ALIBABA CLOUD

Alibaba Cloud

PolarDB Oracle

Product Introduction

Document Version: 20210114

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

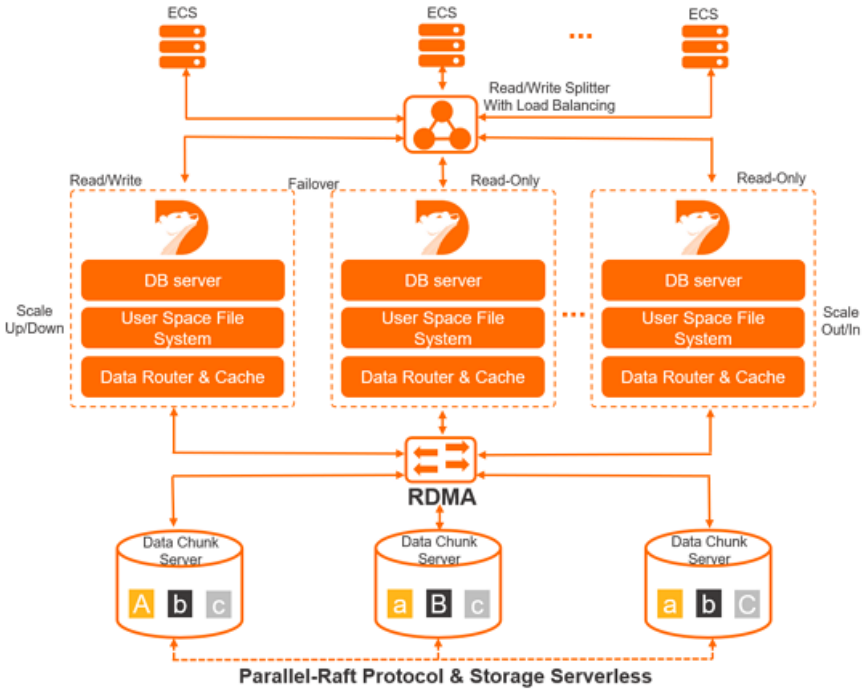
Style	Description	Example
Danger	A danger notice indicates a situation that will cause major system changes, faults, physical injuries, and other adverse results.	**Danger:** Resetting will result in the loss of user configuration data.
Warning	A warning notice indicates a situation that may cause major 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.Architecture

PolarDB uses a cloud native architecture. PolarDB takes advantage of the benefits of commercial databases and open source cloud databases. Commercial databases offer the following benefits: stability, reliability, high performance, and scalability. Open source cloud databases offer the following benefits: simplicity, openness, and rapid iteration. This topic describes the architecture and features of PolarDB.

Architecture



One primary node and multiple read-only nodes

PolarDB uses a distributed cluster-based architecture. Each standard PolarDB cluster consists of a primary node and a maximum of 15 read-only nodes. At least one read-only node must be used to implement failovers to ensure high availability of PolarDB databases. The primary node processes read and write requests and the read-only nodes process only read requests. PolarDB uses an active-active architecture for the primary node and read-only nodes in each cluster. This architecture allows you to implement failovers to ensure high availability of PolarDB databases.

Compute and storage decoupling

PolarDB decouples compute from storage. This allows you to scale clusters that are deployed on Alibaba Cloud to meet your business requirements. Compute nodes store only metadata and remote storage nodes store data files and redo logs. Database engine servers function as compute nodes and database storage servers function as storage nodes. You need only to synchronize the metadata of redo logs among your compute nodes. This reduces the replication delay between the primary node and read-only nodes. If the primary node fails, a read-only node can function as the primary node in a short period.

Read/write splitting

By default, read/write splitting is enabled for PolarDB clusters. The read/write splitting feature is available for free. This feature is transparent to users. This feature provides the capabilities of high availability and self-adaptive load balancing. The read/write splitting feature automatically forwards SQL requests to each node of PolarDB clusters based on cluster endpoints. This allows you to process a large number of concurrent SQL requests in 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 PolarDB databases. The Remote Direct Memory Access (RDMA) protocol is used for data transmission between compute nodes and storage nodes. These two features eliminate the bottlenecks of I/O performance.

Shared distributed storage

In PolarDB, compute nodes share one set of data. This reduces your storage costs. PolarDB uses distributed storage and the distributed file system. This allows you to perform online scale-ups to increase the storage capacity of databases in a smooth manner. The online scale-ups are not affected by the storage capacity of each individual database server. The online scaling allows your databases to process hundreds of terabytes of data.

Multiple data replicas and the Parallel-Raft protocol

Storage nodes of PolarDB databases maintain multiple data replicas to ensure reliability and use the Parallel-Raft protocol to ensure data consistency among these replicas.

2.PolarDB-O terms

This topic introduces terms that are commonly used in PolarDB.

Term	Description
Region	The physical data center where an Apsara PolarDB cluster is deployed.
Zone	Zones are distinct locations within a region that operate on independent power grids and networks. The network latency between instances within the same availability zone is even shorter.
Cluster	Apsara PolarDB runs in a cluster architecture. An Apsara PolarDB cluster contains one writer node (primary node) and multiple reader nodes (read-only nodes). A single Apsara PolarDB cluster can be deployed across zones but not across regions.
Global Database Network (GDN)	It is a network that consists of multiple Apsara PolarDB clusters in different regions across the world. All clusters in the network synchronize with each other to reach data consistency.
Primary cluster	In each GDN, only one cluster is granted the read and write permissions. This read/write cluster is also known as a primary cluster.
Secondary cluster	Clusters to which data from the primary cluster in each GDN is synchronized. These clusters are known as secondary clusters.
Node	An Apsara PolarDB cluster consists of multiple physical nodes. The nodes in each cluster can be divided into two types. Each node type is equivalent and has the same specification. These two types of nodes are known as primary nodes and read-only nodes.
Primary node	Each Apsara 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 an Apsara 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 primary node of an Apsara PolarDB cluster is deployed.
Failover	A read-only node can be promoted to primary node.
Class	Cluster specifications The resources specifications of each node in an Apsara PolarDB, such as 8-core, 64 GB. For more information, see Specifications and pricing.

Term	Description
Endpoint	An endpoint defines the access point of an Apsara 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 only sent to 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 mainly 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, a new primary node is specified through the primary endpoint.
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.
Consistency: eventual consistency	In read-only mode, eventual consistency is enabled by default. Apsara PolarDB clusters can deliver the best performance based on eventual consistency.
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 reads across sessions to meet most application requirements.
Consistency: global consistency	Global consistency is also known as strong consistency, cross-session consistency and highest-level consistency. It ensures the session consistency, but increases the workload on the primary node. We recommend that you do not use global consistency when the replication latency between the primary node and read-only nodes is high.
Transaction splitting	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.
Offload reads from primary node	A configuration item of cluster endpoints. If the session consistency is guaranteed, SQL queries are sent to read-only nodes to reduce the load on the primary node. This ensures the stability of the primary node.
Private address	You can use Apsara 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 Apsara PolarDB can be associated with a private domain name. This private address takes effect only in the specified VPC network within the current region.
Snapshot backup	Apsara PolarDB only allows you to back up data by creating snapshots.

Term	Description
Level -1 backup (snapshot)	A backup file stored locally is a level -1 backup. Level-1 backups are stored on distributed storage clusters. These backups are fast to create and restore, but the costs are high.
Level-2 backup (snapshot)	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 are slow to restore but cost-effective.
Log backup	A log backup stores the Redo logs of a database for point-in-time recovery (PITR). Using log backups can prevent data loss due to user errors. Log backups must be kept at least 7 days. Log backups are cost-effective as they are stored in local storage.

3.Limits

This topic describes the limits of Apsara PolarDB-O.

Node type	Maximum number of files
polar.o.x4.medium	1048576
polar.o.x4.large	2097152
polar.o.x4.xlarge	2097152
polar.o.x8.xlarge	4194304
polar.o.x8.2xlarge	8388608
polar.o.x8.4xlarge	12582912
polar.o.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 PostgreSQL does not support the superuser privilege. Instead, it supports the polar_superuser privilege as a subset of the superuser privilege.
dblink/fdw	Currently, this operation is not supported in Scala.