ALIBABA CLOUD

Alibaba Cloud

PolarDB PostgreSQL
Product Introduction

Document Version: 20220119

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.What is PolarDB?

PolarDB is a next-generation relational database service developed by Alibaba Cloud. uses a compute-storage separation architecture and integrates software with hardware. This way, PolarDB can provide users with extremely flexible, high-performance, secure, and reliable database services that allow you to store large amounts of data. PolarDB is fully compatible with MySQL 5.6, MySQL 5.7, MySQL 8.0, and PostgreSQL 11, and is highly compatible with Oracle.

PolarDB uses an architecture in which computing is decoupled from storage. All compute nodes share the same physical storage. PolarDB allows you to upgrade or downgrade instance specifications within minutes and perform failovers within seconds. PolarDB ensures global data consistency. It also provides data backup and disaster recovery services free of charge. PolarDB has the following benefits of commercial databases: stability, reliability, high performance, and scalability. It also has the following benefits of open source cloud databases: ease of use, high compatibility, and self-iteration.

- Computing and storage are decoupled. A distributed storage system is shared by all compute nodes.

  PolarDB decouples computing from storage. This way, you can use auto scaling to meet the dynamic requirements of your business. All compute nodes share the same underlying physical storage PolarStore by using the distributed file system PolarFileSystem. This significantly reduces your storage costs.

- Each cluster consists of one primary node and multiple read-only nodes. This allows PolarDB to provide read/write splitting.

  PolarDB Cluster uses multi-node clusters to provide services. Each PolarDB cluster consists of one primary node and at least one read-only node. You can perform read and write operations on the primary node. PolarDB uses PolarProxy to provide external services for the applications that are connected to cluster endpoints. PolarProxy forwards the requests from the applications to database nodes. You can use PolarProxy to implement authentication, data protection, and automatic read/write splitting. PolarProxy parses SQL statements, and then sends write requests to the primary node and distributes read requests to read-only nodes for load balancing. PolarProxy allows applications to access PolarDB by using the same method that is used to access a single-node database.

Benefits

You can use PolarDB in the same way that you use MySQL, PostgreSQL, or Oracle. Compared with traditional databases, PolarDB has the following advantages:

- **Large storage capacity**

  Each cluster provides a maximum storage capacity of 100 TB. You do not need to purchase instances for database sharding due to the storage limit of a single host. This simplifies the development of applications and reduces the O&M workloads.

- **Cost-effectiveness**
  - Shared storage: PolarDB decouples computing from storage. Each time you add read-only nodes to a PolarDB cluster, you are charged only for the computing resources. However, in a traditional database solution, you are charged for both computing and storage resources.
  - Elastic storage: You do not need to specify the storage capacity. The storage capacity is automatically scaled based on the data volume. You are charged on an hourly basis only for the storage that you use.

  - Storage plan: PolarDB provides subscription storage plans. If you need to store a large amount of data, we recommend that you purchase storage plans. Storage plans are more cost-effective than the pay-as-you-go billing method. You are offered higher discounts for storage plans that have larger storage capacities.

- **High performance**

  PolarDB significantly improves online transaction processing (OLTP) performance and supports more than 500,000 read requests per second and more than 150,000 write requests per second.

- **Elastic scaling within minutes**

  PolarDB supports rapid scaling for computing. This is based on container virtualization, shared distributed storage, and compute-storage separation. It requires only 5 minutes to add or remove a node. The storage capability is automatically scaled up. During the scale-up process, your services are not interrupted.

- **Read consistency**

  PolarDB uses log sequence numbers (LSNs) for cluster endpoints that have read/write splitting enabled. This ensures global consistency for read operations and prevents the inconsistency that is caused by the replication delay between the primary node and read-only nodes.

- **Millisecond-level latency in physical replication**

  PolarDB performs physical replication from the primary node to read-only nodes based on redo logs. The physical replication replaces the logical replication that is based on binary logs. This way, the replication efficiency and stability are improved. No delays occur even if you perform DDL operations on a large table, such as adding indexes or fields.

- **Data backup within seconds**

  No more than 30 seconds are required to perform a full backup on a database regardless of the data volume of the database. During the backup, the database is not locked. This minimizes the impact of data backup on your applications. You can specify a point in time to back up the database data based on your business requirements.

Pricing

For more information, see Specifications and pricing and visit the buy page of PolarDB clusters.

Use PolarDB

You can use the following methods to manage PolarDB clusters. Cluster management includes creating clusters, databases, and accounts.

- Console: PolarDB provides a visualized web interface. This simplifies management operations.
- CLI: You can use Alibaba Cloud CLI to perform all the operations that are available in the console.
- SDK: You can use SDKs to perform all the operations that are available in the console.
- API: You can call API operations to perform all the operations that are available in the console.

After a PolarDB cluster is created, you can use the following methods to connect to the PolarDB cluster:

- Use Data Management (DMS) to connect to a PolarDB cluster and develop databases in the DMS console. For more information, see Connect to a PolarDB for PostgreSQL cluster.
- Use a common database client to connect to a PolarDB cluster. For example, you can use MySQL-Front or pgAdmin to connect to a PolarDB cluster.

Terms

Before you purchase and use PolarDB, we recommend that you familiarize yourself with the following terms:

- Cluster: PolarDB Cluster uses a cluster architecture. A cluster of consists of 1 primary node and a maximum of 15 read-only nodes.
- Region: A region is a geographic area where a data center is deployed. In most cases, PolarDB clusters must be deployed in the same region as Elastic Compute Service (ECS) instances to ensure optimal access performance.
- Zone: A zone is a geographic area in a region. Each zone has an independent power supply and network. All the zones in a region provide the same services.
- Specification: A specification specifies the resources that are configured for each node, such as 2 CPU cores and 8 GB memory.

Related services

- ECS: ECS instances serve as cloud servers. If you connect to a PolarDB cluster from an ECS instance that is deployed in the same region as the PolarDB cluster over an internal network, the PolarDB cluster provides optimal performance. In a typical service architecture, ECS instances are used in combination with PolarDB clusters.
- ApsaraDB for Redis: ApsaraDB for Redis is a database service that supports in-memory storage and persistent storage. In scenarios in which large numbers of access requests exist, you can use ECS, PolarDB, and ApsaraDB for Redis to improve the read IOPS and reduce the response time.
- ApsaraDB for MongoDB: ApsaraDB for MongoDB is a stable, reliable, and scalable database service that is fully compatible with MongoDB protocols. To meet various business requirements, you can store structured data in PolarDB and unstructured data in ApsaraDB for MongoDB.
- Data Transmission Service (DTS): You can use DTS to migrate on-premises databases to PolarDB clusters.
- Object Storage Service (OSS): OSS is a secure, cost-effective, and reliable cloud storage service that allows you to store large amounts of data.

2.Benefits

This topic describes the benefits of PolarDB.

Ease of use

PolarDB is compatible with a variety of popular relational database engines. It is fully compatible with MySQL and PostgreSQL, and is highly compatible with Oracle syntax, with little or no code and application modification.

Cost efficiency

- Separation of computing and storage: Compute nodes share storage resources. You only pay for compute nodes when you add read-only nodes, which greatly reduces scale-out costs.
- Serverless storage: You do not need to manually configure storage space, because the storage space is automatically scaled based on the data volume. You only need to pay for the database capacity that you have used.

High performance

- With an improved database kernel, PolarDB supports physical replication, RDMA protocol, and shared distributed storage, which greatly improves performance.
- An PolarDB cluster contains one primary node and up to 15 read-only nodes. The cluster meets performance requirements in high concurrency scenarios, particularly suitable for scenarios where read requests are far more than write requests.
- An PolarDB cluster shares storage among the primary node and read-only nodes. To apply data changes to nodes in the cluster, you only need to change data once.

Storage capacity for hundreds of terabytes of data

PolarDB uses distributed block storage and a file system to allow automatic scale-up of database storage capacity, regardless of the storage capacity of each node. This enables your database to handle up to hundreds of terabytes of data.

High availability, reliability, and data security

- Supports shared distributed storage to eliminate data inconsistency in the secondary database caused by asynchronous primary-secondary replication. This ensures zero data loss if a single point of failure occurs in a database cluster.
- Supports a multi-zone architecture. Data replicas are available across multiple zones for database disaster recovery and backup.
- Provides various security measures for your database access, storage, and management. These measures include setting an IP whitelist for database access, using VPC for network isolation, and creating multiple replicas for data storage.

Rapid elastic scaling to handle workload spikes

- Configuration upgraded or downgraded within 5 minutes

  PolarDB supports rapid CPU and memory expansion by using container virtualization and shared distributed block storage.

- Nodes added or removed within 5 minutes

PolarDB can dynamically add or remove nodes to help you improve performance and reduce costs. You can use cluster endpoints to mask changes at the underlying layer. In this case, applications are unaware of the addition or removal of nodes.

Lock-free backup

Based on the snapshot technology of underlying distributed storage, PolarDB requires only a few minutes to back up a database with TB-level data. During the entire backup process, no lock is required, which delivers higher efficiency and minimizes the negative impact.

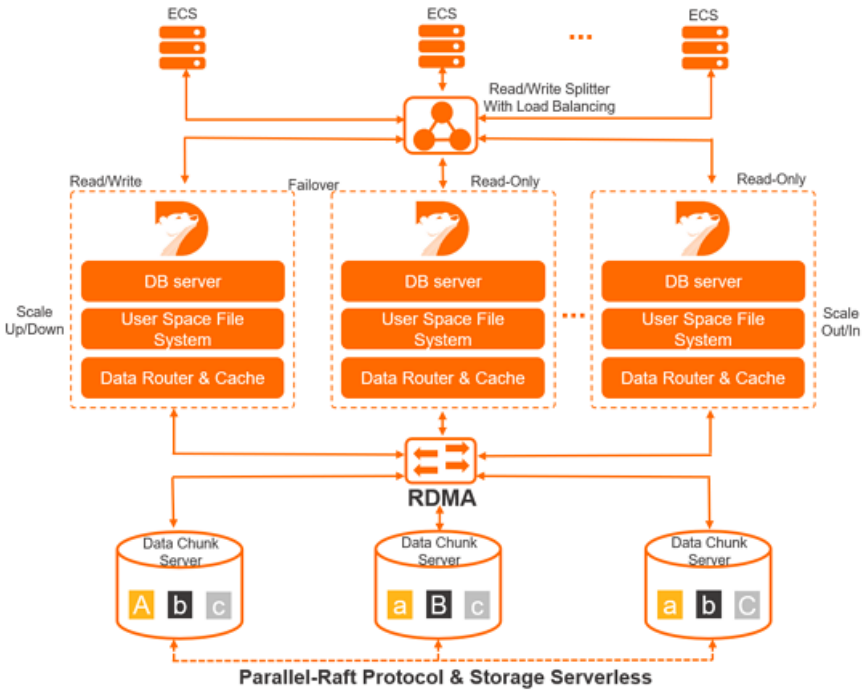
Get started with PolarDB

- PolarDB for MySQL
- PolarDB for PostgreSQL
- PolarDB for Oracle

3.Architecture

PolarDB is based on the cloud native architecture. PolarDB provides the benefits of commercial databases and open source cloud databases. Commercial databases offer the benefits of stability, reliability, high performance, and scalability. Open source cloud databases offer the benefits of simplicity, openness, and rapid iteration. This topic describes the architecture and features of PolarDB.

Product architecture



One primary node and multiple read-only nodes

PolarDB uses a distributed cluster-based architecture. A Cluster cluster consists of a primary node and a maximum of 15 read-only nodes. At least one read-only node is used to ensure high availability. The primary node processes read and write requests and the read-only nodes process only read requests. PolarDB uses the active-active failover method between the primary node and read-only nodes. This method provides the high availability service of databases.

Compute and storage decoupling

PolarDB decouples compute from storage. This allows you to scale clusters that are deployed on Alibaba Cloud based on your business development requirements. The database compute nodes (database engine servers) store only metadata. The remote storage nodes (database storage servers) store data, such as data files and redo logs. You need only to synchronize the metadata that is related to redo logs among your compute nodes. This reduces the replication delay between the primary node and read-only nodes. If the primary node is faulty, a read-only node can serve as the primary node in a short period.

Read/write splitting

By default, read/write splitting is a transparent, high availability, and adaptive load balancing capability that is provided for PolarDB Cluster for free. The read/write splitting feature automatically forwards Structured Query Language (SQL) requests to each node of PolarDB Cluster based on cluster endpoints. This allows you to process concurrent SQL statements in aggregation and high-throughput scenarios. For more information, see Read/write splitting.

High-speed network connections

High-speed network connections are used between compute nodes and storage nodes of databases. The Remote Direct Memory Access (RDMA) protocol is used to transmit data between compute nodes and storage nodes. These two features eliminate the bottlenecks of I/O performance.

Shared distributed storage

Multiple compute nodes share one set of data. Each compute node does not need to store the same set of data. This significantly reduces your storage costs. The storage capacity can be smoothly scaled online by using the new developed distributed storage and distributed file system. This feature is not limited by the storage capacity of a single database server and can cope with hundreds of terabytes of data.

Multiple data replicas and the Parallel-Raft protocol

The data on storage nodes of databases has multiple replicas. This ensures data reliability. In addition, the Parallel-Raft protocol is used to ensure data consistency among these replicas.

4.Cross-node parallel execution

PolarDB for PostgreSQL provides the cross-node parallel execution feature. If this feature is enabled, an SQL query can be run on multiple compute nodes in a distributed manner. This improves the query performance of PolarDB for PostgreSQL, increases the I/O throughput of Polar File System (PolarFS) and the CPU utilization of the compute nodes, and improves the memory usage of the compute nodes.

The cross-node parallel execution feature provides the following benefits:

- Some hybrid transaction/analytical processing (HTAP) capabilities:
  - Analytical queries on transaction processing (TP) data can be run in real time.
  - The read-only nodes on which analytical queries are run are physically isolated from the read-only nodes on which transactional queries are run. This eliminates the impact on TP services.
- Elastic scaling enabled by the compute-storage separation architecture of PolarDB for PostgreSQL:
  - Read-only nodes can be added based on your business requirements to increase computing resources. The added nodes are automatically included in the list of read-only nodes on which parallel execution is performed in a distributed manner. This way, the system does not need to perform resharding on data.
  - No data skew occurs.

For more information about cross-node parallel execution, see Cross-node parallel execution.

5.Glossary

This topic describes the terms that are commonly used in Cloud-native database PolarDB.

Term	Description
Region	The geographic location of a data center where a PolarDB cluster is deployed.
Zone	A physical location with independent power grids and networks within a region. The network latency between instances within the same zone is shorter than that between instances in different zones.
Cluster	PolarDB uses a distributed cluster-based architecture. Each cluster consists of one primary node (read/write node) and multiple read-only nodes. A PolarDB cluster can be deployed across zones but not across regions.
Node	A PolarDB cluster consists of multiple physical nodes. Each cluster consists of two types of nodes. Both node types are equivalent and have the same specifications. These two types of nodes are known as primary nodes and read-only nodes.
Primary Node	Each PolarDB cluster contains one primary node, which is a read/write node.
Read-only Node	You can add up to 15 read-only nodes to a PolarDB cluster.
Cluster Zone	The zone where cluster data is distributed. Cluster data is automatically replicated in two zones for disaster recovery. Node migration can only be performed within these zones.
Primary Zone	The zone where the PolarDB primary node is deployed.
Failover	A read-only node can be specified as the new primary node. For more information, see Automatic failover and manual failover.
Class	The specifications of a PolarDB cluster, which specify the specifications of each node, such as 8-core, 64 GB. For more information, see Specifications and pricing.
Endpoint	An endpoint defines the access point of a PolarDB cluster. Each cluster provides multiple endpoints (access points). Each endpoint can connect to one or more nodes. For example, requests received from a primary endpoint are sent to only the primary node. Cluster endpoints provide the read/write splitting feature. Each cluster endpoint can connect one primary node and multiple read-only nodes. An endpoint contains the attributes of database connections, such as read/write mode, node list, load balancing, and consistency levels.
Address	An address serves as the carrier of an endpoint on different networks. An endpoint may support a VPC-facing address and an Internet-facing address. An address contains network properties, such as the domain name, IP address, VPC, and vSwitch.
Primary Endpoint	The endpoint of the primary node. If a failover occurs, the system automatically connects the primary endpoint to the new primary node.

Term	Description
Cluster Endpoint	A cluster endpoint is a read/write address. Multiple nodes within a cluster use the cluster endpoint to provide services. You can set the cluster endpoint to read-only or read/write mode. Cluster endpoints support features such as auto-scaling, read/write splitting, load balancing, and consistency levels.
Eventual Consistency	By default, eventual consistency is enabled in read-only mode. PolarDB clusters can deliver the best performance based on eventual consistency. For more information, see Eventual consistency.
Session Consistency	Session consistency is also known as causal consistency, which is the default option in read/write mode. It ensures the consistency of read requests across sessions to meet the requirements in most scenarios. For more information, see Session consistency.
Distributed Transaction	A configuration item of cluster endpoints. The transaction splitting feature splits read requests from transactions and forwards these requests to read-only nodes without compromising session consistency. This can reduce the workload on the primary node. For more information, see Features.
Offload Reads from Primary Node	A configuration item of cluster endpoints. SQL statements are sent to read-only nodes while data consistency is achieved. This reduces the loads on the primary node and ensures the stability of the primary node. For more information, see Features.
Private Address	You can use PolarDB with PrivateZone to reserve the connection address (domain name) of your original database. This ensures that each private address of the primary endpoint and cluster endpoint of PolarDB can be associated with a private domain name. This private address takes effect only in the specified VPC within the current region. For more information, see Private domain names.
Snapshot Backup	PolarDB only allows you to back up data only by creating snapshots. For more information, see Overview.
Level-1 Backup	A backup file that is locally stored in the cluster is a level-1 backup. Level-1 backups are stored on distributed storage clusters. These backups allow you to quickly back up and restore the cluster, but cause high costs. For more information, see Overview.
Level-2 Backup	Level-2 backups are backup files stored in local storage media. All data in level-2 backups is archived from level-1 backups and can be permanently stored. Level-2 backups require a longer time to restore the cluster but are cost-effective. For more information, see Overview.
Log Backup	A log backup stores the Redo logs of a database for point-in-time recovery (PITR). Log backups can avoid data loss due to user errors and ensure data security. Log backups must be kept at least 7 days. Log backups are stored in local storage and are cost-effective. For more information, see Overview.

Term	Description
Storage Plan	The resource plan that is provided by PolarDB to reduce the storage costs of clusters. The storage capacity of PolarDB is automatically scaled in or out based on the amount of the stored data. You do not need to manually specify the storage capacity. You are charged for only the used storage. If you want to store a large volume of data, we recommend that you purchase PolarDB storage plans. This reduces the storage costs. For more information, see Purchase a storage plan.

6.Limits

This topic describes the limits of PolarDB for PostgreSQL.

Node type	Maximum number of files
polar.pg.x4.medium	1048576
polar.pg.x4.large	2097152
polar.pg.x4.xlarge	2097152
polar.pg.x8.xlarge	4194304
polar.pg.x8.2xlarge	8388608
polar.pg.x8.4xlarge	12582912
polar.pg.x8.12xlarge	20971520

Maximum number of files: includes user table files, database system table files (approximately 1,000), and log files. An Apsara PolarDB table (non-partition table) occupies three files: data file, visibility map file, and FSM file. Each index indicates a file if indexes are used. The following error message appears when you create a table after the maximum number of files is reached:

```
could not create file
```

In this case, you need to delete some tables or upgrade the specifications of your cluster.

Other limits

Item	Limit
Root privilege of databases	PolarDB for PostgreSQL does not support the superuser privilege. Instead, it supports the polar_superuser privilege as a subset of the superuser privilege.
dblink/fdw	Not supported.