

# 阿里云 数据传输

用户指南

文档版本：20181217

# 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
	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。	 <b>禁止：</b> 重置操作将丢失用户配置数据。
	该类警示信息可能导致系统重大变更甚至故障，或者导致人身伤害等结果。	 <b>警告：</b> 重启操作将导致业务中断，恢复业务所需时间约10分钟。
	用于补充说明、最佳实践、窍门等，不是用户必须了解的内容。	 <b>说明：</b> 您也可以通过按 <b>Ctrl + A</b> 选中全部文件。
>	多级菜单递进。	设置 > 网络 > 设置网络类型
<b>粗体</b>	表示按键、菜单、页面名称等UI元素。	单击 <b>确定</b> 。
<code>courier</code> 字体	命令。	执行 <code>cd /d C:/windows</code> 命令，进入Windows系统文件夹。
斜体	表示参数、变量。	<code>bae log list --instanceid 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>swich {stand   slave}</code>

# 目录

---

法律声明.....	
通用约定.....	
<b>1 创建 RDS for MySQL 实例同POLARDB间的数据实时同步.....</b>	<b>1</b>

# 1 创建 RDS for MySQL 实例同POLARDB间的数据实时同步

---

本小节介绍如何使用数据传输服务DTS快速创建RDS for MySQL同POLARDB 实例间的实时同步作业，实现 RDS for MySQL到POLARDB 增量数据的实时同步。

## 支持数据源

- 支持同一个阿里云账号下RDS for MySQL 同POLARDB实例间的实时同步。
- 支持不同阿里云账号下的 RDS for MySQL同POLARDB 实例间的实时同步。此时，需要使用目标POLARDB所属的阿里云账号配置任务，源RDS实例需要进行权限授权，具体授权方案参考[授权指南](#)。

## 支持同步的SQL操作

对于**MySQL > POLARDB**数据同步，DTS支持同步的SQL操作包括：

Insert、Update、Delete、Replace

ALTER TABLE、ALTER VIEW、ALTER FUNCTION、ALTER PROCEDURE

CREATE DATABASE、CREATE SCHEMA、CREATE INDEX、CREATE TABLE、CREATE PROCEDURE、CREATE FUNCTION、CREATE TRIGGER、CREATE VIEW、CREATE EVENT

DROP FUNCTION、DROP EVENT、DROP INDEX、DROP PROCEDURE、DROP TABLE、DROP TRIGGER、DROP VIEW

RENAME TABLE、TRUNCATE TABLE

## 注意事项

全量初始化过程中，并发insert导致目标实例的表碎片，全量初始化完成后，目标实例的表空间比源实例的表空间大。

## 同步限制

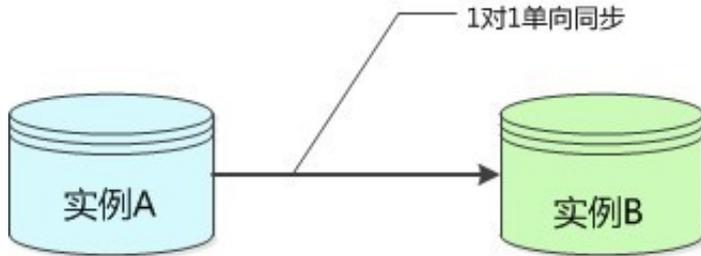
### 数据源

对于 rename table tbl\_name to new\_tbl\_name、create table tbl\_name like new\_tbl\_name、create...select...from new\_tbl\_name、alter table tbl\_name rename to new\_tbl\_name，如果 new\_tbl\_name 不在指定的同步对象中，则不支持对此 DDL 进行复制。

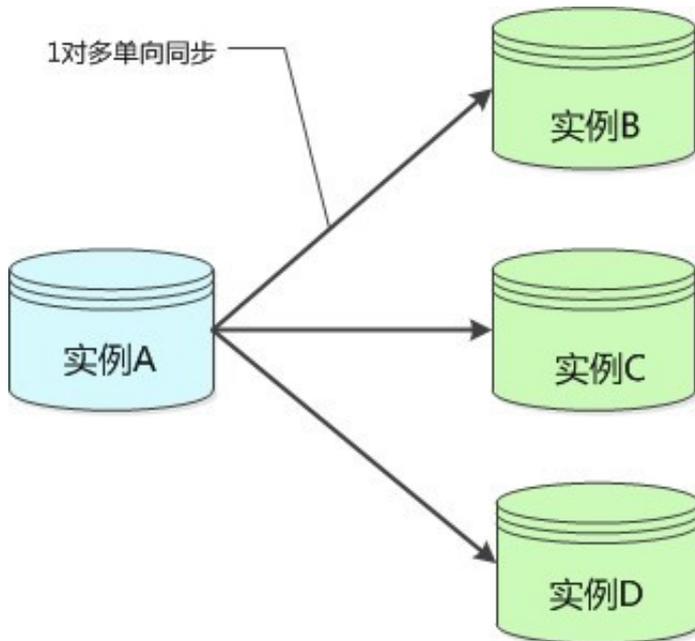
### 同步架构

目前数据传输服务提供的实时同步功能支持的同步架构有限，其仅能支持如下架构：

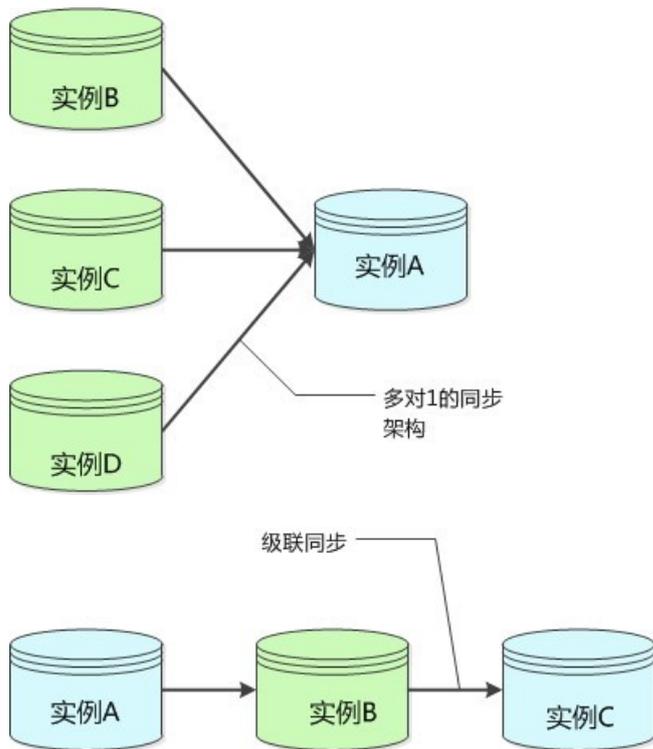
1. **A > B**即两个实例之间的单向同步。且要求实例 B 中同步的对象必须为只读，否则会导致同步链路异常，出现数据不一致的情况。



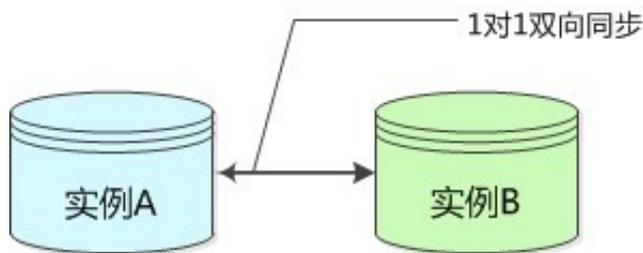
2. **A > B/C/D** 即一对多的分发式同步架构,这个架构对目标POLARDB实例个数没有限制，但是要求目标实例中的同步对象必须为只读，否则会导致同步链路异常，出现数据不一致的情况。



3. **B/C/D > A**即多对一的数据汇总架构。对于这种多对一的同步架构，为了保证同步数据一致性，要求每条同步链路同步的对象不相同。



4.  $A > B > A$ 即实例A和实例B之间的双向同步架构。



说明：

如果需要使用双向同步，需要在购买同步链路时，选择双向同步，并在 [数据传输 DTS 控制台](#) 中根据指引进行配置。

如果用户配置同步链路过程中，配置不在上述支持范围内的的同步架构，那么预检查中的复杂拓扑检查项会检查失败。

### 功能限制

- 不兼容触发器

如果同步对象为整个库且这个库中包含了会更新同步表内容的触发器，会导致同步数据不一致。

例如同步库为A，这个库中存在了两个表 A, B。表 A 上有一个触发器，触发器内容为在 insert 一条数据到表 A 之后，在表 B 中插入一条数据。这种情况下，在同步过程中，如果源实例有表 A 上的 insert 操作，就会导致表 B 在源实例跟目标实例数据不一致。

为了解决这个问题，只能将目标实例中的对应触发器删除掉。表 B 的数据由源实例同步过去。具体解决方案详见最佳实践中的，触发器存在情况下如何配置同步链路。

- rename table 限制

rename table 操作需要满足限制条件方可正常同步，否则会导致同步数据不一致。例如同步对象只包含表 A，不包含表 B，如果同步过程中源实例执行了 rename A to B 的操作，那么改名后的表 B 的操作不会被同步到目标库。为了解决这个问题，可以选择同步表 A、B 对应的整个数据库。

## 准备事项

- 在配置同步作业前，要确保同步作业的源RDS实例及目标POLARDB实例都已经存在。如果不存在，那么请先购买实例 [购买RDS实例](#) [购买POLARDB实例](#)。
- 在配置同步作业前，需要先将POLARDB实例所在区域的DTS IP段添加到POLARDB实例的白名单中。各区域DTS IP段参考[用户手册](#)。

## 配置步骤

下面我们详细介绍下配置 RDS同POLARDB间的同步实例的具体步骤。

### 购买同步链路

1. 进入[数据传输 DTS 控制台](#)，进入数据同步页面，单击控制台右上角“创建同步作业”开始作业配置。
2. 在链路配置之前需要购买一个同步链路。同步链路目前支持包年包月及按量付费两种付费模式，可以根据需要选择不同的付费模式。

在购买页面需要配置参数包括：

- 源实例

源实例目前支持MySQL、DRDS，此处选择MySQL。

- 源实例区域

源地域选择源RDS实例所在地域。

- 目标实例

目标实例支持MySQL、MaxCompute、Datahub、分析型数据库AnalyticDB、Elasticsearch。此处选择MySQL，需要将POLARDB当作通过专线接入的自建MySQL来配置。

目标地域为同步链路目标实例所在地域。

- 目标实例地域

此处选择POLARDB实例所在的区域。

- 同步拓扑

对于MySQL与POLARDB之间的同步拓扑支持：单向同步及双向同步。此处选择单向同步。

- 同步链路规格

实例规格影响了链路的同步性能，实例规格跟性能之间的对应关系详见 [数据同步规格说明](#)。

- 数量

数量为一次性购买的同步链路的数量，如果购买的是按量付费实例，一次最多购买 99 条链路。

当购买完同步实例，返回数据传输控制台，单击新购链路右侧的“配置同步作业”开始链路配置。

### 同步链路连接信息配置

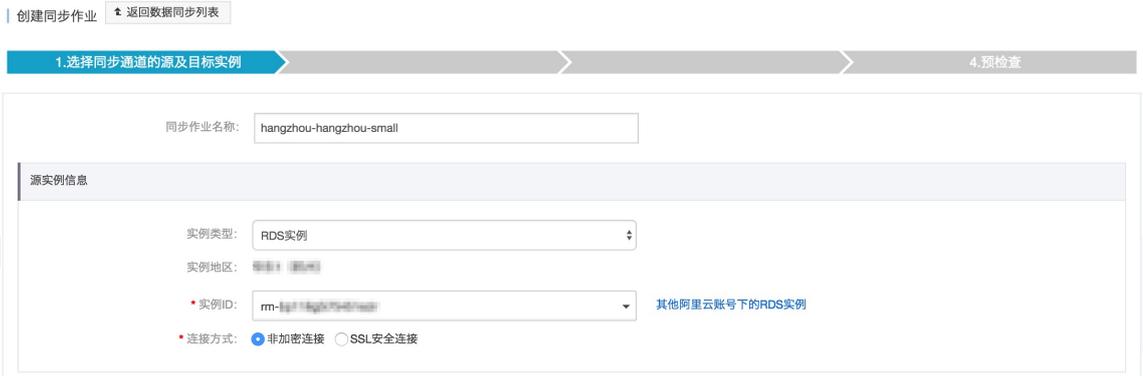
当购买完实例后，开始进行同步实例配置，第一步主要进行同步实例名称及数据源连接信息的配置。具体配置内容如下：

#### 1. 同步实例名称

同步实例名称没有唯一性要求，主要为了更方便识别具体的作业，建议选择一个有业务意义的作业名称，方便后续的链路查找及管理。

#### 2. 源实例连接信息

- 实例类型：此处选择RDS实例。
- 实例 ID：此处配置源RDS实例的实例ID。
- 连接方式：对于RDS实例，支持非加密连接和SSL安全连接两种方式。可以根据需要选择连接方式。如果要选择SSL安全连接，那必须先打开RDS的加密连接，开启方法参考[用户指南](#)。

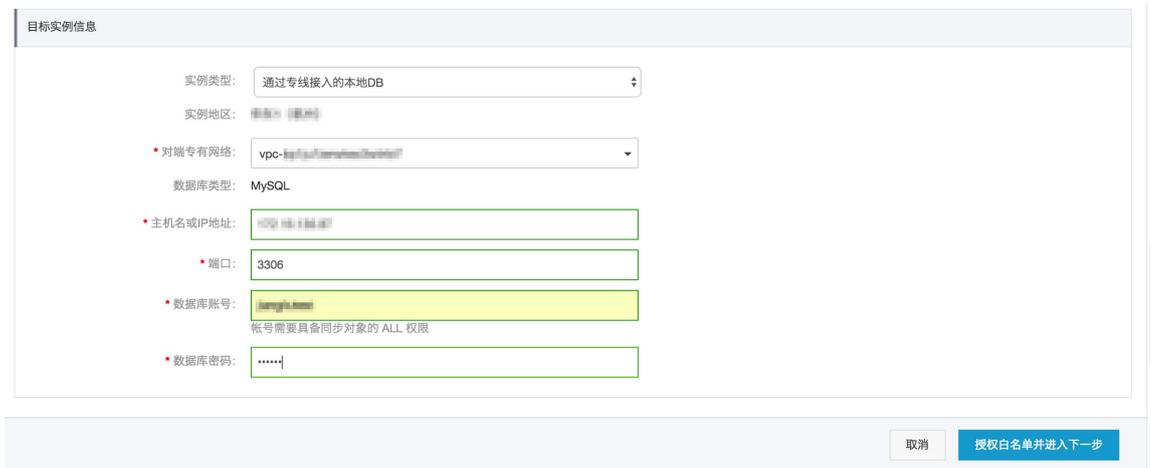


### 3. 目标实例连接信息

- 实例类型：此处选择通过专线接入的本地DB。
- 对端专有网络：此处配置POLARDB的VPC ID。具体VPC ID可以到POLARDB控制台实例信息界面获取。



- 主机名或IP地址：配置POLARDB的IP地址，可以通过ping POLARDB的连接地址，获取对应的IP地址。
- 端口：配置POLARDB的监听端口。
- 数据库账号：配置POLARDB的访问账号。
- 数据库密码：配置POLARDB上面账号对应的数据库密码。



当这些内容配置完成后，可以单击授权白名单并进入下一步。

## 授权实例白名单

这个步骤，主要是将数据传输服务器 IP 添加到同步 RDS 实例的白名单中。避免因为 RDS 设置了白名单，数据传输服务器连接不上 RDS 导致同步作业创建失败。

为了保证同步作业的稳定性，在同步过程中，请勿将这些服务器 IP 从 RDS 实例的白名单中删除。

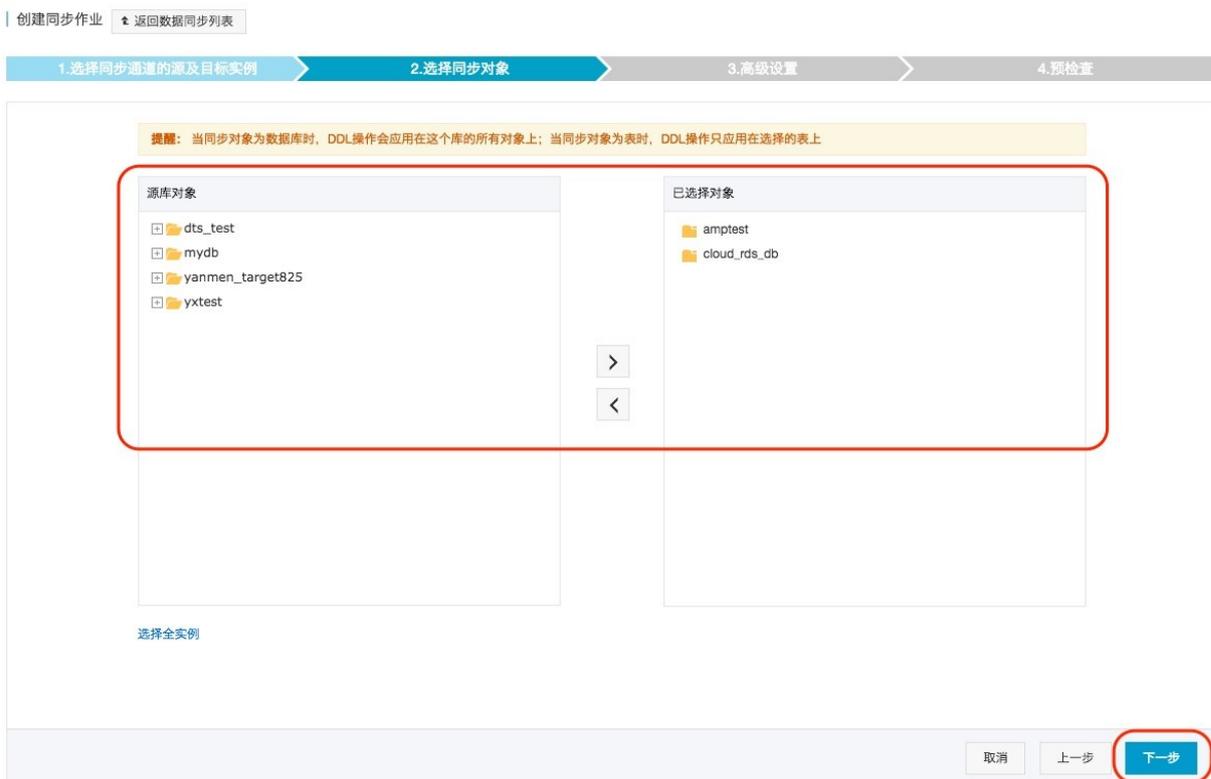
当白名单授权后，单击下一步，进行同步对象的配置。

## 选择同步对象

这个步骤主要进行同步对象配置，实时同步的同步对象的选择粒度可以支持到表级别,即用户可以选择同步某些库或是同步某几张表。

如果选择的同步对象为整个库，那么这个库中所有对象的结构变更操作（例如 create table，drop view 等），都会同步到目标库。

如果选择的某张表，那么只有这个表的 drop/alter/truncate/rename table，create/drop index 的操作会同步到目标库。

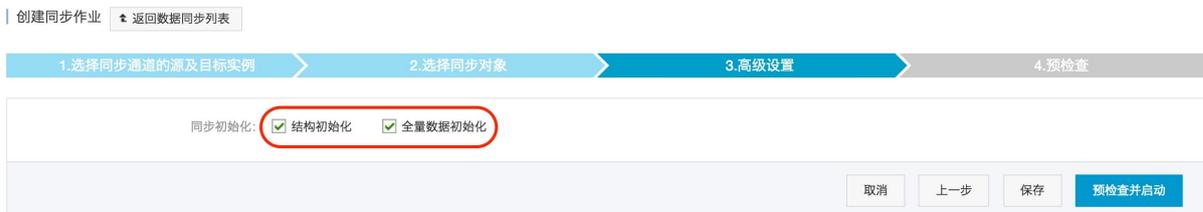


当配置完同步对象后，进入同步初始化配置。

## 同步初始化配置

同步初始化配置，初始化是同步链路启动的第一步，它会将源实例中已经存在同步对象的结构及数据在目标实例中初始化，作为后续增量同步数据的基线数据。

同步初始化类型细分为：结构初始化，全量数据初始化。默认情况下，需要选择结构初始化及全量初始化。



### 预检查

当上面所有选项配置完成后，即进入启动之前的预检查。

当同步作业配置完成后，数据传输服务会进行限制预检查，当预检查通过后，DTS直接启动同步作业。

当同步作业启动之后，即进入同步作业列表。此时刚启动的作业处于同步初始化状态。初始化的时间长度依赖于源实例中同步对象的数据量大小。当初始化完成后同步链路即进入同步中的状态，此时源跟目标实例的同步链路才真正建立完成。