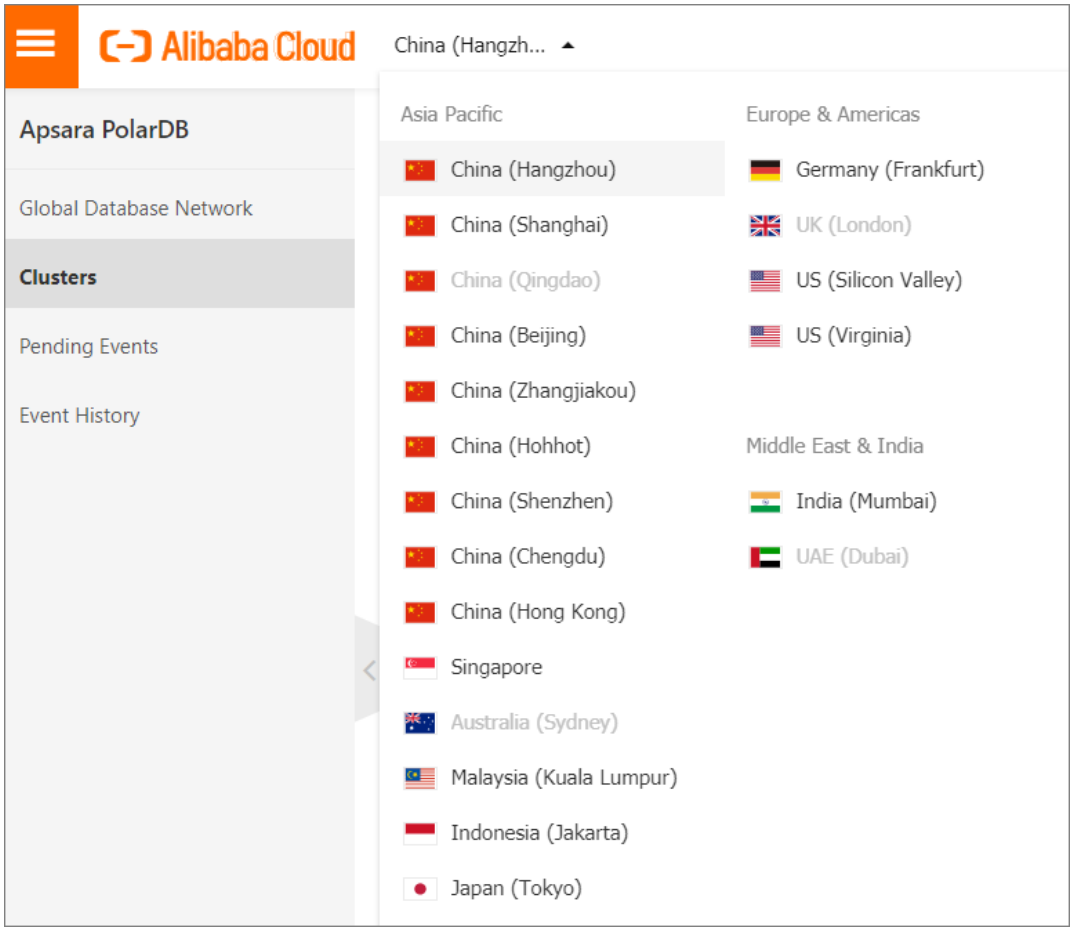
You can manually upgrade the minor kernel version of ApsaraDB for PolarDB cluster that is compatible with Oracle databases. The upgrades improve performance, provide new feature, or fix bugs.

Precautions

- Upgrading the kernel minor version will restart the instance. We recommend that you perform the upgrade during off-peak hours or make sure that your applications can automatically reconnect to the instance.
- You cannot downgrade the minor version after an upgrade.

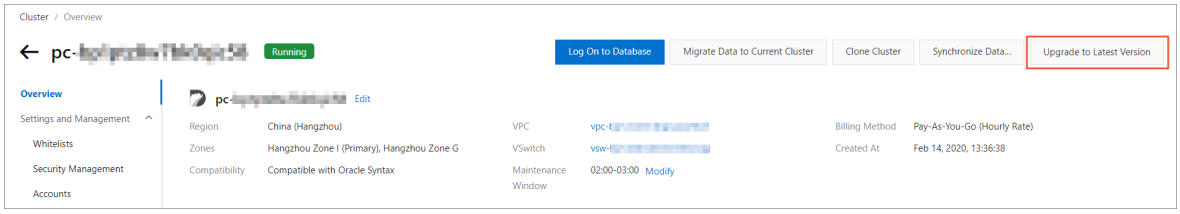
Procedure

1. Log on to the PolarDB console.
2. In the upper-left corner of the page, select the region where the PolarDB cluster is located.



3. Find the target cluster and click the cluster ID.
4. In **Basic Information**, click **Upgrade to Latest Version**.

> **Note** If your cluster kernel version is already the latest, the **Upgrade to Latest Version** button is not displayed.



5. In **Upgrade to Latest Version** dialog box, click **OK**.

> **Note** During the upgrade, services may be interrupted for about 60 seconds. Make sure that your applications can automatically reconnect to the instance.

6.11. Deploy a cluster across zones and change the primary zone

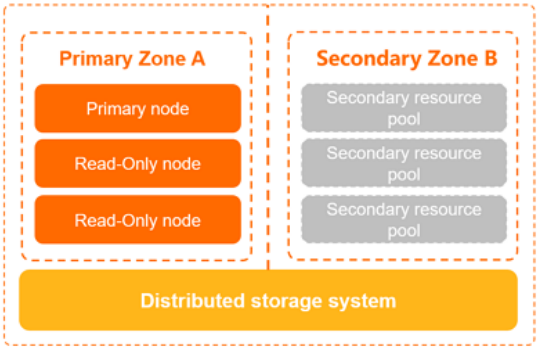
Apsara allows you to create multi-zone clusters. Compared with single-zone clusters, multi-zone clusters have better disaster recovery capabilities and can withstand breakdowns in data centers. This topic describes how to deploy a cluster across multiple zones and change the primary zone.

Prerequisites

- The region must contain at least two zones.
- The zones must have sufficient computing resources.

Multi-zone architecture

When a multi-zone cluster is deployed, data is distributed across zones. Compute nodes must be deployed in the primary zone. reserves sufficient resources in a secondary zone to ensure a successful failover when the primary zone fails. The following figure shows the multi-zone architecture.



Billing

No additional fee is required for multi-zone deployment.

> **Note** You can upgrade a single-zone cluster to a multi-zone cluster for free.

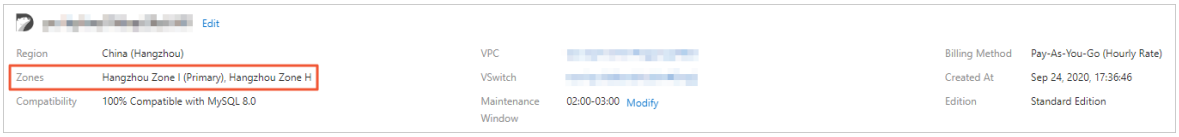
Establish a multi-zone architecture

If the prerequisites are met, a multi-zone cluster is created when you Create a cluster.

You can also upgrade an existing single-zone cluster to a multi-zone cluster. The upgrade is automatically completed by online migration, and does not affect your workloads.

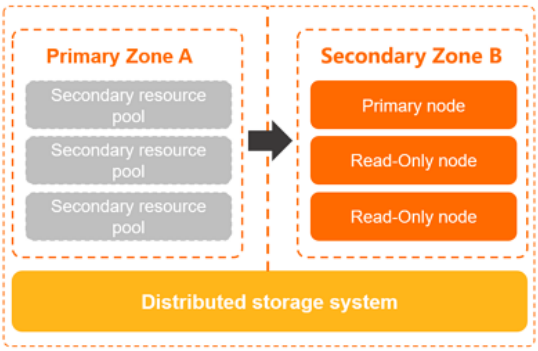
View the zones of a cluster

1. Log on to the Apsara PolarDB console.
2. In the upper-left corner of the console, select the region where the target cluster is deployed.
3. Click the ID of the cluster that you want to manage.
4. On the **Overview** page, view **Zones**.

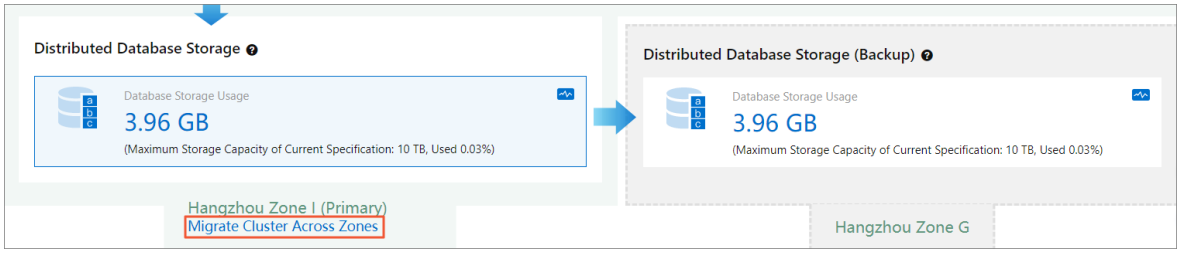


Change the primary zone

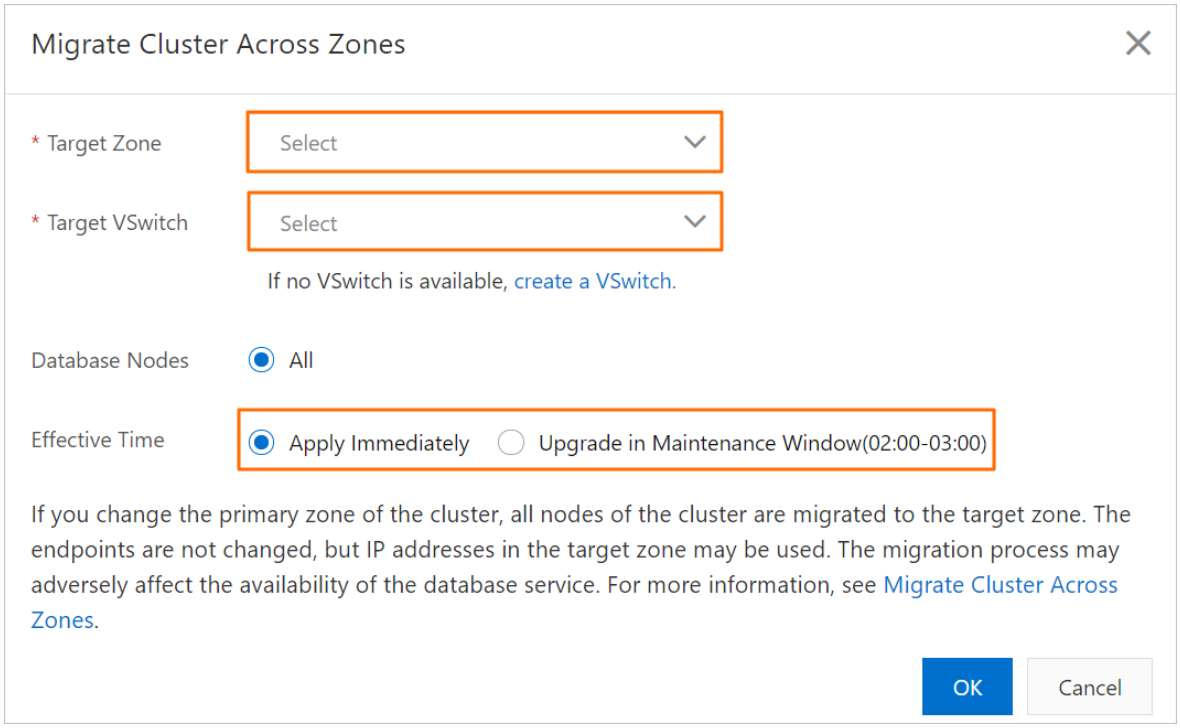
You can change the primary zone of an cluster. This feature allows you to migrate the compute nodes of a database cluster to a different zone. This is applicable to scenarios such as disaster recovery or when an Elastic Compute Service (ECS) instance is required to access the cluster in a nearby zone.



1.
2.
3.
4. On the **Overview** page, click **Migrate Cluster Across Zones**.



5. In the dialog box that appears, specify **Target Zone** and **Target VSwitch**, and set **Effective Time** base on your business requirements.



> **Note**
>
> - If the destination zone is a secondary zone, data migration is not required. Switching to a new secondary zone is fast because only compute nodes are switched. The average time required to migrate a compute node is five minutes. This operation is often performed during disaster recovery drills.
> - If the destination zone is not a secondary zone, data must be migrated. This migration process may take several hours depending on the data size. Proceed with caution. This operation is used to adjust the zones of applications and databases to speed up access from a nearby zone.

6. Click **OK**.

> **Notice** After the primary zone is changed, the primary endpoints and cluster endpoints remain unchanged, but the vSwitch and IP address may be changed. This operation may disrupt your database service for less than 60 seconds. Proceed with caution.

7.Account management

7.1. Account overview

This topic introduces the basic concepts of PolarDB console accounts and PolarDB cluster accounts.

Console accounts

You can use the following accounts to log on to the console:

- Alibaba Cloud account: This account allows flexible control of all your Alibaba Cloud resources and is used for billing purposes. You must create an Alibaba Cloud account before you purchase Alibaba Cloud services.
- RAM user: Optional. You can create and manage RAM users in the Resource Access Management (RAM) console to share resources among multiple users. A RAM user does not have ownership over resources. Charges incurred are billed to the Alibaba Cloud account.

Database cluster accounts

You can use the following accounts to log on to your database cluster. For more information, see Create database accounts.

Account type	Description
Privileged account	• You can use the ApsaraDB for PolarDB console or API operations to create and manage privileged accounts. • You can create multiple privileged accounts for each cluster. You can use the privileged accounts to manage all the standard accounts and databases of the corresponding cluster. • A privileged account has more permissions than before. This allows you to implement fine-grained control over user permissions based on your business requirements. For example, you can grant different users the permissions to query different tables. • A privileged account has all the permissions on the databases in the corresponding cluster. • You can use a privileged account to disconnect accounts from the corresponding databases.
Standard account	• You can use the ApsaraDB for PolarDB console, API operations, or SQL statements to create and manage standard accounts. • You can create multiple standard accounts for each cluster. The maximum number of standard accounts that you can create depends on the database engine. • You must manually grant standard accounts the specific database permissions. • You cannot use a standard account to create, manage, or disconnect other accounts from databases.

Related API operations

API	Description
CreateAccount	Creates an account.
DescribeAccounts	Queries the accounts of a specified cluster.
ModifyAccountDescription	Modifies the description of a database account for a PolarDB cluster.
ModifyAccountPassword	Changes the password of a database account.
DeleteAccount	Deletes an account.

7.2. Register and log on to an Alibaba Cloud account

This topic describes how to register and log on to an Alibaba Cloud account.

Register an Alibaba Cloud account

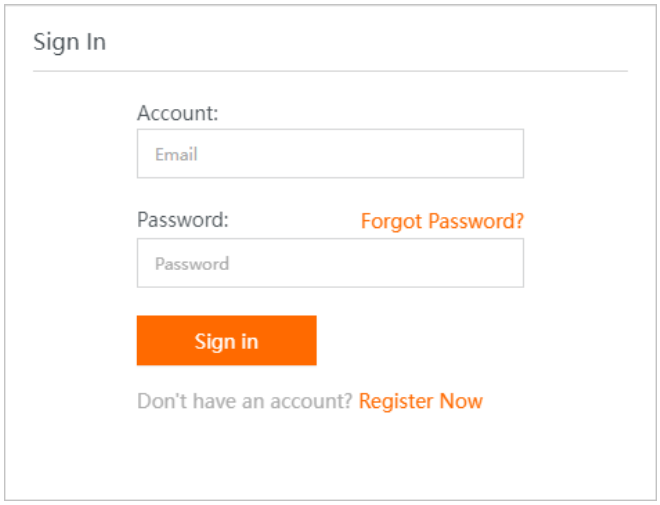
You can register an Alibaba Cloud account by using the following two methods:

- On the Alibaba Cloud International site, click **Free Account** in the upper-right corner.
- Directly go to the Alibaba Cloud account registration page.

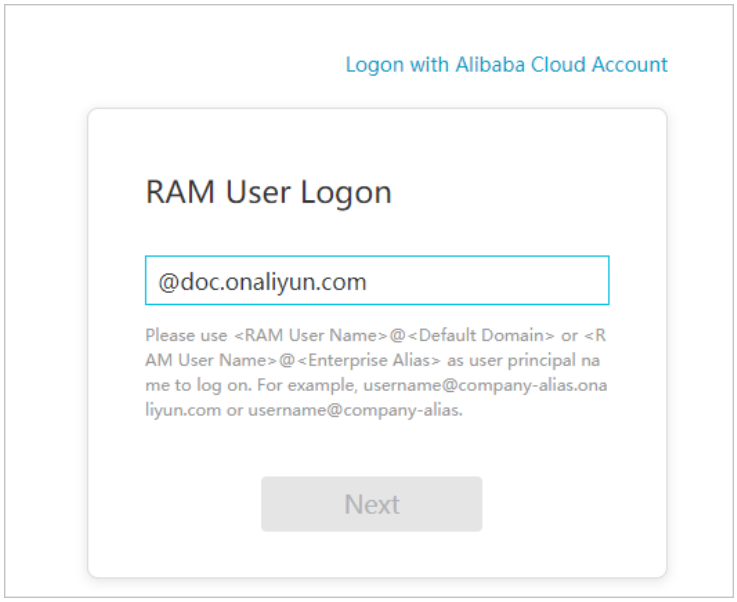
Log on to your Alibaba Cloud account

Your Alibaba Cloud account and Resource Access Management (RAM) user have different logon pages.

- The following figure shows the logon page for an Alibaba Cloud account.



- The following figure shows the logon page for a RAM user.



7.3. Create and authorize a RAM user

You can use your Alibaba Cloud account to access your resources. If you want to share the resources within your Alibaba Cloud account with other users, you must create and authorize Resource Access Management (RAM) users. After the authorization, the RAM users can access the specified resources. This topic describes how to create and authorize a RAM user.

Prerequisites

Log on to the console by using an Alibaba Cloud account or as a RAM user.

- For more information about how to use an Alibaba Cloud account, see Log on to the console with an Alibaba Cloud account.
- For more information about how to log on as a RAM user, see Log on to the console as a RAM user.

> **Note** The username of a RAM user must be in the format of `RAM username@enterprise alias`.

Procedure

- Create a RAM user. For more information, see Create a RAM user.
- Grant permissions to a RAM user on the Users page. For more information, see Grant permissions to a RAM user on the Users page.
- Grant permissions to a RAM user on the Grants page. For more information, see Grant permissions to a RAM user on the Grants page.
- Log on to the console as a RAM user. For more information, see Log on to the Alibaba Cloud Management Console as a RAM user.

Related operations

You can also add a RAM user to a group, assign roles to a RAM user, and authorize a user group or roles. For more information, see RAM User Guide.

7.4. Create database accounts

This topic describes how to create a database account. This topic also explains the difference between a privileged account and a standard account.

Context

You can create two types of database accounts in ApsaraDB for PolarDB: privileged account and standard account. You can use the ApsaraDB for PolarDB console to manage all the database accounts.

> **Note** You cannot create root accounts in ApsaraDB for PolarDB because of security reasons.

Account type	Description
Privileged account	• You can use the ApsaraDB for PolarDB console or API operations to create and manage privileged accounts. • You can create multiple privileged accounts for each cluster. You can use the privileged accounts to manage all the standard accounts and databases of the corresponding cluster. • A privileged account has more permissions than before. This allows you to implement fine-grained control over user permissions based on your business requirements. For example, you can grant different users the permissions to query different tables. • A privileged account has all the permissions on the databases in the corresponding cluster. • You can use a privileged account to disconnect accounts from the corresponding databases.

Account type	Description
Standard account	• You can use the ApsaraDB for PolarDB console, API operations, or SQL statements to create and manage standard accounts. • You can create multiple standard accounts for each cluster. The maximum number of standard accounts that you can create depends on the database engine. • You must manually grant standard accounts the specific database permissions. • You cannot use a standard account to create, manage, or disconnect other accounts from databases.

Create an account

1.
2.
3.
4. In the left-side navigation pane, choose **Settings and Management > Accounts**.
5. On the page that appears, click **Create Account**.
6. In the Create Account pane, configure the following parameters.

Parameter	Description
Account Name	Enter an account name. The account name must meet the following requirements: ◦ It must start with a lowercase letter and end with a letter or a digit. ◦ It can contain lowercase letters, digits, and underscores (_). ◦ It must be 2 to 16 characters in length. ◦ It cannot be a system reserved username, such as root or admin.
Account Type	◦ To create a privileged account, select **Privileged Account**. ◦ To create a standard account, select **Standard Account**.
Password	Enter an account password. The password must meet the following requirements: ◦ It must contain at least three of the following character types: uppercase letters, lowercase letters, digits, and special characters. ◦ It must be 8 to 32 characters in length. ◦ It can contain the following special characters: !@#$%^&*()_+-=
Confirm Password	Enter the password again.
Description	Enter the information about the account to facilitate subsequent account management. The description must meet the following requirements: ◦ It cannot start with http:// or https://. ◦ It must start with a letter. ◦ It can contain letters, digits, underscores (_), and hyphens (-). ◦ It must be 2 to 256 characters in length.

7. Click **OK**.

What to do next

View or apply for an endpoint

Related API operations

API	Description
CreateAccount	Creates a database account for a specified PolarDB cluster.
DescribeAccounts	Queries the database accounts for a specified PolarDB cluster.
ModifyAccountDescription	Changes the description of a database account for a specified PolarDB cluster.
ModifyAccountPassword	Changes the password of a database account for a specified PolarDB cluster.

7.5. Manage database accounts

This topic describes how to manage database accounts, such as how to modify the passwords of accounts and lock, unlock, and delete accounts.

Context

You can create two types of database accounts in PolarDB: privileged and standard accounts. You can manage all the accounts and databases in the PolarDB console.

Considerations

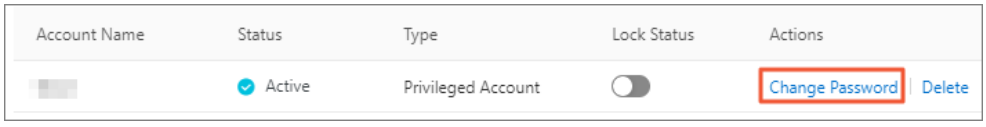
To ensure data security, you cannot create and use a root account in PolarDB.

Create a database account

For more information, see Create database accounts.

Change the password of an account

1.
2.
3.
4. In the left-side navigation pane, choose **Settings and Management > Accounts**.
5. Find the account for which you want to change the password, and click **Change Password** in the **Actions** column.

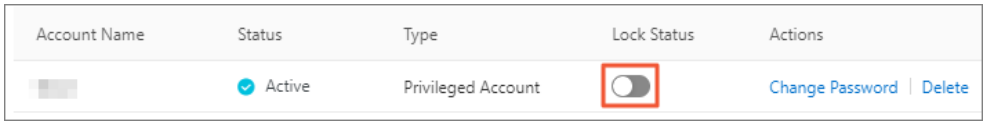


6. In the dialog box that appears, enter and confirm the new password, and click **OK**.

Lock or unlock an account

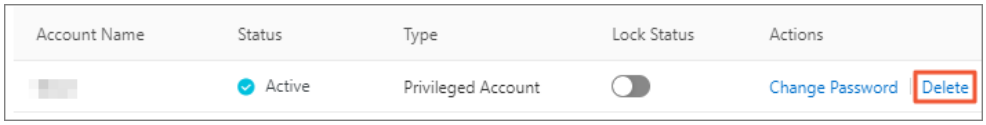
You can lock an account to prevent the account from logging on to the database.

1.
2.
3.
4. In the left-side navigation pane, choose **Settings and Management > Accounts**.
5. Find the account that you want to lock or unlock, and turn on or off the switch in the **Lock Status** column.



Delete an account

1.
2.
3.
4. In the left-side navigation pane, choose **Settings and Management > Accounts**.
5. Find the account that you want to delete, and click **Delete** in the **Actions** column.



6. In the message that appears, click **OK**.

Related API operations

API	Description
CreateAccount	Creates an account.
DescribeAccounts	Queries a list of accounts.
ModifyAccountDescription	Modifies the description of an account.
ModifyAccountPassword	Changes the password of an account.
DeleteAccount	Deletes an account.

8.DBLink

8.1. Overview

provides the database link (DBLink) feature for you to access data across databases.

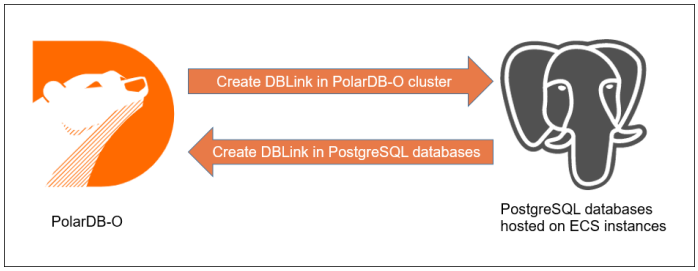
For example, you can create a database link of Database B in Database A. Then, Database A can use the database link to access the data stored in Database B in the same way as Database A accesses its own data. In this case, two-way data access cannot be implemented. If Database B needs to access the data stored in Database A, you must create a database link of Database A in Database B.

Scenarios

supports two types of database links: database links from some clusters to other clusters and database links from to user-created PostgreSQL databases hosted on Elastic Compute Service (ECS) instances.

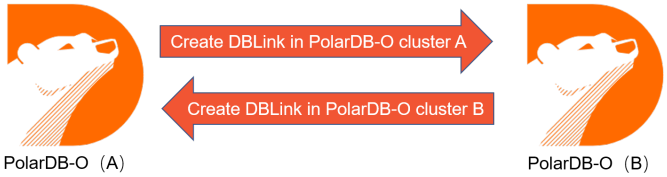
- Scenario 1: Database links from to user-created PostgreSQL databases hosted on ECS instances

  You previously used a user-created PostgreSQL database that is hosted on an ECS instance. When you migrated the services to , only part of the business systems were migrated to due to some reasons. In this scenario, the remaining services are still deployed in the user-created PostgreSQL database that is hosted on the ECS instance. Data access is required between the business systems that were migrated and the remaining services. Therefore, two-way data access must be implemented between the database and the user-created PostgreSQL database hosted on the ECS instance to ensure that your business runs as expected. To implement two-way data access, create a database link from the user-created PostgreSQL database hosted on the ECS instance to . You must also create a database link from to the user-created PostgreSQL database hosted on the ECS instance.

- Scenario 2: Database links from some clusters to other clusters

  Due to business requirements, you use two clusters. Assume that the two clusters are cluster A and cluster B and the two clusters store data sets of different services. Two-way data access is required between the data sets of different services. To implement two-way data access between the two databases, create a database link from cluster A to cluster B and a database link from cluster B to cluster A.

8.2. Create a database link from PolarDB for Oracle to PolarDB for Oracle

This topic describes how to create a database link from a cluster to a cluster.

For more information about how to create a database link from a cluster to a user-created Oracle database hosted on an Elastic Compute Service (ECS) instance, see Create a database link from PolarDB for Oracle to PostgreSQL.

Prerequisites

- The source database and the destination database belong to the same Alibaba Cloud account.

- The source database and the destination database are in the same region.
- The kernel version of the cluster is the latest kernel version. For information about how to upgrade the kernel version, see Upgrade the minor version.

Notes

You can create a maximum of 10 database links for each cluster. Each database link consumes one database link quota of both the source cluster and the destination cluster.

Procedure

1.
2.
3.
4. In the left-side navigation pane, choose **Settings and Management > Databases**.
5. On the **Databases** page, click the **DBLinks** tab in the upper part.
6. Click **Create DBLink** and configure the following parameters.

Parameter	Description
DBLink Name	You can enter a custom name of the database link. The name must meet the following requirements: ○ The name must contain lowercase letters and can contain digits and underscores (_). ○ It must start with a lowercase letter and end with a letter or a digit. ○ It must be 1 to 64 characters in length. This name is required when you use the database link for cross-database queries.
Source Instance Name	The name of the current cluster is used as the fixed value. You cannot modify this parameter.
Source Database Name	Select a database in the current cluster from the drop-down list.
Destination Instance	Select a destination cluster for the database link from the drop-down list. **Note** You can only select a cluster deployed in the same region as the source cluster.
Destination Account	The account that is used to access the destination cluster. For more information about how to create an account, see Create database accounts.
Destination Account Password	The password of the account that is used to access the destination cluster.
Destination Database Name	Enter the database name of the destination cluster. For more information about how to create a database, see Database.

7. Click **OK**.

Related operations

API	Description
CreateDBLink	Creates a database link for a cluster.
DescribeDBLinks	Queries the database link information of a cluster.

8.3. Create a database link from PolarDB for Oracle to PostgreSQL

This topic describes how to create a database link from to a user-created PostgreSQL database hosted on an Elastic Compute Service (ECS) instance.

For more information about how to create a database link from a cluster to a cluster, see Create a database link from PolarDB for Oracle to PolarDB for Oracle.

Prerequisites

- The cluster and the user-created PostgreSQL database hosted on ECS belong to the same Alibaba Cloud account.
- The ECS instance is deployed in a virtual private cloud (VPC).
- The kernel version of the cluster is the latest kernel version. For information about how to upgrade the kernel version, see Upgrade the minor version.
- to apply for creating database links from clusters to user-created PostgreSQL databases hosted on ECS in the console.

> **Note** You can call the CreateDBLink operation to create a database link from a cluster to a user-created PostgreSQL database. This way, you do not need to submit a ticket.

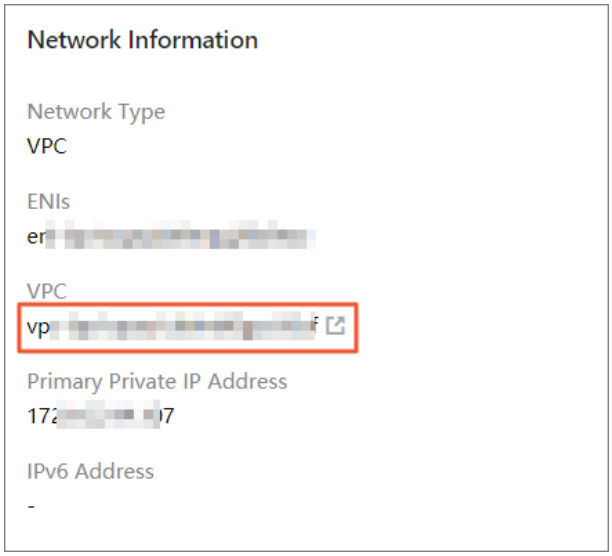
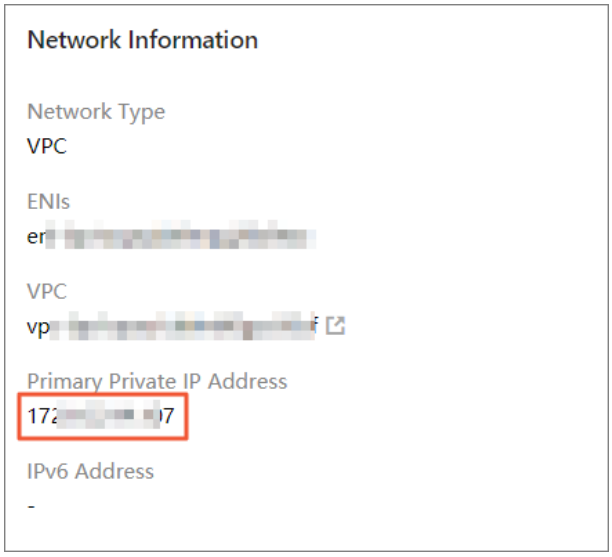
Notes

You can create a maximum of 10 database links for each cluster. Each database link consumes one database link quota of both the source cluster and the destination cluster.

Procedure

1.
2.
3.
4. In the left-side navigation pane, choose **Settings and Management > Databases**.
5. On the **Databases** page, click the **DBLinks** tab in the upper part.
6. Click **Create DBLink to User-create Database** and configure the following parameters.

Parameter	Description
DBLink Name	You can enter a custom name of the database link. The name must meet the following requirements: ○ The name must contain lowercase letters and can contain digits and underscores (_). ○ It must start with a lowercase letter and end with a letter or a digit. ○ It must be 1 to 64 characters in length. This name is required when you use the database link for cross-database queries.
Source Instance Name	The name of the current cluster is used as the fixed value. You cannot modify this parameter.
Source Database Name	Select a database in the current cluster from the drop-down list.

Parameter	Description
VPC of Destination Instance	The VPC where the ECS instance resides. You can view the VPC ID of the destination ECS instance in the **Network Information** section on the **Basic Information** page in the ECS console. 
Region ID of Destination Instance	The ID of the region where the ECS instance resides, such as cn-hangzhou. You can call the DescribeInstances operation to query the region ID of the destination ECS instance.
VPC ID of Destination Instance	The IP address of the ECS instance. You can view the primary private IP address of the destination ECS instance in the **Network Information** section on the **Basic Information** page in the ECS console. 
VPC Port of Destination Instance	The port number of the user-created PostgreSQL database hosted on ECS. The default port number is 1521.
Destination Account	The account of the user-created PostgreSQL database hosted on ECS.
Destination Account Password	The account password of the user-created PostgreSQL database hosted on ECS.

Parameter	Description
Destination Database Name	The name of the user-created PostgreSQL database hosted on ECS.

7. Click **OK**.

Related operations

API	Description
CreateDBLink	Creates a database link for a cluster.
DescribeDBLinks	Queries the database link information of a cluster.

8.4. Use a database link to query data across databases

provides the database link (DBLink) feature to query data across databases. This topic describes how to use a database link to query data across databases.

Prerequisites

- A source cluster and a destination cluster are created. For more information, see Create a cluster.

  > **Note** The DBLink feature of supports the following connection methods:
  > - The source database is in a cluster and the destination database is in a cluster.
  > - The source database is in a cluster. The destination database is a self-managed PostgreSQL database hosted on an ECS instance.

- Databases are created in the source and destination clusters. For more information, see Create a database.
- Tables are created in the destination database.

Note

You can query data of the destination cluster only from the source cluster. You cannot query data of the source cluster from the destination cluster.

Parameter configuration

provides the polar_enable_pushable_unsafe_collate_remote and polar_enable_pushable_all_any_remote parameters to improve query performance.

- polar_enable_pushable_unsafe_collate_remote: specifies whether to push down the functions that do not meet collation requirements. This parameter is a session-level parameter. The default value of this parameter is on.

  If you set this parameter to off, the functions that do not meet collation requirements are not pushed down.

  - To set this parameter to on, run the following command:

    ```
    SET polar_enable_pushable_unsafe_collate_remote = on;
    ```

  - To set this parameter to off, run the following command:

    ```
    SET polar_enable_pushable_unsafe_collate_remote = off;
    ```

- polar_enable_pushable_all_any_remote: specifies whether to push down ANY() and ALL() expressions. This parameter is a session-level parameter. The default value of this parameter is on.

  If you set this parameter to off, the ALL/ANY expressions that do not meet requirements are not pushed down.

- To set this parameter to on, run the following command:

```
SET polar_enable_pushable_all_any_remote = on;
```

- To set this parameter to off, run the following command:

```
SET polar_enable_pushable_all_any_remote = off;
```

Procedure

1. Create a database link.

   provides the following methods to create database links:

   - Create a database link from PolarDB for Oracle to PolarDB for Oracle.
   - Create a database link from PolarDB for Oracle to PostgreSQL.

2. Connect to the source cluster. For more information, see Connect to a PolarDB cluster.
3. Query data across databases from the source cluster.

   Execute the following query statement:

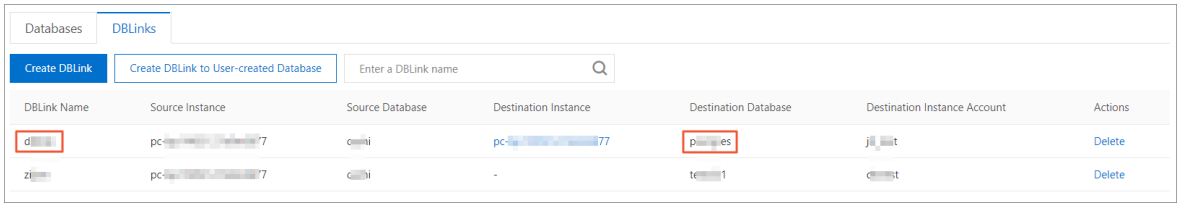
```
SELECT * FROM <dbname>@<dblinkname>;
```

   - *<dbname>*: the table in the database of the destination cluster.

     > **Note** The table must be stored in the destination database to which the database link is connected.

   - *<dblinkname>*: the name of the database link.

   You can view the name of the database link and the destination database to which the database link is connected in the console, as shown in the following figure.



Examples

To create a table named `test` in the destination database for testing and insert test data into the table, execute the following statements:

```
CREATE TABLE test(id int);
INSERT INTO test VALUES(1);
```

Connect to the source database and execute the following query statement:

```
SELECT * FROM test@dblinkname;
```

The following query result is returned:

```
 id
----
  1
(1 row)
```

8.5. Delete a database link

This topic describes how to delete a database link of a cluster.

Notes

After you delete the database link, you cannot access the destination database from the source database. To ensure that your business is not affected, ensure that the source database does not need to access the destination database before you delete the database link.

Procedure

1.
2.
3.
4. In the left-side navigation pane, choose **Settings and Management > Databases**.
5. On the **Databases** page, click the **DBLinks** tab in the upper part.
6. Find the database link that you want to delete and click **Delete** in the **Actions** column.
7. In the **Delete DBLink** dialog box, click **OK** to delete the database link.

Related operations

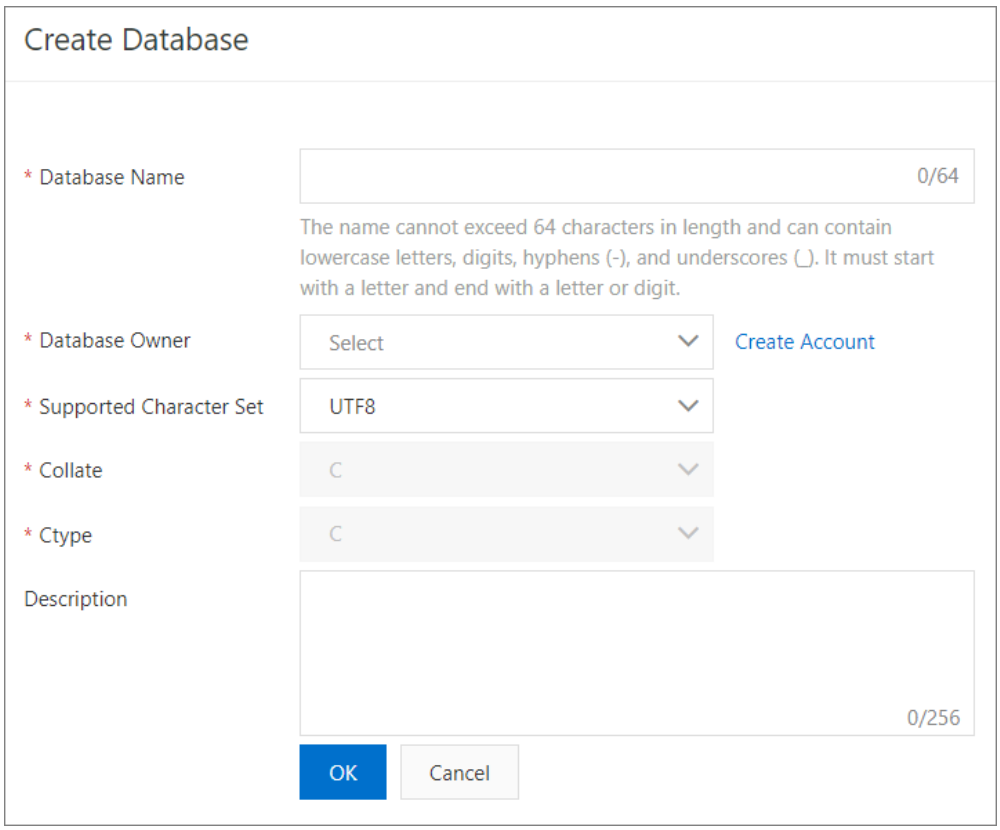
API	Description
DeleteDBLink	Deletes the database links of a cluster.

9.Database

This topic describes how to create and delete a database.

Create a database

1.
2.
3.
4. In the left-side navigation pane, choose **Settings and Management > Databases**.
5. Click **Create Database**.



6. In the Create Database panel, configure the following parameters.

Parameter	Description
Database Name	◦ The name of the database must start with a lowercase letter and end with a lowercase letter or digit. ◦ The database name can contain lowercase letters, digits, underscores (_), and hyphens (-). ◦ The database name must be 2 to 64 characters in length. ◦ The database name must be unique in a cluster. **Note** The database name cannot be `test` or another keyword that is reserved by the system.
Database Owner	The owner of the database. The owner is granted all permissions on the database.

Parameter	Description
Supported Character Set	The character set supported by the database. Default value: UTF8. You can select other required character sets from the drop-down list.
Collate	The rule based on which strings are sorted.
Ctype	The type of characters supported by the database.
Description	Enter a description for the database. The description can help you manage your database. The description that you specify must meet the following requirements: ○ The description cannot start with http:// or https://. ○ The description must start with a letter. ○ The description can contain letters, digits, underscores (_), and hyphens (-). ○ The description must be 2 to 256 characters in length.

7. Click **OK**.

Delete a database

1.
2.
3.
4. In the left-side navigation pane, choose **Settings and Management > Databases**.
5. Find the database that you want to delete, and click **Delete** in the Actions column.
6. In the message that appears, click **OK**.

Related API operations

API	Description
CreateDatabase	Creates a database.
DescribeDatabases	Views the database list.
DeleteDatabase	Deletes a database.

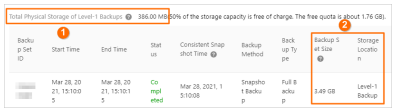
10.Backup and restoration

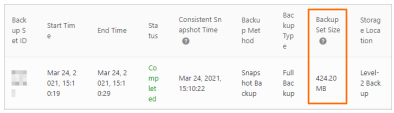
10.1. Overview

This topic describes how to enable to automatically create backups at specified intervals or manually create backups to prevent data loss in a timely manner. allows you to retain backups of a cluster when you delete the cluster.

Data backup

Data backups are divided into level-1 backups and level-2 backups by storage location.

Storage location	Default configuration	Retention period	Benefit	View backup size
Level-1 backup	Yes	3 to 14 days	• Level-1 backups are created based on Redirect-on-Write (ROW) snapshots. These snapshots are stored in the distributed file system of . The system does not replicate data when it saves a data block to a snapshot. When a data block is modified, the system saves one of the previous versions of the data block to a snapshot and creates a new data block that is redirected by the original data block. Therefore, you can create backups within a few seconds regardless of the size of your database storage. • The backup and restoration features of clusters use multi-threading parallel processing and other innovative technologies. This allows you to restore data from a backup set (snapshot) to a new cluster within 10 minutes. **Note** By default, the level-1 backup feature is enabled, and you cannot disable this feature.	The following figure shows the total physical storage of level-1 backups.  **Note** The **total size of level-1 backups** of a cluster is the sum of the dedicated physical storage occupied by all level-1 backups, as shown in part ①. It is not the sum of the logical data sizes of all level-1 backups, as shown in part ②. The data of the cluster and multiple level-1 backups (snapshots) can be stored in the same physical data block that is billed only once. For more information, see FAQ.

Storage location	Default configuration	Retention period	Benefit	View backup size
Level-2 backup	No	• 30 to 7,300 days • Enable the **Retained Before Cluster Is Deleted** feature to save level-2 backups permanently.	• Level-2 backups are level-1 backups that are compressed and then stored in on-premises storage. Level-2 backups are slower to restore than level-1 backups. However, level-2 backups are more cost-effective than level-1 backups. • If you enable this feature, expired level-1 backups are transferred to on-premises storage and stored as level-2 backups. The backups are transferred at a rate of approximately 150 MB/s. **Note** If a level-1 backup expires before the previous one is transferred to a level-2 backup, the level-1 backup is deleted and is not transferred to a level-2 backup. For example, a cluster creates level-1 backups at 01:00 every day and retains the backups for 24 hours. If the cluster creates Level-1 Backup A at 01:00 on January 1 and creates Level-1 Backup B at 01:00 on January 2. Level-1 Backup A expires at 01:00 on January 2 and starts to be transferred to a level-2 backup. However, Level-1 Backup A stores a large amount of data, and the transfer task is not completed by 01:00 on January 3. In this case, Level-1 Backup B is deleted after it expires at 01:00 on January 3 and is not transferred to a level-2 backup.	The following figure shows the total size of level-2 backups. The total size of level-2 backups is the sum of the data sizes of all level-2 backups. 

Physical log backup

- Benefits

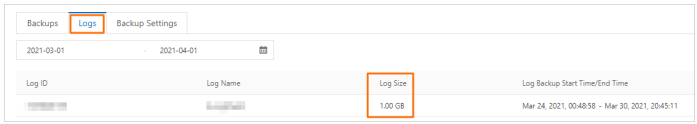
  The log backup feature allows you to create backups by uploading real-time redo logs to Object Storage Service (OSS) in parallel. The feature is enabled by default, and log backups are retained for 3 to 7,300 days. You can save the backups permanently by enabling the **Retained Before Cluster Is Deleted** feature.

  Note By default, log backup is enabled, and you cannot disable this feature.

  Log backups help consistent point-in-time recovery. Based on a full backup set (snapshot) and the redo logs generated after the backup set is created, you can perform point-in-time recovery (PITR) for a cluster. Log backups can prevent data loss caused by user errors and ensure the security of data that is generated within a period of time. If you perform PITR, you must consider the amount of time that is required to query redo logs. Redo logs are queried at a rate of 1 GB every 20 seconds to 70 seconds. The total restoration duration is the sum of the time required to restore backup sets and the time required to query redo logs.

- View backup size

  The following figure shows that the total size of log backups is the sum of the size of each log backup file.



10.2. Billing

This topic describes the billing of the backup and restoration feature.

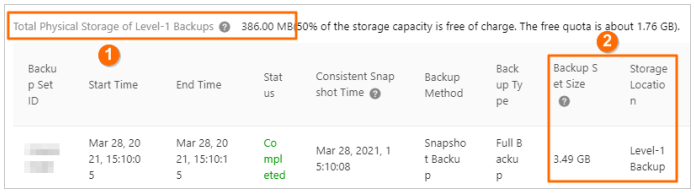
The backup and restoration feature is free of charge. Only storage fees are charged. In , fees are calculated based on the storage consumed by backups (data and logs) and the retention period of these backups.

Pricing

Pricing

Region	Level-1 backup	Level-2 backup	Log backup
Chinese mainland	USD 0.000464 per GB-hour	USD 0.0000325 per GB-hour	USD 0.0000325 per GB-hour
China (Hong Kong) and regions outside China	USD 0.000650 per GB-hour	USD 0.0000455 per GB-hour	USD 0.0000455 per GB-hour

Billing methods

Backup type	Free quota	Billing method
Level-1 backup	Database storage usage × 50% You can view the database storage usage of a cluster on the **Overview** page of the cluster in the PolarDB console.	Storage fee per hour = (Total physical storage of level-1 backups - Free quota) × Unit price per hour • You are not charged if the physical storage of the level-1 backups does not exceed the free quota. • For more information about the unit price per hour, see Pricing. • The following figure shows you how to view **Total Physical Storage of Level-1 Backups** in the console.  **Note** The **total size of level-1 backups** of a cluster is the sum of the dedicated physical storage occupied by all level-1 backups, as shown in part ①. It is not the sum of the logical data sizes of all level-1 backups, as shown in part ②. The data of the cluster and multiple level-1 backups (snapshots) can be stored in the same physical data block that is billed only once. For more information, see FAQ. For example, if the total physical storage of level-1 backups is 700 GB and the database storage usage is 1,000 GB, the storage fee per hour is USD 0.0928. The fee is calculated based on the following formula: [700 GB - (1,000 GB × 50%)] × USD 0.000464/GB/hour = USD 0.0928/hour.

Backup type	Free quota	Billing method
Level-2 backup	No	Storage fee per hour = Total physical storage of level-2 backups × Unit price per hour For example, if the total size of level-2 backups is 1,000 GB, the storage fee per hour is USD 0.0325. The fee is calculated based on the following formula: 1,000 GB × USD 0.0000325/GB/hour = USD 0.0325/hour.
Log backup	100 GB	Storage fee per hour = (Total physical storage of log backups - 100 GB) × Unit price per hour For example, if the total size of log backups is 1,000 GB, the hourly fee is USD 0.02925. The fee is calculated based on the following formula: (1,000 GB - 100 GB) × USD 0.0000325/GB/hour = USD 0.02925/hour.

10.3. Backup methods

10.3.1. Configure a backup policy

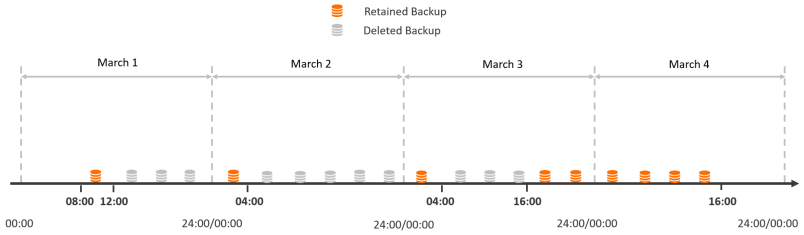
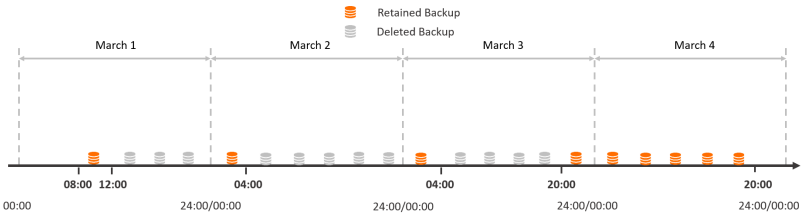
supports data backup and redo log backup. Backing up data is a process of creating a backup set (snapshot) of all data on a cluster at a certain point in time. Backing up redo logs is a process of recording the new data after a backup set is created. You can configure policies for data backup and redo log backup. For example, you can specify the frequency of automatic data backups, the retention period of data backup files, the storage location, and the retention period of log backup files.

Procedure

1.
2.
3.
4. In the left-side navigation pane, choose **Settings and Management > Backup and Restore**.
5. On the **Backup Policy Settings** tab, click **Edit**.
6. In the **Backup Policy Settings** dialog box, configure the parameters in the **Data Backup**, **Log Backup**, and **General** sections.
   - **Parameters in the Data Backup section**

     To configure a data backup policy, specify the frequency of automatic backups and the storage location and retention period of the backup files generated by automatic backups and manual backups.

Parameter	Description

Parameter	Description
Backup Frequency	The frequency of automatic backups. You can select **Standard Backup (at specified intervals)** or **High-frequency Backup**. ■ **Standard Backup (at specified intervals)**: By default, automatic backup is performed once a day. You can set the cycle and start time for automatic data backup. **Note** ■ To prevent data loss, automatic backup must be performed at least twice a week. ■ Automatic backup files cannot be deleted. ■ **High-frequency Backup**: supports enhanced protection in last 24 hours. This feature increases backup frequency to speed up data restoration. You can specify the backup frequency which can be **Last 24 Hours, Every 2 Hours**, **Last 24 Hours, Every 3 Hours**, or **Last 24 Hours, Every 4 Hours**. After you enable enhanced backup, all backups are retained for 24 hours. Backups are automatically deleted when the retention period expires. However, the system permanently retains the first backup that is created after 00:00 every day. For example, if you specify a backup frequency of **every 4 hours** at 08:00 on March 1, the system automatically creates the first backup within four hours from 08:00 to 12:00 on March 1. Then, the system continues to create a backup at an interval of four hours. If the current time is 16:00 on March 4, the system retains the following backups: ■ The backups created within the last 24 hours (from 16:00 on March 3 to 16:00 on March 4). ■ The backups created between 00:00 and 4:00 on March 3. ■ The backups created between 00:00 and 4:00 on March 2. ■ The backups created between 08:00 and 12:00 on March 1.  Then, after four hours or at 20:00 on March 4, the system retains the following backups: ■ The backups created within the last 24 hours from 20:00 on March 3 to 20:00 on March 4. ■ The backups created between 00:00 and 4:00 on March 3. ■ The backups created between 00:00 and 4:00 on March 2. ■ The backups created between 08:00 and 12:00 on March 1. 

Parameter	Description
Data Backup Retention Period	The storage location and retention period of the data backup files generated by automatic backups and manual backups. You can specify **Level-1 Backup** or **Level-2 Backup** as the storage location. For more information, see Data backup. ■ **Level-1 Backup**: Set the retention period for level-1 backups. Note ■ By default, level-1 backup is enabled. The default retention period of level-1 backups is 7 days. ■ A backup can be retained for 3 to 14 days. ■ **Level-2 Backup**: enable or disable the level-2 backup feature. Note ■ By default, the level-2 backup feature is disabled. If you enable the feature, storage fees are incurred. You can delete backup files to reduce costs. For more information about the pricing of level-2 backup, see Billing rules of backup storage that exceeds the free quota. ■ Level-2 backups can be retained for 30 to 7,300 days. ■ If you want to permanently retain level-2 backups, select **Retained Before Cluster Is Deleted**. After you select this option, you cannot specify the retention period of level-2 backups.

- **Parameters in the Log Backup section**

  When you configure a redo log backup policy, you must specify the retention period of redo logs.

Parameter	Description
Log Retention Period (Days)	Specifies the retention period for log backup. Note ■ By default, log backup is enabled, and backup files are retained for seven days. You cannot disable log backup. ■ Log backup files can be retained for 3 to 7,300 days. ■ To retain log backups permanently, select **Retained Before Cluster Is Deleted**. The retention period parameter becomes unavailable after you select this option.

- **Parameters in the General section**

  You can configure a backup retention policy that applies when you delete a cluster.

Parameter	Description

Parameter	Description
When Cluster Is Deleted	The backup retention policy that applies when you delete a cluster. ■ **Permanently Retain All Backups**: retains all backups after you delete a cluster. ■ **Permanently Retain Last Automatic Backup**: retains the most recent backup after you delete a cluster. ■ **Immediately Delete All Backups**: does not retain backups after you delete a cluster. **Note** ■ If you select the **Permanently Retain All Backups** or **Permanently Retain Last Automatic Backup** policy, the system runs an automatic backup task to retain all data when you delete the cluster. ■ After you delete a cluster, level-1 backups are automatically transferred to level-2 backups. You can go to the **Cluster Recycle** page to view all backups. For more information, see Restore a released cluster.

Related API operations

Operation	Description
DescribeBackupPolicy	Queries the backup policy of a specified cluster.
ModifyBackupPolicy	Modifies the backup policy of a specified cluster.

10.3.2. Backup method 1: Automatic backup

By default, automatic backup is enabled and performs automatic backup once a day after a new cluster is created. You can configure parameters such as the frequency of automatic backup and the retention period of backup files in the console based on your business requirements.

Procedure

1.
2.
3.
4. In the left-side navigation pane, choose **Settings and Management > Backup and Restore**.
5. Click **Backup Policy Settings**.
6. On the **Backup Policy Settings** page, click **Edit**. In the dialog box that appears, configure the following parameters.

Parameter	Description
Backup Frequency	You can select **Standard Backup** or **High-frequency Backup**. ○ **Standard Backup**: You can set the cycle and start time for automatic data backup. **Note** To prevent data loss, automatic backup must be performed at least twice a week. ○ **Enhanced Backup**: Set the backup frequency. You can select **Last 24 Hours, Every 2 Hours**, **Last 24 Hours, Every 3 Hours**, or **Last 24 Hours, Every 4 Hours**.

Parameter	Description
Data Backup Retention Period	Specify the retention period for **level-1 backups** and **level-2 backups**. ○ **Level-1 Backup**: Set the retention period for level-1 backups. Note ■ By default, level-1 backup is enabled. The default retention period of level-1 backups is 7 days. ■ Level-1 backups are retained for 3 to 14 days. ○ **Level-2 Backup**: Enable or disable the level-2 backup feature. Note ■ By default, the level-2 backup feature is disabled. If you enable the feature, storage fees are incurred. You can delete backup files to reduce costs. For more information about the pricing of level-2 backup, see Billing rules of backup storage that exceeds the free quota. ■ Level-2 backups can be retained for 30 to 7,300 days. ■ If you want to permanently retain level-2 backups, select **Retained Before Cluster Is Deleted**. After you select this option, you cannot specify the retention period of level-2 backups.

7. Click **OK**.

Related API operations

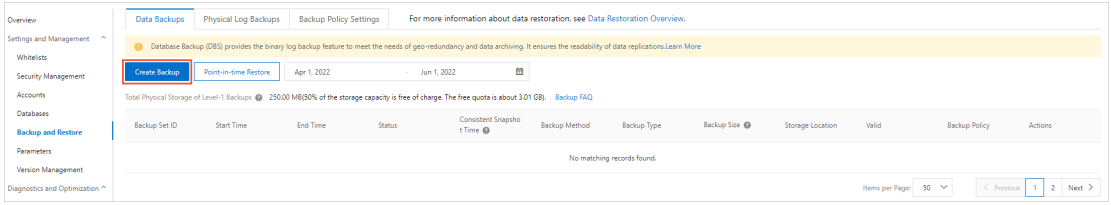
Operation	Description
CreateBackup	Creates a full backup of a specified cluster.
DescribeBackups	Queries the backup information about a specified cluster.
DeleteBackup	Deletes the backups of a specified cluster.

10.3.3. Backup method 2: Manual backup

Manual backups are backups triggered by you. You can manually back up data at any time based on your business requirements to ensure data reliability. This topic describes how to configure the manual backup settings.

Procedure

1.
2.
3.
4. In the left-side navigation pane, choose **Settings and Management > Backup and Restore**.
5. On the **Backups** tab, click **Create Backup**.



6. In the **Create Backup** message, click **OK**.

> Note
> - You can create up to three backups for a cluster.
> - Manual backup files can be deleted. However, a backup file cannot be restored after you delete it. Proceed with caution.

Related API operations

Operation	Description
CreateBackup	Creates a full backup of a specified cluster.
DescribeBackups	Queries the backup information about a specified cluster.
DeleteBackup	Deletes the backups of a specified cluster.

10.4. Restoration methods

10.4.1. Restoration method 1: Restore data to a specific point in time

provides two methods for you to restore historical data to a new cluster: restore data to a specific point in time and restore data from a backup set (snapshot). This topic describes how to restore data to a specific point.

Precautions

Only the data and account information of the original cluster can be restored to a new cluster. The parameters of the original cluster cannot be restored to the new cluster.

Procedure

1.
2.
3.
4. In the left-side navigation pane, choose **Settings and Management > Backup and Restore**.
5. On the **Backup and Restore** page, click **Point-in-time Restore**.
6. On the **Clone Instance** page, select a **billing method** for the new cluster.
   - **Subscription**: When you create a cluster, you must pay for compute nodes. You are charged for the use of storage resources and the costs are deducted from your account balance on an hourly basis.
   - **Pay-As-You-Go**: If you select the pay-as-you-go billing method, you pay for the resources after you use them. You are charged for the compute nodes and the used storage space on an hourly basis. The fee is deducted from your account balance on an hourly basis. We recommend that you select the **Pay-As-You-Go** billing method for the short-term use. You can reduce costs by releasing the cluster based on your business requirements.
7. Configure the parameters that are listed in the following table.

Parameter	Description
Action Mode	Select **Restore to Point in Time**.
Backup Point in Time	The point in time to which you want to restore data. **Note** You can restore your cluster to a particular time only over the past 7 days.

Parameter	Description
Region	This parameter is automatically set to the region of the original cluster. You do not need to change this value.
Primary Zone	Select the **primary zone** where the cluster resides. **Note** In regions that have two or more zones, PolarDB automatically replicates the data to the secondary zone for disaster recovery.
Network Type	The default value is **VPC**.
VPC	Select a **VPC** and a **vSwitch** for the new cluster. We recommend that you select the same VPC and vSwitch that are connected to the original cluster. **Note** Make sure that the cluster is created in the same VPC as the ECS instance to which you want to connect. Otherwise, the cluster and the ECS instance cannot communicate over the internal network to achieve optimal performance.
vSwitch	
Compatibility	The default value is **Compatible with Oracle Syntax**.
Edition	This parameter is automatically set to the **edition** value of the original cluster. You do not need to change this value.
Resource Type	This parameter is automatically set to the **Resource Type** value of the original cluster. You do not need to change this value.
Node Specification	Select a **node specification**. The maximum storage capacity and the performance of clusters vary based on the node specifications. For more information, see Node specifications. **Note** We recommend that you select a **node specification** that is higher than the node specification of the original cluster. This ensures that the new cluster runs as expected.
Nodes	The default value is **2**. **Note** By default, a new cluster has one primary node and one read-only node. After the cluster is created, you can add nodes to the cluster. A cluster can contain one primary node and a maximum of 15 read-only nodes. For more information about how to add nodes, see Add or remove a read-only node.
Storage Cost	You do not need to select the storage capacity when you purchase PolarDB clusters. You are charged for the used storage space on an hourly basis. You can also purchase a storage package to offset storage fees. For more information about how to purchase a storage package, see Purchase a storage plan.

Parameter	Description
Cluster Name	The **name** of the new PolarDB cluster must meet the following requirements: ○ The name must be 2 to 128 characters in length, and can contain letters, digits, periods (.), underscores (_) and hyphens (-). The name must start with a letter. ○ The name must start with a letter. ○ The name can contain digits, periods (.), underscores (_), and hyphens (-). If you leave this field empty, the system automatically generates a cluster name. You can change the cluster name after the cluster is created.
Subscription Duration	Select a **Subscription Duration** value for the PolarDB cluster. **Note** This parameter is valid only when the **Billing Method** parameter is set to **Subscription**.
Quantity	Set the **Quantity** value of the PolarDB cluster.

8. Read and select the terms of service, and then complete the payment based on the selected **billing method**.
   - **Pay-as-you-go**

     Click **Buy Now**.
   - **Subscription**
     a. Click **Buy Now**.
     b. On the **Purchase** page, confirm the order and the payment method, and click **Purchase**.

Note After you complete the payment, it requires 10 to 15 minutes to create the cluster. Then, you can view the new cluster on the **Clusters** page.

Related API operations

Operation	Description
CreateDBCluster	Restores data of a PolarDB cluster. **Note** You must set **CreationOption** to **CloneFromPolarDB**.

10.4.2. Restoration method 2: Restore data from a backup set (snapshot)

provides two methods for you to restore historical data to a new cluster: restore data to a specific point in time and restore data from a backup set (snapshot). This topic describes how to restore data from a backup set (snapshot).

Precautions

Only the data and account information of the original cluster can be restored to a new cluster. The parameters of the original cluster cannot be restored to the new cluster.

Procedure

1.

2.
3.
4. In the left-side navigation pane, choose **Settings and Management > Backup and Restore**.
5. Find the backup set (snapshot) and click **Restore to New Cluster**.
6. On the **Clone Instance** page, select a **billing method** for the new cluster.
   - **Subscription**: When you create a cluster, you must pay for compute nodes. You are charged for the use of storage resources and the costs are deducted from your account balance on an hourly basis.
   - **Pay-As-You-Go**: If you select the pay-as-you-go billing method, you pay for the resources after you use them. You are charged for the compute nodes and the used storage space on an hourly basis. The fee is deducted from your account balance on an hourly basis. We recommend that you select the **Pay-As-You-Go** billing method for the short-term use. You can reduce costs by releasing the cluster based on your business requirements.
7. Configure the parameters that are listed in the following table.

Parameter	Description
Action Mode	Select **Restore from Backup Set**.
Backup Set	The **backup set** from which you want to restore data. **Note** The **Start Time** of each backup set is displayed. You can determine whether to select the backup set based on this backup time.
Region	The **region** where the cluster resides. **Note** The default region is the same as the region of the original cluster. Use the default region.
Primary Zone	Select the **primary zone** where the cluster resides. **Note** In regions that have two or more zones, PolarDB automatically replicates the data to the secondary zone for disaster recovery.
Network Type	The default value is **VPC**.
VPC	Select a **VPC** and a **vSwitch** for the new cluster. We recommend that you select the same VPC and vSwitch that are connected to the original cluster.
vSwitch	
Compatibility	The default value is **Compatible with Oracle Syntax**.
Edition	This parameter is automatically set to the **edition** value of the original cluster. You do not need to change this value.
Resource Type	This parameter is automatically set to the **Resource Type** value of the original cluster. You do not need to change this value.
Node Specification	Select a **node specification**. The maximum storage capacity and the performance of clusters vary based on the node specifications. For more information, see Node specifications. **Note** We recommend that you select a **node specification** that is higher than the node specification of the original cluster. This ensures that the new cluster runs as expected.

Parameter	Description
Nodes	The default value is **2**. **Note** By default, a new cluster has one primary node and one read-only node. After the cluster is created, you can add nodes to the cluster. A cluster can contain one primary node and a maximum of 15 read-only nodes. For more information about how to add nodes, see Add or remove a read-only node.
Storage Cost	You do not need to select the storage capacity when you purchase PolarDB clusters. You are charged for the used storage space on an hourly basis. You can also purchase a storage package to offset storage fees. For more information about how to purchase a storage package, see Purchase a storage plan.
Cluster Name	The **name** of the new PolarDB cluster must meet the following requirements: ○ The name must be 2 to 128 characters in length, and can contain letters, digits, periods (.), underscores (_) and hyphens (-). The name must start with a letter. ○ The name must start with a letter. ○ The name can contain digits, periods (.), underscores (_), and hyphens (-). If you leave this field empty, the system automatically generates a cluster name. You can change the cluster name after the cluster is created.
Subscription Duration	Select a **Subscription Duration** value for the PolarDB cluster. **Note** 该参数只会在**付费模式**为**包年包月**时设置。
Quantity	Set the **number** of PolarDB clusters you want to purchase.

8. Read and select the terms of service, and then complete the payment based on the selected **billing method**.
   - **Pay-as-you-go**

     Click **Buy Now**.
   - **Subscription**
     a. Click **Buy Now**.
     b. On the **Purchase** page, confirm the order and the payment method, and click **Purchase**.

Note After you complete the payment, it requires 10 to 15 minutes to create the cluster. Then, you can view the new cluster on the **Clusters** page.

Related API operations

Operation	Description
CreateDBCluster	Restores data of a PolarDB cluster. **Note** You must set **CreationOption** to **CloneFromPolarDB**.

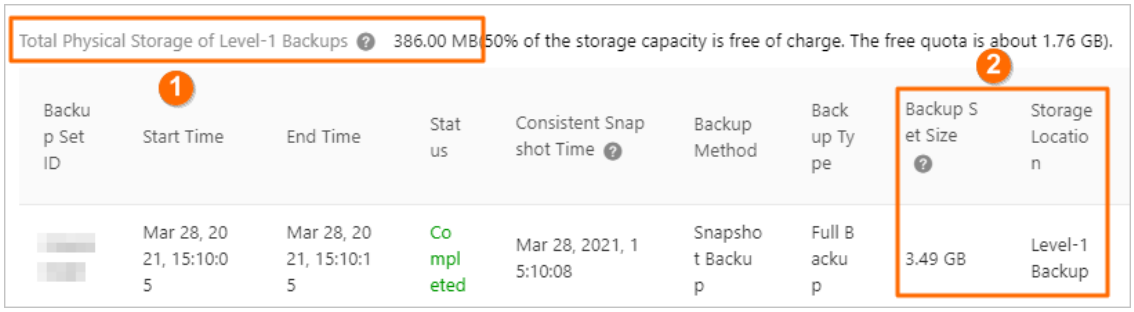
10.5. FAQ

This topic provides answers to frequently asked questions about the backup and restoration features of .

Data backup FAQ

- Is the total size of level-1 backups (snapshots) equal to the sum of the sizes of all level-1 backups (snapshots)?

  No, the total size of level-1 backups (snapshots) is not equal to the sum of the sizes of all level-1 backups (snapshots). The total size of level-1 backups (snapshots) is displayed in part ①, as shown in the following figure.



- Why is the total size of level-1 backups smaller than the sum of the sizes of all level-1 backups?

  The size of level-1 backups in is measured in two forms: the logical size of backups and the total physical storage of backups. uses snapshot chains to store level-1 backups. Only one record is generated for each data block. Therefore, the total physical storage of all level-1 backups is smaller than the total logical size of all level-1 backups. In some cases, the total physical storage of all level-1 backups is smaller than the logical size of a single backup.

- How am I charged for backups in ?

  You are charged for storage space of level-1, level-2, and log backups. By default, the level-1 backup and log backup features are enabled, and a free storage quota is provided. By default, the level-2 backup feature is disabled.

- How are the fees of level-1 backups calculated?

  The fee is calculated based on the following formula: Storage fee per hour = (Total size of level-1 backups - Used database storage space × 50%) × Price per hour. For example, in a region within Chinese mainland, the total size of level-1 backups of a database is 700 GB, and the used database storage space is 1,000 GB. Then, the storage fee per hour is calculated based on the following formula: [700 GB - 500 GB] × USD 0.000464/GB = USD 0.0928.

- Can I use a storage plan to offset the storage fees of backups?

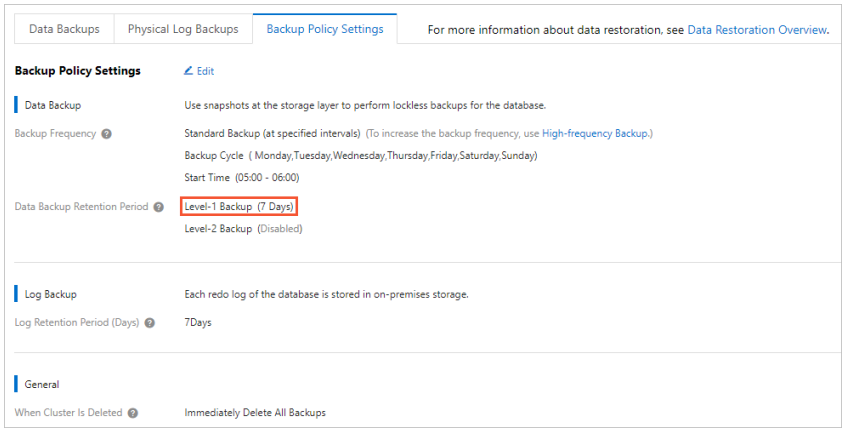
  Yes, you can purchase a storage plan to offset the storage space used by all clusters within your account. The remaining capacity of the storage plan is automatically used to offset the storage space that exceeds the free quota for level-1 backups at a ratio of 1:1.6 until the storage plan is exhausted. If the remaining capacity of the storage plan is insufficient to offset the storage space of level-1 backups, you are charged for additional storage space on a pay-as-you-go basis. For more information, see .Storage plans

- Are level-1 backups the only type of backup that can be manually created?

  A: Yes.

- How long are manually created backups retained?

The retention period of manually created backups is specified by the Level-1 Backup parameter in the Data Backup Retention Period section.



- How do I view the size of a level-2 backup?

  You can view the size of a level-2 backup on the **Backups** tab in the console.

Data restoration FAQ

- How can I restore data that was deleted or modified by accident?

  You can choose different methods to restore data based on your business scenario and database engine version. For more information, see Restoration method 1: Restore data to a specific point in time and Restoration method 2: Restore data from a backup set (snapshot).

- Can I customize the names of restored tables?

  Yes.

- If my cluster does not have a data backup, can I restore the data to a previous point in time?

  No. To restore data to a previous point in time, you must restore the data of a full backup that was created before the specified point in time. Then, you must restore the data generated after the backup that was created and before the specified point in time based on the physical logs.

11.Cluster Recycle

11.1. Pricing

Cluster Recycle stores released clusters. You can restore a cluster in Cluster Recycle to a new cluster, or delete a backup set of the cluster. This topic describes the pricing rules of Cluster Recycle of clusters.

Level-1 backups are provided free of charge. Level-2 backups are paid services.

Region	Fee (USD/GB/hour)
Regions in Chinese mainland	0.0000325
Regions outside Chinese mainland	0.0000455

11.2. Restore a released cluster

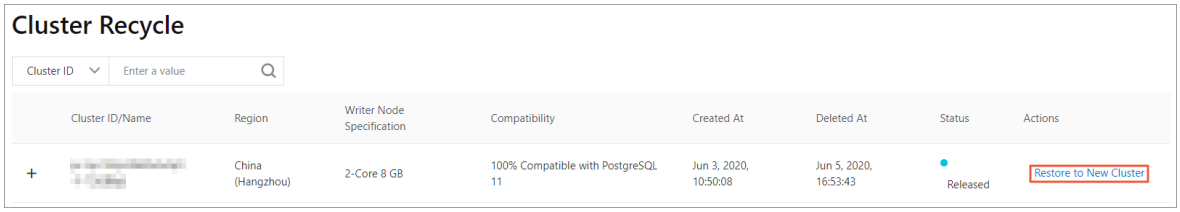
This topic uses a cluster as an example to describe how to restore clusters in Cluster Recycle.

Usage notes

- Released clusters in Cluster Recycle must have at least one backup set. If all backup sets of a cluster have been deleted, the cluster cannot be restored.
- After a cluster is released, the data of all released clusters in Cluster Recycle is archived asynchronously to level-2 backups at a rate of approximately 150 MB/s. For more information about backups, see Overview.

Procedure

1.
2.
3. In the left-side navigation pane, click **Cluster Recycle**.
4. Find the cluster that you want to restore, and click **Restore to New Cluster** in the **Actions** column.



5. Set **Product Type** to **Subscription** or **Pay-As-You-Go**.
   - **Subscription**: When you create a cluster, you must pay for compute nodes. You are charged for the use of storage resources and the costs are deducted from your account balance on an hourly basis.
   - **Pay-As-You-Go**: An upfront payment is not required. You are charged for compute nodes and the amount of storage that is consumed by your data. These costs are deducted from your account balance on an hourly basis.
6. Configure the parameters that are listed in the following table.

Parameter	Description
Region	The region where the cluster is deployed. The region cannot be changed after the cluster is created. **Note** Make sure that the cluster is created in the same region as the Elastic Compute Service (ECS) instance to which you want to connect. Otherwise, the cluster and the ECS instance can communicate only over the Internet. As a result, the performance of the cluster may be compromised.

Parameter	Description
Creation Method	Select **Restore from Recycle**. This value indicates that the deleted database is restored from Cluster Recycle.
Source Version	Select the version of the released cluster.
Deleted Clusters	Select the name of the deleted cluster.
Backup History	Select the backup set to restore. **Note** The timestamps of the backups in the **Backup History** drop-down list are displayed in UTC. The timestamps of backups in the Backups list are displayed in the system time format. Make sure that you choose the correct historical backup. For example, the timestamp of a backup set in the `Backup History` is **2020-05-08T02:00:00Z**. The corresponding timestamp in the backup list is `10:00:00 on May 8, 2020 (UTC+08:00)`.
Primary Zone	The primary zone where the cluster is deployed. ○ Each zone is an independent geographical location in a region. All of the zones in a region provide the same level of service performance. ○ You can choose to create your cluster in the same zone as an ECS instance or in a different zone from the zone of the instance. ○ You must specify only the primary zone. The system automatically selects a secondary zone.
Network Type	This parameter can be set only to **VPC**. You do not need to specify this parameter.
VPC **VSwitch**	Make sure that the cluster is created in the same VPC as the ECS instance to which you want to connect. Otherwise, the cluster and the ECS instance cannot communicate over the internal network to achieve optimal performance. ○ If you have an existing VPC that meets your network requirements, select the VPC. For example, if you have an existing ECS instance and the VPC to which the ECS instance belongs meets your network requirements, select this VPC. ○ Otherwise, use the default VPC and the default vSwitch. ■ Default VPC: ■ Only one VPC is specified as the default VPC in the region that you select. ■ The default VPC uses a 16-bit subnet mask. For example, the CIDR block 172.31.0.0/16 provides up to 65,536 private IP addresses. ■ The default VPC does not count towards the quota of VPCs that you can create on Alibaba Cloud. ■ Default vSwitch: ■ Only one vSwitch is specified as the default vSwitch in the zone that you select. ■ The default VPC uses a 20-bit subnet mask. For example, the CIDR block 172.16.0.0/20 provides up to 4,096 private IP addresses. ■ The default vSwitch does not count towards the quota of vSwitches that you can create in a VPC. ○ If the default VPC and vSwitch cannot meet your requirements, you can create your own VPC and vSwitch. For more information, see Create and manage a VPC.
Compatibility	The database engine version of the cluster. The default version is the same as the version of the deleted cluster and cannot be changed.

Parameter	Description
Edition	The network type of the new cluster. This parameter is automatically set to . You do not need to specify this parameter.
Node Specification	Select a specification based on your requirements. We recommend that you select a specification that is the same or higher than the node specification of the released cluster. For more information about the compute node specifications of , see Specifications and pricing.
Nodes	You do not need to specify this parameter. By default, the system creates two nodes that have the same specification: a primary node and a read-only node. **Note** If the primary node fails, the system uses the read-only node as the primary node and creates another read-only node. For more information about read-only nodes, see Architecture.
Storage Cost	You do not need to specify this parameter. The system charges you on an hourly basis based on the amount of storage that is consumed by your data. For more information, see Specifications and pricing. **Note** You do not need to specify the storage capacity when you create a cluster. The system automatically scales storage resources based on data volume.
Enable TDE	Specify whether to enable Transparent Data Encryption (TDE). After TDE is enabled, encrypts the data files of your cluster. You do not need to modify the code to allow access to your cluster. However, TDE reduces the performance of your cluster by 5% to 10%. **Note** TDE cannot be disabled after it is enabled.
Cluster Name	Enter the name of the cluster. The name must meet the following requirements: ◦ The name cannot start with `http://` or `https://`. ◦ It must be 2 to 256 characters in length. If this parameter is left empty, the system automatically generates a cluster name. You can change the cluster name after the cluster is created.
Resource Group	Select a resource group from available resource groups. For more information, see Create a resource group. **Note** A resource group is a group of resources that belong to an Alibaba Cloud account. Resource groups allow you to manage these resources in a centralized manner. A resource belongs to only one resource group. For more information, see Use RAM to create and authorize resource groups.
Purchase Plan	Select a purchase plan for the new cluster. **Note** This parameter is available only when the **Product Type** parameter is set to **Subscription**.
Number	Select the number of clusters you want to purchase.

7. Complete the rest of the steps based on the product type of the cluster.
   - **Pay-As-You-Go**
     a. Click **Buy Now**.

b. On the **Confirm Order** page, confirm your order information. Read and accept the terms of service, and then click **Buy Now**.

- **Subscription**

a. Click **Buy Now**.

b. On the **Confirm Order** page, confirm your order information. Read and accept the terms of service, and then click **Buy Now**.

c. On the **Purchase** page, confirm the order and the payment method, and click **Purchase**.

After you complete the payment, it requires 10 to 15 minutes to create the cluster. Then, the newly created cluster is displayed on the **Clusters** page.

> **Note** The amount of time required to restore data to a new cluster depends on the size of the backup set. It takes more time for the system to restore data from a larger backup set. After the cluster is created, you can return to the PolarDB console and view the new cluster on the **Clusters** page.

Related API operations

API	Description
CreateDBCluster	Creates a cluster.

11.3. Delete a released cluster

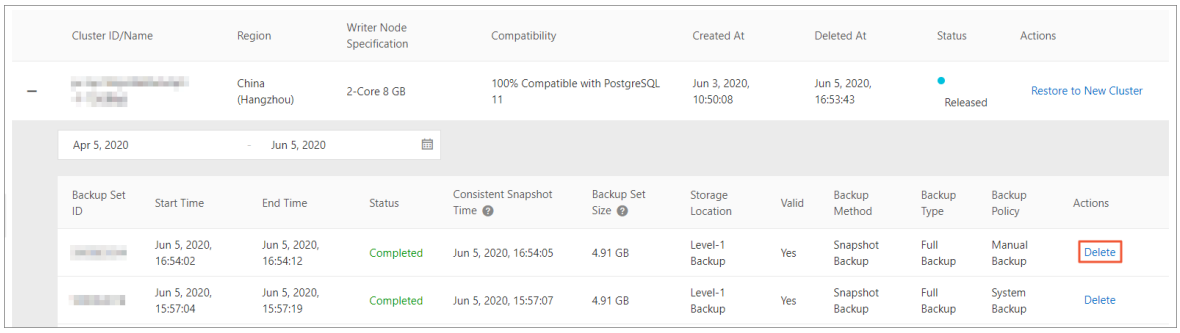
This topic uses a cluster as an example to describe how to delete the backup sets of released clusters.

Usage notes

- Released clusters in Cluster Recycle must have at least one backup set. If all backup sets of a cluster have been deleted, the cluster cannot be restored.
- After a cluster is released, the data of all released clusters in Cluster Recycle is archived asynchronously to level-2 backups at a rate of approximately 150 MB/s. For more information about backups, see Overview.

Procedure

1.
2.
3. In the left-side navigation pane, click **Cluster Recycle**.
4. Find the cluster that you want to manage, and click the + icon next to the cluster to show a list of backup sets.
5. Find the backup set that you want to delete, and click **Delete** in the **Actions** column.



6. In the message that appears, click **OK**.

> **Warning** If you delete all backup sets of a cluster in **Cluster Recycle**, the cluster cannot be restored. Proceed with caution.

12.Data Security and Encryption

12.1. Configure SSL encryption

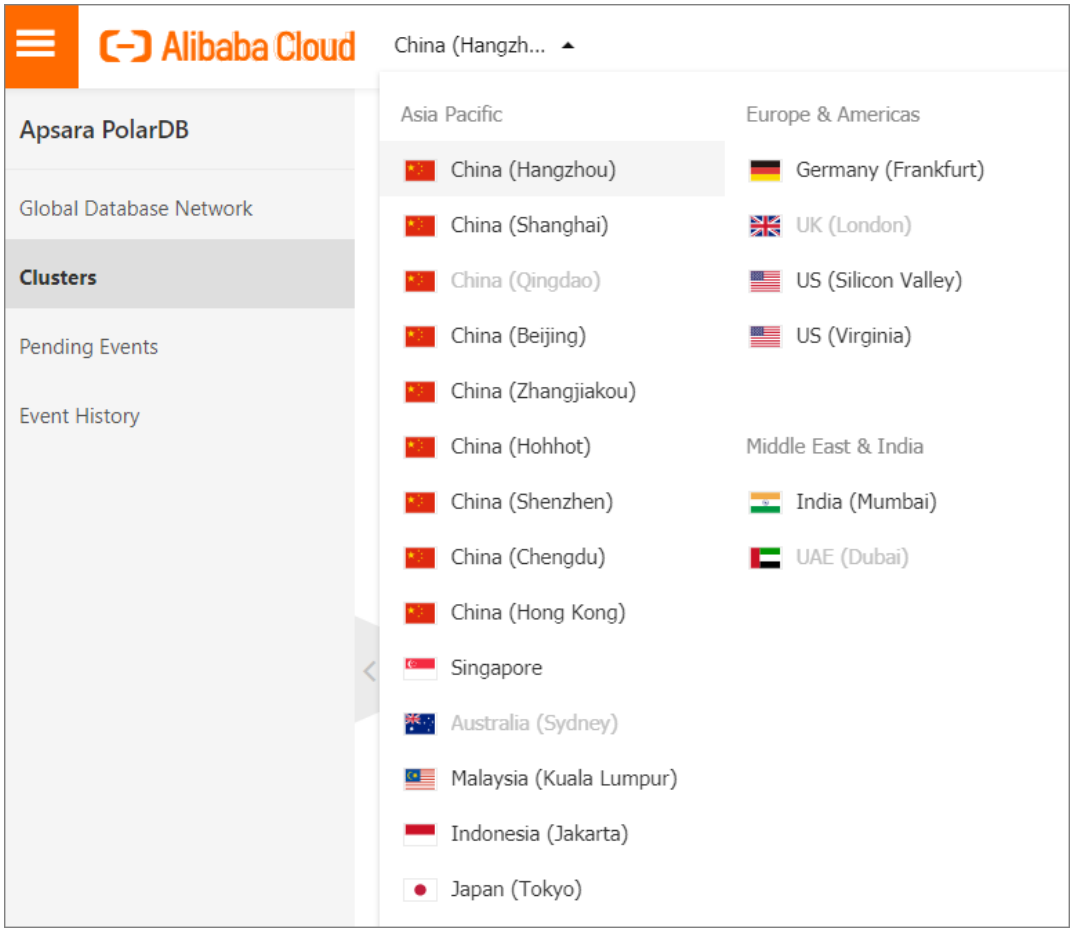
This topic describes how to make data transmission more security by configuring SSL encryption. You must enable SSL encryption and install SSL certificates that are issued by certificate authorities (CAs) in the required applications. SSL is used to encrypt connections at the transport layer and enhance the security and integrity of the transmitted data. However, SSL encryption increases the round-trip time.

Precautions

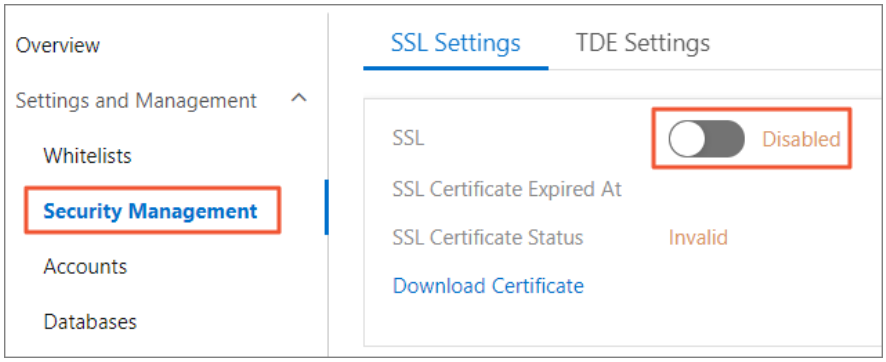
- An SSL certificate is valid for one year. You must Update the validity period of the SSL certificate and then download and configure the certificate again. Otherwise, clients that use encrypted network connections cannot connect to your clusters.
- SSL encryption may cause a sharp increase in CPU utilization. We recommend that you enable SSL encryption only if you want to encrypt the connections that are established to the public endpoint of your cluster. In most cases, connections that are established to the internal endpoint of your cluster are secure and do not require SSL encryption.
- After you disable SSL encryption for a cluster, the cluster is restarted. Proceed with caution.

Enable SSL encryption and download an SSL certificate

1. Log on to the Apsara PolarDB console.
2. In the upper-left corner of the page, select the region where the cluster is deployed.

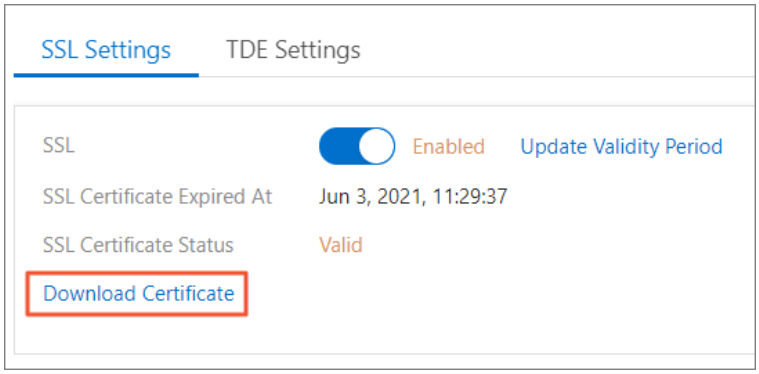


3. Find the cluster and click the cluster ID.
4. In the left-side navigation pane, choose **Settings and Management > Security Management**.
5. On the **SSL Settings** tab, turn on the switch next to **SSL** to enable SSL encryption.



Note You can enable SSL encryption for only the primary endpoints of clusters.

6. In the **Configure SSL** dialog box, click **OK**.
7. After the SSL status changes to **Enabled**, click **Download Certificate**.



The downloaded package contains the following files:

- P7B file: the SSL certificate file that is used for a Windows operating system
- PEM file: used to import CA certificates to other operating systems or applications.
- JKS file: the Java truststore file. The password is apsaradb. It is used to import the CA certificate chain to Java programs.

Note When the JKS file is used in Java, you must modify the default JDK security configuration in JDK 7 and JDK 8. Open the `jre/lib/security/java.security` file on the server that is connected to Apsara PolarDB and modify the following configurations:

```
jdk.tls.disabledAlgorithms=SSLv3, RC4, DH keySize < 224
jdk.certpath.disabledAlgorithms=MD2, RSA keySize < 1024
```

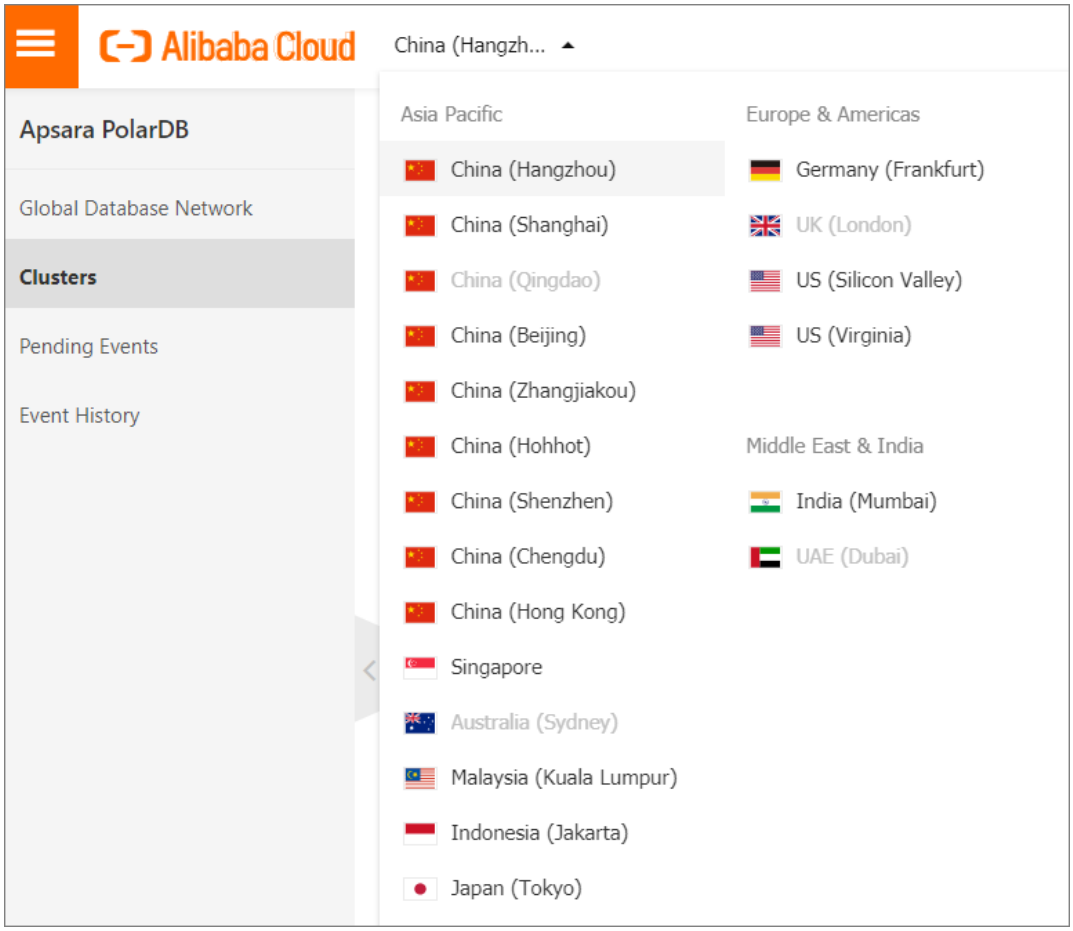
If you do not modify these configurations, the following error is returned. In most cases, similar errors are caused by invalid Java security configurations.

```
javax.net.ssl.SSLHandshakeException: DHPublicKey does not comply to algorithm constraints
```

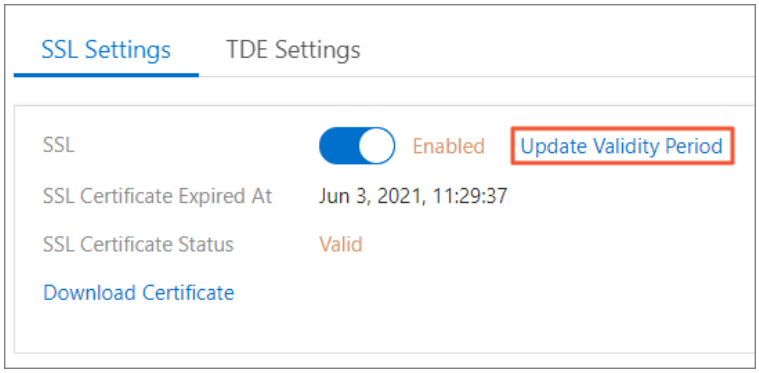
Update the validity period of the SSL certificate

After you change the endpoint that has SSL encryption enabled or when the SSL certificate is about to expire, you must update the validity period of the SSL certificate. This section describes how to update the validity period of an SSL certificate.

1. Log on to the Apsara PolarDB console.
2. In the upper-left corner of the page, select the region where the cluster is deployed.



3. Find the cluster and click the cluster ID.
4. In the left-side navigation pane, choose **Settings and Management > Security Management**.
5. On the **SSL Settings** tab, click **Update Validity Period**.



6. In the **Configure SSL** dialog box, click **OK**.

> **Note** After you update the validity period of the certificate, the cluster is restarted. Proceed with caution.

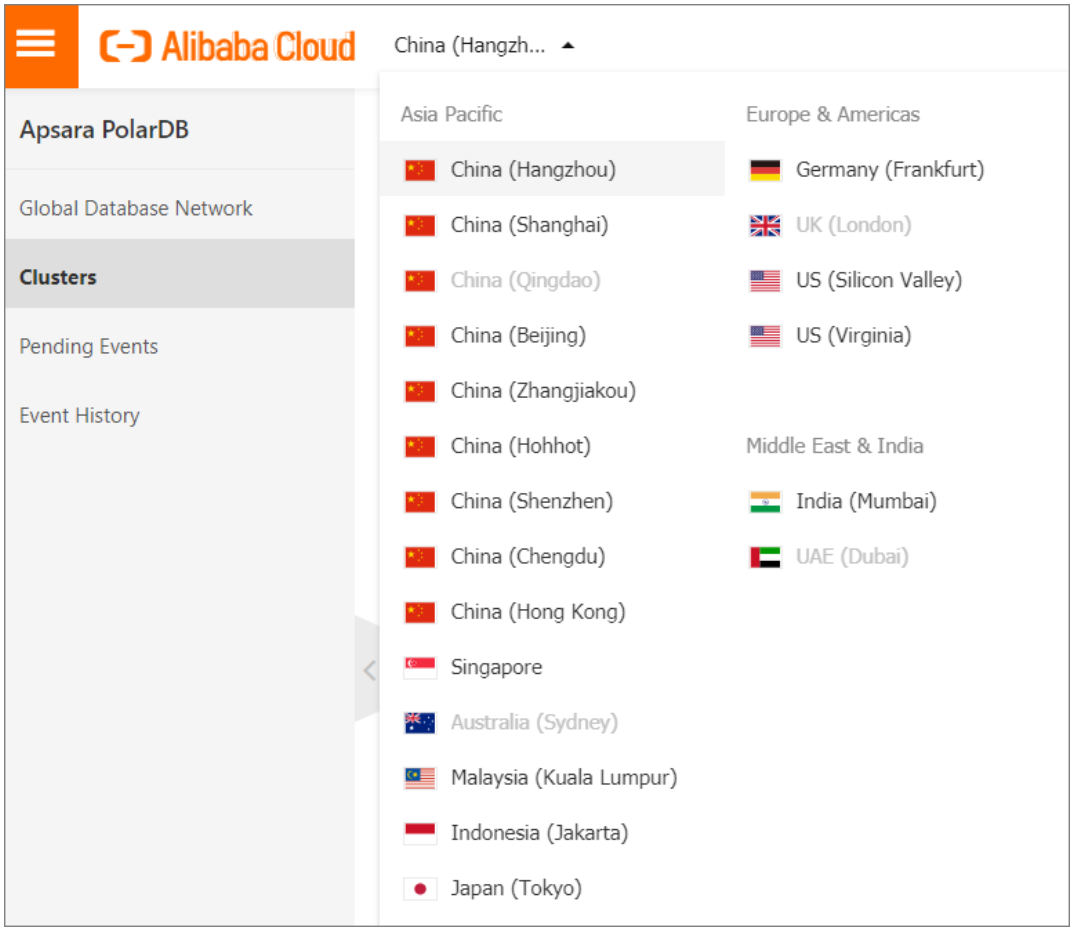
7. After the SSL certificate is renewed, download and configure the SSL certificate again.

> **Note** For more information about how to download a certificate, see Step 7 in the "Enable SSL encryption and download an SSL certificate" section.

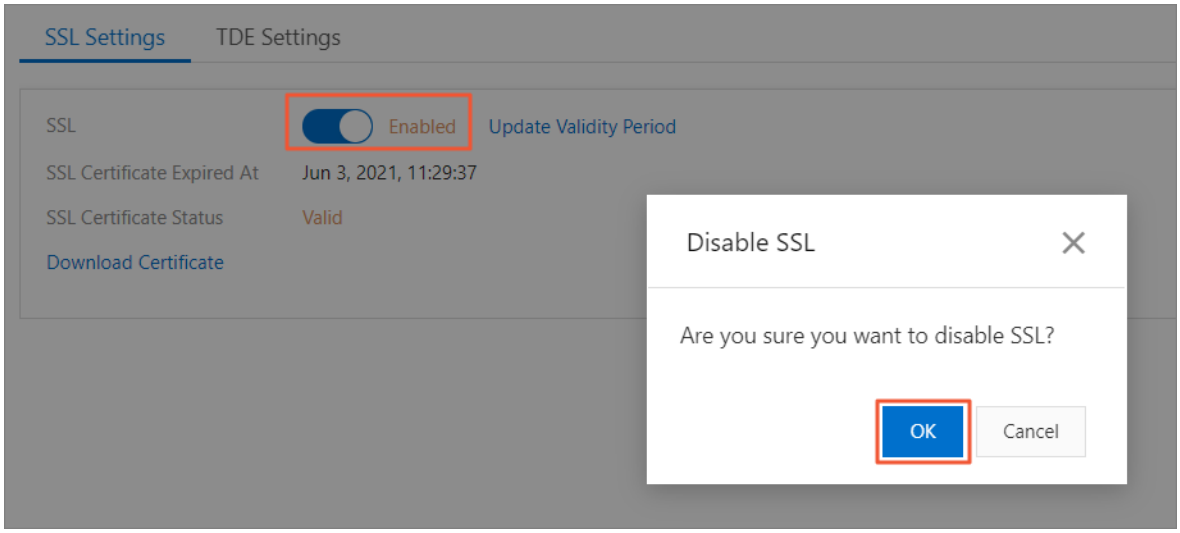
Disable SSL encryption

> **Note**
> - After you disable SSL encryption, the cluster is restarted. We recommend that you perform this operation during off-peak hours.
> - After SSL encryption is disabled, the performance of your cluster is improved but data security is compromised. We recommend that you disable SSL encryption only in secure environments.

1. Log on to the Apsara PolarDB console.
2. In the upper-left corner of the page, select the region where the cluster is deployed.



3. Find the cluster and click the cluster ID.
4. In the left-side navigation pane, choose **Settings and Management > Security Management**.
5. On the **SSL Settings** tab, turn off the switch next to **SSL** to disable SSL encryption.



6. In the **Configure SSL** dialog box, click **OK**.

FAQ

What will happen if I do not renew an expired SSL certificate? Does my cluster malfunction or data security deteriorate?

If you do not renew the SSL certificate after it expires, your cluster can still run as normal and data security is not compromised. However, applications that connect to your cluster over encrypted connections are disconnected.

Related API operations

Operation	Description
DescribeDBClusterSSL	Queries the SSL encryption settings of a specified cluster.
ModifyDBClusterSSL	Enables SSL encryption, disables SSL encryption, or renews the SSL certificate for a specified cluster.

12.2. Configure TDE

provides the Transparent Data Encryption (TDE) feature. TDE performs real-time I/O encryption and decryption on data files. Data can be encrypted before it is written to a disk and decrypted when it is read into memory. TDE does not increase the size of data files. Developers can use TDE without making changes to applications.

Prerequisites

- The version of the cluster is .
- Alibaba Cloud Key Management Service (KMS) is activated. For more information, see Activate KMS.
- ApsaraDB RDS is authorized to access KMS. For more information, see Authorize an ApsaraDB RDS for MySQL instance to access KMS.

Context

TDE performs data-at-rest encryption at the database layer. This prevents potential attackers bypassing the database to read sensitive information from storage. TDE can encrypt sensitive data within tablespaces and data stored in disks and backups. TDE also automatically decrypts data to plaintext for applications and users that have passed the database authentication. OS and unauthorized users are not allowed to access the encrypted data in plaintext form.

TDE keys of PolarDB for PostgreSQL are generated and managed by KMS. PolarDB for PostgreSQL does not provide keys and certificates that are required for encryption. You can use the keys that are automatically generated by Alibaba Cloud, or use your own materials to generate data keys and then authorize PolarDB to use them.

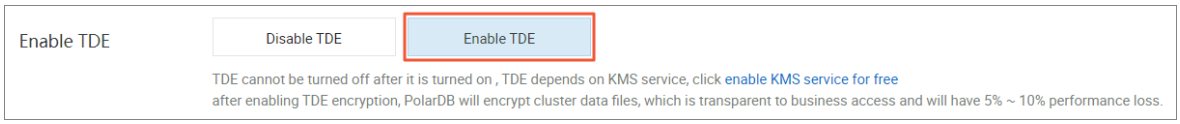
Precautions

- You cannot disable TDE after it is enabled.
- You can enable TDE only when you create a cluster.
- In I/O bound workload scenarios, TDE may affect database performance after it is enabled.
- If you use an existing custom key, pay attention to the following items:
  - If you disable the key, configure a plan to delete the key, or delete the key materials, the key becomes unavailable.
  - If you revoke the authorization to a PolarDB cluster, the cluster becomes unavailable after it is restarted.
  - You must use your Alibaba Cloud account or an account with the AliyunSTSAssumeRoleAccess permission.

Procedure

1.
2.
3. On the **Clusters** page, click **Create Cluster**.
4. On the **PolarDB buy page**, specify PolarDB purchase information and select **Enable TDE**.

   > **Note** For more information, see Create a PolarDB for Oracle cluster.

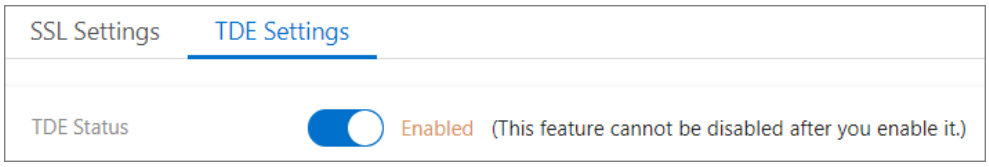
   

5. Click **Buy Now**.
6. On the **Confirm Order** page, confirm the order information, read and accept the agreement of service, and then click **Activate Now**.
7. On the **Purchase** page, confirm the order and the payment method, and click **Purchase**.

   > **Note** After you complete the payment, wait for 10 to 15 minutes. Then, you can view the newly created cluster on the Clusters page.

View the TDE status

1.
2.
3.
4. In the left-side navigation pane, choose **Settings and Management > Security Management**.
5. On the **TDE Settings** tab, view **TDE Status**.

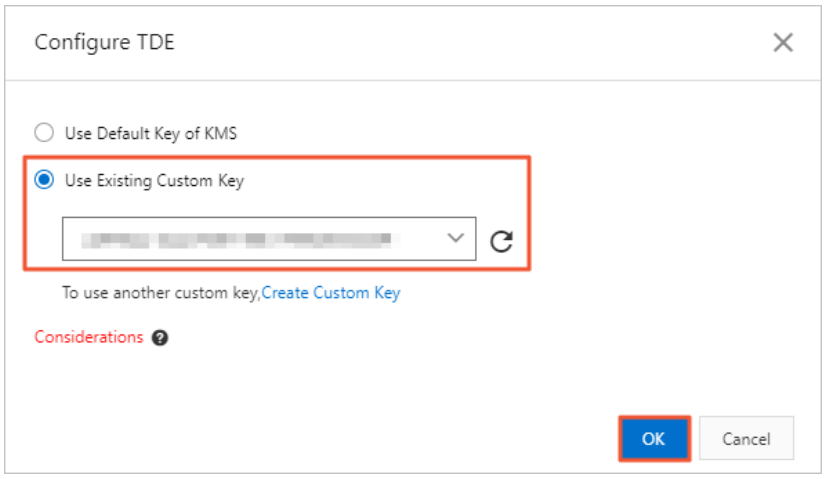
   

Switch to a custom key

1.
2.
3.
4. In the left-side navigation pane, choose **Settings and Management > Security Management**.
5. On the **TDE Settings** tab, click **Switch to Custom Key** on the right side of **TDE Status**.

6. In the **Configure TDE** dialog box, select **Use Existing Custom Key**.

> **Note** If you do not have a custom key, click **Create Custom Key** to create a key in the KMS console and import the key material. For more information, see Create a CMK.



7. In the message that appears, click **OK**.

FAQ

- After I enable TDE, can I still use common database tools, such as Navicat?

  Yes, you can still use common database tools after you enable TDE.

- After I enable TDE, why is my data still in plaintext?

  After TDE is enabled, the stored data is encrypted. When data is queried, it is decrypted and read to the memory. Therefore, it is displayed in plaintext.

Related API operations

API	Description
CreateDBCluster	Creates a PolarDB cluster and enables TDE. **Note** The DBType parameter must be set to PostgreSQL or Oracle.

13.Diagnostics and optimization

13.1. SQL Explorer

Apsara PolarDB provides the SQL Explorer feature. You can use SQL Explorer for database security auditing and performance diagnostics.

Pricing

- The trial edition of Apsara PolarDB is available for free. In the trial edition, audit logs are retained for only one day. You can query only data that is stored in the retained audit logs. The trial edition does not support advanced features. For example, data cannot be exported, and data integrity cannot be ensured.
- If you want to retain the audit logs for 30 days or longer, you can view the pricing details in Pricing of SQL Explorer (optional).

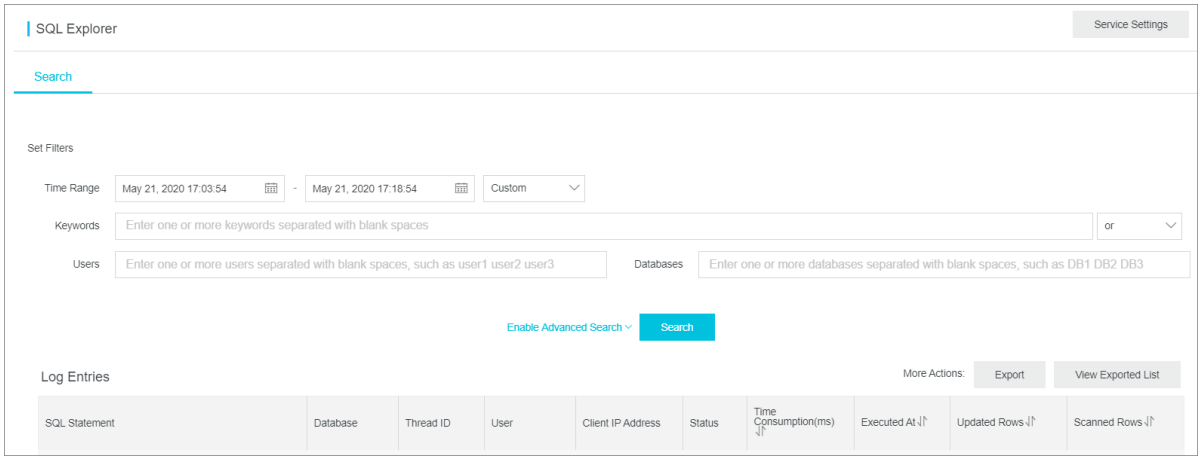
Features

- SQL logging

  SQL audit logs record all operations that are performed on databases. You can use audit logs to identify database failures, analyze behaviors, and perform security auditing.

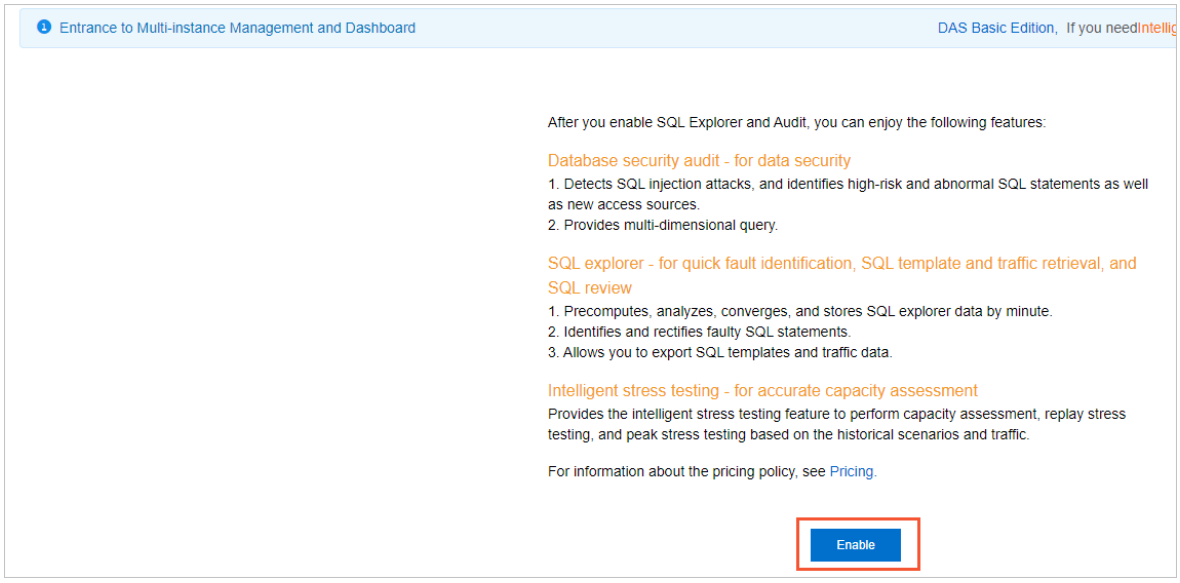
- Advanced search

  SQL Explorer allows you to search data by database, user, client IP, thread ID, execution duration, or execution status. You can also export and download search results.

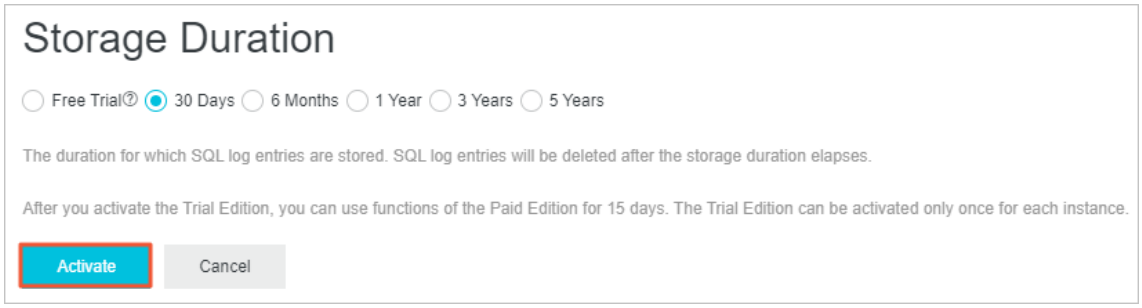


Enable SQL Explorer

1. Log on to the Apsara PolarDB console.
2. In the upper-left corner of the console, select the region where the target cluster resides.
3. Find the target cluster and click its ID.
4. In the left-side navigation pane, choose **Log and Audit > SQL Explorer**.
5. Click **Activate Now**.



6. Specify the storage duration of SQL audit logs, and then click **Activate**.

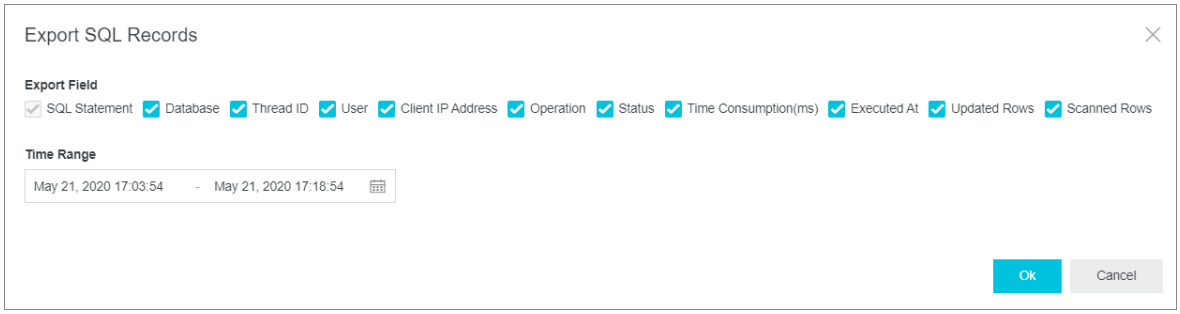


Change the storage duration of SQL audit logs

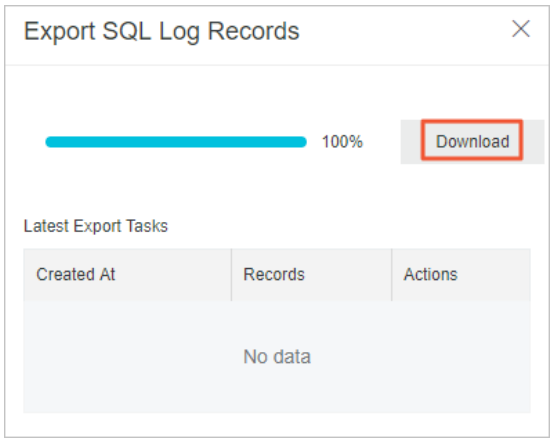
1. Log on to the Apsara PolarDB console.
2. In the upper-left corner of the console, select the region where the target cluster resides.
3. Find the target cluster and click its ID.
4. In the left-side navigation pane, choose **Log and Audit > SQL Explorer**.
5. In the upper-right corner of the page, click **Service Settings**.
6. Change the storage duration and click **OK**.

Export SQL records

1. Log on to the Apsara PolarDB console.
2. In the upper-left corner of the console, select the region where the target cluster resides.
3. Find the target cluster and click its ID.
4. In the left-side navigation pane, choose **Log and Audit > SQL Explorer**.
5. On the right side of the page, click **Export**.
6. In the dialog box that appears, specify the **Export Field** and **Time Range** parameters, and click **OK**.



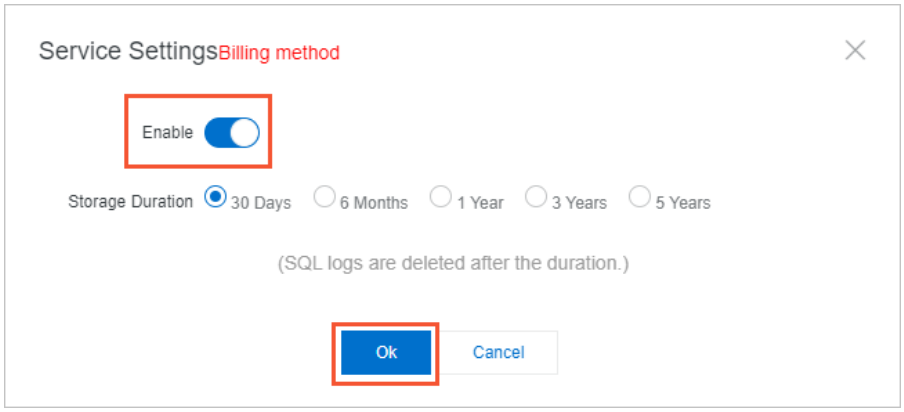
7. After the export is complete, download the log files in the **Export SQL Log Records** dialog box.



Disable SQL Explorer

> **Note**
>
> After SQL Explorer is disabled, SQL audit logs are deleted. We recommend that you export and save SQL log files to your computer before you disable SQL Explorer.

1. Log on to the Apsara PolarDB console.
2. In the upper-left corner of the console, select the region where the target cluster resides.
3. Find the target cluster and click its ID.
4. In the left-side navigation pane, choose **Log and Audit > SQL Explorer**.
5. In the upper-right corner of the page, click **Service Settings**.
6. Change the storage duration and click **OK**.
7. Turn off the Activate SQL Explorer switch.

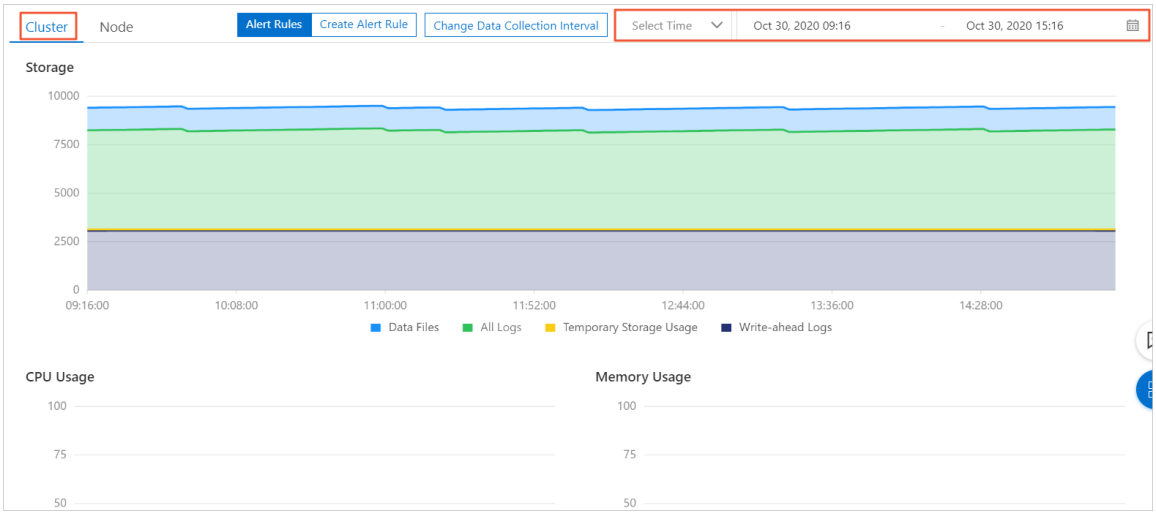


13.2. Performance monitoring

The console allows you to monitor a variety of performance metrics and view monitoring data at intervals of seconds. You can monitor the status of your clusters and locate faults based on the monitoring data.

Performance monitoring

1.
2.
3.
4. In the left-side navigation pane, choose **Diagnostics and Optimization > Monitoring**.
5. View the monitoring information about a **Cluster** or **Compute Node** based on your needs. For more information, see Metrics.
   - To monitor cluster performance, click the **Cluster** tab. Specify a monitoring period in the date and time picker and click **OK**.



   - To monitor node performance, click the **Compute Node** tab and select a node from the drop-down list. Specify a monitoring period in the date and time picker and click **OK**.

Metrics

Category	Metric	Description
Cluster	Storage	Displays the usage of data space, log space, temporary space, and WAL log space.
	CPU	Displays the CPU utilization of each node.

Category	Metric	Description
	Memory Usage	Displays the memory usage of each node.
Node	TPS	Displays the number of transactions per second of the selected node, including the number of committed transactions per second, deadlocked transactions per second, and rollback transactions per second.
	CPU	Displays the CPU utilization of the selected node.
	Memory Usage	Displays the memory usage of the selected node.
	Connections	Displays the total number of current connections, the number of active connections, and the number of idle connections for the selected node.
	Scanned Rows	Displays the numbers of rows that are inserted, read, updated, deleted, and returned per second on the selected node.
	Maximum Database Age	Displays the difference between the transaction IDs of the earliest and latest transactions in the database.
	I/O Throughput	Displays the total I/O throughput, I/O read throughput, and I/O write throughput of the selected node.
	IOPS	Displays the following IOPS types of the selected node: the total IOPS, read IOPS, and write IOPS.
	Cache	Displays the cache reads per second and disk reads per second of the selected node.
	Cache Hit Ratio	Displays the cache hit ratio of the selected node.
	Temporary Files	Displays the number and total size of temporary files on the selected node.

Related operations

API	Description
DescribeDBClusterPerformance	Queries the performance data of a cluster.
DescribeDBNodePerformance	Queries the performance data of a specified node in a specified cluster.
DescribeDBClusterMonitor	Queries the interval for collecting the monitoring data of a specified cluster.
ModifyDBClusterMonitor	Changes the interval for collecting the monitoring data of a specified cluster.

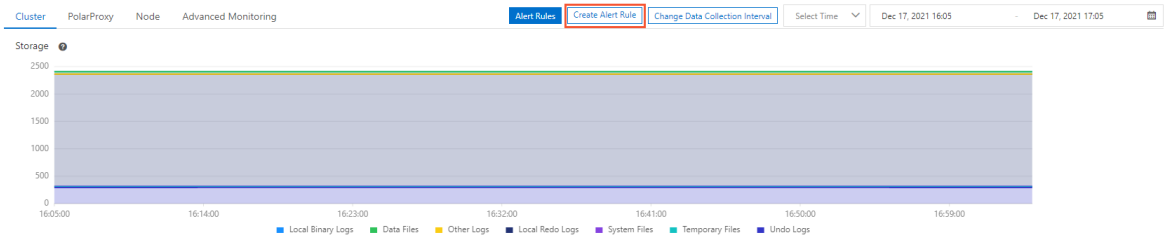
13.3. Create an alert rule

This topic describes how to create and manage rules that can be used to trigger threshold alerts in the console. This helps you identify and handle exceptions of clusters and nodes at the earliest opportunity.

Procedure

1.
2.
3.
4. In the left-side navigation pane, choose **Diagnostics and Optimization > Monitoring**.

5. Click **Create Alert Rule**.



6. On the **Create Alert Rule** page, specify the following parameters.

Step	Parameter	Description
Related Resource	Product	The service that you want to monitor. Use the default value **PolarDB for Oracle**.
	Resource Range	The application scope of the alert rule. Set this parameter to **All Resources** or **Cluster**. **Note** ◦ If the **Resource Range** parameter is set to **All Resources**, the system sends alert notifications if one of the clusters triggers the alert. The Rule Description parameter specifies the conditions that are used to trigger the alert. ◦ If the **Resource Range** parameter is set to **Cluster**, the system sends alert notifications only if the specified cluster triggers the alert. The Rule Description parameter specifies the conditions that are used to trigger the alert.
Set Alert Rules	Alert Rule	The name of the alert rule.
	Rule Description	The content of the alert rule. This parameter specifies the conditions that are used to trigger the alert. **Note** For more information about how to create alert rules, see Create a threshold-triggered alert rule.
	Mute for	The interval at which the system resends the alert notification if the issue that triggers the alert persists. The minimum value is 5 minutes and the maximum value is 24 hours.
	Effective Period	The validity period of the alert rule. **Note** The system sends alert notifications only within the validity period of an alert rule and records events when the validity period expires.
Notification Method	For more information about how to specify **Notification Method**, see Create a threshold-triggered alert rule	

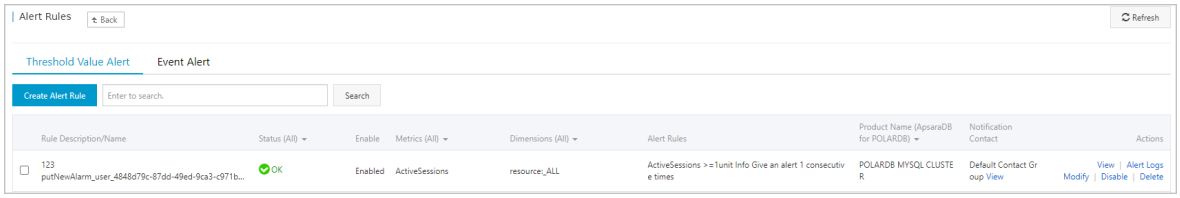
7. Click **Confirm**.

13.4. Manage alert rules

This topic describes how to manage alert rules that are based on threshold values in the console. The alert feature helps you detect exceptions of clusters and nodes and handle the exceptions in time.

Procedure

1.
2.
3.
4. In the left-side navigation pane, choose **Diagnostics and Optimization > Monitoring**.
5. Click **Alert Rules**. The **Alert Rules** page appears.



6. On the **Threshold Value Alert** tab, you can perform the following operations to manage the existing alert rules:
   - To view the basic information about an alert rule, click **View** in the **Actions** column of the alert rule.
   - To view the alert history associated with an alert rule, click **Alert Logs** in the **Actions** column of the alert rule.
   - To modify an alert rule, click **Modify** in the **Actions** column of the alert rule.
   - To disable an alert rule, click **Disable** in the **Actions** column of the alert rule.
   - To delete an alert rule, click **Delete** in the **Actions** column of the alert rule.
   - To view the alert contact group, alert contacts, and alert notification method for an alert rule, click **View** in the **Notification Contact** column of the alert rule.

13.5. Performance insight

provides the diagnostics feature that integrates some features of Database Autonomy Service (DAS). You can use the Performance Insight feature to rapidly evaluate database loads and identify the root causes of performance issues. This helps you improve the database stability.

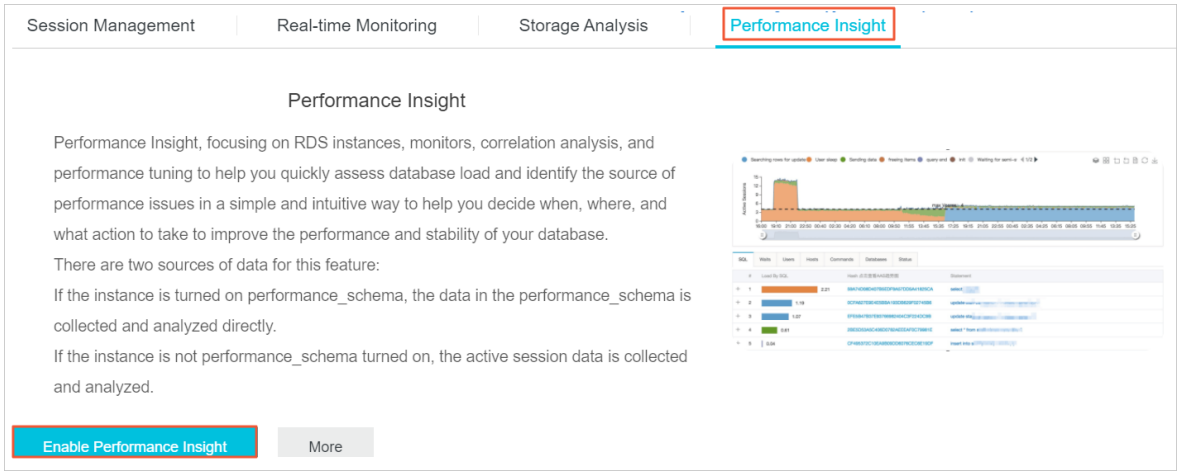
Background

The Performance Insight feature supports the following data sources:

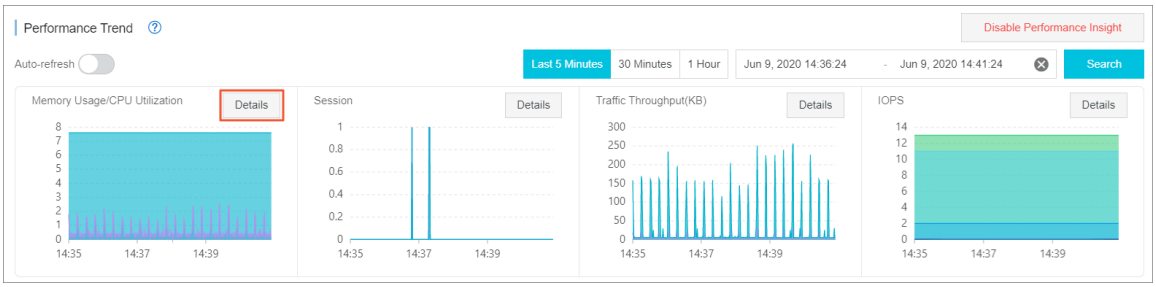
- If performance_schema is enabled for the desired instances, the Performance Insight feature directly collects and analyzes the data stored in performance_schema.
- If performance_schema is disabled for the desired instances, the Performance Insight feature collects and analyzes the data of active sessions.

Procedure

1.
2.
3. On the **Clusters** page, find the cluster for which you want to enable the autonomy service, and click the cluster ID.
4. In the left-side navigation pane, choose **Diagnostics and Optimization > Diagnosis**.
5. Click the **Performance Insight** tab.
6. Click **Enable Performance Insight**.

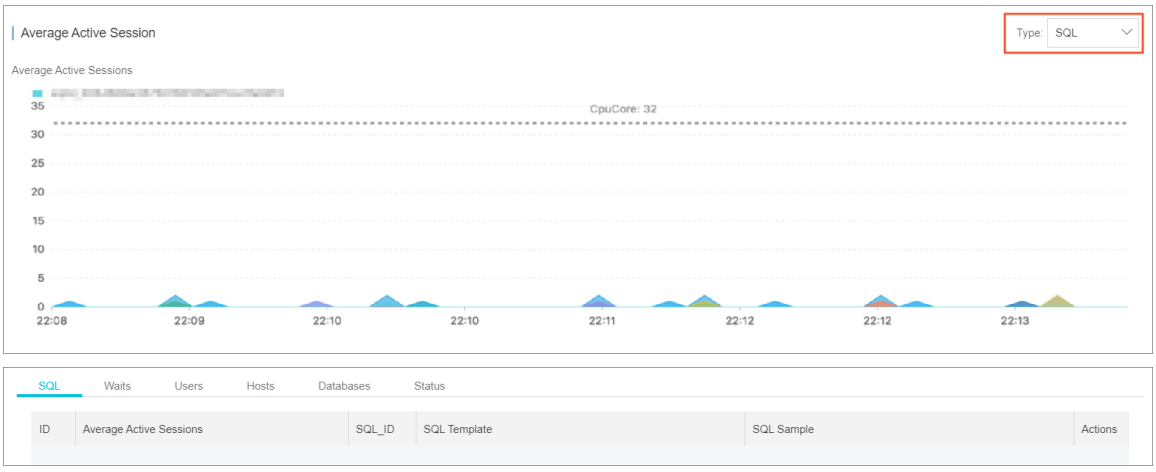


7. In the dialog box that appears, click **Confirm**.
8. On the Performance Insight tab, view and manage the following information:
   - In the **Performance Trend** section, you can specify a time range to view the performance of databases. If you need to view a specific performance metric, such as CPU usage, click **Details** next to the performance metric name.



> **Note** The duration of the specified time range cannot exceed seven days.

   - In the **Average Active Session** section, you can view the trend charts of different types of sessions, such as SQL, and the relevant multidimensional details of service loads. This helps you identify the root causes of performance issues.



14.Configuration parameters

14.1. polar_comp_redwood_date

If DATE appears as the data type of a column in the statements and the polar_comp_redwood_date configuration parameter is set to *TRUE*, DATE is translated to TIMESTAMP when the table definition is stored in the database. In this case, a time component is also stored in the column along with the date. This rule is consistent with the DATE data type of Oracle.

If polar_comp_redwood_date is set to *FALSE*, the data type of the column in a CREATE TABLE or ALTER TABLE statement remains as a native PostgreSQL DATE data type and is stored in the database. PostgreSQL DATE data type stores only the date without a time component in the column.

DATE can appear as a data type in any other context such as the data type of a variable in an SPL declaration section, or the data type of a formal parameter in an SPL procedure or SPL function, or the return type of an SPL function. In this case, regardless of the setting of polar_comp_redwood_date, DATE is always internally translated to a TIMESTAMP and can thus handle an existing time component.

14.2. polar_comp_redwood_raw_names

If polar_comp_redwood_raw_names is set to the default value FALSE, database object names, such as table names, column names, trigger names, program names, and user names, appear in uppercase letters when viewed from Oracle catalogs. In addition, quotation marks enclose names that are created with enclosed quotation marks.

If polar_comp_redwood_raw_names is set to TRUE, the database object names are displayed in the way as they are stored in the PostgreSQL system catalogs when viewed from the Oracle catalogs. Thus, names created without enclosed quotation marks appear in lowercase as expected in PostgreSQL. Names created with enclosed quotation marks appear in the way as they are created, but without the quotation marks.

For example, the following user name is created and then a session is started with that user.

```
CREATE USER reduser IDENTIFIED BY password;
polardb=# \c - reduser
Password for user reduser:
You are now connected to database "polardb" as user "reduser".
```

When you connect to the database as reduser, the following tables are created:

```
CREATE TABLE all_lower (col INTEGER);
CREATE TABLE ALL_UPPER (COL INTEGER);
CREATE TABLE "Mixed_Case" ("Col" INTEGER);
```

When viewed from the Oracle catalog named USER_TABLES, with polar_comp_redwood_raw_names set to the default value FALSE, the names appear in uppercase except for the Mixed_Case name. This name appears in the same way as the name is created and enclosed with quotation marks.

```
polardb=> SELECT * FROM USER_TABLES;
 schema_name |  table_name
| tablespace_name | status | temporary
-------------+--------------+-----------------+--------+-----------
 REDUSER     | ALL_LOWER    |                 | VALID  | N
 REDUSER     | ALL_UPPER    |                 | VALID  | N
 REDUSER     | "Mixed_Case" |                 | VALID  | N
(3 rows)
```

When viewed with polar_comp_redwood_raw_names set to TRUE, the names appear in lowercase except for the Mixed_Case name. This name appears in the same way as the name is created, but the name is not enclosed with quotation marks.

```
polardb=> SET polar_comp_redwood_raw_names TO true;
SET
polardb=> SELECT * FROM USER_TABLES;
 schema_name | table_name |
tablespace_name | status | temporary
-------------+------------+-----------------+--------+-----------
 reduser     | all_lower  |                 | VALID  | N
 reduser     | all_upper  |                 | VALID  | N
 reduser     | Mixed_Case |                 | VALID  | N
(3 rows)
```

These names match the case when viewed from the PostgreSQL pg_tables catalog.

```
polardb=> SELECT schemaname, tablename, tableowner FROM pg_tables WHERE
tableowner = 'reduser';
 schemaname | tablename  | tableowner
------------+------------+------------
 reduser    | all_lower  | reduser
 reduser    | all_upper  | reduser
 reduser    | Mixed_Case | reduser
(3 rows)
```

14.3. polar_comp_redwood_strings

In Oracle, when a string is concatenated with a null variable or null column, the result is the original string. However, in PostgreSQL, concatenation of a string with a null variable or null column generates a null result. If the polar_comp_redwood_strings parameter is set to TRUE, the preceding concatenation operation results in the original string in the same way as Oracle does. If the polar_comp_redwood_strings parameter is set to FALSE, the native PostgreSQL behavior is maintained.

The following example illustrates the difference. The sample application introduced in the next section contains a table of employees. This table has a column named comm that is null for most employees. The following query has polar_comp_redwood_string set to FALSE. The concatenation of a null column with non-empty strings generates a final result of null, so only employees that have a commission appear in the query result. The output line for all other employees is null.

```
SET polar_comp_redwood_strings TO off;
SELECT RPAD(ename,10) || ' ' || TO_CHAR(sal,'99,999.99') || ' ' ||
TO_CHAR(comm,'99,999.99') "EMPLOYEE COMPENSATION" FROM emp;
      EMPLOYEE COMPENSATION
---------------------------------
 ALLEN        1,600.00     300.00
 WARD         1,250.00     500.00
 MARTIN       1,250.00   1,400.00


 TURNER       1,500.00        .00


(14 rows)
```

The following example is the same query executed when polar_comp_redwood_strings is set to TRUE. The value of a null column is treated as an empty string. The concatenation of an empty string with a non-empty string generates the non-empty string. This result is consistent with the results generated by Oracle for the same query.

```
SET polar_comp_redwood_strings TO on;
SELECT RPAD(ename,10) || ' ' || TO_CHAR(sal,'99,999.99') || ' ' ||
TO_CHAR(comm,'99,999.99') "EMPLOYEE COMPENSATION" FROM emp;
      EMPLOYEE COMPENSATION
---------------------------------
 SMITH          800.00
 ALLEN        1,600.00     300.00
 WARD         1,250.00     500.00
 JONES        2,975.00
 MARTIN       1,250.00   1,400.00
 BLAKE        2,850.00
 CLARK        2,450.00
 SCOTT        3,000.00
 KING         5,000.00
 TURNER       1,500.00        .00
 ADAMS        1,100.00
 JAMES          950.00
 FORD         3,000.00
 MILLER       1,300.00
(14 rows)
```

14.4. polar_comp_stmt_level_tx

In Oracle, when a runtime error occurs in a SQL statement, all the updates on the database caused by that single statement are rolled back. This is called statement-level transaction isolation. For example, if a single UPDATE statement updates five rows but an attempt to update a sixth row results in an error, the updates to all six rows made by this UPDATE statement are rolled back. The effects of prior SQL statements that have not yet been committed or rolled back are pending until a COMMIT or ROLLBACK statement is executed.

In PostgreSQL, if an error occurs while executing a SQL statement, all the updates on the database since the start of the transaction are rolled back. In addition, the transaction is left in a terminated state and either a COMMIT or ROLLBACK statement must be executed before another transaction can be started.

If polar_comp_stmt_level_tx is set to TRUE, an error does not automatically roll back prior uncommitted database updates, similar to the Oracle behavior. If polar_comp_stmt_level_tx is set to FALSE, an error rolls back uncommitted database updates.

Notice Set polar_comp_stmt_level_tx to TRUE only when necessary. This setting may decrease the service performance.

As shown in the following example running in PSQL, if polar_comp_stmt_level_tx is set to FALSE, the first INSERT statement is still rolled back after the second INSERT statement is terminated. In PSQL, the statement \set AUTOCOMMIT off must be used. Otherwise every statement commits automatically. This defeats the purpose of this demonstration of the effect of polar_comp_stmt_level_tx.

```
\set AUTOCOMMIT off
SET polar_comp_stmt_level_tx TO off;
INSERT INTO emp (empno,ename,deptno) VALUES (9001, 'JONES', 40);
INSERT INTO emp (empno,ename,deptno) VALUES (9002, 'JONES', 00);
ERROR:  insert or update on table
"emp" violates foreign key constraint "emp_ref_dept_fk"
DETAIL:  Key (deptno)=(0) is not present in table "dept".
COMMIT;
SELECT empno, ename, deptno FROM emp WHERE empno > 9000;
empno | ename | deptno
-------+-------+--------
(0 rows)
```

In the following example, polar_comp_stmt_level_tx is set to TRUE. The first INSERT statement has not been rolled back after an error occurs in the second INSERT statement. At this point, the first INSERT statement can either be committed or rolled back.

```
\set AUTOCOMMIT off
SET polar_comp_stmt_level_tx TO on;
INSERT INTO emp (empno,ename,deptno) VALUES (9001, 'JONES', 40);
INSERT INTO emp (empno,ename,deptno) VALUES (9002, 'JONES', 00);
ERROR:  insert or update on table
"emp" violates foreign key constraint "emp_ref_dept_fk"
DETAIL:  Key (deptno)=(0) is not present in table "dept".
SELECT empno, ename, deptno FROM emp WHERE empno > 9000;
empno | ename | deptno
-------+-------+--------
  9001 | JONES |     40
(1 row)
COMMIT;
```

A ROLLBACK statement may be executed instead of the COMMIT statement. In this case, the insert of employee number 9001 is also rolled back.

14.5. polar_create_table_with_full_replica_identity

When you synchronize tables that have no primary keys in to other databases by using the logical replication method, errors may be reported for the operations on the tables. You can specify the `polar_create_table_with_full_replica_identity` parameter to resolve this issue.

Logical replication of uses a publish and subscribe model. The operations on the publisher side can be executed on the subscriber side in a similar way to Structural Query Language (SQL) so that data can be synchronized. The replica identities of the tables must be configured on the publisher side so that the data to be updated or deleted on the subscriber side can be identified.

The following types of replica identities are supported:

- Primary key
- Unique index
- FULL (a full row of data)

By default, the replica identity is the primary key. If logical replication is implemented on a table that has no primary key, an error occurs for the change operations. As a result, services cannot run as expected. The following error message is returned:

```
ERROR:  cannot delete from table "polardb_test" because it does not have a replica identity and publishes d
eletes
HINT:  To enable deleting from the table, set REPLICA IDENTITY using ALTER TABLE.
```

Notice When you use logical replication, make sure that all the replica identities of the tables that are to be synchronized and have no primary keys are set to `FULL`. For example, when you use Data Transmission Service (DTS) to synchronize data, comply with this rule.

provides the following two methods for you to change the replica identities of tables to `FULL`.

- Run the following command to change the replica identity of an existing table to `FULL`:

  ```
  ALTER TABLE <table_name> REPLICA IDENTITY FULL;
  ```

- Specify the `polar_create_table_with_full_replica_identity` parameter as *on* to set the default replica identity of a newly created table to `FULL`.

> **Note**
>
> The default value of the `polar_create_table_with_full_replica_identity` parameter is *off*. You cannot modify this parameter in the console. If you need to modify this parameter, to contact technical support.

14.6. Custom parameters

The following table describes custom parameters.

Parameter name	Valid value	Restart required	Description
autovacuum_max_workers	[5-20]	Yes	Sets the maximum number of simultaneously running autovacuum worker processes.
autovacuum_vacuum_cost_delay	[-1-100]	No	Vacuum cost delay in milliseconds, for autovacuum.
autovacuum_vacuum_cost_limit	[-1-10000]	No	Vacuum cost amount available before napping, for autovacuum.
auto_explain.log_analyze	[on|off]	No	Use EXPLAIN ANALYZE for plan logging.
auto_explain.log_buffers	[on|off]	No	Log buffers usage.
auto_explain.log_format	[text|xml|json|yaml]	No	EXPLAIN format to be used for plan logging.
auto_explain.log_min_duration	[-1-2147483647]	No	Sets the minimum execution time above which plans will be logged.
auto_explain.log_nested_statements	[on|off]	No	Log nested statements.
auto_explain.log_timing	[on|off]	No	Collect timing data, not just row counts.
auto_explain.log_triggers	[on|off]	No	Include trigger statistics in plans.
auto_explain.log_verbose	[on|off]	No	Use EXPLAIN VERBOSE for plan logging.
auto_explain.sample_rate	[0-1]	No	Fraction of queries to process.
default_transaction_deferrable	[on|off]	No	Sets the default deferrable status of new transactions.
enable_partitionwise_aggregate	[on|off]	No	Enables partitionwise aggregation and grouping.
enable_partitionwise_join	[on|off]	No	Enables partitionwise join.
extra_float_digits	[-15-3]	No	Sets the number of digits displayed for floating-point values.
idle_in_transaction_session_timeout	0 or [1000-2000000000]	No	Sets the maximum allowed duration of any idling transaction.
jit	[on|off]	No	Allow JIT compilation.

Parameter name	Valid value	Restart required	Description
lock_timeout	0 or [1000-2000000000]	No	Sets the maximum allowed duration of any wait for a lock.
log_min_duration_statement	[-1-2147483647]	No	Sets the minimum execution time above which statements will be logged.
log_statement	[none|ddl|mod|all]	No	Sets the type of statements logged.
log_temp_files	[-1-2147483647]	No	Log the use of temporary files larger than this number of kilobytes.
max_parallel_workers	[0-512]	No	Sets the maximum number of parallel workers that can be active at one time.
max_parallel_workers_per_gather	[0-512]	No	Sets the maximum number of parallel processes per executor node.
max_sync_workers_per_subscription	[0-262143]	No	Maximum number of table synchronization workers per subscription.
min_parallel_index_scan_size	[0-715827882]	No	Sets the minimum amount of index data for a parallel scan.
min_parallel_table_scan_size	[0-715827882]	No	Sets the minimum amount of table data for a parallel scan.
old_snapshot_threshold	[-1-86400]	Yes	Time before a snapshot is too old to read pages changed after the snapshot was taken.
polar_comp_dynatune	[0-100]	Yes	Sets the polardb utilization percentage.
polar_comp_dynatune_profile	[oltp|reporting|mixed]	Yes	Sets the workload profile for dynatune.
polar_comp_enable_pruning	[on|off]	No	Enables the planner to early-prune partitioned tables.
polar_comp_redwood_date	[on|off]	No	Determines whether DATE should behave like e TIMESTAMP or not.
polar_comp_redwood_greatest_least	[on|off]	No	Determines whether GREATEST AND LEAST function should behave like Redwood or PG.
polar_comp_redwood_raw_names	[on|off]	No	Return the unmodified name stored in the PostgreSQL system catalogs from Redwood interfaces.
polar_comp_redwood_strings	[on|off]	No	Treat NULL as an empty string when concatenated with a TEXT value.

Parameter name	Valid value	Restart required	Description
polar_comp_stmt_level_tx	[on|off]	No	Allows continuing on errors instead of requiring a transaction abort.
statement_timeout	0 or [1000-2000000000]	No	Sets the maximum allowed duration of any statement.
temp_file_limit	[-1-1048576000]	No	Limits the total size of all temporary files used by each process.
timezone	^'(((UTC)(-){0,1}(d|[1-9]d|1([0-5]d|6[0-7])))|((GMT)(-){0,1}(d|[1-9]d|1([0-5]d|6[0-7])))|CST6CDT|Poland|Kwajalein|MST|NZ|Universal|Libya|Turkey|EST5EDT|Greenwich|NZ-CHAT|MET|Portugal|GMT-0|CET|Eire|PST8PDT|Jamaica|GMT|Zulu|Japan|ROC|GB-Eire|ROK|Navajo|Singapore|posixrules|GB|EST|GMT0|Hongkong|PRC|Iran|MST7MDT|WET|W-SU|UCT|Cuba|Egypt|EET|Israel|UTC|HST|Iceland)'$	No	Sets the time zone for displaying and interpreting time stamps.
track_commit_timestamp	[on|off]	Yes	Collects transaction commit time.
vacuum_defer_cleanup_age	[0-1000000]	No	Number of transactions by which VACUUM and HOT cleanup should be deferred, if any.
wal_level	[replica|logical]	Yes	Set the level of information written to the WAL.
work_mem	[4096-524288]	No	Sets the maximum memory to be used for query workspaces.

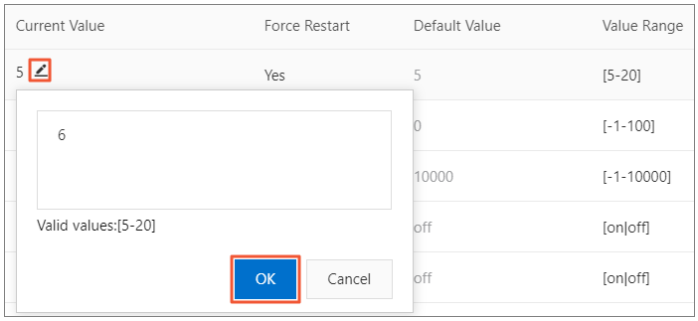
14.7. Configure cluster parameters

This topic shows you how to modify cluster parameters in the console.

> **Note** The console displays the parameters that you can modify.

Procedure

1.
2.
3.
4. In the left-side navigation pane, choose **Settings and Management > Parameters**.
5. Find the parameter that you want to modify, and click the icon in the **Current Value** column. In the dialog box that appears, enter the new parameter value, and click **OK**.

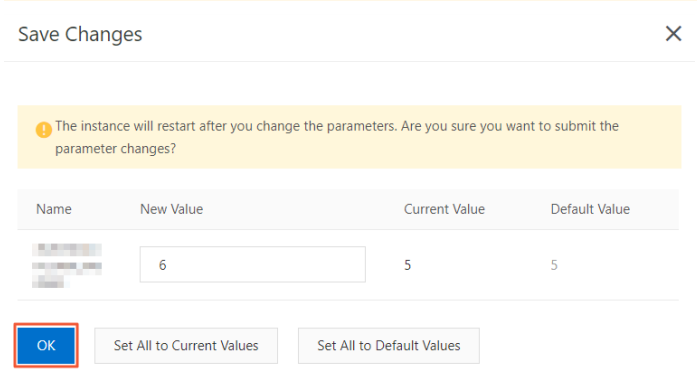


Note

- You must enter a parameter value that is included in the right-side **Value Range** column. Otherwise, an error message appears when you click **Apply Changes**.
- You can move the pointer over the icon of a parameter to view the parameter details.

6. In the upper-left corner of the page, click **Apply Changes**. In the **Save Changes** panel, click **OK**.

Warning If a parameter has a **Yes** value for the **Force Restart** column, the cluster restarts after you click **OK**. We recommend that you make appropriate service arrangements before you modify parameters. Proceed with caution.



Related API operations

API	Description
DescribeDBClusterParameters	Queries cluster parameters.
ModifyDBClusterParameters	Modifies cluster parameters.

15.Version Management

The architecture of a cluster consists of three layers: PolarProxy, the database engine, and the distributed storage. You can upgrade PolarProxy or the database engine separately or upgrade both of them at the same time.

Usage notes

- In most cases, an upgrade requires less than 30 minutes to complete. PolarProxy or the database engine is restarted during the upgrade. This can cause transient connections to occur on the database. We recommend that you perform the upgrade during off-peak hours. Make sure that your application can automatically reconnect to your database.
- During the **Upgrade PolarDB Database Proxy and Kernel** process, transient connections occur for the primary endpoint and the cluster endpoint and last for 30 to 90 seconds. Make sure that your application can automatically reconnect to your database.
- During the **Upgrade PolarDB Database Proxy Only** process, transient connections occur for the cluster endpoint and the custom endpoint and last for 30 seconds. Make sure that your application can automatically reconnect to your database.
- During the **Upgrade Kernel Only** process, clusters with PolarProxy of V2.4.7 or later can use the connection preservation technique to prevent 95% of the database connections from being interrupted.
- During the upgrade, you cannot use some features that are related to cluster changes in the console. For example, you cannot upgrade or downgrade configurations, add or delete nodes, modify parameters, or restart nodes. During this period, features related to data queries are still available, such as performance monitoring.
- You cannot downgrade PolarProxy or the database engine.

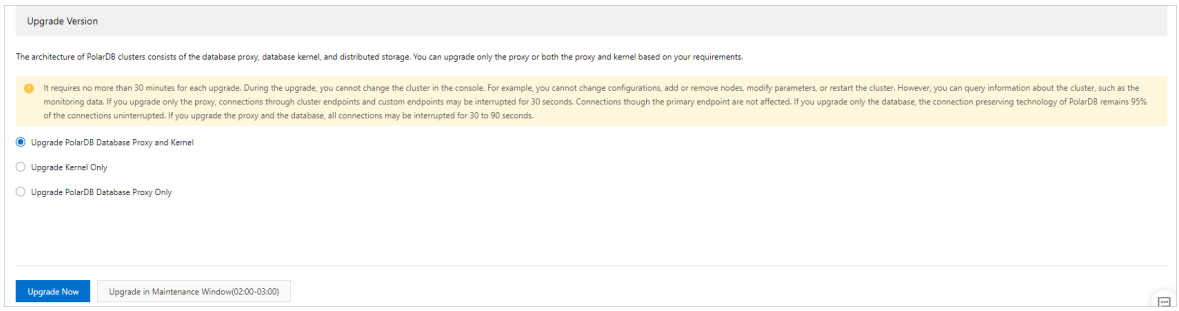
View the version information

1.
2.
3.
4. In the **Version Information** section, view the version information of PolarProxy and the database engine.

Upgrade the version

If PolarProxy or the database engine of the cluster is not of the latest version, you can upgrade the version based on your business requirements.

1. On the details page of the cluster that you want to manage, choose **Settings and Management > Version Management**. In the **Upgrade Version** section, select **Upgrade PolarDB Database Proxy and Kernel**, **Upgrade Kernel Only**, or **Upgrade PolarDB Database Proxy Only** as needed.



> **Note**
>
> - If PolarProxy or the database engine of the cluster is of the latest version, the **Upgrade PolarDB Database Proxy and Kernel**, **Upgrade Kernel Only**, and **Upgrade PolarDB Database Proxy Only** options are unavailable.
> - If you select **Upgrade PolarDB Database Proxy Only**, only the features related to read/write splitting are upgraded, such as the consistency level, transaction splitting, and whether to offload reads from the primary node. The default consistency level is global consistency.

2. Click **Upgrade Now** or **Upgrade in Maintenance Window**.

   If you select Upgrade in Maintenance Window, you can view the details of the task or cancel the task on the **Scheduled Tasks** page.

   > **Notice**
   > - During the **Upgrade PolarDB Database Proxy and Kernel** process, transient connections occur for the primary endpoint and the cluster endpoint and last for 30 to 90 seconds. Make sure that your application can automatically reconnect to your database.
   > - During the **Upgrade PolarDB Database Proxy Only** process, transient connections occur for the cluster endpoint and the custom endpoint and last for 30 seconds. Make sure that your application can automatically reconnect to your database.

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

Related API operations

API	Description
DescribeDBClusterVersion	Queries the details about the database engine version of a cluster.
UpgradeDBClusterVersion	Upgrades a cluster to the latest version.

16.SQL firewalls

This topic describes how to use the SQL/Protect plug-in to protect databases from SQL injection attacks.

Context

Developers are responsible for protecting databases against SQL injection attacks. Database administrators can prevent only a few types of SQL injection attacks. SQL/Protect detects SQL injection attacks based on query requests. If suspicious query requests are identified, SQL/Protect immediately sends alerts to database administrators and prevents the queries from running.

Types of SQL injection attacks

Attack type	Description
Unauthorized relations	Administrators can restrict access to tables. This operation is tedious. SQL/Protect provides a learn mode that dynamically tracks the relationship of tables accessed by a user. In learn mode, SQL/Protect can automatically learn which tables an application can be allowed to access for a user or group. When SQL/Protect is in passive mode or active mode, the incoming queries are checked based on the list of learned tables.
Utility commands	A common technique used in SQL injection attacks is to run utility commands such as typical DDL statements. For example, a user-defined function is created to access the data of other tables. SQL/Protect can prevent some utility commands from being run. In most cases, these commands are not used in applications.
SQL tautology	The most frequent technique used in SQL injection attacks is to issue a tautological WHERE clause. A tautological WHERE clause contains a condition that is always true, such as `WHERE password = 'x' OR 'x'='x'`. In most cases, attackers use this technique to identify security vulnerabilities. SQL/Protect can block queries that contain a tautological conditional clause.
Unbounded DML statements	Unbounded DML statements are database update statements in which no conditions are specified. These statements are UPDATE and DELETE statements that have no WHERE clauses. For example, an attacker may update or delete the passwords of users to initiate a denial-of-service (DoS) attack.

Protected roles

Protected roles are users or groups protected by SQL/Protect. Database administrators can use SQL/Protect to specify protected roles. You can use SQL/Protect to customize different levels of injection attack prevention for different protected roles. The types of SQL injection attacks vary based on the levels.

A role that has the superuser privilege cannot be a protected role. A protected non-superuser role can become a protected superuser role. In this case, SQL/Protect performs operations for the protected superuser role in the following scenarios:

- SQL/Protect generates an alert for each command run by the protected superuser.
- When SQL/Protect is in active mode, SQL/Protect blocks all commands run by the protected superuser.

When SQL/Protect is running, a protected role that has the superuser privilege is changed to a common role or restored to an unprotected role.

In addition, each command run by a protected role is recorded in a statistics view. This view helps you identify the start of a potential SQL injection attack against the role. The statistics are collected based on the type of SQL injection attack.

> **Note** By default, each database supports up to 64 protected roles and up to 1024 protected tables. The maximum number of roles that can be protected is specified by the `max_protected_roles` parameter. The maximum number of tables that can be protected is specified by the `max_protected_relations` parameter.

Use the administrator role to configure SQL/Protect for a database

1. Modify the parameters in the following code block to enable SQL/Protect.

```
set polar_sql_protect.enabled = on; #(The default value is off.)
set polar_sql_protect.level = passive; #(Valid values: learn, active, and passive. The default value is passive.)
```

2. Create a test database named targetdb and a test user named test.

```
CREATE DATABASE targetdb;
CREATE ROLE test;
GRANT ALL ON DATABASE targetdb TO test;
ALTER ROLE test LOGIN;
```

3. Log on to the test database targetdb. Then, execute the following statements to create SQL/Protect and add protected roles:

```
CREATE EXTENSION sqlprotect;
SELECT sqlprotect.protect_role('test');
```

View the list of protected roles.

```
SELECT * FROM sqlprotect.list_protected_users;
SELECT * FROM sqlprotect.polar_sql_protect;
```

4. Change the mode in which SQL/Protect works based on your requirements.

SQL/Protect works in three modes: learn, active, and passive. The default mode is passive. For more information, see Configure the mode in which SQL/Protect works to monitor a protected role.

- Change the mode in which SQL/Protect works to learn.

```
polar_sql_protect.level = learn; #(Valid values: learn, active, and passive. The default value is passive.)
```

a. Log on to the targetdb database as the test user. Then, create a test table named company and execute the SELECT and INSERT statements:

```
CREATE TABLE company(name VARCHAR(100), employee_num INT);
SELECT * FROM company;
INSERT INTO company VALUES('new', 1);
SELECT * FROM company;
```

b. View the learned information about tables used by the test user.

```
SELECT * FROM sqlprotect.polar_sql_protect_rel;
SELECT * FROM sqlprotect.list_protected_rels;
```

- Change the mode in which SQL/Protect works to passive.

```
polar_sql_protect.level = passive; #(Valid values: learn, active, and passive. The default value is passive.)
```

a. Log on to the targetdb database as the test user.

b. Inject SQL statements.

```
SELECT * FROM company WHERE 1 = 1;
DELETE FROM company;
```

> **Note** SQL/Protect returns a message that indicates unauthorized SQL statements. However, SQL/Protect does not prevent the SQL statements from being executed.

- Change the mode in which SQL/Protect works to active.

```
polar_sql_protect.level = active; #(Valid values: learn, active, and passive. The default value is passive.)
```

a. Log on to the targetdb database as the test user.

b. Inject SQL statements.

```
SELECT * FROM company WHERE 1 = 1;
DELETE FROM company;
```

> **Note** SQL/Protect returns a message that indicates unauthorized SQL statements. SQL/Protect also prevents the SQL statements from being executed.

Configure protected roles

Protected roles are stored in the polar_sql_protect table. The database administrator can choose the users and user groups that are protected, and add the users and user groups to the table.

- Invoke the protect_role function to add a user to the table.

```
SELECT sqlprotect.protect_role('userA');
```

- Query the information about the tables that SQL/Protect learned for the protected roles.

```
select * from sqlprotect.list_protected_users;
select * from sqlprotect.polar_sql_protect;
```

- Invoke the unprotect_role function to remove a protected role.

```
SELECT sqlprotect.unprotect_role('userA');
```

Configure the mode in which SQL/Protect works to monitor a protected role

The polar_sql_protect.level parameter specifies the mode in which SQL/Protect works to monitor a protected role. The following three modes are available: learn, passive, and active. The default mode is passive.

Work mode	Description
learn	SQL/Protect tracks the tables that a user accesses and records the tables. This allows you to record the behavior of protected roles.
passive	If a protected role attempts to execute an unauthorized SQL statement, SQL/Protect sends an alert but does not prevent the SQL statement from being executed.
active	SQL/Protect prevents all unauthorized SQL statements from being executed by protected roles. To prevent the SQL statements from being executed, SQL firewalls take effect when attackers perform penetration tests. SQL/Protect also tracks and queries the SQL statements. This way, administrators can identify database vulnerabilities earlier than attackers.

For example, if you want to change the mode in which SQL/Protect works to active, execute the following statement:

```
polar_sql_protect.level = active;  #Set the mode in which SQL/Protect works to active.
```

To modify some fields in the polar_sql_protect table to specify what need to be protected for a role, execute the following statement:

```
targetdb=# \d sqlprotect.polar_sql_protect;
        Table "sqlprotect.polar_sql_protect"
      Column       |  Type   | Collation | Nullable | Default
-------------------+---------+-----------+----------+---------
dbid               | oid     |           | not null |
roleid             | oid     |           | not null |
protect_relations  | boolean |           |          |
allow_utility_cmds | boolean |           |          |
allow_tautology    | boolean |           |          |
allow_empty_dml    | boolean |           |          |
Indexes:
   "polar_sql_protect_pkey" PRIMARY KEY, btree (roleid)
```

For example, if you execute the following statement to set the allow_utility_cmds parameter to TRUE for a protected role named 16480, SQL/Protect blocks the utility commands run by the protected role 16480.

```
UPDATE sqlprotect.polar_sql_protect SET allow_utility_cmds = TRUE WHERE roleid = 16480;
```

Other operations

- To stop SQL/Protect, execute the following statement:

  ```
  polar_sql_protect.enabled = off #(The default value of this parameter is off.)
  polar_sql_protect.level = passive #(Valid values: learn, active, and passive. The default value is passiv
  e.)
  ```

- To view statistics about the SQL statements blocked by SQL/Protect, execute the following statement:

  ```
  SELECT * FROM sqlprotect.polar_sql_protect_stats;
  ```

- To delete statistics about the SQL statements blocked by SQL/Protect blocks for a specified user, execute the following statement:

  ```
  SELECT sqlprotect.drop_stats('username');
  ```

17.More operations

17.1. Clone a cluster

This topic describes how to create a new cluster by cloning the data of a source cluster.

Scenarios

Before you launch a service, the service is deployed in an environment that simulates real-world scenarios for testing, such as stress testing. To achieve this, you can create a new cluster by cloning the data of a source cluster. Then, you can conduct tests on the new cluster. This ensures the accuracy of the tests without affecting normal business operation.

Considerations

- The following data of the source cluster can be cloned:
  - Cluster account information.
  - The transparent data encryption (TDE) configurations can be cloned if the source cluster has TDE enabled.
- The following data of the source cluster cannot be cloned:
  - Parameter settings
  - Whitelist configurations
  - Secure sockets layer (SSL) configurations
- Only the data that exists in the source cluster before the clone operation starts is cloned.

Procedure

1.
2.
3. Find the cluster that you want to clone and choose **More > Clone Cluster** in the **Actions** column.
4. On the **Clone Instance** page, select a **billing method** for the new cluster.
5. Configure the following parameters.

Parameter	Description
Clone Source Type	By default, **Current Cluster** is selected. For this operation, do not change this setting.
Clone Source Cluster	The ID of the source cluster to clone. This setting cannot be changed.
Region	By default, the region of the new cluster is the same as that of the original cluster. This setting cannot be changed.
Primary Availability Zone	Select the primary zone where the cluster is deployed. **Note** In regions that have two or more zones, automatically replicates data to a secondary zone for disaster recovery.
Network Type	This parameter can only be set to **VPC**.

Parameter	Description
VPC	Select a VPC and a vSwitch for the cluster. We recommend that you use the same **VPC** and **VSwitch** that are used for the original cluster. **Note** Make sure that the cluster and the ECS instance you want to connect to the cluster are deployed in the same VPC. Otherwise, the cluster and the ECS instance cannot communicate over the internal network, which results in decreased performance.
VSwitch	
Compatibility	By default, the new cluster has the same compatibility as that of the source cluster. For example, if the **compatibility** of the source cluster is , , and **Oracle syntax**, the **compatibility** of the new cluster is , , and **Oracle syntax**. You do not need to change this parameter value.
Edition	By default, the edition of the new cluster is the same as that of the source cluster. For example, if the **edition** of the source cluster is , the **edition** of the new cluster is also . You do not need to change this parameter value.
Specification Type	has the following two types of specifications: **General Specification** and **Dedicated Specification**. For more information about the two types of specifications, see Comparison between general-purpose and dedicated compute nodes. **Note** This parameter is available only when the **edition** of the source cluster is . The and **editions** do not support this parameter.
Node Specification	Select a **node specification**. The maximum storage capacity and performance of clusters vary based on node specifications. For more information, see Specifications of compute nodes. **Note** We recommend that you select a **node specification** that is the same or higher than the node specification of the original cluster. This ensures that the new cluster runs as expected.
Nodes	◦ The default number of nodes of the **edition** is **2**. You do not need to change this parameter value. **Note** By default, new clusters contain one primary node and one read-only node. After a cluster is created, you can add nodes to the cluster. A cluster can contain one primary node and up to 15 read-only nodes. For more information about how to add nodes, see . ◦ The default number of nodes of the and **editions** is **1**. You do not need to change this parameter value.
Storage Cost	You do not need to select the storage capacity when you purchase clusters. You are charged for the storage capacity used on an hourly basis. You can also purchase a storage plan based on your business requirements. For more information about how to purchase a storage plan, see Purchase a storage plan.
Cluster Name	The name of the cluster. The name must meet the following requirements: ◦ It cannot start with `http://` or `https://` . ◦ It must be 2 to 256 characters in length. If you do not specify this parameter, the system automatically generates a cluster name. You can change the name after the cluster is created.
Purchase Plan	Specify the **purchase plan** for the cluster. **Note** This parameter is available only when the **Billing Method** parameter is set to **Subscription**.

Parameter	Description
Number	Select the **number** of clusters you want to purchase.

6. Read and accept the terms of service, and complete the rest of the steps based on the **billing method** of the cluster.
   - **Pay-as-you-go**

     Click **Buy Now**.
   - **Subscription**
     a. Click **Buy Now**.
     b. On the **Purchase** page, confirm the information of the unpaid order and the payment method and click **Purchase**.

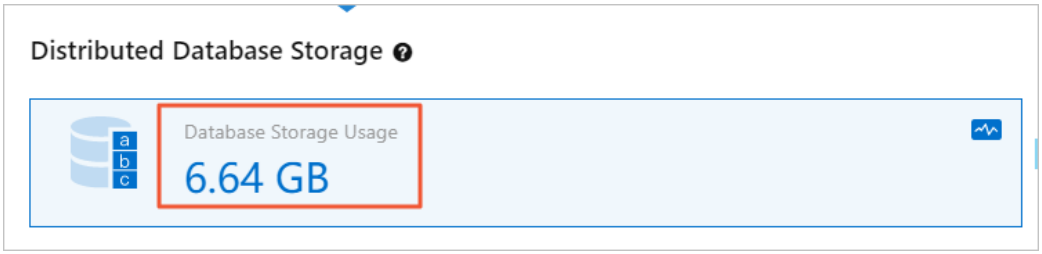
> **Note** After you complete the payment, it requires 10 to 15 minutes to create the cluster. Then, you can view the new cluster on the **Clusters** page.

17.2. View the database storage usage

You can view the database storage usage of a cluster in the console. This topic describes how to view the database storage usage.

Procedure

1.
2.
3.
4. On the **Overview** page, check the value of the **Database Storage Usage** in the **Distributed Database Storage** section.



> **Note** The maximum storage capacity varies based on cluster specifications. If 90% of the maximum storage capacity is used, the system sends SMS messages and emails to notify you on a daily basis. To increase the maximum storage capacity, upgrade your cluster specifications. For more information, see Change the specifications of a PolarDB cluster.

17.3. View or cancel a scheduled task

When you perform operations and management (O&M) tasks, you can customize the execution time of the tasks. For example, you can customize the execution time of tasks for upgrading a cluster, adding nodes, upgrading versions, or changing the primary zone. This topic describes how to view or cancel a scheduled task in the console after you create the task.

Precautions

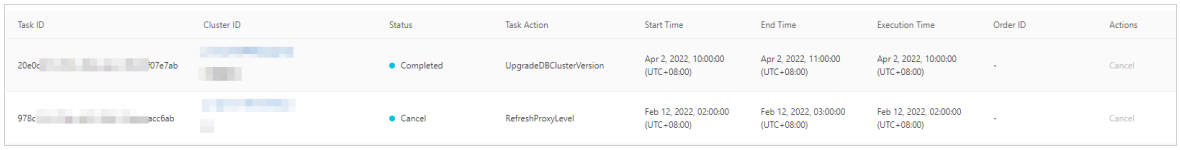
- You can view the details of only the following scheduled tasks:
  - Upgrade a cluster. For more information, see Procedure.
  - Add nodes. For more information, see Add a read-only node.

- Upgrade the version of a cluster. For more information, see Version Management.
- Change the primary zone. For more information, see Deploy a cluster across zones and change the primary zone.

- You can cancel only the tasks whose **Status** is **Pending**. Scheduled tasks for downgrade operations such as node deletion and automatic or manual downgrade cannot be canceled.

View scheduled tasks

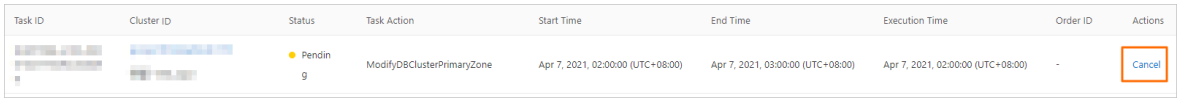
1.
2.
3. In the left-side navigation pane, click **Scheduled Tasks**.
4. On the **Scheduled Tasks** page, you can view the details about all scheduled tasks in the region, such as the **Task ID**, **Status**, **Task Action**, **Start Time**, **End Time**, and **Execution Time**.



> **Note**
> - The API of the task is displayed in the **Task Action** column. The following **Task Action** are supported:
>   - **ModifyDBClusterPrimaryZone**: changes the primary zone.
>   - **ModifyDBNodeClass**: upgrades a cluster.
>   - **CreateDBNodes**: adds nodes.
>   - **UpgradeDBClusterVersion**: upgrades the version of a cluster.
> - You can view the **Order ID** of the cluster only when **Task Action** is **ModifyDBNodeClass** or **CreateDBNodes**.

Cancel a scheduled task

1.
2.
3. In the left-side navigation pane, click **Scheduled Tasks**.
4. On the **Scheduled Tasks** page, find the scheduled task that you want to cancel, and click **Cancel** in the **Actions** column.



> **Note** You can cancel only the tasks whose **Status** is **Pending**. Scheduled tasks for downgrade operations such as node deletion and automatic or manual downgrade cannot be canceled.

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

Related API operations

Operation	Description
DescribeScheduleTasks	Queries the details of all scheduled tasks or a specified scheduled task that belongs to the current account.
CancelScheduleTasks	Cancels a specified scheduled task.