

ALIBABA CLOUD

阿里云

PolarDB PostgreSQL

产品简介

文档版本：20220321

 阿里云

法律声明

阿里云提醒您在阅读或使用本文档之前仔细阅读、充分理解本法律声明各条款的内容。如果您阅读或使用本文档，您的阅读或使用行为将被视为对本声明全部内容的认可。

1. 您应当通过阿里云网站或阿里云提供的其他授权通道下载、获取本文档，且仅能用于自身的合法合规的业务活动。本文档的内容视为阿里云的保密信息，您应当严格遵守保密义务；未经阿里云事先书面同意，您不得向任何第三方披露本手册内容或提供给任何第三方使用。
2. 未经阿里云事先书面许可，任何单位、公司或个人不得擅自摘抄、翻译、复制本文档内容的部分或全部，不得以任何方式或途径进行传播和宣传。
3. 由于产品版本升级、调整或其他原因，本文档内容有可能变更。阿里云保留在没有任何通知或者提示下对本文档的内容进行修改的权利，并在阿里云授权通道中不时发布更新后的用户文档。您应当实时关注用户文档的版本变更并通过阿里云授权渠道下载、获取最新版的用户文档。
4. 本文档仅作为用户使用阿里云产品及服务的参考性指引，阿里云以产品及服务的“现状”、“有缺陷”和“当前功能”的状态提供本文档。阿里云在现有技术的基础上尽最大努力提供相应的介绍及操作指引，但阿里云在此明确声明对本文档内容的准确性、完整性、适用性、可靠性等不作任何明示或暗示的保证。任何单位、公司或个人因为下载、使用或信赖本文档而发生任何差错或经济损失的，阿里云不承担任何法律责任。在任何情况下，阿里云均不对任何间接性、后果性、惩戒性、偶然性、特殊性或刑罚性的损害，包括用户使用或信赖本文档而遭受的利润损失，承担责任（即使阿里云已被告知该等损失的可能性）。
5. 阿里云网站上所有内容，包括但不限于著作、产品、图片、档案、资讯、资料、网站架构、网站画面的安排、网页设计，均由阿里云和/或其关联公司依法拥有其知识产权，包括但不限于商标权、专利权、著作权、商业秘密等。非经阿里云和/或其关联公司书面同意，任何人不得擅自使用、修改、复制、公开传播、改变、散布、发行或公开发表阿里云网站、产品程序或内容。此外，未经阿里云事先书面同意，任何人不得为了任何营销、广告、促销或其他目的使用、公布或复制阿里云的名称（包括但不限于单独为或以组合形式包含“阿里云”、“Aliyun”、“万网”等阿里云和/或其关联公司品牌，上述品牌的附属标志及图案或任何类似公司名称、商号、商标、产品或服务名称、域名、图案标示、标志、标识或通过特定描述使第三方能够识别阿里云和/或其关联公司）。
6. 如若发现本文档存在任何错误，请与阿里云取得直接联系。

通用约定

格式	说明	样例
 危险	该类警示信息将导致系统重大变更甚至故障，或者导致人身伤害等结果。	 危险 重置操作将丢失用户配置数据。
 警告	该类警示信息可能会导致系统重大变更甚至故障，或者导致人身伤害等结果。	 警告 重启操作将导致业务中断，恢复业务时间约十分钟。
 注意	用于警示信息、补充说明等，是用户必须了解的内容。	 注意 权重设置为0，该服务器不会再接受新请求。
 说明	用于补充说明、最佳实践、窍门等，不是用户必须了解的内容。	 说明 您也可以通过按Ctrl+A选中全部文件。
>	多级菜单递进。	单击设置> 网络> 设置网络类型。
粗体	表示按键、菜单、页面名称等UI元素。	在结果确认页面，单击确定。
Courier字体	命令或代码。	执行 <code>cd /d C:/window</code> 命令，进入Windows系统文件夹。
斜体	表示参数、变量。	<code>bae log list --instanceid</code> <code>Instance_ID</code>
[] 或者 [a b]	表示可选项，至多选择一个。	<code>ipconfig [-all -t]</code>
{ } 或者 {a b}	表示必选项，至多选择一个。	<code>switch {active stand}</code>

目录

1.什么是PolarDB	05
2.产品优势	07
3.产品架构	08
4.跨机并行查询	10
5.术语	11
6.使用限制	14

1.什么是PolarDB

PolarDB是阿里巴巴自研的新一代云原生关系型数据库，在存储计算分离架构下，利用了软硬件结合的优势，为用户提供具备极致弹性、高性能、海量存储、安全可靠的数据库服务。PolarDB 100%兼容MySQL 5.6/5.7/8.0，PostgreSQL 11，高度兼容Oracle。

PolarDB采用存储和计算分离的架构，所有计算节点共享一份数据，提供分钟级的配置升降级、秒级的故障恢复、全局数据一致性和免费的数据备份容灾服务。PolarDB既融合了商业数据库稳定可靠、高性能、可扩展的特征，又具有开源云数据库简单开放、自我迭代的优势。

- 计算与存储分离，共享分布式存储。

采用计算与存储分离的设计理念，满足业务弹性扩展的需求。各计算节点通过分布式文件系统（PolarFileSystem）共享底层的存储（PolarStore），极大降低了用户的存储成本。

- 一写多读，读写分离。

PolarDB集群版采用多节点集群的架构，集群中有一个主节点（可读可写）和至少一个只读节点。当应用程序使用集群地址时，PolarDB通过内部的代理层（PolarProxy）对外提供服务，应用程序的请求都先经过代理，然后才访问到数据库节点。代理层不仅可以做安全认证和保护，还可以解析SQL，把写操作发送到主节点，把读操作均衡地分发到多个只读节点，实现自动的读写分离。对于应用程序来说，就像使用一个单点的数据库一样简单。

产品优势

您可以像使用MySQL、PostgreSQL、Oracle一样使用PolarDB。此外，PolarDB还有传统数据库不具备的优势：

- 大容量

最高100 TB，您不再需要因为单机容量的天花板而去购买多个实例做分片，由此简化应用开发，降低运维负担。

- 低成本

- 共享存储：计算与存储分离，每增加一个只读节点只收取计算资源费用，而传统的只读节点同时包含计算和存储资源，每增加一个只读节点需要支付相应的存储费用。
- 弹性存储：存储空间无需配置，根据数据量自动伸缩，您只需为实际使用的数据量按小时付费。
- **存储包**：PolarDB推出了预付费形式的存储包。当您的数据量较大时，推荐您使用存储包，相比按小时付费，预付费购买存储包有折扣，购买的容量越大，折扣力度越大。

- 高性能

大幅提升OLTP性能，支持超过50万次/秒的读请求以及超过15万次/秒的写请求。

- 分钟级扩缩容

存储与计算分离的架构，配合容器虚拟化技术和共享存储，增减节点只需5分钟。存储容量自动在线扩容，无需中断业务。

- 读一致性

集群地址利用LSN（Log Sequence Number）确保读取数据时的全局一致性，避免因主备延迟引起的不一致。

- 毫秒级延迟（物理复制）

利用基于Redo的物理复制代替基于Binlog的逻辑复制，提升主备复制的效率和稳定性。即使对大表进行加索引、加字段等DDL操作，也不会造成数据库的延迟。

- **秒级快速备份**

不论多大的数据量，全库备份只需30秒，而且备份过程不会对数据库加锁，对应用程序几乎无影响，全天24小时均可进行备份。

PolarDB定价

详情请参见[规格与定价](#)和[购买PolarDB集群](#)。

如何使用PolarDB

您可以通过以下方式管理PolarDB集群，包括创建集群、创建数据库、创建账号等。

- **控制台**：提供图形化的Web界面，操作方便。
- **CLI**：控制台上所有的操作都可以通过CLI实现。
- **SDK**：控制台上所有的操作都可以通过SDK实现。
- **API**：控制台上所有的操作都可以通过API实现。

创建PolarDB集群后，您可以通过以下方式连接PolarDB集群：

- **DMS**：您可以通过[DMS连接PolarDB集群](#)，在Web界面进行数据库开发工作。
- **客户端**：您可以使用通用的数据库客户端工具连接PolarDB集群。例如，MySQL-Front、pgAdmin等。

相关概念

了解以下概念，将帮助您更好地选购和使用PolarDB：

- **集群**：PolarDB集群版采用集群架构，一个集群中可包含一个主节点和最多15个只读节点。
- **地域**：地域是指物理的数据中心。一般情况下，PolarDB集群应该和ECS实例位于同一地域，以实现最高的访问性能。
- **可用区**：可用区是指在某个地域内拥有独立电力和网络的物理区域。同一地域的不同可用区之间没有实质性区别。
- **规格**：每个节点的资源配置，例如2核8 GB。

相关服务

- **ECS**：ECS是云服务器，通过内网访问同一地域的PolarDB集群时，可实现PolarDB集群的最佳性能。ECS搭配PolarDB集群是典型的业务访问架构。
- **Redis**：Redis提供持久化的内存数据库服务。当业务访问量较大时，ECS、PolarDB和Redis的组合可以支持更多的读请求，同时减少响应时间。
- **MongoDB**：提供稳定可靠、弹性伸缩、完全兼容MongoDB协议的数据库服务。数据结构多样时，可以选择将结构化数据存储存储在PolarDB，将非结构化数据存储存储在MongoDB，满足业务的多样化存储需求。
- **DTS**：您可以使用数据传输服务DTS将本地数据库迁移到云上的PolarDB。
- **OSS**：对象存储服务OSS是阿里云提供的海量、安全、低成本、高可靠的云存储服务。

2. 产品优势

本文介绍云原生关系型数据库PolarDB的产品优势，帮助您更好地了解PolarDB。

简单易用

PolarDB兼容多款流行的关系型数据库引擎，完全兼容MySQL和PostgreSQL，高度兼容Oracle语法，代码/应用无需修改或只需少量修改。

降低成本

- 计算节点和存储分离：多个计算节点共享存储，新增只读节点时只需支付计算节点费用，大大降低扩容成本。
- Serverless存储：存储空间无需手动配置，根据数据量自动伸缩，您只需为实际使用的数据库容量付费。

极致性能

- 深度优化数据库内核，同时采用物理复制、RDMA高速网络和分布式共享存储，大幅提高性能。
- 集群包含一个主节点和最多15个只读节点，满足高并发场景对性能的要求，尤其适用于读多写少的场景。
- 基于共享存储的一写多读集群，数据只需要一次修改，所有节点立即生效。

海量存储，支持上百TB级别数据

采用分布式块存储设计和文件系统，使得存储容量不限制于单节点的规格，能够轻松扩展，应对上百TB级别的数据规模。

高可用和高可靠保障，数据安全可靠

- 共享分布式存储的设计，彻底解决了主从（Master-Slave）异步复制所带来的备库数据非强一致的缺陷，使得整个数据库集群在应对任何单点故障时，可以保证数据零丢失。
- 多可用区架构，在多个可用区内都有数据备份，为数据库提供容灾和备份。
- 采用白名单、VPC网络、数据多副本存储等全方位的手段，对数据库数据访问、存储、管理等各个环节提供安全保障。

快速弹性，应对不确定的业务增长

- 配置升降级，5分钟生效。

采用容器虚拟化技术和共享的分布式块存储技术，使得数据库服务器的CPU、内存能够快速扩容。

- 增减节点，5分钟生效。

通过动态增减节点提升性能或节省成本。通过使用集群地址，可屏蔽底层的变化，应用对于增减节点无感知。

无锁备份

利用底层分布式存储的快照技术，只需分钟级别即可完成对上TB数据量大小的数据库的备份，且整个备份过程不需要加锁，效率更高，影响更小。

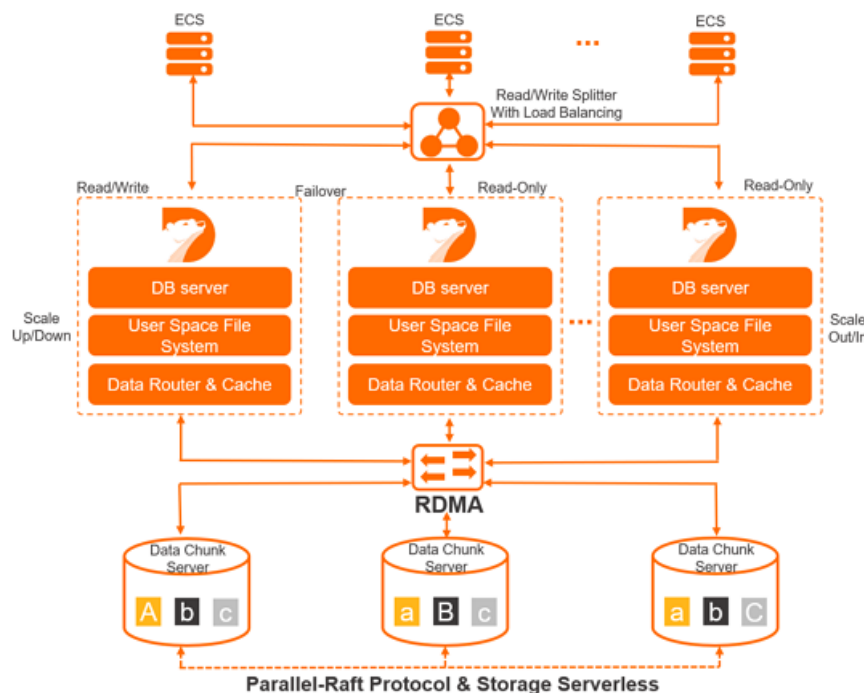
开始使用PolarDB

- [PolarDB MySQL引擎快速入门](#)
- [PolarDB PostgreSQL引擎快速入门](#)
- [PolarDB O引擎快速入门](#)

3. 产品架构

云原生关系型数据库PolarDB基于Cloud Native设计理念，既融合了商业数据库稳定可靠、高性能、可扩展的特征，又具有开源云数据库简单开放、快速迭代的优势。本文将介绍PolarDB的产品架构及特点。

产品架构图



一写多读

PolarDB采用分布式集群架构，一个集群版集群包含一个主节点和最多15个只读节点（至少一个，用于保障高可用）。主节点处理读写请求，只读节点仅处理读请求。主节点和只读节点之间采用Active-Active的Failover方式，提供数据库的高可用服务。

计算与存储分离

PolarDB采用计算与存储分离的设计理念，满足公共云计算环境下根据业务发展弹性扩展集群的刚性需求。数据库的计算节点（Database Engine Server）仅存储元数据，而将数据文件、Redo Log等存储于远端的存储节点（Database Storage Server）。各计算节点之间仅需同步Redo Log相关的元数据信息，极大降低了主节点和只读节点间的复制延迟，而且在主节点故障时，只读节点可以快速切换为主节点。

读写分离

读写分离是PolarDB集群版默认免费提供的一个透明、高可用、自适应的负载均衡能力。通过集群地址，SQL请求自动转发到PolarDB集群版的各个节点，提供聚合、高吞吐的并发SQL处理能力。请参见[读写分离](#)。

高速链路互联

数据库的计算节点和存储节点之间采用高速网络互联，并通过RDMA协议进行数据传输，使I/O性能不再成为瓶颈。

共享分布式存储

多个计算节点共享一份数据，而不是每个计算节点都存储一份数据，极大降低了用户的存储成本。基于全新打造的分布式块存储（Distributed Storage）和文件系统（Distributed Filesystem），存储容量可以在线平滑扩展，不会受到单个数据库服务器的存储容量限制，可应对上百TB级别的数据规模。

数据多副本、Parallel-Raft协议

数据库存储节点的数据采用多副本形式，确保数据的可靠性，并通过Parallel-Raft协议保证数据的一致性。

4.跨机并行查询

PolarDB PostgreSQL引擎提供了跨机并行查询（Parallel Execution）的功能，支持多个计算节点分布式地执行SQL查询，加速PolarDB PostgreSQL引擎的分析型查询性能，充分发挥存储层PolarFileSystem的高I/O吞吐能力，以及提高所有计算节点的CPU和内存资源的使用率。

跨机并行查询功能具有如下优势：

- 具备一定的HTAP能力：
 - 能在TP数据上实时执行分析型查询。
 - 执行分析查询的只读节点和执行TP型查询的只读节点可以物理隔离，避免影响TP业务。
- 结合PolarDB PostgreSQL引擎存储计算分离的架构，可以做到弹性扩展：
 - 当算力不够时，可以弹性地增加只读节点，新增的只读节点加入分布式的并行计算中，而不需数据重新分片（Reshard）。
 - 不会出现数据倾斜问题。

关于更多跨机并行查询的介绍和功能说明，请参见[跨机并行查询](#)。

5. 术语

本文介绍云原生关系型数据库PolarDB使用过程中遇到的术语。

名词	描述
地域 (Region)	数据中心所在的地理位置。
可用区 (Zone)	可用区是指在某一地域内，具有独立电力和网络的物理区域。同一可用区内实例之间的网络延时更小。
集群 (Cluster)	PolarDB采用多节点集群的架构，集群中有一个主节点和多个只读节点。单个PolarDB集群支持跨可用区，但不能跨地域。
节点 (Node)	PolarDB集群由多个物理节点构成，每个集群中的节点可分为两类，每一类节点关系对等，规格相同。这两类节点分别叫主节点和只读节点。
主节点 (Primary node)	PolarDB主节点，也叫读写节点，一个集群中只有一个主节点。
只读节点 (Read-only node)	PolarDB只读节点，一个集群中最多可添加至15个只读节点。
集群可用区 (Cluster zone)	集群数据分布的可用区。集群的数据会自动在两个可用区间实现冗余，用于灾备恢复，节点迁移只支持在这两个可用区间进行。
主可用区 (Primary zone)	PolarDB的主节点所在可用区。
故障切换 (Failover)	故障切换（也称主备切换）是指提升一个只读节点为主节点。更多详情，请参见 自动/手动主备切换 。
规格 (Class)	集群规格。PolarDB每个节点的资源配置，例如8核64 GB，更多规格请参见 规格与定价 。
访问点 (Endpoint)	访问点定义了数据库的访问入口，访问点也称为接入点。每个集群都提供了多个访问点，每个访问点可以连接1个或多个节点。例如，主访问点永远指向主节点，集群访问点提供了读写分离能力，连接了主节点和多个只读节点。访问点中包含的主要是数据库链路属性，例如读写状态、节点列表、负载均衡、一致性级别等。

名词	描述
访问地址 (Address)	访问地址是访问点在不同网络平面中的载体，一个访问点可能包含私网和公网两种访问地址。访问地址中包含了一些网络属性，例如域（Domain）、IP地址、专有网络（VPC）、交换机（VSwitch）等。
主地址 (Primary Endpoint)	主节点的访问点，当发生故障切换（Failover）后，系统会将访问点自动指向新的主节点。
集群地址 (Cluster Endpoint)	整合集群下的多个节点，对外提供一个统一的读写地址，可以设置为只读或读写。集群地址具有自动弹性、读写分离、负载均衡、一致性协调等能力。更多详情，请参见 集群地址和主地址 。
最终一致性 (Eventual Consistency)	只读模式下默认选项为最终一致性。最终一致性下PolarDB集群将提供最优的性能。更多详情，请参见 最终一致性 。
会话一致性 (Session Consistency)	会话一致性也叫因果一致性，读写模式下的默认选项，提供Session级的读一致性保证，可以满足大部分应用场景。更多详情，请参见 会话一致性 。
事务拆分 (Distributed Transaction)	集群地址的一个配置项。在保证一致性的前提下，通过拆分一个事务（Transaction）内的读请求（Select Queries）到只读节点，可以在一定程度上降低主节点的负载。更多详情，请参见 功能特性 。
主库不接受读 (Offload Reads from Primary Node)	集群地址的一个配置项。在确保一致性的前提下，将查询SQL发送到只读节点，来降低主节点的负载，确保主节点稳定。更多详情，请参见 功能特性 。
私有域名 (Private Address)	为了保留用户原来数据库的连接地址（域名），PolarDB联手PrivateZone，保证PolarDB主地址和集群地址中的每一个内网地址，均可以绑定一个私有域名。该私有域名仅在当前地域内指定的VPC中生效。更多详情，请参见 私有域名 。
快照备份 (Snapshot Backup)	PolarDB数据的备份方式，目前仅支持快照备份。更多详情，请参见 备份数据 。
一级备份 (Level-1 Backup)	保存在本地的备份文件叫一级备份。一级备份直接存储在分布式存储集群中，备份和恢复速度最快，但成本高。更多详情，请参见 备份数据 。
二级备份 (Level-2 Backup)	保存在其他离线存储介质中的备份文件叫二级备份。二级备份的数据全部来自于一级备份，可以永久保存，成本低但恢复速度较慢。更多详情，请参见 备份数据 。
日志备份 (Log Backup)	日志备份是把数据库的Redo日志保存下来，用于按时间点恢复，为了保证最近一段时间数据的安全性，避免误操作导致的数据丢失，因此日志备份最少保留7天。日志备份采用离线存储备份，成本较低。更多详情，请参见 备份数据 。

名词	描述
存储包 (Storage Plan)	存储包PolarDB推出的一款降低集群存储成本的资源包。PolarDB的存储空间可根据数据量自动伸缩，无需您手动配置，您只需为实际使用的数据量付费。当您的数据量较大时，推荐使用PolarDB存储包以降低存储成本。更多详情，请参见 购买存储包 。

6.使用限制

本文为您介绍PolarDB PostgreSQL引擎的相关使用限制。

节点规格	文件个数上限
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

文件个数上限：包括用户表文件、数据库系统表文件（大约1000个）、日志文件等。一个普通的PolarDB表（非分区表）占用3个文件（数据文件、visibilitymap文件、fsm文件，使用索引的话，每个索引一个文件）。文件数量达到上限后，继续创建表会提示类似以下错误：

```
could not create file
```

此时需要删除部分表或者升级集群规格。

其他限制

操作	使用约束
数据库的root权限	不提供superuser权限，但提供polar_superuser权限，为superuser权限的子集。
dblink/fdw	不支持。