# 第九章：云原生数据库应用实践：PolarDB

## 9.1 创建云上实例

自建数据库涉及购买服务器、安装和部署数据库软件等，需要投入大量的人力物力。对于云数据库，只需设置若干参数，提交后即可在数分钟内完成数据库实例的创建。

**相关概念**

* 实例：实例是虚拟化的数据库服务器。用户可以在一个实例中创建和管理多个数据库。
* 系列：创建PolarDB实例时，可按需选择适合业务的系列（例如集群版或单节点版）。
* 集群：PolarDB主要采用集群版的架构，集群中有一个主节点和多个只读节点。
* 规格：每个节点的资源配置，例如2核8GB。
* 地域：地域是指物理的数据中心。一般情况下，PolarDB实例应该和ECS实例位于同一地域，以实现最高的访问性能。

注：ECS是阿里云提供的云服务器，ECS搭配云数据库是典型的业务访问架构

* 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。
* 数据库引擎：PolarDB有三个独立的引擎，分别100%兼容MySQL、100%兼容PostgreSQL、高度兼容Oracle语法。

**前提条件**

1.注册阿里云账号或获取由主账号管理员分配的子账号。

2.登陆PolarDB控制台([https://polardb.console.aliyun.com](https://polardb.console.aliyun.com/)）。

3.进入PolarDB购买页

**选择付费方式**

* **包年包月**：在创建集群时支付计算节点的费用，而存储空间会根据实际数据量按小时计费，并从账户中按小时扣除。
* **按量付费**：无需预先支付费用，计算节点和存储空间（根据实际数据量）均按小时计费，并从账户中按小时扣除。

PolarDB已推出计算包。推荐购买按量付费的集群，因为按量付费的集群支持配合计算包使用，比包年包月付费方式更划算和灵活。

**选择地域和可用区**

集群所在的地理位置。购买后无法更换地域。

**说明** 请确保PolarDB与需要连接的ECS创建于同一个地域，否则它们无法通过内网互通，只能通过外网互通，无法发挥最佳性能。

**选择创建方式**

按需选择一种创建方式：

* 创建一个全新的PolarDB实例。
* 如果已有RDS MySQL，可以从RDS MySQL一键升级到PolarDB MySQL。
* 如果回收站里有已删除集群的备份，可以通过恢复该备份来创建新集群。

**选择网络类型**

固定为VPC专有网络，无需选择。

请确保PolarDB与需要连接的ECS创建于同一个VPC，否则它们无法通过内网互通，无 法发挥最佳性能。

**选择系列**

* 集群版：是推荐的主流系列，免费提供[快速备份数据](https://help.aliyun.com/document_detail/72672.html#task-1580301)、[恢复数据](https://help.aliyun.com/document_detail/87897.html#task-1580301)和[全球数据库部署](https://help.aliyun.com/document_detail/160381.html#task-2449685)功能，同时还支持快速[弹性升降级](https://help.aliyun.com/document_detail/72674.html#task-1580301)、并行[查询加速](https://help.aliyun.com/document_detail/128615.html#concept-1563422)等企业级功能，推荐在生产环境中使用。
* 单节点版：是个人用户测试、学习的最佳选择，也可作为初创企业的入门级产品。
* 历史库版：定位为归档数据库，具有较高的数据压缩率，非常适合对计算诉求不高但需要存储一些归档类数据（如钉钉消息等数据）的业务。

**选择计算节点规格**

按需选择计算节点的规格，所有节点均为独享型，性能稳定可靠。

每种规格有相应的CPU和内存、最大存储容量、最大连接数、内网带宽、最大IOPS等。

**关于存储空间**

PolarDB采用计算与存储分离的架构，而且存储容量随数据量的增减而自动弹性伸缩，因此创建集群时无需选择存储容量。

注：

* + - 系统会根据实际数据使用量按小时计费。
    - 最大存储容量取决于选择的计算节点规格。

**完成创建**

支付成功后，需要10~15分钟创建集群，之后就可以在集群列表中看到新创建的集群。

请确认已选中正确的地域，否则无法看到创建的集群。

**相关资源**

* PolarDB官方帮助文档：<https://help.aliyun.com/product/58609.html>
* 也可以通过调用API的方式创建PolarDB实例，在阿里云官网搜索[CreateDBCluster](https://help.aliyun.com/document_detail/98169.html?spm=a2c4g.11186623.2.40.9eeb6c79nubyjp&doc-api-polardb-CreateDBCluster)可以查看具体说明。

## 9.2 数据库接入

### 9.2.1 相关账号的创建

**高权限账号与普通账号**

|  |  |
| --- | --- |
| 账号类型 | 说明 |
| 高权限账号 | * 只能通过控制创建和管理。 * 一个集群只能有一个高权限账号，可以管理所有普通账号和数据库。 * 开放了更多权限，可满足个性化和精细化的权限管理需求，例如可为不同用户分配不同表的查询权限。 * 拥有集群中所有数据库的所有权限。 * 可以断开任意账号的连接。 |
| 普通账号 | * 可以通过控制台或者SQL语句创建和管理。 * 一个集群可以创建多个普通账号，具体的数量与数据库内核有关 。 * 需要手动给普通账号授予特定数据库的权限。 * 普通账号不能创建和管理其他账号，也不能断开其他账号的连接。 |

**如何创建**

1. 登录PolarDB控制台。

2. 在控制台左上脚，选择集群所在地域。

3. 找到目标集群，单击集群ID。

4. 在左侧导航栏中，进入账号管理页面。

5. 单击创建账号。

6. 选择要创建的账号类型，并填写账号的密码。

* + 如果已经创建过高权限账号，则无法再选择高权限账号，因为每个集群只能有一个高权限账号。
  + 对于高权限账号，无需选择授权的数据库，因为高权限账号拥有集群中所有数据库的所有权限。
  + 对于普通账号，需要选择授权的数据库。

### 9.2.2 图形化访问

DMS是阿里云提供的图形化的数据管理工具，可用于管理关系型数据库和NoSQL数据库，支持数据管理、结构管理、用户授权、安全审计、数据趋势、数据追踪、BI图表、性能与优化等功能。

1. 找到目标集群，单击集群ID，进入基本信息页面。

2. 单击页面右上角的登录数据库。

3. 在登陆数据库的对话框中，输入数据库的账号和密码。

* + 对于PolarDB PostgreSQL和兼容Oracle版本，还需填写要登陆的数据库。在PolarDB集群的数据库管理页可以创建数据库。
  + 如果是第一次通过DMS访问该实例，系统会弹出设置白名单的对话框，确认添加即可。

### 9.2.3 连接方式访问

#### 9.2.3.1 设置白名单

创建PolarDB数据库集群后，需要设置集群的IP白名单，并创建集群的初始账号，才能连接和使用该集群。

**白名单的两种形式**

* IP白名单

在IP白名单中添加IP地址，允许这些IP地址访问该集群。 默认的IP白名单只包含默认IP地址127.0.0.1，表示任何设备均无法访问该集群。只有已添加到IP白名单中的IP地址才可以访问该集群。

* ECS安全组

ECS安全组是一种虚拟防火墙，用于控制安全组中的ECS实例的出入流量。在PolarDB集群白名单中 添加ECS安全组后，该安全组中的ECS实例就可以访问PolarDB集群。

可以同时设置IP白名单和ECS安全组。IP白名单中的IP地址和安全组中的ECS实例都可以 访问该PolarDB集群。

**设置IP白名单**

1. 登录PolarDB控制台。

2. 在控制台左上角，选择集群所在地域。

3. 找到目标集群，单击集群ID。

4. 在左侧导航栏，进入白名单设置页面。

5. 在白名单设置页面，新增IP白名单分组，或者打开已有的分组。

ali\_dms\_group（DMS产品IP地址白名单分组）、hdm\_security\_ips（DAS产品IP地址白名 单分组）、dtspolardb（DTS产品IP地址白名单分组）等分组为使用相关产品时系统自动 生成。请勿修改或删除分组，避免影响相关产品的使用。

6. 在白名单分组里，添加需要访问PolarDB的设备的IP地址。

* 如果ECS实例需要访问PolarDB，可在ECS实例详情页面的配置信息区域，查看ECS实例的IP地址，然后填写到白名单中。

**说明** 如果ECS与PolarDB位于同一地域（例如，华东1），填写ECS的私网IP地址；如 果ECS与PolarDB位于不同的地域，填写ECS的公网IP地址，或者将ECS迁移到PolarDB 所在地域后填写ECS私网IP地址。

* 如果本地的服务器、电脑或其它云服务器需要访问PolarDB，请将其IP地址添加到白名单中。

**设置安全组**

7. 登录PolarDB控制台。

8. 在控制台左上角，选择集群所在地域。

9. 找到目标集群，单击集群ID。

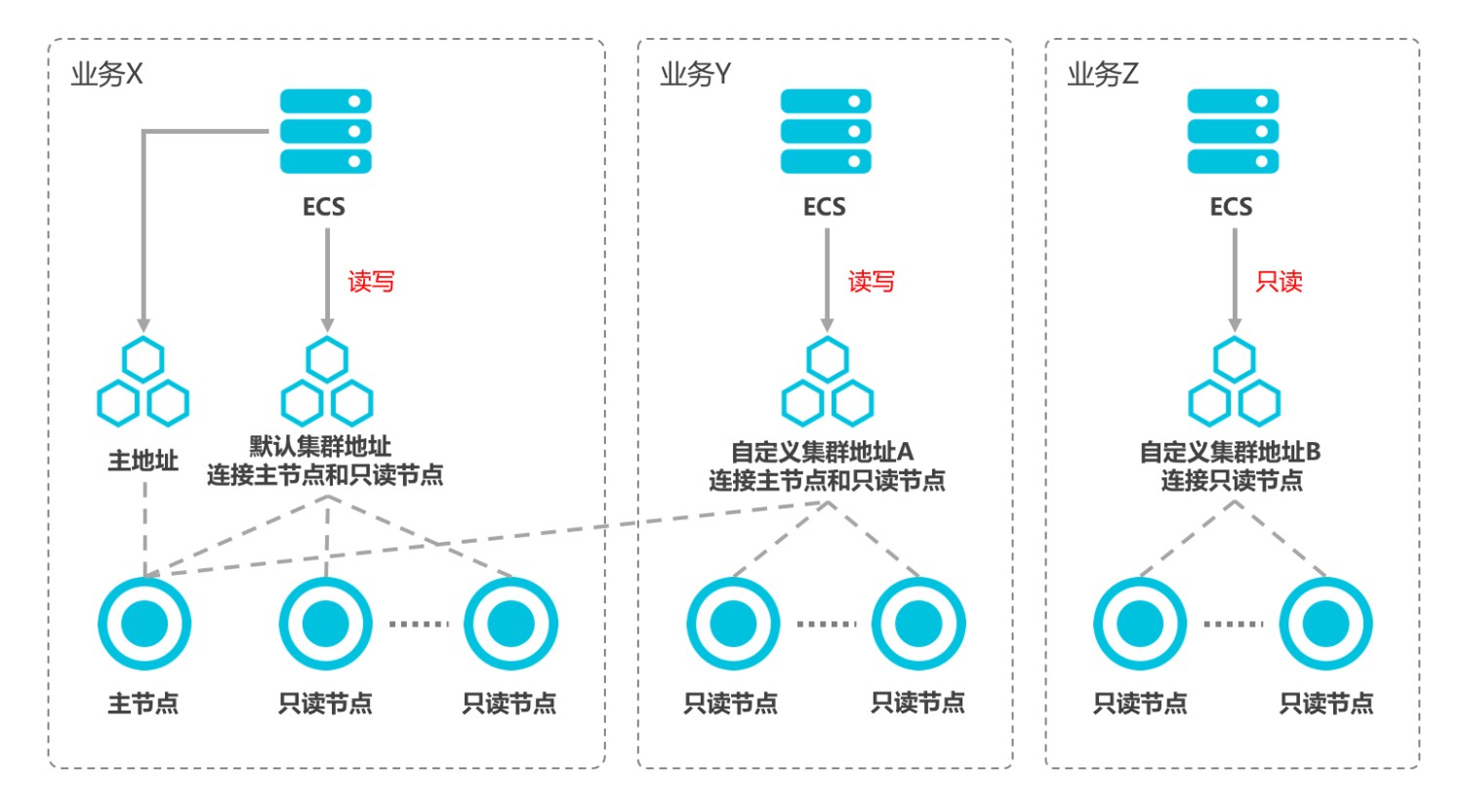
10. 在左侧导航栏，进入白名单设置页面。

11. 打开选择安全组的对话框，选中目标安全组，单击确定即可。

#### 9.2.3.2 获取连接地址

PolarDB集群的连接地址包括集群地址和主地址。

**集群地址和主地址**



|  |  |  |
| --- | --- | --- |
| 地址类型 | 地址说明 | 支持的⽹络类型 |
| 集群地址（推荐） | * 应用程序只需连接一个集群地址，即可连接到多个节点。 * 带有读写分离功能，写请求会自动发往主节点，读请求会自动根据各节点的负载发往主节点或只读节点。   **说明** PolarDB包含一个默认的集群地址，还支持根据业务需求  创建多个自定义的集群地址，自定义集群地址可以连接到指定的节点，以及设置读写模式等。 | * 私网 * 公网 |
| 主地址 | * 总是连接到主节点，支持读和写操作。 * 当主节点发生故障时，主访问地址会自动切换到新的主节点。 |

**公网和私网地址**

集群地址和主地址都有公网和私网两种形式。私网也称为内网。

* 选择私网地址
  + 如果应用或客户端部署在ECS实例，且该ECS实例与PolarDB实例在同一地域，且网络类型相同，则PolarDB实例与ECS实例可以通过内网互通，无需申请公网地址。
  + 通过私网的连接地址访问可以发挥PolarDB的最佳性能。
* 选择公网地址
* 无法通过内网访问PolarDB实例时，需要申请公网地址。具体场景如下：
  + - ECS实例访问PolarDB实例，且ECS实例与PolarDB实例位于不同地域，或者网络类型不同。
    - 阿里云以外的设备访问PolarDB实例。
* 使用公网地址会降低实例的安全性，需谨慎使用。

#### 9.2.3.3 连接数据库

用户可以通过应用程序、客户端或命令行连接数据库。以下表格列出需要填写的信息。

|  |  |
| --- | --- |
| 参数 | 说明 |
| 主机名称/地址 | 输入PolarDB集群的连接地址。查看PolarDB集群的地址及端口信息的步骤如下：   1. 登录[PolarDB控制台](https://polardb.console.aliyun.com/)。 2. 在控制台左上角，选择集群所在地域。 3. 单击目标集群ID。 4. 在基本信息页面，查看PolarDB连接地址。 |
| 端口 | 在查看连接地址的地方，可以查看到对应的端口号。以下是默认端口号：   * PolarDB MySQL：默认为3306。 * PolarDB PostgreSQL，默认为1921。 * PolarDB 兼容Oracle版，默认为1521。 |
| 数据库 | 在数据库管理页面可以查看或创建数据库。  仅PolarDB PostgreSQL和兼容Oracle版需要填写。 |
| 用户名 | 在账号管理页面可以查看或创建数据库账号。 |
| 密码 | 上述账号对应的密码。 |

## 9.3 基本操作

### 数据库与表创建

**创建数据库**

一个PolarDB集群中可以创建一个或多个数据库。

1. 登录PolarDB控制台。
2. 在控制台左上角，选择集群所在地域。
3. 单击目标集群的ID。
4. 在左侧导航栏中，进入数据库管理页面。
5. 单击创建数据库。
6. 设置数据库参数。

* PolarDB MySQL

|  |  |
| --- | --- |
| 参数 | 说明 |
| 数据库（DB）名称 | * 以字母开头，以字母或数字结尾； * 由小写字母、数字、下划线或中划线组成； * 长度为2~64个字符。 * 数据库名称在实例内必须是唯一的。 |
| 支持字符集 | 选择utf8mb4、utf8、gbk或latin1。  如果需要其他字符集，请在右侧下拉菜单中选择需要的字符集。 |
| 授权账号 | 选中需要授权访问本数据库的账号。本参数可以留空，在创建数据库后再绑定账号。  **说明** 此处只会显示普通账号，因为高权限账号拥有所有数据库的所有权限，不需要授权。 |
| 账号权限 | 选择要授予账号的权限：读写、只读或仅DML。 |
| 备注说明 | 用于备注该数据库的相关信息，便于后续数据库管理。要求如下：   * 不能以http://或<https://开头>。 * 必须以大小写字母或中文开头。 * 可以包含大小写字母、中文、数字、下划线“\_”或连字符“-”。 * 长度为2~256个字符。 |

* PolarDB PostgreSQL或兼容Oracle版

|  |  |
| --- | --- |
| 参数 | 说明 |
| 数据库（DB）名称 | * 以字母开头，以字母或数字结尾； * 由小写字母、数字、下划线或中划线组成； * 长度为2~64个字符。 * 数据库名称在实例内必须是唯一的。 |
| 数据库Owner | 数据库的所有者，对数据库拥有ALL权限。 |
| 支持字符集 | 数据库支持的字符集，默认为UTF8。 如果需要其他字符集，请在下拉列表中  选择需要的字符集。 |
| Collate | 字符串排序规则。 |
| Ctype | 字符分类 |
| 备注说明 | 用于备注该数据库的相关信息，便于后续数据库管理。要求如下：   * 不能以http://或<https://开头>。 * 必须以大小写字母或中文开头。 * 可以包含大小写字母、中文、数字、下划线“\_”或连字符“-”。 * 长度为2~256个字符。 |

**创建表**

下面介绍如何通过DMS的SQLConsole创建表。

1. 登录PolarDB控制台。

2. 在控制台左上角，选择集群所在地域。

3. 单击目标集群的ID。

4. 在左侧导航栏中，进入数据库管理页面。

5. 找到目标数据库，单击SQL查询。

如果弹出登陆数据库的对话框，则输入数据库的账号和密码。对于PolarDB PostgreSQL和兼容Oracle版本，还需填写要登陆的数据库。在PolarDB集群的数据 库管理页可以创建数据库。

6. 在SQLConsole页面，输入创建表的命令并单击执行。

例如，执行以下命令创建big\_table表。

1 CREATE TABLE `big\_table` (

2 `id` bigint(20) unsigned NOT NULL AUTO\_INCREMENT COMMENT '主键',

3 `name` varchar(64) NOT NULL COMMENT '名称',

4 `long\_text\_a` varchar(1024) DEFAULT NULL COMMENT 'A文本',

5 `long\_text\_b` varchar(1024) DEFAULT NULL COMMENT 'B文本',

6 PRIMARY KEY (`id`)

7 ) ENGINE=InnoDB DEFAULT CHARSET=utf8 COMMENT='大表变更';

### 9.3.2 创建测试数据

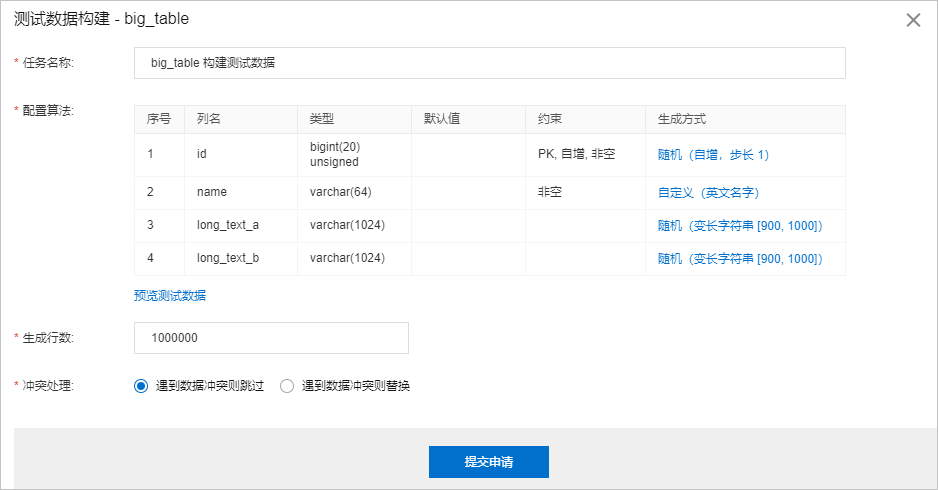
本示例通过测试数据构建功能批量生成big\_table 表的100万条测试数据。

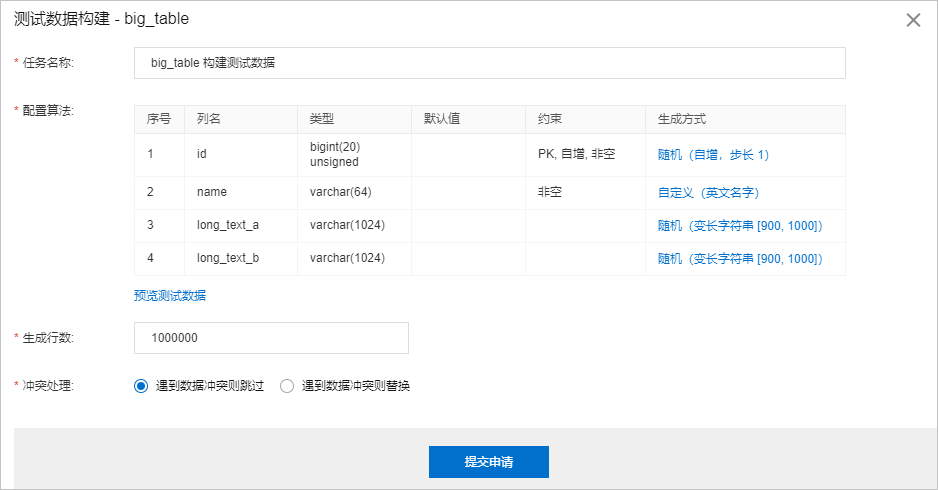
1. 登录DMS控制台。

2. 在DMS对象列表中展开目标PolarDB实例，双击目标数据库，即可进入SQLConsole页。

3. 在SQLConsole窗口中，右键单击big\_table 表，在弹出的列表中选择数据方案 > 测试数据构建。

4. 填写相关参数并提交，等待审批完成。





5. 待审批通过后，系统会自动生成SQL语句并执行，可以在工单详情页中查看执行进度。

6. 系统执行完毕后，可以到该数据库中执行以下命令查看数据生成情况。

SELECT COUNT(\*) FROM '表名'

### 9.3.3 用户与权限管理

### 9.3.4 数据查询

### 9.3.5 事务处理

## 9.4 数据导入与导出

### 9.4.1 数据导入

### 9.4.2 数据导出

## 9.5 云上数据迁移

### 9.5.1 云上数据的迁入

国际权威研究机构Gartner发布报告指出，数据库的未来趋势就是上云。下面以MySQL为例，介绍如何将自建MySQL迁移至云数据库PolarDB MySQL。

**DTS简介**

数据传输服务DTS（Data Transmission Service）是阿里云提供的实时数据流服务，支持RDBMS、NoSQL、OLAP等数据源，集数据迁移/订阅/同步于一体，提供稳定安全的传输链路。

**准备工作**

* 已有自建MySQL数据库（5.1、5.5、5.6、5.7或8.0版本）。
* 已创建目标PolarDB MySQL集群。
* 如果MySQL数据库部署在本地，那么需要将DTS服务器的IP地址设置为该数据库远程连接的白名单，
* 允许其访问数据库。详情请参见迁移、同步或订阅本地数据库时需添加的IP白名单。
* 为自建MySQL创建账号并设置binlog
* 数据库账号满足以下权限要求。

|  |  |  |
| --- | --- | --- |
| 数据库 | 结构/全量迁移 | 增量迁移 |
| 自建MySQL数据库 | SELECT权限 | REPLICATION CLIENT、REPLICATION SLAVE、SHOW VIEW和SELECT权限 |
| PolarDB集群 | 读写权限 | 读写权限 |

**注意事项**

* DTS在执行全量数据迁移时将占用源库和目标库一定的读写资源，可能会导致数据库的负载上升，在数据库性能较差、规格较低或业务量较大的情况下（例如源库有大量慢SQL、存在无主键表或目标库存在死锁等），可能会加重数据库压力，甚至导致数据库服务不可用。因此需要在执行数据迁移前评估源库和目标库的性能，同时建议在业务低峰期执行数据迁移（例如源库和目标库的CPU负载在30%以下）。
* 如果源数据库没有主键或唯一约束，且所有字段没有唯一性，可能会导致目标数据库中出现重复数据。
* 对于数据类型为FLOAT或DOUBLE的列，DTS会通过ROUND(COLUMN,PRECISION) 来读取该列的值。如果没有明确定义其精度，DTS对FLOAT的迁移精度为38位，对DOUBLE的迁移精度为308位，请确认迁移精度是否符合业务预期。
* 对于迁移失败的任务，DTS会触发自动恢复。在将业务切换至目标集群前，请务必先结束或释放该任务，避免该任务被自动恢复后，导致源端数据覆盖目标集群的数据。

**迁移类型**

支持结构迁移、全量数据迁移和增量数据迁移。

同时使用这三种迁移类型可实现在应用不停服的情况下，平滑地完成数据库迁移。

**费用说明**

|  |  |  |
| --- | --- | --- |
| 迁移类型 | 链路配置费用 | 公网流量费用 |
| 结构迁移和全量数据迁移 | 不收费。 | 通过公网将数据迁移出阿里云时将收费，详情请参见[产品定价](https://help.aliyun.com/document_detail/117780.html#concept-261679)。 |
| 增量数据迁移 | [收费，详情请参见产品定价。](https://help.aliyun.com/document_detail/117780.html#concept-261679) |

**如何操作**

1. 登录数据传输控制台。

2. 选择迁移的目标集群所属地域，并打开创建迁移任务的页面。

3. 配置迁移任务的源库和目标库连接信息

|  |  |  |
| --- | --- | --- |
| 类别 | 配置 | 说明 |
| 无 | 任务名称 | DTS会自动生成一个任务名称，建议配置具有业务意义的名称（无唯一性要求），便于后续识别。 |
| 源库信息 | 实例类型 | 根据源库的部署位置进行选择，本文以有公网IP的自建数据库为例介绍配置流程。  **说明** 当自建数据库为其他实例类型时，还需要执行相应的准备工作，详情请参见[准备工作概览](https://help.aliyun.com/document_detail/146958.html#concept-2364477)。 |
| 实例地区 | 当实例类型选择为有公网IP的自建数据库时，实例地区无需设置。  **说明** 如果自建MySQL数据库具有白名单安全设置，需要在实例地区配置项后，单击获取DTS IP段来获取到DTS服务器的IP地址， 并将获取到的IP地址加入自建MySQL数据库的白名单安全设置中。 |
| 数据库类型 | 选择MySQL |
| 主机名或IP地址 | 填入自建MySQL数据库的访问地址，本案例中填入公网地址。 |
| 端口 | 填入自建MySQL数据库的服务端口，默认为3306。 |
| 数据库账号 | [填入自建MySQL的数据库账号，权限要求请参见数据库账号的权限要求。](https://help.aliyun.com/document_detail/170538.html?spm=a2c4g.11174283.6.582.7df06121yn74jc&section-3jt-90x-lna) |
| 数据库密码 | 填入该账号对应的密码。  **说明** 源库信息填写完毕后，可以单击数据库密码后的测试连接来  验证填入的信息是否正确。如果填写正确则提示测试通过；如果提  示测试失败，单击测试失败后的诊断，根据提示调整填写的源库信  息。 |
| 目标实例信息 | 实例类型 | 选择PolarDB。 |
| 实例地区 | 选择目标PolarDB集群所属的地域。 |
| PolarDB实例ID | 选择目标PolarDB集群ID。 |
| 数据库账号 | 填入目标PolarDB集群的数据库账号，权限要求请参见数据库账号  的权限要求。 |
| 数据库密码 | 填入该账号对应的密码。  **说明** 目标库信息填写完毕后，可以单击数据库密码后的测试连接  来验证填入的信息是否正确。如果填写正确则提示测试通过；如果  提示测试失败，单击测试失败后的诊断，根据提示调整填写的目标  库信息。 |

1. 配置完成后，单击页面右下角的授权白名单并进入下一步。

**说明** 此步骤会将DTS服务器的IP地址自动添加到目标PolarDB MySQL的白名单中，用于 保障DTS服务器能够正常连接目标集群。

2. 选择迁移类型和迁移对象。

|  |  |
| --- | --- |
| 配置 | 说明 |
| 迁移类型 | * 如果只需要进行全量迁移，则同时选中结构迁移和全量数据迁移。 * 如果需要进行不停机迁移，则同时选中结构迁移、全量数据迁移和增量数据迁移。   **注意** 如果未选中增量数据迁移，为保障数据一致性，全量数据迁移期间请勿在源库中写入新的数据。 |
| 迁移对象 | 在迁移对象框中单击待迁移的对象，然后单击>图标将其移动至已选择对象框。  **注意**   * 迁移对象选择的粒度为库、表、列。 * 默认情况下，迁移对象在目标库中的名称与源库保持一致。如果需要改变迁移对象在目标库中的名称，需要使用对象名映射功能，详情请参见库表列映射 * 如果使用了对象名映射功能，可能会导致依赖这个对象的其他对象迁移失败。 |

1. 单击页面右下角的预检查并启动。
   * 在迁移任务正式启动之前，会先进行预检查。只有通过预检查，DTS才能迁移数据。
   * 如果预检查失败，单击具体检查项后的图标，查看失败详情。根据提示修复后，重新进行预检查。
2. 预检查通过后，确认购买并启动迁移任务。
   * 结构迁移+全量数据迁移请勿手动结束迁移任务，否则可能会导致数据不完整。只需等待迁移任务完成即可，迁移任务会自动结束。
   * 结构迁移+全量数据迁移+增量数据迁移迁移任务不会自动结束，需要手动结束迁移任务。注意请选择合适的时间手动结束迁移任务，例如业务低峰期或准备将业务切换至目标集群时。
     1. 观察迁移任务的进度变更为增量迁移，并显示为无延迟状态时，将源库停写几分钟，此时增量迁移的状态可能会显示延迟的时间。
     2. 等待迁移任务的增量迁移再次进入无延迟状态后，手动结束迁移任务。
3. 将业务切换至PolarDB集群。

**附：增量数据迁移阶段支持同步的SQL操作**

|  |  |
| --- | --- |
| 操作类型 | SQL操作语句 |
| DML | INSERT、UPDATE、DELETE、REPLACE |
| DDL | * ALTER TABLE、ALTER VIEW * CREATE FUNCTION、CREATE INDEX、CREATE PROCEDURE、CREATE TABLE、CREATE VIEW * DROP INDEX、DROP TABLE * RENAME TABLE * TRUNCATE TABLE |

### 9.5.2 云上数据的迁出