

ALIBABA CLOUD

# Alibaba Cloud

数据传输  
数据同步

文档版本：20220711

 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<br>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<br>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<br>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<br>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。   |
| <b>粗体</b>  | 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><br><i>Instance_ID</i>  |
| [ ] 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.数据同步拓扑介绍                       | 08  |
| 2.结构初始化涉及的数据类型映射关系               | 10  |
| 3.注意事项及限制                        | 28  |
| 3.1. 同步时源库为MySQL的注意事项及限制         | 28  |
| 3.2. 同步时源库为PolarDB MySQL的注意事项及限制 | 45  |
| 3.3. 同步时源库为PolarDB-X 1.0的注意事项及限制 | 52  |
| 3.4. 同步时源库为PolarDB-X 2.0的注意事项及限制 | 56  |
| 3.5. 源为PolarDB-X的功能规范和约束说明       | 62  |
| 3.6. 源为PolarDB-X的任务配置方案          | 64  |
| 3.7. 同步时源库为Oracle的注意事项及限制        | 66  |
| 3.8. 同步时源库为PostgreSQL的注意事项及限制    | 70  |
| 3.9. 同步时源库为SQL Server的注意事项及限制    | 82  |
| 3.10. 同步时源库为MongoDB的注意事项及限制      | 87  |
| 3.11. 同步时源库为DB2 for LUW的注意事项及限制  | 90  |
| 3.12. 同步时源库为Redis的注意事项及限制        | 93  |
| 4.准备用于数据同步的数据库账号                 | 96  |
| 5.同步作业管理                         | 99  |
| 5.1. 设置同步对象在目标实例中的名称             | 99  |
| 5.2. 通过SQL条件过滤任务数据               | 101 |
| 5.3. 目标库对象名称大小写策略                | 102 |
| 5.4. 修改数据同步的附加列规则                | 103 |
| 5.5. 查看同步链路状态和性能                 | 105 |
| 5.6. 重新配置数据同步作业                  | 107 |
| 5.7. 新增同步对象                      | 108 |
| 5.8. 移除同步对象                      | 109 |
| 5.9. 结束数据同步作业                    | 110 |

---

|   |     |
|---|-----|
| 5.10. 修复同步作业预检查失败的问题                              | 111 |
| 5.11. Kafka集群的数据存储格式                              | 112 |
| 5.12. Kafka Partition同步策略说明                       | 117 |
| 6. MySQL间数据同步                                     | 119 |
| 6.1. MySQL实例间的双向同步                                | 119 |
| 6.2. RDS MySQL实例间的单向同步                            | 125 |
| 6.3. 从ECS上的自建MySQL同步至RDS MySQL                    | 130 |
| 6.4. 从通过专线、VPN网关或智能接入网关接入的自建MySQL同步至RD...         | 135 |
| 6.5. 从RDS MySQL同步至通过专线、VPN网关或智能接入网关接入的自...        | 140 |
| 6.6. 不同阿里云账号下RDS MySQL实例间的数据同步                    | 145 |
| 7. MySQL同步至其他数据库                                  | 151 |
| 7.1. 从RDS MySQL同步至PolarDB MySQL                   | 151 |
| 7.2. 从ECS上的自建MySQL同步至PolarDB MySQL                | 156 |
| 7.3. 从通过专线、VPN网关或智能接入网关接入的自建MySQL同步至Pol...        | 161 |
| 7.4. 从RDS MySQL同步到云原生数据仓库AnalyticDB MySQL         | 167 |
| 7.5. 从RDS MySQL同步至云原生数据仓库AnalyticDB PostgreSQL    | 173 |
| 7.6. 从ECS上的自建MySQL同步至云原生数据仓库AnalyticDB Postgre... | 178 |
| 7.7. 从通过专线、VPN网关或智能接入网关接入的自建MySQL同步至云原...         | 183 |
| 7.8. 从RDS MySQL同步至DataHub                         | 189 |
| 7.9. 从ECS上的自建MySQL同步至Elasticsearch                | 195 |
| 7.10. 从RDS MySQL同步至MaxCompute                     | 201 |
| 7.11. 从RDS MySQL同步至Tablestore                     | 210 |
| 7.12. 从自建MySQL同步至阿里云消息队列Kafka版                    | 218 |
| 7.13. 从RDS MySQL同步至自建Kafka集群                      | 222 |
| 7.14. 从ECS上的自建MySQL同步至自建Kafka集群                   | 227 |
| 7.15. 从通过专线、VPN网关或智能接入网关接入的自建MySQL同步至自...         | 232 |
| 7.16. 从RDS MySQL同步至PolarDB-X                      | 236 |
| 8. PolarDB MySQL数据同步                              | 241 |

---

---

|   |     |
|---|-----|
| 8.1. PolarDB MySQL集群间的双向同步                              | 241 |
| 8.2. PolarDB MySQL集群间的单向同步                              | 248 |
| 8.3. 从PolarDB MySQL同步至RDS MySQL                         | 253 |
| 8.4. 从PolarDB MySQL同步至Datahub                           | 258 |
| 8.5. 从PolarDB MySQL同步至Elasticsearch                     | 264 |
| 8.6. 从PolarDB MySQL同步到Kafka                             | 270 |
| 8.7. 从PolarDB MySQL同步至云原生数据仓库AnalyticDB MySQL           | 274 |
| 8.8. 从PolarDB MySQL同步至云原生数据仓库AnalyticDB PostgreSQL...   | 280 |
| 8.9. 从PolarDB MySQL同步至PolarDB-X                         | 285 |
| 9.PolarDB O引擎数据同步                                       | 290 |
| 9.1. PolarDB O引擎集群间的单向同步                                | 290 |
| 9.2. 从PolarDB O引擎集群同步至自建Kafka                           | 295 |
| 10.PolarDB-X数据同步  | 300 |
| 10.1. PolarDB-X间的数据实时同步                                 | 300 |
| 10.2. 从PolarDB-X同步至云原生数据仓库AnalyticDB MySQL              | 303 |
| 10.3. 从PolarDB-X同步至AnalyticDB for PostgreSQL            | 308 |
| 11.Redis数据同步  | 314 |
| 11.1. Redis实例间的单向数据同步                                   | 314 |
| 11.2. Redis企业版实例间的双向同步                                  | 319 |
| 11.3. 调用OpenAPI配置Redis企业版实例间单向或双向数据同步                   | 325 |
| 11.4. 从ECS上的自建Redis同步至Redis实例                           | 331 |
| 11.5. 从通过专线、VPN网关或智能接入网关接入的自建Redis同步至ECS...             | 337 |
| 11.6. 从自建Redis集群同步至Redis集群实例                            | 343 |
| 11.7. 从ECS上的Codis集群同步至Redis实例                           | 348 |
| 11.8. 从ECS上的Twemproxy Redis集群同步至Redis实例                 | 355 |
| 12.SQL Server数据同步                                       | 363 |
| 12.1. 从RDS SQL Server同步至云原生数据仓库AnalyticDB PostgreSQL... | 363 |
| 12.2. 从ECS上的自建SQL Server同步至云原生数据仓库AnalyticDB Po...      | 369 |

---

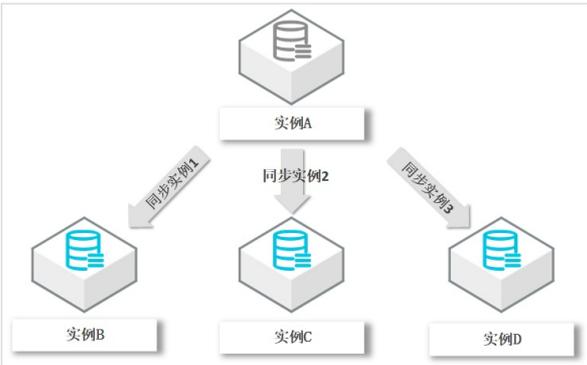
|  |     |
|--|-----|
| 13.PostgreSQL数据同步 .....                                    | 377 |
| 13.1. 从RDS PostgreSQL同步至云原生数据仓库AnalyticDB PostgreSQL ..... | 377 |
| 13.2. 自建PostgreSQL同步到云原生数据仓库AnalyticDB PostgreSQL .....    | 381 |
| 14.TiDB数据同步 .....  | 387 |
| 14.1. 从自建TiDB同步至云原生数据仓库AnalyticDB MySQL .....              | 387 |

# 1.数据同步拓扑介绍

数据同步功能支持多种同步拓扑，您可以根据业务需求规划您的同步实例。本文介绍DTS支持的数据同步类型和使用说明。

## 单向同步

在单向同步的拓扑中，为保障同步数据的一致性，对目标实例中的同步对象应只进行读操作，请勿进行写操作。

| 拓扑类型    | 拓扑图   | 说明  |
|---------|---|---|
| 一对一单向同步 |    | 无   |
| 一对多单向同步 |   | 您需要购买多个同步实例来实现一对多单向同步。<br>例如，从实例A同步到实例B、实例C和实例D，那么您需要购买三个同步实例。    |
| 级联单向同步  |  | 您需要购买多个同步实例来实现级联单向同步。<br>例如，从实例A同步到实例B，再从实例B同步到实例C，那么您需要购买两个同步实例。 |

| 拓扑类型    | 拓扑图 | 说明  |
|---------|-----|---|
| 多对一单向同步 |     | <p>您需要购买多个同步实例来实现多对一单向同步。<br/>例如，从实例B、实例C和实例D同步到实例A，那么您需要购买三个同步实例。</p> <p><b>说明</b> 为保障同步数据的一致性，每个同步实例应选择不同的同步对象。</p> |

### 双向同步

目前DTS仅支持两个MySQL数据库之间的双向同步，暂不支持多个MySQL数据库之间的双向同步。

**说明**

- 关于双向同步的配置案例请参见[MySQL实例间的双向同步](#)。
- 您可以将现有的单向同步升级为双向同步，详情请参见[单向同步升级至双向同步](#)。

| 拓扑类型    | 拓扑图 | 说明   |
|---------|-----|--|
| 一对一双向同步 |     | <p>为保障同步数据的一致性，您需要确保同一个主键、业务主键、唯一键的记录只在一个实例中更新。</p> <p><b>说明</b> 如果同时更新则会按照您在数据同步作业中配置的冲突修复策略进行响应。</p> |

### 相关文档

- [DTS数据同步方案概览](#)
- [同步方案概览](#)

## 2. 结构初始化涉及的数据类型映射关系

不同类型的数据库（异构数据库）支持的数据类型不一样，DTS在进行异构数据库之间的数据同步时，会在结构初始化阶段进行数据类型映射，即将源库中的数据类型转为目标库支持的数据类型。本文为您列出详细数据类型映射关系，便于您查阅和评估数据同步对业务的影响。

### 概览

根据如下同步方案，查看异构数据库间的数据类型映射关系：

- 以MySQL数据库为源的数据同步
- 以Oracle为源的数据同步
- 以PostgreSQL为源的数据同步
- 以SQL Server为源的数据同步
- 以DB2 for LUW为源的数据同步
- 以Db2 for i为源的数据同步

### 以MySQL数据库为源的数据同步

当源数据库类型为MySQL（如RDS MySQL、自建MySQL、PolarDB MySQL）、且目标实例为异构数据库时（包括云原生数据仓库AnalyticDB MySQL版和2.0、云原生数据仓库AnalyticDB PostgreSQL版），两者间的数据类型映射关系如下：

 **说明** 如源实例中待同步数据范围超出DTS所支持的范围，则会导致目标实例中的所写入数据的精度降低。

目标实例为云原生数据仓库AnalyticDB MySQL版、云原生数据仓库AnalyticDB PostgreSQL版

| 类型   | 源实例的数据类型                  | 数值范围                                       | 云原生数据仓库AnalyticDB MySQL版的数据类型 | 云原生数据仓库AnalyticDB PostgreSQL版的数据类型 |
|------|---------------------------|--|-------------------------------|------------------------------------|
| 整数类型 | BIT[(M)]                  | 1 ~ 64                                     | VARCHAR                       | BIT                                |
|      | TINYINT[(M)]              | -128 ~ 127                                 | TINYINT                       | SMALLINT                           |
|      | TINYINT[(M)] [UNSIGNED]   | 0 ~ 255                                    | SMALLINT                      | SMALLINT                           |
|      | SMALLINT[(M)]             | -32768 ~ 32767                             | SMALLINT                      | SMALLINT                           |
|      | SMALLINT[(M)] [UNSIGNED]  | 0 ~ 65535                                  | INT                           | INTEGER                            |
|      | MEDIUMINT[(M)]            | -8388608 ~ 8388607                         | INT                           | INTEGER                            |
|      | MEDIUMINT[(M)] [UNSIGNED] | 0 ~ 16777215                               | INT                           | INTEGER                            |
|      | INT[(M)]                  | -2147483648 ~ 2147483647                   | INT                           | INTEGER                            |
|      | INT[(M)] [UNSIGNED]       | 0 ~ 4294967295                             | BIGINT                        | BIGINT                             |
|      | BIGINT[(M)]               | -9223372036854775808 ~ 9223372036854775807 | BIGINT                        | BIGINT                             |

| 类型   | 源实例的数据类型                  | 数值范围   | 云原生数据库 AnalyticDB MySQL版的数据类型 | 云原生数据库 AnalyticDB PostgreSQL版的数据类型 |
|------|---------------------------|--|-------------------------------|------------------------------------|
|      | BIGINT[(M)]<br>[UNSIGNED] | 0 ~<br>18446744073709551<br>615  | DECIMAL(20,0)                 | NUMERIC(20)                        |
| 小数类型 | DECIMAL[(M,D)]            | M: 0~65 ;<br>D: 0~30   | DECIMAL[(M,D)]                | DECIMAL                            |
|      | FLOAT(p)                  | 1.175494351E-38 ~<br>3.402823466E+38   | FLOAT                         | REAL                               |
|      | DOUBLE[(M,D)]             | 2.2250738585072014<br>E-308 ~<br>1.7976931348623157<br>E+308   | DOUBLE                        | DOUBLE PRECISION                   |
| 时间类型 | DATE                      | 1000-01-01~9999-<br>12-31<br><br>❓ 说明 格式<br>为YYYY-MM-<br>DD (UTC时<br>间)。   | DATE                          | DATE                               |
|      | DATETIME[(fsp)]           | 1000-01-01<br>00:00:00.000000 ~<br>9999-12-31<br>23:59:59.999999<br><br>❓ 说明 格式<br>为YYYY-MM-DD<br>hh:mm:ss[.fracti<br>on] (UTC时<br>间)。 | DATETIME                      | TIMESTAMP                          |
|      | TIMESTAMP[(fsp)]          | 1970-01-01<br>00:00:01.000000 ~<br>2038-01-19<br>03:14:07.999999<br><br>❓ 说明 格式<br>为YYYY-MM-DD<br>hh:mm:ss[.fracti<br>on] (UTC时<br>间)。 | TIMESTAMP                     | TIMESTAMP WITH<br>TIME ZONE        |

| 类型    | 源实例的数据类型                    | 数值范围  | 云原生数据库 AnalyticDB MySQL版的数据类型 | 云原生数据库 AnalyticDB PostgreSQL版的数据类型 |
|-------|-----------------------------|---|-------------------------------|------------------------------------|
|       | TIME[(fsp)]                 | -838:59:59.000000 ~ 838:59:59.000000<br><br>② 说明 格式为 hh:mm:ss[.fraction] (UTC时间)。 | TIME                          | TIME                               |
|       | YEAR[(4)]                   | 1901 ~ 2155, 或 0000.  | INT                           | INTEGER                            |
| 字符串类型 | CHAR[(M)]                   | 0 ~ 255 字符  | VARCHAR                       | CHAR                               |
|       | VARCHAR(M)                  | 0 ~ 65,535 字符   | VARCHAR                       | VARCHAR                            |
|       | BINARY[(M)]                 | 0 ~ 255 字节  | VARBINARY                     | BYTEA                              |
|       | VARBINARY(M)                | 0 ~ 65,535 字节   | VARBINARY                     | BYTEA                              |
|       | TINYBLOB                    | 255 (2 <sup>8</sup> - 1) 字节   | VARBINARY                     | BYTEA                              |
|       | TINYTEXT                    | 255 (2 <sup>8</sup> - 1) 字符   | VARCHAR                       | TEXT                               |
|       | BLOB                        | 65,535 (2 <sup>16</sup> - 1) 字节   | VARBINARY                     | BYTEA                              |
|       | TEXT                        | 65,535 (2 <sup>16</sup> - 1) 字符   | VARCHAR                       | TEXT                               |
|       | MEDIUMBLOB                  | 16,777,215 (2 <sup>24</sup> - 1) 字节   | VARBINARY                     | BYTEA                              |
|       | MEDIUMTEXT                  | 16,777,215 (2 <sup>24</sup> - 1) 字符   | VARCHAR                       | TEXT                               |
|       | LOBLOB                      | 4,294,967,295 or 4GB (2 <sup>32</sup> - 1) 字节                                     | VARBINARY                     | BYTEA                              |
|       | LONGTEXT                    | 4,294,967,295 or 4GB (2 <sup>32</sup> - 1) 字符                                     | VARCHAR                       | TEXT                               |
|       | ENUM('value1','value2',...) | 最多可包含65,535枚枚举值   | VARCHAR                       | VARCHAR(128)                       |
|       | SET('value1','value2',...)  | 最多可包含64个元素  | VARCHAR                       | VARCHAR(128)                       |
|       | GEOMETRY                    | 任意几何类型的值  | VARBINARY                     | POLYGON                            |
|       | POINT                       | 无   | VARBINARY                     | POINT                              |
|       | LINESTRING                  | 无   | VARBINARY                     | PATH                               |

| 类型<br>空间类型 | 源实例的数据类型           | 数值范围        | 云原生数据仓库<br>AnalyticDB MySQL版<br>的数据类型 | 云原生数据仓库<br>AnalyticDB<br>PostgreSQL版的数据<br>类型 |
|------------|--------------------|-------------|---------------------------------------|---|
|            | POLYGON            | 无           | VARBINARY                             | POLYGON                                       |
|            | MULTIPOINT         | 无           | VARBINARY                             | POLYGON                                       |
|            | MULTILINESTRING    | 无           | VARBINARY                             | PATH  |
|            | MULTIPOLYGON       | 无           | VARBINARY                             | POLYGON                                       |
|            | GEOMETRYCOLLECTION | 任何几何类型的值的集合 | VARBINARY                             | POLYGON                                       |
| JSON类型     | JSON               | 无           | JSON                                  | JSON  |

目标实例为DataHub、Kafka（阿里云消息队列Kafka、自建Kafka）

| 类型                        | 源实例的数据类型                        | 数值范围  | DataHub的数据类型     | 阿里云消息队列<br>Kafka、自建Kafka的<br>数据类型 |
|---------------------------|---------------------------------|---|------------------|-----------------------------------|
| 整数类型                      | BIT[(M)]                        | 1 ~ 64  | BOOLEAN   STRING |                                   |
|                           | TINYINT[(M)]                    | -128 ~ 127  | BIGINT           |                                   |
|                           | TINYINT[(M)]<br>[UNSIGNED]      | 0 ~ 255   | BIGINT           |                                   |
|                           | SMALLINT[(M)]                   | -32768 ~ 32767  | BIGINT           |                                   |
|                           | SMALLINT[(M)]<br>[UNSIGNED]     | 0 ~ 65535   | BIGINT           |                                   |
|                           | MEDIUMINT[(M)]                  | -8388608 ~ 8388607  | BIGINT           |                                   |
|                           | MEDIUMINT[(M)]<br>[UNSIGNED]    | 0 ~ 16777215  | BIGINT           |                                   |
|                           | INT[(M)]                        | -2147483648 ~<br>2147483647                               | BIGINT           |                                   |
|                           | INT[(M)] [UNSIGNED]             | 0 ~ 4294967295  | BIGINT           |                                   |
|                           | BIGINT[(M)]                     | -<br>92233720368547758<br>08 ~<br>92233720368547758<br>07 | BIGINT           |                                   |
| BIGINT[(M)]<br>[UNSIGNED] | 0 ~<br>18446744073709551<br>615 | BIGINT  |                  |                                   |
|                           | DECIMAL[(M[,D])]                | M: 0~65 ;<br>D: 0~30                                      | DECIMAL          |                                   |
|                           | FLOAT(p)                        | 1.175494351E-38 ~<br>3.402823466E+38                      | DOUBLE           |                                   |

| 小数类型<br>类型 | 源实例的数据类型       | 数值范围   | DataHub的数据类型 | 阿里云消息队列<br>Kafka、自建Kafka的<br>数据类型    |
|------------|----------------|--|--------------|--------------------------------------|
|            | DOUBLE(M,D)    | 2.2250738585072014<br>E-308 ~<br>1.7976931348623157<br>E+308             | DOUBLE       |                                      |
| 时间类型       | DATE           | 1000-01-01~9999-<br>12-31<br><br>  | TIMESTAMP    | 与MySQL、PolarDB<br>MySQL的数据类型保持<br>一致 |
|            | DATETIME(fsp)  | 1000-01-01<br>00:00:00.000000 ~<br>9999-12-31<br>23:59:59.999999<br><br> | TIMESTAMP    |                                      |
|            | TIMESTAMP(fsp) | 1970-01-01<br>00:00:01.000000 ~<br>2038-01-19<br>03:14:07.999999<br><br> | TIMESTAMP    |                                      |
|            | TIME(fsp)      | -838:59:59.000000 ~<br>838:59:59.000000<br><br>                          | STRING       |                                      |
|            | YEAR(4)        | 1901 ~ 2155, 或<br>0000.  | STRING       |                                      |
|            | CHAR(M)        | 0 ~ 255 字符   | STRING       |                                      |
|            | VARCHAR(M)     | 0 ~ 65,535 字符  | STRING       |                                      |

| 类型                        | 源实例的数据类型                    | 数值范围  | DataHub的数据类型 | 阿里云消息队列 Kafka、自建Kafka的数据类型 |
|---------------------------|-----------------------------|---|--------------|----------------------------|
| 字符串类型                     | BINARY[(M)]                 | 0 ~ 255 字节                                    | STRING       |                            |
|                           | VARBINARY(M)                | 0 ~ 65,535 字节                                 | STRING       |                            |
|                           | TINYBLOB                    | 255 (2 <sup>8</sup> - 1) 字节                   | STRING       |                            |
|                           | TINYTEXT                    | 255 (2 <sup>8</sup> - 1) 字符                   | STRING       |                            |
|                           | BLOB                        | 65,535 (2 <sup>16</sup> - 1) 字节               | STRING       |                            |
|                           | TEXT                        | 65,535 (2 <sup>16</sup> - 1) 字符               | STRING       |                            |
|                           | MEDIUMBLOB                  | 16,777,215 (2 <sup>24</sup> - 1) 字节           | STRING       |                            |
|                           | MEDIUMTEXT                  | 16,777,215 (2 <sup>24</sup> - 1) 字符           | STRING       |                            |
|                           | LOBLOB                      | 4,294,967,295 or 4GB (2 <sup>32</sup> - 1) 字节 | STRING       |                            |
|                           | LONGTEXT                    | 4,294,967,295 or 4GB (2 <sup>32</sup> - 1) 字符 | STRING       |                            |
|                           | ENUM('value1','value2',...) | 最多可包含65,535元素                                 | STRING       |                            |
| SET('value1','value2',..) | 最多可包含64个元素                  | STRING  |              |                            |
| 空间类型                      | GEOMETRY                    | 任意几何类型的值                                      | STRING       |                            |
|                           | POINT                       | 无   | STRING       |                            |
|                           | LINestring                  | 无   | STRING       |                            |
|                           | POLYGON                     | 无   | STRING       |                            |
|                           | MULTIPOINT                  | 无   | STRING       |                            |
|                           | MULTILINESTRING             | 无   | STRING       |                            |
|                           | MULTIPOLYGON                | 无   | STRING       |                            |
|                           | GEOMETRYCOLLECTION          | 任何几何类型的值的集合                                   | STRING       |                            |
| JSON类型                    | JSON                        | 无   | STRING       |                            |

目标实例为MaxCompute、Elasticsearch

| 类型 | 源实例中的数据类型 | 数值范围 | MaxCompute | Elasticsearch |
|----|-----------|------|------------|---------------|
|----|-----------|------|------------|---------------|

| 类型                     | 源实例中的数据类型                 | 数值范围  | MaxCompute       | Elasticsearch   |
|------------------------|---------------------------|---|------------------|---|
| 整数类型                   | BIT[(M)]                  | 1 ~ 64  | BOOLEAN   STRING | BOOLEAN   LONG<br><br><span style="background-color: #e0f2f7; padding: 5px;">? 说明 如果只有一个字节, 建议在Elasticsearch中使用BOOLEAN。</span>          |
|                        | TINYINT[(M)]              | -128 ~ 127  | BIGINT           | SHORT   |
|                        | TINYINT[(M)] [UNSIGNED]   | 0 ~ 255   | BIGINT           | INTEGER   |
|                        | SMALLINT[(M)]             | -32768 ~ 32767                                    | BIGINT           | SHORT   |
|                        | SMALLINT[(M)] [UNSIGNED]  | 0 ~ 65535   | BIGINT           | INTEGER   |
|                        | MEDIUMINT[(M)]            | -8388608 ~ 8388607                                | BIGINT           | INTEGER   |
|                        | MEDIUMINT[(M)] [UNSIGNED] | 0 ~ 16777215                                      | BIGINT           | INTEGER   |
|                        | INT[(M)]                  | -2147483648 ~ 2147483647                          | BIGINT           | INTEGER   |
|                        | INT[(M)] [UNSIGNED]       | 0 ~ 4294967295                                    | BIGINT           | LONG  |
|                        | BIGINT[(M)]               | -9223372036854775808 ~ 9223372036854775807        | BIGINT           | LONG  |
| BIGINT[(M)] [UNSIGNED] | 0 ~ 18446744073709551615  | BIGINT  | LONG             |   |
| 小数类型                   | DECIMAL[(M,D)]            | M: 0~65 ;<br>D: 0~30                              | DOUBLE           | DOUBLE<br><br><span style="background-color: #e0f2f7; padding: 5px;">? 说明 如果DECIMAL的值带有小数点, 那么为保障数据一致性, 建议在Elasticsearch中使用TEXT。</span> |
|                        | FLOAT(p)                  | 1.175494351E-38 ~ 3.402823466E+38                 | DOUBLE           | FLOAT   |
|                        | DOUBLE[(M,D)]             | 2.2250738585072014E-308 ~ 1.7976931348623157E+308 | DOUBLE           | DOUBLE  |

| 类型   | 源实例中的数据类型      | 数值范围   | MaxCompute | Elasticsearch  |
|------|----------------|--|------------|--|
| 时间类型 | DATE           | 1000-01-01~9999-12-31<br><br>② 说明 格式为YYYY-MM-DD (UTC时间)。   | DATETIME   | DATE<br><br>② 说明 格式为YYYY-MM-DD (UTC时间), 详情请参见date format的mapping定义。  |
|      | DATETIME(fsp)  | 1000-01-01 00:00:00.000000 ~ 9999-12-31 23:59:59.999999<br><br>② 说明 格式为YYYY-MM-DD hh:mm:ss[.fraction] (UTC时间)。 | DATETIME   | DATE<br><br>② 说明 DATE格式为yyyy-MM-dd'T'HH:mm:ss (UTC时间)。如果精度为微秒, 那么格式为yyyy-MM-dd'T'HH:mm:ss.S, 详情请参见date format的mapping定义。 |
|      | TIMESTAMP(fsp) | 1970-01-01 00:00:01.000000 ~ 2038-01-19 03:14:07.999999<br><br>② 说明 格式为YYYY-MM-DD hh:mm:ss[.fraction] (UTC时间)。 | DATETIME   | DATE<br><br>② 说明 DATE格式为yyyy-MM-dd'T'HH:mm:ss (UTC时间)。如果精度为微秒, 那么格式为yyyy-MM-dd'T'HH:mm:ss.S, 详情请参见date format的mapping定义。 |
|      | TIME(fsp)      | -838:59:59.000000 ~ 838:59:59.000000<br><br>② 说明 格式为hh:mm:ss[.fraction] (UTC时间)。                               | STRING     | DATE<br><br>② 说明 格式为YYYY-MM-DD (UTC时间), 详情请参见date format的mapping定义。  |

| 类型    | 源实例中的数据类型                   | 数值范围  | MaxCompute | Elasticsearch   |
|-------|-----------------------------|---|------------|---|
|       | YEAR(4)                     | 1901 ~ 2155, 或 0000.                          | STRING     | DATE<br><br> 说明 DATE 格式为yyyy (UTC 时间), 详情请参见date format的mapping定义。 |
| 字符串类型 | CHAR(M)                     | 0 ~ 255 字符                                    | STRING     | TEXT  |
|       | VARCHAR(M)                  | 0 ~ 65,535 字符                                 | STRING     | TEXT  |
|       | BINARY[(M)]                 | 0 ~ 255 字节                                    | STRING     | BINARY  |
|       | VARBINARY(M)                | 0 ~ 65,535 字节                                 | STRING     | BINARY  |
|       | TINYBLOB                    | 255 (2 <sup>8</sup> - 1) 字节                   | STRING     | BINARY  |
|       | TINYTEXT                    | 255 (2 <sup>8</sup> - 1) 字符                   | STRING     | TEXT  |
|       | BLOB                        | 65,535 (2 <sup>16</sup> - 1) 字节               | STRING     | BINARY  |
|       | TEXT                        | 65,535 (2 <sup>16</sup> - 1) 字符               | STRING     | TEXT  |
|       | MEDIUMBLOB                  | 16,777,215 (2 <sup>24</sup> - 1) 字节           | STRING     | BINARY  |
|       | MEDIUMTEXT                  | 16,777,215 (2 <sup>24</sup> - 1) 字符           | STRING     | TEXT  |
|       | LOB                         | 4,294,967,295 or 4GB (2 <sup>32</sup> - 1) 字节 | STRING     | BINARY  |
|       | LONGTEXT                    | 4,294,967,295 or 4GB (2 <sup>32</sup> - 1) 字符 | STRING     | TEXT  |
|       | ENUM('value1','value2',...) | 最多可包含65,535枚举值                                | STRING     | KEYWORD   |
|       | SET('value1','value2',..)   | 最多可包含64个元素                                    | STRING     | KEYWORD   |
|       | GEOMETRY                    | 任意几何类型的值                                      | STRING     | GEO_SHAPE   |
|       | POINT                       | 无   | STRING     | GEO_POINT   |
|       | LINestring                  | 无   | STRING     | GEO_SHAPE   |
|       | POLYGON                     | 无   | STRING     | GEO_SHAPE   |

| 类型<br>空间类型 | 源实例中的数据类型          | 数值范围        | MaxCompute | Elasticsearch  |
|------------|--------------------|-------------|------------|--|
|            | MULTIPOINT         | 无           | STRING     | GEO_SHAPE<br><br><span style="border: 1px solid #ADD8E6; padding: 5px;">? 说明 如果只有一个字节，建议在Elasticsearch中使用BOOLEAN。</span> |
|            | MULTILINESTRING    | 无           | STRING     | GEO_SHAPE  |
|            | MULTIPOLYGON       | 无           | STRING     | GEO_SHAPE  |
|            | GEOMETRYCOLLECTION | 任何几何类型的值的集合 | STRING     | GEO_SHAPE  |
| JSON类型     | JSON               | 无           | STRING     | OBJECT<br><br><span style="border: 1px solid #ADD8E6; padding: 5px;">? 说明 如果只有一个字节，建议在Elasticsearch中使用BOOLEAN。</span>    |

### 以Oracle为源的数据同步

当源实例为自建Oracle，且目标实例为异构数据库时（包括云原生数据仓库AnalyticDB PostgreSQL版），两者间的数据类型映射关系如下：

? 说明 如源实例中待同步数据范围超出DTS所支持的范围，则会导致目标实例中的所写入数据的精度降低。

| 类型   | Oracle的数据类型                                | 数值范围  | 云原生数据仓库AnalyticDB PostgreSQL版的数据类型              |
|------|--|---|---|
| 数字类型 | NUMBER(p,s)                                | 1 ~ 22 字节。<br>p代表精度位，取值范围是1 ~ 38。<br>s代表小数位，取值范围是-84 ~ 127。 | DECIMAL   TINYINT   SMALLINT   INTEGER   BIGINT |
|      | FLOAT(p)                                   | 1 ~ 22 字节。<br>p代表指针变量，取值范围是1 ~ 126 bit。                     | DOUBLE PRECISION                                |
|      | BINARY_FLOAT                               | 32-bit的浮点数，即4字节。  | DOUBLE PRECISION                                |
|      | BINARY_DOUBLE                              | 64-bit的浮点数，即8字节。  | DOUBLE PRECISION                                |
|      | DATE                                       | 无   | TIMESTAMP(0)                                    |
|      | TIMESTAMP [(fractional_seconds_precision)] | 无   | TIMESTAMP                                       |

| 类型      | Oracle的数据类型   | 数值范围   | 云原生数据库AnalyticDB PostgreSQL版的数据类型 |
|---------|---|--|-----------------------------------|
| 日期类型    | TIMESTAMP [(fractional_seconds_precision)] WITH TIME ZONE                 | 无  | TIMESTAMP WITH TIME ZONE          |
|         | TIMESTAMP [(fractional_seconds_precision)] WITH LOCAL TIME ZONE           | 无  | TIMESTAMP WITH TIME ZONE          |
|         | INTERVAL YEAR [(year_precision)] TO MONTH                                 | 无  | VARCHAR(32)                       |
|         | INTERVAL DAY [(day_precision)] TO SECOND [(fractional_seconds_precision)] | 无  | VARCHAR(32)                       |
| 字符串类型   | CHAR [(size [BYTE   CHAR])]   | 2000 字节。   | CHAR                              |
|         | NCHAR[(size)]   | 2000 字节。   | VARCHAR                           |
|         | VARCHAR2(size [BYTE   CHAR])  | 当MAX_STRING_SIZE = EXTENDED时，最大长度为32767字节；<br>当MAX_STRING_SIZE = STANDARD，最大长度为4000字节。 | VARCHAR                           |
|         | NVARCHAR2(size)   | 当MAX_STRING_SIZE = EXTENDED时，最大长度为32767字节；<br>当MAX_STRING_SIZE = STANDARD，最大长度为4000字节。 | VARCHAR                           |
|         | LONG  | 最大长度2G (2 <sup>31</sup> -1)。   | TEXT                              |
|         | RAW(size)   | 最大长度32767 字节或2000 字节。  | BYTEA                             |
|         | LONG RAW  | 最大长度2G。  | BYTEA                             |
|         | CLOB  | 最大长度 (4 GB - 1) *DB_BLOCK_SIZE。  | TEXT                              |
|         | NCLOB   | 最大长度(4 GB - 1) *DB_BLOCK_SIZE。   | TEXT                              |
| BLOB    | 最大长度 (4 GB - 1) *DB_BLOCK_SIZE。   | BYTEA  |                                   |
| BFILE   | 4G。   | 不支持  |                                   |
| JSON类型  | JSON  | 最大长度32MB。  | JSON                              |
| ROWID类型 | ROWID   | 64字符。  | OID                               |

| 类型   | Oracle的数据类型 | 数值范围 | 云原生数据库AnalyticDB PostgreSQL版的数据类型 |
|------|-------------|------|-----------------------------------|
| 空间类型 | 需要自定义       | 不支持  |                                   |

**说明** 对于云原生数据库AnalyticDB PostgreSQL版不支持字段类型，DTS会将其直接转为BYTEA，如果转换不了则将数据内容置为NULL。

### 以PostgreSQL为源的数据同步

当源实例为PostgreSQL（包括自建PostgreSQL、RDS PostgreSQL），且目标实例为异构数据库时（包括云原生数据库AnalyticDB PostgreSQL版），两者间的数据类型映射关系如下：

**说明** 如源实例中待同步数据范围超出DTS所支持的范围，则会导致目标实例中的所写入数据的精度降低。

| 类型      | PostgreSQL的数据类型                     | 数值范围  | 云原生数据库AnalyticDB PostgreSQL版的数据类型   |
|---------|-------------------------------------|---|-------------------------------------|
| 整数类型    | SMALLINT                            | -32767至+32767                               | SMALLINT                            |
|         | INTEGER                             | -2147483648至+2147483647                     | INTEGER                             |
|         | BIGINT                              | -9223372036854775808至+9223372036854775807   | BIGINT                              |
| 小数类型    | DECIMAL                             | 最高小数点前131072位，以及小数点后16383位                  | DECIMAL                             |
|         | NUMERIC                             | 最高小数点前131072位，以及小数点后16383位                  | NUMERIC                             |
|         | REAL                                | 6位十进制精度                                     | REAL                                |
|         | DOUBLE PRECISION                    | 15位十进制精度                                    | DOUBLE PRECISION                    |
| 货币类型    | MONEY                               | -92233720368547758.08至+92233720368547758.07 | MONEY                               |
| 字符串类型   | CHARACTER VARYING(n)                | 无   | CHARACTER VARYING(n)                |
|         | CHARACTER(n)                        | 无   | CHARACTER(n)                        |
|         | TEXT                                | 无   | TEXT                                |
|         | CHAR                                | 默认为1字节                                      | CHAR                                |
|         | NAME                                | 最大长度为64字节                                   | NAME                                |
| 文本搜索类型  | TSQUERY                             | 表示一个文本查询                                    | TEXT                                |
|         | TSVECTOR                            | 表示一个为文本搜索优化的形式下的文档                          | TEXT                                |
| 二进制数据类型 | BYTEA                               | 1或4字节外加真正的二进制串                              | BYTEA                               |
|         | TIMESTAMP [(p)] [WITHOUT TIME ZONE] | 表示日期和时间（无时区），占用8字节                          | TIMESTAMP [(p)] [WITHOUT TIME ZONE] |

| 类型     | PostgreSQL的数据类型                | 数值范围                        | 云原生数据库AnalyticDB PostgreSQL版的数据类型 |
|--------|--------------------------------|-----------------------------|-----------------------------------|
| 日期类型   | TIMESTAMP [(p)] WITH TIME ZONE | 表示日期和时间（有时区），占用8字节          | TIMESTAMP [(p)] WITH TIME ZONE    |
|        | DATE                           | 表示日期，占用4字节                  | DATE                              |
|        | TIME [(p)] [WITHOUT TIME ZONE] | 表示时间（无时区），占用8字节             | TIME [(p)] [WITHOUT TIME ZONE]    |
|        | TIME [(p)] WITH TIME ZONE      | 表示时间（有时区），占用12字节            | TIME [(p)] WITH TIME ZONE         |
|        | interval [ fields ] [(p)]      | 表示时间间隔，占用16字节               | interval [ fields ] [(p)]         |
| 布尔类型   | BOOLEAN                        | 1字节                         | BOOLEAN                           |
| 枚举类型   | 自定义枚举                          | 无                           | VARCHAR(128)                      |
| 空间类型   | POINT                          | 平面上的点，最大长度为16字节             | POINT                             |
|        | LINE                           | 无限长的线，最大长度为32字节             | LINE                              |
|        | LSEG                           | 有限线段，最大长度为32字节              | LSEG                              |
|        | BOX                            | 矩形框，最大长度为32字节               | BOX                               |
|        | PATH                           | 路径，最大长度为16+16n字节            | PATH                              |
|        | POLYGON                        | 多边形（类似于封闭路径），最大长度为40+16n字节  | POLYGON                           |
|        | CIRCLE                         | 圆，最大长度为24字节                 | CIRCLE                            |
| 网络地址类型 | CIDR                           | IPv4和IPv6网络，最大长度为7或19字节     | CIDR                              |
|        | INET                           | IPv4和IPv6主机以及网络，最大长度为7或19字节 | INET                              |
|        | MACADDR                        | MAC地址，最大长度为6字节              | MACADDR                           |
|        | MACADDR8                       | MAC地址（EUI-64格式），最大长度为8字节    | MACADDR8                          |
| 位串类型   | Bit (n)                        | 无                           | Bit (n)                           |
|        | BIT VARYING (n)                | 无                           | BIT VARYING (n)                   |
| UUID类型 | UUID                           | 无                           | VARCHAR(64)                       |
| XML类型  | XML                            | 无                           | XML                               |
| JSON类型 | JSON                           | 无                           | JSON                              |
|        | JSONB                          | 无                           | JSONB                             |

### 以SQL Server为源的数据同步

当源实例为SQL Server（包括自建SQL Server、RDS SQL Server），且目标实例为异构数据库时（包括云原生数据仓库AnalyticDB MySQL版、云原生数据仓库AnalyticDB PostgreSQL版），两者间的数据类型映射关系如下：

 **说明** 如源实例中待同步数据范围超出DTS所支持的范围，则会导致目标实例中的所写入数据的精度降低。

| 类型   | SQL Server的数据类型      | 数值范围  | 云原生数据仓库AnalyticDB MySQL版的数据类型 | 云原生数据仓库AnalyticDB PostgreSQL版的数据类型 |
|------|----------------------|---|-------------------------------|------------------------------------|
| 整数类型 | BIT                  | 可以取值为1、0或NULL的INTEGER数据类型   | BOOLEAN                       | BIT(1)                             |
|      | TINYINT              | 0至255   | TINYINT                       | SMALLINT                           |
|      | SMALLINT             | $-2^{15}$ (-32768) 至 $2^{15}-1$ (32767)                             | SMALLINT                      | SMALLINT                           |
|      | INT                  | $-2^{31}$ (-2147483648) 至 $2^{31}-1$ (2147483647)                   | INTEGER                       | INTEGER                            |
|      | BIGINT               | $-2^{63}$ (-9223372036854775808) 至 $2^{63}-1$ (9223372036854775807) | BIGINT                        | BIGINT                             |
| 小数类型 | NUMERIC[ (p[ ,s] ) ] | $-10^{38+1}$ 至 $10^{38}-1$ ;<br>$1 \leq p \leq 38$                  | DECIMAL                       | DECIMAL                            |
|      | DECIMAL[ (p[ ,s] ) ] | $-10^{38+1}$ 至 $10^{38}-1$ ;<br>$1 \leq p \leq 38$                  | DECIMAL                       | DECIMAL                            |
|      | FLOAT                | $-1.79E+308$ 至 $-2.23E-308$ 、0 以及 $2.23E-308$ 至 $1.79E+308$         | DOUBLE                        | DOUBLE PRECISION                   |
|      | REAL                 | $-3.40E+38$ 至 $-1.18E-38$ 、0 以及 $1.18E-38$ 至 $3.40E+38$             | FLOAT                         | REAL                               |
| 货币类型 | MONEY                | -922,337,203,685,477.5808至922,337,203,685,477.5807                  | DECIMAL(19, 4)                | DECIMAL(19, 4)                     |
|      | SMALLMONEY           | -214,748.3648至214,748.3647  | DECIMAL(10, 4)                | DECIMAL(10, 4)                     |
|      | DATE                 | 0001-01-01至9999-12-31   | DATE                          | DATE                               |

| 类型    | SQL Server的数据类型                                   | 数值范围   | 云原生数据库 AnalyticDB MySQL版的数据类型 | 云原生数据库 AnalyticDB PostgreSQL版的数据类型 |
|-------|---|--|-------------------------------|------------------------------------|
| 日期类型  | DATETIME  | 日期范围为1753年1月1日至9999年12月31日；<br>时间范围为00:00:00至23:59:59.997                              | DATETIME                      | TIMESTAMP(3) WITHOUT TIME ZONE     |
|       | DATETIME2 [ (fractional seconds precision) ]      | 日期范围为公元1年1月1日至公元9999年12月31日；<br>时间范围为00:00:00至23:59:59.999999                          | DATETIME                      | TIMESTAMP(7) WITHOUT TIME ZONE     |
|       | DATETIMEOFFSET [ (fractional seconds precision) ] | 日期范围公元1年1月1日至公元9999年12月31日；<br>时间范围为00:00:00至23:59:59.9999999；<br>时区偏移量范围-14:00至+14:00 | TIMESTAMP                     | TIMESTAMP(7) WITH TIME ZONE        |
|       | SMALLDATETIME                                     | 秒始终为零 (:00)，并且不带秒小数部分  | DATETIME                      | TIMESTAMP WITHOUT TIME ZONE        |
|       | TIME [ (fractional second scale) ]                | 00:00:00.0000000至23:59:59.9999999  | TIME                          | TIME(7) WITH TIME ZONE             |
| 字符串类型 | BINARY [ ( n ) ]                                  | n取值范围为1~8,000  | VARBINARY                     | BYTEA                              |
|       | VARBINARY [ ( n   max ) ]                         | n取值范围为1~8,000；max代表最大存储大小是2 <sup>31</sup> -1个字节  | VARBINARY                     | BYTEA                              |
|       | CHAR [ ( n ) ]                                    | n取值范围为1~8,000，存储大小为n个字节  | VARCHAR                       | CHARACTER                          |
|       | VARCHAR [ ( n   max ) ]                           | n取值范围为1~8,000；max代表最大存储大小是 2 <sup>31</sup> -1个字节（2GB）                                  | VARCHAR                       | CHARACTER                          |
|       | NCHAR [ ( n ) ]                                   | 以双字节为单位，n取值范围为1至4,000，存储大小为 n字节的两倍   | VARCHAR                       | CHARACTER VARYING                  |
|       | NVARCHAR [ ( n   max ) ]                          | 以双字节为单位，n取值范围为1至4,000，max代表最大存储大小是2 <sup>30</sup> -1个字（2 GB）                           | VARCHAR                       | CHARACTER VARYING                  |

| 类型      | SQL Server的数据类型                                       | 数值范围  | 云原生数据库AnalyticDB MySQL版的数据类型 | 云原生数据库AnalyticDB PostgreSQL版的数据类型 |
|---------|---|---|------------------------------|-----------------------------------|
|         | NTEXT   | 长度可变的Unicode 数据，字符串最大长度为2 <sup>30</sup> -1 (1073741823) 个字节 | VARCHAR                      | TEXT                              |
|         | TEXT  | 字符串最大长度为2 <sup>31</sup> -1 (2147483647) 个字节                 | VARCHAR                      | TEXT                              |
|         | IMAGE   | 长度可变的二进制数据，从0到2 <sup>31</sup> -1 (2147483647) 个字节。          | VARBINARY                    | BYTEA                             |
| 空间和几何类型 | GEOGRAPHY   | 无   | VARCHAR                      | 不支持                               |
|         | GEOMETRY  | 无   | VARCHAR                      | 不支持                               |
| XML类型   | XML ( [ CONTENT   DOCUMENT ] xml_schema_collecti on ) | 无   | VARCHAR                      | XML                               |
| 其它类型    | UNIQUEIDENTIFIER                                      | 无   | VARCHAR                      | CHARACTER(36)                     |
|         | SQL_VARIANT   | 无   | 不支持                          | 不支持                               |
|         | HIERARCHYID   | 无   | 不支持                          | 不支持                               |
|         | SYSNAME   | 无   | VARCHAR                      | CHARACTER VARYING(128)            |

### 以DB2 for LUW为源的数据同步

当源实例为DB2 for LUW，且目标实例为异构数据库时（包括MySQL），两者间的数据类型映射关系如下：

 **说明** 如源实例中待同步数据范围超出DTS所支持的范围，则会导致目标实例中的所写入数据的精度降低。

| 类型   | DB2 for LUW数据类型                           | 数值范围  | MySQL的数据类型 |
|------|---|---|------------|
| 整数类型 | SMALLINT                                  | -32,768~+32,767                                       | SMALLINT   |
|      | INTEGER                                   | -2,147,483,648~+2,147,483,647                         | INT        |
|      | BIGINT                                    | -9,223,372,036,854,775,808~+9,223,372,036,854,775,807 | BIGINT     |
|      | DECIMAL(precision-integer, scale-integer) | p<= 38  | DECIMAL    |

| 类型    | DB2 for LUW数据类型                 | 数值范围  | MySQL的数据类型         |
|-------|---------------------------------|---|--------------------|
| 小数类型  | FLOAT(integer)                  | 取值范围为1~53，取值在1~24表示为单精度，取值在25~53表示为双精度                                      | FLOAT              |
|       | DECFLOAT(precision-integer)     | 无   | DECIMAL(65,10)     |
| 日期类型  | DATE                            | 0001-01-01~9999-12-31   | DATE               |
|       | TIME                            | 00:00:00~24:00:00   | TIME               |
|       | TIMESTAMP(integer)              | 0001-01-01-00.00.00.000000000000~9999-12-31-24.00.00.000000000000;0<=p<= 12 | DATETIME           |
| 字符串类型 | CHARACTER(integer)              | 254   | CHAR   VARCHAR     |
|       | VARCHAR(integer)                | 32,672  | VARCHAR            |
|       | CHARACTER(integer) FOR BIT DATA | 254   | BLOB               |
|       | CLOB                            | 2,147,483,647   | LONGTEXT           |
|       | GRAPHIC(integer)                | 127   | CHAR(length*4)     |
|       | VARGRAPHIC(integer)             | 16,336  | CHAR(length*4)     |
|       | DBCLOB(integer)                 | 1,073,741,823   | VARCHAR   LONGTEXT |
|       | BLOB                            | 2,147,483,647   | LONGBLOB           |
| 其它类型  | XML                             | 2,147,483,647   | VARCHAR   LONGTEXT |

### 以Db2 for i为源的数据同步

当源实例为DB2 for i，且目标实例为异构数据库时（包括MySQL），两者间的数据类型映射关系如下：

 **说明** 如源实例中待同步数据范围超出DTS所支持的范围，则会导致目标实例中的所写入数据的精度降低。

| 类型   | Db2 for i数据类型                             | 数值范围  | MySQL中的数据类型 |
|------|---|---|-------------|
| 整数类型 | SMALLINT                                  | -32,768~+32,767                                       | SMALLINT    |
|      | INTEGER                                   | -2,147,483,648~+2,147,483,647                         | INT         |
|      | BIGINT                                    | -9,223,372,036,854,775,808~+9,223,372,036,854,775,807 | BIGINT      |
|      | DECIMAL(precision-integer, scale-integer) | p<=63   | DECIMAL     |
|      | NUMERIC                                   | 无   | DECIMAL     |

| 数据类型  | Db2 for i数据类型               | 数值范围  | MySQL中的数据类型        |
|-------|-----------------------------|---|--------------------|
|       | FLOAT(integer)              | 无   | FLOAT              |
|       | DECFLOAT(precision-integer) | 无   | DECIMAL(65,10)     |
| 日期类型  | DATE                        | 0001-01-0~9999-12-31  | DATE               |
|       | TIME                        | 00:00:00~24:00:00   | TIME               |
|       | TIMESTAMP(integer)          | 0001-01-01-00.00.00.000000000000~9999-12-31-24.00.00.000000000000; 0 <=p <=12 | DATETIME           |
| 字符串类型 | CHAR(integer)               | 32,765  | CHAR   VARCHAR     |
|       | VARCHAR(integer)            | 32,739  | VARCHAR            |
|       | CHAR(integer) FOR BIT DATA  | 无   | BLOB               |
|       | CLOB                        | 2,147,483,647   | LONGTEXT           |
|       | GRAPHIC(integer)            | 16,382  | CHAR               |
|       | VARGRAPHIC(integer)         | 16,369  | VARCHAR            |
|       | DBCLOB(integer)             | 1,073,741,823   | LONGTEXT           |
|       | BINARY                      | 32,765  | BINARY             |
|       | VARBIN                      | 32,739  | VARBINARY          |
|       | BLOB                        | 2,147,483,647   | LOB                |
| 其它类型  | DATALINK                    | 无   | VARCHAR   LONGTEXT |
|       | ROWID                       | 40  | VARCHAR   LONGTEXT |
|       | XML                         | 2,147,483,647   | VARCHAR   LONGTEXT |

# 3. 注意事项及限制

## 3.1. 同步时源库为MySQL的注意事项及限制

如果同步的源数据库类型为MySQL，如自建MySQL、RDS MySQL，您需要在配置具体的同步任务前，参考本文的注意事项及限制，以保障数据同步任务的正常运行。

### 源库为MySQL的同步方案概览

根据如下同步方案，查看同步任务的注意事项及限制：

#### 说明

DTS默认同步到目标数据库中时会取消外键约束，因此源数据库的级联、删除等操作不会同步到如下目标数据库：

- MySQL (RDS MySQL、自建MySQL)
- PolarDB MySQL
- 云原生数据仓库AnalyticDB MySQL
- 云原生数据仓库AnalyticDB PostgreSQL
- Elasticsearch
- ClickHouse集群
- Tablestore
- PolarDB-X 2.0
- Oracle
- PostgreSQL

- MySQL间的同步
- MySQL同步至PolarDB MySQL
- MySQL同步至云原生数据仓库AnalyticDB MySQL
- MySQL同步至云原生数据仓库AnalyticDB PostgreSQL
- MySQL同步至阿里云流式数据服务DataHub
- MySQL同步至Elasticsearch
- MySQL同步至MaxCompute
- MySQL同步至ClickHouse集群
- MySQL同步至Tablestore
- MySQL同步至阿里云消息队列Kafka版、自建Kafka集群
- MySQL同步至PolarDB-X
- MySQL同步至Oracle
- MySQL同步至PostgreSQL

### MySQL间的同步

如果同步的目标数据库类型为MySQL，如RDS MySQL、自建MySQL，具体注意事项及限制如下：

- MySQL间单向同步

| 类型 | 说明 |
|----|----|
|----|----|

| 类型     | 说明   |
|--------|--|
| 源库限制   | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> <li>Binlog日志：                             <ul style="list-style-type: none"> <li>需开启Binlog，开启方法请参见<a href="#">设置实例参数</a>，并且binlog_row_image为full。否则预检查阶段提示报错，且无法成功启动数据同步任务。</li> </ul> </li> </ul> <div style="border: 1px solid #ccc; padding: 10px; margin: 10px 0;"> <p> <b>注意</b></p> <ul style="list-style-type: none"> <li>如源实例为自建MySQL，则需开启Binlog，并且binlog_format为row、binlog_row_image为full。</li> <li>如源实例自建MySQL是双主集群（两者互为主从），为保障DTS能获取全部的Binlog日志，则您需开启参数log_slave_updates。具体操作请参见<a href="#">自建MySQL创建账号并设置binlog</a>。</li> </ul> </div> <ul style="list-style-type: none"> <li>如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在全量同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。如源为RDS MySQL，具体操作请参见<a href="#">本地日志（Binlog）</a>。</li> </ul> |
| 其他限制   | <ul style="list-style-type: none"> <li>建议源和目标库版本保持一致，或者从低版本同步到高版本以保障兼容性。如为高版本同步至低版本，可能存在数据库兼容性问题。</li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>如同步对象为单个或多个表（非整库），那么在数据同步时，勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，请参见<a href="#">不锁表结构变更</a>。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul>   |
| 其他注意事项 | <p>当源库为自建MySQL时</p> <ul style="list-style-type: none"> <li>在同步时，如果源库进行主备切换，将会导致同步任务失败。</li> <li>由于DTS的延迟时间是根据同步到目标库最后一条数据的时间戳和当前时间戳对比得出，源库长时间未执行DML操作可能导致延迟信息不准确。如果任务显示的延迟时间过大，您可以在源库执行一个DML操作来更新延迟信息。</li> </ul> <div style="border: 1px solid #ccc; padding: 10px; margin: 10px 0;"> <p> <b>说明</b> 如果同步对象选择为整库，您还可以创建心跳表，心跳表每秒定期更新或者写入数据。</p> </div> <ul style="list-style-type: none"> <li>DTS会在源库定时执CREATE DATABASE IF NOT EXISTS `test`命令以推进Binlog位点。</li> </ul>   |

● MySQL间双向同步

| 类型 | 说明 |
|----|----|
|    |    |

| 类型      | 说明  |
|---------|---|
| 源和目标库限制 | <ul style="list-style-type: none"> <li>◦ 待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>◦ 如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> <li>◦ Binlog日志：             <ul style="list-style-type: none"> <li>■ 需开启Binlog，开启方法请参见<a href="#">设置实例参数</a>，并且binlog_row_image为full。否则预检查阶段提示报错，且无法成功启动数据同步任务。</li> </ul> <div style="border: 1px solid #ccc; padding: 10px; margin: 10px 0;"> <p> <b>注意</b></p> <ul style="list-style-type: none"> <li>■ 如源实例为自建MySQL，则需开启Binlog，并且binlog_format为row、binlog_row_image为full。</li> <li>■ 如源实例自建MySQL是双主集群（两者互为主从），为保障DTS能获取全部的Binlog日志，则您需开启参数log_slave_updates。具体操作请参见<a href="#">为自建MySQL创建账号并设置binlog</a>。</li> </ul> </div> <li>■ 如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在全量同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。如源为RDS MySQL，具体操作请参见<a href="#">本地日志（Binlog）</a>。</li> </li></ul> |
| 其他限制    | <ul style="list-style-type: none"> <li>◦ 建议源和目标库版本保持一致，或者从低版本同步到高版本以保障兼容性。如为高版本同步至低版本，可能存在数据库兼容性问题。</li> <li>◦ 双向同步任务时，DTS会在目标库中建立数据库dts，以避免数据循环同步，任务运行期间，请勿修改该库。</li> <li>◦ 如双向同步任务的源实例或目标实例位于海外地域，则仅支持同地域的双向同步，不支持跨地域的双向同步。例如，支持日本地域间的双向同步，不支持日本地域与法兰克福地域间的双向同步。</li> <li>◦ 执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>◦ 全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>◦ 如同步对象为单个或多个表（非整库），那么在数据同步时，勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。<br/>您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，请参见<a href="#">不锁表结构变更</a>。</li> <li>◦ 在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> <li>◦ DTS默认同步到目标数据库中时会取消外键约束，因此源数据库的级联、删除等操作不会同步到目标数据库。</li> <li>◦ 若DTS正向和反向同步任务均同步了某个表，且正向同步任务同步了该表的全量和增量数据，则反向同步任务只同步该表的增量数据。</li> </ul>   |

| 类型   | 说明   |
|------|--|
| 特殊情况 | <p>当源库为自建MySQL时</p> <ul style="list-style-type: none"> <li>在同步时，如果源库进行主备切换，将会导致同步任务失败。</li> <li>由于DTS的延迟时间是根据同步到目标库最后一条数据的时间戳和当前时间戳对比得出，源库长时间未执行DML操作可能导致延迟信息不准确。如果任务显示的延迟时间过大，您可以在源库执行一个DML操作来更新延迟信息。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin: 10px 0;"> <p> <b>说明</b> 如果同步对象选择为整库，您还可以创建心跳表，心跳表每秒定期更新或者写入数据。</p> </div> <ul style="list-style-type: none"> <li>DTS会在源库定时执行CREATE DATABASE IF NOT EXISTS `test`命令以推进Binlog位点。</li> </ul> |

## MySQL同步至PolarDB MySQL

- MySQL单向同步至PolarDB MySQL

| 类型   | 说明  |
|------|---|
| 源库限制 | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> <li>Binlog日志：             <ul style="list-style-type: none"> <li>需开启Binlog，开启方法请参见<a href="#">设置实例参数</a>，并且binlog_row_image为full。否则预检查阶段提示报错，且无法成功启动数据同步任务。</li> </ul> </li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin: 10px 0;"> <p> <b>注意</b></p> <ul style="list-style-type: none"> <li>如源实例为自建MySQL，则需开启Binlog，并且binlog_format为row、binlog_row_image为full。</li> <li>如源实例自建MySQL是双主集群（两者互为主从），为保障DTS能获取全部的Binlog日志，则您需开启参数log_slave_updates。具体操作请参见<a href="#">自建MySQL创建账号并设置binlog</a>。</li> </ul> </div> <ul style="list-style-type: none"> <li>如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在全量同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。如源为RDS MySQL，具体操作请参见<a href="#">本地日志（Binlog）</a>。</li> </ul> |
| 其他限制 | <ul style="list-style-type: none"> <li>建议源和目标库版本保持一致，或者从低版本同步到高版本以保障兼容性。如为高版本同步至低版本，可能存在数据库兼容性问题。</li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>如同步对象为单个或多个表（非整库），那么在数据同步时，勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，请参见<a href="#">不锁表结构变更</a>。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul>  |

| 类型   | 说明  |
|------|---|
| 特殊情况 | <p>当源库为自建MySQL时</p> <ul style="list-style-type: none"> <li>在同步时，如果源库进行主备切换，将会导致同步任务失败。</li> <li>由于DTS的延迟时间是根据同步到目标库最后一条数据的时间戳和当前时间戳对比得出，源库长时间未执行DML操作可能导致延迟信息不准确。如果任务显示的延迟时间过大，您可以在源库执行一个DML操作来更新延迟信息。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin: 5px 0;"> <p> <b>说明</b> 如果同步对象选择为整库，您还可以创建心跳表，心跳表每秒定期更新或者写入数据。</p> </div> <ul style="list-style-type: none"> <li>DTS会在源库定时执行CREATE DATABASE IF NOT EXISTS `test`命令以推进Binlog位点。</li> </ul> |

● MySQL双向同步至PolarDB MySQL

| 类型      | 说明  |
|---------|---|
| 源和目标库限制 | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> <li>Binlog日志：             <ul style="list-style-type: none"> <li>MySQL：                 <ul style="list-style-type: none"> <li>需开启Binlog，开启方法请参见<a href="#">设置实例参数</a>，并且binlog_row_image为full。否则预检查阶段提示报错，且无法成功启动数据同步任务。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin: 5px 0;"> <p> <b>注意</b></p> <ul style="list-style-type: none"> <li>如源实例为自建MySQL，则需开启Binlog，并且binlog_format为row、binlog_row_image为full。</li> <li>如源实例自建MySQL是双主集群（两者互为主从），为保障DTS能获取全部的Binlog日志，则您需开启参数log_slave_updates。具体操作请参见<a href="#">为自建MySQL创建账号并设置binlog</a>。</li> </ul> </div> </li> <li>PolarDB MySQL：                 <ul style="list-style-type: none"> <li>需开启Binlog，并且loose_polar_log_bin为on。否则预检查阶段显示报错提示，且无法成功启动数据同步任务。</li> </ul> </li> <li>如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在全量同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。如源为RDS MySQL，具体操作请参见<a href="#">本地日志（Binlog）</a>。</li> </ul> </li> </ul> |

| 类型   | 说明  |
|------|---|
| 其他限制 | <ul style="list-style-type: none"> <li>◦</li> <li>◦ 双向同步任务时，DTS会在目标库中建立数据库dts，以避免数据循环同步，任务运行期间，请勿修改该库。</li> <li>◦ 如双向同步任务的源实例或目标实例位于海外地域，则仅支持同地域的双向同步，不支持跨地域的双向同步。例如，支持日本地域间的双向同步，不支持日本地域与法兰克福地域间的双向同步。</li> <li>◦</li> <li>◦</li> <li>◦</li> <li>◦ 在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> <li>◦ DTS默认同步到目标数据库中时会取消外键约束，因此源数据库的级联、删除等操作不会同步到目标数据库。</li> <li>◦ 若DTS正向和反向同步任务均同步了某个表，且正向同步任务同步了该表的全量和增量数据，则反向同步任务只同步该表的增量数据。</li> </ul>  |
| 特殊情况 | <p>当源库为自建MySQL时</p> <ul style="list-style-type: none"> <li>◦ 在同步时，如果源库进行主备切换，将会导致同步任务失败。</li> <li>◦ 由于DTS的延迟时间是根据同步到目标库最后一条数据的时间戳和当前时间戳对比得出，源库长时间未执行DML操作可能导致延迟信息不准确。如果任务显示的延迟时间过大，您可以在源库执行一个DML操作来更新延迟信息。</li> </ul> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin: 5px 0;"> <p> <b>说明</b> 如果同步对象选择为整库，您还可以创建心跳表，心跳表每秒定期更新或者写入数据。</p> </div> <ul style="list-style-type: none"> <li>◦ DTS会在源库定时执CREATE DATABASE IF NOT EXISTS `test`命令以推进Binlog位点。</li> </ul> |

### MySQL同步至云原生数据仓库AnalyticDB MySQL

| 类型 | 说明 |
|----|----|
|    |    |

| 类型   | 说明  |
|------|---|
| 源库限制 | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出現重复数据。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> <li>Binlog日志：             <ul style="list-style-type: none"> <li>需开启Binlog，开启方法请参见<a href="#">设置实例参数</a>，并且binlog_row_image为full。否则预检查阶段提示报错，且无法成功启动数据同步任务。</li> </ul> </li> </ul> <div style="border: 1px solid #ccc; padding: 10px; margin: 10px 0;"> <p> <b>注意</b></p> <ul style="list-style-type: none"> <li>如源实例为自建MySQL，则需开启Binlog，并且binlog_format为row、binlog_row_image为full。</li> <li>如源实例自建MySQL是双主集群（两者互为主从），为保障DTS能获取全部的Binlog日志，则您需开启参数log_slave_updates。具体操作请参见<a href="#">自建MySQL创建账号并设置binlog</a>。</li> </ul> </div> <ul style="list-style-type: none"> <li>如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在全量同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。如源为RDS MySQL，具体操作请参见<a href="#">本地日志（Binlog）</a>。</li> <li>同步期间，请勿执行修改主键和添加注释的DDL操作（如 <code>ALTER TABLE table_name COMMENT='表的注释';</code>），否则数据同步过程中会导致执行该DDL失败。</li> </ul> |
| 其他限制 | <ul style="list-style-type: none"> <li>暂不支持同步前缀索引，如果源库存在前缀索引可能导致数据同步失败。</li> <li>由于云原生数据仓库AnalyticDB MySQL版本本身的使用限制，当云原生数据仓库AnalyticDB MySQL版集群中的节点磁盘空间使用量超过80%，会导致DTS任务异常，产生延迟。请提前根据待同步的对象预估所需空间，确保目标集群具备充足的存储空间。</li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>如同步对象为单个或多个表（非整库），那么在数据同步时，勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，请参见<a href="#">不锁表结构变更</a>。</li> <li>对于表级别的数据同步，如果除DTS以外的数据没有写入目标库云原生数据仓库AnalyticDB MySQL版，您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，请参见<a href="#">不锁表结构变更</a>。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul>  |
| 特殊情况 | <p>当源库为自建MySQL时</p> <ul style="list-style-type: none"> <li>在同步时，如果源库进行主备切换，将会导致同步任务失败。</li> <li>由于DTS的延迟时间是根据同步到目标库最后一条数据的时间戳和当前时间戳对比得出，源库长时间未执行DML操作可能导致延迟信息不准确。如果任务显示的延迟时间过大，您可以在源库执行一个DML操作来更新延迟信息。</li> </ul> <div style="border: 1px solid #ccc; padding: 10px; margin: 10px 0;"> <p> <b>说明</b> 如果同步对象选择为整库，您还可以创建心跳表，心跳表每秒定期更新或者写入数据。</p> </div> <ul style="list-style-type: none"> <li>DTS会在源库定时执行CREATE DATABASE IF NOT EXISTS `test`命令以推进Binlog位点。</li> </ul>   |

## MySQL同步至云原生数据仓库AnalyticDB PostgreSQL

| 类型   | 说明   |
|------|--|
| 源库限制 | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> <li>Binlog日志：             <ul style="list-style-type: none"> <li>需开启Binlog，开启方法请参见<a href="#">设置实例参数</a>，并且binlog_row_image为full。否则预检查阶段提示报错，且无法成功启动数据同步任务。</li> </ul> </li> </ul> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 10px; margin: 10px 0;"> <p> <b>注意</b></p> <ul style="list-style-type: none"> <li>如源实例为自建MySQL，则需开启Binlog，并且binlog_format为row、binlog_row_image为full。</li> <li>如源实例自建MySQL是双主集群（两者互为主从），为保障DTS能获取全部的Binlog日志，则您需开启参数log_slave_updates。具体操作请参见<a href="#">自建MySQL创建账号并设置binlog</a>。</li> </ul> </div> <ul style="list-style-type: none"> <li>如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在全量同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。如源为RDS MySQL，具体操作请参见<a href="#">本地日志（Binlog）</a>。</li> <li>同步期间，请勿执行修改主键和添加注释的DDL操作（如 <code>ALTER TABLE table_name COMMENT='表的注释';</code>），否则数据同步过程中会导致执行该DDL失败。</li> </ul> |
| 其他限制 | <ul style="list-style-type: none"> <li>同步对象要求：             <ul style="list-style-type: none"> <li>仅支持表级同步。</li> <li>不支持BIT、VARBIT、GEOMETRY、ARRAY、UUID、TSQUERY、TSVECTOR、TXID_SNAPSHOT类型的数据同步。</li> <li>暂不支持同步前缀索引，如果源库存在前缀索引可能导致数据同步失败。</li> </ul> </li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>如同步对象为单个或多个表（非整库），那么在数据同步时，勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，请参见<a href="#">不锁表结构变更</a>。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul>  |
| 特殊情况 | <p>当源库为自建MySQL时</p> <ul style="list-style-type: none"> <li>在同步时，如果源库进行主备切换，将会导致同步任务失败。</li> <li>由于DTS的延迟时间是根据同步到目标库最后一条数据的时间戳和当前时间戳对比得出，源库长时间未执行DML操作可能导致延迟信息不准确。如果任务显示的延迟时间过大，您可以在源库执行一个DML操作来更新延迟信息。</li> </ul> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 10px; margin: 10px 0;"> <p> <b>说明</b> 如果同步对象选择为整库，您还可以创建心跳表，心跳表每秒定期更新或者写入数据。</p> </div> <ul style="list-style-type: none"> <li>DTS会在源库定时执CREATE DATABASE IF NOT EXISTS `test`命令以推进Binlog位点。</li> </ul>  |

## MySQL同步至阿里云流式数据服务DataHub

| 类型   | 说明   |
|------|--|
| 源库限制 | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> <li>Binlog日志：                         <ul style="list-style-type: none"> <li>需开启Binlog，开启方法请参见<a href="#">设置实例参数</a>，并且binlog_row_image为full。否则预检查阶段提示报错，且无法成功启动数据同步任务。</li> </ul> </li> </ul> <div style="border: 1px solid #ccc; padding: 10px; margin: 10px 0;"> <p> <b>注意</b></p> <ul style="list-style-type: none"> <li>如源实例为自建MySQL，则需开启Binlog，并且binlog_format为row、binlog_row_image为full。</li> <li>如源实例自建MySQL是双主集群（两者互为主从），为保障DTS能获取全部的Binlog日志，则您需开启参数log_slave_updates。具体操作请参见<a href="#">自建MySQL创建账号并设置binlog</a>。</li> </ul> </div> <ul style="list-style-type: none"> <li>如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为库表结构同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在库表结构同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。如源为RDS MySQL，具体操作请参见<a href="#">本地日志（Binlog）</a>。</li> </ul> |
| 其他限制 | <ul style="list-style-type: none"> <li>不支持全量数据初始化，即DTS不会将源RDS实例中同步对象的存量数据同步至目标DataHub实例。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul>  |
| 特殊情况 | <p>当源库为自建MySQL时</p> <ul style="list-style-type: none"> <li>在同步时，如果源库进行主备切换，将会导致同步任务失败。</li> <li>由于DTS的延迟时间是根据同步到目标库最后一条数据的时间戳和当前时间戳对比得出，源库长时间未执行DML操作可能导致延迟信息不准确。如果任务显示的延迟时间过大，您可以在源库执行一个DML操作来更新延迟信息。</li> </ul> <div style="border: 1px solid #ccc; padding: 10px; margin: 10px 0;"> <p> <b>说明</b> 如果同步对象选择为整库，您还可以创建心跳表，心跳表每秒定期更新或者写入数据。</p> </div> <ul style="list-style-type: none"> <li>DTS会在源库定时执行CREATE DATABASE IF NOT EXISTS `test`命令以推进Binlog位点。</li> </ul>  |

## MySQL同步至Elasticsearch

| 类型 | 说明 |
|----|----|
|    |    |

| 类型   | 说明  |
|------|---|
| 源库限制 | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> <li>Binlog日志：             <ul style="list-style-type: none"> <li>需开启Binlog，开启方法请参见<a href="#">设置实例参数</a>，并且binlog_row_image为full。否则预检查阶段提示报错，且无法成功启动数据同步任务。</li> </ul> </li> </ul> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 10px; margin: 10px 0;"> <p> <b>注意</b></p> <ul style="list-style-type: none"> <li>如源实例为自建MySQL，则需开启Binlog，并且binlog_format为row、binlog_row_image为full。</li> <li>如源实例自建MySQL是双主集群（两者互为主从），为保障DTS能获取全部的Binlog日志，则您需开启参数log_slave_updates。具体操作请参见<a href="#">自建MySQL创建账号并设置binlog</a>。</li> </ul> </div> <ul style="list-style-type: none"> <li>如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在全量同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。如源为RDS MySQL，具体操作请参见<a href="#">本地日志（Binlog）</a>。</li> </ul> |
| 其他限制 | <ul style="list-style-type: none"> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>如同步对象为单个或多个表（非整库），那么在数据同步时，勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。<br/>您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，请参见<a href="#">不锁表结构变更</a>。</li> <li>如果源库中待同步的表需要执行增加列的操作，您只需先在Elasticsearch实例中修改对应表的mapping，然后在源MySQL数据库中执行相应的DDL操作，最后暂停并启动DTS同步实例。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul>  |
| 特殊情况 | <p>当源库为自建MySQL时</p> <ul style="list-style-type: none"> <li>在同步时，如果源库进行主备切换，将会导致同步任务失败。</li> <li>由于DTS的延迟时间是根据同步到目标库最后一条数据的时间戳和当前时间戳对比得出，源库长时间未执行DML操作可能导致延迟信息不准确。如果任务显示的延迟时间过大，您可以在源库执行一个DML操作来更新延迟信息。</li> </ul> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 10px; margin: 10px 0;"> <p> <b>说明</b> 如果同步对象选择为整库，您还可以创建心跳表，心跳表每秒定期更新或者写入数据。</p> </div> <ul style="list-style-type: none"> <li>DTS会在源库定时执行CREATE DATABASE IF NOT EXISTS `test`命令以推进Binlog位点。</li> </ul>  |

## MySQL同步至MaxCompute

| 类型   | 说明   |
|------|--|
| 源库限制 | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> <li>Binlog日志： <ul style="list-style-type: none"> <li>需开启Binlog，开启方法请参见<a href="#">设置实例参数</a>，并且binlog_row_image为full。否则预检查阶段提示报错，且无法成功启动数据同步任务。</li> </ul> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 10px; margin: 10px 0;"> <p> <b>注意</b></p> <ul style="list-style-type: none"> <li>如源实例为自建MySQL，则需开启Binlog，并且binlog_format为row、binlog_row_image为full。</li> <li>如源实例自建MySQL是双主集群（两者互为主从），为保障DTS能获取全部的Binlog日志，则您需开启参数log_slave_updates。具体操作请参见<a href="#">自建MySQL创建账号并设置binlog</a>。</li> </ul> </div> </li> <li>如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在全量同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。如源为RDS MySQL，具体操作请参见<a href="#">本地日志（Binlog）</a>。</li> </ul> |
| 其他限制 | <ul style="list-style-type: none"> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>如同步对象为单个或多个表（非整库），那么在数据同步时，勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，请参见<a href="#">不锁表结构变更</a>。</li> <li>由于MaxCompute不支持主键约束，当DTS在同步数据时因网络等原因触发重传，可能会导致MaxCompute中出现重复记录。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul>   |
| 特殊情况 | <p>当源库为自建MySQL时</p> <ul style="list-style-type: none"> <li>在同步时，如果源库进行主备切换，将会导致同步任务失败。</li> <li>由于DTS的延迟时间是根据同步到目标库最后一条数据的时间戳和当前时间戳对比得出，源库长时间未执行DML操作可能导致延迟信息不准确。如果任务显示的延迟时间过大，您可以在源库执行一个DML操作来更新延迟信息。</li> </ul> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 10px; margin: 10px 0;"> <p> <b>说明</b> 如果同步对象选择为整库，您还可以创建心跳表，心跳表每秒定期更新或者写入数据。</p> </div> <ul style="list-style-type: none"> <li>DTS会在源库定时执CREATE DATABASE IF NOT EXISTS `test`命令以推进Binlog位点。</li> </ul>  |

## MySQL同步至ClickHouse集群

| 类型 | 说明 |
|----|----|
|    |    |

| 类型   | 说明   |
|------|--|
| 源库限制 | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> <li>Binlog日志：                         <ul style="list-style-type: none"> <li>需开启Binlog，开启方法请参见<a href="#">设置实例参数</a>，并且binlog_row_image为full。否则预检查阶段提示报错，且无法成功启动数据同步任务。</li> </ul> </li> </ul> <div style="border: 1px solid #ccc; padding: 10px; margin: 10px 0;"> <p> <b>注意</b></p> <ul style="list-style-type: none"> <li>如源实例为自建MySQL，则需开启Binlog，并且binlog_format为row、binlog_row_image为full。</li> <li>如源实例自建MySQL是双主集群（两者互为主从），为保障DTS能获取全部的Binlog日志，则您需开启参数log_slave_updates。具体操作请参见<a href="#">自建MySQL创建账号并设置binlog</a>。</li> </ul> </div> <ul style="list-style-type: none"> <li>如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在全量同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。如源为RDS MySQL，具体操作请参见<a href="#">本地日志（Binlog）</a>。</li> </ul> |
| 注意事项 | <ul style="list-style-type: none"> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>如同步对象为单个或多个表（非整库），那么在数据同步时，勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，请参见<a href="#">不锁表结构变更</a>。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> <li>待同步数据库、表和列的名称符合云数据库ClickHouse的命名规范，相关规范，请参见<a href="#">对象命名规范限制</a>。</li> </ul>   |
| 特殊情况 | <p>当源库为自建MySQL时</p> <ul style="list-style-type: none"> <li>在同步时，如果源库进行主备切换，将会导致同步任务失败。</li> <li>由于DTS的延迟时间是根据同步到目标库最后一条数据的时间戳和当前时间戳对比得出，源库长时间未执行DML操作可能导致延迟信息不准确。如果任务显示的延迟时间过大，您可以在源库执行一个DML操作来更新延迟信息。</li> </ul> <div style="border: 1px solid #ccc; padding: 10px; margin: 10px 0;"> <p> <b>说明</b> 如果同步对象选择为整库，您还可以创建心跳表，心跳表每秒定期更新或者写入数据。</p> </div> <ul style="list-style-type: none"> <li>DTS会在源库定时执行CREATE DATABASE IF NOT EXISTS `test`命令以推进Binlog位点。</li> </ul>  |

## MySQL同步至Tablestore

| 类型 | 说明 |
|----|----|
|    |    |

| 类型   | 说明   |
|------|--|
| 源库限制 | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> <li>Binlog日志：             <ul style="list-style-type: none"> <li>需开启Binlog，开启方法请参见<a href="#">设置实例参数</a>，并且binlog_row_image为full。否则预检查阶段提示报错，且无法成功启动数据同步任务。</li> </ul> </li> </ul> <div style="border: 1px solid #ccc; padding: 10px; margin: 10px 0;"> <p> <b>注意</b></p> <ul style="list-style-type: none"> <li>如源实例为自建MySQL，则需开启Binlog，并且binlog_format为row、binlog_row_image为full。</li> <li>如源实例自建MySQL是双主集群（两者互为主从），为保障DTS能获取全部的Binlog日志，则您需开启参数log_slave_updates。具体操作请参见<a href="#">自建MySQL创建账号并设置binlog</a>。</li> </ul> </div> <ul style="list-style-type: none"> <li>如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在全量同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。如源为RDS MySQL，具体操作请参见<a href="#">本地日志（Binlog）</a>。</li> </ul> |
| 其他限制 | <ul style="list-style-type: none"> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>如同步对象为单个或多个表（非整库），那么在数据同步时，勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，请参见<a href="#">不锁表结构变更</a>。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> <li>待同步的表或列名称符合Tablestore实例的命名规范：             <ul style="list-style-type: none"> <li>表或列的名称由大小写字母、数字或下划线（_）组成，且只能以字母或下划线开头。</li> <li>表或列的名称长度为1~255个字符。</li> </ul> </li> </ul>  |
| 特殊情况 | <p>当源库为自建MySQL时</p> <ul style="list-style-type: none"> <li>在同步时，如果源库进行主备切换，将会导致同步任务失败。</li> <li>由于DTS的延迟时间是根据同步到目标库最后一条数据的时间戳和当前时间戳对比得出，源库长时间未执行DML操作可能导致延迟信息不准确。如果任务显示的延迟时间过大，您可以在源库执行一个DML操作来更新延迟信息。</li> </ul> <div style="border: 1px solid #ccc; padding: 10px; margin: 10px 0;"> <p> <b>说明</b> 如果同步对象选择为整库，您还可以创建心跳表，心跳表每秒定期更新或者写入数据。</p> </div> <ul style="list-style-type: none"> <li>DTS会在源库定时执行CREATE DATABASE IF NOT EXISTS `test`命令以推进Binlog位点。</li> </ul>  |

### MySQL同步至阿里云消息队列Kafka版、自建Kafka集群

| 类型 | 说明 |
|----|----|
|    |    |

| 类型      | 说明  |
|---------|---|
| 源库的环境要求 | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> <li>Binlog日志：                         <ul style="list-style-type: none"> <li>需开启Binlog，开启方法请参见<a href="#">设置实例参数</a>，并且binlog_row_image为full。否则预检查阶段提示报错，且无法成功启动数据同步任务。</li> </ul> </li> </ul> <div style="border: 1px solid #ccc; padding: 10px; margin: 10px 0;"> <p> <b>注意</b></p> <ul style="list-style-type: none"> <li>如源实例为自建MySQL，则需开启Binlog，并且binlog_format为row、binlog_row_image为full。</li> <li>如源实例自建MySQL是双主集群（两者互为主从），为保障DTS能获取全部的Binlog日志，则您需开启参数log_slave_updates。具体操作请参见<a href="#">为自建MySQL创建账号并设置binlog</a>。</li> </ul> </div> <ul style="list-style-type: none"> <li>如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在全量同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。如源为RDS MySQL，具体操作请参见<a href="#">本地日志（Binlog）</a>。</li> </ul> |
| 注意事项    | <ul style="list-style-type: none"> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>如同步对象为单个或多个表（非整库），那么在数据同步时，勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，请参见<a href="#">不锁表结构变更</a>。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul>  |
| 特殊情况    | <p>当源库为自建MySQL时</p> <ul style="list-style-type: none"> <li>在同步时，如果源库进行主备切换，将会导致同步任务失败。</li> <li>由于DTS的延迟时间是根据同步到目标库最后一条数据的时间戳和当前时间戳对比得出，源库长时间未执行DML操作可能导致延迟信息不准确。如果任务显示的延迟时间过大，您可以在源库执行一个DML操作来更新延迟信息。</li> </ul> <div style="border: 1px solid #ccc; padding: 10px; margin: 10px 0;"> <p> <b>说明</b> 如果同步对象选择为整库，您还可以创建心跳表，心跳表每秒定期更新或者写入数据。</p> </div> <ul style="list-style-type: none"> <li>DTS会在源库定时执行CREATE DATABASE IF NOT EXISTS `test`命令以推进Binlog位点。</li> </ul>   |

## MySQL同步至PolarDB-X 2.0

| 类型 | 说明 |
|----|----|
|    |    |

| 类型      | 说明  |
|---------|---|
| 源库的环境要求 | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> <li>Binlog日志：             <ul style="list-style-type: none"> <li>需开启Binlog，开启方法请参见<a href="#">设置实例参数</a>，并且binlog_row_image为full。否则预检查阶段提示报错，且无法成功启动数据同步任务。</li> </ul> </li> </ul> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 10px; margin: 10px 0;"> <p> <b>注意</b></p> <ul style="list-style-type: none"> <li>如源实例为自建MySQL，则需开启Binlog，并且binlog_format为row、binlog_row_image为full。</li> <li>如源实例自建MySQL是双主集群（两者互为主从），为保障DTS能获取全部的Binlog日志，则您需开启参数log_slave_updates。具体操作请参见<a href="#">自建MySQL创建账号并设置binlog</a>。</li> </ul> </div> <ul style="list-style-type: none"> <li>如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在全量同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。如源为RDS MySQL，具体操作请参见<a href="#">本地日志（Binlog）</a>。</li> </ul> |
| 注意事项    | <ul style="list-style-type: none"> <li>暂不支持结构初始化，在配置同步作业前，您需要在目标实例中创建对应的库和表。</li> <li>同步对象要求：             <ul style="list-style-type: none"> <li>不支持BIT、VARBIT、GEOMETRY、ARRAY、UUID、TSQUERY、TSVECTOR、TXID_SNAPSHOT类型的数据同步。</li> <li>暂不支持同步前缀索引，如果源库存在前缀索引可能导致数据同步失败。</li> </ul> </li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>如同步对象为单个或多个表（非整库），那么在数据同步时，勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，请参见<a href="#">不锁表结构变更</a>。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul>  |
| 特殊情况    | <p>当源库为自建MySQL时</p> <ul style="list-style-type: none"> <li>在同步时，如果源库进行主备切换，将会导致同步任务失败。</li> <li>由于DTS的延迟时间是根据同步到目标库最后一条数据的时间戳和当前时间戳对比得出，源库长时间未执行DML操作可能导致延迟信息不准确。如果任务显示的延迟时间过大，您可以在源库执行一个DML操作来更新延迟信息。</li> </ul> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 10px; margin: 10px 0;"> <p> <b>说明</b> 如果同步对象选择为整库，您还可以创建心跳表，心跳表每秒定期更新或者写入数据。</p> </div> <ul style="list-style-type: none"> <li>DTS会在源库定时执CREATE DATABASE IF NOT EXISTS `test`命令以推进Binlog位点。</li> </ul>   |

## MySQL同步至Oracle

| 类型   | 说明   |
|------|--|
| 源库限制 | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> <li>Binlog日志：                         <ul style="list-style-type: none"> <li>需开启Binlog，开启方法请参见<a href="#">设置实例参数</a>，并且binlog_row_image为full。否则预检查阶段提示报错，且无法成功启动数据同步任务。</li> </ul> </li> </ul> <div style="border: 1px solid #ccc; padding: 10px; margin: 10px 0;"> <p> <b>注意</b></p> <ul style="list-style-type: none"> <li>如源实例为自建MySQL，则需开启Binlog，并且binlog_format为row、binlog_row_image为full。</li> <li>如源实例自建MySQL是双主集群（两者互为主从），为保障DTS能获取全部的Binlog日志，则您需开启参数log_slave_updates。具体操作请参见<a href="#">自建MySQL创建账号并设置binlog</a>。</li> </ul> </div> <ul style="list-style-type: none"> <li>如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为库表结构同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在库表结构同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。如源为RDS MySQL，具体操作请参见<a href="#">本地日志（Binlog）</a>。</li> </ul> |
| 其他限制 | <ul style="list-style-type: none"> <li>不支持多列属性变更。</li> <li>不支持基本数据类型的变更。</li> <li>若自建Oracle为RAC架构，不支持配置ScanIP，仅支持将其中1个VIP配置到连接信息中，该方式配置后不支持RAC的节点切换。</li> </ul>  |
| 特殊情况 | <p>当源库为自建MySQL时</p> <ul style="list-style-type: none"> <li>在同步时，如果源库进行主备切换，将会导致同步任务失败。</li> <li>由于DTS的延迟时间是根据同步到目标库最后一条数据的时间戳和当前时间戳对比得出，源库长时间未执行DML操作可能导致延迟信息不准确。如果任务显示的延迟时间过大，您可以在源库执行一个DML操作来更新延迟信息。</li> </ul> <div style="border: 1px solid #ccc; padding: 10px; margin: 10px 0;"> <p> <b>说明</b> 如果同步对象选择为整库，您还可以创建心跳表，心跳表每秒定期更新或者写入数据。</p> </div> <ul style="list-style-type: none"> <li>DTS会在源库定时执行CREATE DATABASE IF NOT EXISTS `test`命令以推进Binlog位点。</li> </ul>  |

## MySQL同步至PostgreSQL

| 类型 | 说明 |
|----|----|
|    |    |

| 类型   | 说明  |
|------|---|
| 源库限制 | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> <li>Binlog日志：             <ul style="list-style-type: none"> <li>需开启Binlog，开启方法请参见<a href="#">设置实例参数</a>，并且binlog_row_image为full。否则预检查阶段提示报错，且无法成功启动数据同步任务。</li> </ul> </li> </ul> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 10px; margin: 10px 0;"> <p> <b>注意</b></p> <ul style="list-style-type: none"> <li>如源实例为自建MySQL，则需开启Binlog，并且binlog_format为row、binlog_row_image为full。</li> <li>如源实例自建MySQL是双主集群（两者互为主从），为保障DTS能获取全部的Binlog日志，则您需开启参数log_slave_updates。具体操作请参见<a href="#">自建MySQL创建账号并设置binlog</a>。</li> </ul> </div> <ul style="list-style-type: none"> <li>如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为库表结构同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在库表结构同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。如源为RDS MySQL，具体操作请参见<a href="#">本地日志（Binlog）</a>。</li> </ul> |
| 其他限制 | <ul style="list-style-type: none"> <li>同步对象要求：             <ul style="list-style-type: none"> <li>仅支持表级同步。</li> <li>不支持BIT、VARBIT、GEOMETRY、ARRAY、UUID、TSQUERY、TSVECTOR、TXID_SNAPSHOT类型的数据同步。</li> <li>暂不支持同步前缀索引，如果源库存在前缀索引可能导致数据同步失败。</li> </ul> </li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul>   |
| 特殊情况 | <p>当源库为自建MySQL时</p> <ul style="list-style-type: none"> <li>在同步时，如果源库进行主备切换，将会导致同步任务失败。</li> <li>由于DTS的延迟时间是根据同步到目标库最后一条数据的时间戳和当前时间戳对比得出，源库长时间未执行DML操作可能导致延迟信息不准确。如果任务显示的延迟时间过大，您可以在源库执行一个DML操作来更新延迟信息。</li> </ul> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 10px; margin: 10px 0;"> <p> <b>说明</b> 如果同步对象选择为整库，您还可以创建心跳表，心跳表每秒定期更新或者写入数据。</p> </div> <ul style="list-style-type: none"> <li>DTS会在源库定时执行CREATE DATABASE IF NOT EXISTS `test`命令以推进Binlog位点。</li> </ul>  |

## 3.2. 同步时源库为PolarDB MySQL的注意事项及限制

如果同步的源集群为PolarDB MySQL集群，您需要在配置具体的同步任务前，参考本文的注意事项及限制，以保障数据同步任务的正常运行。

### 源实例为PolarDB MySQL的同步方案概览

根据如下同步方案，查看同步任务的注意事项及限制：

#### 说明

DTS默认同步到目标数据库中时会取消外键约束，因此源数据库的级联、删除等操作不会同步到如下目标数据库：

- MySQL (RDS MySQL、自建MySQL)
- PolarDB MySQL
- PolarDB-X 1.0
- 云原生数据仓库AnalyticDB MySQL
- 云原生数据仓库AnalyticDB PostgreSQL
- Elasticsearch

- [PolarDB MySQL间的同步](#)
- [PolarDB MySQL同步至RDS MySQL、自建MySQL](#)
- [PolarDB MySQL同步至PolarDB-X 1.0](#)
- [PolarDB MySQL同步至云原生数据仓库AnalyticDB MySQL](#)
- [PolarDB MySQL同步至云原生数据仓库AnalyticDB PostgreSQL](#)
- [PolarDB MySQL同步至阿里云流式数据服务DataHub](#)
- [PolarDB MySQL同步至Elasticsearch](#)
- [PolarDB MySQL同步至阿里云消息队列Kafka、自建Kafka](#)
- [PolarDB MySQL同步至云原生数据仓库AnalyticDB PostgreSQL](#)
- [PolarDB MySQL同步至MaxCompute](#)
- [PolarDB MySQL同步至自建Oracle](#)

### PolarDB MySQL间的同步

具体注意事项及限制如下：

| 类型   | 说明   |
|------|--|
| 源库限制 | <ul style="list-style-type: none"> <li>• 待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>• 如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> <li>• Binlog日志：               <ul style="list-style-type: none"> <li>◦ 需开启，并且loose_polar_log_bin为on。否则预检查阶段显示报错提示，且无法成功启动数据同步任务。</li> <li>◦ 如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在全量同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> </ul> |

| 类型     | 说明  |
|--------|---|
| 其他限制   | <ul style="list-style-type: none"> <li>不支持同步源PolarDB MySQL引擎只读节点。</li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据迁移。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>对于表级别的数据同步，请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>对于表级别的数据同步，如果除DTS以外的数据没有写入目标库，您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，请参见<a href="#">不锁表结构变更</a>。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul> |
| 其他注意事项 | <p>PolarDB MySQL集群间的双向同步</p> <ul style="list-style-type: none"> <li>目前DTS仅支持两个PolarDB MySQL集群之间的双向同步，暂不支持多个PolarDB MySQL集群间的双向同步。</li> <li>DDL语法同步方向限制。为保障双向同步链路的稳定性和数据一致性，只支持正向同步DDL，不支持反向同步DDL。</li> <li>双向同步任务时，DTS会在目标库中建立数据库dts，以避免数据循环同步，任务运行期间，请勿修改该库。</li> <li>如双向同步任务的源实例或目标实例位于海外地域，则仅支持同地域的双向同步，不支持跨地域的双向同步。例如，支持日本地域间的双向同步，不支持日本地域与法兰克福地域间的双向同步。</li> <li>若DTS正向和反向同步任务均同步了某个表，且正向同步任务同步了该表的全量和增量数据，则反向同步任务只同步该表的增量数据。</li> <li>DTS会在源库定时执行CREATE DATABASE IF NOT EXISTS `test`命令以推进Binlog位点。</li> </ul>                                      |

## PolarDB MySQL同步至RDS MySQL、自建MySQL

具体注意事项及限制如下：

| 类型   | 说明   |
|------|--|
| 源库限制 | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出現重复数据。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> <li>Binlog日志：                         <ul style="list-style-type: none"> <li>需开启，并且loose_polar_log_bin为on。否则预检查阶段显示报错提示，且无法成功启动数据同步任务。</li> <li>如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在全量同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> </ul> |
| 其他限制 | <ul style="list-style-type: none"> <li>不支持同步源PolarDB MySQL引擎只读节点。</li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据迁移。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>对于表级别的数据同步，请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>对于表级别的数据同步，如果除DTS以外的数据没有写入目标库，您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，请参见<a href="#">不锁表结构变更</a>。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul>  |

| 类型     | 说明  |
|--------|---|
| 其他注意事项 | DTS会在源库定时执行CREATE DATABASE IF NOT EXISTS `test`命令以推进Binlog位点。 |

## PolarDB MySQL同步至PolarDB-X 1.0

具体注意事项及限制如下：

| 类型     | 说明  |
|--------|---|
| 源库限制   | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> <li>Binlog日志：                         <ul style="list-style-type: none"> <li>需开启，并且loose_polar_log_bin为on。否则预检查阶段显示报错提示，且无法成功启动数据同步任务。</li> <li>如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在全量同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> </ul>  |
| 其他限制   | <ul style="list-style-type: none"> <li>同步对象要求：                         <ul style="list-style-type: none"> <li>不支持BIT、VARBIT、GEOMETRY、ARRAY、UUID、TSQUERY、TSVECTOR、TXID_SNAPSHOT类型的数据同步。</li> <li>暂不支持同步前缀索引，如果源库存在前缀索引可能导致数据同步失败。</li> <li>不支持同步源PolarDB MySQL引擎只读节点。</li> </ul> </li> <li>暂不支持结构初始化，在配置同步任务前，您需要在目标实例中创建对应的库和表。</li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据迁移。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>对于表级别的数据同步，请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>对于表级别的数据同步，如果除DTS以外的数据没有写入目标库，您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，请参见<a href="#">不锁表结构变更</a>。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul> |
| 其他注意事项 | DTS会在源库定时执行CREATE DATABASE IF NOT EXISTS `test`命令以推进Binlog位点。   |

## PolarDB MySQL同步至云原生数据仓库AnalyticDB MySQL

具体注意事项及限制如下：

| 类型 | 说明 |
|----|----|
|----|----|

| 类型     | 说明  |
|--------|---|
| 源库限制   | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> <li>Binlog日志：                         <ul style="list-style-type: none"> <li>需开启，并且loose_polar_log_bin为on。否则预检查阶段显示报错提示，且无法成功启动数据同步任务。</li> <li>如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在全量同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> <li>同步期间，请勿执行修改主键和添加注释的DDL操作（如 <code>ALTER TABLE table_name COMMENT='表的注释';</code>），否则数据同步过程中会导致执行该DDL失败。</li> </ul>                         |
| 其他限制   | <ul style="list-style-type: none"> <li>暂不支持同步前缀索引，如果源库存在前缀索引可能导致数据同步失败。</li> <li>不支持同步源PolarDB MySQL引擎只读节点。</li> <li>由于云原生数据仓库AnalyticDB MySQL本身的使用限制，当云原生数据仓库AnalyticDB MySQL集群中的节点磁盘空间使用量超过80%，会导致DTS任务异常，产生延迟。请提前根据待同步的对象预估所需空间，确保目标集群具备充足的存储空间。</li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据迁移。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>对于表级别的数据同步，请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>对于表级别的数据同步，如果除DTS以外的数据没有写入目标库，您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，请参见<a href="#">不锁表结构变更</a>。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul> |
| 其他注意事项 | DTS会在源库定时执行CREATE DATABASE IF NOT EXISTS `test`命令以推进Binlog位点。   |

## PolarDB MySQL同步至云原生数据仓库AnalyticDB PostgreSQL

具体注意事项及限制如下：

| 类型   | 说明  |
|------|---|
| 源库限制 | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> <li>Binlog日志：                         <ul style="list-style-type: none"> <li>需开启，并且loose_polar_log_bin为on。否则预检查阶段显示报错提示，且无法成功启动数据同步任务。</li> <li>如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在全量同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> <li>同步期间，请勿执行修改主键和添加注释的DDL操作（如 <code>ALTER TABLE table_name COMMENT='表的注释';</code>），否则数据同步过程中会导致执行该DDL失败。</li> </ul> |

| 类型     | 说明   |
|--------|--|
| 其他限制   | <ul style="list-style-type: none"> <li>同步对象要求：                             <ul style="list-style-type: none"> <li>仅支持表级同步。</li> <li>不支持BIT、VARBIT、GEOMETRY、ARRAY、UUID、TSQUERY、TSVECTOR、TXID_SNAPSHOT类型的数据同步。</li> <li>暂不支持同步前缀索引，如果源库存在前缀索引可能导致数据同步失败。</li> <li>不支持同步源PolarDB MySQL引擎只读节点。</li> </ul> </li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据迁移。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>对于表级别的数据同步，请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>对于表级别的数据同步，如果除DTS以外的数据没有写入目标库，您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，请参见<a href="#">不锁表结构变更</a>。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul> |
| 其他注意事项 | DTS会在源库定时执CREATE DATABASE IF NOT EXISTS `test`命令以推进Binlog位点。   |

### PolarDB MySQL同步至阿里云流式数据服务DataHub

具体注意事项及限制如下：

| 类型     | 说明  |
|--------|---|
| 源库限制   | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出現重复数据。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> <li>Binlog日志：                             <ul style="list-style-type: none"> <li>需开启，并且loose_polar_log_bin为on。否则预检查阶段显示报错提示，且无法成功启动数据同步任务。</li> <li>如为同步任务场景，DTS要求源数据库的本地Binlog日志至少保存24小时。</li> </ul> </li> </ul>  |
| 其他限制   | <ul style="list-style-type: none"> <li>不支持全量数据初始化，即DTS不会将源RDS实例中同步对象的存量数据同步至目标DataHub实例。</li> <li>支持库级和表级的数据同步。</li> <li>不支持同步源PolarDB MySQL引擎只读节点。</li> <li>对于表级别的数据同步，请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>对于表级别的数据同步，如果除DTS以外的数据没有写入目标库，您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，请参见<a href="#">不锁表结构变更</a>。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul> |
| 其他注意事项 | DTS会在源库定时执CREATE DATABASE IF NOT EXISTS `test`命令以推进Binlog位点。  |

### PolarDB MySQL同步至Elasticsearch

具体注意事项及限制如下：

| 类型     | 说明   |
|--------|--|
| 源库限制   | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> <li>Binlog日志： <ul style="list-style-type: none"> <li>需开启，并且loose_polar_log_bin为on。否则预检查阶段显示报错提示，且无法成功启动数据同步任务。</li> <li>如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在全量同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> </ul>   |
| 其他限制   | <ul style="list-style-type: none"> <li>不支持同步源PolarDB MySQL引擎只读节点。</li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据迁移。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>对于表级别的数据同步，请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>对于表级别的数据同步，如果除DTS以外的数据没有写入目标库，您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，请参见<a href="#">不锁表结构变更</a>。</li> <li>如果源库中待同步的表需要执行增加列的操作，您只需先在Elasticsearch实例中修改对应表的mapping，然后在源MySQL数据库中执行相应的DDL操作，最后暂停并启动DTS同步实例。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul> |
| 其他注意事项 | DTS会在源库定时执行CREATE DATABASE IF NOT EXISTS `test`命令以推进Binlog位点。  |

## PolarDB MySQL同步至阿里云消息队列Kafka、自建Kafka

具体注意事项及限制如下：

| 类型   | 说明   |
|------|--|
| 源库限制 | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> <li>Binlog日志： <ul style="list-style-type: none"> <li>需开启，并且loose_polar_log_bin为on。否则预检查阶段显示报错提示，且无法成功启动数据同步任务。</li> <li>如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在全量同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> </ul> |

| 类型     | 说明  |
|--------|---|
| 其他限制   | <ul style="list-style-type: none"> <li>不支持同步源PolarDB MySQL引擎只读节点。</li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据迁移。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>对于表级别的数据同步，请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>对于表级别的数据同步，如果除DTS以外的数据没有写入目标库，您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，请参见<a href="#">不锁表结构变更</a>。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul> |
| 其他注意事项 | DTS会在源库定时执CREATE DATABASE IF NOT EXISTS `test`命令以推进Binlog位点。  |

### PolarDB MySQL同步至MaxCompute

具体注意事项及限制如下：

| 类型     | 说明  |
|--------|---|
| 源库限制   | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出現重复数据。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> <li>Binlog日志：                     <ul style="list-style-type: none"> <li>需开启，并且loose_polar_log_bin为on。否则预检查阶段显示报错提示，且无法成功启动数据同步任务。</li> <li>如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在全量同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> </ul>                                    |
| 其他限制   | <ul style="list-style-type: none"> <li>不支持同步源PolarDB MySQL引擎只读节点。</li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据迁移。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>对于表级别的数据同步，请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>对于表级别的数据同步，如果除DTS以外的数据没有写入目标库，您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，请参见<a href="#">不锁表结构变更</a>。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> <li>由于MaxCompute不支持主键约束，当DTS在同步数据时因网络等原因触发重传，可能会导致MaxCompute中出現重复记录。</li> </ul> |
| 其他注意事项 | DTS会在源库定时执CREATE DATABASE IF NOT EXISTS `test`命令以推进Binlog位点。  |

### PolarDB MySQL同步至自建Oracle

具体注意事项及限制如下：

| 类型     | 说明  |
|--------|---|
| 源库限制   | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出現重复数据。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> <li>Binlog日志： <ul style="list-style-type: none"> <li>需开启，并且loose_polar_log_bin为on。否则预检查阶段显示报错提示，且无法成功启动数据同步任务。</li> <li>如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在全量同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> </ul>  |
| 其他限制   | <ul style="list-style-type: none"> <li>不支持同步源PolarDB MySQL引擎只读节点。</li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据迁移。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>对于表级别的数据同步，请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>对于表级别的数据同步，如果除DTS以外的数据没有写入目标库，您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，请参见<a href="#">不锁表结构变更</a>。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> <li>若自建Oracle为RAC架构，不支持配置ScanIP，仅支持将其中任意1个VIP配置到连接信息中，该方式配置后不支持RAC的节点切换。</li> </ul> |
| 其他注意事项 | DTS会在源库定时执行CREATE DATABASE IF NOT EXISTS `test`命令以推进Binlog位点。   |

### 3.3. 同步时源库为PolarDB-X 1.0的注意事项及限制

如果同步的源实例为PolarDB-X 1.0，您需要在配置具体的同步任务前，参考本文的注意事项及限制，以保障数据同步任务的正常运行。

#### 注意

更多功能规范和约束说明，请参见[源为PolarDB-X的功能规范和约束说明](#)。  
推荐的配置方案，请参见[源为PolarDB-X的任务配置方案](#)。

#### 源库为PolarDB-X 1.0的同步方案概览

根据如下同步方案，查看同步任务的注意事项及限制：

#### 说明

DTS默认同步到目标数据库中时会取消外键约束，因此源数据库的级联、删除等操作不会同步到如下目标数据库：

- PolarDB-X 1.0
- MySQL（RDS MySQL、自建MySQL）
- PolarDB MySQL
- 云原生数据仓库AnalyticDB MySQL

- [PolarDB-X 1.0间同步](#)

- PolarDB-X 1.0同步至MySQL或PolarDB MySQL
- PolarDB-X 1.0同步至云原生数据仓库AnalyticDB MySQL 3.0

### PolarDB-X 1.0间同步

| 类型   | 说明   |
|------|--|
| 源库限制 | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据；不支持二级索引表同步。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），单次同步任务的表数量超过5000时，建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务，否则任务提交后可能会显示请求报错。</li> <li>PolarDB-X 1.0下面挂载的RDS MySQL的Binlog日志：                     <ul style="list-style-type: none"> <li>需开启，并且binlog_row_image为full。否则预检查阶段提示报错，且无法成功启动数据同步任务。</li> <li>如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在全量同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> <li>源库的操作限制：                     <ul style="list-style-type: none"> <li>同步期间如需切换PolarDB-X 1.0的网络类型，切换成功后，请提交工单调整同步链路的网络连接信息。</li> <li>在同步期间，请勿对源实例执行扩缩容（比如扩缩容其下挂载的RDS MySQL，或者虽未扩缩容RDS MySQL，但RDS MySQL中逻辑库表对应的物理库表的分布发生变化）、迁移热点表、变更拆分键操作和变更DDL等操作，否则数据同步任务会失败，或者数据不一致。</li> </ul> </li> </ul> |
| 其他限制 | <ul style="list-style-type: none"> <li>暂不支持结构初始化，在配置同步作业前，您需要在目标实例中创建对应的库和表。</li> <li>PolarDB-X 1.0的存储类型支持为RDS MySQL（包括私有定制RDS和单独购买的RDS），暂不支持为PolarDB MySQL。</li> <li>不支持为PolarDB-X 1.0计算资源只读实例。</li> <li>仅支持PolarDB-X 1.0存储资源通过水平拆分（分库分表），不支持为垂直拆分。</li> <li>由于PolarDB-X 1.0的同步任务为分布式同步，其下挂的一个RDS MySQL对应一个同步子任务。子任务运行情况，可在任务拓扑进行查询。</li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul>  |

### PolarDB-X 1.0同步至MySQL或PolarDB MySQL

| 类型 | 说明 |
|----|----|
|----|----|

| 类型   | 说明   |
|------|--|
| 源库限制 | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据；不支持二级索引表同步。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），单次同步任务的表数量超过5000时，建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务，否则任务提交后可能会显示请求报错。</li> <li>PolarDB-X 1.0下面挂载的RDS MySQL的Binlog日志：                         <ul style="list-style-type: none"> <li>需开启，并且binlog_row_image为full。否则预检查阶段提示报错，且无法成功启动数据同步任务。</li> <li>如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在全量同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> <li>源库的操作限制：                         <ul style="list-style-type: none"> <li>同步期间如需切换PolarDB-X 1.0的网络类型，切换成功后，请<a href="#">提交工单</a>调整同步链路的网络连接信息。</li> <li>在同步期间，请勿对源实例执行扩缩容（比如扩缩容其下挂载的RDS MySQL，或者虽未扩缩容RDS MySQL，但RDS MySQL中逻辑库表对应的物理库表的分布发生变化）、迁移热点表、变更拆分键操作和变更DDL等操作，否则数据同步任务会失败，或者数据不一致。</li> </ul> </li> <li>PolarDB-X 1.0的存储类型支持为RDS MySQL（包括私有定制RDS和单独购买的RDS），暂不支持为PolarDB MySQL。</li> <li>仅支持PolarDB-X 1.0存储资源通过水平拆分（分库分表），不支持为垂直拆分。</li> <li>不支持为PolarDB-X 1.0计算资源只读实例。</li> </ul> |
| 其他限制 | <ul style="list-style-type: none"> <li>由于PolarDB-X 1.0的同步任务为分布式同步，其下挂的一个RDS MySQL对应一个同步子任务。子任务运行情况，可在<a href="#">任务拓扑</a>进行查询。</li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul>   |

### PolarDB-X 1.0同步至云原生数据仓库AnalyticDB MySQL

| 类型 | 说明 |
|----|----|
|----|----|

| 类型   | 说明   |
|------|--|
| 源库限制 | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据；不支持二级索引表同步。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），单次同步任务的表数量超过5000时，建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务，否则任务提交后可能会显示请求报错。</li> <li>PolarDB-X 1.0下面挂载的RDS MySQL的Binlog日志：                         <ul style="list-style-type: none"> <li>需开启，并且binlog_row_image为full。否则预检查阶段提示报错，且无法成功启动数据同步任务。</li> <li>如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在全量同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> <li>源库的操作限制：                         <ul style="list-style-type: none"> <li>同步期间如需切换PolarDB-X 1.0的网络类型，切换成功后，请<a href="#">提交工单</a>调整同步链路的网络连接信息。</li> <li>在同步期间，请勿对源实例执行扩缩容（比如扩缩容其下挂载的RDS MySQL，或者虽未扩缩容RDS MySQL，但RDS MySQL中逻辑库表对应的物理库表的分布发生变化）、迁移热点表、变更拆分键操作和变更DDL等操作，否则数据同步任务会失败，或者数据不一致。</li> </ul> </li> <li>PolarDB-X 1.0的存储类型支持为RDS MySQL（包括私有定制RDS和单独购买的RDS），暂不支持为PolarDB MySQL。</li> <li>仅支持PolarDB-X 1.0存储资源通过水平拆分（分库分表），不支持为垂直拆分。</li> <li>不支持为PolarDB-X 1.0计算资源只读实例。</li> </ul> |
| 其他限制 | <ul style="list-style-type: none"> <li>由于PolarDB-X 1.0的同步任务为分布式同步，其下挂的一个RDS MySQL对应一个同步子任务。子任务运行情况，可在<a href="#">任务拓扑</a>进行查询。</li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul>   |

### PolarDB-X 1.0同步至阿里云流式数据服务DataHub

| 类型 | 说明 |
|----|----|
|----|----|

| 类型   | 说明   |
|------|--|
| 源库限制 | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据；不支持二级索引表同步。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），单次同步任务的表数量超过5000时，建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务，否则任务提交后可能会显示请求报错。</li> <li>PolarDB-X 1.0下面挂载的RDS MySQL的Binlog日志： <ul style="list-style-type: none"> <li>需开启，并且binlog_row_image为full。否则预检查阶段提示报错，且无法成功启动数据同步任务。</li> <li>如为增量同步任务，DTS要求源数据库的本地Binlog日志至少保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在全量同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> <li>源库的操作限制： <ul style="list-style-type: none"> <li>同步期间如需切换PolarDB-X 1.0的网络类型，切换成功后，请提交工单调整同步链路的网络连接信息。</li> <li>在同步期间，请勿对源实例执行扩缩容（比如扩缩容其下挂载的RDS MySQL，或者虽未扩缩容RDS MySQL，但RDS MySQL中逻辑库表对应的物理库表的分布发生变化）、迁移热点表、变更拆分键操作和变更DDL等操作，否则数据同步任务会失败，或者数据不一致。</li> </ul> </li> <li>PolarDB-X 1.0的存储类型支持为RDS MySQL（包括私有定制RDS和单独购买的RDS），暂不支持为PolarDB MySQL。</li> <li>仅支持PolarDB-X 1.0存储资源通过水平拆分（分库分表），不支持为垂直拆分。</li> <li>不支持为PolarDB-X 1.0计算资源只读实例。</li> </ul> |
| 其他限制 | <ul style="list-style-type: none"> <li>仅支持增量同步和库表结构同步，不支持全量同步。</li> <li>同步对象的限制： <ul style="list-style-type: none"> <li>仅支持表级别的数据同步。</li> <li>不支持新增列的数据同步，即源库中同步对象表新增了某个列，该列的数据不会同步至目标DataHub实例。</li> </ul> </li> <li>由于PolarDB-X 1.0的同步任务为分布式同步，其下挂的一个RDS MySQL对应一个同步子任务。子任务运行情况，可在任务拓扑进行查询。</li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul>  |

## 3.4. 同步时源库为PolarDB-X 2.0的注意事项及限制

如果同步的源实例为PolarDB-X，您需要在配置具体的同步任务前，参考本文的注意事项及限制，以保障数据同步任务的正常运行。

### 源库为PolarDB-X的同步方案概览

根据如下同步方案，查看同步任务的注意事项及限制：

**说明**

DTS默认同步到目标数据库中时会取消外键约束，因此源数据库的级联、删除等操作不会同步到如下目标数据库：

- MySQL (RDS MySQL、自建MySQL)
- PolarDB MySQL
- PolarDB-X 2.0
- 云原生数据仓库AnalyticDB MySQL
- 云原生数据仓库AnalyticDB PostgreSQL
- Elasticsearch

- [PolarDB-X 2.0同步至RDS MySQL、自建MySQL](#)
- [PolarDB-X 2.0同步至PolarDB MySQL](#)
- [PolarDB-X 2.0间的同步](#)
- [PolarDB-X 2.0同步至云原生数据仓库AnalyticDB MySQL](#)
- [PolarDB-X 2.0同步至阿里云消息队列Kafka](#)
- [PolarDB-X 2.0同步至DataHub](#)
- [PolarDB-X 2.0同步至Elasticsearch](#)

### PolarDB-X 2.0同步至RDS MySQL、自建MySQL

| 类型     | 说明   |
|--------|--|
| 源库限制   | <ul style="list-style-type: none"> <li>• 待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>• 若同步对象为表级别，且需进行编辑（如表列名映射），单次同步任务的表数量超过5000时，建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务，否则任务提交后可能会显示请求报错。</li> <li>• Binlog日志：                         <ul style="list-style-type: none"> <li>◦ 在PolarDB-X 2.0控制台参数设置界面开启binlog，开启方法请参见<a href="#">参数设置</a>。并且binlog_row_image为full。否则预检查阶段提示报错，且无法成功启动数据同步任务。</li> <li>◦ 如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在全量同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> <li>• PolarDB-X需兼容MySQL 5.7版本。</li> </ul> |
| 其他限制   | <ul style="list-style-type: none"> <li>• 执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>• 全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>• 请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>• 在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul>  |
| 其他注意事项 | DTS会在源库定时更新`dts_health_check`.`ha_health_check`表以推进binlog位点。   |

### PolarDB-X 2.0同步至PolarDB MySQL

| 类型 | 说明 |
|----|----|
|----|----|

| 类型     | 说明   |
|--------|--|
| 源库限制   | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>若同步对象为表级别，且需进行编辑（如表列名映射），单次同步任务的表数量超过5000时，建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务，否则任务提交后可能会显示请求报错。</li> <li>Binlog日志： <ul style="list-style-type: none"> <li>在PolarDB-X 2.0控制台参数设置界面开启binlog，开启方法请参见<a href="#">参数设置</a>。并且binlog_row_image为full。否则预检查阶段提示报错，且无法成功启动数据同步任务。</li> <li>如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在全量同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> <li>PolarDB-X需兼容MySQL 5.7版本。</li> </ul> |
| 其他限制   | <ul style="list-style-type: none"> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul>  |
| 其他注意事项 | DTS会在源库定时更新`dts_health_check`.`ha_health_check`表以推进binlog位点。   |

## PolarDB-X 2.0间的同步

| 类型   | 说明   |
|------|--|
| 源库限制 | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>若同步对象为表级别，且需进行编辑（如表列名映射），单次同步任务的表数量超过5000时，建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务，否则任务提交后可能会显示请求报错。</li> <li>Binlog日志： <ul style="list-style-type: none"> <li>在PolarDB-X 2.0控制台参数设置界面开启binlog，开启方法请参见<a href="#">参数设置</a>。并且binlog_row_image为full。否则预检查阶段提示报错，且无法成功启动数据同步任务。</li> <li>如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在全量同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> <li>PolarDB-X需兼容MySQL 5.7版本。</li> </ul> |

| 类型     | 说明   |
|--------|--|
| 其他限制   | <ul style="list-style-type: none"> <li>暂不支持结构初始化，在配置同步作业前，您需要在目标实例中创建对应的库和表。</li> <li>目标PolarDB-X需兼容MySQL 5.7版本。</li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul> |
| 其他注意事项 | DTS会在源库定时更新`dts_health_check`.`ha_health_check`表以推进binlog位点。   |

### PolarDB-X 2.0同步至云原生数据仓库AnalyticDB MySQL

| 类型     | 说明   |
|--------|--|
| 源库限制   | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>若同步对象为表级别，且需进行编辑（如表列名映射），单次同步任务的表数量超过5000时，建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务，否则任务提交后可能会显示请求报错。</li> <li>Binlog日志：                     <ul style="list-style-type: none"> <li>在PolarDB-X 2.0控制台参数设置界面开启binlog，开启方法请参见<a href="#">参数设置</a>。并且binlog_row_image为full。否则预检查阶段提示报错，且无法成功启动数据同步任务。</li> <li>如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在全量同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> <li>PolarDB-X需兼容MySQL 5.7版本。</li> </ul> |
| 其他限制   | <ul style="list-style-type: none"> <li>暂不支持同步前缀索引，如果源库存在前缀索引可能导致数据同步失败。</li> <li>由于云原生数据仓库AnalyticDB MySQL版本本身的使用限制，当云原生数据仓库AnalyticDB MySQL版中的节点磁盘空间使用量超过80%，会导致DTS任务异常，产生延迟。请提前根据待同步的对象预估所需空间，确保目标集群具备充足的存储空间。</li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul>   |
| 其他注意事项 | DTS会在源库定时更新`dts_health_check`.`ha_health_check`表以推进binlog位点。   |

### PolarDB-X 2.0同步至阿里云消息队列Kafka

| 类型 | 说明 |
|----|----|
|----|----|

| 类型     | 说明   |
|--------|--|
| 源库限制   | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>若同步对象为表级别，且需进行编辑（如表列名映射），单次同步任务的表数量超过5000时，建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务，否则任务提交后可能会显示请求报错。</li> <li>Binlog日志：                             <ul style="list-style-type: none"> <li>在PolarDB-X 2.0控制台参数设置界面开启binlog，开启方法请参见<a href="#">参数设置</a>。并且binlog_row_image为full。否则预检查阶段提示报错，且无法成功启动数据同步任务。</li> <li>如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在全量同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> <li>PolarDB-X需兼容MySQL 5.7版本。</li> </ul> |
| 其他限制   | <ul style="list-style-type: none"> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> <li>同步对象仅支持数据表，不支持非数据表的对象。</li> <li>不支持自动调整同步对象，如果对同步对象中的数据表进行重命名操作，且重命名后的名称不在同步对象中，那么这部分数据将不再同步到目标Kafka集群中。如需将修改后的数据表继续数据同步至目标Kafka集群中，您需要进行修改同步对象操作，详情请参见<a href="#">新增同步对象</a>。</li> </ul>  |
| 其他注意事项 | DTS会在源库定时更新`dts_health_check`.`ha_health_check`表以推进binlog位点。   |

## PolarDB-X 2.0同步至DataHub

| 类型   | 说明   |
|------|--|
| 源库限制 | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>若同步对象为表级别，且需进行编辑（如表列名映射），单次同步任务的表数量超过5000时，建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务，否则任务提交后可能会显示请求报错。</li> <li>Binlog日志：                             <ul style="list-style-type: none"> <li>在PolarDB-X 2.0控制台参数设置界面开启binlog，开启方法请参见<a href="#">参数设置</a>。并且binlog_row_image为full。否则预检查阶段提示报错，且无法成功启动数据同步任务。</li> <li>如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为库表结构同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在库表结构同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> <li>PolarDB-X需兼容MySQL 5.7版本。</li> </ul> |

| 类型     | 说明   |
|--------|--|
| 其他限制   | <ul style="list-style-type: none"> <li>同步对象的限制：                             <ul style="list-style-type: none"> <li>仅支持表级别的数据同步。</li> <li>不支持新增列的数据同步，即源库中同步对象表新增了某个列，该列的数据不会同步至目标DataHub实例。</li> </ul> </li> <li>请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> <li>不支持全量数据同步，即DTS不会将源PolarDB-X中同步对象的存量数据同步至目标DataHub实例中。</li> </ul> |
| 其他注意事项 | DTS会在源库定时更新`dts_health_check`.`ha_health_check`表以推进binlog位点。   |

### PolarDB-X 2.0同步至Elasticsearch

| 类型     | 说明   |
|--------|--|
| 源库限制   | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>若同步对象为表级别，且需进行编辑（如表列名映射），单次同步任务的表数量超过5000时，建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务，否则任务提交后可能会显示请求报错。</li> <li>Binlog日志：                             <ul style="list-style-type: none"> <li>在PolarDB-X 2.0控制台参数设置界面开启binlog，开启方法请参见<a href="#">参数设置</a>。并且binlog_row_image为full。否则预检查阶段提示报错，且无法成功启动数据同步任务。</li> <li>如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在全量同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> <li>PolarDB-X需兼容MySQL 5.7版本。</li> </ul> |
| 其他限制   | <ul style="list-style-type: none"> <li>不支持同步DDL操作，如果源库中待同步的表在同步的过程中，已经执行了DDL操作，您需要先移除同步对象，然后在Elasticsearch实例中移除该表对应的索引，最后新增同步对象。请参见<a href="#">移除同步对象</a>和<a href="#">新增同步对象</a>。</li> <li>如果源库中待同步的表需要执行增加列的操作，您只需先在Elasticsearch实例中修改对应表的mapping，然后在源数据库中执行相应的DDL操作，最后暂停并启动同步任务。</li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul>   |
| 其他注意事项 | DTS会在源库定时更新`dts_health_check`.`ha_health_check`表以推进binlog位点。   |

### PolarDB-X 2.0同步至MaxCompute

| 类型 | 说明 |
|----|----|
|----|----|

| 类型     | 说明   |
|--------|--|
| 源库限制   | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>若同步对象为表级别，且需进行编辑（如表列名映射），单次同步任务的表数量超过5000时，建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务，否则任务提交后可能会显示请求报错。</li> <li>Binlog日志：                     <ul style="list-style-type: none"> <li>在PolarDB-X 2.0控制台参数设置界面开启binlog，开启方法请参见<a href="#">参数设置</a>。并且binlog_row_image为full。否则预检查阶段提示报错，且无法成功启动数据同步任务。</li> <li>如为增量同步任务，DTS要求源数据库的本地Binlog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的本地Binlog日志至少保留7天以上（您可在全量同步完成后将Binlog保存时间设置为24小时以上），否则DTS可能因无法获取Binlog而导致任务失败，极端情况下可能会导致数据不一致或丢失。由于您所设置的Binlog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> <li>PolarDB-X需兼容MySQL 5.7版本。</li> </ul> |
| 其他限制   | <ul style="list-style-type: none"> <li>仅支持表级别的数据同步。</li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>若在DTS同步数据期间，有除DTS同步任务外的数据写入目标库，则会导致源库与目标库数据不一致。例如，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> <li>由于MaxCompute不支持主键约束，当DTS在同步数据时因网络等原因触发重传，可能会导致MaxCompute中出现重复记录。</li> </ul>   |
| 其他注意事项 | DTS会在源库定时更新`dts_health_check`.`ha_health_check`表以推进binlog位点。   |

### 3.5. 源为PolarDB-X的功能规范和约束说明

由于PolarDB-X 1.0自身不提供Binlog以及其他的一些限制，在业务设计、运维变更、数据质量以及业务开发时，会受到如下规范的约束，请您在实际使用中注意。

#### 概览

- [业务设计规范](#)
- [数据库架构规范](#)
- [运维变更规范](#)
- [数据质量风险声明](#)
- [业务开发的数据质量工作](#)

#### 业务设计规范

- 表都需要有主键，否则可能会造成数据不一致（造成目标库有重复数据）。
- 因为PolarDB-X 1.0的GSI全局二级索引（Global Secondary Index）具有异步特性，不推荐您使用该功能。如果您仍选择使用该功能，DTS只能保证最终数据的一致性。
- 数据库不支持混合模式，即unit模式和copy模式混合的同步链路。

 **说明** unit模式是多地用户分别在各自地域读写本地域的数据，且本地域的数据会和中心数据双向同步。copy模式是指此集群数据在中心数据库写入，完成后全量同步到各个单元。

- 如果是使用底层MySQL进行链路配置，PolarDB-X 1.0间双向同步，业务表不支持使用FLOAT、DOUBLE类型字段，需要改为decimal类型。如果为PolarDB-X 1.0间单向同步、迁移或订阅，则允许使用这两种数据类型。

- DTS不支持同步、迁移或订阅PolarDB-X 1.0中存储过程、触发器、函数、视图、事件等对象。
- 不支持PolarDB-X 1.0的结构初始化，需要在目标库手动创建好对应的库表等对象。
- PolarDB-X 1.0留够足够支撑业务增长的容量。
- 如PolarDB-X 1.0实例下挂载的MySQL版本为5.7和8.0，则不支持直接订阅该PolarDB-X 1.0实例，需要对该实例下的多个MySQL配置单独的数据订阅任务来实现对PolarDB-X 1.0数据的订阅和消费。

### 数据库架构规范

- 一个PolarDB-X 1.0实例使用到的RDS MySQL实例，不能再被其他PolarDB-X 1.0实例使用。
- PolarDB-X 1.0间的同步或迁移，两端对应的RDS MySQL实例需保持对等部署，比如源PolarDB-X 1.0使用了4个RDS MySQL实例，目标PolarDB-X 1.0也需要使用4个规格配置相同的RDS MySQL实例。
- 源和目标PolarDB-X 1.0的分库分表规则需要保持一致，否则DTS同步或迁移任务无法建立。
- 只能同步、迁移或订阅PolarDB-X 1.0实例的业务表，无法同步、迁移或订阅该实例的元数据表和系统表。

### 运维变更规范

| 变更类型            | 具体变更                                 | 影响及应对规范  |
|-----------------|--------------------------------------|--|
| PolarDB-X 1.0方面 | 分库分表变化（如变更分库分表键、或变更分库分表数量）场景。        | 暂不支持，需按如下步骤重新创建任务： <ol style="list-style-type: none"> <li>1. 停止并删除原先的DTS任务。</li> <li>2. 待源库变更完成后，清空目标库已经同步或迁移的数据。</li> <li>3. 对PolarDB-X 1.0下的多个RDS MySQL实例单独配置到目标库的同步或迁移任务。</li> </ol>                                |
|                 | 存储层实例个数发生变化场景（如扩容、热点表迁移等）。           |  |
| 存储层方面           | 存储层实例级别的规格变更、切换等。                    | 不影响DTS任务。  |
|                 | 参数修改。                                | 源库和目标库的参数需一致。存储层实例级别参数修改只允许做向下兼容的参数修改，即新参数不会影响老参数的行为和参数。<br><span style="background-color: #e6f2ff; padding: 2px;">❓ 说明 如不确认需联系数据库专家服务组。</span>  |
|                 | 存储层实例内备份恢复策略、开启审计、诊断等。               | 对当前实例有效，不涉及有复制关系的其他实例。   |
| DTS任务方面         | DDL操作。                               | 如果是配置PolarDB-X 1.0下的多个RDS MySQL到目标库的DTS任务，受限于MySQL的实现逻辑，执行DDL操作可能会导致DTS任务延迟。   |
| 库表级别的DDL操作      | 新增表。                                 | 不支持，需要按照如下步骤操作： <ol style="list-style-type: none"> <li>1. 在目标库执行建表DDL操作之后，在源库执行同样的建表DDL操作。</li> <li>2. 修改同步任务，将新表分别加入到同步对象中。</li> </ol> 上述操作全部完成后建表操作才算正式完成，业务才可以写入。如果是基于表级别的同步，在源和目标存储层实例上需要把新增的表加到同步队列里。             |
|                 | 增加字段、增加二级索引、删除索引、修改索引（二级索引改成唯一索引除外）。 | <ul style="list-style-type: none"> <li>● 如果是基于PolarDB-X 1.0级别配置任务，需要您先在目标库先做对应的DDL操作，然后再去源库执行对应的DDL操作。</li> <li>● 如果是基于PolarDB-X 1.0下挂载的RDS MySQL实例配置的任务，DTS会自动更新对应的新增字段，增加二级索引、删除索引、修改索引（二级索引改成唯一索引除外）的操作。</li> </ul> |
|                 | 除上述操作以外的其他DDL操作。                     | 禁止此类DDL操作。   |

| 变更类型  | 具体变更   | 影响及应对规范   |
|---|--|---|
| 切流操作<br><div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p><b>?</b> 说明 切流操作是指，通过DTS将数据从源库同步或迁移至目标库后，您将业务流量从源库切换至目标库。</p> </div> | 正常的切流。   | DTS提供的延迟检测功能判断无延时后，您才能进行安全切流，否则会产生数据质量问题。   |
|   | 满足恢复点目标RPO (Recovery Point Objective) 的异常切流。<br><div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p><b>?</b> 说明 RPO代表故障恢复后业务允许丢失的数据量，可用时间表示。</p> </div> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p><b>!</b> 警告 异常切流是指源实例或者源实例所在的数据中心发生故障时进行的切流操作。这类操作都是有损操作。</p> </div> | 在出现故障（如网络中断、机房批量设备故障或互联网数据中心IDC故障），且DTS任务存在延迟的情况下，此时如更新至目标库的最后一条数据的时间与故障发生时的时间之差小于RPO（如5分钟），则可以业务优先恢复为准则切流。切流后可能有5分钟以内的数据质量问题，需要您的业务开发配合订正，确保数据的一致性。                                |
|   | 不满足RPO的异常切流。   | 源端执行大量DDL操作、网络、目的库性能等问题，可能会导致DTS任务存在延迟，此时如刚好遇到机房故障，且同步或迁移至目标库的最后一条数据的时间戳与故障发生时的时间戳之差大于RPO（如5分钟），那么切流需要非常谨慎，建议您暂时不要执行切流操作，选择等待机房恢复正常。如果遇到异常切流，会存在延时窗口内的数据质量问题，需要您的业务开发配合订正，确保数据的一致性。 |

### 数据质量风险声明

一些变更或切流操作可能会导致源库和目标库结构不一致等数据质量问题，具体示例如下：

- 当源实例主备之间存在数据延迟时，新写入主库的数据未能及时更新至备库。此时，如源实例进行主备切换，DTS会使用源实例的备库作为源数据库进行数据同步、迁移或订阅，从而导致丢失未能及时更新至备库的部分数据。
- 在断网、业务切流后，如DTS任务恢复正常，会自动启动重试机制，重新同步、迁移或订阅故障发生前一段时间的数据，以避免目标库数据丢失。在这种情况下，如目标表缺少主键，会导致源目库数据不一致；如目标表存在主键，则在重试机制过程中源目库数据不一定能保持一致，但在重试结束后数据将保持一致。
- 网络问题或业务DDL，导致DTS任务延时。
- 源库变更、目标库性能、任何原因使表结构不一致等问题导致DTS任务延时或中断。

阿里云无法解决以上问题，需要您重构链路或者自行调整源库和目标库。

### 业务开发的数据质量工作

- 请您慎重执行所有的DDL操作，所有的DDL操作都要经过驻场同学的确认，以遵守上述日常DDL变更规范。
- 请勿在程序代码中直接进行DDL操作。

## 3.6. 源为PolarDB-X的任务配置方案

DTS支持将源PolarDB-X 1.0实例同步或迁移至目标库。但是由于直接使用PolarDB-X 1.0作为源实例配置DTS任务，当该实例下挂载的RDS MySQL数量超过2个时，可能会产生性能瓶颈，稳定性风险，最终可能会对业务造成影响。为了更好地保障任务的性能和稳定性，DTS建议您在配置时，根据PolarDB-X 1.0下挂载的RDS MySQL实例，分别配置DTS任务。

### 前提条件

您已阅读[源为PolarDB-X的功能规范和约束说明](#)，已确保遵循相关规范，了解注意事项。

### 任务配置方案

| 方案 | 配置方式 | 特点总结 |
|----|------|------|
|    |      |      |

| 方案  | 配置方式   | 特点总结   |
|-----|--|--|
| 方案一 | <p>针对PolarDB-X 1.0下的多个RDS MySQL实例，配置多个同步或迁移至目标实例的DTS任务，并且在配置过程中，需将多个RDS MySQL的数据库表名都映射成目标实例的库表名。</p> <p>该方案相比于直接配置PolarDB-X 1.0至目标实例的同步或迁移任务，更能保障性能和稳定性。</p> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin-top: 10px;"> <p> <b>注意</b> 该方案仅支持表级别同步或迁移。</p> </div> | <p>推荐您使用方案一，其在性能、稳定性方面均优于方案二，且对PolarDB-X 1.0下挂载的RDS MySQL数量没有限制。</p> |
| 方案二 | <p>直接使用PolarDB-X 1.0作为源实例配置同步或迁移至目标实例任务，当该实例下挂载的RDS MySQL实例数量超过2个后，可能会对DTS任务的稳定性、操作性产生较大影响，严重时甚至将影响业务。</p>   | <p>使用方案二时，如源PolarDB-X 1.0下挂载的RDS MySQL实例超过2个时，DTS任务可能会有性能和稳定性风险。</p> |

### 方案对比

| 对比项  | 方案一   | 方案二  |
|------|---|--|
| 性能   | 拆分成多个DTS任务，性能成倍增长，可承载大规模的PolarDB-X 1.0数据写入，   | 仅配置一个以PolarDB-X 1.0为源的DTS任务，当业务系统写入源实例数据量较大时，会存在性能瓶颈。          |
| 稳定性  | 稳定性较强。<br>配置PolarDB-X 1.0下多个RDS MySQL至目标库的DTS任务，如其中一个DTS任务发生故障，则不影响其他DTS任务的运行，且只需恢复发生故障的DTS任务即可。            | 稳定性一般。<br>仅配置一个以PolarDB-X 1.0为源的DTS任务，如该DTS任务发生故障，则整个DTS任务会失败。 |
| 易用性  | 需配置多次，相对较繁琐。需要拆分成独立的任务，且每次配置均需做库表名映射，将源实例多个RDS MySQL的库表名映射成目标实例中库表的名称，如源实例中RDS MySQL实例的数量和库表数量较多，则需配置多次的映射。 | 配置方便，仅需配置一个PolarDB-X 1.0为源的DTS任务。                              |
| 资源占用 | 拆分成n个独立的DTS任务，消耗n个DTS实例。  | 消耗1个DTS实例。   |

### 两种方案如何实现DDL同步或增量迁移

目前PolarDB-X 1.0为源的同步或迁移任务，暂不支持同步或增量迁移DDL操作，如源PolarDB-X 1.0在同步或迁移时执行DDL操作，为避免数据写入目标端失败，一般解决方案如下：

1. 释放DTS任务。
2. 清空目标库。
3. 重新配置任务。

在PolarDB-X 1.0为源的部分同步或迁移场景中，支持在保留DTS任务的同时，实现同步或增量迁移DDL操作，具体场景及对应的操作，如下所示：

| 场景              | 操作   |
|-----------------|--|
| 使用方案一，仅支持表级别同步。 | <ul style="list-style-type: none"> <li>• 支持通过修改同步对象来实现加表的操作。</li> <li>• 目前仅PolarDB-X 1.0间的同步，支持加列、减列的操作，且需按如下操作执行：                         <ol style="list-style-type: none"> <li>i. 修改同步对象，使其包含或删除列。</li> <li>ii. 先在目标库手动进行加列或减列，然后在源端执行对应的加列或减列操作。当目标库已经存在这个列时，则DTS会忽略这个报错，且不会显示写入失败。</li> </ol> </li> </ul> |

| 场景                 | 操作   |
|--------------------|--|
| 使用方案二，且同步或迁移对象为整库。 | <ul style="list-style-type: none"> <li>仅支持加表的操作，且需先在目标库手动加上新表，然后在源端执行加表的操作。</li> <li>不支持加列、减列的操作。</li> </ul> <div style="border: 1px solid #ccc; background-color: #fff9c4; padding: 5px; margin-top: 10px;"> <p> <b>警告</b> 在源PolarDB-X 1.0执行加列操作，可能会导致底层多个物理表不一致，有的缺少该列，有的包含该列。从而DTS在拼装的SQL语句，可能会找不到该列，或者丢失该列的数据。</p> </div> |
| 使用方案二，且同步对象为非整库。   | <ul style="list-style-type: none"> <li>仅支持加表的操作，且需按如下操作执行：                             <ol style="list-style-type: none"> <li>i. 修改同步对象，使其包含该新表。</li> <li>ii. 先在目标库手动加上新表，然后在源端执行加表的操作。</li> </ol> </li> <li>不支持加列、减列的操作。</li> </ul>   |

### 3.7. 同步时源库为Oracle的注意事项及限制

如果同步的源数据库类型为自建Oracle，您需要在配置具体的同步任务前，参考本文的注意事项及限制，以保障数据同步任务的正常运行。

#### 源库为Oracle的同步方案概览

根据如下同步方案，查看同步任务的注意事项及限制：

 **说明**

DTS默认同步到目标数据库中时会取消外键约束，因此源数据库的级联、删除等操作不会同步到云原生数据仓库AnalyticDB PostgreSQL。

- [自建Oracle同步至云原生数据仓库AnalyticDB PostgreSQL版](#)
- [自建Oracle同步至阿里云消息队列Kafka版或自建Kafka](#)
- [自建Oracle同步至阿里云流式数据服务DataHub](#)
- [自建Oracle同步至PolarDB-X 2.0](#)

#### 自建Oracle同步至云原生数据仓库AnalyticDB PostgreSQL版

| 类型 | 说明 |
|----|----|
|----|----|

| 类型   | 说明   |
|------|--|
| 源库限制 | <ul style="list-style-type: none"> <li>同步对象要求：                             <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>如果您的自建Oracle版本为12c及以上，待同步表的名称长度需不超过30个字节。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> </ul> </li> <li>如果源库通过专线的方式接入，那么需要将其中任意1个VIP配置到连接信息中，实现Oracle RAC通过专线接入同步任务。</li> <li>如自建Oracle为RAC架构，不支持配置ScanIP，仅支持将其中任意1个VIP配置到连接信息中，该方式配置后不支持RAC的节点切换。</li> <li>需开启Redo Log和Archive Log。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p> <b>说明</b></p> <p>如为增量同步任务，DTS要求源数据库的Redo Log和Archive Log保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的Redo Log和Archive Log至少保留7天以上（您可在全量同步完成后将Redo Log和Archive Log保存时间设置为24小时以上），否则DTS可能因无法获取Redo Log或Archive Log而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Redo Log或Archive Log保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</p> </div> <ul style="list-style-type: none"> <li>如果源库在同步阶段进行主备切换，将会导致同步任务失败。</li> </ul>   |
| 其他限制 | <ul style="list-style-type: none"> <li>结构初始化的对象：table、index、constraint、function、sequence和view。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p> <b>警告</b></p> <ul style="list-style-type: none"> <li>此场景属于异构数据库间的数据同步，DTS在执行结构初始化时数据类型无法完全对应，请谨慎评估数据类型的映射关系对业务的影响，详情请参见<a href="#">结构初始化涉及的数据类型映射关系</a>。</li> <li>不兼容触发器。建议您删除源库的触发器，以避免因触发器而导致数据不一致的问题。详情请参见<a href="#">源库存在触发器时如何配置同步作业</a>。</li> <li>对于分区表，DTS会丢弃分区定义，您需要在目标库自行定义。</li> </ul> </div> <ul style="list-style-type: none"> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>由于DTS的延迟时间是根据同步到目标库最后一条数据的时间戳和当前时间戳对比得出，源库长时间未执行DML操作可能导致延迟信息不准确。如果任务显示的延迟时间过大，您可以在源库执行一个DML操作来更新延迟信息。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p> <b>说明</b> 如果同步对象选择为整库，您还可以创建心跳表，心跳表每秒定期更新或者写入数据。</p> </div> <ul style="list-style-type: none"> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul> |

### 自建Oracle同步至阿里云消息队列Kafka版或自建Kafka

| 类型   | 说明   |
|------|--|
| 源库限制 | <ul style="list-style-type: none"> <li>同步对象要求：                             <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出現重复数据。</li> <li>如果您的自建Oracle版本为12c及以上，待同步表的名称长度需不超过30个字节。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> </ul> </li> <li>如果源库通过专线的方式接入，那么需要将其中任意1个VIP配置到连接信息中，实现Oracle RAC通过专线接入同步任务。</li> <li>如自建Oracle为RAC架构，不支持配置ScanIP，仅支持将其中任意1个VIP配置到连接信息中，该方式配置后不支持RAC的节点切换。</li> <li>需开启Redo Log和Archive Log。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p> <b>说明</b></p> <p>如为增量同步任务，DTS要求源数据库的Redo Log和Archive Log保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的Redo Log和Archive Log至少保留7天以上（您可在全量同步完成后将Redo Log和Archive Log保存时间设置为24小时以上），否则DTS可能因无法获取Redo Log或Archive Log而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Redo Log或Archive Log保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</p> </div> <ul style="list-style-type: none"> <li>如果源库在同步阶段进行主备切换，将会导致同步任务失败。</li> </ul> |
| 其他限制 | <ul style="list-style-type: none"> <li>不支持自动调整同步对象，如果对同步对象中的数据表进行重命名操作，且重命名后的名称不在同步对象中，那么这部分数据将不再同步到目标Kafka集群中。如需将修改后的数据表继续数据同步至目标Kafka集群中，您需要进行<b>修改同步对象</b>操作，详情请参见<a href="#">新增同步对象</a>。</li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>由于DTS的延迟时间是根据同步到目标库最后一条数据的时间戳和当前时间戳对比得出，源库长时间未执行DML操作可能导致延迟信息不准确。如果任务显示的延迟时间过大，您可以在源库执行一个DML操作来更新延迟信息。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p> <b>说明</b> 如果同步对象选择为整库，您还可以创建心跳表，心跳表每秒定期更新或者写入数据。</p> </div> <ul style="list-style-type: none"> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul>   |

### 自建Oracle同步至阿里云流式数据服务DataHub

| 类型 | 说明 |
|----|----|
|    |    |

| 类型   | 说明   |
|------|--|
| 源库限制 | <ul style="list-style-type: none"> <li>同步对象要求：                             <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>如果您的自建Oracle版本为12c及以上，待同步表的名称长度需不超过30个字节。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> </ul> </li> <li>如果源库通过专线的方式接入，那么需要将其中任意1个VIP配置到连接信息中，实现Oracle RAC通过专线接入同步任务。</li> <li>如自建Oracle为RAC架构，不支持配置ScanIP，仅支持将其中任意1个VIP配置到连接信息中，该方式配置后不支持RAC的节点切换。</li> <li>需开启Redo Log和Archive Log。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin: 10px 0;"> <p> <b>说明</b></p> <p>如为增量同步任务，DTS要求源数据库的Redo Log和Archive Log保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的Redo Log和Archive Log至少保留7天以上（您可在全量同步完成后将Redo Log和Archive Log保存时间设置为24小时以上），否则DTS可能因无法获取Redo Log或Archive Log而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Redo Log或Archive Log保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</p> </div> <ul style="list-style-type: none"> <li>如果源库在同步阶段进行主备切换，将会导致同步任务失败。</li> </ul> |
| 其他限制 | <ul style="list-style-type: none"> <li>结构初始化对象为table。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin: 10px 0;"> <p> <b>警告</b> 不兼容触发器。建议您删除源库的触发器，以避免因触发器而导致数据不一致的问题。详情请参见<a href="#">源库存在触发器时如何配置同步作业</a>。</p> </div> <ul style="list-style-type: none"> <li>暂不支持全量同步，即DTS不会将源实例中同步对象的存量数据同步至目标DataHub实例。</li> <li>仅支持表级别的数据同步。</li> <li>请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>由于DTS的延迟时间是根据同步到目标库最后一条数据的时间戳和当前时间戳对比得出，源库长时间未执行DML操作可能导致延迟信息不准确。如果任务显示的延迟时间过大，您可以在源库执行一个DML操作来更新延迟信息。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin: 10px 0;"> <p> <b>说明</b> 如果同步对象选择为整库，您还可以创建心跳表，心跳表每秒定期更新或者写入数据。</p> </div> <ul style="list-style-type: none"> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul>  |

### 自建Oracle同步至PolarDB-X 2.0

| 类型 | 说明 |
|----|----|
|----|----|

| 类型   | 说明   |
|------|--|
| 源库限制 | <ul style="list-style-type: none"> <li>同步对象要求： <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>如果您的自建Oracle版本为12c及以上，待同步表的名称长度需不超过30个字节。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> </ul> </li> <li>如果源库通过专线的方式接入，那么需要将其中任意1个VIP配置到连接信息中，实现Oracle RAC通过专线接入同步任务。</li> <li>如自建Oracle为RAC架构，不支持配置ScanIP，仅支持将其中任意1个VIP配置到连接信息中，该方式配置后不支持RAC的节点切换。</li> <li>需开启Redo Log和Archive Log。 <div style="border: 1px solid #ccc; padding: 5px; margin: 5px 0;"> <p> 说明</p> <p>如为增量同步任务，DTS要求源数据库的Redo Log和Archive Log保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的Redo Log和Archive Log至少保留7天以上（您可在全量同步完成后将Redo Log和Archive Log保存时间设置为24小时以上），否则DTS可能因无法获取Redo Log或Archive Log而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Redo Log或Archive Log保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</p> </div> </li> <li>如果源库在同步阶段进行主备切换，将会导致同步任务失败。</li> </ul> |
| 其他限制 | <ul style="list-style-type: none"> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>由于DTS的延迟时间是根据同步到目标库最后一条数据的时间戳和当前时间戳对比得出，源库长时间未执行DML操作可能导致延迟信息不准确。如果任务显示的延迟时间过大，您可以在源库执行一个DML操作来更新延迟信息。 <div style="border: 1px solid #ccc; padding: 5px; margin: 5px 0;"> <p> 说明 如果同步对象选择为整库，您还可以创建心跳表，心跳表每秒定期更新或者写入数据。</p> </div> </li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul>  |

## 3.8. 同步时源库为PostgreSQL的注意事项及限制

如果同步的源数据库类型为PostgreSQL，如RDS PostgreSQL、自建PostgreSQL，您需要在配置具体的同步任务前，参考本文的注意事项及限制，以保障数据同步任务的正常运行。

### 源库为PostgreSQL的同步方案概览

根据同步方案，查看同步任务的注意事项及限制：

说明

DTS默认同步到目标数据库中时会取消外键约束，因此源数据库的级联、删除等操作不会同步到如下目标数据库：

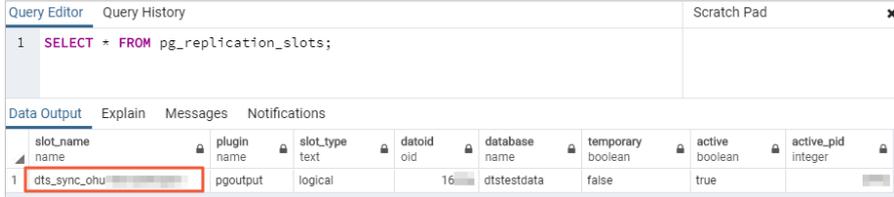
- RDS PostgreSQL
- 云原生数据仓库AnalyticDB PostgreSQL
- PolarDB O
- RDS MySQL

- PostgreSQL间同步
- 自建PostgreSQL或RDS PostgreSQL同步至云原生数据仓库AnalyticDB PostgreSQL
- 自建PostgreSQL同步至PolarDB O引擎
- RDS PostgreSQL同步至RDS MySQL

### PostgreSQL间同步

- RDS PostgreSQL间的单向同步

| 类型   | 说明   |
|------|--|
| 源库限制 | <ul style="list-style-type: none"> <li>◦ 待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>◦ 若同步对象为表级别，且需进行编辑（如表列名映射），单次同步任务的表数量超过5000时，建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务，否则任务提交后可能会显示请求报错。</li> <li>◦ WAL日志：             <ul style="list-style-type: none"> <li>▪ 需开启，即设置wal_level参数的值为logical。</li> <li>▪ 如为增量同步任务，DTS要求源数据库的WAL日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的WAL日志至少保留7天以上（您可在全量同步完成后将WAL日志保存时间设置为24小时以上），否则DTS可能因无法获取WAL日志而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的WAL日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> <li>◦ 源库的操作限制<br/>为保障同步任务的正常进行，避免主备切换导致的逻辑订阅中断，需要RDS PostgreSQL支持并开启Logical Replication Slot Failover，设置方式，请参见<a href="#">逻辑复制槽故障转移（Logical Replication Slot Failover）</a>。</li> </ul> |
|      | <ul style="list-style-type: none"> <li>◦ 单个数据同步任务只能同步一个数据库，如需同步多个数据库，您需要为每个数据库配置数据同步任务。</li> <li>◦ 同步期间，如果同步对象的选择粒度为Schema，您需要在对该表写入数据前执行 <code>ALTER TABLE schema.table REPLICA IDENTITY FULL;</code> 命令，以保障该表数据同步的一致性。执行该命令期间，建议您不要有锁表操作，否则会导致表锁死。</li> </ul> <div style="background-color: #e6f2ff; padding: 10px; margin: 10px 0;"> <p>说明</p> <ul style="list-style-type: none"> <li>▪ 将上述命令中的 <code>schema</code> 和 <code>table</code> 替换成真实的Schema名和表名。</li> <li>▪ 建议您在业务低峰期进行操作。</li> </ul> </div> <ul style="list-style-type: none"> <li>◦ DTS的校验对象为数据内容，暂不支持Sequence等元数据的校验，您需要自行校验。</li> </ul>   |

|             |  |
|-------------|--|
| <p>类型</p>   | <p>说明</p> <p>由于业务切换到目标端后，新写入的Sequence不会按照源库的Sequence最大值作为初始值去递增，您需要在业务切换前，在源库中查询对应Sequence的最大值，然后在目标库中将其作为对应Sequence的初始值。查询源库Sequence值的相关命令如下：</p>   |
| <p>其他限制</p> | <pre>do language plpgsql \$\$ declare   nsp name;   rel name;   val int8; begin   for nsp,rel in select nspname,relname from pg_class t2 , pg_namespace t3   where t2.relnamespace=t3.oid and t2.relkind='S'   loop     execute format(\$\$_select last_value from %I.%I\$_\$, nsp, rel) into val;     raise notice '%',       format(\$\$_select setval('%I.%I':regclass, %s);\$_\$, nsp, rel, val+1);   end loop; end; \$\$;</pre> <ul style="list-style-type: none"> <li>○ DTS会在源库中创建以下临时表，用于获取增量数据的DDL、增量表的结构、心跳信息等。在同步期间，请勿删除源库中的临时表，否则会导致DTS任务异常。临时表会在DTS实例释放后自动删除。<br/>public.DTS_PG_CLASS、public.DTS_PG_ATTRIBUTE、public.DTS_PG_TYPE、public.DTS_PG_ENUM、public.DTS_POSTGRES_HEARTBEAT、public.DTS_DDL_COMMA ND、public.DTS_ARGS_SESSION。</li> <li>○ 为保障同步延迟时间展示的准确性，DTS会在源库中新增一个名为dts_postgres_heartbeat的心跳表。</li> <li>○ 同步期间，DTS会在源库中创建前缀为dts_sync_的replication slot用于复制数据。DTS会每隔120分钟自动清理一次历史replication slot，避免其持续累积占用磁盘空间。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin: 10px 0;"> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>■ DTS释放实例后会主动删除该replication slot，如果您在同步期间修改了数据库密码，或者删除了DTS的访问IP白名单，则会导致该replication slot无法自动删除，此时需要您在源库手动删除，避免其持续累积占用磁盘空间导致RDS PostgreSQL实例不可用。</li> <li>■ 当释放同步任务或同步失败时，DTS会主动清理该replication slot；如果RDS PostgreSQL发生了主备切换，则需要您登录备库来手动清理。</li> </ul> </div>  <ul style="list-style-type: none"> <li>○ 执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>○ 全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>○ 对于表级别的数据同步，请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>○ 对于表级别的数据同步，如果除DTS以外的数据没有写入目标库，您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，请参见<a href="#">不锁表结构变更</a>。</li> <li>○ 在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> <li>○ 若为全量同步和增量同步任务，且源数据库中要同步的表包含外键、触发器、事件触发器（PostgreSQL 11.5及以上版本），则目标数据库中的session_replication_role参数必须设置为replica，在DTS同步任务释放后可以将session_replication_role参数改回origin。</li> </ul> |

| 类型   | 说明  |
|------|---|
| 特殊情况 | 当源实例为RDS PostgreSQL时<br>同步期间，请勿修改RDS PostgreSQL的连接地址和可用区，否则会导致同步失败。 |

● 自建PostgreSQL同步至RDS PostgreSQL

| 类型   | 说明   |
|------|--|
| 源库限制 | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>若同步对象为表级别，且需进行编辑（如表列名映射），单次同步任务的表数量超过5000时，建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务，否则任务提交后可能会显示请求报错。</li> <li>WAL日志：                             <ul style="list-style-type: none"> <li>需开启，即设置wal_level参数的值为logical。</li> <li>如为增量同步任务，DTS要求源数据库的WAL日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的WAL日志至少保留7天以上（您可在全量同步完成后将WAL日志保存时间设置为24小时以上），否则DTS可能因无法获取WAL日志而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的WAL日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> <li>如自建PostgreSQL进行主备切换，会导致同步失败。</li> </ul>  |
|      | <ul style="list-style-type: none"> <li>单个数据同步任务只能同步一个数据库，如需同步多个数据库，您需要为每个数据库配置数据同步任务。</li> <li>同步期间，如果同步对象的选择粒度为Schema，您需要在对该表写入数据前执行 <code>ALTER TABLE schema.table REPLICA IDENTITY FULL;</code> 命令，以保障该表数据同步的一致性。执行该命令期间，建议您不要有锁表操作，否则会导致表锁死。</li> </ul> <div style="background-color: #e6f2ff; padding: 10px; margin: 10px 0;"> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>将上述命令中的 <code>schema</code> 和 <code>table</code> 替换成真实的Schema名和表名。</li> <li>建议您在业务低峰期进行操作。</li> </ul> </div> <ul style="list-style-type: none"> <li>DTS的校验对象为数据内容，暂不支持Sequence等元数据的校验，您需要自行校验。</li> <li>由于业务切换到目标端后，新写入的Sequence不会按照源库的Sequence最大值作为初始值去递增，您需要在业务切换前，在源库中查询对应Sequence的最大值，然后在目标库中将其作为对应Sequence的初始值。查询源库Sequence值的相关命令如下：</li> </ul> <pre style="background-color: #f0f0f0; padding: 10px; margin: 10px 0;">do language plpgsql \$\$ declare     nsp name;     rel name;     val int8; begin     for nsp,rel in select nspname,relname from pg_class t2 , pg_namespace t3     where t2.relnamespace=t3.oid and t2.relkind='S'     loop         execute format(\$_\$select last_value from %I.%I\$_\$, nsp, rel) into val;         raise notice '%',             format(\$_\$select setval('%I.%I'::regclass, %s);\$_\$, nsp, rel, val+1);     end loop; end; \$\$;</pre> |

其他限制

| 其他限制类型         | <p>说明</p> <ul style="list-style-type: none"> <li>DTS会在源库中创建以下临时表，用于获取增量数据的DDL、增量表的结构、心跳信息等。在同步期间，请勿删除源库中的临时表，否则会导致DTS任务异常。临时表会在DTS实例释放后自动删除。<br/> <code>public.DTS_PG_CLASS</code>、<code>public.DTS_PG_ATTRIBUTE</code>、<code>public.DTS_PG_TYPE</code>、<code>public.DTS_PG_ENUM</code>、<code>public.DTS_POSTGRES_HEARTBEAT</code>、<code>public.DTS_DDL_COMMA</code>、<code>public.DTS_ARGS_SESSION</code>。</li> <li>为保障同步延迟时间展示的准确性，DTS会在源库中新增一个名为<code>dts_postgres_heartbeat</code>的心跳表。</li> <li>同步期间，DTS会在源库中创建前缀为<code>dts_sync_</code>的replication slot用于复制数据。DTS会每隔120分钟自动清理一次历史replication slot，避免其持续累积占用磁盘空间。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin: 10px 0;"> <p><span style="color: #00aaff;">?</span> 说明 当释放同步任务或同步失败时，DTS会主动清理该replication slot；如果RDS PostgreSQL发生了主备切换，则需要您登录备库来手动清理。</p> </div> <div style="border: 1px solid #ccc; padding: 5px; margin: 10px 0;"> <p>Query Editor Query History Scratch Pad</p> <pre>1 SELECT * FROM pg_replication_slots;</pre> <p>Data Output Explain Messages Notifications</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>slot_name</th> <th>plugin_name</th> <th>slot_type</th> <th>datoid</th> <th>database</th> <th>temporary</th> <th>active</th> <th>active_pid</th> </tr> <tr> <th>name</th> <th>name</th> <th>text</th> <th>oid</th> <th>name</th> <th>boolean</th> <th>boolean</th> <th>integer</th> </tr> </thead> <tbody> <tr> <td>1 dts_sync_ohu</td> <td>pgoutput</td> <td>logical</td> <td>16</td> <td>dtstestdata</td> <td>false</td> <td>true</td> <td></td> </tr> </tbody> </table> </div> <ul style="list-style-type: none"> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>对于表级别的数据同步，请勿对源库的同步对象使用<code>gh-ost</code>或<code>pt-online-schema-change</code>等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>对于表级别的数据同步，如果除DTS以外的数据没有写入目标库，您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，请参见<a href="#">不锁表结构变更</a>。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> <li>若为全量同步和增量同步任务，且源数据库中要同步的表包含外键、触发器、事件触发器（PostgreSQL 11.5及以上版本），则目标数据库中的<code>session_replication_role</code>参数必须设置为<code>replica</code>，在DTS同步任务释放后可以将<code>session_replication_role</code>参数改回<code>origin</code>。</li> </ul> | slot_name | plugin_name | slot_type   | datoid    | database | temporary  | active | active_pid | name | name | text | oid | name | boolean | boolean | integer | 1 dts_sync_ohu | pgoutput | logical | 16 | dtstestdata | false | true |  |
|----------------|--|-----------|-------------|-------------|-----------|----------|------------|--------|------------|------|------|------|-----|------|---------|---------|---------|----------------|----------|---------|----|-------------|-------|------|--|
| slot_name      | plugin_name  | slot_type | datoid      | database    | temporary | active   | active_pid |        |            |      |      |      |     |      |         |         |         |                |          |         |    |             |       |      |  |
| name           | name   | text      | oid         | name        | boolean   | boolean  | integer    |        |            |      |      |      |     |      |         |         |         |                |          |         |    |             |       |      |  |
| 1 dts_sync_ohu | pgoutput   | logical   | 16          | dtstestdata | false     | true     |            |        |            |      |      |      |     |      |         |         |         |                |          |         |    |             |       |      |  |

● RDS PostgreSQL间的双向同步

|          |  |
|----------|--|
| 类型       | 说明   |
| 源和目标库的限制 | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>若同步对象为表级别，且需进行编辑（如表列名映射），单次同步任务的表数量超过5000时，建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务，否则任务提交后可能会显示请求报错。</li> <li>WAL日志： <ul style="list-style-type: none"> <li>需开启，即设置<code>wal_level</code>参数的值为<code>logical</code>。</li> <li>如为增量同步任务，DTS要求源数据库的WAL日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的WAL日志至少保留7天以上（您可在全量同步完成后将WAL日志保存时间设置为24小时以上），否则DTS可能因无法获取WAL日志而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的WAL日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> <li>为保障同步任务的正常进行，避免主备切换导致的逻辑订阅中断，需要RDS PostgreSQL支持并开启Logical Replication Slot Failover，设置方式，请参见<a href="#">逻辑复制槽故障转移（Logical Replication Slot Failover）</a>。</li> </ul> |
|          | <ul style="list-style-type: none"> <li>目前仅支持在华东1（杭州）、华东2（上海）、华北1（青岛）地域配置。</li> </ul>   |

| <p>类型</p>      | <p>○ 单个数据同步任务只能同步一个数据库，如需同步多个数据库，您需要为每个数据库配置数据同步任务。</p>  |           |             |             |           |          |            |        |            |      |  |      |     |      |         |         |         |                |          |         |    |             |       |      |  |
|----------------|--|-----------|-------------|-------------|-----------|----------|------------|--------|------------|------|--|------|-----|------|---------|---------|---------|----------------|----------|---------|----|-------------|-------|------|--|
| <p>其他限制</p>    | <p>○ 同步期间，如果同步对象的选择粒度为Schema，您需要在该表写入数据前执行 <code>ALTER TABLE schema.table REPLICA IDENTITY FULL;</code> 命令，以保障该表数据同步的一致性。执行该命令期间，建议您不要有锁表操作，否则会导致表锁死。</p> <div style="background-color: #e6f2ff; padding: 10px; border: 1px solid #d9e1f2;"> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>■ 将上述命令中的 <code>schema</code> 和 <code>table</code> 替换成真实的Schema名和表名。</li> <li>■ 建议您在业务低峰期进行操作。</li> </ul> </div> <p>○ DTS的校验对象为数据内容，暂不支持Sequence等元数据的校验，您需要自行校验。</p> <p>○ 由于业务切换到目标端后，新写入的Sequence不会按照源库的Sequence最大值作为初始值去递增，您需要在业务切换前，在源库中查询对应Sequence的最大值，然后在目标库中将其作为对应Sequence的初始值。查询源库Sequence值的相关命令如下：</p> <pre>do language plpgsql \$\$ declare   nsp name;   rel name;   val int8; begin   for nsp,rel in select nspname,relname from pg_class t2 , pg_namespace t3   where t2.relnamespace=t3.oid and t2.relkind='S'   loop     execute format(\$\$_select last_value from %I.%I\$_\$, nsp, rel) into val;     raise notice '%',       format(\$\$_select setval('%I.%I':regclass, %s);\$_\$, nsp, rel, val+1);   end loop; end; \$\$;</pre> <p>○ DTS会在源库中创建以下临时表，用于获取增量数据的DDL、增量表的结构、心跳信息等。在同步期间，请勿删除源库中的临时表，否则会导致DTS任务异常。临时表会在DTS实例释放后自动删除。</p> <p><code>public.DTS_PG_CLASS</code>、<code>public.DTS_PG_ATTRIBUTE</code>、<code>public.DTS_PG_TYPE</code>、<code>public.DTS_PG_ENUM</code>、<code>public.DTS_POSTGRES_HEARTBEAT</code>、<code>public.DTS_DDL_COMMAND</code>、<code>public.DTS_ARGS_SESSION</code>。</p> <p>○ 为保障同步延迟时间展示的准确性，DTS会在源库中新增一个名为 <code>dts_postgres_heartbeat</code> 的心跳表。</p> <p>○ 同步期间，DTS会在源库中创建前缀为 <code>dts_sync_</code> 的replication slot用于复制数据。DTS会每隔120分钟自动清理一次历史replication slot，避免其持续累积占用磁盘空间。</p> <div style="background-color: #e6f2ff; padding: 10px; border: 1px solid #d9e1f2;"> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>■ DTS释放实例后会主动删除该replication slot，如果您在同步期间修改了数据库密码，或者删除了DTS的访问IP白名单，则会导致该replication slot无法自动删除，此时需要您在源库手动删除，避免其持续累积占用磁盘空间导致RDS PostgreSQL实例不可用。</li> <li>■ 当释放同步任务或同步失败时，DTS会主动清理该replication slot；如果RDS PostgreSQL发生了主备切换，则需要您登录备库来手动清理。</li> </ul> </div> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p>Query Editor Query History Scratch Pad</p> <pre>1 SELECT * FROM pg_replication_slots;</pre> <p>Data Output Explain Messages Notifications</p> <table border="1"> <thead> <tr> <th>slot_name</th> <th>plugin_name</th> <th>slot_type</th> <th>datoid</th> <th>database</th> <th>temporary</th> <th>active</th> <th>active_pid</th> </tr> <tr> <th>name</th> <th></th> <th>text</th> <th>oid</th> <th>name</th> <th>boolean</th> <th>boolean</th> <th>integer</th> </tr> </thead> <tbody> <tr> <td>1 dts_sync_ohu</td> <td>pgoutput</td> <td>logical</td> <td>16</td> <td>dtstestdata</td> <td>false</td> <td>true</td> <td></td> </tr> </tbody> </table> </div> <p>○ 执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</p> | slot_name | plugin_name | slot_type   | datoid    | database | temporary  | active | active_pid | name |  | text | oid | name | boolean | boolean | integer | 1 dts_sync_ohu | pgoutput | logical | 16 | dtstestdata | false | true |  |
| slot_name      | plugin_name  | slot_type | datoid      | database    | temporary | active   | active_pid |        |            |      |  |      |     |      |         |         |         |                |          |         |    |             |       |      |  |
| name           |  | text      | oid         | name        | boolean   | boolean  | integer    |        |            |      |  |      |     |      |         |         |         |                |          |         |    |             |       |      |  |
| 1 dts_sync_ohu | pgoutput   | logical   | 16          | dtstestdata | false     | true     |            |        |            |      |  |      |     |      |         |         |         |                |          |         |    |             |       |      |  |

|      |  |
|------|--|
| 类型   | <ul style="list-style-type: none"> <li>◦ 全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>◦ 对于表级别的数据同步，请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>◦ 对于表级别的数据同步，如果除DTS以外的数据没有写入目标库，您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，请参见<a href="#">不锁表结构变更</a>。</li> <li>◦ 在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> <li>◦ 若为全量同步和增量同步任务，且源数据库中要同步的表包含外键、触发器、事件触发器（PostgreSQL 11.5及以上版本），则目标数据库中的session_replication_role参数必须设置为replica，在DTS同步任务释放后可以将session_replication_role参数改回origin。</li> <li>◦ 若DTS正向和反向同步任务均同步了某个表，且正向同步任务同步了该表的全量和增量数据，则反向同步任务只同步该表的增量数据。</li> </ul> |
| 特殊情况 | <p>当源实例为RDS PostgreSQL时</p> <p>同步期间，请勿修改RDS PostgreSQL的连接地址和可用区，否则会导致同步失败。</p>   |

### 自建PostgreSQL或RDS PostgreSQL同步至云原生数据仓库AnalyticDB PostgreSQL

| 类型   | 说明  |
|------|---|
| 源库限制 | <ul style="list-style-type: none"> <li>• 待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>• 若同步对象为表级别，且需进行编辑（如表列名映射），单次同步任务的表数量超过5000时，建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务，否则任务提交后可能会显示请求报错。</li> <li>• WAL日志： <ul style="list-style-type: none"> <li>◦ 需开启，即设置wal_level参数的值为logical。</li> <li>◦ 如为增量同步任务，DTS要求源数据库的WAL日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的WAL日志至少保留7天以上（您可在全量同步完成后将WAL日志保存时间设置为24小时以上），否则DTS可能因无法获取WAL日志而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的WAL日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> <li>• 源库的操作限制 <ul style="list-style-type: none"> <li>◦ 如自建PostgreSQL进行主备切换，会导致同步失败。</li> <li>◦ 为保障同步任务的正常进行，避免主备切换导致的逻辑订阅中断，需要RDS PostgreSQL支持并开启Logical Replication Slot Failover，设置方式，请参见<a href="#">逻辑复制槽故障转移（Logical Replication Slot Failover）</a>。</li> </ul> </li> </ul>                             |
|      | <ul style="list-style-type: none"> <li>• 暂不支持结构初始化，在配置同步任务前，您需要在目标实例中创建对应的库和表。</li> <li>• 同步对象要求： <ul style="list-style-type: none"> <li>◦ 仅支持表级同步。</li> <li>◦ 不支持BIT、VARBIT、GEOMETRY、UUID、TSQUERY、TSVECTOR、TXID_SNAPSHOT类型的数据同步。</li> </ul> </li> <li>• 单个数据同步任务只能同步一个数据库，如需同步多个数据库，您需要为每个数据库配置数据同步任务。</li> <li>• 同步期间，如果同步对象的选择粒度为Schema，您需要在对该表写入数据前执行 <code>ALTER TABLE schema.table REPLICA IDENTITY FULL;</code> 命令，以保障该表数据同步的一致性。执行该命令期间，建议您不要有锁表操作，否则会导致表锁死。</li> </ul> <div style="background-color: #e6f2ff; padding: 10px; margin-top: 10px;"> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>◦ 将上述命令中的 <code>schema</code> 和 <code>table</code> 替换成真实的Schema名和表名。</li> <li>◦ 建议您在业务低峰期进行操作。</li> </ul> </div> <ul style="list-style-type: none"> <li>• DTS会在源库中创建以下临时表，用于获取增量数据的DDL、增量表的结构、心跳信息等。在同步期</li> </ul> |

| 类型           | <p>说明 请勿删除源库中的临时表，否则会导致DTS任务异常。临时表会在DTS实例释放后自动删除。</p> <p>public.DTS_PG_CLASS 、 public.DTS_PG_ATTRIBUTE 、 public.DTS_PG_TYPE 、 public.DTS_PG_ENUM 、 public.DTS_POSTGRES_HEARTBEAT 、 public.DTS_DDL_COMMAND 、 public.DTS_ARGS_SESSION 。</p>   |           |             |               |           |               |            |        |            |      |  |      |     |  |         |         |         |              |          |         |    |             |       |      |  |           |             |           |        |               |           |        |            |      |  |      |     |  |         |         |         |              |          |         |    |             |       |      |  |
|--------------|---|-----------|-------------|---------------|-----------|---------------|------------|--------|------------|------|--|------|-----|--|---------|---------|---------|--------------|----------|---------|----|-------------|-------|------|--|-----------|-------------|-----------|--------|---------------|-----------|--------|------------|------|--|------|-----|--|---------|---------|---------|--------------|----------|---------|----|-------------|-------|------|--|
| 其他限制         | <ul style="list-style-type: none"> <li>为保障同步延迟时间展示的准确性，DTS会在源库中新增一个名为 <code>dts_postgres_heartbeat</code> 的心跳表。</li> <li>同步期间，DTS会在源库中创建前缀为 <code>dts_sync_</code> 的replication slot用于复制数据。DTS会每隔120分钟自动清理一次历史replication slot，避免其持续累积占用磁盘空间。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin: 10px 0;"> <p><b>说明</b> 当释放同步任务或同步失败时，DTS会主动清理该replication slot；如果RDS PostgreSQL发生了主备切换，则需要您登录备库来手动清理。</p> </div> <div style="border: 1px solid #ccc; padding: 5px; margin: 10px 0;"> <p>Query Editor Query History Scratch Pad</p> <pre>1 SELECT * FROM pg_replication_slots;</pre> <p>Data Output Explain Messages Notifications</p> <table border="1"> <thead> <tr> <th>slot_name</th> <th>plugin_name</th> <th>slot_type</th> <th>datoid</th> <th>database_name</th> <th>temporary</th> <th>active</th> <th>active_pid</th> </tr> <tr> <th>name</th> <th></th> <th>text</th> <th>oid</th> <th></th> <th>boolean</th> <th>boolean</th> <th>integer</th> </tr> </thead> <tbody> <tr> <td>dts_sync_ohu</td> <td>pgoutput</td> <td>logical</td> <td>16</td> <td>dtstestdata</td> <td>false</td> <td>true</td> <td></td> </tr> </tbody> </table> </div> <ul style="list-style-type: none"> <li>同步期间，DTS会在源库中创建前缀为 <code>dts_sync_</code> 的replication slot用于复制数据。DTS会每隔120分钟自动清理一次历史replication slot，避免其持续累积占用磁盘空间。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin: 10px 0;"> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>DTS释放实例后会主动删除该replication slot，如果您在同步期间修改了数据库密码，或者删除了DTS的访问IP白名单，则会导致该replication slot无法自动删除，此时需要您在源库手动删除，避免其持续累积占用磁盘空间导致RDS PostgreSQL实例不可用。</li> <li>当释放同步任务或同步失败时，DTS会主动清理该replication slot；如果RDS PostgreSQL发生了主备切换，则需要您登录备库来手动清理。</li> </ul> </div> <div style="border: 1px solid #ccc; padding: 5px; margin: 10px 0;"> <p>Query Editor Query History Scratch Pad</p> <pre>1 SELECT * FROM pg_replication_slots;</pre> <p>Data Output Explain Messages Notifications</p> <table border="1"> <thead> <tr> <th>slot_name</th> <th>plugin_name</th> <th>slot_type</th> <th>datoid</th> <th>database_name</th> <th>temporary</th> <th>active</th> <th>active_pid</th> </tr> <tr> <th>name</th> <th></th> <th>text</th> <th>oid</th> <th></th> <th>boolean</th> <th>boolean</th> <th>integer</th> </tr> </thead> <tbody> <tr> <td>dts_sync_ohu</td> <td>pgoutput</td> <td>logical</td> <td>16</td> <td>dtstestdata</td> <td>false</td> <td>true</td> <td></td> </tr> </tbody> </table> </div> <ul style="list-style-type: none"> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>对于表级别的数据同步，请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>对于表级别的数据同步，如果除DTS以外的数据没有写入目标库，您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，请参见 <a href="#">不锁表结构变更</a>。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul> | slot_name | plugin_name | slot_type     | datoid    | database_name | temporary  | active | active_pid | name |  | text | oid |  | boolean | boolean | integer | dts_sync_ohu | pgoutput | logical | 16 | dtstestdata | false | true |  | slot_name | plugin_name | slot_type | datoid | database_name | temporary | active | active_pid | name |  | text | oid |  | boolean | boolean | integer | dts_sync_ohu | pgoutput | logical | 16 | dtstestdata | false | true |  |
| slot_name    | plugin_name   | slot_type | datoid      | database_name | temporary | active        | active_pid |        |            |      |  |      |     |  |         |         |         |              |          |         |    |             |       |      |  |           |             |           |        |               |           |        |            |      |  |      |     |  |         |         |         |              |          |         |    |             |       |      |  |
| name         |   | text      | oid         |               | boolean   | boolean       | integer    |        |            |      |  |      |     |  |         |         |         |              |          |         |    |             |       |      |  |           |             |           |        |               |           |        |            |      |  |      |     |  |         |         |         |              |          |         |    |             |       |      |  |
| dts_sync_ohu | pgoutput  | logical   | 16          | dtstestdata   | false     | true          |            |        |            |      |  |      |     |  |         |         |         |              |          |         |    |             |       |      |  |           |             |           |        |               |           |        |            |      |  |      |     |  |         |         |         |              |          |         |    |             |       |      |  |
| slot_name    | plugin_name   | slot_type | datoid      | database_name | temporary | active        | active_pid |        |            |      |  |      |     |  |         |         |         |              |          |         |    |             |       |      |  |           |             |           |        |               |           |        |            |      |  |      |     |  |         |         |         |              |          |         |    |             |       |      |  |
| name         |   | text      | oid         |               | boolean   | boolean       | integer    |        |            |      |  |      |     |  |         |         |         |              |          |         |    |             |       |      |  |           |             |           |        |               |           |        |            |      |  |      |     |  |         |         |         |              |          |         |    |             |       |      |  |
| dts_sync_ohu | pgoutput  | logical   | 16          | dtstestdata   | false     | true          |            |        |            |      |  |      |     |  |         |         |         |              |          |         |    |             |       |      |  |           |             |           |        |               |           |        |            |      |  |      |     |  |         |         |         |              |          |         |    |             |       |      |  |
| 特殊情况         | <ul style="list-style-type: none"> <li>当源实例为RDS PostgreSQL时同步期间，请勿修改RDS PostgreSQL的连接地址和可用区，否则会导致同步失败。</li> </ul>   |           |             |               |           |               |            |        |            |      |  |      |     |  |         |         |         |              |          |         |    |             |       |      |  |           |             |           |        |               |           |        |            |      |  |      |     |  |         |         |         |              |          |         |    |             |       |      |  |

### 自建PostgreSQL同步至PolarDB O引擎

| 类型   | 说明  |
|------|---|
| 源库限制 | <ul style="list-style-type: none"> <li>•</li> <li>•</li> <li>•</li> <li>• 如自建PostgreSQL进行主备切换，会导致同步失败。</li> </ul>   |
| 其他限制 | <ul style="list-style-type: none"> <li>• 暂不支持结构初始化，在配置同步任务前，您需要在目标实例中创建对应的库和表。</li> <li>• 单个数据同步任务只能同步一个数据库，如需同步多个数据库，您需要为每个数据库配置数据同步任务。</li> <li>• 同步期间，如果同步对象的选择粒度为Schema，您需要在对该表写入数据前执行 <code>ALTER TABLE schema.table REPLICA IDENTITY FULL;</code> 命令，以保障该表数据同步的一致性。执行该命令期间，建议您不要有锁表操作，否则会导致表锁死。</li> </ul> <div style="background-color: #e6f2ff; padding: 10px; margin: 10px 0;"> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>◦ 将上述命令中的 <code>schema</code> 和 <code>table</code> 替换成真实的Schema名和表名。</li> <li>◦ 建议您在业务低峰期进行操作。</li> </ul> </div> <ul style="list-style-type: none"> <li>• DTS的校验对象为数据内容，暂不支持Sequence等元数据的校验，您需要自行校验。</li> <li>• 由于业务切换到目标端后，新写入的Sequence不会按照源库的Sequence最大值作为初始值去递增，您需要在业务切换前，在源库中查询对应Sequence的最大值，然后在目标库中将其作为对应Sequence的初始值。查询源库Sequence值的相关命令如下：</li> </ul> <pre style="background-color: #f0f0f0; padding: 10px; margin: 10px 0;">do language plpgsql \$\$ declare   nsp name;   rel name;   val int8; begin   for nsp,rel in select nspname,relname from pg_class t2 , pg_namespace t3   where t2.relnamespace=t3.oid and t2.relkind='S'   loop     execute format(\$select last_value from %I.%I\$, nsp, rel) into val;     raise notice '%',       format(\$select setval('%I.%I':regclass, %s);\$ , nsp, rel, val+1);   end loop; end; \$\$;</pre> <ul style="list-style-type: none"> <li>• DTS会在源库中创建以下临时表，用于获取增量数据的DDL、增量表的结构、心跳信息等。在同步期间，请勿删除源库中的临时表，否则会导致DTS任务异常。临时表会在DTS实例释放后自动删除。<br/><code>public.DTS_PG_CLASS</code>、<code>public.DTS_PG_ATTRIBUTE</code>、<code>public.DTS_PG_TYPE</code>、<code>public.DTS_PG_ENUM</code>、<code>public.DTS_POSTGRES_HEARTBEAT</code>、<code>public.DTS_DDL_COMMAND</code>、<code>public.DTS_ARGS_SESSION</code>。</li> <li>• 为保障同步延迟时间展示的准确性，DTS会在源库中新增一个名为 <code>dts_postgres_heartbeat</code> 的心跳表。</li> <li>• 同步期间，DTS会在源库中创建前缀为 <code>dts_sync_</code> 的replication slot用于复制数据。DTS会每隔120分钟自动清理一次历史replication slot，避免其持续累积占用磁盘空间。</li> </ul> <div style="background-color: #e6f2ff; padding: 10px; margin: 10px 0;"> <p><b>说明</b> 当释放同步任务或同步失败时，DTS会主动清理该replication slot；如果RDS PostgreSQL发生了主备切换，则需要您登录备库来手动清理。</p> </div> |

Query Editor Query History Scratch Pad

```
1 SELECT * FROM pg_replication_slots;
```

| slot_name    | plugin_name | slot_type | datoid | database    | temporary | active | active_pid |
|--------------|-------------|-----------|--------|-------------|-----------|--------|------------|
| dts_sync_ohu | pgoutput    | logical   | 16     | dtsTestData | false     | true   |            |

- 执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。
- 全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。
- 对于表级别的数据同步，请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。
- 对于表级别的数据同步，如果除DTS以外的数据没有写入目标库，您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，请参见[不锁表结构变更](#)。

**RDS PostgreSQL同步至RDS MySQL**

| 类型   | 说明   |
|------|--|
| 源库限制 | <ul style="list-style-type: none"> <li>• 待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>• 若同步对象为表级别，且需进行编辑（如表列名映射），单次同步任务的表数量超过5000时，建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务，否则任务提交后可能会显示请求报错。</li> <li>• WAL日志：             <ul style="list-style-type: none"> <li>◦ 需开启，即设置wal_level参数的值为logical。</li> <li>◦ 如为增量同步任务，DTS要求源数据库的WAL日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的WAL日志至少保留7天以上（您可在全量同步完成后将WAL日志保存时间设置为24小时以上），否则DTS可能因无法获取WAL日志而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的WAL日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> <li>• 源库的操作限制<br/>为保障同步任务的正常进行，避免主备切换导致的逻辑订阅中断，需要RDS PostgreSQL支持并开启Logical Replication Slot Failover，设置方式，请参见<a href="#">逻辑复制槽故障转移（Logical Replication Slot Failover）</a>。</li> </ul> |

| 类型             | 说明  |           |             |             |           |          |            |        |            |      |      |      |     |      |         |         |         |                |          |         |    |             |       |      |  |
|----------------|---|-----------|-------------|-------------|-----------|----------|------------|--------|------------|------|------|------|-----|------|---------|---------|---------|----------------|----------|---------|----|-------------|-------|------|--|
| 其他限制           | <ul style="list-style-type: none"> <li>• 单个数据同步任务只能同步一个数据库，如需同步多个数据库，您需要为每个数据库配置数据同步任务。</li> <li>• 同步期间，如果同步对象的选择粒度为Schema，您需要在对该表写入数据前执行 <code>ALTER TABLE schema.table REPLICA IDENTITY FULL;</code> 命令，以保障该表数据同步的一致性。执行该命令期间，建议您不要有锁表操作，否则会导致表锁死。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin: 10px 0;"> <p><b>? 说明</b></p> <ul style="list-style-type: none"> <li>◦ 将上述命令中的 <code>schema</code> 和 <code>table</code> 替换成真实的Schema名和表名。</li> <li>◦ 建议您在业务低峰期进行操作。</li> </ul> </div> <ul style="list-style-type: none"> <li>• DTS会在源库中创建以下临时表，用于获取增量数据的DDL、增量表的结构、心跳信息等。在同步期间，请勿删除源库中的临时表，否则会导致DTS任务异常。临时表会在DTS实例释放后自动删除。<br/><code>public.DTS_PG_CLASS</code>、<code>public.DTS_PG_ATTRIBUTE</code>、<code>public.DTS_PG_TYPE</code>、<code>public.DTS_PG_ENUM</code>、<code>public.DTS_POSTGRES_HEARTBEAT</code>、<code>public.DTS_DDL_COMMAND</code>、<code>public.DTS_ARGS_SESSION</code>。</li> <li>• 为保障同步延迟时间展示的准确性，DTS会在源库中新增一个名为 <code>dts_postgres_heartbeat</code> 的心跳表。</li> <li>• 同步期间，DTS会在源库中创建前缀为 <code>dts_sync_</code> 的replication slot用于复制数据。DTS会每隔120分钟自动清理一次历史replication slot，避免其持续累积占用磁盘空间。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin: 10px 0;"> <p><b>? 说明</b></p> <ul style="list-style-type: none"> <li>◦ DTS释放实例后会主动删除该replication slot，如果您在同步期间修改了数据库密码，或者删除了DTS的访问IP白名单，则会导致该replication slot无法自动删除，此时需要您在源库手动删除，避免其持续累积占用磁盘空间导致RDS PostgreSQL实例不可用。</li> <li>◦ 当释放同步任务或同步失败时，DTS会主动清理该replication slot；如果RDS PostgreSQL发生了主备切换，则需要您登录备库来手动清理。</li> </ul> </div> <div style="border: 1px solid #ccc; padding: 5px; margin: 10px 0;"> <p>Query Editor Query History Scratch Pad</p> <pre>1 SELECT * FROM pg_replication_slots;</pre> <p>Data Output Explain Messages Notifications</p> <table border="1"> <thead> <tr> <th>slot_name</th> <th>plugin_name</th> <th>slot_type</th> <th>datoid</th> <th>database</th> <th>temporary</th> <th>active</th> <th>active_pid</th> </tr> <tr> <th>name</th> <th>name</th> <th>text</th> <th>oid</th> <th>name</th> <th>boolean</th> <th>boolean</th> <th>integer</th> </tr> </thead> <tbody> <tr> <td>1 dts_sync_ohu</td> <td>pgoutput</td> <td>logical</td> <td>16</td> <td>dtstestdata</td> <td>false</td> <td>true</td> <td></td> </tr> </tbody> </table> </div> <ul style="list-style-type: none"> <li>• 执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>• 全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>• 对于表级别的数据同步，请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>• 对于表级别的数据同步，如果除DTS以外的数据没有写入目标库，您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，请参见<a href="#">不锁表结构变更</a>。</li> <li>• 在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul> | slot_name | plugin_name | slot_type   | datoid    | database | temporary  | active | active_pid | name | name | text | oid | name | boolean | boolean | integer | 1 dts_sync_ohu | pgoutput | logical | 16 | dtstestdata | false | true |  |
| slot_name      | plugin_name   | slot_type | datoid      | database    | temporary | active   | active_pid |        |            |      |      |      |     |      |         |         |         |                |          |         |    |             |       |      |  |
| name           | name  | text      | oid         | name        | boolean   | boolean  | integer    |        |            |      |      |      |     |      |         |         |         |                |          |         |    |             |       |      |  |
| 1 dts_sync_ohu | pgoutput  | logical   | 16          | dtstestdata | false     | true     |            |        |            |      |      |      |     |      |         |         |         |                |          |         |    |             |       |      |  |
| 特殊情况           | <ul style="list-style-type: none"> <li>• 当源实例为RDS PostgreSQL时同步期间，请勿修改RDS PostgreSQL的连接地址和可用区，否则会导致同步失败。</li> </ul>   |           |             |             |           |          |            |        |            |      |      |      |     |      |         |         |         |                |          |         |    |             |       |      |  |

## RDS PostgreSQL同步至PolarDB PostgreSQL引擎

| 类型   | 说明  |
|------|---|
| 源库限制 | <ul style="list-style-type: none"><li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li><li>若同步对象为表级别，且需进行编辑（如表列名映射），单次同步任务的表数量超过5000时，建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务，否则任务提交后可能会显示请求报错。</li><li>WAL日志：<ul style="list-style-type: none"><li>需开启，即设置wal_level参数的值为<i>logical</i>。</li><li>如为增量同步任务，DTS要求源数据库的WAL日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的WAL日志至少保留7天以上（您可在全量同步完成后将WAL日志保存时间设置为24小时以上），否则DTS可能因无法获取WAL日志而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的WAL日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li></ul></li><li>同步期间，请勿修改RDS PostgreSQL的连接地址和可用区，否则会导致同步失败。</li></ul> |

| 类型             | 说明   |           |             |               |           |               |            |        |            |      |      |      |     |      |         |         |         |                |          |         |    |             |       |      |  |
|----------------|--|-----------|-------------|---------------|-----------|---------------|------------|--------|------------|------|------|------|-----|------|---------|---------|---------|----------------|----------|---------|----|-------------|-------|------|--|
| 其他限制           | <ul style="list-style-type: none"> <li>• 单个数据同步任务只能同步一个数据库，如需同步多个数据库，您需要为每个数据库配置数据同步任务。</li> <li>• 同步期间，如果同步对象的选择粒度为Schema，您需要在对该表写入数据前执行 <code>ALTER TABLE schema.table REPLICA IDENTITY FULL;</code> 命令，以保障该表数据同步的一致性。执行该命令期间，建议您不要有锁表操作，否则会导致表锁死。</li> </ul> <div data-bbox="475 465 1386 624" style="border: 1px solid #ccc; padding: 5px; margin: 10px 0;"> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>◦ 将上述命令中的 <code>schema</code> 和 <code>table</code> 替换成真实的Schema名和表名。</li> <li>◦ 建议您在业务低峰期进行操作。</li> </ul> </div> <ul style="list-style-type: none"> <li>• DTS会在源库中创建以下临时表，用于获取增量数据的DDL、增量表的结构、心跳信息等。在同步期间，请勿删除源库中的临时表，否则会导致DTS任务异常。临时表会在DTS实例释放后自动删除。<br/><code>public.DTS_PG_CLASS</code>、<code>public.DTS_PG_ATTRIBUTE</code>、<code>public.DTS_PG_TYPE</code>、<code>public.DTS_PG_ENUM</code>、<code>public.DTS_POSTGRES_HEARTBEAT</code>、<code>public.DTS_DDL_COMMAND</code>、<code>public.DTS_ARGS_SESSION</code>。</li> <li>• 为保障同步延迟时间展示的准确性，DTS会在源库中新增一个名为 <code>dts_postgres_heartbeat</code> 的心跳表。</li> <li>• 同步期间，DTS会在源库中创建前缀为 <code>dts_sync_</code> 的replication slot用于复制数据。DTS会每隔120分钟自动清理一次历史replication slot，避免其持续累积占用磁盘空间。</li> </ul> <div data-bbox="475 943 1386 1182" style="border: 1px solid #ccc; padding: 5px; margin: 10px 0;"> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>◦ DTS释放实例后会主动删除该replication slot，如果您在同步期间修改了数据库密码，或者删除了DTS的访问IP白名单，则会导致该replication slot无法自动删除，此时需要您在源库手动删除，避免其持续累积占用磁盘空间导致RDS PostgreSQL实例不可用。</li> <li>◦ 当释放同步任务或同步失败时，DTS会主动清理该replication slot；如果RDS PostgreSQL发生了主备切换，则需要您登录备库来手动清理。</li> </ul> </div> <div data-bbox="475 1200 1386 1402" style="border: 1px solid #ccc; padding: 5px; margin: 10px 0;"> <p>Query Editor Query History Scratch Pad</p> <pre>1 SELECT * FROM pg_replication_slots;</pre> <p>Data Output Explain Messages Notifications</p> <table border="1"> <thead> <tr> <th>slot_name</th> <th>plugin_name</th> <th>slot_type</th> <th>datoid</th> <th>database_name</th> <th>temporary</th> <th>active</th> <th>active_pid</th> </tr> <tr> <th>name</th> <th>name</th> <th>text</th> <th>oid</th> <th>name</th> <th>boolean</th> <th>boolean</th> <th>integer</th> </tr> </thead> <tbody> <tr> <td>1 dts_sync_ohu</td> <td>pgoutput</td> <td>logical</td> <td>16</td> <td>dtstestdata</td> <td>false</td> <td>true</td> <td></td> </tr> </tbody> </table> </div> <ul style="list-style-type: none"> <li>• 执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>• 全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>• 对于表级别的数据同步，请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>• 对于表级别的数据同步，如果除DTS以外的数据没有写入目标库，您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，请参见<a href="#">不锁表结构变更</a>。</li> <li>• 在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul> | slot_name | plugin_name | slot_type     | datoid    | database_name | temporary  | active | active_pid | name | name | text | oid | name | boolean | boolean | integer | 1 dts_sync_ohu | pgoutput | logical | 16 | dtstestdata | false | true |  |
| slot_name      | plugin_name  | slot_type | datoid      | database_name | temporary | active        | active_pid |        |            |      |      |      |     |      |         |         |         |                |          |         |    |             |       |      |  |
| name           | name   | text      | oid         | name          | boolean   | boolean       | integer    |        |            |      |      |      |     |      |         |         |         |                |          |         |    |             |       |      |  |
| 1 dts_sync_ohu | pgoutput   | logical   | 16          | dtstestdata   | false     | true          |            |        |            |      |      |      |     |      |         |         |         |                |          |         |    |             |       |      |  |

### 3.9. 同步时源库为SQL Server的注意事项及限制

如果同步的源数据库类型为SQL Server，如自建SQL Server、RDS SQL Server，您需要在配置具体的同步任务前，参考本文的注意事项及限制，以保障数据同步任务的正常运行。

### 源库为SQL Server的同步方案概览

根据同步方案，查看同步任务的注意事项及限制：

**说明**

DTS默认同步到目标数据库中时会取消外键约束，因此源数据库的级联、删除等操作不会同步到如下目标数据库：

- SQL Server
- MySQL (RDS MySQL、自建MySQL)
- PolarDB MySQL
- 云原生数据仓库AnalyticDB MySQL
- 云原生数据仓库AnalyticDB PostgreSQL

- [SQL Server间的同步](#)
- [RDS SQL Server同步至MySQL](#)
- [SQL Server同步至PolarDB MySQL](#)
- [SQL Server同步至云原生数据仓库AnalyticDB MySQL版](#)
- [SQL Server同步至云原生数据仓库AnalyticDB PostgreSQL版](#)

### SQL Server间的同步

具体注意事项及限制如下：

| 类型   | 说明   |
|------|--|
| 源库限制 | <ul style="list-style-type: none"> <li>• 待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>• 如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> <li>• 单次同步任务仅支持同步至多10个数据库。当超出数量限制，会出现稳定性和性能问题的风险，此时建议您拆分待同步的表，分批配置任务。</li> <li>• 数据日志：                         <ul style="list-style-type: none"> <li>◦ 需开启，备份模式设置为Full，且已成功执行过全量逻辑备份。</li> <li>◦ 如为增量同步任务，DTS要求源数据库的数据日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的数据日志至少保留7天以上（您可在全量同步完成后将数据日志保存时间设置为24小时以上），否则DTS可能因无法获取数据日志而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的数据日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> </ul> |

| 类型   | 说明   |
|------|--|
| 其他限制 | <ul style="list-style-type: none"> <li>不支持同步TEXT、CURSOR、ROWVERSION、SQL_VARIANT、HIERACHYID和GEOMETRY。</li> <li>如需跨版本同步，请提前确认兼容性。</li> <li>如在配置任务对象及高级配置时选择SQLServer增量同步模式为解析源库日志做增量同步，则待同步的表需具备聚集索引，且聚集索引中需包含主键列；待同步表不能为压缩表；待同步表不能包含计算列。在混合式日志解析模式下，没有如上限制。</li> <li>为保障同步延迟时间展示的准确性，解析源端日志做增量同步模式，DTS会在源库中新增一个名为 <i>dts_log_heart_beat</i> 的心跳表；混合式增量同步模式，DTS会在源库中创建触发器 <i>dts_cdc_sync_ddl</i>、心跳表 <i>dts_sync_progress</i>、DDL存储表 <i>dts_cdc_ddl_history</i> 以及会开启库级别CDC和部分表CDC。</li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>对于表级别的数据同步，请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul> |

### RDS SQL Server同步至MySQL

当目标为MySQL，包括自建MySQL、RDS MySQL时，具体注意事项及限制如下：

| 类型   | 说明   |
|------|--|
| 源库限制 | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> <li>单次同步任务仅支持同步至多10个数据库。当超出数量限制，会出现稳定性和性能问题的风险，此时建议您拆分待同步的表，分批配置任务。</li> <li>数据日志：             <ul style="list-style-type: none"> <li>需开启，备份模式设置为Full，且已成功执行过全量逻辑备份。</li> <li>如为增量同步任务，DTS要求源数据库的数据日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的数据日志至少保留7天以上（您可在全量同步完成后将数据日志保存时间设置为24小时以上），否则DTS可能因无法获取数据日志而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的数据日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> </ul> |

| 类型   | 说明  |
|------|---|
| 其他限制 | <ul style="list-style-type: none"> <li>同步对象的要求：                             <ul style="list-style-type: none"> <li>支持结构初始化的对象有database、schema、table。</li> <li>不支持同步数据类型TEXT、CURSOR、ROWVERSION、HIERACHYID、SQL_VARIANT、SPATIAL GEOMETRY、SPATIAL GEOGRAPHY。</li> </ul> </li> <li>如在配置任务对象及高级配置时选择SQLServer增量同步模式为解析源库日志做增量同步，则待同步的表需具备聚集索引，且聚集索引中需包含主键列；待同步表不能为压缩表；待同步表不能包含计算列。在混合式日志解析模式下，没有如上限制。</li> <li>暂不支持复杂的同步DDL操作。</li> <li>为保障同步延迟时间展示的准确性，解析源端日志做增量同步模式，DTS会在源库中新增一个名为 <i>dts_log_heart_beat</i> 的心跳表；混合式增量同步模式，DTS会在源库中创建触发器 <i>dts_cdc_sync_ddl</i>、心跳表 <i>dts_sync_progress</i>、DDL存储表 <i>dts_cdc_ddl_history</i> 以及会开启库级别CDC和部分表CDC。</li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>对于表级别的数据同步，请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul> |

### SQL Server同步至PolarDB MySQL

具体注意事项及限制如下：

| 类型   | 说明   |
|------|--|
| 源库限制 | <ul style="list-style-type: none"> <li></li> <li></li> <li></li> <li></li> </ul>   |
| 其他限制 | <ul style="list-style-type: none"> <li>同步对象的要求：                             <ul style="list-style-type: none"> <li>支持结构初始化的对象有database、schema、table。</li> <li>不支持同步数据类型TEXT、CURSOR、ROWVERSION、HIERACHYID、SQL_VARIANT、SPATIAL GEOMETRY、SPATIAL GEOGRAPHY。</li> </ul> </li> <li></li> <li>暂不支持复杂的同步DDL操作。</li> <li></li> <li></li> <li></li> <li></li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul> |

### SQL Server同步至云原生数据仓库AnalyticDB MySQL版

| 类型 | 说明 |
|----|----|
|    |    |

| 类型   | 说明  |
|------|---|
| 源库限制 | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> <li>单次同步任务仅支持同步至多10个数据库。当超出数量限制，会出现稳定性和性能问题的风险，此时建议您拆分待同步的表，分批配置任务。</li> <li>数据日志：                         <ul style="list-style-type: none"> <li>需开启，备份模式设置为Full，且已成功执行过全量逻辑备份。</li> <li>如为增量同步任务，DTS要求源数据库的数据日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的数据日志至少保留7天以上（您可在全量同步完成后将数据日志保存时间设置为24小时以上），否则DTS可能因无法获取数据日志而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的数据日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> </ul>  |
| 其他限制 | <ul style="list-style-type: none"> <li>同步对象的要求：                         <ul style="list-style-type: none"> <li>支持结构初始化的对象有Schema、Table、View、Function和Procedure。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin: 5px 0;"> <p> <b>警告</b> 由于此场景属于异构数据库间的数据同步，数据类型无法一一对应，请谨慎评估数据类型映射关系对业务的影响，详情请参见<a href="#">结构初始化涉及的数据类型映射关系</a>。</p> </div> <ul style="list-style-type: none"> <li>不支持结构初始化的结构对象有：assemblies、service broker、全文索引、全文目录、分布式schema、分布式函数、CLR存储过程、CLR标量函数、CLR表值函数、内部表、系统、聚合函数。</li> <li>不支持同步数据类型为TIMESTAMP、CURSOR、ROWVERSION、HIERACHYID、SQL_VARIANT、SPATIAL GEOMETRY、SPATIAL GEOGRAPHY、TABLE的数据。</li> <li>不支持同步含有计算列的表。</li> </ul> </li> <li>如在配置任务对象及高级配置时选择SQLServer增量同步模式为解析源库日志做增量同步，则待同步的表需具备聚集索引，且聚集索引中需包含主键列；待同步表不能为压缩表；待同步表不能包含计算列。在混合式日志解析模式下，没有如上限制。</li> <li>为保障同步延迟时间展示的准确性，解析源端日志做增量同步模式，DTS会在源库中新增一个名为 <i>dts_log_heart_beat</i> 的心跳表；混合式增量同步模式，DTS会在源库中创建触发器 <i>dts_cdc_sync_ddl</i>、心跳表 <i>dts_sync_progress</i>、DDL存储表 <i>dts_cdc_ddl_history</i> 以及会开启库级别CDC和部分表CDC。</li> <li>由于云原生数据库AnalyticDB MySQL版本本身的使用限制，当云原生数据库AnalyticDB MySQL版中的节点磁盘空间使用量超过80%，数据写入目标库的性能变慢，导致DTS任务延迟；当使用量超过90%，数据无法写入目标库，导致DTS任务异常。请提前根据待同步的对象预估所需空间，确保目标集群具备充足的存储空间。</li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>对于表级别的数据同步，请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul> |

### SQL Server同步至云原生数据库AnalyticDB PostgreSQL版

具体注意事项及限制如下：

| 类型 | 说明 |
|----|----|
|----|----|

| 类型   | 说明   |
|------|--|
| 源库限制 | <ul style="list-style-type: none"> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> <li>单次同步任务仅支持同步至多10个数据库。当超出数量限制，会出现稳定性和性能问题的风险，此时建议您拆分待同步的表，分批配置任务。</li> <li>数据日志：                         <ul style="list-style-type: none"> <li>需开启，备份模式设置为Full，且已成功执行过全量逻辑备份。</li> <li>如为增量同步任务，DTS要求源数据库的数据日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的数据日志至少保留7天以上（您可在全量同步完成后将数据日志保存时间设置为24小时以上），否则DTS可能因无法获取数据日志而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的数据日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> </ul>   |
| 其他限制 | <ul style="list-style-type: none"> <li>同步对象的要求：                         <ul style="list-style-type: none"> <li>支持结构初始化的对象有Schema、Table、View、Function和Procedure。</li> </ul> </li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin: 10px 0;"> <p> <b>警告</b> 由于此场景属于异构数据库间的数据同步，数据类型无法一一对应，请谨慎评估数据类型映射关系对业务的影响，详情请参见<a href="#">结构初始化涉及的数据类型映射关系</a>。</p> </div> <ul style="list-style-type: none"> <li>不支持结构初始化的结构对象有：assemblies、service broker、全文索引、全文目录、分布式schema、分布式函数、CLR存储过程、CLR标量函数、CLR表值函数、内部表、系统、聚合函数。</li> <li>不支持同步数据类型为TIMESTAMP、CURSOR、ROWVERSION、HIERACHYID、SQL_VARIANT、SPATIAL GEOMETRY、SPATIAL GEOGRAPHY、TABLE的数据。</li> <li>不支持同步含有计算列的表。</li> <li>如在配置任务对象及高级配置时选择SQLServer增量同步模式为解析源库日志做增量同步，则待同步的表需具备聚集索引，且聚集索引中需包含主键列；待同步表不能为压缩表；待同步表不能包含计算列。在混合式日志解析模式下，没有如上限制。</li> <li>为保障同步延迟时间展示的准确性，解析源端日志做增量同步模式，DTS会在源库中新增一个名为 <i>dts_log_heart_beat</i> 的心跳表；混合式增量同步模式，DTS会在源库中创建触发器 <i>dts_cdc_sync_ddl</i>、心跳表 <i>dts_sync_progress</i>、DDL存储表 <i>dts_cdc_ddl_history</i> 以及会开启库级别CDC和部分表CDC。</li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>对于表级别的数据同步，请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul> |

### 3.10. 同步时源库为MongoDB的注意事项及限制

如果同步的源数据库类型为MongoDB，如自建MongoDB、云数据库MongoDB，您需要在配置具体的同步任务前，参考本文的注意事项及限制，以保障数据同步任务的正常运行。

#### 源库为MongoDB的同步方案概览

根据如下同步方案，查看同步任务的注意事项及限制：

- [MongoDB（副本集架构）同步至MongoDB（副本集架构、分片集群架构）](#)
- [MongoDB（分片集群架构）间双向同步](#)

#### MongoDB（副本集架构）同步至MongoDB（副本集架构、分片集群架构）

如果同步的目标数据库类型为MongoDB，如自建MongoDB、云数据库MongoDB，具体注意事项及限制如下：

| 类型   | 说明   |
|------|--|
| 源库限制 | <ul style="list-style-type: none"> <li>带宽要求：源库所属的服务器需具备足够出口带宽，否则将影响数据同步速率。</li> <li>待同步的集合需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>如同步对象为集合级别，且需进行编辑（如集合的名称映射），则单次同步任务仅支持同步至多1000张集合。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的集合，分批配置多个任务，或者配置整库的同步任务。</li> <li>Olog日志：                         <ul style="list-style-type: none"> <li>需开启。</li> <li>如为增量同步任务，DTS要求源数据库的Olog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的Olog日志至少保留7天以上（您可在全量同步完成后将Olog日志保存时间设置为24小时以上），否则DTS可能因无法获取Olog日志而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Olog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> </ul> |
| 其他限制 | <ul style="list-style-type: none"> <li>建议源和目标库的MongoDB的数据库版本保持一致，或者从低版本同步到高版本以保障兼容性。如为高版本同步至低版本，可能存在数据库兼容性问题。</li> <li>不支持同步admin和local库中的数据。</li> <li>不保留事务信息，即源库中的事务同步到目标库时会转变为单条的记录。</li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的集合产生碎片，因此全量初始化完成后目标实例的集合空间比源实例的集合空间大。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul>  |
| 特殊情况 | <p>当源库为自建MongoDB时：</p> <ul style="list-style-type: none"> <li>在同步时，如果源库进行主备切换，将会导致同步任务失败。</li> <li>由于DTS的延迟时间是根据同步到目标库最后一条数据的时间戳和当前时间戳对比得出，源库长时间未执行更新操作可能导致延迟信息不准确。如果任务显示的延迟时间过大，您可以在源库执行一个更新操作来更新延迟信息。</li> </ul> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin-top: 10px;"> <p> 说明 如果同步对象选择为整库，您还可以创建心跳，心跳每秒定期更新或者写入数据。</p> </div>  |

### MongoDB（分片集群架构）间双向同步

如果同步的目标数据库类型为MongoDB，如自建MongoDB、云数据库MongoDB，具体注意事项及限制如下：

| 类型 | 说明 |
|----|----|
|    |    |

| 类型       | 说明  |
|----------|---|
| 源和目标库的限制 | <ul style="list-style-type: none"> <li>• 带宽要求：源库所属的服务器需具备足够出口带宽，否则将影响数据同步速率。</li> <li>• 待同步的集合需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出現重复数据。</li> <li>• 如同步对象为集合级别，且需进行编辑（如集合的名称映射），则单次同步任务仅支持同步至多1000张集合。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的集合，分批配置多个任务，或者配置整库的同步任务。</li> <li>• Olog日志：                         <ul style="list-style-type: none"> <li>◦ 需开启。</li> <li>◦ 如为增量同步任务，DTS要求源数据库的Olog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的Olog日志至少保留7天以上（您可在全量同步完成后将Olog日志保存时间设置为24小时以上），否则DTS可能因无法获取Olog日志而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Olog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> <li>• MongoDB分片集群为源的单向同步不支持源端做分片的扩缩容，双向同步源和目标都不支持分片数量的扩缩容，否则会导致DTS任务失败。</li> </ul> |
| 其他限制     | <ul style="list-style-type: none"> <li>• 建议源和目标库的MongoDB的数据库版本保持一致，或者从低版本同步到高版本以保障兼容性。如为高版本同步至低版本，可能存在数据库兼容性问题。</li> <li>• 如双向同步任务的源实例或目标实例位于海外地域，则仅支持同地域的双向同步，不支持跨地域的双向同步。例如，支持日本地域间的双向同步，不支持日本地域与法兰克福地域间的双向同步。</li> <li>• 不支持同步admin和local库中的数据。</li> <li>• 不保留事务信息，即源库中的事务同步到目标库时会转变为单条的记录。</li> <li>• 执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>• 全量初始化会并发执行INSERT操作，导致目标数据库的集合产生碎片，因此全量初始化完成后目标实例的集合空间比源实例的集合空间大。</li> <li>• 在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> <li>• 若DTS正向和反向同步任务均同步了某个表，且正向同步任务同步了该表的全量和增量数据，则反向同步任务只同步该表的增量数据。</li> </ul>                    |

### MongoDB（分片集群架构）间单向同步

如果同步的目标数据库类型为MongoDB，如自建MongoDB、云数据库MongoDB，具体注意事项及限制如下：

| 类型 | 说明 |
|----|----|
|----|----|

| 类型       | 说明  |
|----------|---|
| 源和目标库的限制 | <ul style="list-style-type: none"> <li>带宽要求：源库所属的服务器需具备足够出口带宽，否则将影响数据同步速率。</li> <li>待同步的集合需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出現重复数据。</li> <li>如同步对象为集合级别，且需进行编辑（如集合的名称映射），则单次同步任务仅支持同步至多1000张集合。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的集合，分批配置多个任务，或者配置整库的同步任务。</li> <li>Olog日志： <ul style="list-style-type: none"> <li>需开启。</li> <li>如为增量同步任务，DTS要求源数据库的Olog日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的Olog日志至少保留7天以上（您可在全量同步完成后将Olog日志保存时间设置为24小时以上），否则DTS可能因无法获取Olog日志而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的Olog日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> <li>在DTS同步期间，不支持MongoDB分片集群进行分片的扩缩容，否则会导致DTS任务失败。</li> </ul> |
| 其他限制     | <ul style="list-style-type: none"> <li>建议源和目标库的MongoDB的数据库版本保持一致，或者从低版本同步到高版本以保障兼容性。如为高版本同步至低版本，可能存在数据库兼容性问题。</li> <li>不支持同步admin和local库中的数据。</li> <li>不保留事务信息，即源库中的事务同步到目标库时会转变为单条的记录。</li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的集合产生碎片，因此全量初始化完成后目标实例的集合空间比源实例的集合空间大。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul>   |

## 3.11. 同步时源库为DB2 for LUW的注意事项及限制

如果同步的源库为Db2 for LUW，您需要在配置具体的同步任务前，参考本文的注意事项及限制，以保障数据同步任务的正常运行。

### DB2 for LUW同步至PolarDB-X 2.0

 **说明** DTS默认同步到目标数据库中时会取消外键约束，因此源数据库的级联、删除等操作不会同步到目标数据库。

| 类型 | 说明 |
|----|----|
|----|----|

| 类型   | 说明  |
|------|---|
| 源库限制 | <ul style="list-style-type: none"> <li>带宽要求：源库所属的服务器需具备足够出口带宽，否则将影响数据同步速率。</li> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> <li>数据日志： <ul style="list-style-type: none"> <li>需开启，否则预检查阶段提示报错，且无法成功启动数据同步任务。</li> <li>如为增量同步任务，DTS要求源数据库的数据日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的数据日志至少保留7天以上（您可在全量同步完成后将数据日志保存时间设置为24小时以上），否则DTS可能因无法获取数据日志而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的数据日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> </ul>  |
| 其他限制 | <ul style="list-style-type: none"> <li>暂不支持库表结构同步，在配置同步任务前，您需要在目标实例中创建对应的库和表。</li> <li>由于DTS基于Db2 for LUW的CDC复制技术，将Db2 for LUW数据库的增量数据同步到目标库中，但是CDC复制技术自身具有限制，请参见<a href="#">CDC复制技术所支持数据类型的限制</a>。</li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>如果除DTS以外的数据没有写入目标库，您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，请参见<a href="#">不锁表结构变更</a>。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul> |
| 特殊情况 | <p>由于源库Db2 for LUW为自建数据库，您还请注意如下事项：</p> <ul style="list-style-type: none"> <li>在同步时，如果源库进行主备切换，将会导致同步任务失败。</li> <li>由于DTS的延迟时间是根据同步到目标库最后一条数据的时间戳和当前时间戳对比得出，源库长时间未执行DML操作可能导致延迟信息不准确。如果任务显示的延迟时间过大，您可以在源库执行一个DML操作来更新延迟信息。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p> <b>说明</b> 如果同步对象选择为整库，您还可以创建心跳表，心跳表每秒定期更新或者写入数据。</p> </div>  |

### DB2 for LUW同步至AnalyticDB MySQL版 3.0

 **说明** DTS默认同步到目标数据库中时会取消外键约束，因此源数据库的级联、删除等操作不会同步到目标数据库。

| 类型 | 说明 |
|----|----|
|----|----|

| 类型   | 说明   |
|------|--|
| 源库限制 | <ul style="list-style-type: none"> <li>带宽要求：源库所属的服务器需具备足够出口带宽，否则将影响数据同步速率。</li> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> <li>数据日志：                     <ul style="list-style-type: none"> <li>需开启，否则预检查阶段提示报错，且无法成功启动数据同步任务。</li> <li>如为增量同步任务，DTS要求源数据库的数据日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的数据日志至少保留7天以上（您可在全量同步完成后将数据日志保存时间设置为24小时以上），否则DTS可能因无法获取数据日志而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的数据日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> </ul>   |
| 其他限制 | <ul style="list-style-type: none"> <li>暂不支持库表结构同步，在配置同步任务前，您需要在目标实例中创建对应的库和表。</li> <li>由于DTS基于Db2 for LUW的CDC复制技术，将Db2 for LUW数据库的增量数据同步到目标库中，但是CDC复制技术自身具有限制，请参见<a href="#">CDC复制技术所支持数据类型的限制</a>。</li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>如果除DTS以外的数据没有写入目标库，您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，请参见<a href="#">不锁表结构变更</a>。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> <li>由于云原生数据仓库AnalyticDB MySQL版 3.0本身的使用限制，当云原生数据仓库AnalyticDB MySQL版 3.0集群中的节点磁盘空间使用量超过80%，会导致DTS任务异常，产生延迟。请提前根据待同步的对象预估所需空间，确保目标集群具备充足的存储空间。</li> <li>云原生数据仓库AnalyticDB MySQL版 3.0数据大小写为严格敏感。</li> </ul> |
| 特殊情况 | <p>由于源库Db2 for LUW为自建数据库，您还请注意如下事项：</p> <ul style="list-style-type: none"> <li>在同步时，如果源库进行主备切换，将会导致同步任务失败。</li> <li>由于DTS的延迟时间是根据同步到目标库最后一条数据的时间戳和当前时间戳对比得出，源库长时间未执行DML操作可能导致延迟信息不准确。如果任务显示的延迟时间过大，您可以在源库执行一个DML操作来更新延迟信息。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p> <b>说明</b> 如果同步对象选择为整库，您还可以创建心跳表，心跳表每秒定期更新或者写入数据。</p> </div>   |

### DB2 for LUW同步至PolarDB MySQL集群

 **说明** DTS默认同步到目标数据库中时会取消外键约束，因此源数据库的级联、删除等操作不会同步到目标数据库。

| 类型 | 说明 |
|----|----|
|----|----|

| 类型   | 说明  |
|------|---|
| 源库限制 | <ul style="list-style-type: none"> <li>带宽要求：源库所属的服务器需具备足够出口带宽，否则将影响数据同步速率。</li> <li>待同步的表需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出现重复数据。</li> <li>如同步对象为表级别，且需进行编辑（如表列名映射），则单次同步任务仅支持同步至多1000张表。当超出数量限制，任务提交后会显示请求报错，此时建议您拆分待同步的表，分批配置多个任务，或者配置整库的同步任务。</li> <li>数据日志： <ul style="list-style-type: none"> <li>需开启，否则预检查阶段提示报错，且无法成功启动数据同步任务。</li> <li>如为增量同步任务，DTS要求源数据库的数据日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的数据日志至少保留7天以上（您可在全量同步完成后将数据日志保存时间设置为24小时以上），否则DTS可能因无法获取数据日志而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的数据日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> </ul>  |
| 其他限制 | <ul style="list-style-type: none"> <li>暂不支持库表结构同步，在配置同步任务前，您需要在目标实例中创建对应的库和表。</li> <li>由于DTS基于Db2 for LUW的CDC复制技术，将Db2 for LUW数据库的增量数据同步到目标库中，但是CDC复制技术自身具有限制，请参见<a href="#">CDC复制技术所支持数据类型的限制</a>。</li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>全量初始化会并发执行INSERT操作，导致目标数据库的表产生碎片，因此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。</li> <li>请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。</li> <li>如果除DTS以外的数据没有写入目标库，您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，请参见<a href="#">不锁表结构变更</a>。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul> |
| 特殊情况 | <p>由于源库Db2 for LUW为自建数据库，您还需要注意如下事项：</p> <ul style="list-style-type: none"> <li>在同步时，如果源库进行主备切换，将会导致同步任务失败。</li> <li>由于DTS的延迟时间是根据同步到目标库最后一条数据的时间戳和当前时间戳对比得出，源库长时间未执行DML操作可能导致延迟信息不准确。如果任务显示的延迟时间过大，您可以在源库执行一个DML操作来更新延迟信息。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p> <b>说明</b> 如果同步对象选择为整库，您还可以创建心跳表，心跳表每秒定期更新或者写入数据。</p> </div>   |

### 3.12. 同步时源库为Redis的注意事项及限制

如果同步的源数据库类型为Redis，如自建Redis、云数据库Redis，您需要在配置具体的同步任务前，参考本文的注意事项及限制，以保障数据同步任务的正常运行。

#### 源库为Redis的同步方案概览

根据如下同步方案，查看同步任务的注意事项及限制：

- Redis间的双向同步
- Redis间的单向同步

#### Redis间的双向同步

| 类型   | 说明  |
|------|---|
| 源库限制 | <ul style="list-style-type: none"> <li>待同步的集合需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出現重复数据。</li> <li>为保障同步质量，DTS会在源库中插入一个key: DTS_REDIS_TIMESTAMP_HEARTBEAT，用于记录更新时间点。</li> <li>为保障同步链路稳定性，建议将配置文件redis.conf中 repl-backlog-size 参数的值适当调大。</li> <li>请勿在源库中执行 FLUSHDB 和 FLUSHALL 命令，否则将导致源和目标的数据不一致。</li> <li>AOF日志：             <ul style="list-style-type: none"> <li>需开启。</li> <li>如为增量同步任务，DTS要求源数据库的AOF日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的AOF日志至少保留7天以上（您可在全量同步完成后将AOF日志保存时间设置为24小时以上），否则DTS可能因无法获取AOF日志而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的AOF日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin: 10px 0;"> <p><span style="color: #00aaff;">?</span> 说明 云数据库Redis不存在该限制。</p> </div> <ul style="list-style-type: none"> <li>如果源库中的某些Key使用了过期（expire）策略，由于可能存在Key已过期但未被及时删除的情况，所以在目标库中查看到的Key数量（例如通过info命令查看）会比源库的Key数量少。如果源库中的某些Key使用了过期（expire）策略，由于可能存在Key已过期但未被及时删除的情况，所以在目标库中查看到的Key数量（例如通过info命令查看）会比源库的Key数量少。</li> <li>如果目标库的数据逐出策略（ maxmemory-policy ）配置为 noeviction 以外的值，可能导致目标库的数据与源库不一致。关于数据逐出策略详情，请参见<a href="#">Redis数据逐出策略介绍</a>。</li> <li>源库单机版Redis同步到目标库集群版Redis的操作限制：由于集群cluster只允许单个命令操作单个slot，若在源库执行多Key操作时，Key不在同一个slot或涉及多个slot，则会出现以下报错：             <pre style="background-color: #f0f0f0; padding: 5px; margin: 5px 0;">CROSSSLOT Keys in request don't hash to the same slot</pre> <p>建议在DTS同步过程中仅执行单Key操作，以免导致链路中断。</p> </li> </ul> |
| 其他限制 | <ul style="list-style-type: none"> <li>同步期间，如源Redis发生扩缩容（如增加或者减少分片）、规格变配（如扩大内存），则您需重新配置任务。且为保障数据一致性，建议重新配置任务前，先清空已同步至目标Redis的数据。</li> <li>同步期间，如源Redis连接地址变化，则需要 <a href="#">提交工单</a> 进行处理，如果长时间未处理，源Redis会重置AOF文件，届时您需重新配置任务。</li> <li>建议源和目标库版本保持一致，或者从低版本同步到高版本以保障兼容性。如为高版本同步至低版本，可能存在数据库兼容性问题。</li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>如双向同步任务的源实例或目标实例位于海外地域，则仅支持同地域的双向同步，不支持跨地域的双向同步。例如，支持日本地域间的双向同步，不支持日本地域与法兰克福地域间的双向同步。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> <li>若DTS正向和反向同步任务均同步了某个表，且正向同步任务同步了该表的全量和增量数据，则反向同步任务只同步该表的增量数据。</li> </ul>   |

### Redis间的单向同步

| 类型 | 说明 |
|----|----|
|    |    |

| 类型   | 说明   |
|------|--|
| 源库限制 | <ul style="list-style-type: none"> <li>待同步的集合需具备主键或唯一约束，且字段具有唯一性，否则可能会导致目标数据库中出現重复数据。</li> <li>为保障同步质量，DTS会在源库中插入一个key: DTS_REDIS_TIMESTAMP_HEARTBEAT，用于记录更新时间点。</li> <li>为保障同步链路稳定性，建议将配置文件redis.conf中 repl-backlog-size 参数的值适当调大。</li> <li>AOF日志：             <ul style="list-style-type: none"> <li>需开启。</li> <li>如为增量同步任务，DTS要求源数据库的AOF日志保存24小时以上，如为全量同步和增量同步任务，DTS要求源数据库的AOF日志至少保留7天以上（您可在全量同步完成后将AOF日志保存时间设置为24小时以上），否则DTS可能因无法获取AOF日志而导致任务失败，极端情况下甚至可能会导致数据不一致或丢失。由于您所设置的AOF日志保存时间低于DTS要求的时间进而导致的问题，不在DTS的SLA保障范围内。</li> </ul> </li> </ul> <div style="background-color: #e6f2ff; padding: 5px; margin: 10px 0;"> <p> 说明 云数据库Redis不存在该限制。</p> </div> <ul style="list-style-type: none"> <li>如果源库中的某些Key使用了过期（expire）策略，由于可能存在Key已过期但未被及时删除的情况，所以在目标库中查看到的Key数量（例如通过info命令查看）会比源库的Key数量少。如果源库中的某些Key使用了过期（expire）策略，由于可能存在Key已过期但未被及时删除的情况，所以在目标库中查看到的Key数量（例如通过info命令查看）会比源库的Key数量少。</li> <li>如果目标库的数据逐出策略（ maxmemory-policy ）配置为 noeviction 以外的值，可能导致目标库的数据与源库不一致。关于数据逐出策略详情，请参见<a href="#">Redis数据逐出策略介绍</a>。</li> <li>源库单机版Redis同步到目标库集群版Redis的操作限制：由于集群cluster只允许单个命令操作单个slot，若在源库执行多Key操作时，Key不在同一个slot或涉及多个slot，则会出现以下报错：             <pre style="background-color: #f0f0f0; padding: 5px; margin: 5px 0;">CROSSSLOT Keys in request don't hash to the same slot</pre> <p>建议在DTS同步过程中仅执行单Key操作，以免导致链路中断。</p> </li> </ul> |
| 其他限制 | <ul style="list-style-type: none"> <li>同步期间，如源Redis发生扩缩容（如增加或者减少分片）、规格变配（如扩大内存），则您需重新配置任务。且为保障数据一致性，建议重新配置任务前，先清空已同步至目标Redis的数据。</li> <li>同步期间，如源Redis连接地址变化，则需要<a href="#">提交工单</a>进行处理，如果长时间未处理，源Redis会重置AOF文件，届时您需重新配置任务。</li> <li>建议源和目标库版本保持一致，或者从低版本同步到高版本以保障兼容性。如为高版本同步至低版本，可能存在数据库兼容性问题。</li> <li>执行数据同步前需评估源库和目标库的性能，同时建议业务低峰期执行数据同步。否则全量数据初始化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。</li> <li>在DTS同步期间，不允许有除DTS外的数据写入目标库，否则会导致源库与目标库数据不一致。例如，有除DTS外的数据写入目标库时，使用DMS执行在线DDL变更，可能引起目标库数据丢失。</li> </ul>   |

## 4. 准备用于数据同步的数据库账号

在配置数据同步任务时，您需要在任务配置页面分别为源库和目标库填入数据库账号，该数据库账号将用于数据同步。由于不同的数据库类型和同步类型对权限要求有所不同，需要您在配置同步任务前完成数据库账号的创建与授权操作。

### 源库的数据库账号所需权限

| 数据库            | 所需权限  | 创建及授权方法                                     |
|----------------|---|---|
| RDS MySQL      | 待同步对象的读权限。  | 创建数据库和账号和修改账号权限。                            |
| 自建MySQL        | 待同步对象的SELECT权限。<br>REPLICATION CLIENT、REPLICATION SLAVE、<br>SHOW VIEW。<br>建库建表的权限，以允许DTS创建库dts，用于记录同步期间的心跳数据。   | 为自建MySQL创建账号并设置binlog。                      |
| PolarDB MySQL  | 待同步对象的读权限。  | 创建数据库账号。                                    |
| PolarDB O引擎    | 高权限账号。  | 创建数据库账号。                                    |
| DRDS           | 待同步对象的读权限。  | 账号管理  |
| RDS SQL Server | 待同步对象所属数据库的所有者权限。<br> 说明 高权限账号满足要求。   | 修改账号权限。                                     |
| 自建SQL Server   | sysadmin权限。   | CREATE USER和用户权限管理。                         |
| RDS PostgreSQL | 高权限账号，且需为所选数据库的owner（授权账号）。<br> 说明 当源库为RDS PostgreSQL 9.4，且仅需同步DML操作时，账号具备replication权限即可。 | 创建账号和创建数据库。                                 |
| 自建PostgreSQL   | superuser权限。  | CREATE USER和GRANT语法。                        |
| 阿里云Redis实例     | 待同步对象的读写权限。   | 创建与管理账号。                                    |
| 自建Redis        | 无用户和权限概念，可正常执行 <code>psync</code> 或 <code>sync</code> 命令即可。   | 无   |
| 云数据库MongoDB    | <ul style="list-style-type: none"> <li>全量数据迁移：待迁移库的read权限。</li> <li>增量数据迁移：待迁移库、admin库和local库的read权限。</li> </ul>  | 请参见 <a href="#">使用DMS管理MongoDB数据库用户</a> 。   |
| 自建MongoDB      | <ul style="list-style-type: none"> <li>全量数据迁移：待迁移库的read权限。</li> <li>增量数据迁移：待迁移库、admin库和local库的read权限。</li> </ul>  | 请参见 <a href="#">MongoDB Create User说明</a> 。 |
| 自建TiDB         | SHOW VIEW和待迁移对象的SELECT权限。   | 权限管理  |

### 目标库的数据库账号所需权限

| 数据库                          | 所需权限   | 创建及授权方法  |
|------------------------------|--|--|
| RDS MySQL                    | 目标库的读写权限。  | 创建数据库和账号和修改账号权限。   |
| 自建MySQL                      | 目标库的ALL权限。   | 为自建MySQL创建账号并设置binlog。   |
| PolarDB MySQL                | 目标库的ALL权限。   | 创建数据库账号。   |
| PolarDB O引擎                  | 具备数据库Owner权限。  | 数据库Owner在创建数据库时已指定。  |
| DRDS                         | 待同步对象所属数据库的写权限。  | 账号管理   |
| 云数据库Redis                    | 使用实例的密码：无需执行授权。  | 无  |
|                              | 使用自定义账号：读写权限。  | 创建与管理账号。   |
| 自建Redis                      | 无用户和权限概念，只需提供正确的数据库密码。   | 无  |
| 云数据库MongoDB                  | dbAdminAnyDatabase权限、目标库的readWrite权限、local库的read权限。  | 请参见使用DMS管理MongoDB数据库用户。  |
| 自建MongoDB                    | dbAdminAnyDatabase权限、目标库的readWrite权限、local库的read权限。  | 请参见MongoDB Create User说明。  |
| 云原生数据仓库AnalyticDB MySQL      | <ul style="list-style-type: none"> <li>2.0版本：无需填写数据库账号信息，DTS会自动创建账号并授权。</li> <li>3.0版本：读写权限。</li> </ul>  | 3.0版本：创建数据库账号  |
| 云原生数据仓库AnalyticDB PostgreSQL | 初始账号或具备RDS_SUPERUSER权限。  | <ul style="list-style-type: none"> <li>初始账号：创建数据库账号。</li> <li>RDS_SUPERUSER权限：用户权限管理。</li> </ul> |
| 阿里云消息队列Kafka版                | 非必填项，无。<br><span style="background-color: #e1f5fe; padding: 5px; border: 1px solid #ccc;"> <span style="font-size: 1.2em; color: #0070c0;">?</span> 说明 如果Kafka实例的实例类型为VPC实例，无需配置数据库账号和数据库密码。                     </span> | 无  |
| 自建Kafka                      | 非必填项，无。<br><span style="background-color: #e1f5fe; padding: 5px; border: 1px solid #ccc;"> <span style="font-size: 1.2em; color: #0070c0;">?</span> 说明 如Kafka集群未开启验证可不填写。                     </span>                    | 无  |
| DataHub                      | 无需配置数据库账号。   | 无  |
| Elasticsearch                | 创建时设置的登录名（默认为elastic）和登录密码。  | 创建实例。  |
| MaxCompute                   | 待同步项目的CREATE TABLE、CREATE INSTANCE、CREATE RESOURCE、CREATE JOB、List权限。  | 同步任务配置时系统将自动授权。  |
| 表格存储(Tablestore)             | 无需配置数据库账号。   | 无  |
| ClickHouse                   | 读写权限。  | 创建账号。  |

## 双向同步任务对账号权限的要求

双向同步任务要求源和目标数据库具备如下权限，以允许DTS在源和目标数据库中创建库dts，用于做数据防循环复制。

| 数据库            | 所需权限   | 创建及授权方式                |
|----------------|--|------------------------|
| RDS MySQL      | 高权限账号。   | 创建数据库和账号。              |
| 自建MySQL        | 待同步对象的SELECT权限。<br>REPLICATION CLIENT、REPLICATION SLAVE、<br>SHOW VIEW。<br>建库建表的权限，以允许DTS创建库dts，用于做<br>数据防循环复制。 | 为自建MySQL创建账号并设置binlog。 |
| RDS PostgreSQL | 高权限账号，且需为所选数据库的owner（授权账号）。  | 创建账号和创建数据库。            |
| 自建PostgreSQL   | superuser权限。   | CREATE USER和GRANT语法。   |
| PolarDB MySQL  | 高权限账号。   | 创建数据库账号。               |
| 云数据库Redis      | 使用实例的密码：无需执行授权。  | 无                      |
|                | 使用自定义账号：读写权限。  | 创建与管理账号。               |
| 自建Redis        | 无用户和权限概念，可正常执行 <code>psync</code> 或 <code>sync</code> 命令即可。  | 无                      |

# 5.同步作业管理

## 5.1. 设置同步对象在目标实例中的名称

默认情况下，被同步的对象（如库、表等）在源实例和目标实例中的名称相同，您可以通过DTS提供的对象名映射功能，设置被同步的对象在目标实例中的名称。

### 注意事项

仅支持在配置数据同步作业时，在设置选择同步对象环节中执行该操作。

**说明** 数据同步作业一旦启动，请勿再执行该操作，否则可能导致同步失败。

### 操作步骤

1. 在选择同步对象环节，将需要同步的对象移动至已选择对象区域框中后，把鼠标指针放置在要进行名称映射的库或表上，并单击对象后出现的编辑。

**说明** 由于不同的数据库类型支持的对象有所不同，您可以将鼠标指针放置在目标对象上，如果出现编辑则表示支持该操作。



2. 在弹出的对话框中，设置该对象在目标实例中名称。
  - o 库名映射  
在弹出的编辑库对话框中，设置该数据库在目标实例中的名称。

编辑库 ✕

**注意：** 编辑库名后，目标数据库的库名为修改后的名称。

\* 数据库名称:

原库名: dtstestdata

**DML&DDL过滤** 请选中您需要的DDL或者DML的复选框! ⓘ

DML Filter:  insert  update  delete

Table:  create  alter  drop  rename  truncate

View:  create  alter  drop

Procedure:  create  alter  drop

Trig&Func:  create trigger  drop trigger  create function  drop function

Index:  create  drop

确定

- 表名映射  
在弹出的编辑表对话框中，设置该数据表在目标实例中的名称。

编辑表 ✕

**注意：** 编辑表名或列名后，目标数据库的表名或列名为修改后的名称。

\* 数据库表名称:

原表名: test1

过滤条件:

验证语法

**DML&DDL过滤** 请选中您需要的DDL或者DML的复选框! ⓘ

DML Filter:  insert  update  delete

Table:  create  alter  drop  rename  truncate

View:  create  alter  drop

Procedure:  create  alter  drop

Trig&Func:  create trigger  drop trigger  create function  drop function

Index:  create  drop

- 列名映射  
在弹出的编辑表对话框中，设置对应的列在目标实例中的名称。



说明 在此步骤中，您可以通过取消对应列的复选框来过滤不需要同步的列。

3. 单击**确定**。
4. 根据提示完成后续的数据同步作业配置。

## 5.2. 通过SQL条件过滤任务数据

在配置任务对象时，您可以设置SQL过滤条件过滤任务数据。通过设置后，只有满足过滤条件的数据才会被同步或迁移到目标数据库。该功能可应用于数据的定期同步或迁移、拆分数据表等多种应用场景。

### 前提条件

已启动任务配置，并处于配置任务对象及高级配置环节。关于如何启动任务配置，请参见[配置同步任务（新控制台）](#)。

### 功能限制

仅支持过滤当前表的字段，不支持跨表过滤。

### 操作步骤

1. 在配置任务对象及高级配置环节，将任务对象移动到已选择对象区域框中后，将鼠标指针放置在要修改的数据表上并单击右键。
2. 在弹出的编辑表对话框中，填入过滤条件。

编辑表
✕

**注意：** 编辑表名或列名后，目标数据库的表名列名称将修改后的名称

表名称:

过滤条件: 

order>100

**勾选所需同步的DML&DDL操作**

勾选所需同步的DML操作  全选

insert  update  delete

勾选所需同步的DDL操作  全选

Table

create  alter  drop  rename  truncate

---

View

create  alter  drop

---

Procedure

create  alter  drop

---

Trig&Func

create function  drop function  create trigger  drop trigger

---

Index

create  drop

确定
取消

#### 说明

- 过滤条件支持标准的SQL WHERE语句（仅支持 = 、 != 、 < 、 > 、 in 操作符），只有满足 WHERE条件的数据才会被同步或迁移到目标数据库中。本案例填入 `order>100`。
- 过滤条件支持按时间条件过滤，但SQL WHERE语句中时间的格式有一定要求，例如过滤2020年之后创建的增量数据，您需要填入 `create_time>'2020-01-01'` 或 `create_time>'2020-01-01 00:00:00'`，不支持填入 `create_time>'2020'`。
- 过滤条件中如需使用引号，请使用单引号（'），例如 `address in('hangzhou','shanghai')`。
- 如源库中同一张表内存在两个列A和a，由于过滤条件忽略列名的英文大小写，则SQL WHERE语句仅支持过滤列A，不支持过滤列a，例如支持 `WHERE A=10`，不支持 `WHERE a=10`。

3. 单击**确定**。

4. 根据提示完成后续的任务配置。

## 5.3. 目标库对象名称大小写策略

在配置任务时，DTS支持配置任务到目标库对象的名称大小写策略，来设置目标实例中任务对象的库名、表名和列名的英文大小写规则。

### 注意事项

如果您指定目标库对象名称大小写策略，则任务对象具有如下限制：

- 不支持配置任务中视图、函数、存储过程。否则目标库对象在目标库中的表名变化后，对应的视图、函数和存储过程可能会创建失败。
- 不支持任务对象包含检查约束和计算列。

### 策略说明

DTS目前支持配置如下四种目标库对象名称大小写策略：

- **DTS默认策略**  
DTS会根据目标数据库类型和相关参数（如 `lower_case_table_names` ），来配置目标实例中迁移对象的库名、表名、列名等英文大小写规则。

| 目标数据库类型   | 目标库参数                                     | 目标实例中迁移对象的英文大小写规则              |
|---|---|--------------------------------|
| <ul style="list-style-type: none"> <li>◦ 自建MySQL、RDS MySQL</li> <li>◦ RDS MariaDB</li> <li>◦ PolarDB MySQL</li> <li>◦ PolarDB-X</li> <li>◦ 云原生数据仓库AnalyticDB MySQL版 3.0</li> </ul>  | <code>lower_case_table_names</code> 为1或2。 | 库名和表名的英文映射为小写，列名的英文大小写和源库保持一致。 |
| <ul style="list-style-type: none"> <li>◦ 云原生数据仓库AnalyticDB MySQL版 2.0</li> <li>◦ DataHub</li> <li>◦ MaxCompute</li> <li>◦ HybridDB MySQL</li> </ul>   | 无   | 库名、表名和列名的英文映射为小写。              |
| 自建Oracle  | 无   | 库名、表名和列名的英文映射成大写。              |
| <ul style="list-style-type: none"> <li>◦ 自建SQL Server、RDS SQL Server</li> <li>◦ PolarDB O引擎</li> <li>◦ 自建PostgreSQL、RDS PostgreSQL</li> <li>◦ RDS PPAS</li> <li>◦ 云原生数据仓库AnalyticDB PostgreSQL</li> <li>◦ 自建Db2</li> <li>◦ 自建MongoDB、阿里云MongoDB实例</li> <li>◦ 自建Redis、阿里云Redis实例</li> <li>◦ TableStore</li> <li>◦ Elasticsearch</li> </ul> | 无   | 库名和表名的英文映射为小写，列名的英文大小写和源库保持一致。 |

- **和源库保持一致**  
迁移或同步到目标库对象的库名、表名和列名的英文大小写与源库保持一致。
- **与目标库默认策略保持一致（大写）**  
迁移或同步到目标库对象的库名、表名和列名的英文全部大写。
- **与目标库默认策略保持一致（小写）**  
迁移或同步到目标库对象的库名、表名和列名的英文全部小写。

## 5.4. 修改数据同步的附加列规则

DTS在将数据同步到MaxCompute时，会在同步的目标表添加一些附加列，用于元数据管理、排序去重等。如果附加列和目标表中已有的列出现名称冲突将会导致数据同步失败，为避免出现此类问题，您可以修改同步附加列的命名规则。

### 背景信息

在修改数据同步的附加列规则前，您需要评估附加列和目标表中已有的列是否会出现名称冲突。关于附加列名称及其定义请参见下表。

**说明** 您可以在操作步骤中的**步骤2**中选择附加列规则是否为新版。

#### 附加列名称和定义说明

| 目标实例       | 旧版附加列名称  | 新版附加列名称  | 附加列定义                   |
|------------|--|--|-------------------------|
| MaxCompute | <ul style="list-style-type: none"> <li>record_id</li> <li>operation_flag</li> <li>utc_timestamp</li> <li>before_flag</li> <li>after_flag</li> <li>modifytime_year</li> <li>modifytime_month</li> <li>modifytime_day</li> <li>modifytime_hour</li> <li>modifytime_minute</li> </ul> | <ul style="list-style-type: none"> <li>new_dts_sync_record_id</li> <li>new_dts_sync_operation_flag</li> <li>new_dts_sync_utc_timestamp</li> <li>new_dts_sync_before_flag</li> <li>new_dts_sync_after_flag</li> <li>new_dts_sync_modifytime_year</li> <li>new_dts_sync_modifytime_month</li> <li>new_dts_sync_modifytime_day</li> <li>new_dts_sync_modifytime_hour</li> <li>new_dts_sync_modifytime_minute</li> </ul> | 参见 <b>增量日志表结构定义说明</b> 。 |

### 操作步骤

1. 开始配置数据同步作业，详情请参见**配置同步任务（旧控制台）**中的第1步到第7步。
2. 在**选择同步对象**环节的页面底部，根据业务需求选择**是否启用新的附加列规则**。

同步初始化:  结构初始化

源库对象

若全局搜索，请先展开树

- dtstestdata
  - Tables
    - vip
    - customer
- dtstestdata617
- dtstestdatanew

已选择对象 (鼠标移到对象行,点击编辑可修改对象名或过滤条件) 详情点我

- dtstestdata (1个对象)
  - order

\*映射名称更改:  不进行库表名称批量更改  要进行库表名称批量更改

**是否启用新的附加列规则  是  否**

了解该功能

| 选项 | 说明 |
|----|----|
|----|----|

| 选项 | 说明   |
|----|--|
| 是  | <p>选择为是，则启用新的附加列规则。</p> <p> 说明 如果旧版附加列的名称和目标表中已有的列出现名称冲突，请选择为是来避免。</p> <p>DTS在旧版附加列添加名为 <code>new_dts_sync_</code> 的前缀组成新的附加列。<br/>例如旧版附加列的名称为 <code>record_id</code>，那么新的附加列的名称即为 <code>new_dts_sync_record_id</code>。</p> |
| 否  | 使用旧版附加列的名称。  |

 说明 关于数据同步涉及的附加列名称和定义，请参见[附加列名称和定义说明](#)。

3. 根据提示，完成后续的数据同步作业配置。

## 相关文档

- [同步方案概览](#)
- [DTS数据同步方案概览](#)

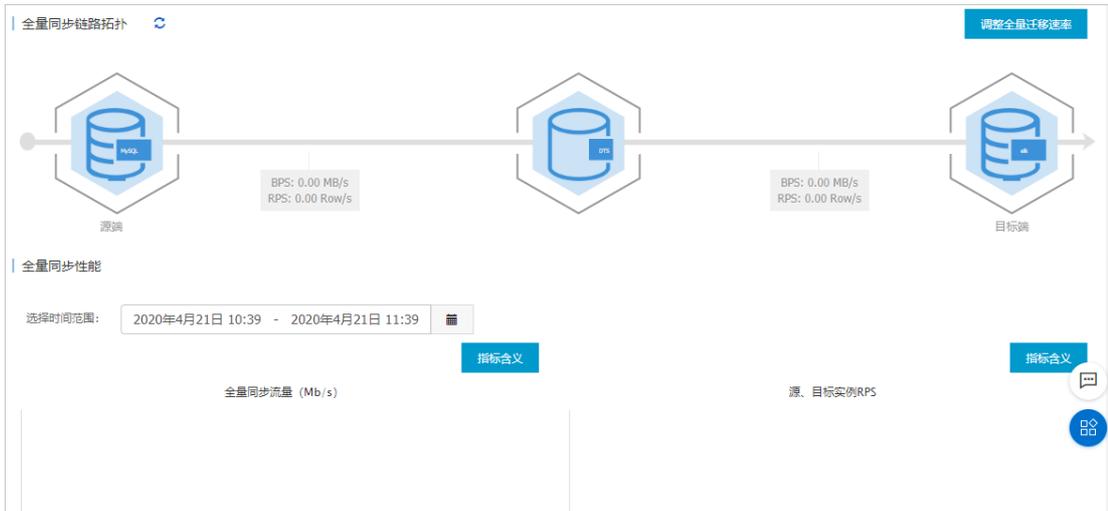
## 5.5. 查看同步链路状态和性能

DTS可以展示全量同步性能、增量同步性能，使您可以了解数据同步作业的链路状态信息和多种性能指标，为您管理数据同步作业提供参考依据。

### 操作步骤

1. 登录[数据传输控制台](#)。
2. 在左侧导航栏，单击[数据同步](#)。
3. 在同步作业列表页面顶部，选择数据同步实例所属地域。
4. 在数据同步列表页面，单击目标实例ID。
5. 根据要执行的操作，选择下述操作步骤：
  - o 查看全量同步链路状态和性能
    - a. 在左侧导航栏，选择[性能监控](#) > [全量同步性能](#)。

b. 在本页面，您可以查看到DTS链路的拓扑信息，同时您可以选择时间范围，查看全量同步性能的各指标趋势图。

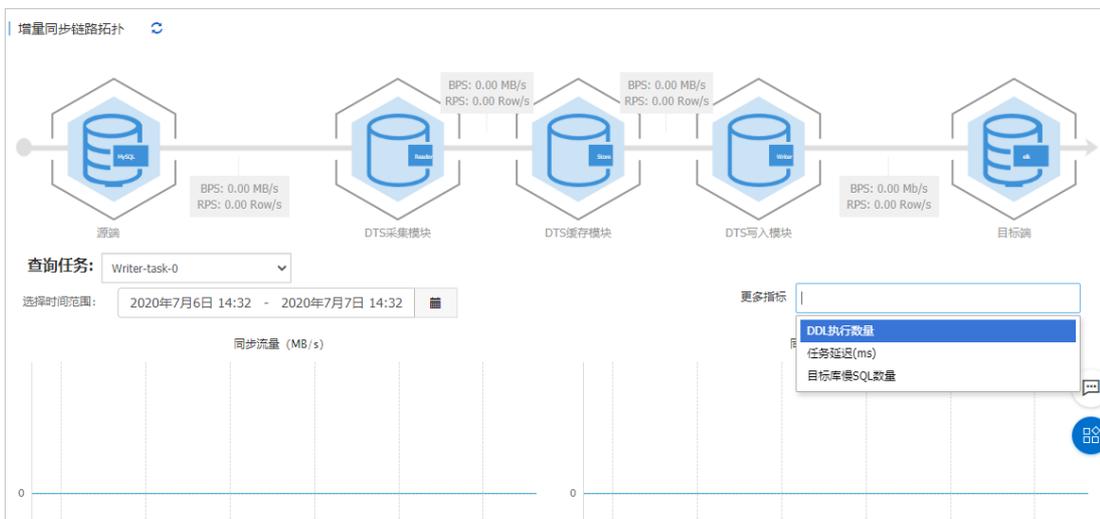


| 功能       | 说明  |
|----------|---|
| 全量同步链路拓扑 | <p>全量同步链路主要展示源库到DTS、DTS到目标库两部分的读写情况和网络信息，相关参数解释如下：</p> <ul style="list-style-type: none"><li>■ 源库到DTS的链路<ul style="list-style-type: none"><li>■ BPS：DTS每秒从源库读取数据量，单位为MB/s。</li><li>■ RPS：DTS每秒从源库读取的行数。</li><li>■ 网络延迟：DTS与源库间的网络延迟。</li></ul></li><li>■ DTS到目标库的链路<ul style="list-style-type: none"><li>■ BPS：DTS每秒写入到目标库的数据量，单位为MB/s。</li><li>■ RPS：DTS每秒写入到目标库的行数。</li><li>■ 网络延迟：DTS与目标库间的网络延迟。</li></ul></li></ul> |
| 全量同步性能   | <p>可查看全量同步流量、源和目标实例的RPS、读写响应时间以及网络延时等信息。</p> <p><b>?</b> 说明 将鼠标指针放置在对应监控指标的指标含义按钮上，即可查看性能指标的详细解释和判断参考。</p>  |

o 查看增量同步链路状态和性能

a. 在左侧导航栏，选择性能监控 > 增量同步性能。

b. 在本页面，您可以查看到DTS链路的拓扑信息，同时您可以选择时间范围，查看增量同步性能的各指标趋势图。



| 功能       | 说明   |
|----------|--|
| 增量同步链路拓扑 | <p>增量同步链路拓扑主要展示链路中各模块间的数据传输情况和网络延迟信息，模块包含：源库、DTS采集模块、DTS缓存模块、DTS写入模块、目标库。相关参数解释如下：</p> <ul style="list-style-type: none"> <li>■ <b>BPS</b>：链路中模块间每秒传输的数据量，单位为MB/s。</li> <li>■ <b>RPS</b>：链路中模块间每秒传输的行数。</li> <li>■ <b>网络延迟</b>：链路中模块间的网络延迟。</li> </ul>   |
| 增量同步性能   | <p>可查看同步流量、同步速度和同步延迟等信息，单击<b>更多指标</b>对应的下拉列表，可选择展示更多的性能指标，例如DDL执行数量、任务延迟、目标库慢SQL数量。相关性能指标说明如下：</p> <ul style="list-style-type: none"> <li>■ <b>同步流量</b>：数据写入模块每秒从数据拉取模块拉取到的数据量，单位为MB/s。</li> <li>■ <b>同步速度</b>：每秒同步到目标库的事务数。</li> <li>■ <b>同步延迟</b>：目标库同步的最新数据的时间戳和源库当前时间戳的时间差，单位为毫秒。</li> <li>■ <b>DDL执行数量</b>：选定的时间范围内，目标库执行DDL语句的数量。</li> <li>■ <b>任务延迟 (ms)</b>：数据在源库产生的时间点与被写入到目标库的时间点的差值。例如某条数据在源库产生的时间是7点，DTS将这条数据写入到目标端的时间是8点，则该任务存在1个小时的延迟。</li> <li>■ <b>目标库慢SQL数量</b>：选定的时间范围内，目标库产生的慢SQL数量。</li> </ul> |

## 5.6. 重新配置数据同步作业

当您需要停止数据同步或重新配置数据同步时，您可以通过本文的方法来重置数据同步作业。

### 前提条件

数据同步作业不处于未配置状态。

### 费用影响

- 包年包月：无影响。
- 按量付费：数据同步作业将转变为未配置状态并停止计费，再次完成数据同步作业的配置且成功启动后才会开始计费。

### 操作步骤

1. 登录[数据传输控制台](#)。

2. 在左侧导航栏，单击数据同步。
3. 在同步作业列表页面顶部，选择数据同步实例所属地域。
4. 找到目标同步作业，在其操作列选择更多 > 重新配置同步。

**警告** 执行该操作后，将造成下述影响，请谨慎操作：

- 数据同步作业将停止同步数据，即不再将源库的增量数据变更和结构变更同步至目标库。
- 数据同步作业的配置信息将被清空，并转变为未配置状态。

5. 在弹出的对话框中，单击确定。

## 相关文档

如需重新配置数据同步作业，请参见[DTS数据同步方案概览](#)。

## 5.7. 新增同步对象

DTS支持在数据同步的过程中新增同步对象，无需重新配置同步任务，以在降低对业务影响的情况下变更同步对象。本文将介绍新增同步对象的操作流程。

### 前提条件

- 数据同步作业处于同步中、已暂停或同步失败的状态。
- 为保障DTS可以正常连接源和目标库、读取源库的库表信息，在配置数据同步作业时，请勿对源库和目标库执行升级、变配、网络切换、跨可用区迁移等操作，即源库和目标库须处于正常运行的状态。

### 注意事项

新增对象的同步起始时间点与同步作业是否配置了同步初始化有关，具体如下：

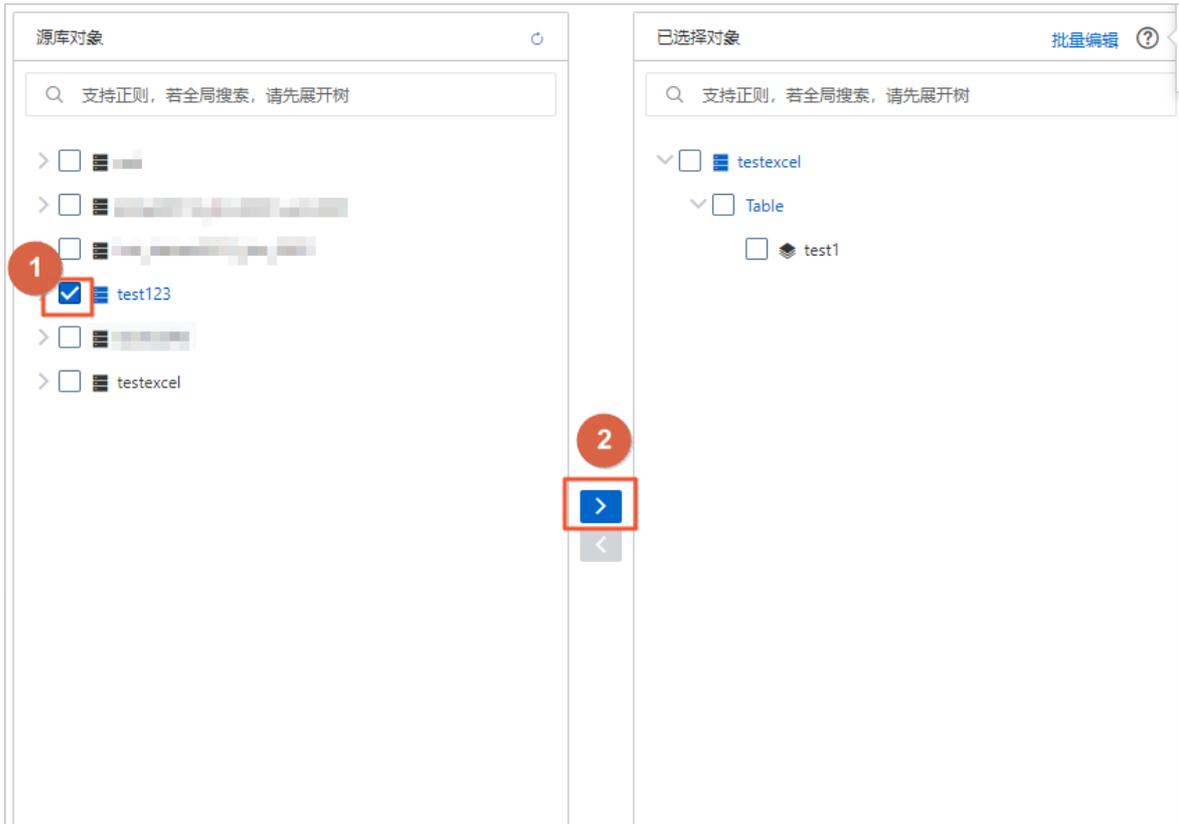
- 如果未配置同步初始化，那么同步对象会从源实例中当前时间点产生的增量数据开始同步。
- 如果配置了同步初始化，那么新增对象会先进行结构初始化和全量数据初始化，然后开始同步增量数据。

### 操作步骤

1. 登录[数据传输控制台](#)。
2. 在左侧导航栏，单击数据同步。
3. 在同步作业列表页面顶部，选择数据同步实例所属地域。
4. 找到目标同步作业，单击其操作列的更多 > 修改同步对象。



5. 在源库对象区域框中，单击需要新增的同步对象，并单击 > 图标将其移动至已选择对象区域框中。



6. 单击预检查并启动。

**说明**

- 在修改同步对象后，会对同步作业进行预检查。只有预检查通过后，才能成功启动同步作业。
- 如果预检查失败，单击具体检查项后的图标，查看具体的失败详情。根据失败原因修复后，重新进行预检查。

7. 预检查完成后开始同步数据。

## 5.8. 移除同步对象

DTS支持在数据同步的过程中移除同步对象，无需重新配置同步任务，以在降低对业务影响的情况下变更同步对象。本文将介绍移除同步对象的操作流程。

### 前提条件

- 数据同步作业处于同步中、已暂停或同步失败的状态。
- 为保障DTS可以正常连接源和目标库、读取源库的库表信息，在配置数据同步作业时，请勿对源库和目标库执行升级、变配、网络切换、跨可用区迁移等操作，即源库和目标库须处于正常运行的状态。

### 注意事项

将同步对象移除后，同步作业将不再同步该对象的增量数据到目标数据库。

### 操作步骤

- 登录[数据传输控制台](#)。
- 在左侧导航栏，单击数据同步。
- 在同步作业列表页面顶部，选择目标数据同步作业所属的地域。

4. 找到目标同步作业，单击其操作列的更多 > 修改同步对象。



5. 在已选择对象区域框中，单击需要移除的同步对象，并单击 < 图标将其移动至源库对象区域框中。



6. 单击预检查并启动。

#### 说明

- 在修改同步对象后，会对同步作业进行预检查。只有预检查通过后，才能成功启动同步作业。
- 如果预检查失败，单击具体检查项后的  图标，查看具体的失败详情。根据失败原因修复后，重新进行预检查。

7. 预检查完成后开始同步数据。

## 5.9. 结束数据同步作业

如果不再需要数据同步，您可以结束数据同步作业，停止将源库的数据变更同步至目标库。

### 前提条件

数据同步作业处于同步中、已暂停、同步失败、同步初始化中或同步初始化失败状态。

### 注意事项

- 结束数据同步作业后，数据同步作业将处于已完成状态，不再将源库的数据变更同步至目标库中。
- 结束同步作业后，不能重新启动。如需再次同步，您需要重新配置同步作业。

### 操作步骤

1. 登录 [数据传输控制台](#)。

2. 在同步作业列表页面顶部，选择数据同步实例所属地域。
3. 在数据同步列表页面，单击目标实例ID。
4. 找到目标同步作业，在其操作列选择更多 > 结束同步。



相关文档

释放实例

## 5.10. 修复同步作业预检查失败的问题

DTS在执行数据同步作业之前，会对源库和目标库执行预检查，本文介绍各类检查项失败可能的原因及修复方法。

| 预检查项目               | 失败可能的原因   | 修复方法  |
|---------------------|---|---|
| 源库连接性检查<br>目标库连接性检查 | 数据库账号或密码不正确。  | 获取到正确的密码，然后修改同步作业中配置的源或目标库的数据库密码，详情请参见 <a href="#">修改数据库密码</a> 。  |
|                     | 源或目标库对来源IP进行了限制。  | <ul style="list-style-type: none"> <li>源或目标库为MySQL：为数据库账号重新授权，授权IP地址列表需包含DTS服务器的IP地址。关于如何选择DTS服务器的IP地址，请参见<a href="#">迁移、同步或订阅本地数据库时需添加的IP白名单</a>。</li> <li>源或目标库库为SQL Server：关闭防火墙或禁用trigger。</li> <li>源或目标库库为Oracle：修改 TCP.VALIDN ODE_CHECKING 的值为 no 并重启进程。</li> </ul> |
|                     | 源或目标库所属的服务器设置了防火墙。  | 关闭相关防火墙限制。  |
|                     | DTS服务器与源或目标库的网络互通问题。  | 您可以通过 <a href="#">提交工单</a> 的形式联系阿里云工程师协助解决。   |
| 源库版本检查              | DTS不支持该版本的数据库，DTS支持数据库版本信息请参见 <a href="#">同步方案概览</a> 。  | 升级或降级数据库版本。   |
| 数据库存在性检查            | 当待同步的数据库在目标实例中不存在，由DTS自动创建时，可能由于下述原因导致创建失败： <ul style="list-style-type: none"> <li>数据库名包含小写字母、数字、下划线（_）或中划线（-）以外的其他特殊字符。</li> <li>数据库的字符集为utf8、GBK、latin1或utf8mb4以外的字符集。</li> <li>数据同步作业中配置的目标数据库账号权限不足。</li> </ul> | 手动在目标实例中创建满足要求的数据库或调整数据库账号。   |

| 预检查项目         | 失败可能的原因   | 修复方法   |
|---------------|---|--|
| 源库权限检查        | 数据同步作业中配置的数据库账号的权限不足。   | 调整数据库账号的权限。  |
| 目标库权限检查       | <p><b>说明</b> 不同类型数据库所需的权限不同，详情请参见具体的配置案例：<a href="#">DTS数据同步方案概览</a>。</p> |  |
| 同名对象存在性检查     | 目标库存在和待同步对象同名的对象，例如库名或表名相同。   | <ul style="list-style-type: none"> <li>登录目标库，调整名称冲突的库或表名。</li> <li>使用DTS提供的库表映射功能，将同名的待同步对象映射为目标库中的其他对象名，详情请参见<a href="#">设置同步对象在目标实例中的名称</a>。</li> <li>从同步对象中移除同名对象，详情请参见<a href="#">移除同步对象</a>。</li> </ul> |
| 源库server_id检查 | <code>server-id</code> 设置的不是大于等于2的整数。                                     | 登录源库，调整 <code>server-id</code> 的值，详情请参见 <a href="#">源库server-id检查</a> 。  |
| 源库Binlog开启检查  | 源库没有开启Binlog功能。   | 登录源库，开启Binlog功能，详情请参见 <a href="#">源库binlog是否开启检查</a> 。   |
| 源库Binlog模式检查  | 源库的Binlog模式没有设置为ROW。  | 登录源库执行命令 <code>set global binlog_format='ROW'</code> ，然后重启MySQL进程。   |
| 约束完整性检查       | 待同步子表依赖的父表没有被选择为同步对象，从而破坏了外键约束完整性。  | <ul style="list-style-type: none"> <li>删除约束完整性检查失败的子表的外键依赖。</li> <li>在同步对象中加入相应父表，详情请参见<a href="#">新增同步对象</a>。</li> <li>从同步对象中移除相应子表，详情请参见<a href="#">移除同步对象</a>。</li> </ul>                                 |
| 存储引擎检查        | 待同步表的存储引擎为FEDERATED、MRG_MyISAM或TokuDB。                                    | 登录源库，将表的存储引擎修改为InnoDB。   |
| 字符集检查         | DTS不支持同步对象的字符集（例如UCS2）。   | 登录源库，将表的字符集修改为utf8、GBK、latin1或utf8mb4。   |
| 复杂拓扑存在性检查     | DTS不支持源和目标实例组成的拓扑，DTS支持的拓扑及注意事项请参见 <a href="#">数据同步拓扑介绍</a> 。             | 结束或等待冲突的任务完成后再次执行数据同步。   |
| MySQL密码格式检查   | 源库使用数据库密码格式是否为老版。   | 调整数据库密码的格式，详情请参见 <a href="#">old_passwords</a> 。   |

## 5.11. Kafka集群的数据存储格式

DTS支持选择迁移或同步到Kafka集群的数据存储格式，本文为您介绍数据格式的定义说明，方便您根据定义解析数据。

### 数据存储格式

DTS支持将写入至Kafka集群的数据存储为如下三种格式：

- DTS Avro**：一种数据序列化格式，可以将数据结构或对象转化成便于存储或传输的格式。
- Shareplex Json**：数据复制软件Shareplex读取源库中的数据，将数据写入至Kafka集群时，数据存储格式为Shareplex Json。
- Canal Json**：Canal解析数据库增量日志，并将增量数据传输至Kafka集群，数据存储格式为Canal Json。

## DTS Avro

DTS Avro为默认存储格式。您需要根据DTS Avro的schema定义进行数据解析，schema定义详情请参见迁移或同步至Kafka集群中的数据均[DTS Avro的schema定义](#)。

## Shareplex Json

参数说明

| 参数       | 说明   |
|----------|--|
| time     | 数据库中事务的提交时间，格式为yyyy-MM-ddTHH:mm:ssZ（UTC时间）。  |
| userid   | 提交事务的用户ID。   |
| op       | 数据操作类型，包括INSERT, UPDATE, DELETE, TRUNCATE, DROP COLUMN, UPDATE BEFORE, UPDATE AFTER。 |
| scn      | 系统变化编号SCN（System Change Number），用以标识数据库在某个确切时刻提交事务的版本。每个已提交的事务分配一个唯一的SCN。            |
| rowid    | 用于定位数据库中一条记录的一个相对唯一地址值。  |
| trans    | 事务ID。  |
| seq      | 事务内部的操作序号，从1开始记录。  |
| size     | 事务内部的操作总数。   |
| table    | 表名。  |
| idx      | 事务内部操作的索引，格式为 seq/size ，例如 1/11 表示，在操作总数为11的事务内部，该操作的序号为1。                           |
| posttime | 事务提交至目标库的时间。   |

示例如下：

- 插入数据：

```
{
  "meta": {
    "time": "2017-06-16T14:24:34",
    "userid": 84,
    "op": "ins",
    "scn": "14589063118712",
    "rowid": "AAATGpAAIAAItcIAAA",
    "trans": "7.0.411499",
    "seq": 1,
    "size": 11,
    "table": "CL_BIZ1.MIO_LOG",
    "idx": "1/11",
    "posttime": "2017-06-16T14:33:52"
  },
  "data": {
    "MIO_LOG_ID": "32539737"
  }
}
```

- 更新数据：

```
{
  "meta": {
    "time": "2017-06-16T15:38:13",
    "userid": 84,
    "op": "upd",
    "table": "CL_BIZ1.MIO_LOG"
  },
  "data": {
    "CNTR_NO": "1171201606"
  },
  "key": {
    "MIO_LOG_ID": "32537893",
    "PLNMIO_REC_ID": "31557806",
    "POL_CODE": null,
    "CNTR_TYPE": null,
    "CNTR_NO": "1171201606syui26"
  }
}
```

● 删除数据:

```
{
  "meta": {
    "time": "2017-06-16T15:51:35",
    "userid": 84,
    "op": "del",
  },
  "data": {
    "MIO_LOG_ID": "32539739",
    "PLNMIO_REC_ID": "31557806",
    "POL_CODE": null,
    "CNTR_TYPE": null,
    "CG_NO": null
  }
}
```

## Canal json

### 参数说明

| 参数        | 说明   |
|-----------|--|
| database  | 数据库名称。   |
| es        | 操作在源库的执行时间，13位Unix时间戳，单位为毫秒。<br><br><div style="border: 1px solid #ccc; background-color: #e0f2f1; padding: 5px; border-radius: 4px;"> <span style="font-size: 1em;">?</span> 说明 Unix时间戳转换工具可用搜索引擎获取。         </div> |
| id        | 操作的序列号。  |
| isDdl     | 是否是DDL操作。<br><ul style="list-style-type: none"> <li>• true: 是。</li> <li>• false: 否。</li> </ul>   |
| mysqlType | 字段的数据类型。   |

| 参数         | 说明   |
|------------|--|
| old 与 data | 变更前的数据。<br>? 说明 2022年3月20日之前创建的DTS订阅实例，old 的值是数据，data 的值是null。为了和开源社区保持一致，2022年3月20日起创建或重启的DTS订阅实例，data 的值是数据，old 的值是null。 |
| pkNames    | 主键名称。  |
| sql        | SQL语句。   |
| sqlType    | 经转换处理后的字段类型，如unsigned int会被转化为Long，unsigned long会被转换为BigDecimal。   |
| table      | 表名。  |
| ts         | 操作写入到目标库的时间，13位Unix时间戳，单位为毫秒。<br>? 说明 Unix时间戳转换工具可用搜索引擎获取。   |
| type       | 操作的类型，比如DELETE、UPDATE、INSERT。  |

更新数据的示例如下：

? 说明 2022年3月20日之前创建的DTS订阅实例，源表的DELETE 语句同步到kafka，其中 old 的值是数据，data 的值是null。为了和开源社区保持一致，2022年3月20日起创建或重启的DTS订阅实例，data 的值是数据，old 的值是null。

**2022年3月20日之前创建的DTS订阅实例**

```
{
  "old": [
    {
      "shipping_type": "aaa"
    }
  ],
  "database": "dbname",
  "es": 1600161894000,
  "id": 58,
  "isDdl": false,
  "mysqlType": {
    "id": "bigint(20)",
    "shipping_type": "varchar(50)"
  },
  "pkNames": [
    "id"
  ],
  "sql": "",
  "sqlType": {
    "id": -5,
    "shipping_type": 12
  },
  "table": "tablename",
  "ts": 1600161894771,
  "type": "DELETE"
}
```

#### 2022年3月20日起创建或重启的DTS订阅实例

```
{
  "data": [
    {
      "id": "500000287",
      "shipping_type": null
    }
  ],
  "database": "dbname",
  "es": 1600161894000,
  "id": 58,
  "isDdl": false,
  "mysqlType": {
    "id": "bigint(20)",
    "shipping_type": "varchar(50)"
  },
  "pkNames": [
    "id"
  ],
  "sql": "",
  "sqlType": {
    "id": -5,
    "shipping_type": 12
  },
  "table": "tablename",
  "ts": 1600161894771,
  "type": "DELETE"
}
```

DDL操作示例如下：

```

{
  "database": "dbname",    表示同步的数据库名称
  "es": 1600161894000,    表示源库数据写入到binlog的时间
  "id": 58,                DTS缓存的偏移量
  "isDdl": true,          是否同步DDL
  "sql": "eg: create xxx", Binlog的DDL语句
  "table": "tablename",   同步的表名
  "ts": 1600161894771,    DTS将数据写入目标的时间
  "type": "DDL"
}

```

## 5.12. Kafka Partition同步策略说明

在配置同步到Kafka的数据同步作业时，为提升同步的性能，您可以调整Kafka Partition（分区）同步策略，例如根据Hash结果将数据同步至不同的Partition。

### Hash算法

DTS采用Java默认的hashCode算法来计算Hash值。

### 配置方法

您需要在配置数据同步作业的选择同步对象环节中配置Kafka Partition同步策略。配置案例请参见[从RDS MySQL同步至自建Kafka集群](#)，更多相关案例请参见[DTS数据同步方案概览](#)。

 **警告** 数据同步作业启动后，请勿修改目标Topic的Partition数量，否则将导致数据同步失败。

### 策略说明

| 策略名称                       | 说明  | 优缺点   |
|----------------------------|---|---|
| 全部递到Partition 0            | 将所有数据和DDL信息都投递到目标Topic的Partition 0。   | <ul style="list-style-type: none"> <li>优势：所有对象的创建、更改顺序都和源库保持一致。</li> <li>缺点：性能一般。</li> </ul>                                    |
| 按库名+表名的hash值投递到不同Partition | <p>将库名与表名合并作为Partition Key来计算Hash值，然后将各个表的数据和DDL信息投递到目标Topic中的各个Partition。</p> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin-top: 10px;"> <p> <b>说明</b></p> <ul style="list-style-type: none"> <li>同一个表的数据和DDL信息会投递到同一个Partition。</li> <li>与表无关的DDL信息（例如CREATE DATABASE）则投递到Partition 0。</li> </ul> </div> | <ul style="list-style-type: none"> <li>优势：单个表的创建、变更顺序和源端保持一致，性能较好。</li> <li>缺点：由于不同的表投递到不同的Partition，表与表之间的先后顺序无法保证。</li> </ul> |

| 策略名称                    | 说明  | 优缺点   |
|-------------------------|---|---|
| 按主键的hash值投递到不同Partition | <p>将表中的列（默认为主键，若无主键则使用唯一键）作为Partition Key来计算Hash值，然后将不同的行投递到目标Topic的各个Partition中。您也可以指定单个或多个列作为Partition Key来计算Hash值。</p> <p> 说明</p> <ul style="list-style-type: none"><li>• 此策略下，DDL信息默认投递到目标Topic的Partition 0中。</li><li>• 如果表没有主键和唯一键，DTS会将其数据和DDL信息都投递到目标Topic的Partition 0。</li></ul> | <ul style="list-style-type: none"><li>• 优势：该策略性能最好。</li><li>• 缺点：只能保证单条记录的变更的先后顺序，无主键表以及多张表之间的先后顺序无法保证。</li></ul> |

# 6. MySQL间数据同步

## 6.1. MySQL实例间的双向同步

数据传输服务DTS (Data Transmission Service) 支持两个MySQL数据库之间 (比如RDS MySQL、自建MySQL) 的双向数据实时同步, 适用于异地多活 (单元化)、数据异地容灾等多种应用场景。本文以RDS MySQL实例为例, 介绍双向数据同步的配置步骤, 其他数据源的配置流程与本案例类似。

### 前提条件

数据同步的源和目标RDS MySQL实例已存在, 如不存在请[创建RDS实例](#)。

### 注意事项

- DTS在执行全量数据初始化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。因此您需要在执行数据同步前评估源库和目标库的性能, 同时建议您在业务低峰期执行数据同步 (例如源库和目标库的CPU负载在30%以下)。
- 数据同步时, 请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更, 否则会导致同步失败。
- 如果除DTS以外的数据没有写入目标库, 您可以使用数据管理DMS (Data Management Service) 来执行在线DDL变更, 详情请参见[不锁表结构变更](#)。

**警告** 如果有除DTS外的数据写入目标库, 请勿使用DMS执行在线DDL变更, 否则可能引起目标库数据丢失。

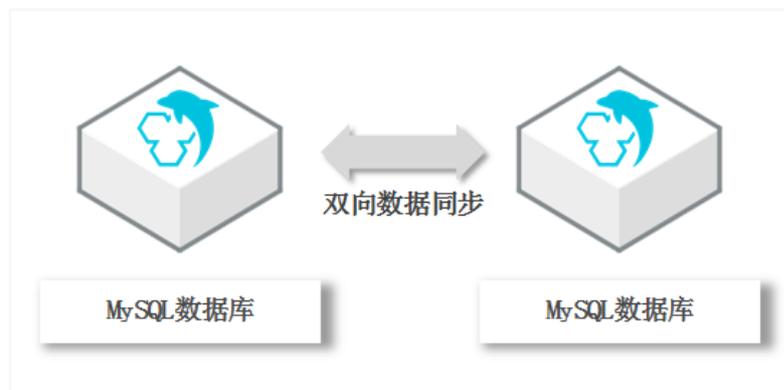
- 如双向同步任务的源实例或目标实例位于海外地域, 则仅支持同地域的双向同步, 不支持跨地域的双向同步。例如, 支持日本地域间的双向同步, 不支持日本地域与法兰克福地域间的双向同步。

### 费用说明

| 同步类型          | 链路配置费用                           |
|---------------|----------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                             |
| 增量数据同步        | 收费, 详情请参见 <a href="#">产品定价</a> 。 |

### 支持的同步架构

目前DTS仅支持两个MySQL数据库之间的双向同步, 暂不支持多个MySQL数据库之间的双向同步。



### 支持的数据源

MySQL间的双向数据同步支持以下数据源, 本文以RDS MySQL实例为数据源介绍配置流程, 其他数据源的配置流程与该案例类似。

| 同步源数据库   | 同步目的数据库  |
|--|--|
| <ul style="list-style-type: none"> <li>• RDS MySQL实例</li> <li>• ECS上的自建数据库</li> <li>• 通过专线、VPN网关或智能网关接入的自建数据库</li> <li>• 通过数据库网关接入的自建数据库</li> <li>• 通过云企业网CEN接入的自建数据库</li> </ul> | <ul style="list-style-type: none"> <li>• RDS MySQL实例</li> <li>• ECS上的自建数据库</li> <li>• 通过专线、VPN网关或智能网关接入的自建数据库</li> <li>• 通过数据库网关接入的自建数据库</li> <li>• 通过云企业网CEN接入的自建数据库</li> </ul> |

## 支持同步的SQL操作

| 操作类型 | SQL操作语句   |
|------|---|
| DML  | INSERT、UPDATE、DELETE、REPLACE  |
| DDL  | <ul style="list-style-type: none"> <li>• ALTER TABLE、ALTER VIEW</li> <li>• CREATE FUNCTION、CREATE INDEX、CREATE PROCEDURE、CREATE TABLE、CREATE VIEW</li> <li>• DROP INDEX、DROP TABLE</li> <li>• RENAME TABLE</li> <li>• TRUNCATE TABLE</li> </ul> |

## 支持的冲突检测

为保障同步数据的一致性，您需要确保同一个主键、业务主键或唯一键的记录只在双向同步的一个节点进行更新。如果同时更新则会按照您在数据同步作业中配置的冲突修复策略进行响应。

DTS通过冲突检测和修复最大程度地维护双向同步实例的稳定性。目前DTS支持进行检测的冲突类型包括：

- **INSERT导致的唯一性冲突**  
同步INSERT语句时违背了唯一性约束，例如双向同步的两个节点同时或者在极为接近的时间INSERT某个主键值相同的记录，那么同步到对端时，会因为已经存在相同主键值的记录，导致Insert同步失败。
- **UPDATE更新的记录不完全匹配**
  - UPDATE要更新的记录在同步目标实例中不存在时，DTS会自动转化为INSERT，此时可能会出现唯一键的唯一性冲突。
  - UPDATE要更新的记录出现主键或唯一键冲突。
- **DELETE对应的记录不存在**  
DELETE要删除的记录在同步的目标实例中不存在。出现这种冲突时，不论配置何种冲突修复策略，DTS都会自动忽略DELETE操作。

### 注意

- 由于数据同步两端的系统时间可能存在差异、同步存在延时等多种因素，DTS无法完全保证冲突检测机制能够完全防止数据的冲突。在使用双向同步时，您需要在业务层面配合进行相应的改造，保证同一个主键、业务主键或唯一键的记录只在双向同步的某个节点进行更新。
- 对于上述数据同步的冲突，DTS提供了修复策略，您可以在配置双向同步时选择。

## 功能限制

- **不兼容触发器**  
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，那么可能导致同步数据不一致。例如数据库中存在着两个表A和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。这种情况在同步过程中，如果源实例表A上进行了INSERT操作，则会导致表B在源实例跟目标实例数据不一致。此类情况须要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，详情请参见[源库存在触发器时如何配置同步作业](#)。
- **RENAME TABLE限制**

RENAME TABLE操作可能导致同步数据不一致。例如同步对象只包含某个表，如果同步过程中源实例对该表执行了重命名操作，那么该表的数据将不会同步到目标库。为避免该问题，您可以在数据同步配置时将该表所属的整个数据库作为同步对象。

• DDL语法同步方向限制

为保障双向同步链路的稳定性，DDL更新只能在其中一个同步方向进行同步。即一旦某个同步方向配置了DDL同步，则在反方向上不支持DDL同步，只进行DML同步。

### 操作步骤

- 1. 购买双向数据同步实例，详情请参见[购买数据同步实例](#)。

**注意** 购买时，源实例和目标实例均选择为MySQL，并选择同步拓扑为双向同步。

- 2. 登录[数据传输控制台](#)。
- 3. 在左侧导航栏，单击[数据同步](#)。
- 4. 在同步作业列表页面顶部，选择同步的目标实例所在地域。
- 5. 定位至已购买的数据同步实例，单击该实例下第一个同步作业的[配置同步链路](#)。

**注意** 一个双向数据同步实例会包含两个同步作业，需要分别进行配置。

- 6. 配置同步通道的源实例及目标实例信息。

| 类别     | 配置     | 说明   |
|--------|--------|--|
| 无      | 同步作业名称 | DTS会自动生成一个同步作业名称，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。   |
| 源实例信息  | 实例类型   | 选择RDS实例。   |
|        | 实例地区   | 购买数据同步实例时选择的源实例地域信息，不可变更。  |
|        | 实例ID   | 选择作为数据同步源的RDS实例ID。   |
|        | 数据库账号  | 填入源RDS的数据库账号。<br><br> <b>注意</b> 当源RDS实例的数据库类型为MySQL 5.5或MySQL 5.6时，没有数据库账号和数据库密码配置选项。   |
|        | 数据库密码  | 填入数据库账号对应的密码。  |
|        | 连接方式   | 根据需求选择非加密连接或SSL安全连接。如果设置为SSL安全连接，您需要提前开启RDS实例的SSL加密功能，详情请参见 <a href="#">设置SSL加密</a> 。<br><br> <b>注意</b> 目前仅中国内地及中国香港地域支持设置连接方式。   |
| 目标实例信息 | 实例类型   | 选择RDS实例。   |
|        | 实例地区   | 购买数据同步实例时选择的目标实例地域信息，不可变更。   |
|        | 实例ID   | 选择作为数据同步目标的RDS实例ID。  |
|        | 数据库账号  | 填入目标RDS的数据库账号。<br><br> <b>注意</b> 当目标RDS实例的数据库类型为MySQL 5.5或MySQL 5.6时，没有数据库账号和数据库密码配置选项。   |
|        | 数据库密码  | 填入数据库账号对应的密码。  |
|        | 连接方式   | 根据需求选择非加密连接或SSL安全连接。如果设置为SSL安全连接，您需要提前开启RDS实例的SSL加密功能，详情请参见 <a href="#">设置SSL加密</a> 。<br><br> <b>注意</b> 目前仅中国内地及中国香港地域支持设置连接方式。 |

7. 单击页面右下角的授权白名单并进入下一步。

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置同步策略及对象信息。

1.选择同步通道的源及目标实例
2.选择同步对象
3.高级设置
4.预检查

同步架构: 双向同步 (DML+DDL)

是否过滤DDL:  是  否

DML同步类型:  Insert  Delete  Update

冲突修复策略: TaskFailed(遇到冲突, 任务报错退出)

**源库对象**

若全局搜索, 请先展开树

- dtstestdata
  - Tables
  - Views

全选

已选择对象 (鼠标移到对象行, 点击编辑可修改对象名或过滤条件) 详情点我

- dtstestdata (2个对象)
  - customer
  - order

全选

\* 映射名称更改:  不进行库表名称批量更改  要进行库表名称批量更改

\* 源表DMS\_ONLINE\_DDL过程中是否复制临时表到目标库:  是  否

\* 源库、目标库无法连接后的重试时间:  分钟

取消 上一步 下一步

| 类别     | 配置      | 说明  |
|--------|---------|---|
| 同步策略配置 | 是否过滤DDL | <ul style="list-style-type: none"> <li>◦ 选择为是: 不同步DDL操作。</li> <li>◦ 选择为否: 同步DDL操作。</li> </ul> <div style="border: 1px solid #add8e6; padding: 5px; margin-top: 5px;"> <p><b>注意</b> DDL语法同步方向限制。为保障双向同步链路的稳定性, 只支持正向同步DDL, 不支持反向同步DDL。</p> </div>  |
|        | DML同步类型 | 选择需要同步的DML类型, 默认为Insert、Update、Delete, 您可以根据业务需求调整。   |
|        | 冲突修复策略  | <p>选择同步冲突的修复策略, 默认为TaskFailed, 您可以根据业务情况选择合适的冲突修复策略。</p> <ul style="list-style-type: none"> <li>◦ <b>TaskFailed</b> (遇到冲突, 任务报错退出)<br/>默认的冲突修复策略。当数据同步遇到上述冲突类型时, 同步任务直接报错并退出, 同步任务进入失败状态, 需要用户介入修复任务。</li> <li>◦ <b>Ignore</b> (遇到冲突, 直接使用目标实例中的冲突记录)<br/>当数据同步遇到上述的冲突类型时, 直接跳过当前同步语句, 继续往下执行, 选择使用目标库中的冲突记录。</li> <li>◦ <b>Overwrite</b> (遇到冲突, 直接覆盖目标实例中的冲突记录)<br/>当数据同步遇到上述的冲突类型时, 直接覆盖目标库中的冲突记录。</li> </ul> |

123

> 文档版本: 20220711

| 类别                                     | 配置          | 说明  |
|--|-------------|---|
|  | 目标已存在表的处理模式 | <ul style="list-style-type: none"> <li>◦ <b>预检查并报错拦截</b>：检查目标库中是否有同名的表。如果目标库中没有同名的表，则通过该检查项目；如果目标库中有同名的表，则在预检查阶段提示错误，数据同步作业不会被启动。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin: 5px 0;"> <p> <b>注意</b> 如果目标库中同名的表不方便删除或重命名，您可以<b>设置同步对象在目标实例中的名称</b>来避免表名冲突。</p> </div> <ul style="list-style-type: none"> <li>◦ <b>忽略报错并继续执行</b>：跳过目标库中是否有同名表的检查项。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin: 5px 0;"> <p> <b>警告</b> 选择为忽略报错并继续执行，可能导致数据不一致，给业务带来风险，例如：</p> <ul style="list-style-type: none"> <li>■ 表结构一致的情况下，如果在目标库遇到与源库主键的值相同的记录，在初始化阶段会保留目标库中的该条记录；在增量同步阶段则会覆盖目标库的该条记录。</li> <li>■ 表结构不一致的情况下，可能会导致无法初始化数据、只能同步部分列的数据或同步失败。</li> </ul> </div> |
| 选择同步对象                                 | 无           | <p>在源库对象框中单击待同步的对象（选择粒度为库或表），然后单击  将其移动到已选择对象框。</p> <div style="border: 1px solid #ccc; padding: 5px; margin: 5px 0;"> <p> <b>注意</b></p> <ul style="list-style-type: none"> <li>◦ 如果选择整个库作为同步对象，该库中所有对象的结构变更操作都会同步至目标库。</li> <li>◦ 默认情况下，同步对象的名称保持不变。如果您需要改变同步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<b>设置同步对象在目标实例中的名称</b>。</li> </ul> </div>   |
| 映射名称更改                                 | 无           | <p>如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见<b>库表列映射</b>。</p>  |
| 源表<br>DMS_ONLINE_DDL<br>过程中是否复制临时表到目标库 | 无           | <p>如源库使用<b>数据管理DMS（Data Management Service）</b>执行Online DDL变更，您可以选择是否同步Online DDL变更产生的临时表数据。</p> <ul style="list-style-type: none"> <li>◦ <b>是</b>：同步Online DDL变更产生的临时表数据。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin: 5px 0;"> <p> <b>说明</b> Online DDL变更产生的临时表数据过大，可能会导致同步任务延迟。</p> </div> <ul style="list-style-type: none"> <li>◦ <b>否</b>：不同步Online DDL变更产生的临时表数据，只同步源库的原始DDL数据。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin: 5px 0;"> <p> <b>说明</b> 该方案会导致目标库锁表。</p> </div>  |
| 源、目标库无法连接重试时间                          | 无           | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <div style="border: 1px solid #ccc; padding: 5px; margin: 5px 0;"> <p> <b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p> </div>   |

- 9. 上述配置完成后，单击页面右下角的下一步。
- 10. 配置同步初始化的高级配置信息。



此步骤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**和**全量数据初始化**。

**注意** 如果同步对象中有部分表包含在另外一个同步方向的同步对象中，那么这部分表不会进行同步初始化。

- 11. 上述配置完成后，单击页面右下角的预检查并启动。

**注意**

- 在数据同步任务正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动数据同步任务。
- 如果预检查失败，单击具体检查项后的，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择**确认屏蔽**、**忽略告警项**并重新进行预检查，跳过告警检测项重新进行预检查。

- 12. 在**预检查**对话框中显示**预检查通过**后，关闭**预检查**对话框，同步作业将正式开始。
- 13. 等待该同步作业的链路初始化完成，直至状态处于**同步中**。  
您可以在**数据同步**页面，查看数据同步状态。
- 14. 定位至第二个同步作业，单击**配置同步链路**，配置流程参见步骤5到步骤12。

| 实例ID/作业名称 | 状态  | 同步概况                                | 付费方式     | 同步架构(全部)  | 操作                |
|-----------|-----|-------------------------------------|----------|-----------|-------------------|
| [实例ID]    | --  | --                                  | 按量付费     | 双向同步      | 转包年包月   升级   查看更多 |
| 作业名称      | 状态  | 同步概况                                | 源实例/目标实例 | 操作        |                   |
| MySQL正向同步 | 同步中 | 延时: 0 毫秒<br>速度: 0.00RPS/(0.000MB/s) | [实例ID]   | 暂停同步   更多 |                   |
| MySQL反向同步 | 未配置 |                                     | [实例ID]   | 配置同步链路    |                   |

- 15. 第二个同步作业配置完成后，等待两个同步作业的链路状态均处于**同步中**，即完成双向数据同步的配置流程。

| 实例ID/作业名称 | 状态  | 同步概况                                | 付费方式     | 同步架构(全部)  | 操作                |
|-----------|-----|-------------------------------------|----------|-----------|-------------------|
| [实例ID]    | --  | --                                  | 按量付费     | 双向同步      | 转包年包月   升级   查看更多 |
| 作业名称      | 状态  | 同步概况                                | 源实例/目标实例 | 操作        |                   |
| MySQL正向同步 | 同步中 | 延时: 0 毫秒<br>速度: 0.00RPS/(0.000MB/s) | [实例ID]   | 暂停同步   更多 |                   |
| MySQL反向同步 | 同步中 | 延时: 0 毫秒<br>速度: 0.00RPS/(0.000MB/s) | [实例ID]   | 暂停同步   更多 |                   |

## 6.2. RDS MySQL实例间的单向同步

数据传输服务DTS（Data Transmission Service）支持两个MySQL数据库之间的数据同步，本文以RDS MySQL实例之间的单向同步为例，介绍配置流程。

### 前提条件

- 数据同步的源RDS实例和目标RDS实例已存在，详情请参见[创建RDS实例](#)。
- 源RDS实例和目标RDS实例的数据库类型为MySQL。

### 注意事项

- DTS在执行全量数据初始化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。因此您需要在执行数据同步前评估源库和目标库的性能，同时建议您在业务低峰期执行数据同步（例如源库和目标库的CPU负载在30%以下）。
- 数据同步时，请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。
- 如果除DTS以外的数据没有写入目标库，您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，详情请参见[不锁表结构变更](#)。

 **警告** 如果有除DTS外的数据写入目标库，请勿使用DMS执行在线DDL变更，否则可能引起目标库数据丢失。

- 如果源库中待迁移的表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。
- 全量初始化过程中，并发INSERT会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。

### 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

### 支持的同步架构

- 一对一单向同步
- 一对多单向同步
- 级联单向同步
- 多对一单向同步
- 一对一双向同步

关于各类同步架构的介绍及注意事项，请参见[数据同步拓扑介绍](#)。

### 支持同步的SQL操作

| 操作类型 | SQL操作语句   |
|------|---|
| DML  | INSERT、UPDATE、DELETE、REPLACE  |
| DDL  | <ul style="list-style-type: none"> <li>• ALTER TABLE、ALTER VIEW</li> <li>• CREATE FUNCTION、CREATE INDEX、CREATE PROCEDURE、CREATE TABLE、CREATE VIEW</li> <li>• DROP INDEX、DROP TABLE</li> <li>• RENAME TABLE</li> <li>• TRUNCATE TABLE</li> </ul> |

### 功能限制

- 不兼容触发器  
当同步对象为整个库，且库中的触发器（TRIGGER）会更新库内某个表时，可能导致源和目标库的数据不一致。相关解决方案请参见[源库存在触发器时如何配置同步作业](#)。
- RENAME TABLE限制

RENAME TABLE操作可能导致同步数据不一致。例如同步对象只包含表A，如果同步过程中源实例将表A重命名为表B，那么表B将不会被同步到目标库。为避免该问题，您可以在数据同步配置时，选择同步表A和表B所在的整个数据库作为同步对象。

### 操作步骤

1. 购买数据同步实例，详情请参见[购买数据同步实例](#)。

? **说明** 购买时，选择源实例和目标实例均选择为MySQL，并选择同步拓扑为单向同步。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击[数据同步](#)。
4. 在同步作业列表页面顶部，选择数据同步实例所属地域。
5. 定位至已购买的数据同步实例，单击[配置同步链路](#)。
6. 配置同步通道的源实例及目标实例信息。

1. 选择同步通道的源及目标实例
2. 选择同步对象
3. 高级设置
4. 预检查

同步作业名称：

**源实例信息**

实例类型：

实例地区：

\* 实例ID： [其他阿里云账号下的RDS实例](#)

\* 数据库账号：

\* 数据库密码：

\* 连接方式： 非加密连接  SSL安全连接

**目标实例信息**

实例类型：

实例地区：

\* 实例ID：

\* 数据库账号：

\* 数据库密码：

\* 连接方式： 非加密连接  SSL安全连接

| 类别    | 配置     | 说明   |
|-------|--------|--|
| 无     | 同步作业名称 | DTS会自动生成一个同步作业名称，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。   |
| 源实例信息 | 实例类型   | 选择RDS实例。   |
|       | 实例地区   | 购买数据同步实例时选择的源实例地域信息，不可变更。  |
|       | 实例ID   | 选择作为数据同步源的RDS实例ID。   |
|       | 数据库账号  | 填入源RDS的数据库账号。<br><div style="background-color: #e6f2ff; padding: 5px; margin-top: 5px;"> <span style="font-size: 1.2em;">?</span> <b>说明</b> 当源RDS实例的数据库类型为MySQL 5.5或MySQL 5.6时，无需配置数据库账号和数据库密码。         </div> |

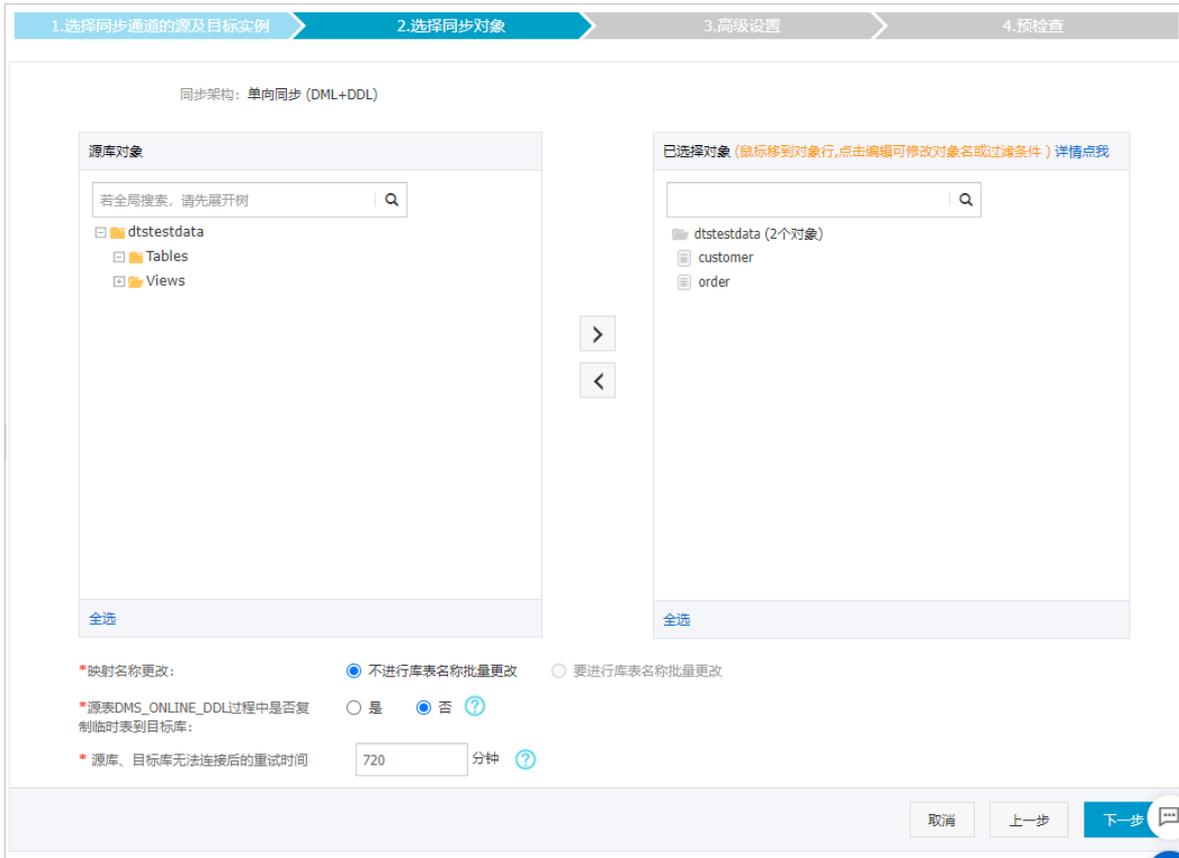
| 类别     | 配置    | 说明  |
|--------|-------|---|
|        | 数据库密码 | 填入数据库账号对应的密码。   |
|        | 连接方式  | <p>根据需求选择非加密连接或SSL安全连接。如果设置为SSL安全连接，您需要提前开启RDS实例的SSL加密功能，详情请参见<a href="#">设置SSL加密</a>。</p> <p> <b>注意</b> 目前仅中国内地及中国香港地域支持设置连接方式。</p>   |
| 目标实例信息 | 实例类型  | 选择RDS实例。  |
|        | 实例地区  | 购买数据同步实例时选择的目标实例地域信息，不可变更。  |
|        | 实例ID  | 选择作为数据同步目标的RDS实例ID。   |
|        | 数据库账号 | <p>填入目标RDS的数据库账号。</p> <p> <b>说明</b> 当目标RDS实例的数据库类型为MySQL 5.5或MySQL 5.6时，无需配置数据库账号和数据库密码。</p>   |
|        | 数据库密码 | 填入数据库账号对应的密码。   |
|        | 连接方式  | <p>根据需求选择非加密连接或SSL安全连接。如果设置为SSL安全连接，您需要提前开启RDS实例的SSL加密功能，详情请参见<a href="#">设置SSL加密</a>。</p> <p> <b>注意</b> 目前仅中国内地及中国香港地域支持设置连接方式。</p> |

7. 单击页面右下角的授权白名单并进入下一步。

 **说明**

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置同步策略及对象信息。



| 配置项目                           | 配置说明   |
|--------------------------------|--|
| 选择同步对象                         | <p>在源库对象框中单击待同步的对象，然后单击  图标将其移动至已选择对象框。</p> <p>同步对象的选择粒度为库、表。</p> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>如果选择整个库作为同步对象，该库中所有对象的结构变更操作都会同步至目标库。</li> <li>默认情况下，同步对象在目标库中的名称与源库保持一致。如果您需要改变同步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 <a href="#">设置同步对象在目标实例中的名称</a>。</li> </ul> </div>  |
| 映射名称更改                         | <p>如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见 <a href="#">库表映射</a>。</p>  |
| 源表DMS_ONLINE_DDL过程中是否复制临时表到目标库 | <p>如源库使用 <a href="#">数据管理DMS (Data Management Service)</a> 执行Online DDL变更，您可以选择是否同步Online DDL变更产生的临时表数据。</p> <ul style="list-style-type: none"> <li>是：同步Online DDL变更产生的临时表数据。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p><b>说明</b> Online DDL变更产生的临时表数据过大，可能会导致同步任务延迟。</p> </div> <ul style="list-style-type: none"> <li>否：不同步Online DDL变更产生的临时表数据，只同步源库的原始DDL数据。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p><b>说明</b> 该方案会导致目标库锁表。</p> </div> |

| 配置项目          | 配置说明  |
|---------------|---|
| 源、目标库无法连接重试时间 | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <p><b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p> |

9. 上述配置完成后，单击页面右下角的下一步。

10. 配置同步初始化的高级配置信息。



**说明**

- 此步骤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，需要选中结构初始化和全量数据初始化。

11. 上述配置完成后，单击页面右下角的预检查并启动。

**说明**

- 在同步作业正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步作业。
- 如果预检查失败，单击具体检查项后的 ，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择**确认屏蔽**、**忽略告警项**并重新进行预检查，跳过告警检测项重新进行预检查。

12. 在预检查对话框中显示预检查通过后，关闭预检查对话框，同步作业将正式开始。

13. 等待同步作业的链路初始化完成，直至处于同步中状态。

您可以在数据同步页面，查看数据同步作业的状态。



## 6.3. 从ECS上的自建MySQL同步至RDS MySQL

数据传输服务DTS（Data Transmission Service）支持ECS上的自建MySQL同步至RDS MySQL实例，实现增量数据的实时同步。

## 前提条件

- 自建MySQL数据库版本为5.1、5.5、5.6、5.7或8.0版本。
- 数据同步的目标RDS实例已存在，RDS实例创建方法请参见[创建RDS实例](#)。

## 注意事项

- DTS在执行全量数据初始化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。因此您需要在执行数据同步前评估源库和目标库的性能，同时建议您在业务低峰期执行数据同步（例如源库和目标库的CPU负载在30%以下）。
- 数据同步时，请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。
- 如果除DTS以外的数据没有写入目标库，您可以使用数据管理DMS（Data Management Service）来执行在线DDL变更，详情请参见[不锁表结构变更](#)。

 **警告** 如果有除DTS外的数据写入目标库，请勿使用DMS执行在线DDL变更，否则可能引起目标库数据丢失。

- 如果源库中待迁移的表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。
- 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。

## 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

## 支持的同步架构

- 一对一单向同步
- 一对多单向同步
- 多对一单向同步
- 级联单向同步
- 一对一双向同步

 **说明** 如需实现双向同步，请参见[MySQL实例间的双向同步](#)。

## 支持同步的SQL操作

| 操作类型 | SQL操作语句   |
|------|---|
| DML  | INSERT、UPDATE、DELETE、REPLACE  |
| DDL  | <ul style="list-style-type: none"> <li>• ALTER TABLE、ALTER VIEW</li> <li>• CREATE FUNCTION、CREATE INDEX、CREATE PROCEDURE、CREATE TABLE、CREATE VIEW</li> <li>• DROP INDEX、DROP TABLE</li> <li>• RENAME TABLE</li> <li>• TRUNCATE TABLE</li> </ul> |

## 功能限制

- 不兼容触发器

当同步对象为整个库，且库中的触发器（TRIGGER）会更新库内某个表时，可能导致源和目标库的数据不一致。相关解决方案请参见[源库存在触发器时如何配置同步作业](#)。

- **RENAME TABLE限制**  
RENAME TABLE操作可能导致同步数据不一致。例如同步对象只包含表A，如果同步过程中源实例将表A重命名为表B，那么表B将不会被同步到目标库。为避免该问题，您可以在数据同步配置时，选择同步表A和表B所在的整个数据库作为同步对象。

## 准备工作

在正式配置数据同步作业之前，您需要[为自建MySQL创建账号并设置binlog](#)。

## 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

 **说明** 购买时，选择源实例和目标实例均为MySQL，并选择同步拓扑为单向同步。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击[数据同步](#)。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击该实例的[配置同步链路](#)。

| 同步作业名称 | 实例ID/作业名称         | 状态  | 同步概况 | 付费方式 | 同步架构(全部) | 操作   |
|--------|-------------------|-----|------|------|----------|--|
|        | RDS for MySQL单向同步 | 未配置 |      | 按量付费 | 单向同步     | <a href="#">配置同步链路</a> <a href="#">转包年包月</a> <a href="#">升级</a> <a href="#">更多</a> |

6. 配置同步通道的源实例及目标实例信息。

1. 选择同步通道的源及目标实例 > 2. 选择同步对象 > 3. 高级设置

同步作业名称: RDS for MySQL单向同步

### 源实例信息

实例类型: ECS上的自建数据库

实例地区: 华东1 (杭州)

\* ECS实例ID: i-bp

数据库类型: MySQL

\* 端口: 3306

\* 数据库账号:

\* 数据库密码:

### 目标实例信息

实例类型: RDS实例

实例地区: 华东1 (杭州)

\* 实例ID: m-

\* 数据库账号:

\* 数据库密码:

\* 连接方式:  非加密连接  SSL安全连接

[取消](#) [授权白名单并进入下一步](#)

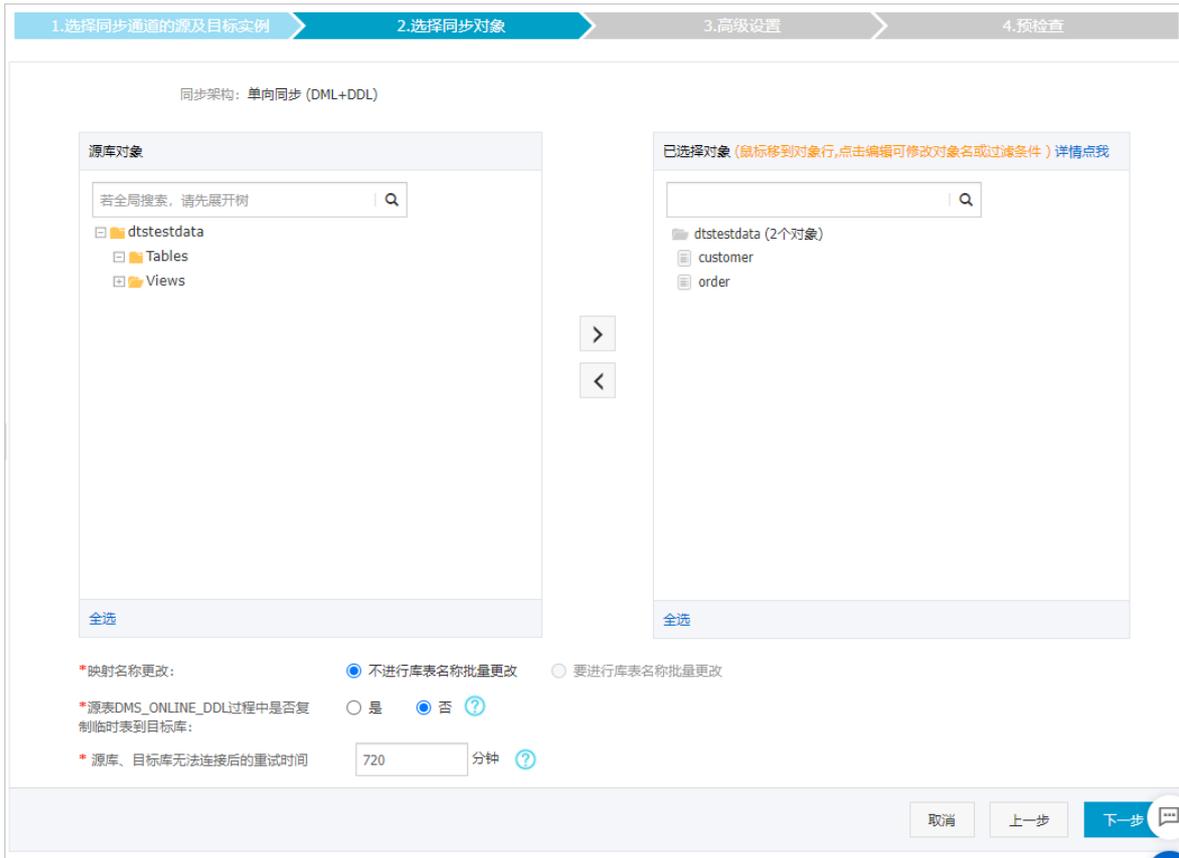
| 类别     | 配置      | 说明   |
|--------|---------|--|
| 无      | 同步作业名称  | DTS会自动生成一个同步作业名称，建议配置具有业务意义的名称（无唯一性要求），便于后续识别。   |
| 源实例信息  | 实例类型    | 选择ECS上的自建数据库。  |
|        | 实例地区    | 购买数据同步实例时选择的源实例地域信息，不可变更。  |
|        | ECS实例ID | 选择作为同步数据源的ECS实例ID。   |
|        | 数据库类型   | 购买数据同步实例时选择的数据库类型： <b>MySQL</b> ，不可变更。   |
|        | 端口      | 填入自建MySQL数据库的服务端口，默认为 <b>3306</b> 。  |
|        | 数据库账号   | 填入自建MySQL数据库的账号，需具备REPLICATION CLIENT、REPLICATION SLAVE、SHOW VIEW和所有同步对象的SELECT权限。   |
|        | 数据库密码   | 填入自建MySQL数据库账号的密码。   |
| 目标实例信息 | 实例类型    | 选择RDS实例。   |
|        | 实例地区    | 购买数据同步实例时选择的目标实例地域信息，不可变更。   |
|        | 实例ID    | 选择作为数据同步目标的RDS实例ID。  |
|        | 数据库账号   | 填入目标RDS的数据库账号。<br><br><div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px;"> <p> <b>说明</b> 当目标RDS实例的数据库类型为MySQL5.5或MySQL5.6时，无需配置数据库账号和数据库密码。</p> </div>  |
|        | 数据库密码   | 填入数据库账号的密码。  |
|        | 连接方式    | 根据需求选择 <b>非加密连接</b> 或 <b>SSL安全连接</b> 。如果设置为 <b>SSL安全连接</b> ，您需要提前开启RDS实例的SSL加密功能，详情请参见 <a href="#">设置SSL加密</a> 。<br><br><div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px;"> <p> <b>注意</b> 目前仅中国内地及中国香港地域支持设置连接方式。</p> </div> |

7. 单击页面右下角的**授权白名单**并进入下一步。

 **说明**

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置同步策略及对象信息。



| 配置项目                           | 配置说明   |
|--------------------------------|--|
| 选择同步对象                         | <p>在源库对象框中单击待同步的对象，然后单击  图标将其移动至已选择对象框。</p> <p>同步对象的选择粒度为库、表。</p> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>如果选择整个库作为同步对象，该库中所有对象的结构变更操作都会同步至目标库。</li> <li>默认情况下，同步对象在目标库中的名称与源库保持一致。如果您需要改变同步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 <a href="#">设置同步对象在目标实例中的名称</a>。</li> </ul> </div>  |
| 映射名称更改                         | <p>如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见 <a href="#">库表映射</a>。</p>  |
| 源表DMS_ONLINE_DDL过程中是否复制临时表到目标库 | <p>如源库使用 <a href="#">数据管理DMS (Data Management Service)</a> 执行Online DDL变更，您可以选择是否同步Online DDL变更产生的临时表数据。</p> <ul style="list-style-type: none"> <li><b>是</b>：同步Online DDL变更产生的临时表数据。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p><b>说明</b> Online DDL变更产生的临时表数据过大，可能会导致同步任务延迟。</p> </div> <ul style="list-style-type: none"> <li><b>否</b>：不同步Online DDL变更产生的临时表数据，只同步源库的原始DDL数据。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p><b>说明</b> 该方案会导致目标库锁表。</p> </div> |

| 配置项目          | 配置说明  |
|---------------|---|
| 源、目标库无法连接重试时间 | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <p><b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p> |

9. 上述配置完成后，单击页面右下角的下一步。

10. 配置同步初始化的高级配置信息。



- 此步骤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和全量数据初始化。

11. 上述配置完成后，单击页面右下角的预检查并启动。

**注意**

- 在数据同步任务正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动数据同步任务。
- 如果预检查失败，单击具体检查项后的 ，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择确认屏蔽、忽略告警项并重新进行预检查，跳过告警检测项重新进行预检查。

12. 在预检查对话框中显示预检查通过后，关闭预检查对话框，同步作业将正式开始。

13. 等待该同步作业的链路初始化完成，直至状态处于同步中。

您可以在数据同步页面，查看数据同步状态。



## 6.4. 从通过专线、VPN网关或智能接入网关接入的自建MySQL同步至RDS MySQL

数据传输服务DTS（Data Transmission Service）支持将通过专线、VPN网关或智能接入网关接入的自建MySQL同步至RDS MySQL，实现增量数据的实时同步。

### 前提条件

- 数据同步的目标RDS实例已存在，如不存在请[创建RDS实例](#)。
- 自建MySQL数据库版本为5.1、5.5、5.6、5.7或8.0版本。
- 已经将自建MySQL数据库接入至阿里云专有网络，详情请参见[通过CEN实现本地IDC与DTS云服务互通](#)。

 **说明** 接入至阿里云专有网络后，还需要放通DTS的IP地址访问自建数据库所属的网络，详情请参见[通过VPN网关实现本地IDC与DTS云服务互通](#)。

## 注意事项

- DTS在执行全量数据初始化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。因此您需要在执行数据同步前评估源库和目标库的性能，同时建议您在业务低峰期执行数据同步（例如源库和目标库的CPU负载在30%以下）。
- 数据同步时，请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。
- 如果源库中待迁移的表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。
- 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。

## 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

## 支持的同步架构

- 一对一单向同步
- 一对多单向同步
- 级联单向同步
- 多对一单向同步
- 一对一双向同步

关于各类同步架构的介绍及注意事项，请参见[数据同步拓扑介绍](#)。

## 支持同步的SQL操作

| 操作类型 | SQL操作语句   |
|------|---|
| DML  | INSERT、UPDATE、DELETE、REPLACE  |
| DDL  | <ul style="list-style-type: none"> <li>● ALTER TABLE、ALTER VIEW</li> <li>● CREATE FUNCTION、CREATE INDEX、CREATE PROCEDURE、CREATE TABLE、CREATE VIEW</li> <li>● DROP INDEX、DROP TABLE</li> <li>● RENAME TABLE</li> <li>● TRUNCATE TABLE</li> </ul> |

## 功能限制

- 不兼容触发器  
当同步对象为整个库，且库中的触发器（TRIGGER）会更新库内某个表时，可能导致源和目标库的数据不一致。相关解决方案请参见[源库存在触发器时如何配置同步作业](#)。
- RENAME TABLE限制  
RENAME TABLE操作可能导致同步数据不一致。例如同步对象只包含表A，如果同步过程中源实例将表A重命名为表B，那么表B将不会被同步到目标库。为避免该问题，您可以在数据同步配置时，选择同步表A和表B所在的整个数据库作为同步对象。

### 准备工作

在正式配置数据同步作业之前，您需要为自建MySQL创建账号并设置binlog。

### 操作步骤

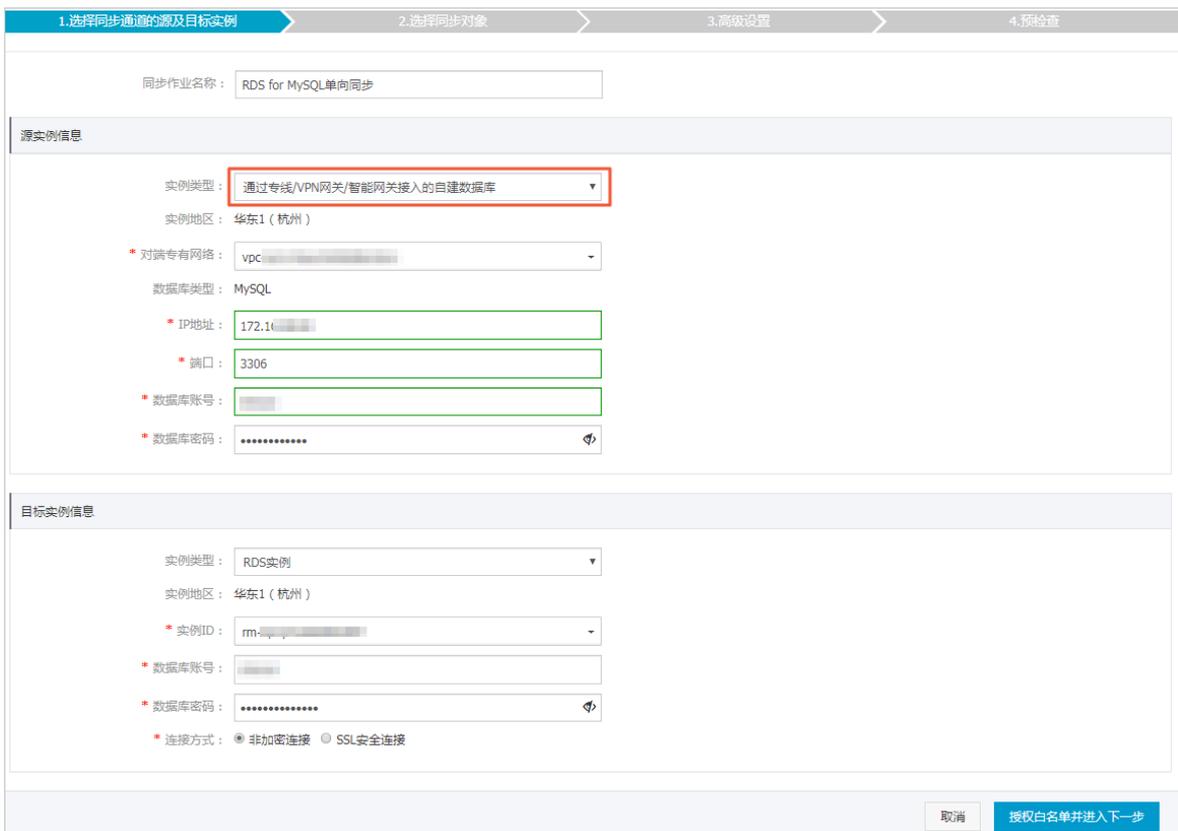
1. 购买数据同步作业，详情请参见[购买流程](#)。

? 说明 购买时，选择源实例和目标实例均为MySQL，并选择同步拓扑为单向同步。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击数据同步。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击该实例的[配置同步链路](#)。



6. 配置同步通道的源实例及目标实例信息。



| 类别 | 配置     | 说明   |
|----|--------|--|
| 无  | 同步作业名称 | DTS会自动生成一个同步作业名称，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。 |
|    | 实例类型   | 选择通过专线/VPN网关/智能接入网关接入的自建数据库。                   |

| 类别     | 配置     | 说明   |
|--------|--------|--|
| 源实例信息  | 实例地区   | 购买数据同步实例时选择的源实例地域信息，不可变更。  |
|        | 对端专有网络 | 选择自建数据库接入的VPC ID。  |
|        | 数据库类型  | 购买数据同步实例时选择的数据库类型：MySQL，不可变更。  |
|        | IP地址   | 填入自建MySQL数据库的服务器IP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，需要具备REPLICATION CLIENT、REPLICATION SLAVE、SHOW VIEW和所有同步对象的SELECT权限。  |
|        | 数据库密码  | 填入该数据库账号对应的密码。   |
| 目标实例信息 | 实例类型   | 选择RDS实例。   |
|        | 实例地区   | 购买数据同步实例时选择的目标实例地域信息，不可变更。   |
|        | 实例ID   | 选择作为数据同步目标的RDS实例ID。  |
|        | 数据库账号  | 填入目标RDS的数据库账号。<br><br><div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px;"> <p> 说明 当目标RDS实例的数据库类型为MySQL 5.5或MySQL 5.6时，无需配置数据库账号和数据库密码。</p> </div>   |
|        | 数据库密码  | 填入该数据库账号对应的密码。   |
|        | 连接方式   | 根据需求选择非加密连接或SSL安全连接。如果设置为SSL安全连接，您需要提前开启RDS实例的SSL加密功能，详情请参见 <a href="#">设置SSL加密</a> 。<br><br><div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px;"> <p> 注意 目前仅中国内地及中国香港地域支持设置连接方式。</p> </div> |

#### 7. 单击页面右下角的授权白名单并进入下一步。

##### 说明

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

#### 8. 配置同步策略及对象信息。

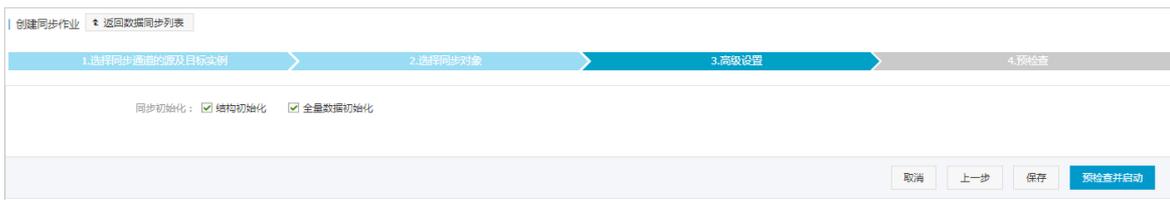


| 配置项目                           | 配置说明   |
|--------------------------------|--|
| 选择同步对象                         | <p>在源库对象框中单击待同步的对象，然后单击  图标将其移动至已选择对象框。</p> <p>同步对象的选择粒度为库、表。</p> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>如果选择整个库作为同步对象，该库中所有对象的结构变更操作都会同步至目标库。</li> <li>默认情况下，同步对象在目标库中的名称与源库保持一致。如果您需要改变同步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 <a href="#">设置同步对象在目标实例中的名称</a>。</li> </ul> </div>  |
| 映射名称更改                         | <p>如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见 <a href="#">库表映射</a>。</p>  |
| 源表DMS_ONLINE_DDL过程中是否复制临时表到目标库 | <p>如源库使用 <a href="#">数据管理DMS (Data Management Service)</a> 执行Online DDL变更，您可以选择是否同步Online DDL变更产生的临时表数据。</p> <ul style="list-style-type: none"> <li>是：同步Online DDL变更产生的临时表数据。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p><b>说明</b> Online DDL变更产生的临时表数据过大，可能会导致同步任务延迟。</p> </div> <ul style="list-style-type: none"> <li>否：不同步Online DDL变更产生的临时表数据，只同步源库的原始DDL数据。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p><b>说明</b> 该方案会导致目标库锁表。</p> </div> |

| 配置项目          | 配置说明  |
|---------------|---|
| 源、目标库无法连接重试时间 | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <p><b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p> |

9. 上述配置完成后，单击页面右下角的下一步。

10. 配置同步初始化的高级配置信息。



- 此步骤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和全量数据初始化。

11. 上述配置完成后，单击页面右下角的预检查并启动。

**注意**

- 在数据同步任务正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动数据同步任务。
- 如果预检查失败，单击具体检查项后的 ，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择确认屏蔽、忽略告警项并重新进行预检查，跳过告警检测项重新进行预检查。

12. 在预检查对话框中显示预检查通过后，关闭预检查对话框，同步作业将正式开始。

13. 等待该同步作业的链路初始化完成，直至状态处于同步中。

您可以在数据同步页面，查看数据同步状态。



## 6.5. 从RDS MySQL同步至通过专线、VPN网关或智能接入网关接入的自建MySQL

数据传输服务DTS（Data Transmission Service）支持将RDS MySQL同步至通过专线、VPN网关或智能接入网关接入的自建MySQL，实现增量数据的实时同步。

### 前提条件

- 自建MySQL数据库版本为5.1、5.5、5.6、5.7或8.0版本。

**说明** 建议与RDS MySQL的数据库版本一致。

- 已经将自建MySQL数据库通过专线、VPN网关或智能接入网关接入至阿里云专有网络，详情请参见[本地IDC接入至阿里云方案概览](#)。

 **说明** 接入至阿里云专有网络后，还需要放通DTS的IP地址访问自建数据库所属的网络，详情请参见[通过VPN网关实现本地IDC与DTS云服务互通](#)。

### 注意事项

- DTS在执行全量数据初始化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。因此您需要在执行数据同步前评估源库和目标库的性能，同时建议您在业务低峰期执行数据同步（例如源库和目标库的CPU负载在30%以下）。
- 全量初始化过程中，并发写入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。

### 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

### 同步限制

如果源库中待迁移的表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。

### 支持的同步架构

- 一对一单向同步
- 一对多单向同步
- 级联单向同步
- 多对一单向同步
- 一对一双向同步

关于各类同步架构的介绍及注意事项，请参见[数据同步拓扑介绍](#)。

### 支持同步的SQL操作

| 操作类型 | SQL操作语句   |
|------|---|
| DML  | INSERT、UPDATE、DELETE、REPLACE  |
| DDL  | <ul style="list-style-type: none"> <li>● ALTER TABLE、ALTER VIEW</li> <li>● CREATE FUNCTION、CREATE INDEX、CREATE PROCEDURE、CREATE TABLE、CREATE VIEW</li> <li>● DROP INDEX、DROP TABLE</li> <li>● RENAME TABLE</li> <li>● TRUNCATE TABLE</li> </ul> |

### 功能限制

- 不兼容触发器  
当同步对象为整个库，且库中的触发器（TRIGGER）会更新库内某个表时，可能导致源和目标库的数据不一致。相关解决方案请参见[源库存在触发器时如何配置同步作业](#)。
- RENAME TABLE限制  
RENAME TABLE操作可能导致同步数据不一致。例如同步对象只包含表A，如果同步过程中源实例将表A重命名为表B，那么表B将不会被同步到目标库。为避免该问题，您可以在数据同步配置时，选择同步表A和表B所在的整个数据库作为同步对象。

## 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

? **说明** 购买时，选择源实例和目标实例均为MySQL，并选择同步拓扑为单向同步。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击[数据同步](#)。
4. 在同步作业列表页面顶部，选择数据同步实例所属地域。
5. 定位至已购买的数据同步实例，单击[配置同步链路](#)。
6. 配置同步作业的源和目标实例信息。

1.选择同步通道的源及目标实例
4.预检查

同步作业名称:

**源实例信息**

实例类型:

实例地区: 华东1 (杭州)

\* 实例ID:  [其他阿里云账号下的RDS实例](#)

\* 数据库账号:

\* 数据库密码:

\* 连接方式:  非加密连接  SSL安全连接

**目标实例信息**

实例类型:

实例地区: 华东1 (杭州)

\* 对端专有网络:

数据库类型:

\* IP地址:

\* 端口:

\* 数据库账号:

\* 数据库密码:

| 类别    | 配置     | 配置说明   |
|-------|--------|--|
| 无     | 同步作业名称 | DTS会自动生成一个同步作业名称，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。   |
| 源实例信息 | 实例类型   | 选择RDS实例。   |
|       | 实例地区   | 购买数据同步实例时选择的源实例地域信息，不可变更。  |
|       | 实例ID   | 选择源RDS实例ID。  |
|       | 数据库账号  | 填入RDS的数据库账号。<br><div style="background-color: #e1f5fe; padding: 5px; margin-top: 5px;"><span style="color: #0070c0;">?</span> <b>说明</b> 当RDS实例的数据库类型为MySQL 5.5或MySQL 5.6时，无需配置数据库账号和数据库密码。</div> |
|       | 数据库密码  | 填入该数据库账号对应的密码。   |

| 类别     | 配置             | 配置说明  |
|--------|----------------|---|
|        | 连接方式           | <p>根据需求选择非加密连接或SSL安全连接。如果设置为SSL安全连接，您需要提前开启RDS实例的SSL加密功能，详情请参见<a href="#">设置SSL加密</a>。</p> <p> <b>注意</b> 目前仅中国内地及中国香港地域支持设置连接方式。</p> |
| 目标实例信息 | 实例类型           | 选择通过专线/VPN网关/智能接入网关接入的自建数据库。  |
|        | 实例地区           | 购买数据同步实例时选择的目标实例地域信息，不可变更。  |
|        | 对端专有网络         | 选择自建数据库接入的VPC ID。   |
|        | 数据库类型          | 固定为MySQL，不可变更。  |
|        | IP地址           | 填入自建MySQL数据库的服务器IP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          | <p>填入自建MySQL的数据库账号。</p> <p> <b>说明</b> 该账号需具备REPLICATION CLIENT、REPLICATION SLAVE、SHOW VIEW和所有同步对象的SELECT权限。</p>                      |
| 数据库密码  | 填入该数据库账号对应的密码。 |   |

7. 单击页面右下角的授权白名单并进入下一步。

 **说明**

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置同步策略及对象信息。

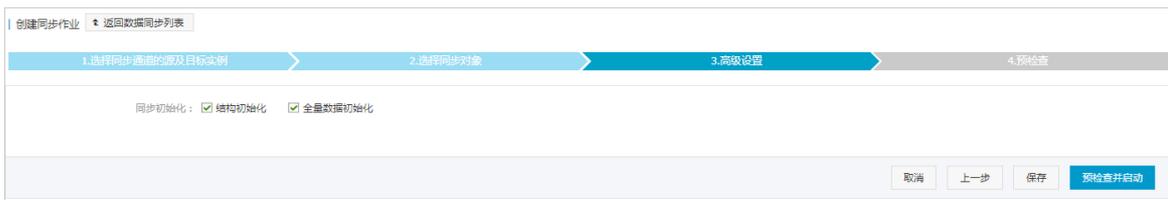


| 配置项目                           | 配置说明   |
|--------------------------------|--|
| 选择同步对象                         | <p>在源库对象框中单击待同步的对象，然后单击  图标将其移动至已选择对象框。</p> <p>同步对象的选择粒度为库、表。</p> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>如果选择整个库作为同步对象，该库中所有对象的结构变更操作都会同步至目标库。</li> <li>默认情况下，同步对象在目标库中的名称与源库保持一致。如果您需要改变同步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 <a href="#">设置同步对象在目标实例中的名称</a>。</li> </ul>   |
| 映射名称更改                         | <p>如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见 <a href="#">库表映射</a>。</p>  |
| 源表DMS_ONLINE_DDL过程中是否复制临时表到目标库 | <p>如源库使用 <a href="#">数据管理DMS (Data Management Service)</a> 执行Online DDL变更，您可以选择是否同步Online DDL变更产生的临时表数据。</p> <ul style="list-style-type: none"> <li>是：同步Online DDL变更产生的临时表数据。</li> </ul> <p><b>说明</b> Online DDL变更产生的临时表数据过大，可能会导致同步任务延迟。</p> <ul style="list-style-type: none"> <li>否：不同步Online DDL变更产生的临时表数据，只同步源库的原始DDL数据。</li> </ul> <p><b>说明</b> 该方案会导致目标库锁表。</p> |

| 配置项目          | 配置说明  |
|---------------|---|
| 源、目标库无法连接重试时间 | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <p><b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p> |

9. 上述配置完成后，单击页面右下角的下一步。

10. 配置同步初始化的高级配置信息。



**说明**

- 此步骤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，需要选中结构初始化和全量数据初始化。

11. 上述配置完成后，单击页面右下角的预检查并启动。

**说明**

- 在同步作业正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步作业。
- 如果预检查失败，单击具体检查项后的 ，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择确认屏蔽、忽略告警项并重新进行预检查，跳过告警检测项重新进行预检查。

12. 在预检查对话框中显示预检查通过后，关闭预检查对话框，同步作业将正式开始。

13. 等待同步作业的链路初始化完成，直至处于同步中状态。

您可以在数据同步页面，查看数据同步作业的状态。



## 6.6. 不同阿里云账号下RDS MySQL实例间的数据同步

数据传输服务DTS（Data Transmission Service）支持对不同阿里云账号下的RDS MySQL实例配置数据同步，实现跨阿里云账号的数据同步。

## 前提条件

- 数据同步的源RDS实例和目标RDS实例已存在，如不存在请[创建RDS实例](#)。
- 源RDS实例和目标RDS实例的数据库类型为MySQL。
- RDS MySQL实例必须具备内网地址。

## 注意事项

- DTS在执行全量数据初始化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。因此您需要在执行数据同步前评估源库和目标库的性能，同时建议您在业务低峰期执行数据同步（例如源库和目标库的CPU负载在30%以下）。
- 如果源库中待迁移的表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。
- 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。

## 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

## 支持的同步架构

- 一对一单向同步
- 一对多单向同步
- 级联单向同步
- 多对一单向同步

关于各类同步架构的介绍及注意事项，请参见[数据同步拓扑介绍](#)。

## 支持同步的SQL操作

| 操作类型 | SQL操作语句   |
|------|---|
| DML  | INSERT、UPDATE、DELETE、REPLACE  |
| DDL  | <ul style="list-style-type: none"> <li>• ALTER TABLE、ALTER VIEW</li> <li>• CREATE FUNCTION、CREATE INDEX、CREATE PROCEDURE、CREATE TABLE、CREATE VIEW</li> <li>• DROP INDEX、DROP TABLE</li> <li>• RENAME TABLE</li> <li>• TRUNCATE TABLE</li> </ul> |

## 功能限制

- 不兼容触发器  
当同步对象为整个库，且库中的触发器（TRIGGER）会更新库内某个表时，可能导致源和目标库的数据不一致。相关解决方案请参见[源库存在触发器时如何配置同步作业](#)。
- RENAME TABLE限制  
RENAME TABLE操作可能导致同步数据不一致。例如同步对象只包含表A，如果同步过程中源实例将表A重命名为表B，那么表B将不会被同步到目标库。为避免该问题，您可以在数据同步配置时，选择同步表A和表B所在的整个数据库作为同步对象。

## 准备工作

将目标RDS实例所属的云账号设置为授信云账号，允许通过数据传输服务访问源RDS实例所属云账号的相关云资源。详情请参见[跨阿里云账号数据迁移或同步时如何配置RAM授权](#)。

**说明** 请使用源实例所属的阿里云账号登录RAM控制台执行授权操作，授权操作完成后，请在目标实例所属的阿里云账号中创建数据迁移或同步任务。

### 操作步骤

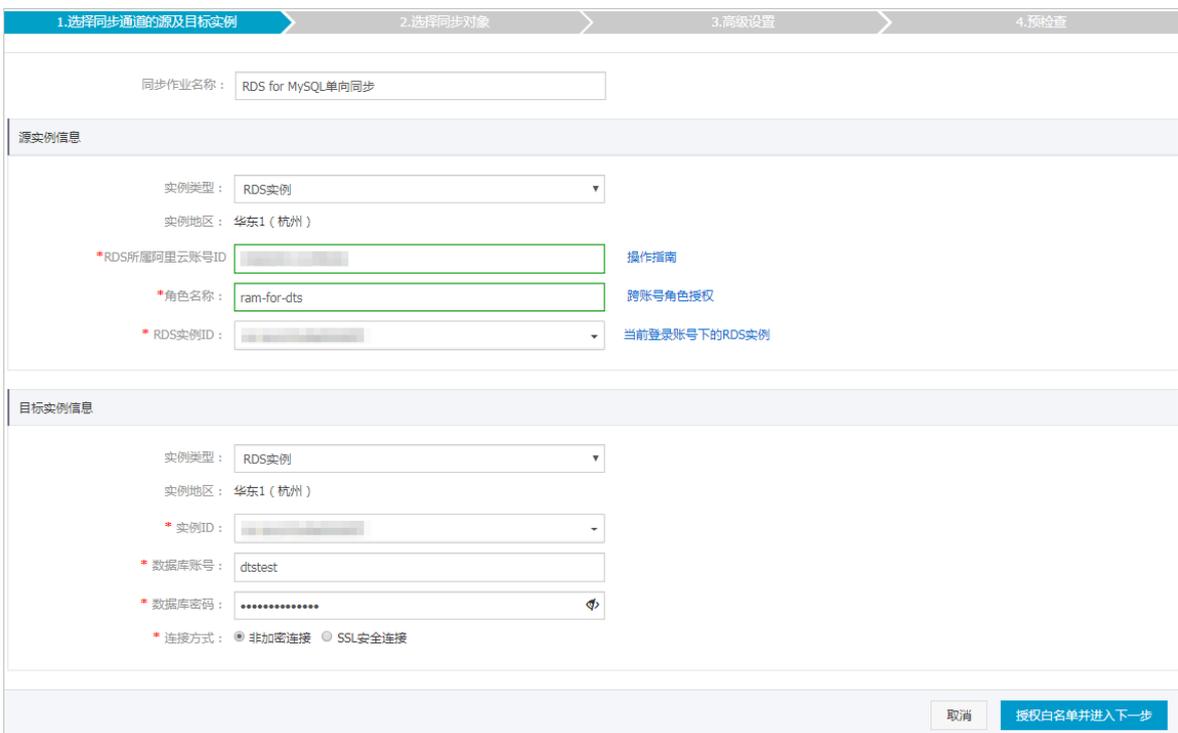
1. 使用目标RDS实例所属的阿里云账号[购买数据同步实例](#)。

**说明** 购买时，选择源实例和目标实例均为MySQL，并选择同步拓扑为单向同步。

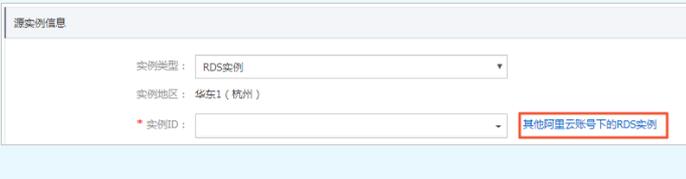
2. 使用目标RDS实例所属的阿里云账号登录[数据传输控制台](#)。
3. 在左侧导航栏，单击[数据同步](#)。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击该实例的[配置同步链路](#)。



6. 配置同步通道的源实例及目标实例信息。



| 类别 | 配置     | 说明   |
|----|--------|--|
| 无  | 同步作业名称 | DTS会自动生成一个同步作业名称，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。 |
|    | 实例类型   | 选择RDS实例。                                       |

| 类别     | 配置           | 说明  |
|--------|--------------|---|
| 源实例信息  | 实例地区         | 购买数据同步实例时选择的源实例地域信息，不可变更。   |
|        | RDS所属阿里云账号ID | 填入源RDS实例所属的阿里云账号ID。<br><br><div style="border: 1px solid #ccc; padding: 5px; background-color: #e6f2ff;"> <p><b>说明</b> 在配置该选项之前，需要先单击源实例信息页签中的其他阿里云账号下的RDS实例。</p>  </div> |
|        | 角色名称         | 填入准备工作步骤中配置的RAM角色名称。  |
|        | RDS实例ID      | 选择源实例ID。  |
| 目标实例信息 | 实例类型         | 选择RDS实例。  |
|        | 实例地区         | 购买数据同步实例时选择的目标实例地域信息，不可变更。  |
|        | 实例ID         | 选择作为数据同步目标的RDS实例ID。   |
|        | 数据库账号        | 填入目标RDS的数据库账号。<br><br><div style="border: 1px solid #ccc; padding: 5px; background-color: #e6f2ff;"> <p><b>说明</b> 当目标RDS实例的数据库类型为MySQL 5.5或MySQL 5.6时，无需配置数据库账号和数据库密码。</p> </div>   |
|        | 数据库密码        | 填入该数据库账号对应的密码。  |
|        | 连接方式         | 根据需求选择非加密连接或SSL安全连接。如果设置为SSL安全连接，您需要提前开启RDS实例的SSL加密功能，详情请参见设置SSL加密。<br><br><div style="border: 1px solid #ccc; padding: 5px; background-color: #e6f2ff;"> <p><b>注意</b> 目前仅中国内地及中国香港地域支持设置连接方式。</p> </div>   |

7. 单击页面右下角的授权白名单并进入下一步。

**说明**

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见DTS服务器的IP地址段。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置同步策略及对象信息。



| 配置项目                           | 配置说明   |
|--------------------------------|--|
| 选择同步对象                         | <p>在源库对象框中单击待同步的对象，然后单击  图标将其移动至已选择对象框。</p> <p>同步对象的选择粒度为库、表。</p> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p> 说明</p> <ul style="list-style-type: none"> <li>如果选择整个库作为同步对象，该库中所有对象的结构变更操作都会同步至目标库。</li> <li>默认情况下，同步对象在目标库中的名称与源库保持一致。如果您需要改变同步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 <a href="#">设置同步对象在目标实例中的名称</a>。</li> </ul> </div>  |
| 映射名称更改                         | <p>如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见 <a href="#">库表映射</a>。</p>  |
| 源表DMS_ONLINE_DDL过程中是否复制临时表到目标库 | <p>如源库使用 <a href="#">数据管理DMS (Data Management Service)</a> 执行Online DDL变更，您可以选择是否同步Online DDL变更产生的临时表数据。</p> <ul style="list-style-type: none"> <li>是：同步Online DDL变更产生的临时表数据。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p> 说明 Online DDL变更产生的临时表数据过大，可能会导致同步任务延迟。</p> </div> <ul style="list-style-type: none"> <li>否：不同步Online DDL变更产生的临时表数据，只同步源库的原始DDL数据。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p> 说明 该方案会导致目标库锁表。</p> </div> |

| 配置项目          | 配置说明  |
|---------------|---|
| 源、目标库无法连接重试时间 | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <p><b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p> |

9. 上述配置完成后，单击页面右下角的下一步。

10. 配置同步初始化的高级配置信息。



- 此步骤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和全量数据初始化。

11. 上述配置完成后，单击页面右下角的预检查并启动。

**注意**

- 在数据同步任务正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动数据同步任务。
- 如果预检查失败，单击具体检查项后的 ，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择确认屏蔽、忽略告警项并重新进行预检查，跳过告警检测项重新进行预检查。

12. 在预检查对话框中显示预检查通过后，关闭预检查对话框，同步作业将正式开始。

13. 等待该同步作业的链路初始化完成，直至状态处于同步中。

您可以在数据同步页面，查看数据同步状态。



# 7. MySQL同步至其他数据库

## 7.1. 从RDS MySQL同步至PolarDB MySQL

PolarDB是阿里巴巴自主研发的下一代关系型分布式云原生数据库，可完全兼容MySQL，具备简单易用、高性能、高可靠、高可用等优势。通过数据传输服务DTS（Data Transmission Service），您可以将RDS MySQL同步至PolarDB MySQL。

### 前提条件

已创建PolarDB MySQL集群，详情请参见[创建PolarDB MySQL集群](#)。

### 注意事项

- DTS在执行全量数据初始化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。因此您需要在执行数据同步前评估源库和目标库的性能，同时建议您在业务低峰期执行数据同步（例如源库和目标库的CPU负载在30%以下）。
- 全量初始化过程中，并发INSERT会导致目标集群的表碎片，全量初始化完成后，目标集群的表空间比源库的表空间大。
- 如果数据同步的源库没有主键或唯一约束，且记录的全字段没有唯一性，可能会出现重复数据。

### 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

### 支持同步的SQL操作

| 操作类型 | SQL操作语句   |
|------|---|
| DML  | INSERT、UPDATE、DELETE、REPLACE  |
| DDL  | <ul style="list-style-type: none"> <li>• ALTER TABLE、ALTER VIEW</li> <li>• CREATE FUNCTION、CREATE INDEX、CREATE PROCEDURE、CREATE TABLE、CREATE VIEW</li> <li>• DROP INDEX、DROP TABLE</li> <li>• RENAME TABLE</li> <li>• TRUNCATE TABLE</li> </ul> |

### 支持的同步架构

- 一对一单向同步
- 一对多单向同步
- 级联单向同步
- 多对一单向同步

关于各类同步架构的介绍及注意事项，请参见[数据同步拓扑介绍](#)。

### 功能限制

- 不兼容触发器  
当同步对象为整个库，且库中的触发器（TRIGGER）会更新库内某个表时，可能导致源和目标库的数据不一致。相关解决方案请参见[源库存在触发器时如何配置同步作业](#)。

● RENAME TABLE限制

RENAME TABLE操作可能导致同步数据不一致。例如同步对象只包含某个表，如果同步过程中源实例对该表执行了重命名操作，那么该表的数据将不会同步到目标库。为避免该问题，您可以在数据同步配置时将该表所属的整个数据库作为同步对象。

### 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

? 说明 购买时，选择源实例为MySQL、目标实例为PolarDB，并选择同步拓扑为单向同步。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击[数据同步](#)。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击[配置同步链路](#)。
6. 配置同步通道的源实例及目标实例信息。

1.选择同步通道的源及目标实例
2.选择同步对象
3.高级设置
4.预检查

同步作业名称:

**源实例信息**

实例类型:

实例地区: 华东1 (杭州)

\* 实例ID:  其他阿里云账号下的RDS实例

\* 数据库账号:

\* 数据库密码:

\* 连接方式:  非加密连接  SSL安全连接

**目标实例信息**

实例类型: PolarDB

实例地区: 华东1 (杭州)

\* PolarDB实例ID:

\* 数据库账号:

\* 数据库密码:

| 类别 | 配置     | 说明   |
|----|--------|--|
| 无  | 同步作业名称 | DTS会自动生成一个同步作业名称，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。 |
|    | 实例类型   | 选择RDS实例。                                       |
|    | 实例地区   | 购买数据同步实例时选择的源实例地域信息，不可变更。                      |

| 类别     | 配置          | 说明  |
|--------|-------------|---|
| 源实例信息  | 数据库账号       | 填入连接RDS实例的数据库账号。<br><div style="background-color: #e6f2ff; padding: 5px; margin-top: 10px;"> <span style="font-size: 1.2em;">?</span> <b>说明</b> <ul style="list-style-type: none"> <li>◦ 该账号需具备REPLICATION CLIENT、REPLICATION SLAVE、SHOW VIEW和所有同步对象的SELECT权限。</li> <li>◦ 当源RDS实例的数据库类型为MySQL 5.5或MySQL 5.6时，无需配置数据库账号和数据库密码。</li> </ul> </div> |
|        | 数据库密码       | 填入数据库账号的密码。   |
|        | 连接方式        | 根据需求选择 <b>非加密连接</b> 或 <b>SSL安全连接</b> 。如果设置为 <b>SSL安全连接</b> ，您需要提前开启RDS实例的SSL加密功能，详情请参见 <a href="#">设置SSL加密</a> 。<br><div style="background-color: #e6f2ff; padding: 5px; margin-top: 10px;"> <span style="font-size: 1.2em;">🔊</span> <b>注意</b> 目前仅中国内地及中国香港地域支持设置连接方式。                 </div>  |
| 目标实例信息 | 实例类型        | 固定为PolarDB，不可变更。  |
|        | 实例地区        | 购买数据同步实例时选择的目标实例地域信息，不可变更。  |
|        | PolarDB实例ID | 选择目标PolarDB集群ID。  |
|        | 数据库账号       | 填入PolarDB集群的数据库账号。<br><div style="background-color: #e6f2ff; padding: 5px; margin-top: 10px;"> <span style="font-size: 1.2em;">?</span> <b>说明</b> 用于数据同步的数据库账号需具备目标同步对象的ALL权限。                 </div>   |
|        | 数据库密码       | 填入数据库账号的密码。   |

7. 单击页面右下角的授权白名单并进入下一步。

? **说明**

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置目标已存在表的处理模式和同步对象。



| 配置项目        | 配置说明   |
|-------------|--|
| 目标已存在表的处理模式 | <ul style="list-style-type: none"><li><b>预检查并报错拦截:</b> 检查目标数据库中是否有同名的表。如果目标数据库中没有同名的表, 则通过该检查项目; 如果目标数据库中有同名的表, 则在预检查阶段提示错误, 数据同步作业不会被启动。<br/><p><b>说明</b> 如果目标库中同名的表不方便删除或重命名, 您可以<b>设置同步对象在目标实例中的名称</b>来避免表名冲突。</p></li><li><b>忽略报错并继续执行:</b> 跳过目标数据库中是否有同名表的检查项。<br/><p><b>警告</b> 选择为<b>忽略报错并继续执行</b>, 可能导致数据不一致, 给业务带来风险, 例如:</p><ul style="list-style-type: none"><li>表结构一致的情况下, 如果在目标库遇到与源库主键的值相同的记录, 在初始化阶段会保留目标库中的该条记录; 在增量同步阶段则会覆盖目标库的该条记录。</li><li>表结构不一致的情况下, 可能会导致无法初始化数据、只能同步部分列的数据或同步失败。</li></ul></li></ul> |

| 配置项目                           | 配置说明  |
|--------------------------------|---|
| 选择同步对象                         | <p>在源库对象框中单击待同步的对象，然后单击  图标将其移动至已选择对象框。</p> <p>同步对象的选择粒度为库、表。</p> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>如果选择整个库作为同步对象，那么该库中所有对象的结构变更操作都会同步至目标库。</li> <li>默认情况下，同步对象的名称保持不变。如果您需要改变同步对象在目标集群中的名称，请使用对象名映射功能，详情请参见<a href="#">设置同步对象在目标实例中的名称</a>。</li> </ul> </div>   |
| 映射名称更改                         | <p>如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见<a href="#">库表列映射</a>。</p>   |
| 源表DMS_ONLINE_DDL过程中是否复制临时表到目标库 | <p>如源库使用<a href="#">数据管理DMS (Data Management Service)</a> 执行Online DDL变更，您可以选择是否同步Online DDL变更产生的临时表数据。</p> <ul style="list-style-type: none"> <li><b>是</b>：同步Online DDL变更产生的临时表数据。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p><b>说明</b> Online DDL变更产生的临时表数据过大，可能会导致同步任务延迟。</p> </div> <ul style="list-style-type: none"> <li><b>否</b>：不同步Online DDL变更产生的临时表数据，只同步源库的原始DDL数据。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p><b>说明</b> 该方案会导致目标库锁表。</p> </div> |
| 源、目标库无法连接重试时间                  | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p><b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p> </div>  |

9. 上述配置完成后，单击页面右下角的下一步。

10. 配置同步初始化的高级配置信息。



**说明** 同步初始化类型细分为：结构初始化，全量数据初始化。选中结构初始化和全量数据初始化后，DTS会在增量数据同步之前，将源数据库中待同步对象的结构和存量数据，同步到目标数据库。

11. 上述配置完成后，单击页面右下角的预检查并启动。

**说明**

- 在同步作业正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步作业。
- 如果预检查失败，单击具体检查项后的，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择**确认屏蔽**、**忽略告警项**并重新进行预检查，跳过告警检测项重新进行预检查。

12. 在**预检查**对话框中显示**预检查通过**后，关闭**预检查**对话框，同步作业将正式开始。

13. 等待同步作业的链路初始化完成，直至处于**同步中**状态。

您可以在**数据同步**页面，查看数据同步作业的状态。



| 实例ID/作业名称               | 状态  | 同步概况                                   | 付费方式 | 同步架构(全部) | 操作                  |
|-------------------------|-----|--|------|----------|---------------------|
| hangzhou-hangzhou-small | 同步中 | 耗时: 1376 毫秒<br>速率: 0.00RPS/(0.000MB/s) | 按量付费 | 单向同步     | 暂停同步   转包年包月   升级更多 |

## 7.2. 从ECS上的自建MySQL同步至PolarDB MySQL

PolarDB是阿里巴巴自主研发的下一代关系型分布式云原生数据库，可完全兼容MySQL，具备简单易用、高性能、高可靠、高可用等优势。通过数据传输服务DTS（Data Transmission Service），您可以将自建的MySQL数据库同步至PolarDB MySQL，本文以ECS上的自建MySQL为例介绍配置流程。

### 前提条件

已创建PolarDB MySQL集群，详情请参见[创建PolarDB MySQL集群](#)。

### 注意事项

- DTS在执行全量数据初始化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。因此您需要在执行数据同步前评估源库和目标库的性能，同时建议您在业务低峰期执行数据同步（例如源库和目标库的CPU负载在30%以下）。
- 全量初始化过程中，并发INSERT会导致目标集群的表碎片，全量初始化完成后，目标集群的表空间比源库的表空间大。
- 如果数据同步的源库没有主键或唯一约束，且记录的全字段没有唯一性，可能会出现重复数据。

### 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

### 支持同步的SQL操作

| 操作类型 | SQL操作语句 |
|------|---------|
|------|---------|

| 操作类型 | SQL操作语句   |
|------|---|
| DML  | INSERT、UPDATE、DELETE、REPLACE  |
| DDL  | <ul style="list-style-type: none"> <li>ALTER TABLE、ALTER VIEW</li> <li>CREATE FUNCTION、CREATE INDEX、CREATE PROCEDURE、CREATE TABLE、CREATE VIEW</li> <li>DROP INDEX、DROP TABLE</li> <li>RENAME TABLE</li> <li>TRUNCATE TABLE</li> </ul> |

## 功能限制

- 不兼容触发器  
当同步对象为整个库，且库中的触发器（TRIGGER）会更新库内某个表时，可能导致源和目标库的数据不一致。相关解决方案请参见[源库存在触发器时如何配置同步作业](#)。
- RENAME TABLE限制  
RENAME TABLE操作可能导致同步数据不一致。例如同步对象只包含某个表，如果同步过程中源实例对该表执行了重命名操作，那么该表的数据将不会同步到目标库。为避免该问题，您可以在数据同步配置时将该表所属的整个数据库作为同步对象。

## 准备工作

为自建MySQL创建账号并设置binlog

 **说明** 用于数据同步的数据库账号需具备待同步对象的REPLICATION CLIENT、REPLICATION SLAVE、SHOW VIEW和所有同步对象的SELECT权限。

## 支持的同步架构

- 一对一单向同步
- 一对多单向同步
- 级联单向同步
- 多对一单向同步

关于各类同步架构的介绍及注意事项，请参见[数据同步拓扑介绍](#)。

## 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

 **说明** 购买时，选择源实例为MySQL、目标实例为PolarDB，并选择同步拓扑为单向同步。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击[数据同步](#)。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击[配置同步链路](#)。
6. 配置同步通道的源实例及目标实例信息。

1. 选择同步通道的源及目标实例
2. 选择同步对象
3. 高级设置
4. 预检查

同步作业名称:

**源实例信息**

实例类型:

实例地区:

\* ECS实例ID:

数据库类型:

\* 端口:

\* 数据库账号:

\* 数据库密码:

**目标实例信息**

实例类型:

实例地区:

\* PolarDB实例ID:

\* 数据库账号:

\* 数据库密码:

| 类别     | 配置          | 说明  |
|--------|-------------|---|
| 无      | 同步作业名称      | DT S会自动生成一个同步作业名称，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。   |
| 源实例信息  | 实例类型        | 选择ECS上的自建数据库。   |
|        | 实例地区        | 购买数据同步实例时选择的源实例地域信息，不可变更。   |
|        | ECS实例ID     | 选择自建MySQL数据库所属的ECS实例ID。   |
|        | 数据库类型       | 固定为MySQL，不可变更。  |
|        | 端口          | 填入自建MySQL的数据库服务端口。  |
|        | 数据库账号       | 填入连接自建MySQL的数据库账号。<br><br><span style="background-color: #e0f2f7; padding: 5px; border: 1px solid #ccc;">? 说明 用于数据同步的数据库账号需具备REPLICATION CLIENT、REPLICATION SLAVE、SHOW VIEW和所有同步对象的SELECT权限。</span> |
|        | 数据库密码       | 填入数据库账号的密码。   |
| 目标实例信息 | 实例类型        | 固定为PolarDB实例，不可变更。  |
|        | 实例地区        | 购买数据同步实例时选择的目标实例地域信息，不可变更。  |
|        | PolarDB实例ID | 选择目标PolarDB集群ID。  |
|        | 数据库账号       | 填入连接PolarDB集群的数据库账号。<br><br><span style="background-color: #e0f2f7; padding: 5px; border: 1px solid #ccc;">? 说明 用于数据同步的数据库账号需具备目标同步对象的ALL权限。</span>   |

| 类别 | 配置    | 说明          |
|----|-------|-------------|
|    | 数据库密码 | 填入数据库账号的密码。 |

7. 单击页面右下角的授权白名单并进入下一步。

**说明**

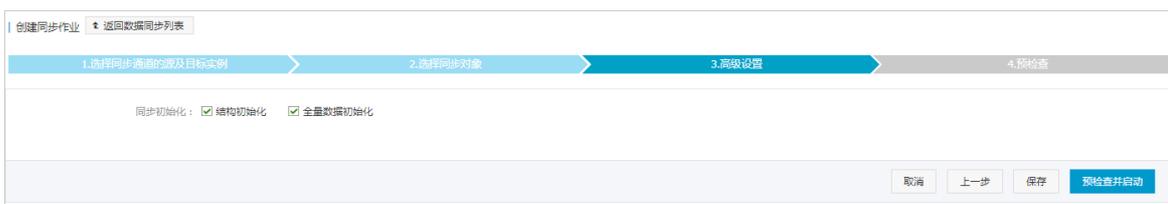
- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置目标已存在表的处理模式和同步对象。

| 配置项目                           | 配置说明  |
|--------------------------------|---|
| 目标已存在表的处理模式                    | <ul style="list-style-type: none"> <li> <b>预检查并报错拦截</b>：检查目标数据库中是否有同名的表。如果目标数据库中没有同名的表，则通过该检查项目；如果目标数据库中有同名的表，则在预检查阶段提示错误，数据同步作业不会被启动。                     <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin-top: 5px;"> <p> <b>说明</b> 如果目标库中同名的表不方便删除或重命名，您可以<b>设置同步对象在目标实例中的名称</b>来避免表名冲突。</p> </div> </li> <li> <b>忽略报错并继续执行</b>：跳过目标数据库中是否有同名表的检查项。                     <div style="border: 1px solid #ccc; background-color: #fff9c4; padding: 5px; margin-top: 5px;"> <p> <b>警告</b> 选择为<b>忽略报错并继续执行</b>，可能导致数据不一致，给业务带来风险，例如：</p> <ul style="list-style-type: none"> <li>表结构一致的情况下，如果在目标库遇到与源库主键的值相同的记录，在初始化阶段会保留目标库中的该条记录；在增量同步阶段则会覆盖目标库的该条记录。</li> <li>表结构不一致的情况下，可能会导致无法初始化数据、只能同步部分列的数据或同步失败。</li> </ul> </div> </li> </ul> |
| 选择同步对象                         | <p>在<b>源库对象框</b>中单击待同步的对象，然后单击  图标将其移动至<b>已选择对象框</b>。</p> <p>同步对象的选择粒度为库、表。</p> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin-top: 5px;"> <p> <b>说明</b></p> <ul style="list-style-type: none"> <li>如果选择整个库作为同步对象，那么该库中所有对象的结构变更操作都会同步至目标库。</li> <li>默认情况下，同步对象的名称保持不变。如果您需要改变同步对象在目标集群中的名称，请使用对象名映射功能，详情请参见<b>设置同步对象在目标实例中的名称</b>。</li> </ul> </div>  |
| 映射名称更改                         | <p>如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见<b>库表列映射</b>。</p>  |
| 源表DMS_ONLINE_DDL过程中是否复制临时表到目标库 | <p>如源库使用<b>数据管理DMS (Data Management Service)</b>执行Online DDL变更，您可以选择是否同步Online DDL变更产生的临时表数据。</p> <ul style="list-style-type: none"> <li> <b>是</b>：同步Online DDL变更产生的临时表数据。                     <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin-top: 5px;"> <p> <b>说明</b> Online DDL变更产生的临时表数据过大，可能会导致同步任务延迟。</p> </div> </li> <li> <b>否</b>：不同步Online DDL变更产生的临时表数据，只同步源库的原始DDL数据。                     <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin-top: 5px;"> <p> <b>说明</b> 该方案会导致目标库锁表。</p> </div> </li> </ul>  |

| 配置项目          | 配置说明  |
|---------------|---|
| 源、目标库无法连接重试时间 | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <p><b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p> |

- 上述配置完成后，单击页面右下角的下一步。
- 配置同步初始化的高级配置信息。



**说明** 同步初始化类型细分为：结构初始化，全量数据初始化。选中结构初始化和全量数据初始化后，DTS会在增量数据同步之前，将源数据库中待同步对象的结构和存量数据，同步到目标数据库。

- 上述配置完成后，单击页面右下角的预检查并启动。

**说明**

- 在同步作业正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步作业。
- 如果预检查失败，单击具体检查项后的 ，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择确认屏蔽、忽略告警项并重新进行预检查，跳过告警检测项重新进行预检查。

- 在预检查对话框中显示预检查通过后，关闭预检查对话框，同步作业将正式开始。
- 等待同步作业的链路初始化完成，直至处于同步中状态。

您可以在数据同步页面，查看数据同步作业的状态。



## 7.3. 从通过专线、VPN网关或智能接入网关接入的自建MySQL同步至PolarDB MySQL

PolarDB是阿里巴巴自主研发的下一代关系型分布式云原生数据库，可完全兼容MySQL，具备简单易用、高性能、高可靠、高可用等优势。通过数据传输服务DTS（Data Transmission Service），您可以将自建的MySQL数据库同步至PolarDB MySQL，本文以通过专线、VPN网关或智能接入网关接入的自建MySQL为例介绍配置流程。

### 前提条件

- 自建MySQL数据库版本为5.1、5.5、5.6、5.7或8.0版本。
- 已经将自建MySQL数据库通过专线、VPN网关或智能接入网关接入至阿里云专有网络，详情请参见[本地IDC接入至阿里云方案概览](#)。

 **说明** 接入至阿里云专有网络后，还需要允许DTS服务器的IP地址访问自建数据库所属的网络，详情请参见[通过VPN网关实现本地IDC与DTS云服务互通](#)。

- 已创建PolarDB MySQL集群，详情请参见[创建PolarDB MySQL集群](#)。

 **说明** PolarDB MySQL集群的存储空间须大于自建MySQL数据库占用的存储空间。

## 注意事项

- DTS在执行全量数据初始化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。因此您需要在执行数据同步前评估源库和目标库的性能，同时建议您在业务低峰期执行数据同步（例如源库和目标库的CPU负载在30%以下）。
- 全量初始化过程中，并发INSERT会导致目标集群的表碎片，全量初始化完成后，目标集群的表空间比源库的表空间大。
- 如果数据同步的源集群没有主键或唯一约束，且记录的全字段没有唯一性，可能会出现重复数据。

## 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

## 支持同步的SQL操作

| 操作类型 | SQL操作语句   |
|------|---|
| DML  | INSERT、UPDATE、DELETE、REPLACE  |
| DDL  | <ul style="list-style-type: none"> <li>• ALTER TABLE、ALTER VIEW</li> <li>• CREATE FUNCTION、CREATE INDEX、CREATE PROCEDURE、CREATE TABLE、CREATE VIEW</li> <li>• DROP INDEX、DROP TABLE</li> <li>• RENAME TABLE</li> <li>• TRUNCATE TABLE</li> </ul> |

## 支持的同步架构

- 一对一单向同步
- 一对多单向同步
- 级联单向同步
- 多对一单向同步

关于各类同步架构的介绍及注意事项，请参见[数据同步拓扑介绍](#)。

## 功能限制

- 不兼容触发器

当同步对象为整个库，且库中的触发器（TRIGGER）会更新库内某个表时，可能导致源和目标库的数据不一致。相关解决方案请参见[源库存在触发器时如何配置同步作业](#)。

- **RENAME TABLE限制**  
RENAME TABLE操作可能导致同步数据不一致。例如同步对象只包含某个表，如果同步过程中源实例对该表执行了重命名操作，那么该表的数据将不会同步到目标库。为避免该问题，您可以在数据同步配置时将该表所属的整个数据库作为同步对象。

## 准备工作

### 为自建MySQL创建账号并设置binlog

**说明** 用于数据同步的数据库账号需具备REPLICATION CLIENT、REPLICATION SLAVE、SHOW VIEW和所有同步对象的SELECT权限。

## 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

**说明** 购买时，选择源实例为MySQL、目标实例为PolarDB，并选择同步拓扑为单向同步。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击[数据同步](#)。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击[配置同步链路](#)。
6. 配置同步通道的源实例及目标实例信息。

1.选择同步通道的源及目标实例
2.选择同步对象
3.高级设置
4.预检查

同步作业名称:

**源实例信息**

实例类型:

实例地区:

\* 已和源端数据库联通的VPC:

数据库类型:

\* IP地址:

\* 端口:

\* 数据库账号:

\* 数据库密码:

**目标实例信息**

实例类型:

实例地区:

\* PolarDB实例ID:

\* 数据库账号:

\* 数据库密码:

| 类别 | 配置     | 说明   |
|----|--------|--|
| 无  | 同步作业名称 | DTS会自动生成一个同步作业名称，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。 |

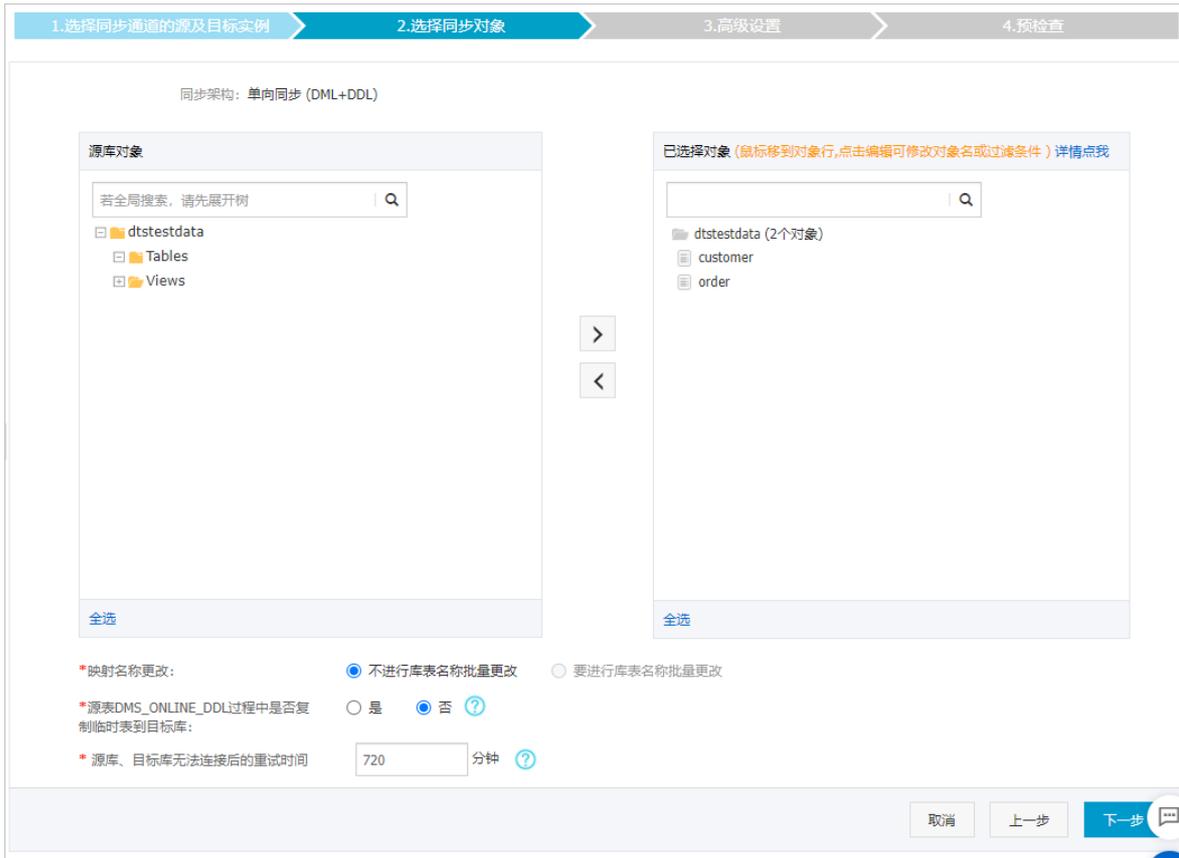
| 类别     | 配置            | 说明  |
|--------|---------------|---|
| 源实例信息  | 实例类型          | 选择通过专线/VPN网关/智能接入网关接入的自建数据库。  |
|        | 实例地区          | 购买数据同步实例时选择的源实例地域信息，不可变更。   |
|        | 已和源端数据库联通的VPC | 选择自建数据库接入的VPC ID。   |
|        | 数据库类型         | 固定为MySQL，不可变更。  |
|        | IP地址          | 填入自建MySQL数据库的服务器IP地址。   |
|        | 端口            | 填入自建MySQL的数据库服务端口。  |
|        | 数据库账号         | 填入为自建MySQL创建的数据库账号，详情请参见 <a href="#">准备工作</a> 。   |
|        | 数据库密码         | 填入该数据库账号的密码。  |
| 目标实例信息 | 实例类型          | 固定为PolarDB，不可变更。  |
|        | 实例地区          | 购买数据同步实例时选择的目标实例地域信息，不可变更。  |
|        | PolarDB实例ID   | 选择目标PolarDB集群ID。  |
|        | 数据库账号         | 填入PolarDB集群的数据库账号。<br><br><div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin-top: 5px;"> <span style="font-size: 1.2em; color: #007bff;">?</span> 说明 该数据库账号需具备目标同步对象的ALL权限。 </div> |
|        | 数据库密码         | 填入该数据库账号的密码。  |

7. 单击页面右下角的授权白名单并进入下一步。

? 说明

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置目标已存在表的处理模式和同步对象。



| 配置项目        | 配置说明  |
|-------------|---|
| 目标已存在表的处理模式 | <ul style="list-style-type: none"> <li> <b>预检查并报错拦截：</b>检查目标数据库中是否有同名的表。如果目标数据库中没有同名的表，则通过该检查项目；如果目标数据库中有同名的表，则在预检查阶段提示错误，数据同步作业不会被启动。           <div style="border: 1px solid #ccc; padding: 5px; margin: 5px 0;"> <p><b>说明</b> 如果目标库中同名的表不方便删除或重命名，您可以<b>设置同步对象在目标实例中的名称</b>来避免表名冲突。</p> </div> </li> <li> <b>忽略报错并继续执行：</b>跳过目标数据库中是否有同名表的检查项。           <div style="border: 1px solid #ccc; padding: 5px; margin: 5px 0;"> <p><b>警告</b> 选择为<b>忽略报错并继续执行</b>，可能导致数据不一致，给业务带来风险，例如：</p> <ul style="list-style-type: none"> <li>表结构一致的情况下，如果在目标库遇到与源库主键的值相同的记录，在初始化阶段会保留目标库中的该条记录；在增量同步阶段则会覆盖目标库的该条记录。</li> <li>表结构不一致的情况下，可能会导致无法初始化数据、只能同步部分列的数据或同步失败。</li> </ul> </div> </li> </ul> |

| 配置项目                           | 配置说明  |
|--------------------------------|---|
| 选择同步对象                         | <p>在源库对象框中单击待同步的对象，然后单击  图标将其移动至已选择对象框。</p> <p>同步对象的选择粒度为库、表。</p> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>如果选择整个库作为同步对象，那么该库中所有对象的结构变更操作都会同步至目标库。</li> <li>默认情况下，同步对象的名称保持不变。如果您需要改变同步对象在目标集群中的名称，请使用对象名映射功能，详情请参见<a href="#">设置同步对象在目标实例中的名称</a>。</li> </ul> </div>   |
| 映射名称更改                         | <p>如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见<a href="#">库表映射</a>。</p>  |
| 源表DMS_ONLINE_DDL过程中是否复制临时表到目标库 | <p>如源库使用<a href="#">数据管理DMS (Data Management Service)</a> 执行Online DDL变更，您可以选择是否同步Online DDL变更产生的临时表数据。</p> <ul style="list-style-type: none"> <li><b>是</b>：同步Online DDL变更产生的临时表数据。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p><b>说明</b> Online DDL变更产生的临时表数据过大，可能会导致同步任务延迟。</p> </div> <ul style="list-style-type: none"> <li><b>否</b>：不同步Online DDL变更产生的临时表数据，只同步源库的原始DDL数据。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p><b>说明</b> 该方案会导致目标库锁表。</p> </div> |
| 源、目标库无法连接重试时间                  | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p><b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p> </div>  |

9. 上述配置完成后，单击页面右下角的下一步。

10. 配置同步初始化的高级配置信息。

创建同步作业 返回数据同步列表

---

1. 选择同步源和目标实例 2. 选择同步对象 3. 高级设置 4. 预检查

---

同步初始化： 结构初始化  全量数据初始化

---

**说明** 同步初始化类型细分为：结构初始化，全量数据初始化。选中结构初始化和全量数据初始化后，DTS会在增量数据同步之前，将源数据库中待同步对象的结构和存量数据，同步到目标数据库。

11. 上述配置完成后，单击页面右下角的预检查并启动。

**说明**

- 在同步作业正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步作业。
- 如果预检查失败，单击具体检查项后的 ，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择**确认屏蔽**、**忽略告警项**并重新进行预检查，跳过告警检测项重新进行预检查。

12. 在预检查对话框中显示预检查通过后，关闭预检查对话框，同步作业将正式开始。

13. 等待同步作业的链路初始化完成，直至处于同步中状态。

您可以在数据同步页面，查看数据同步作业的状态。



## 7.4. 从RDS MySQL同步到云原生数据仓库AnalyticDB MySQL

云原生数据仓库AnalyticDB MySQL是阿里巴巴自主研发的海量数据实时高并发在线分析（Realtime OLAP）云计算服务，使得您可以在毫秒级针对千亿级数据进行即时的多维分析透视和业务探索。通过数据传输服务DTS（Data Transmission Service），您可以将RDS MySQL同步到云原生数据仓库AnalyticDB MySQL，帮助您快速构建企业内部BI、交互查询、实时报表等系统。

### 前提条件

- RDS MySQL中待同步的数据表必须具备主键。
- 已创建目标云原生数据仓库AnalyticDB MySQL集群，详情请参见[创建云原生数据仓库AnalyticDB MySQL集群](#)。
- 确保目标云原生数据仓库AnalyticDB MySQL具备充足的存储空间。

### 注意事项

- DTS在执行全量数据初始化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。因此您需要在执行数据同步前评估源库和目标库的性能，同时建议您在业务低峰期执行数据同步（例如源库和目标库的CPU负载在30%以下）。
- 请勿在数据同步时，对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。
- 由于云原生数据仓库AnalyticDB MySQL本身的使用限制，当云原生数据仓库AnalyticDB MySQL集群中的节点磁盘空间使用量超过80%，该集群将被锁定。请提前根据待同步的对象预估所需空间，确保目标集群具备充足的存储空间。
- 暂不支持同步前缀索引，如果源库存在前缀索引可能导致数据同步失败。

### 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

## 支持同步的SQL操作

- DDL操作：CREATE TABLE、DROP TABLE、RENAME TABLE、TRUNCATE TABLE、ADD COLUMN、DROP COLUMN、MODIFY COLUMN
- DML操作：INSERT、UPDATE、DELETE

 **说明** 如果在数据同步的过程中变更了源表的字段类型，同步作业将报错并中断。您可以[提交工单](#)处理或参照文末的方法来手动修复，详情请参见[修复因变更字段类型导致的同步失败](#)。

## 数据库账号的权限要求

| 数据库                     | 所需权限  |
|-------------------------|---|
| RDS MySQL               | REPLICATION CLIENT、REPLICATION SLAVE、SHOW VIEW和所有同步对象的SELECT权限。 |
| 云原生数据仓库AnalyticDB MySQL | 读写权限。   |

## 数据类型映射关系

由于MySQL和云原生数据仓库AnalyticDB MySQL的数据类型并不是一一对应的，所以DTS在进行结构初始化时，会根据数据类型定义进行类型映射，详情请参见[结构初始化涉及的数据类型映射关系](#)。

## 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

 **说明** 购买时，选择源实例为MySQL，目标实例为AnalyticDB MySQL，并选择同步拓扑为单向同步。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击[数据同步](#)。
4. 在同步作业列表页面顶部，选择数据同步实例所属地域。
5. 定位至已购买的数据同步实例，单击[配置同步链路](#)。
6. 配置同步通道的源实例及目标实例信息。

1.选择同步通道的源及目标实例
2.ADS账号授权
3.选择同步对象
4.预检查

同步作业名称:

**源实例信息**

实例类型:

实例地区: 华东1 (杭州)

\* 实例ID:  [其他阿里云账号下的RDS实例](#)

\* 数据库账号:

\* 数据库密码:

\* 连接方式:  非加密连接  SSL安全连接

**目标实例信息**

实例类型: ADS

实例地区: 华东1 (杭州)

\* 版本:  2.0  3.0

\* 数据库:

\* 数据库账号:

\* 数据库密码:

| 类别     | 配置     | 说明   |
|--------|--------|--|
| 无      | 同步作业名称 | DTS会自动生成一个同步作业名称，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。   |
| 源实例信息  | 实例类型   | 选择RDS实例。   |
|        | 实例地区   | 购买数据同步实例时选择的源实例地域信息，不可变更。  |
|        | 实例ID   | 选择源RDS实例ID。  |
|        | 数据库账号  | 填入源RDS的数据库账号，权限要求请参见 <a href="#">数据库账号的权限要求</a> 。<br><br><span style="background-color: #e1f5fe; padding: 5px; border: 1px solid #cfe2f3;">? 说明 当源RDS实例的数据库类型为MySQL 5.5或MySQL 5.6时，无需配置数据库账号和数据库密码。</span>         |
|        | 数据库密码  | 填入该数据库账号对应的密码。   |
| 目标实例信息 | 连接方式   | 根据需求选择非加密连接或SSL安全连接。如果设置为SSL安全连接，您需要提前开启RDS实例的SSL加密功能，详情请参见 <a href="#">设置SSL加密</a> 。<br><br><span style="background-color: #e1f5fe; padding: 5px; border: 1px solid #cfe2f3;">🔊 注意 目前仅中国内地及中国香港地域支持设置连接方式。</span> |
|        | 实例类型   | 固定为ADS，不可变更。   |
|        | 实例地区   | 购买数据同步实例时选择的目标实例地域信息，不可变更。   |
|        | 版本     | 选择为3.0。  |
|        | 数据库    | 选择目标云原生数据仓库AnalyticDB MySQL的集群ID。  |

| 类别 | 配置    | 说明   |
|----|-------|--|
|    | 数据库账号 | 填入云原生数据仓库AnalyticDB MySQL的数据库账号，权限要求请参见 <a href="#">数据库账号的权限要求</a> 。 |
|    | 数据库密码 | 填入该数据库账号对应的密码。   |

7. 单击页面右下角的授权白名单并进入下一步。

**说明**

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置同步策略及对象信息。

The screenshot shows the '3. 选择同步对象' (Select synchronization objects) step in the DTS configuration wizard. It includes the following elements:

- 同步初始化:**  结构初始化  全量数据初始化
- 注:** DTS全量任务运行期间，不要清理DTS任务启动后源库产生的增量数据日志。源库如果过早清理日志，可能会导致DTS增量任务失败
- 目标已存在表的处理模式:**  预检查并报错拦截  忽略报错并继续执行
- 多表归并:**  是  否
- 同步操作类型:**  Insert  Update  Delete  Alter Table  Truncate Table  Create Table  Drop Table
- 源库对象:** 树状结构显示 dtstestdata (Tables, Views)
- 已选择对象:** dtstestdata (2个对象) (customer, order)
- \*映射名称更改:**  不进行库表名称批量更改  要进行库表名称批量更改
- \*源表DMS\_ONLINE\_DDL过程中是否复制临时表到目标库:**  是  否
- \*源库、目标库无法连接后的重试时间:** 720 分钟
- 底部按钮:** 取消, 上一步, 下一步, 预检查并启动

| 配置    | 说明   |
|-------|--|
| 同步初始化 | 默认情况下，您需要同时选中 <b>结构初始化</b> 和 <b>全量数据初始化</b> 。预检查完成后，DTS会将源实例中待同步对象的结构及数据在目标集群中初始化，作为后续增量同步数据的基线数据。 |

| 配置          | 说明   |
|-------------|--|
| 目标已存在表的处理模式 | <ul style="list-style-type: none"> <li>○ <b>预检查并报错拦截</b>：检查目标数据库中是否有同名的表。如果目标数据库中没有同名的表，则通过该检查项目；如果目标数据库中有同名的表，则在预检查阶段提示错误，数据同步作业不会被启动。</li> </ul> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin: 5px 0;"> <p> <b>说明</b> 如果目标库中同名的表不方便删除或重命名，您可以更改该表在目标库中的名称，详情请参见<a href="#">设置同步对象在目标实例中的名称</a>。</p> </div> <ul style="list-style-type: none"> <li>○ <b>忽略报错并继续执行</b>：跳过目标数据库中是否有同名表的检查项。</li> </ul> <div style="border: 1px solid #ccc; background-color: #fff9c4; padding: 5px; margin: 5px 0;"> <p> <b>警告</b> 选择为忽略报错并继续执行，可能导致数据不一致，给业务带来风险，例如：</p> <ul style="list-style-type: none"> <li>■ 表结构一致的情况下，在目标库遇到与源库主键的值相同的记录，则会保留目标集群中的该条记录，即源库中的该条记录不会同步至目标数据库中。</li> <li>■ 表结构不一致的情况下，可能会导致无法初始化数据、只能同步部分列的数据或同步失败。</li> </ul> </div> |
| 多表归并        | <ul style="list-style-type: none"> <li>○ 选择为是：DTS将在每个表中增加 <code>__dts_data_source</code> 列来存储数据来源，且不再支持DDL同步。</li> <li>○ 选择为否：默认选项，支持DDL同步。</li> </ul> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin: 5px 0;"> <p> <b>说明</b> 多表归并功能基于任务级别，即不支持基于表级别执行多表归并。如果需要让部分表执行多表归并，另一部分不执行多表归并，您可以创建两个数据同步作业。</p> </div>   |
| 同步操作类型      | <p>根据业务选中需要同步的操作类型，支持的同步操作详情请参见<a href="#">支持同步的SQL操作</a>，默认情况下都处于选中状态。</p>  |
| 选择同步对象      | <p>在源库对象框中单击待同步的对象，然后单击  图标将其移动至已选择对象框。</p> <p>同步对象的选择粒度为库、表。</p> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin: 5px 0;"> <p> <b>说明</b></p> <ul style="list-style-type: none"> <li>○ 如果选择整个库作为同步对象，那么该库中所有对象的结构变更操作会同步至目标库。</li> <li>○ 如果选择某个表作为同步对象，那么只有这个表的ADD COLUMN操作会同步至目标库。</li> <li>○ 默认情况下，同步对象的名称保持不变。如果您需要同步对象在目标集群上名称不同，请使用对象名映射功能，详情请参见<a href="#">设置同步对象在目标实例中的名称</a>。</li> </ul> </div>  |
| 映射名称更改      | <p>如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见<a href="#">库表列映射</a>。</p>  |

| 配置                             | 说明   |
|--------------------------------|--|
| 源表DMS_ONLINE_DDL过程中是否复制临时表到目标库 | <p>如源库使用数据管理DMS (Data Management Service) 执行Online DDL变更, 您可以选择是否同步Online DDL变更产生的临时表数据。</p> <ul style="list-style-type: none"> <li>是: 同步Online DDL变更产生的临时表数据。</li> </ul> <p><b>说明</b> Online DDL变更产生的临时表数据过大, 可能会导致同步任务延迟。</p> <ul style="list-style-type: none"> <li>否: 不同步Online DDL变更产生的临时表数据, 只同步源库的原始DDL数据。</li> </ul> <p><b>说明</b> 该方案会导致目标库锁表。</p> |
| 源、目标库无法连接重试时间                  | <p>当源、目标库无法连接时, DTS默认重试720分钟(即12小时), 您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库, 同步任务将自动恢复。否则, 同步任务将失败。</p> <p><b>说明</b> 由于连接重试期间, DTS将收取任务运行费用, 建议您根据业务需要自定义重试时间, 或者在源和目标库实例释放后尽快释放DTS实例。</p>   |

9. 上述配置完成后, 单击页面右下角的下一步。

10. 设置待同步的表在目标库中类型。

**说明** 选择了结构初始化后, 您需要定义待同步的表在云原生数据仓库AnalyticDB MySQL中的类型、主键列、分区列等信息, 详情请参见CREATE TABLE操作手册。

11. 上述配置完成后, 单击页面右下角的预检查并启动。

**说明**

- 在同步作业正式启动之前, 会先进行预检查。只有预检查通过后, 才能成功启动同步作业。
- 如果预检查失败, 单击具体检查项后的 , 查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项, 您也可以选择确认屏蔽、忽略告警项并重新进行预检查, 跳过告警检测项重新进行预检查。

12. 在预检查对话框中显示预检查通过后, 关闭预检查对话框, 同步作业将正式开始。

13. 等待同步作业的链路初始化完成, 直至处于同步中状态。

您可以在数据同步页面, 查看数据同步作业的状态。

### 修复因变更字段类型导致的同步失败

如果在数据同步的过程中变更了源表的字段类型，同步作业将报错并中断。您可参照如下的方法来手动修复，或者[提交工单](#)处理。

1. 在目标实例中，根据同步失败的表A（表名以customer为例），重新创建一个新表B（表名以customer\_new为例），确保两张表的表结构保持一致。
2. 通过INSERT INTO SELECT命令，将表A的数据复制并插入到新创建的表B中，确保两张表的数据保持一致。
3. 重命名或删除同步失败的表A，然后将表B的名称修改为customer。
4. 在DTS控制台，重新启动数据同步作业。

## 7.5. 从RDS MySQL同步至云原生数据仓库AnalyticDB PostgreSQL

数据传输服务DTS（Data Transmission Service）支持将RDS MySQL同步至云原生数据仓库AnalyticDB PostgreSQL。通过DTS提供的数据同步功能，可以轻松实现数据的流转，将企业数据集中分析。

### 前提条件

- RDS MySQL中待同步的数据表必须具备主键。
- 已创建目标云原生数据仓库AnalyticDB PostgreSQL实例，如未创建请参见[创建云原生数据仓库AnalyticDB PostgreSQL实例](#)。

### 注意事项

DTS在执行全量数据初始化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。因此您需要在执行数据同步前评估源库和目标库的性能，同时建议您在业务低峰期执行数据同步（例如源库和目标库的CPU负载在30%以下）。

### 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

### 同步限制

- 同步对象仅支持数据表。
- 不支持BIT、VARBIT、GEOMETRY、ARRAY、UUID、TSQUERY、TSVECTOR、TXID\_SNAPSHOT类型的数据同步。
- 暂不支持同步前缀索引，如果源库存在前缀索引可能导致数据同步失败。
- 在数据同步时，请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。

### 支持同步的SQL操作

- DML操作：INSERT、UPDATE、DELETE。
- DDL操作：ADD COLUMN。

 **说明** 不支持CREATE TABLE操作，如果您需要将新增的表作为同步对象，则需要执行[新增同步对象](#)操作。

### 支持的同步架构

- 1对1单向同步。
- 1对多单向同步。
- 多对1单向同步。

### 术语及概念对应关系

|          |                              |
|----------|------------------------------|
| MySQL    | 云原生数据仓库AnalyticDB PostgreSQL |
| Database | Schema                       |
| Table    | Table                        |

### 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

? **说明** 购买时，选择源实例为MySQL，目标实例为AnalyticDB for PostgreSQL，并选择同步拓扑为单向同步。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击数据同步。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击配置同步链路。
6. 配置同步作业的源实例及目标实例信息。

1.选择同步通道的源及目标实例
2.选择同步对象
3.预检查

同步作业名称：

**源实例信息**

实例类型：

实例地区：

\* 实例ID： [其他阿里云账号下的RDS实例](#)

\* 数据库账号：

\* 数据库密码：

\* 连接方式： 非加密连接  SSL安全连接

**目标实例信息**

实例类型：

实例地区：

\* 实例ID：

\* 数据库名称：

\* 数据库账号：

\* 数据库密码：

| 类别 | 配置     | 说明   |
|----|--------|--|
| 无  | 同步作业名称 | DTS会自动生成一个同步作业名称，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。 |
|    | 实例类型   | 选择RDS实例。                                       |
|    | 实例地区   | 购买数据同步实例时选择的源实例地域信息，不可变更。                      |
|    | 实例ID   | 选择源RDS实例ID。                                    |

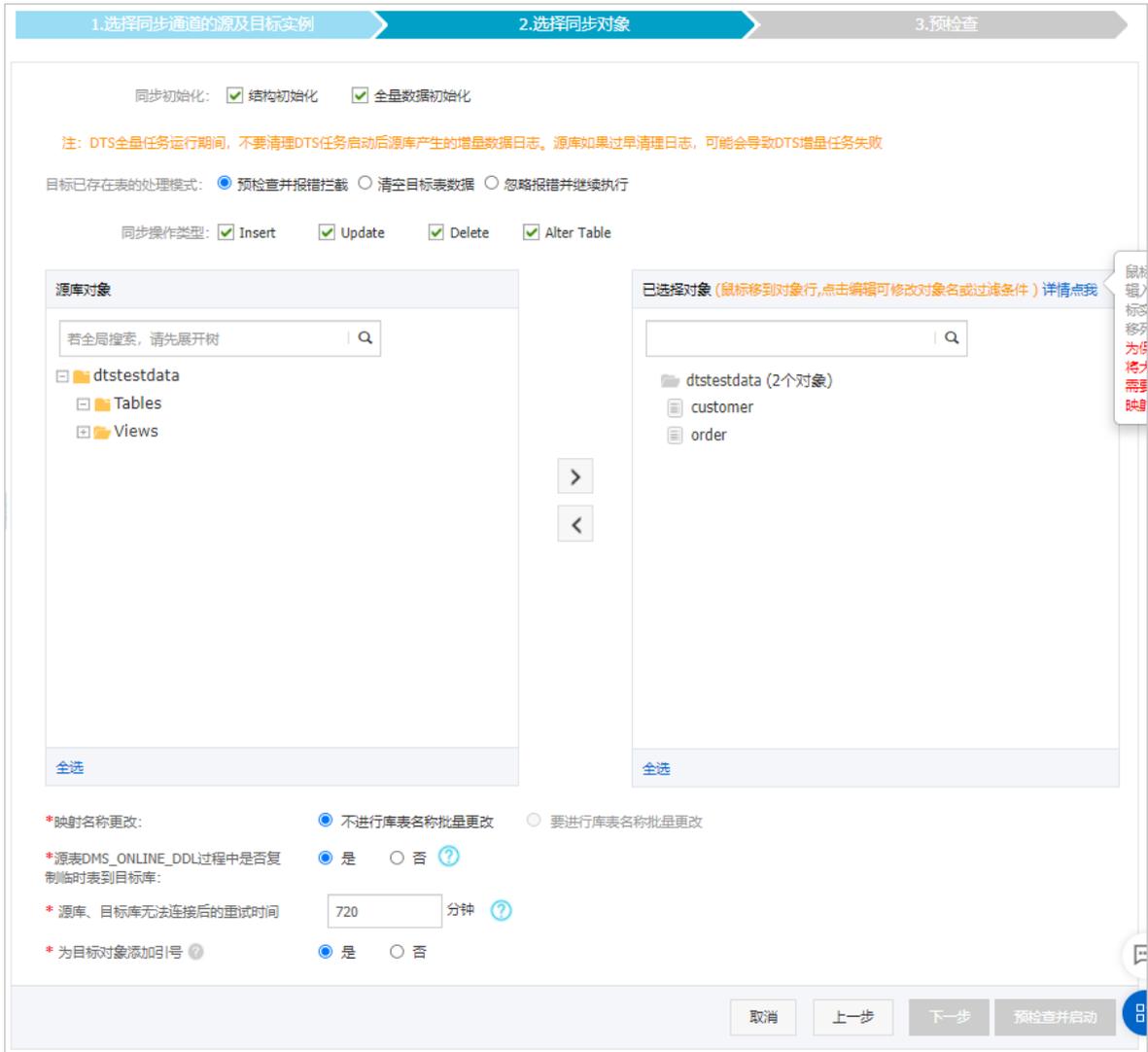
| 类别     | 配置    | 说明   |
|--------|-------|--|
| 源实例信息  | 数据库账号 | 填入RDS MySQL的数据库账号。<br><br><b>说明</b> 当源RDS实例的数据库类型为MySQL 5.5或MySQL 5.6时，没有数据库账号和数据库密码的配置选项。   |
|        | 数据库密码 | 填入该数据库账号对应的密码。   |
|        | 连接方式  | 根据需求选择非加密连接或SSL安全连接。如果设置为SSL安全连接，您需要提前开启RDS实例的SSL加密功能，详情请参见 <a href="#">设置SSL加密</a> 。<br><br><b>说明</b> 目前仅中国内地及中国香港地域支持设置连接方式。                   |
| 目标实例信息 | 实例类型  | 固定为AnalyticDB for PostgreSQL，无需设置。   |
|        | 实例地区  | 购买数据同步实例时选择的目标实例地域信息，不可变更。   |
|        | 实例ID  | 选择目标云原生数据仓库AnalyticDB PostgreSQL实例ID。  |
|        | 数据库名称 | 填入同步目标表所属的数据库名称。   |
|        | 数据库账号 | 填入云原生数据仓库AnalyticDB PostgreSQL的初始账号，详情请参见 <a href="#">创建数据库账号</a> 。<br><br><b>说明</b> 您也可以填入具备RDS_SUPERUSER权限的账号，创建方法请参见 <a href="#">用户权限管理</a> 。 |
|        | 数据库密码 | 填入该数据库账号对应的密码。   |

7. 单击页面右下角的授权白名单并进入下一步。

**说明**

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置同步策略及同步对象。



| 类别     | 配置          | 说明  |
|--------|-------------|---|
| 同步策略配置 | 同步初始化       | 默认情况下，您需要同时选中 <b>结构初始化</b> 和 <b>全量数据初始化</b> 。预检查完成后，DTS会将源实例中待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。  |
|        | 目标已存在表的处理模式 | <ul style="list-style-type: none"> <li><b>清空目标表的数据</b><br/>在预检查阶段跳过<b>同名对象存在性检查</b>的检查项目。全量初始化之前将目标表的数据清空。适用于完成同步任务测试后的正式同步场景。</li> <li><b>忽略报错并继续执行</b><br/>在预检查阶段跳过<b>同名对象存在性检查</b>的检查项目。全量初始化时直接追加数据。适用于多张表同步到一张表的汇总同步场景。</li> </ul> |
|        | 同步操作类型      | 根据业务需求选择需要同步的操作类型： <ul style="list-style-type: none"> <li>Insert</li> <li>Update</li> <li>Delete</li> <li>AlterTable</li> </ul>   |

| 类别                              | 配置 | 说明  |
|---------------------------------|----|---|
| 选择同步对象                          | 无  | <p>在源库对象框中单击待同步的表，然后单击  图标将其移动至已选择对象框。</p> <div style="border: 1px solid #ccc; padding: 5px; background-color: #e6f2ff;"> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>同步对象的选择粒度为表。</li> <li>如果需要目标表中的列名称与源表不同，需要使用DTS的字段映射功能，详情请参见<a href="#">设置同步对象在目标实例中的名称</a>。</li> </ul> </div>   |
| 映射名称更改                          | 无  | <p>如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见<a href="#">库表列映射</a>。</p>   |
| 源表 DMS_ONLINE_DDL过程中是否复制临时表到目标库 | 无  | <p>如源库使用<a href="#">数据管理DMS (Data Management Service)</a> 执行Online DDL变更，您可以选择是否同步Online DDL变更产生的临时表数据。</p> <ul style="list-style-type: none"> <li><b>是</b>：同步Online DDL变更产生的临时表数据。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; background-color: #e6f2ff;"> <p><b>说明</b> Online DDL变更产生的临时表数据过大，可能会导致同步任务延迟。</p> </div> <ul style="list-style-type: none"> <li><b>否</b>：不同步Online DDL变更产生的临时表数据，只同步源库的原始DDL数据。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; background-color: #e6f2ff;"> <p><b>说明</b> 该方案会导致目标库锁表。</p> </div> |
| 源、目标库无法连接重试时间                   | 无  | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <div style="border: 1px solid #ccc; padding: 5px; background-color: #e6f2ff;"> <p><b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p> </div>   |

9. 设置待同步的表在云原生数据库AnalyticDB PostgreSQL中的主键列和分布列信息。



**说明** 当您在上一步中选择了结构初始化才会出现该页面。关于主键列和分布列的详细说明，请参见[表的约束定义](#)和[表分布键定义](#)。

10. 上述配置完成后，单击页面右下角的预检查并启动。

**说明**

- 在同步作业正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步作业。
- 如果预检查失败，单击具体检查项后的，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择**确认屏蔽**、**忽略告警项**并重新进行预检查，跳过告警检测项重新进行预检查。

11. 在预检查对话框中显示预检查通过后，关闭预检查对话框，同步作业将正式开始。

12. 等待同步作业的链路初始化完成，直至处于同步中状态。

您可以在数据同步页面，查看数据同步作业的状态。



| 实例ID/作业名称               | 状态  | 同步概况                                   | 付费方式 | 同步架构(全部) | 操作                  |
|-------------------------|-----|--|------|----------|---------------------|
| hangzhou-hangzhou-small | 同步中 | 耗时: 1376 毫秒<br>速率: 0.00RPS/(0.000MB/s) | 按量付费 | 单向同步     | 暂停同步   转包年包月   升级更多 |

## 7.6. 从ECS上的自建MySQL同步至云原生数据仓库AnalyticDB PostgreSQL

数据传输服务DTS（Data Transmission Service）支持将ECS上的自建MySQL数据同步至云原生数据仓库AnalyticDB PostgreSQL，帮助您轻松实现数据的流转，将企业数据集中分析。

### 前提条件

- ECS上的自建MySQL数据库版本为5.1、5.5、5.6、5.7或8.0版本。
- 源库必须开启binlog，同时建议创建一个账号用于数据同步，详情请参见[自建MySQL创建账号并设置binlog](#)。

**说明** 该账号需具备REPLICATION CLIENT、REPLICATION SLAVE、SHOW VIEW和所有同步对象的SELECT权限。

- 源库中待同步的数据表必须具备主键。
- 已创建目标云原生数据仓库AnalyticDB PostgreSQL实例，如未创建请参见[创建云原生数据仓库AnalyticDB PostgreSQL实例](#)。

### 注意事项

DTS在执行全量数据初始化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。因此您需要在执行数据同步前评估源库和目标库的性能，同时建议您在业务低峰期执行数据同步（例如源库和目标库的CPU负载在30%以下）。

### 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

### 同步限制

- 同步对象仅支持数据表。
- 不支持BIT、VARBIT、GEOMETRY、ARRAY、UUID、TSQUERY、TSVECTOR、TXID\_SNAPSHOT类型的数据同步。
- 暂不支持同步前缀索引，如果源库存在前缀索引可能导致数据同步失败。
- 在数据同步时，请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。

### 支持同步的SQL操作

- DML操作：INSERT、UPDATE、DELETE。
- DDL操作：ADD COLUMN。

 **说明** 不支持CREATE TABLE操作，如果您需要将新增的表作为同步对象，则需要执行[新增同步对象](#)操作。

### 支持的同步架构

- 1对1单向同步。
- 1对多单向同步。
- 多对1单向同步。

### 术语及概念对应关系

|          |                              |
|----------|------------------------------|
| MySQL    | 云原生数据仓库AnalyticDB PostgreSQL |
| Database | Schema                       |
| Table    | Table                        |

### 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

 **说明** 购买时，选择源实例为MySQL，目标实例为AnalyticDB for PostgreSQL，并选择同步拓扑为单向同步。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击[数据同步](#)。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击[配置同步链路](#)。
6. 配置同步作业的源实例及目标实例信息。

1.选择同步通道的源及目标实例
2.选择同步对象
3.预检查

同步作业名称：

**源实例信息**

实例类型：ECS上的自建数据库

实例地区：华东1（杭州）

\* ECS实例ID：

数据库类型：MySQL

\* 端口：

\* 数据库账号：

\* 数据库密码：

**目标实例信息**

实例类型：AnalyticDB for PostgreSQL

实例地区：华东1（杭州）

\* 实例ID：

\* 数据库名称：

\* 数据库账号：

\* 数据库密码：

取消
授权白名单并进入下一步

| 类别     | 配置             | 说明  |
|--------|----------------|---|
| 无      | 同步作业名称         | DTS会自动生成一个同步作业名称，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。  |
| 源实例信息  | 实例类型           | 选择ECS上的自建数据库。   |
|        | 实例地区           | 购买数据同步实例时选择的源实例地域信息，不可变更。   |
|        | 实例ID           | 选择作为自建MySQL所属的ECS实例ID。  |
|        | 数据库类型          | 固定为MySQL，不可变更。  |
|        | 端口             | 填入自建数据库的服务端口，默认为3306。   |
|        | 数据库账号          | 填入ECS上的自建MySQL的数据库账号。<br><div style="background-color: #e1f5fe; padding: 5px; border: 1px solid #cfe2f3; margin-top: 5px;"> <span style="font-size: 1.2em; color: #0070c0; float: left; margin-right: 5px;">?</span> <b>说明</b> 该账号需具备REPLICATION CLIENT、REPLICATION SLAVE、SHOW VIEW和所有同步对象的SELECT权限。                 </div> |
| 数据库密码  | 填入该数据库账号对应的密码。 |   |
| 目标实例信息 | 实例类型           | 固定为AnalyticDB for PostgreSQL，无需设置。  |
|        | 实例地区           | 购买数据同步实例时选择的目标实例地域信息，不可变更。  |
|        | 实例ID           | 选择目标云原生数据仓库AnalyticDB PostgreSQL实例ID。   |
|        | 数据库名称          | 填入同步目标表所属的数据库名称。  |

| 目标类别信息类别 | 配置    | 说明   |
|----------|-------|--|
|          | 数据库账号 | 填入云原生数据仓库AnalyticDB PostgreSQL的初始账号，详情请参见 <a href="#">创建数据库账号</a> 。<br><br><div style="border: 1px solid #ccc; padding: 5px; background-color: #e6f2ff;"> <p><span style="color: #00aaff;">?</span> <b>说明</b> 您也可以填入具备RDS_SUPERUSER权限的账号，创建方法请参见<a href="#">用户权限管理</a>。</p> </div> |
|          | 数据库密码 | 填入该数据库账号对应的密码。   |

7. 单击页面右下角的授权白名单并进入下一步。

? **说明**

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置同步策略及同步对象。

1.选择同步通道的源及目标实例
2.选择同步对象
3.预检查

同步初始化:  结构初始化  全量数据初始化

注: DTS全量任务运行期间, 不要清理DTS任务启动后源库产生的增量数据日志。源库如果过早清理日志, 可能会导致DTS增量任务失败

目标已存在表的处理模式:  预检查并报错拦截  清空目标表数据  忽略报错并继续执行

同步操作类型:  Insert  Update  Delete  Alter Table

**源库对象**

看全局搜索, 请先展开树

- dtstestdata
  - Tables
  - Views

全选

**已选择对象** (鼠标移到对象行, 点击编辑可修改对象名或过滤条件) [详情点我](#)

- dtstestdata (2个对象)
  - customer
  - order

全选

\*映射名称更改:  不进行库表名称批量更改  要进行库表名称批量更改

\*源表DMS\_ONLINE\_DDL过程中是否复制临时表到目标库:  是  否 ?

\*源库、目标库无法连接后的重试时间:  分钟 ?

\*为目标对象添加引号:  是  否

取消
上一步
下一步
预检查并启动

| 类别                              | 配置          | 说明  |
|---------------------------------|-------------|---|
| 同步策略配置                          | 同步初始化       | 默认情况下，您需要同时选中 <b>结构初始化</b> 和 <b>全量数据初始化</b> 。预检查完成后，DTS会将源实例中待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。  |
|                                 | 目标已存在表的处理模式 | <ul style="list-style-type: none"> <li>◦ <b>清空目标表的数据</b><br/>在预检查阶段跳过<b>同名对象存在性检查</b>的检查项目。全量初始化之前将目标表的数据清空。适用于完成同步任务测试后的正式同步场景。</li> <li>◦ <b>忽略报错并继续执行</b><br/>在预检查阶段跳过<b>同名对象存在性检查</b>的检查项目。全量初始化时直接追加数据。适用于多张表同步到一张表的汇总同步场景。</li> </ul>   |
|                                 | 同步操作类型      | 根据业务需求选择需要同步的操作类型： <ul style="list-style-type: none"> <li>◦ <b>Insert</b></li> <li>◦ <b>Update</b></li> <li>◦ <b>Delete</b></li> <li>◦ <b>AlterTable</b></li> </ul>   |
| 选择同步对象                          | 无           | 在源库对象框中单击待同步的表，然后单击  图标将其移动至已选择对象框。 <div style="background-color: #e6f2ff; padding: 5px; margin-top: 10px;"> <p> <b>说明</b></p> <ul style="list-style-type: none"> <li>◦ 同步对象的选择粒度为表。</li> <li>◦ 如果需要目标表中的列名称与源表不同，需要使用DTS的字段映射功能，详情请参见<a href="#">设置同步对象在目标实例中的名称</a>。</li> </ul> </div>  |
| 映射名称更改                          | 无           | 如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见 <a href="#">库表列映射</a> 。  |
| 源表 DMS_ONLINE_DDL过程中是否复制临时表到目标库 | 无           | 如源库使用 <a href="#">数据管理DMS (Data Management Service)</a> 执行Online DDL变更，您可以选择是否同步Online DDL变更产生的临时表数据。 <ul style="list-style-type: none"> <li>◦ <b>是</b>：同步Online DDL变更产生的临时表数据。</li> </ul> <div style="background-color: #e6f2ff; padding: 5px; margin-top: 10px;"> <p> <b>说明</b> Online DDL变更产生的临时表数据过大，可能会导致同步任务延迟。</p> </div> <ul style="list-style-type: none"> <li>◦ <b>否</b>：不同步Online DDL变更产生的临时表数据，只同步源库的原始DDL数据。</li> </ul> <div style="background-color: #e6f2ff; padding: 5px; margin-top: 10px;"> <p> <b>说明</b> 该方案会导致目标库锁表。</p> </div> |

| 类别            | 配置 | 说明  |
|---------------|----|---|
| 源、目标库无法连接重试时间 | 无  | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <p><b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p> |

9. 设置待同步的表在云原生数据仓库AnalyticDB PostgreSQL中的主键列和分布列信息。



**说明** 当您在上一步中选择了结构初始化才会出现该页面。关于主键列和分布列的详细说明，请参见表的约束定义和表分布键定义。

10. 上述配置完成后，单击页面右下角的预检查并启动。

**说明**

- 在同步作业正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步作业。
- 如果预检查失败，单击具体检查项后的 ，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择确认屏蔽、忽略告警项并重新进行预检查，跳过告警检测项重新进行预检查。

11. 在预检查对话框中显示预检查通过后，关闭预检查对话框，同步作业将正式开始。

12. 等待同步作业的链路初始化完成，直至处于同步中状态。

您可以在数据同步页面，查看数据同步作业的状态。



## 7.7. 从通过专线、VPN网关或智能接入网关接入的自建MySQL同步至云原生数据仓库AnalyticDB PostgreSQL

数据传输服务DTS（Data Transmission Service）支持将通过专线、VPN网关或智能接入网关接入的自建MySQL数据同步至云原生数据仓库AnalyticDB PostgreSQL，帮助您轻松实现数据的流转，将企业数据集中分析。

## 前提条件

- 自建MySQL数据库版本为5.1、5.5、5.6、5.7或8.0版本。
- 源库中待同步的数据表必须具备主键。
- 源库必须开启binlog，同时建议您创建一个账号用于数据同步，详情请参见[自建MySQL创建账号并设置binlog](#)。

 **说明** 该账号需具备REPLICATION CLIENT、REPLICATION SLAVE、SHOW VIEW和所有同步对象的SELECT权限。

- 自建数据库所属的本地网络已通过专线、VPN网关或智能接入网关的方式接入至阿里云专有网络，并且已经[通过VPN网关实现本地IDC与DTS云服务互通](#)。

 **说明** 相关接入方案请参见[本地IDC接入至阿里云方案概览](#)，本文不做详细介绍。

- 已创建目标云原生数据仓库AnalyticDB PostgreSQL实例，如未创建请参见[创建云原生数据仓库AnalyticDB PostgreSQL实例](#)。

## 注意事项

DTS在执行全量数据初始化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。因此您需要在执行数据同步前评估源库和目标库的性能，同时建议您在业务低峰期执行数据同步（例如源库和目标库的CPU负载在30%以下）。

## 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

## 同步限制

- 同步对象仅支持数据表。
- 不支持BIT、VARBIT、GEOMETRY、ARRAY、UUID、TSQUERY、TSVECTOR、TXID\_SNAPSHOT类型的数据同步。
- 暂不支持同步前缀索引，如果源库存在前缀索引可能导致数据同步失败。
- 在数据同步时，请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。

## 支持同步的SQL操作

- DML操作：INSERT、UPDATE、DELETE。
- DDL操作：ADD COLUMN。

 **说明** 不支持CREATE TABLE操作，如果您需要将新增的表作为同步对象，则需要执行[新增同步对象](#)操作。

## 支持的同步架构

- 1对1单向同步。
- 1对多单向同步。
- 多对1单向同步。

## 术语及概念对应关系

|          |                              |
|----------|------------------------------|
| MySQL    | 云原生数据仓库AnalyticDB PostgreSQL |
| Database | Schema                       |
| Table    | Table                        |

### 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

**说明** 购买时，选择源实例为MySQL，目标实例为AnalyticDB for PostgreSQL，并选择同步拓扑为单向同步。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击[数据同步](#)。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击[配置同步链路](#)。
6. 配置同步作业的源实例及目标实例信息。

1.选择同步通道的源及目标实例
2.选择同步对象
3.预检查

同步作业名称：

**源实例信息**

实例类型：通过专线/VPN网关/智能网关接入的自建数据库

实例地区：华东1（杭州）

\* 对端专有网络：

数据库类型：MySQL

\* IP地址：

\* 端口：

\* 数据库账号：

\* 数据库密码：

**目标实例信息**

实例类型：AnalyticDB for PostgreSQL

实例地区：华东1（杭州）

\* 实例ID：

\* 数据库名称：

\* 数据库账号：

\* 数据库密码：

取消
授权白名单并进入下一步

| 类别 | 配置     | 说明  |
|----|--------|---|
| 无  | 同步作业名称 | DT S会自动生成一个同步作业名称，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。 |
|    | 实例类型   | 选择通过专线/VPN网关/智能接入网关接入的自建数据库。                    |
|    | 实例地区   | 购买数据同步实例时选择的源实例地域信息，不可变更。                       |
|    | 对端专有网络 | 选择自建数据库接入的VPC ID。                               |

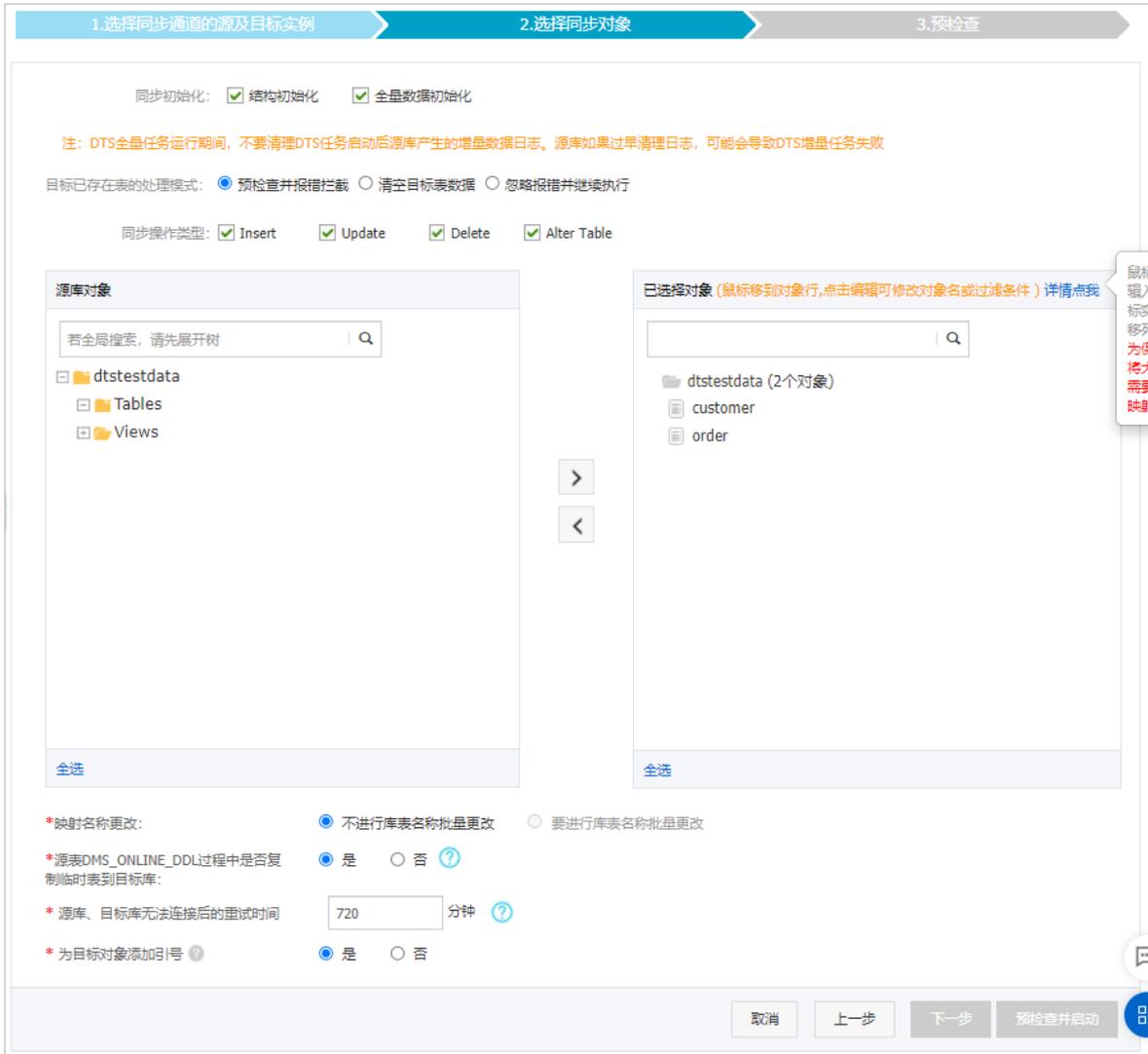
| 类别     | 配置    | 说明   |
|--------|-------|--|
| 源实例信息  | 数据库类型 | 固定为MySQL，不可变更。   |
|        | IP地址  | 填入自建MySQL数据库的服务器IP地址。  |
|        | 端口    | 填入自建数据库的服务端口，默认为3306。  |
|        | 数据库账号 | 填入自建MySQL的数据库账号。<br><br><span style="color: #00AEEF;">?</span> 说明 该账号需具备REPLICATION CLIENT、REPLICATION SLAVE、SHOW VIEW和所有同步对象的SELECT权限。  |
|        | 数据库密码 | 填入该数据库账号对应的密码。   |
| 目标实例信息 | 实例类型  | 固定为AnalyticDB for PostgreSQL，无需设置。   |
|        | 实例地区  | 购买数据同步实例时选择的目标实例地域信息，不可变更。   |
|        | 实例ID  | 选择目标云原生数据仓库AnalyticDB PostgreSQL实例ID。  |
|        | 数据库名称 | 填入同步目标表所属的数据库名称。   |
|        | 数据库账号 | 填入云原生数据仓库AnalyticDB PostgreSQL的初始账号，详情请参见 <a href="#">创建数据库账号</a> 。<br><br><span style="color: #00AEEF;">?</span> 说明 您也可以填入具备RDS_SUPERUSER权限的账号，创建方法请参见 <a href="#">用户权限管理</a> 。 |
|        | 数据库密码 | 填入该数据库账号对应的密码。   |

#### 7. 单击页面右下角的授权白名单并进入下一步。

##### ? 说明

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

#### 8. 配置同步策略及同步对象。



| 类别     | 配置          | 说明  |
|--------|-------------|---|
| 同步策略配置 | 同步初始化       | 默认情况下，您需要同时选中 <b>结构初始化</b> 和 <b>全量数据初始化</b> 。预检查完成后，DTS会将源实例中待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。  |
|        | 目标已存在表的处理模式 | <ul style="list-style-type: none"> <li><b>清空目标表的数据</b><br/>在预检查阶段跳过<b>同名对象存在性检查</b>的检查项目。全量初始化之前将目标表的数据清空。适用于完成同步任务测试后的正式同步场景。</li> <li><b>忽略报错并继续执行</b><br/>在预检查阶段跳过<b>同名对象存在性检查</b>的检查项目。全量初始化时直接追加数据。适用于多张表同步到一张表的汇总同步场景。</li> </ul> |
|        | 同步操作类型      | 根据业务需求选择需要同步的操作类型： <ul style="list-style-type: none"> <li>Insert</li> <li>Update</li> <li>Delete</li> <li>AlterTable</li> </ul>   |

| 类别                              | 配置 | 说明  |
|---------------------------------|----|---|
| 选择同步对象                          | 无  | <p>在源库对象框中单击待同步的表，然后单击  图标将其移动至已选择对象框。</p> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>同步对象的选择粒度为表。</li> <li>如果需要目标表中的列名称与源表不同，需要使用DTS的字段映射功能，详情请参见<a href="#">设置同步对象在目标实例中的名称</a>。</li> </ul>   |
| 映射名称更改                          | 无  | <p>如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见<a href="#">库表列映射</a>。</p>   |
| 源表 DMS_ONLINE_DDL过程中是否复制临时表到目标库 | 无  | <p>如源库使用<a href="#">数据管理DMS (Data Management Service)</a> 执行Online DDL变更，您可以选择是否同步Online DDL变更产生的临时表数据。</p> <ul style="list-style-type: none"> <li><b>是</b>：同步Online DDL变更产生的临时表数据。</li> </ul> <p><b>说明</b> Online DDL变更产生的临时表数据过大，可能会导致同步任务延迟。</p> <ul style="list-style-type: none"> <li><b>否</b>：不同步Online DDL变更产生的临时表数据，只同步源库的原始DDL数据。</li> </ul> <p><b>说明</b> 该方案会导致目标库锁表。</p> |
| 源、目标库无法连接重试时间                   | 无  | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <p><b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p>   |

9. 设置待同步的表在云原生数据库AnalyticDB PostgreSQL中的主键列和分布列信息。



**说明** 当您在上一步中选择了结构初始化才会出现该页面。关于主键列和分布列的详细说明，请参见[表的约束定义](#)和[表分布键定义](#)。

10. 上述配置完成后，单击页面右下角的**预检查并启动**。

**说明**

- 在同步作业正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步作业。
- 如果预检查失败，单击具体检查项后的，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择**确认屏蔽**、**忽略告警项**并重新进行预检查，跳过告警检测项重新进行预检查。

11. 在**预检查**对话框中显示**预检查通过**后，关闭**预检查**对话框，同步作业将正式开始。

12. 等待同步作业的链路初始化完成，直至处于**同步中**状态。

您可以在**数据同步**页面，查看数据同步作业的状态。



## 7.8. 从RDS MySQL同步至DataHub

阿里云流式数据服务DataHub是流式数据（Streaming Data）的处理平台，提供对流式数据的发布、订阅和分发功能，让您可以轻松构建基于流式数据的分析和应用。通过数据传输服务DTS（Data Transmission Service），您可以将RDS MySQL同步至DataHub，帮助您快速实现使用流计算等大数据产品对数据实时分析。

### 前提条件

- DataHub实例的地域为华东1、华东2、华北2或华南1。
- DataHub实例中，已创建用作接收同步数据的项目（Project），详情请参见[创建项目](#)。
- RDS MySQL中待同步的表需具备主键或唯一约束。

### 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

### 功能限制

- 不支持全量数据初始化，即DTS不会将源RDS实例中同步对象的存量数据同步至目标DataHub实例。
- 仅支持表级别的数据同步。
- 不支持新增列的数据同步，即源数据表新增了某个列，该列的数据不会同步至目标DataHub实例。
- 数据同步的过程中，请勿对源库中待同步的表执行DDL变更，否则会导致同步失败。

### 支持同步的SQL操作

INSERT、UPDATE、DELETE。

### 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

**说明** 购买时，选择源实例为MySQL、目标实例为DataHub，并选择同步拓扑为单向同步。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击[数据同步](#)。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击[配置同步链路](#)。
6. 配置同步作业的源实例及目标实例信息。

| 类别    | 配置     | 说明  |
|-------|--------|---|
| 无     | 同步作业名称 | DTS会自动生成一个同步作业名称，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。  |
| 源实例信息 | 实例类型   | 根据源库的部署位置进行选择，本文以RDS实例为例介绍配置流程。<br><b>说明</b> 如果源库为自建MySQL数据库，您还需要执行相应的准备工作，详情请参见 <a href="#">准备工作概览</a> 。 |
|       | 实例地区   | 购买数据同步实例时选择的源实例地域信息，不可变更。   |
|       | 实例ID   | 选择作为数据同步源的RDS实例ID。  |
|       | 数据库账号  | 填入源RDS的数据库账号。<br><b>说明</b> 当源RDS实例的数据库类型为MySQL 5.5或MySQL 5.6时，无需配置数据库账号和数据库密码。                            |
|       | 数据库密码  | 填入数据库账号对应的密码。   |

| 类别     | 配置      | 说明   |
|--------|---------|--|
|        | 连接方式    | 根据需求选择非加密连接或SSL安全连接。如果设置为SSL安全连接，您需要提前开启RDS实例的SSL加密功能，详情请参见 <a href="#">设置SSL加密</a> 。<br><div style="border: 1px solid #ccc; padding: 5px; margin-top: 5px;">  <b>注意</b> 目前仅中国内地及中国香港地域支持设置连接方式。           </div> |
| 目标实例信息 | 实例类型    | 固定为DataHub，不可变更。   |
|        | 实例地区    | 购买数据同步实例时选择的目标实例地域信息，不可变更。   |
|        | Project | 选择DataHub实例的Project。   |

7. 单击页面右下角的授权白名单并进入下一步。

**说明**

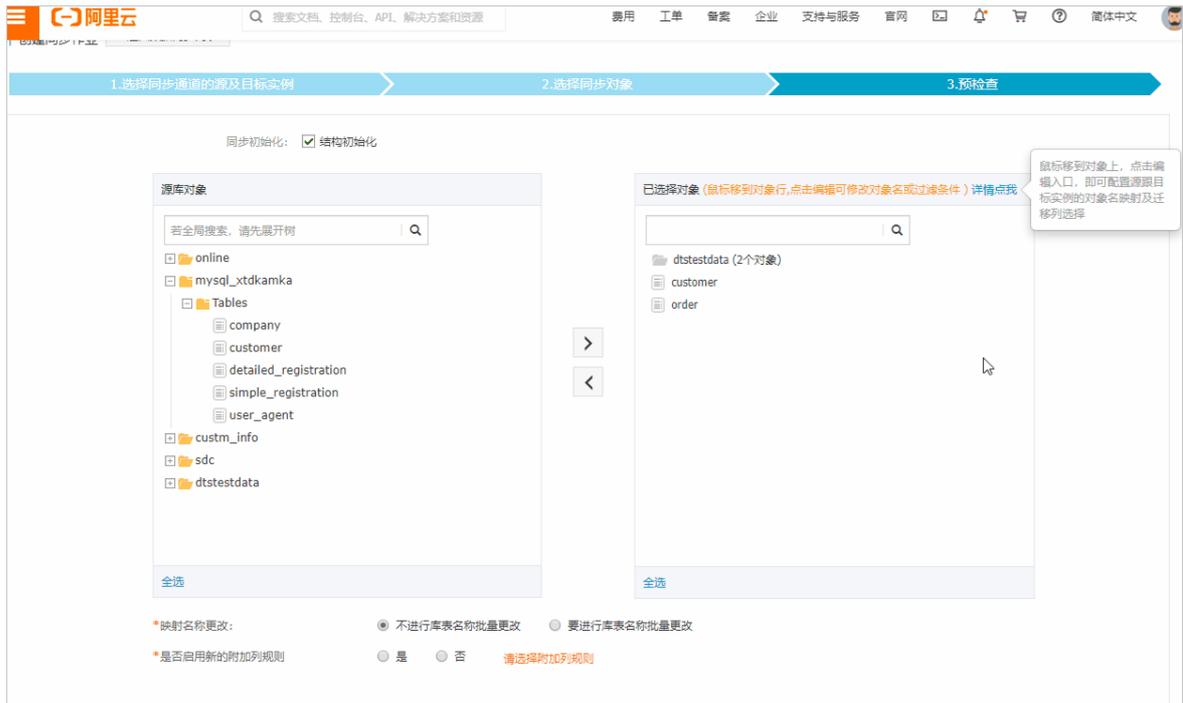
- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置同步策略和同步对象。



| 配置                             | 说明  |
|--------------------------------|---|
| 同步初始化                          | <p>勾选结构初始化。</p> <p> <b>说明</b> 勾选结构初始化后，在数据同步作业的初始化阶段，DTS会将同步对象的结构信息（例如表结构）同步至目标DataHub实例。</p>  |
| 选择同步对象                         | <p>在源库对象框中单击待迁移的对象，然后单击  将其移动至已选择对象框。</p> <p> <b>说明</b></p> <ul style="list-style-type: none"> <li>同步对象的选择粒度为表。</li> <li>默认情况下，同步对象的名称保持不变。如果您需要改变同步对象在目标实例中的名称，需要使用对象名映射功能，详情请参见<a href="#">设置同步对象在目标实例中的名称</a>。</li> </ul>  |
| 选择附加列规则                        | <p>DTS在将数据同步到DataHub时，会在同步的目标Topic中添加一些附加列。如果附加列和目标Topic中已有的列出现名称冲突将会导致数据同步失败。您需要根据业务需求选择是否启用新的附加列规则为是或否。</p> <p> <b>警告</b> 在选择附加列规则前，您需要评估附加列和目标Topic中已有的列是否会出现名称冲突。关于附加列的规则和定义说明，请参见<a href="#">附加列名称和定义说明</a>。</p>  |
| 映射名称更改                         | <p>如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见<a href="#">库表列映射</a>。</p>   |
| 源表DMS_ONLINE_DDL过程中是否复制临时表到目标库 | <p>如源库使用<a href="#">数据管理DMS (Data Management Service)</a> 执行Online DDL变更，您可以选择是否同步Online DDL变更产生的临时表数据。</p> <ul style="list-style-type: none"> <li>是：同步Online DDL变更产生的临时表数据。</li> </ul> <p> <b>说明</b> Online DDL变更产生的临时表数据过大，可能会导致同步任务延迟。</p> <ul style="list-style-type: none"> <li>否：不同步Online DDL变更产生的临时表数据，只同步源库的原始DDL数据。</li> </ul> <p> <b>说明</b> 该方案会导致目标库锁表。</p> |
| 源、目标库无法连接重试时间                  | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <p> <b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p>   |

- （可选）将鼠标指针放置在已选择对象框中待同步的Topic名上，单击对象后出现的编辑，然后在弹出的对话框中设置Shardkey（即用于分区的key）。



10. 上述配置完成后，单击页面右下角的预检查并启动。

**说明**

- 在同步作业正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步作业。
- 如果预检查失败，单击具体检查项后的，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择**确认屏蔽**、**忽略告警项**并重新进行预检查，跳过告警检测项重新进行预检查。

11. 在预检查对话框中显示预检查通过后，关闭预检查对话框，同步作业将正式开始。

12. 等待同步作业的链路初始化完成，直至处于同步中状态。

您可以在数据同步页面，查看数据同步作业的状态。



### Topic结构定义说明

DTS在将数据变更同步至DataHub实例的Topic时，目标Topic中除了存储变更数据外，还会新增一些附加列用于存储元信息，示例如下。

**说明** 本案例中的业务字段为 `id`、`name`、`address`，由于在配置数据同步时选用的是旧版附加列规则，DTS会为业务字段添加 `dtst_` 的前缀。

| dts_id | dts_name | dts_address | dts_record_id      | dts_operation_flag | dts_instance_id | dts_db_name | dts_table_name | dts_utc_timestamp | dts_before_flag | dts_after_flag |
|--------|----------|-------------|--------------------|--------------------|-----------------|-------------|----------------|-------------------|-----------------|----------------|
| 10006  |          | 曲靖市         | 157483213000000000 | U                  |                 | dtstestdata | customer       | 1574832130        | Y               | N              |
| 10006  |          | 杭州市         | 157483213000000000 | U                  |                 | dtstestdata | customer       | 1574832130        | N               | Y              |
| 10009  |          | 马鞍山市        | 157483291900000000 | D                  |                 | dtstestdata | customer       | 1574832919        | Y               | N              |
| 10112  |          | 北京市         | 157483291900000000 | I                  |                 | dtstestdata | customer       | 1574832919        | N               | Y              |

结构定义说明:

| 旧版附加列名称                         | 新版附加列名称                                      | 数据类型   | 说明   |
|---------------------------------|--|--------|--|
| <code>dts_record_id</code>      | <code>new_dts_sync_dts_record_id</code>      | String | 增量日志的记录ID, 为该日志唯一标识。<br><div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>• 正常情况下是全局唯一自增的, 在容灾的情况下会有回退且无法保证自增和唯一。</li> <li>• 如果增量日志的操作类型为UPDATE, 那么增量更新会被拆分成两条记录 (分别记录更新前和更新后的值), 且 <code>dts_record_id</code> 的值相同。</li> </ul> </div> |
| <code>dts_operation_flag</code> | <code>new_dts_sync_dts_operation_flag</code> | String | 操作类型, 取值: <ul style="list-style-type: none"> <li>• I: INSERT操作。</li> <li>• D: DELETE操作。</li> <li>• U: UPDATE操作。</li> </ul>   |
| <code>dts_instance_id</code>    | <code>new_dts_sync_dts_instance_id</code>    | String | 数据库的server ID。暂不支持显示实际的值, 目前固定为 <code>null</code> 。  |
| <code>dts_db_name</code>        | <code>new_dts_sync_dts_db_name</code>        | String | 数据库名称。   |
| <code>dts_table_name</code>     | <code>new_dts_sync_dts_table_name</code>     | String | 表名。  |
| <code>dts_utc_timestamp</code>  | <code>new_dts_sync_dts_utc_timestamp</code>  | String | 操作时间戳, 即binlog的时间戳 (UTC 时间)。   |
| <code>dts_before_flag</code>    | <code>new_dts_sync_dts_before_flag</code>    | String | 所有列的值是否更新前的值, 取值: Y或N。   |
| <code>dts_after_flag</code>     | <code>new_dts_sync_dts_after_flag</code>     | String | 所有列的值是否更新后的值, 取值: Y或N。   |

### 关于dts\_before\_flag和dts\_after\_flag的补充说明

对于不同的操作类型, 增量日志中的 `dts_before_flag` 和 `dts_after_flag` 定义如下:

• INSERT

当操作类型为INSERT时, 所有列的值为新插入的值, 即为更新后的值, 所以 `dts_before_flag` 取值为N, `dts_after_flag` 取值为Y, 示例如下。

| dts_id | dts_name | dts_address | dts_record_id      | dts_operation_flag | dts_instance_id | dts_db_name | dts_table_name | dts_utc_timestamp | dts_before_flag | dts_after_flag |
|--------|----------|-------------|--------------------|--------------------|-----------------|-------------|----------------|-------------------|-----------------|----------------|
| 10112  |          | 北京市         | 157483291900000000 | I                  |                 | dtstestdata | customer       | 1574832919        | N               | Y              |

• UPDATE

当操作类型为UPDATE时, DTS会将UPDATE操作拆为两条增量日志。这两条增量日志的 `dts_record_id`、`dts_operation_flag` 及 `dts_utc_timestamp` 对应的值相同。

第一条增量日志记录了更新前的值，所以 `dts_before_flag` 取值为Y，`dts_after_flag` 取值为N。第二条增量日志记录了更新后的值，所以 `dts_before_flag` 取值为N，`dts_after_flag` 取值为Y，示例如下。

| dts_id | dts_name | dts_address | dts_record_id       | dts_operation_flag | dts_instance_id | dts_db_name | dts_table_name | dts_utc_timestamp | dts_before_flag | dts_after_flag |
|--------|----------|-------------|---------------------|--------------------|-----------------|-------------|----------------|-------------------|-----------------|----------------|
| 10006  |          | 曲靖市         | 1574832130000000000 | U                  |                 | dtstestdata | customer       | 1574832130        | Y               | N              |
| 10006  |          | 杭州市         | 1574832130000000000 | U                  |                 | dtstestdata | customer       | 1574832130        | N               | Y              |

• DELETE

当操作类型为DELETE时，增量日志中所有的列值为被删除的值，即列值不变，所以 `dts_before_flag` 取值为Y，`dts_after_flag` 取值为N，示例如下。

| dts_id | dts_name | dts_address | dts_record_id       | dts_operation_flag | dts_instance_id | dts_db_name | dts_table_name | dts_utc_timestamp | dts_before_flag | dts_after_flag |
|--------|----------|-------------|---------------------|--------------------|-----------------|-------------|----------------|-------------------|-----------------|----------------|
| 10009  |          | 马鞍山市        | 1574832919000000000 | D                  |                 | dtstestdata | customer       | 1574832919        | Y               | N              |

后续操作

配置完数据同步作业后，您可以对同步到DataHub实例中的数据执行计算分析。更多详情，请参见[阿里云实时计算](#)。

## 7.9. 从ECS上的自建MySQL同步至Elasticsearch

阿里云Elasticsearch完全兼容开源Elasticsearch的功能及Security、Machine Learning、Graph、APM等商业功能，致力于数据分析、数据搜索等场景服务，支持企业级权限管控、安全监控告警、自动报表生成等功能。通过数据传输服务DTS（Data Transmission Service），您可以将ECS上的自建MySQL同步至Elasticsearch，帮助您快速构建数据。

前提条件

- 已创建目标Elasticsearch实例（5.5、5.6、6.3、6.7或7.4版本），详情请参见[开通阿里云Elasticsearch服务](#)。
- 自建MySQL数据库版本为5.1、5.5、5.6、5.7或8.0版本。

注意事项

- DTS在执行全量数据初始化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。因此您需要在执行数据同步前评估源库和目标库的性能，同时建议您在业务低峰期执行数据同步（例如源库和目标库的CPU负载在30%以下）。
- 不支持同步DDL操作，如果源库中待同步的表在同步的过程中，已经执行了DDL操作，您需要先移除同步对象，然后在Elasticsearch实例中移除该表对应的索引，最后新增同步对象。详情请参见[移除同步对象](#)和[新增同步对象](#)。
- 如果源库中待同步的表需要执行增加列的操作，您只需先在Elasticsearch实例中修改对应表的mapping，然后在MySQL数据库中执行相应的DDL操作，最后暂停并启动DTS同步实例。

费用说明

|               |                                 |
|---------------|---------------------------------|
| 同步类型          | 链路配置费用                          |
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

支持同步的SQL操作

INSERT、DELETE、UPDATE

数据类型映射关系

由于MySQL和Elasticsearch实例支持的数据类型不同，数据类型无法一一对应。所以DTS在进行结构初始化时，会根据目标库支持的数据类型进行类型映射，详情请参见[结构初始化涉及的数据类型映射关系](#)。

准备工作

[为自建MySQL创建账号并设置binlog](#)

### 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

? **说明** 购买时，选择源实例为MySQL、目标实例为Elasticsearch，并选择同步拓扑为单向同步。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击[数据同步](#)。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击[配置同步链路](#)。
6. 配置同步通道的源实例及目标实例信息。

1.选择同步通道的源及目标实例
2.选择同步对象
3.高级设置
4.预检查

同步作业名称:

**源实例信息**

实例类型:

实例地区: 华东1 (杭州)

\* ECS实例ID:

数据库类型: MySQL

\* 端口:

\* 数据库账号:

\* 数据库密码:

**目标实例信息**

实例类型: Elasticsearch

实例地区: 华东1 (杭州)

\* Elasticsearch:

\* 数据库账号:

\* 数据库密码:

| 类别    | 配置      | 说明   |
|-------|---------|--|
| 无     | 同步作业名称  | DTS会自动生成一个同步作业名称，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。   |
| 源实例信息 | 实例类型    | 选择ECS上的自建数据库。  |
|       | 实例地区    | 购买数据同步实例时选择的源实例地域信息，不可变更。  |
|       | ECS实例ID | 选择自建MySQL数据库所属的ECS实例ID。  |
|       | 数据库类型   | 固定为MySQL，不可变更。   |
|       | 端口      | 填入自建MySQL的数据库服务端口。   |
|       | 数据库账号   | 填入自建MySQL的数据库账号。<br><br><span style="font-size: 1.2em;">?</span> <b>说明</b> 该账号需具备REPLICATION CLIENT、REPLICATION SLAVE、SHOW VIEW和所有同步对象的SELECT权限。 |

| 类别     | 配置            | 说明                                    |
|--------|---------------|---------------------------------------|
|        | 数据库密码         | 填入数据库账号的密码。                           |
| 目标实例信息 | 实例类型          | 固定为Elasticsearch，不可变更。                |
|        | 实例地区          | 购买数据同步实例时选择的目标实例地域信息，不可变更。            |
|        | Elasticsearch | 选择Elasticsearch实例ID。                  |
|        | 数据库账号         | 填入连接Elasticsearch实例的数据库账号，默认为elastic。 |
|        | 数据库密码         | 填入该数据库账号的密码。                          |

7. 单击页面右下角的授权白名单并进入下一步。

**说明**

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置索引名称、目标已存在表的处理模式和同步对象。

| 配置                             | 说明  |
|--------------------------------|---|
| 索引名称                           | <ul style="list-style-type: none"> <li>◦ <b>表名</b><br/>选择为表名后，在目标Elasticsearch实例中创建的索引名称和表名一致，在本案例中即为customer。</li> <li>◦ <b>库名_表名</b><br/>选择为库名_表名后，在目标Elasticsearch实例中创建的索引名称为库名_表名，在本案例中即为dtstestdata_customer。</li> </ul>   |
| 目标已存在表的处理模式                    | <ul style="list-style-type: none"> <li>◦ <b>预检查并报错拦截</b>：检查目标数据库中是否有同名的索引。如果目标数据库中没有同名的索引，则通过该检查项目；如果目标数据库中有同名的索引，则在预检查阶段提示错误，数据同步作业不会被启动。<br/><br/> <div style="border: 1px solid #ccc; padding: 5px; background-color: #e6f2ff;"> <p> <b>说明</b> 如果目标库中同名的索引不方便删除或重命名，您可以<b>设置同步对象在目标实例中的名称</b>来避免表名冲突。</p> </div> </li> <li>◦ <b>忽略报错并继续执行</b>：跳过目标数据库中是否有同名索引的检查项。<br/><br/> <div style="border: 1px solid #ccc; padding: 5px; background-color: #fff9c4;"> <p> <b>警告</b> 选择为忽略报错并继续执行，可能导致数据不一致，给业务带来风险，例如：</p> <ul style="list-style-type: none"> <li>■ mapping结构一致的情况下，如果在目标库遇到与源库主键的值相同的记录，在初始化阶段会保留目标库中的该条记录；在增量同步阶段则会覆盖目标库的该条记录。</li> <li>■ mapping结构不一致的情况下，可能会导致无法初始化数据、只能同步部分列的数据或同步失败。</li> </ul> </div> </li> </ul> |
| 选择同步对象                         | <p>在<b>源库对象</b>框中单击待同步的对象，然后单击  将其移动至<b>已选择对象</b>框。</p> <p>同步对象的选择粒度为库、表。</p>  |
| 映射名称更改                         | <p>如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见<b>库表列映射</b>。</p>  |
| 源表DMS_ONLINE_DDL过程中是否复制临时表到目标库 | <p>如源库使用<b>数据管理DMS (Data Management Service)</b>执行Online DDL变更，您可以选择是否同步Online DDL变更产生的临时表数据。</p> <ul style="list-style-type: none"> <li>◦ <b>是</b>：同步Online DDL变更产生的临时表数据。<br/><br/> <div style="border: 1px solid #ccc; padding: 5px; background-color: #e6f2ff;"> <p> <b>说明</b> Online DDL变更产生的临时表数据过大，可能会导致同步任务延迟。</p> </div> </li> <li>◦ <b>否</b>：不同步Online DDL变更产生的临时表数据，只同步源库的原始DDL数据。<br/><br/> <div style="border: 1px solid #ccc; padding: 5px; background-color: #e6f2ff;"> <p> <b>说明</b> 该方案会导致目标库锁表。</p> </div> </li> </ul>  |
| 源、目标库无法连接重试时间                  | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <div style="border: 1px solid #ccc; padding: 5px; background-color: #e6f2ff;"> <p> <b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p> </div>   |

9. 在**已选择对象**区域框中，将鼠标指针放置在待同步的表上，并单击表名后出现的**编辑**，设置该表在目标

Elasticsearch实例中的索引名称、Type名称等信息。

✕

**编辑表**

注意：编辑表名或列名后，目标数据库的表名列名将为修改后的名称。

\* 索引名称:  ⓘ

\* Type名称:  ⓘ

过滤条件:

设置\_routing:  是  否 ⓘ

\_id取值:

| <input checked="" type="checkbox"/> | 列名                                   | 类型            | 字段参数    | 字段参数值   |                      |
|-------------------------------------|--------------------------------------|---------------|---------|---------|----------------------|
| <input checked="" type="checkbox"/> | <input type="text" value="Name"/>    | varchar(32... | index ▾ | false ▾ | <a href="#">添加参数</a> |
| <input checked="" type="checkbox"/> | <input type="text" value="address"/> | varchar(32... | index ▾ | false ▾ | <a href="#">添加参数</a> |
| <input checked="" type="checkbox"/> | <input type="text" value="id"/>      | int(11)       | index ▾ | false ▾ |                      |

[确定](#)

| 配置     | 说明  |
|--------|---|
| 索引名称   | 详情请参见 <a href="#">Elasticsearch基本概念</a> 。   |
| Type名称 | <p><b>警告</b></p> <ul style="list-style-type: none"> <li>◦ 索引名称和Type名称支持的特殊字符仅为下划线（_）。</li> <li>◦ 如需将多个相同结构的源表同步到一个目标对象，您需要重复本步骤为这些表设置相同的索引名称和Type名称。</li> </ul> |
| 过滤条件   | 您可以设置SQL过滤条件，过滤待同步的数据，只有满足过滤条件的数据才会被同步到目标实例，详情请参见 <a href="#">通过SQL条件过滤任务数据</a> 。   |
| 是否分区   | 选择是否设置分区，如果您选择为是，您还需要设置分区列和分区数量。  |

| 配置         | 说明   |
|------------|--|
| 设置_routing | 设置_routing可以将文档路由存储在目标Elasticsearch实例的指定分片上，详情请参见_routing。<br><ul style="list-style-type: none"> <li>选择为是，您可以自定义列进行路由。</li> <li>选择为否，则用_id进行路由。</li> </ul> <div style="background-color: #e6f2ff; padding: 5px; margin-top: 10px;"> <p><span>?</span> 说明 创建的目标Elasticsearch实例为7.4版本时，您必须选择为否。</p> </div> |
| _id取值      | <ul style="list-style-type: none"> <li><b>表的主键列</b><br/>联合主键合并为一列。</li> <li><b>业务主键</b><br/>如果选择为业务主键，那么您还需要设置对应的业务主键列。</li> </ul>   |
| 添加参数       | 选择所需的字段参数和字段参数值，字段参数及取值介绍请参见Elasticsearch官方文档。<br><div style="background-color: #e6f2ff; padding: 5px; margin-top: 10px;"> <p><span>?</span> 说明 DTS目前仅支持可以选中的参数。</p> </div>  |

10. 上述配置完成后，单击页面右下角的预检查并启动。

? 说明

- 在同步作业正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步作业。
- 如果预检查失败，单击具体检查项后的，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择确认屏蔽、忽略告警项并重新进行预检查，跳过告警检测项重新进行预检查。

11. 在预检查对话框中显示预检查通过后，关闭预检查对话框，同步作业将正式开始。

12. 等待同步作业的链路初始化完成，直至处于同步中状态。

您可以在数据同步页面，查看数据同步作业的状态。

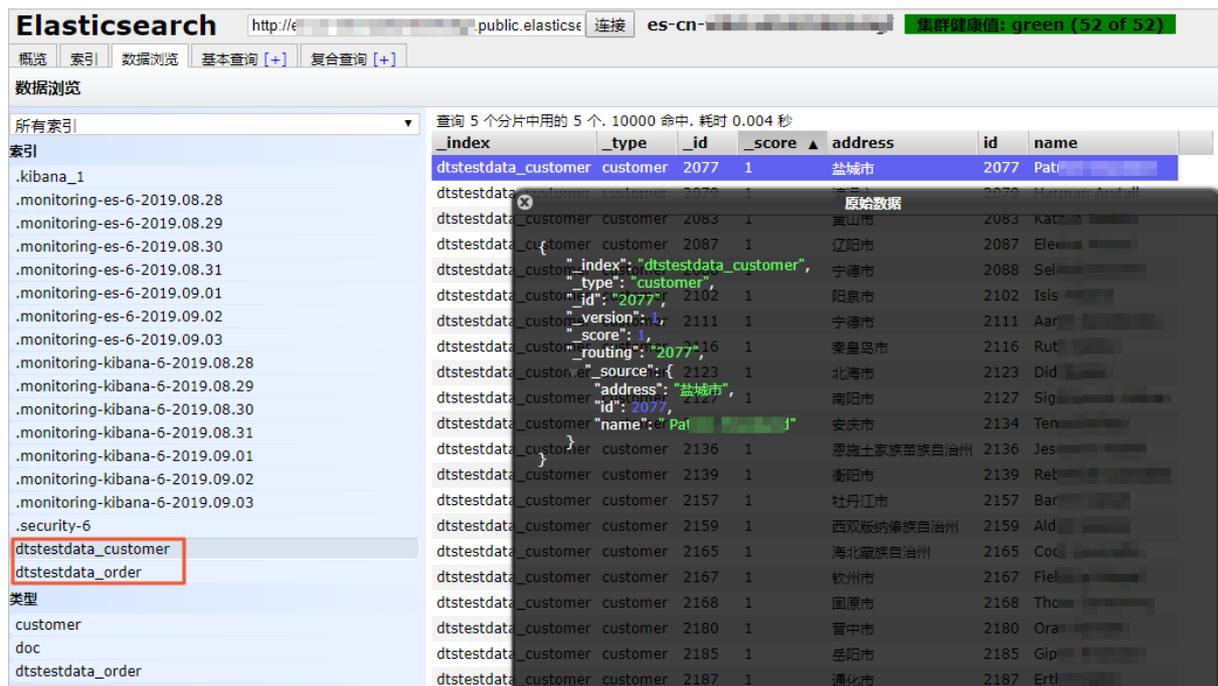
| 实例ID/作业名称               | 状态  | 同步概况                                   | 付费方式 | 同步策略(全部) | 操作              |
|-------------------------|-----|--|------|----------|-----------------|
| hangzhou-hangzhou-small | 同步中 | 耗时: 1376 毫秒<br>速度: 0.00RPS/(0.000MB/s) | 按量付费 | 单向同步     | 暂停同步 转包年包月 升级更多 |
| 暂停同步 释放同步               |     |  |      |          |                 |

共有1条，每页显示：20条

### 查看同步后的索引和数据

数据同步作业处于同步中状态后，您可以连接Elasticsearch实例（本案例使用Head插件进行连接），确认创建的索引和同步的数据是否符合业务的预期。

? 说明 如果不符合业务预期，您可以删除该索引及对应的数据，然后重新配置数据同步作业。



## 7.10. 从RDS MySQL同步至MaxCompute

大数据计算服务（MaxCompute，原名ODPS）是一种快速、完全托管的EB级数据仓库解决方案。通过数据传输服务DTS（Data Transmission Service），您可以将MySQL的数据同步至MaxCompute，帮助您快速搭建数据实时分析系统。

### 前提条件

您已完成以下操作：

- 开通MaxCompute
- 创建项目

### 注意事项

- DTS在执行全量数据初始化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。因此您需要在执行数据同步前评估源库和目标库的性能，同时建议您在业务低峰期执行数据同步（例如源库和目标库的CPU负载在30%以下）。
- 仅支持表级别的数据同步。
- 数据同步期间，请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。
- 由于MaxCompute不支持主键约束，当DTS在同步数据时因网络等原因触发重传，可能会导致MaxCompute中出现重复记录。

### 费用说明

|               |                                 |
|---------------|---------------------------------|
| 同步类型          | 链路配置费用                          |
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

### 源库支持的实例类型

执行数据同步操作的MySQL数据库支持以下实例类型：

- ECS上的自建数据库

- 通过专线、VPN网关或智能网关接入的自建数据库
- 通过数据库网关接入的自建数据库
- 同一或不同阿里云账号下的RDS MySQL实例

本文以RDS实例为例介绍配置流程，当源库为其他实例类型时，配置流程与该案例类似。

 **说明** 如果源库为自建MySQL数据库，您还需要执行相应的准备工作，详情请参见[准备工作概览](#)。

## 支持同步的SQL操作

- DDL操作：ADD COLUMN

 **说明** 仅支持下述数据类型：INTEGER、BIGINTEGER、IGINT、LONGSTRING、YEAR、TIME、DATA、TIME STAMP、DATA TIME、BYTE、BOOLEAN、DECIMAL、DOUBLE、FLOAT。

- DML操作：INSERT、UPDATE、DELETE

## 同步过程介绍

### 1. 结构初始化。

DTS将源库中待同步表的结构定义信息同步至MaxCompute中，初始化时DTS会为表名增加\_base后缀。例如源表为customer，那么MaxCompute中的表即为customer\_base。

### 2. 全量数据初始化。

DTS将源库中待同步表的存量数据，全部同步至MaxCompute中的目标表名\_base表中（例如从源库的customer表同步至MaxCompute的customer\_base表），作为后续增量同步数据的基线数据。

 **说明** 该表也被称为全量基线表。

### 3. 增量数据同步。

DTS在MaxCompute中创建一个增量日志表，表名为同步的目标表名\_log，例如customer\_log，然后将源库产生的增量数据实时同步到该表中。

 **说明** 关于增量日志表结构的详细信息，请参见[增量日志表结构定义说明](#)。

## 操作步骤

 **警告** 为保障DTS同步账号能够授权成功，请使用主账号完成下述操作步骤。

1. 购买数据同步作业，详情请参见[购买流程](#)。

 **说明** 购买时，选择源实例为MySQL，目标实例为MaxCompute，并选择同步拓扑为单向同步。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击[数据同步](#)。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击[配置同步链路](#)。
6. 配置同步通道的源实例及目标实例信息。

1.选择同步通道的源及目标实例
2.ODPS账号授权
3.选择同步对象
4.预检查

同步作业名称:

---

**源实例信息**

实例类型:

实例地区:

\* 实例ID:  [其他阿里云账号下的RDS实例](#)

\* 数据库账号:

\* 数据库密码:

---

**目标实例信息**

实例类型:

实例地区:

\* Project:

| 类别     | 配置     | 说明  |
|--------|--------|---|
| 无      | 同步作业名称 | DTS会自动生成一个同步作业名称，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。  |
| 源实例信息  | 实例类型   | 选择RDS实例。  |
|        | 实例地区   | 购买数据同步实例时选择的源实例地域信息，不可变更。   |
|        | 实例ID   | 选择作为数据同步源的RDS实例ID。  |
|        | 数据库账号  | 填入源RDS的数据库账号。<br><br><div style="border: 1px solid #add8e6; padding: 5px; margin-top: 5px;"> <span style="font-size: 1.2em;">?</span> 说明 当源RDS实例的数据库类型为MySQL 5.5或MySQL 5.6时，无需配置数据库账号和数据库密码。                 </div>   |
|        | 数据库密码  | 填入数据库账号对应的密码。   |
|        | 连接方式   | 根据需求选择非加密连接或SSL安全连接。如果设置为SSL安全连接，您需要提前开启RDS实例的SSL加密功能，详情请参见 <a href="#">设置SSL加密</a> 。<br><br><div style="border: 1px solid #add8e6; padding: 5px; margin-top: 5px;"> <span style="font-size: 1.2em;">!</span> 注意 目前仅中国内地及中国香港地域支持设置连接方式。                 </div> |
| 目标实例信息 | 实例类型   | 固定为MaxCompute，不可变更。   |
|        | 实例地区   | 购买数据同步实例时选择的目标实例地域信息，不可变更。  |

| 类别 | 配置      | 说明  |
|----|---------|---|
|    | Project | <p>填入MaxCompute实例的Project，您可以在MaxCompute工作空间列表页面中查询。</p>  |

7. 单击页面右下角的授权白名单并进入下一步。

**说明**

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 单击页面右下角的下一步，允许将MaxCompute中项目的下述权限授予DTS同步账号，详情如下图所示。



9. 配置同步策略和同步对象。

1.选择同步通道的源及目标实例 2.Maxcompute账号授权 3.选择同步对象 4.预检查

### 增量日志表分区定义

| 选择                                  | 分区名称              | 字段类型   | 描述                        |
|-------------------------------------|-------------------|--------|---------------------------|
| <input checked="" type="checkbox"/> | modifytime_year   | String | 增量更新时间对应的年份               |
| <input checked="" type="checkbox"/> | modifytime_month  | String | 增量更新时间对应的月份               |
| <input checked="" type="checkbox"/> | modifytime_day    | String | 增量更新时间对应的日期               |
| <input checked="" type="checkbox"/> | modifytime_hour   | String | 增量更新时间对应的小时               |
| <input type="checkbox"/>            | modifytime_minute | String | 增量更新时间对应的分钟 (每15分钟写入一个分区) |

同步初始化:  结构初始化  全量数据初始化

注: DTS全量任务运行期间, 不要清理DTS任务启动后源库产生的增量数据日志。源库如果过早清理日志, 可能会导致DTS增量任务失败

目标已存在表的处理模式:  预检查并报错拦截  忽略报错并继续执行

#### 源库对象

若全局搜索, 请先展开树

- chw
- chw02
- test\_polar2

全选

#### 已选择对象 (鼠标移到对象行, 点击编辑可修改对象名或过滤条件) 详情点我

- dtstest123 源库名:chw (1个对象)
- tw02

全选

\*映射名称更改:  不进行库表名称批量更改  要进行库表名称批量更改

\*源库、目标库无法连接后的重试时间: 720 分钟

\*是否启用新的附加列规则:  是  否

取消 上一步 预检查并启动

| 配置        | 说明   |
|-----------|--|
| 增量日志表分区定义 | 根据业务需求, 选择分区名称。关于分区的相关介绍请参见分区。   |
| 同步初始化     | 同步初始化类型细分为: 结构初始化、全量数据初始化。此处同时勾选结构初始化和全量数据初始化, DTS会在增量数据同步之前, 将源数据库中待同步对象的结构和存量数据同步到目标数据库。 |

| 配置                             | 说明   |
|--------------------------------|--|
| 目标已存在表的处理模式                    | <ul style="list-style-type: none"> <li> <b>预检查并报错拦截</b>：检查目标数据库中是否有同名的表。如果目标数据库中没有同名的表，则通过该检查项目；如果目标数据库中有同名的表，则在预检查阶段提示错误，数据同步作业不会被启动。                     </li> </ul> <div style="background-color: #e6f2ff; padding: 5px; margin: 5px 0;"> <p> <b>说明</b> 如果目标库中同名的表不方便删除或重命名，您可以<a href="#">设置同步对象在目标实例中的名称</a>来避免表名冲突。</p> </div> <ul style="list-style-type: none"> <li> <b>忽略报错并继续执行</b>：跳过目标数据库中是否有同名表的检查项。                     </li> </ul> <div style="background-color: #fff9c4; padding: 5px; margin: 5px 0;"> <p> <b>警告</b> 选择为忽略报错并继续执行，可能导致数据不一致，给业务带来风险，例如：</p> <ul style="list-style-type: none"> <li>表结构一致的情况下，如果在目标库遇到与源库主键的值相同的记录，在初始化阶段会保留目标库中的该条记录；在增量同步阶段则会覆盖目标库的该条记录。</li> <li>表结构不一致的情况下，可能会导致无法初始化数据、只能同步部分列的数据或同步失败。</li> </ul> </div> |
| 选择同步对象                         | <p>在源库对象框中单击待同步的表，然后单击  将其移动至已选择对象框。</p> <div style="background-color: #e6f2ff; padding: 5px; margin: 5px 0;"> <p> <b>说明</b></p> <ul style="list-style-type: none"> <li>同步对象支持选择的粒度仅为表，您可以从多个库中选择表作为同步对象。</li> <li>默认情况下，同步对象的名称保持不变。如果您需要在目标实例上名称不同，那么需要使用DTS提供的对象名映射功能，详情请参见<a href="#">设置同步对象在目标实例中的名称</a>。</li> </ul> </div>  |
| 选择附加列规则                        | <p>DTS在将数据同步到MaxCompute时，会在同步的目标表中添加一些附加列。如果附加列和目标表中已有的列出现名称冲突将会导致数据同步失败。您需要根据业务需求选择是否启用新的附加列规则为是或否。</p> <div style="background-color: #fff9c4; padding: 5px; margin: 5px 0;"> <p> <b>警告</b> 在选择附加列规则前，您需要评估附加列和目标表中已有的列是否会出现名称冲突。关于附加列的规则和定义说明，请参见<a href="#">附加列名称和定义说明</a>。</p> </div>   |
| 映射名称更改                         | <p>如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见<a href="#">库表列映射</a>。</p>  |
| 源表DMS_ONLINE_DDL过程中是否复制临时表到目标库 | <p>如源库使用<a href="#">数据管理DMS (Data Management Service)</a> 执行Online DDL变更，您可以选择是否同步Online DDL变更产生的临时表数据。</p> <ul style="list-style-type: none"> <li>是：同步Online DDL变更产生的临时表数据。</li> </ul> <div style="background-color: #e6f2ff; padding: 5px; margin: 5px 0;"> <p> <b>说明</b> Online DDL变更产生的临时表数据过大，可能会导致同步任务延迟。</p> </div> <ul style="list-style-type: none"> <li>否：不同步Online DDL变更产生的临时表数据，只同步源库的原始DDL数据。</li> </ul> <div style="background-color: #e6f2ff; padding: 5px; margin: 5px 0;"> <p> <b>说明</b> 该方案会导致目标库锁表。</p> </div>  |

| 配置            | 说明  |
|---------------|---|
| 源、目标库无法连接重试时间 | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <p><b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p> |

10. 上述配置完成后，单击页面右下角的**预检查并启动**。

**说明**

- 在同步作业正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步作业。
- 如果预检查失败，单击具体检查项后的，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择**确认屏蔽**、**忽略告警项**并重新进行预检查，跳过告警检测项重新进行预检查。

11. 在**预检查**对话框中显示**预检查通过**后，关闭**预检查**对话框，同步作业将正式开始。

12. 等待同步作业的链路初始化完成，直至处于**同步中**状态。

您可以在**数据同步**页面，查看数据同步作业的状态。



| 实例ID/作业名称               | 状态  | 同步概况                                   | 付费方式 | 同步架构(全部) | 操作                  |
|-------------------------|-----|--|------|----------|---------------------|
| hangzhou-hangzhou-small | 同步中 | 耗时: 1376 毫秒<br>速度: 0.00RPS/(0.000MB/s) | 按量付费 | 单向同步     | 暂停同步   转包年包月   升级更多 |

### 增量日志表结构定义说明

**说明** 你需要在MaxCompute中执行 `set odps.sql.allow.fullscan=true;`，设置项目空间属性，允许进行全表扫描。

DTS在将MySQL产生的增量数据同步至MaxCompute的增量日志表时，除了存储增量数据，还会存储一些元信息，示例如下。

|   | A     | B                   | C       | D         | E              | F             | G           | H          | I               | J                | K              | L               |
|---|-------|---------------------|---------|-----------|----------------|---------------|-------------|------------|-----------------|------------------|----------------|-----------------|
| 1 | id    | register_time       | address | record_id | operation_flag | utc_timestamp | before_flag | after_flag | modifytime_year | modifytime_month | modifytime_day | modifytime_hour |
| 2 | 10000 | 2018-02-03 01:38:01 | 广东省     | 15650333  | U              | 15650333      | Y           | N          | 2019            | 08               | 16             | 16              |
| 3 | 10000 | 2018-02-03 01:38:01 | 广东省广州市  | 15650333  | U              | 15650333      | N           | Y          | 2019            | 08               | 16             | 16              |
| 4 | 9999  | 2016-11-18 11:44:54 | 河北省     | 15650419  | D              | 15650419      | Y           | N          | 2019            | 08               | 16             | 16              |
| 5 | 10001 | 2018-12-23 05:11:59 | 浙江省     | 15650435  | I              | 15650435      | N           | Y          | 2019            | 08               | 16             | 16              |

**说明** 示例中的 `modifytime_year`、`modifytime_month`、`modifytime_day`、`modifytime_hour`、`modifytime_minute` 为分区字段，是在**配置同步策略**和**同步对象**步骤中指定的。

### 结构定义说明

| 字段 | 说明 |
|----|----|
|    |    |

| 字段             | 说明   |
|----------------|--|
| record_id      | 增量日志的记录ID，为该日志唯一标识。<br><div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p><span style="color: #00aaff;">?</span> 说明</p> <ul style="list-style-type: none"> <li>ID的值唯一且递增。</li> <li>如果增量日志的操作类型为UPDATE，那么增量更新会被拆分成两条记录（分别记录更新前和更新后的值），且record_id的值相同。</li> </ul> </div> |
| operation_flag | 操作类型，取值： <ul style="list-style-type: none"> <li>I: INSERT操作。</li> <li>D: DELETE操作。</li> <li>U: UPDATE操作。</li> </ul>  |
| utc_timestamp  | 操作时间戳，即binlog的时间戳（UTC时间）。  |
| before_flag    | 所有列的值是否为更新前的值，取值：Y或N。  |
| after_flag     | 所有列的值是否为更新后的值，取值：Y或N。  |

### 关于before\_flag和after\_flag的补充说明

对于不同的操作类型，增量日志中的before\_flag和after\_flag定义如下：

• INSERT

当操作类型为INSERT时，所有列的值为新插入的记录值，即为更新后的值，所以before\_flag取值为N，after\_flag取值为Y，示例如下。

| A | B     | C                   | D       | E           | F              | G             | H           | I          | J               | K                | L              |                 |
|---|-------|---------------------|---------|-------------|----------------|---------------|-------------|------------|-----------------|------------------|----------------|-----------------|
| 1 | id    | register_time       | address | record_id   | operation_flag | utc_timestamp | before_flag | after_flag | modifytime_year | modifytime_month | modifytime_day | modifytime_hour |
| 5 | 10001 | 2018-12-23 05:11:59 | 浙江省     | 15600000435 | I              | 1560878       | N           | Y          | 2019            | 08               | 16             | 16              |

• UPDATE

当操作类型为UPDATE时，DTS会将UPDATE操作拆为两条增量日志。这两条增量日志的record\_id、operation\_flag及utc\_timestamp对应的值相同。

第一条增量日志记录了更新前的值，所以before\_flag取值为Y，after\_flag取值为N。第二条增量日志记录了更新后的值，所以before\_flag取值为N，after\_flag取值为Y，示例如下。

| A | B     | C                   | D       | E         | F              | G             | H           | I          | J               | K                | L              |                 |
|---|-------|---------------------|---------|-----------|----------------|---------------|-------------|------------|-----------------|------------------|----------------|-----------------|
| 1 | id    | register_time       | address | record_id | operation_flag | utc_timestamp | before_flag | after_flag | modifytime_year | modifytime_month | modifytime_day | modifytime_hour |
| 2 | 10000 | 2018-02-03 01:38:01 | 广东省     | 156000333 | U              | 15604655      | Y           | N          | 2019            | 08               | 16             | 16              |
| 3 | 10000 | 2018-02-03 01:38:01 | 广东省广州市  | 156000333 | U              | 15604655      | N           | Y          | 2019            | 08               | 16             | 16              |

• DELETE

当操作类型为DELETE时，增量日志中所有的列值为被删除的值，即列值不变，所以before\_flag取值为Y，after\_flag取值为N，示例如下。

| A | B    | C                   | D       | E           | F              | G             | H           | I          | J               | K                | L              |                 |
|---|------|---------------------|---------|-------------|----------------|---------------|-------------|------------|-----------------|------------------|----------------|-----------------|
| 1 | id   | register_time       | address | record_id   | operation_flag | utc_timestamp | before_flag | after_flag | modifytime_year | modifytime_month | modifytime_day | modifytime_hour |
| 4 | 9999 | 2016-11-18 11:44:54 | 河北省     | 15600000419 | D              | 1560845       | Y           | N          | 2019            | 08               | 16             | 16              |

### 全量数据合并示例

执行数据同步的操作后，DTS会在MaxCompute中分别创建该表的全量基线表和增量日志表。您可以通过MaxCompute的SQL命令，对这两个表执行合并操作，得到某个时间点的全量数据。

本案例以customer表为例（表结构如下），介绍操作流程。

|   | Field         | Type        | Null | Key | Default | Extra |
|---|---------------|-------------|------|-----|---------|-------|
| 1 | id            | int(11)     | NO   | PRI | null    |       |
| 2 | register_time | timestamp   | YES  |     | null    |       |
| 3 | address       | varchar(32) | YES  |     | null    |       |

1. 根据源库中待同步表的结构，在MaxCompute中创建用于存储合并结果的表。

例如，需要获取customer表在 1565944878 时间点的全量数据。为方便业务识别，创建如下数据表：

```
CREATE TABLE `customer_1565944878` (
  `id` bigint NULL,
  `register_time` datetime NULL,
  `address` string);
```

#### ② 说明

- 您可以在MaxCompute的临时查询中，运行SQL命令。
- 关于MaxCompute支持的数据类型与相关说明，请参见数据类型。

2. 在MaxCompute中执行如下SQL命令，合并全量基线表和增量日志表，获取该表在某一时间点的全量数据。

```
set odps.sql.allow.fullscan=true;
insert overwrite table <result_storage_table>
select <col1>,
       <col2>,
       <colN>
  from(
select row_number() over(partition by t.<primary_key_column>
  order by record_id desc, after_flag desc) as row_number, record_id, operation_flag, after_flag, <c
oll1>, <col2>, <colN>
  from(
select incr.record_id, incr.operation_flag, incr.after_flag, incr.<col1>, incr.<col2>,incr.<colN>
  from <table_log> incr
 where utc_timestamp< <timestamp>
 union all
select 0 as record_id, 'I' as operation_flag, 'Y' as after_flag, base.<col1>, base.<col2>,base.<col
N>
  from <table_base> base) t) gt
where record_num=1
   and after_flag='Y'
```

#### ② 说明

- <result\_storage\_table>：存储全量merge结果集的表名。
- <col1>/<col2>/<colN>：同步表中的列名。
- <primary\_key\_column>：同步表中的主键列名。
- <table\_log>：增量日志表名。
- <table\_base>：全量基线表名。
- <timestamp>：需要获取全量数据的时间点。

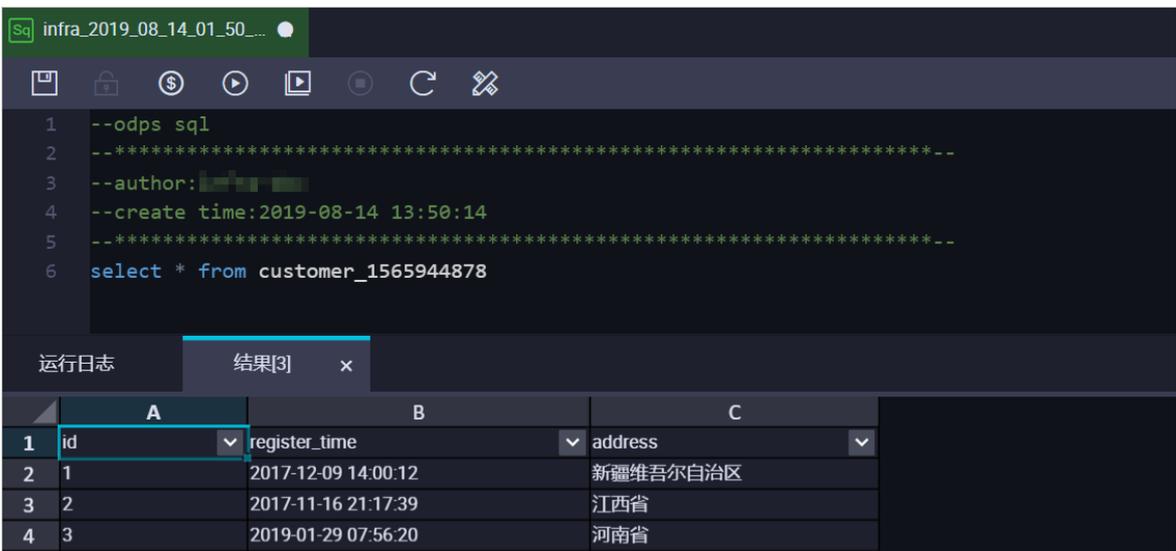
合并数据表，获取customer表在 1565944878 时间点的全量数据，示例如下：

```

set odps.sql.allow.fullscan=true;
insert overwrite table customer_1565944878
select id,
       register_time,
       address
  from(
select row_number() over(partition by t.id
order by record_id desc, after_flag desc) as row_number, record_id, operation_flag, after_flag, id
, register_time, address
  from(
select incr.record_id, incr.operation_flag, incr.after_flag, incr.id, incr.register_time, incr.address
  from customer_log incr
 where utc_timestamp< 1565944878
 union all
select 0 as record_id, 'I' as operation_flag, 'Y' as after_flag, base.id, base.register_time, base.address
  from customer_base base) t) gt
 where gt.row_number= 1
       and gt.after_flag= 'Y';

```

3. 上述命令执行完成后，可在customer\_1565944878表中查询合并后的数据。



## 7.11. 从RDS MySQL同步至Tablestore

表格存储（Tablestore）是阿里云自研的多模型结构化数据存储，提供海量结构化数据存储以及快速的查询和分析服务。通过数据传输服务DTS（Data Transmission Service），您可以将MySQL数据库（例如自建MySQL或RDS MySQL）同步至Tablestore实例，帮助您轻松实现数据的流转。

### 前提条件

已创建目标表格存储（Tablestore）实例，详情请参见[创建表格存储实例](#)。

### 注意事项

- DTS在执行全量数据初始化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，您需要在执行数据同步前评估数据同步对源库和目标库性能的影响，同时建议您在业务低峰期执行数据同步。
- 不支持同步DDL操作。如果同步过程中执行了DDL操作，您需要先[移除同步对象](#)，然后在Tablestore实例中删除该表，最后[新增同步对象](#)。
- 待同步的表数量符合Tablestore实例的限制（不超过64个）。如业务确实会超过此限制，请为目标Tablestore实例[提交工单](#)放开此限制。

- 待同步的表或列名称符合Tablestore实例的命名规范：
  - 表或列的名称由大小写字母、数字或下划线（\_）组成，且只能以字母或下划线开头。
  - 表或列的名称长度为1~255个字符。

### 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

### 同步初始化类型说明

| 同步初始化类型 | 说明   |
|---------|--|
| 结构初始化   | <p>DTS将源库中待同步对象的结构定义信息同步至目标库，目前支持的对象为表。</p> <p> <b>警告</b> 此场景属于异构数据库间的数据同步，DTS在执行结构初始化时数据类型无法完全对应，请谨慎评估数据类型的映射关系对业务的影响，详情请参见<a href="#">结构初始化涉及的数据类型映射关系</a>。</p> |
| 全量数据初始化 | DTS将源库中待同步对象的存量数据，全部同步到目标库中，作为后续增量同步数据的基线数据。   |
| 增量数据初始化 | <p>DTS在全量数据初始化的基础上，将源库的增量更新数据实时同步至目标库。</p> <p>在增量数据初始化阶段，DTS支持同步的SQL语句为INSERT、UPDATE、DELETE。</p> <p> <b>警告</b> 请勿在源库执行DDL语句，否则将导致数据同步失败。</p>                    |

### 准备工作

 **注意** 由于在配置目标库时需要填写访问密钥AccessKey（AK）信息来执行授权，为避免主账号泄露AccessKey带来的安全风险，建议您通过RAM用户来完成授权和AccessKey的创建。

- 创建RAM用户并授予其AliyunOTSFULLAccess权限（管理表格存储服务的权限）。
  - 登录[RAM控制台](#)。
  - 创建RAM用户。
  - 在左侧导航栏的人员管理菜单下，单击用户。
  - 找到目标RAM用户，单击其操作列的添加权限。



- v. 在对话框的文本框中输入`AliyunOTSFullAccess`，然后单击选择搜索结果中的权限策略名称。

- vi. 单击**确定**。
- vii. 单击**完成**。
2. 为RAM用户创建访问密钥AccessKey (AK)，详情请参见[创建AccessKey](#)。
  3. (可选) 如需使用RAM用户配置数据同步作业，您还需要为该用户授予AliyunDTSDefaultRole权限，详情请参见[授予DTS访问云资源的权限](#)。

**说明** 如果使用主账号配置数据同步作业，无需执行本步骤。

## 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

**说明** 购买时，选择源实例为MySQL，目标实例为Tablestore，并选择同步拓扑为**单向同步**。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击**数据同步**。
4. 在**同步作业列表**页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击**配置同步链路**。
6. 配置数据同步作业的源库和目标库信息。

1. 选择同步通道的源及目标实例
2. 选择同步对象
3. 高级设置
4. 预检查

同步作业名称:

**源实例信息**

实例类型:

实例地区: 华东1 (杭州)

\* 实例ID:  其他阿里云账号下的RDS实例

\* 数据库账号:

\* 数据库密码:

\* 连接方式:  非加密连接  SSL安全连接

**目标实例信息**

实例类型:

实例地区: 华东1 (杭州)

\* 实例ID:

\* AccessKeyId:

\* AccessKeySecret:

| 类别    | 配置     | 说明  |
|-------|--------|---|
| 无     | 同步作业名称 | DTS会自动生成一个同步作业名称，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。  |
| 源实例信息 | 实例类型   | 选择RDS实例。  |
|       | 实例地区   | 购买数据同步实例时选择的源实例地域信息，不可变更。   |
|       | 实例ID   | 选择源RDS实例ID。   |
|       | 数据库账号  | 填入源RDS的数据库账号，该账号需具备REPLICATION CLIENT、REPLICATION SLAVE、SHOW VIEW和所有同步对象的SELECT权限。<br><br><span style="background-color: #e1f5fe; padding: 5px; border: 1px solid #cfcfcf; font-size: small;">? 说明 当源RDS实例的数据库类型为MySQL 5.5或MySQL 5.6时，无需配置数据库账号和数据库密码。</span> |
|       | 数据库密码  | 填入该数据库账号的密码。  |
|       | 连接方式   | 根据需求选择非加密连接或SSL安全连接。如果设置为SSL安全连接，您需要提前开启RDS实例的SSL加密功能，详情请参见 <a href="#">设置SSL加密</a> 。<br><br><span style="background-color: #e1f5fe; padding: 5px; border: 1px solid #cfcfcf; font-size: small;">🔔 注意 目前仅中国内地及中国香港地域支持设置连接方式。</span>                          |
|       | 实例类型   | 选择Tablestore。   |
|       | 实例地区   | 购买数据同步实例时选择的目标实例地域信息，不可变更。  |

| 类别     | 配置              | 说明  |
|--------|-----------------|---|
| 目标实例信息 | 实例ID            | 选择目标Tablestore实例ID。                                   |
|        | AccessKeyId     | 填入AccessKeyId信息，创建方法请参见 <a href="#">创建AccessKey</a> 。 |
|        | AccessKeySecret | 填入AccessKeyId对应的AccessKeySecret信息。                    |

7. 单击页面右下角的授权白名单并进入下一步。

**说明**

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置同步策略及对象信息。

i. 配置同步策略。

1.选择同步通道的源及目标实例
2.选择同步对象
3.高级设置
4.预检查

同步初始化:  结构初始化  全量数据初始化  增量数据初始化

目标已存在表的处理模式:  预检查并报错拦截  忽略报错并继续执行

多表归并:  是  否

同步操作类型:  Insert  Update  Delete

脏数据处理策略: 跳过

数据写入模式: 行覆盖

批量写入方式: BulkImportRequest

队列大小: 1024 收起

线程数: 8

并发数: 8

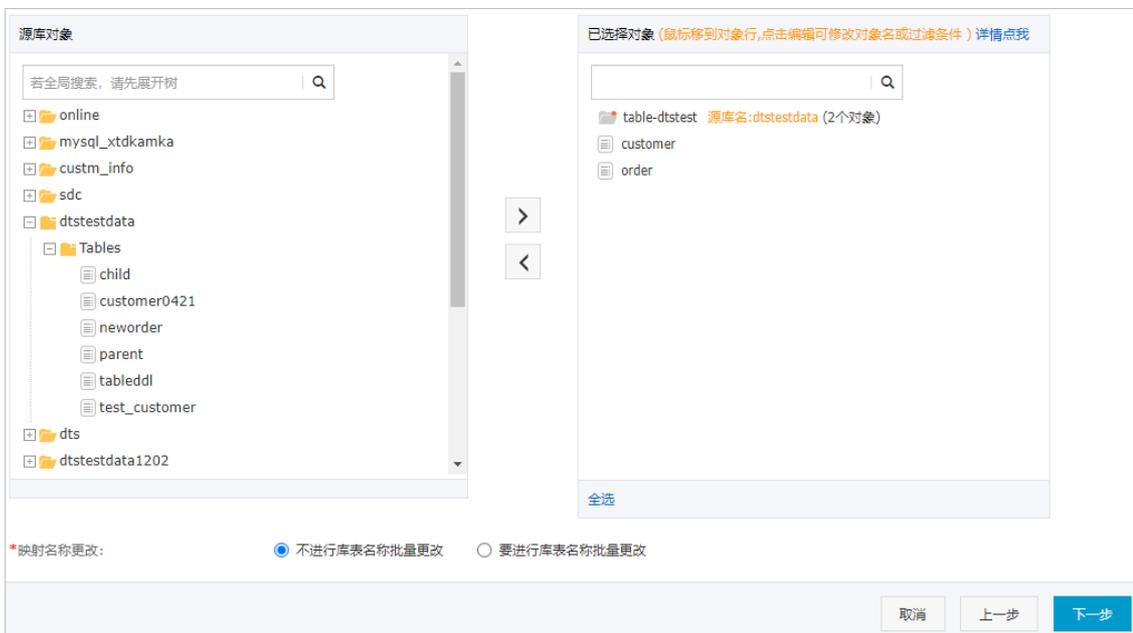
分桶数: 2

| 配置    | 说明  |
|-------|---|
| 同步初始化 | 选中结构初始化、全量数据初始化和增量数据初始化，相关说明请参见 <a href="#">同步初始化类型说明</a> 。 |

| 配置          | 说明  |
|-------------|---|
| 目标已存在表的处理模式 | <ul style="list-style-type: none"> <li>■ <b>预检查并报错拦截</b>：检查目标数据库中是否有同名的表。如果目标数据库中没有同名的表，则通过该检查项目；如果目标数据库中有同名的表，则在预检查阶段提示错误，数据同步作业不会被启动。</li> </ul> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin: 5px 0;"> <p> <b>说明</b> 如果目标库中同名的表不方便删除或重命名，您可以<b>设置同步对象在目标实例中的名称</b>来避免表名冲突。</p> </div> <ul style="list-style-type: none"> <li>■ <b>忽略报错并继续执行</b>：跳过目标数据库中是否有同名表的检查项。</li> </ul> <div style="border: 1px solid #ccc; background-color: #fff9c4; padding: 5px; margin: 5px 0;"> <p> <b>警告</b> 选择为忽略报错并继续执行，可能导致数据不一致，给业务带来风险，例如：</p> <ul style="list-style-type: none"> <li>■ 表结构一致的情况下，如果在目标库遇到与源库主键的值相同的记录，在初始化阶段会保留目标库中的该条记录；在增量同步阶段则会覆盖目标库的该条记录。</li> <li>■ 表结构不一致的情况下，可能会导致无法初始化数据、只能同步部分列的数据或同步失败。</li> </ul> </div> |
| 多表归并        | <ul style="list-style-type: none"> <li>■ <b>选择为是</b>：通常在OLTP场景中，为提高业务表响应速度，通常会做分库分表处理。此类场景中，您可以借助DTS的多表归并功能将源库中多个表结构相同的表（即各分表）同步至Tablestore实例的同一个表中。</li> </ul> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin: 5px 0;"> <p> <b>说明</b></p> <ul style="list-style-type: none"> <li>■ DTS会在Tablestore实例的同步目标表中增加 <code>__dts_data_source</code> 列（类型为varchar）来存储数据来源。DTS将以 <code>&lt;dts数据同步实例ID&gt;:&lt;源数据库名&gt;.&lt;源表名&gt;</code> 的格式写入列值用于区分表的来源，例如 <code>dts*****:dts testdata.customer1</code>。</li> <li>■ 多表归并功能基于实例级别，如果需要让部分表执行多表归并，另一部分不执行多表归并，您需要为这两批表分别创建数据同步作业。</li> </ul> </div> <ul style="list-style-type: none"> <li>■ <b>选择为否</b>：默认选项。</li> </ul>   |
| 同步操作类型      | 根据业务选中需要同步的操作类型，默认情况下都处于选中状态。   |
| 脏数据处理策略     | <p>选择数据写入错误时的处理策略：</p> <ul style="list-style-type: none"> <li>■ 跳过</li> <li>■ 阻塞</li> </ul>   |
| 数据写入模式      | <ul style="list-style-type: none"> <li>■ <b>行更新</b>：使用PutRowChange会做行级别更新。</li> <li>■ <b>行覆盖</b>：使用UpdateRowChange会做行级别覆盖。</li> </ul>   |
| 批量写入方式      | <p>批量写入调用接口。</p> <ul style="list-style-type: none"> <li>■ <b>BulkImportRequest</b>：离线写入。</li> <li>■ <b>BatchWriteRowRequest</b>：批量写入。</li> </ul> <p>建议选择BulkImportRequest，读写效率更高，Tablestore实例计费便宜。</p>  |

| 配置     | 说明  |
|--------|---|
| 更多设置   |   |
| └ 队列大小 | Tablestore实例数据写入进程的队列长度。  |
| └ 线程数  | Tablestore实例数据写入进程的回调处理线程数。   |
| └ 并发数  | Tablestore实例的并发请求限制数。   |
| └ 分桶数  | 增量按序写入时的分桶并发数，适当调大可提升并发写入能力。<br><span style="background-color: #e0f2f7; padding: 5px; border: 1px solid #ccc;"> <span style="color: #0070c0; font-size: 1.2em;">?</span> <b>说明</b> 该值需小于或等于并发数。                 </span> |

ii. 配置同步对象。



| 配置     | 说明   |
|--------|--|
| 选择同步对象 | <p>在源库对象框中单击待迁移的表（不能超过64个表），然后单击  图标将其移动至已选择对象框。</p> <p>默认情况下，同步对象的名称保持不变。如果您需要同步对象在目标实例上名称不同，请使用对象名映射功能，详情请参见 <a href="#">设置同步对象在目标实例中的名称</a>。</p> <div style="background-color: #fff9c4; padding: 5px; border: 1px solid #ccc;"> <p> <b>警告</b> 如果在 <a href="#">步骤8</a> 选择多表归并为是，选择源库的多个表后，您需要通过对对象名映射功能，将其改为Tablestore实例中的同一个表名，否则将导致数据不一致。</p> </div> |

iii. 单击下一步。

9. (可选) 在已选择对象区域框中，将鼠标指针放置在待同步的表上，并单击表名后出现的编辑，设置该表中各列在Tablestore实例中的数据类型。



10. 配置待同步的表在Tablestore实例中的主键列信息。



说明 关于Tablestore实例中主键的相关介绍，请参见主键。

11. 上述配置完成后，单击页面右下角的预检查并启动。

说明

- 在数据同步作业正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动数据同步作业。
- 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。

12. 在预检查对话框中显示预检查通过后，关闭预检查对话框，同步作业将正式开始。

13. 等待同步作业的链路初始化完成，直至处于同步中状态。

您可以在数据同步页面，查看数据同步作业的状态。



## 7.12. 从自建MySQL同步至阿里云消息队列Kafka版

通过数据传输服务DTS（Data Transmission Service），您可以将自建MySQL同步至消息队列Kafka版，扩展消息处理能力。

### 前提条件

- 自建MySQL的数据库版本为5.1、5.5、5.6、5.7或8.0版本。
- 目标Kafka实例的版本为0.10.1.0-2.x版本。
- 目标Kafka实例中已创建用于接收同步数据的Topic，详情请参见[创建Topic](#)。

### 背景信息

[消息队列Kafka版](#)是阿里云提供的分布式、高吞吐、可扩展的消息队列服务，针对开源的Apache Kafka提供全托管服务，彻底解决开源产品长期以来的痛点，您只需专注于业务开发，无需部署运维。消息队列Kafka版广泛用于日志收集、监控数据聚合、流式数据处理、在线和离线分析等大数据领域，已成为大数据生态中不可或缺的部分。

### 注意事项

- DTS在执行全量数据初始化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。因此您需要在执行数据同步前评估源库和目标库的性能，同时建议您在业务低峰期执行数据同步（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。

### 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

### 功能限制

- 同步对象仅支持数据表，不支持非数据表的对象。
- 不支持自动调整同步对象，如果对同步对象中的数据表进行重命名操作，且重命名后的名称不在同步对象中，那么这部分数据将不再同步到目标Kafka集群中。如需将修改后的数据表继续数据同步至目标Kafka集群中，您需要进行修改同步对象操作，详情请参见[新增同步对象](#)。

### 准备工作

为自建MySQL创建账号并设置binlog

### 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

 说明 购买时，选择源实例为MySQL，目标实例为Kafka，选择同步拓扑为单向同步。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击[数据同步](#)。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击[配置同步链路](#)。
6. 配置同步作业的源实例及目标实例信息。

同步作业名称:

---

**源实例信息**

实例类型:

实例地区:

\* ECS实例ID:

数据库类型:

\* 端口:

\* 数据库账号:

\* 数据库密码:

---

**目标实例信息**

实例类型:

实例地区:

\* 对端专有网络:

数据库类型:

\* IP地址:

\* 端口:

数据库账号:  非必填项

数据库密码:  非必填项

\* Topic:

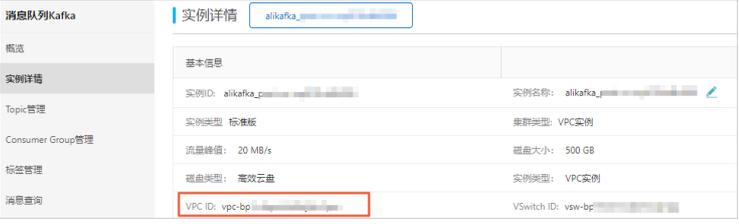
请先点击右侧按钮，获取Topic列表后选择具体的Topic

\* Kafka版本:

\* 连接方式:  非加密连接  SCRAM-SHA-256

---

| 类别    | 配置      | 说明  |
|-------|---------|---|
| 无     | 同步作业名称  | DTS会自动生成一个同步作业名称，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。  |
| 源实例信息 | 实例类型    | 根据源库部署位置选择，本文以ECS上的自建数据库为例介绍配置流程，其他类型的配置流程与该案例类似。                                     |
|       | 实例地区    | 购买数据同步实例时选择的源实例地域信息，不可变更。   |
|       | ECS实例ID | 选择自建MySQL所属的ECS实例ID。  |
|       | 数据库类型   | 固定为MySQL，不可变更。  |
|       | 端口      | 填入自建MySQL的数据库服务端口。  |
|       | 数据库账号   | 填入自建MySQL的数据库账号，该账号需具备REPLICATION CLIENT、REPLICATION SLAVE、SHOW VIEW和所有同步对象的SELECT权限。 |
|       | 数据库密码   | 填入该数据库账号对应的密码。  |

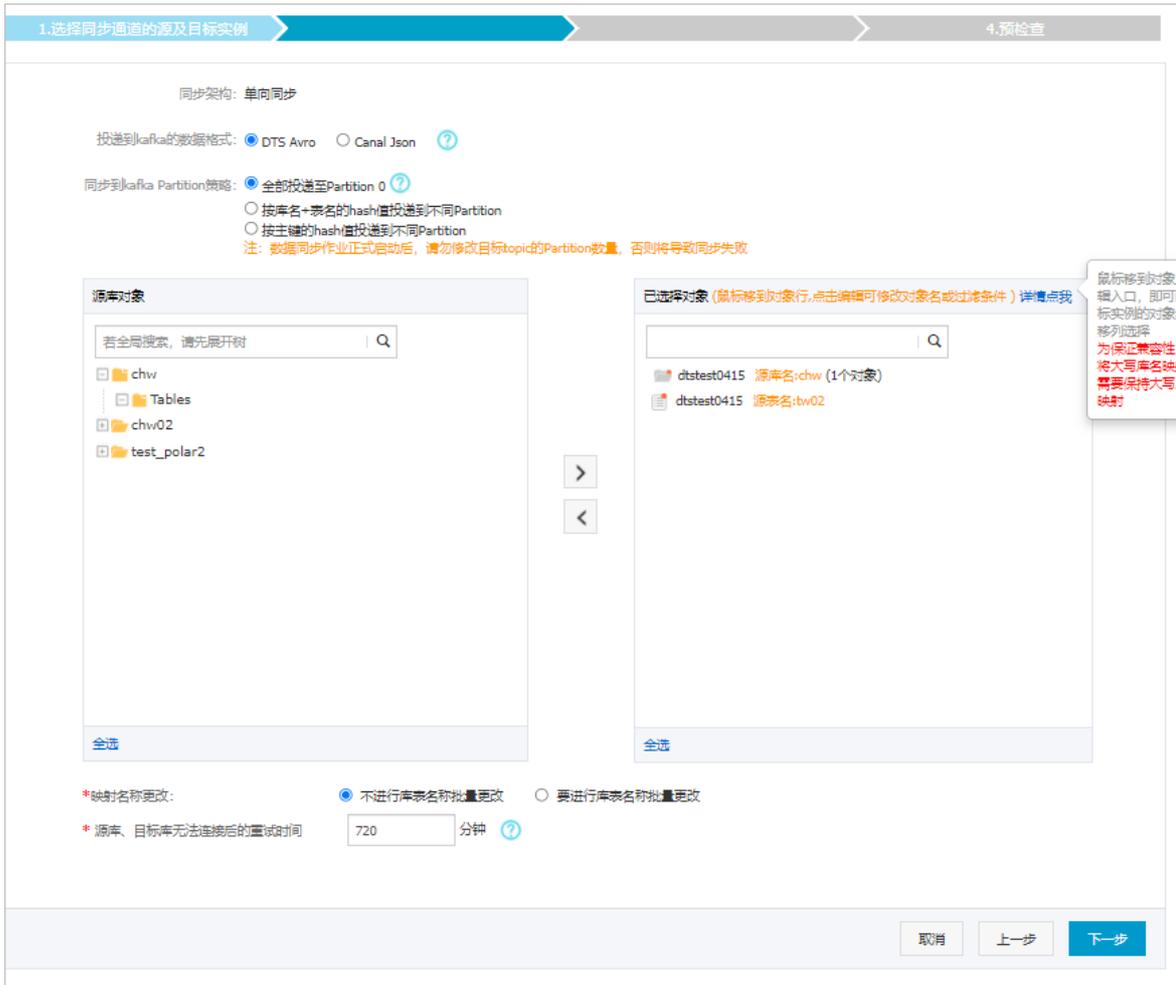
| 类别      | 配置                               | 说明  |
|---------|----------------------------------|---|
| 目标实例信息  | 实例类型                             | 选择通过专线/VPN网关/智能接入网关接入的自建数据库。<br><br><span style="background-color: #e6f2ff; padding: 5px;"> <span style="font-size: 1.2em;">?</span> <b>说明</b> 由于DTS暂时不支持直接选择消息队列Kafka版，此处将其作为自建Kafka来配置数据同步。                 </span> |
|         | 实例地区                             | 购买数据同步实例时选择的目标实例地域信息，不可变更。  |
|         | 对端专有网络                           | 选择目标Kafka实例所属的专有网络ID。您可以在Kafka实例的基本信息页面中查看到专有网络ID。<br><br>  |
|         | 数据库类型                            | 选择为Kafka。   |
|         | IP地址                             | 填入Kafka实例默认接入点中的任意一个IP地址。<br><br><span style="background-color: #e6f2ff; padding: 5px;"> <span style="font-size: 1.2em;">?</span> <b>说明</b> 您可以在Kafka实例的基本信息页面中，获取默认接入点对应的IP地址。                 </span>               |
|         | 端口                               | Kafka实例的服务端口，默认为9092。   |
|         | 数据库账号                            | 填入Kafka实例的用户名。<br><br><span style="background-color: #e6f2ff; padding: 5px;"> <span style="font-size: 1.2em;">?</span> <b>说明</b> 如果Kafka实例的实例类型为VPC实例，无需配置数据库账号和数据库密码。                 </span>                        |
|         | 数据库密码                            | 填入该用户名对应的密码。  |
|         | Topic                            | 单击右侧的 <b>获取Topic列表</b> ，然后在下拉框中选择具体的Topic。  |
| Kafka版本 | 根据Kafka实例版本，选择对应的版本信息。           |   |
| 连接方式    | 根据业务及安全需求，选择非加密连接或SCRAM-SHA-256。 |   |

7. 单击页面右下角的**授权白名单**并进入下一步。

? **说明**

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置同步对象信息。



| 配置                   | 说明   |
|----------------------|--|
| 投递到kafka的数据格式        | 同步到Kafka集群中的数据以avro格式或者Canal Json格式存储，定义详情请参见 <a href="#">Kafka集群的数据存储格式</a> 。   |
| 同步到Kafka Partition策略 | 根据业务需求选择同步的策略，详细介绍请参见 <a href="#">Kafka Partition同步策略说明</a> 。  |
| 同步对象                 | <p>在源库对象区域框中，选择需要同步的对象（选择的粒度为表），然后单击  图标将其移动到已选对象区域框中。</p> <p> 说明 DTS会自动将表名映射为步骤6选择的Topic名称。如果需要更换同步的目标Topic，您需要使用库表列名映射功能，详情请参见<a href="#">设置同步对象在目标实例中的名称</a>。</p> |
| 映射名称更改               | 如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见 <a href="#">库表列映射</a> 。   |

| 配置            | 说明  |
|---------------|---|
| 源、目标库无法连接重试时间 | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <p><b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p> |

9. 上述配置完成后单击页面右下角的下一步。

10. 配置同步初始化的高级配置信息。



| 配置    | 说明  |
|-------|---|
| 同步初始化 | 默认选择结构初始化和全量数据初始化，DTS会在增量数据同步之前，将源库中待同步对象的结构和存量数据，同步到目标库。 |
| 过滤选项  | 默认选择忽略增量同步阶段的 DDL，即增量同步阶段源库执行的DDL操作不会被DTS同步至目标库。          |

11. 上述配置完成后，单击页面右下角的预检查并启动。

**说明**

- 在同步作业正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步作业。
- 如果预检查失败，单击具体检查项后的 ，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择**确认屏蔽**、**忽略告警项**并重新进行预检查，跳过告警检测项重新进行预检查。

12. 在预检查对话框中显示预检查通过后，关闭预检查对话框，数据同步作业正式开始。

您可以在数据同步页面，查看数据同步状态。



## 7.13. 从RDS MySQL同步至自建Kafka集群

Kafka是应用较为广泛的分布式、高吞吐量、高可扩展性消息队列服务，普遍用于日志收集、监控数据聚合、流式数据处理、在线和离线分析等大数据领域，是大数据生态中不可或缺的产品之一。通过数据传输服务DTS（Data Transmission Service），您可以将RDS MySQL的数据库同步至自建Kafka集群，扩展消息处理能力。

## 前提条件

- Kafka集群的版本为0.10.1.0-2.7.0版本。
- 已创建RDS MySQL实例，详情请参见[创建RDS MySQL实例](#)。

## 注意事项

- DTS在执行全量数据初始化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。因此您需要在执行数据同步前评估源库和目标库的性能，同时建议您在业务低峰期执行数据同步（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。

## 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

## 功能限制

- 同步对象仅支持数据表，不支持非数据表的对象。
- 不支持自动调整同步对象，如果对同步对象中的数据表进行重命名操作，且重命名后的名称不在同步对象中，那么这部分数据将不再同步到目标Kafka集群中。如需将修改后的数据表继续数据同步至目标Kafka集群中，您需要进行[修改同步对象](#)操作，详情请参见[新增同步对象](#)。

## 支持的同步架构

- 1对1单向同步。
- 1对多单向同步。
- 多对1单向同步。
- 级联单向同步。

## 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

 **说明** 购买时，选择源实例为MySQL、目标实例为Kafka，并选择同步拓扑为单向同步。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击[数据同步](#)。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击[配置同步链路](#)。
6. 配置同步作业的源实例及目标实例信息。

同步作业名称: MySQL\_TO\_Kafka

---

**源实例信息**

实例类型: RDS实例

实例地区: 华东1 (杭州)

\* 实例ID: rm-bp-... 其他阿里云账号下的RDS实例

\* 数据库账号: dtstest

\* 数据库密码: .....

\* 连接方式:  非加密连接  SSL安全连接

---

**目标实例信息**

实例类型: ECS上的自建数据库

实例地区: 华东1 (杭州)

\* ECS实例ID: i-bp-...

数据库类型: Kafka

\* 端口: 9092

数据库账号: 非必填项

数据库密码: 非必填项

\* Topic: rdstest 请先点击右侧按钮, 获取Topic列表后选择具体的Topic

\* Kafka版本: 0.10

\* 连接方式:  非加密连接  SCRAM-SHA-256

| 类别     | 配置      | 说明  |
|--------|---------|---|
| 无      | 同步作业名称  | DTs会自动生成一个同步作业名称, 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。  |
| 源实例信息  | 实例类型    | 选择RDS实例。  |
|        | 实例地区    | 购买数据同步实例时选择的源实例地域, 不可变更。  |
|        | 实例ID    | 选择源RDS实例的实例ID。  |
|        | 数据库账号   | 填入连接源RDS实例数据库的账号, 需要具备REPLICATION CLIENT、REPLICATION SLAVE、SHOW VIEW和所有同步对象的SELECT权限。   |
|        | 数据库密码   | 填入连接源RDS实例数据库账号的密码。   |
|        | 连接方式    | 根据需求选择非加密连接或SSL安全连接。如果设置为SSL安全连接, 您需要提前开启RDS实例的SSL加密功能, 详情请参见 <a href="#">设置SSL加密</a> 。<br><div style="border: 1px solid #ccc; padding: 5px; margin-top: 5px;"> <span style="font-size: 1.2em;">🔔</span> <b>注意</b> 目前仅中国内地及中国香港地域支持设置连接方式。                 </div> |
| 目标实例信息 | 实例类型    | 根据Kafka集群的部署位置选择, 本文以ECS上的自建数据库为例介绍配置流程。<br><div style="border: 1px solid #ccc; padding: 5px; margin-top: 5px;"> <span style="font-size: 1.2em;">❓</span> <b>说明</b> 当选择为其他实例类型时, 您还需要执行相应的准备工作, 详情请参见<a href="#">准备工作概览</a>。                 </div>               |
|        | 实例地区    | 购买数据同步实例时选择的目标实例地域, 不可变更。   |
|        | ECS实例ID | 选择部署了Kafka集群的ECS实例ID。   |

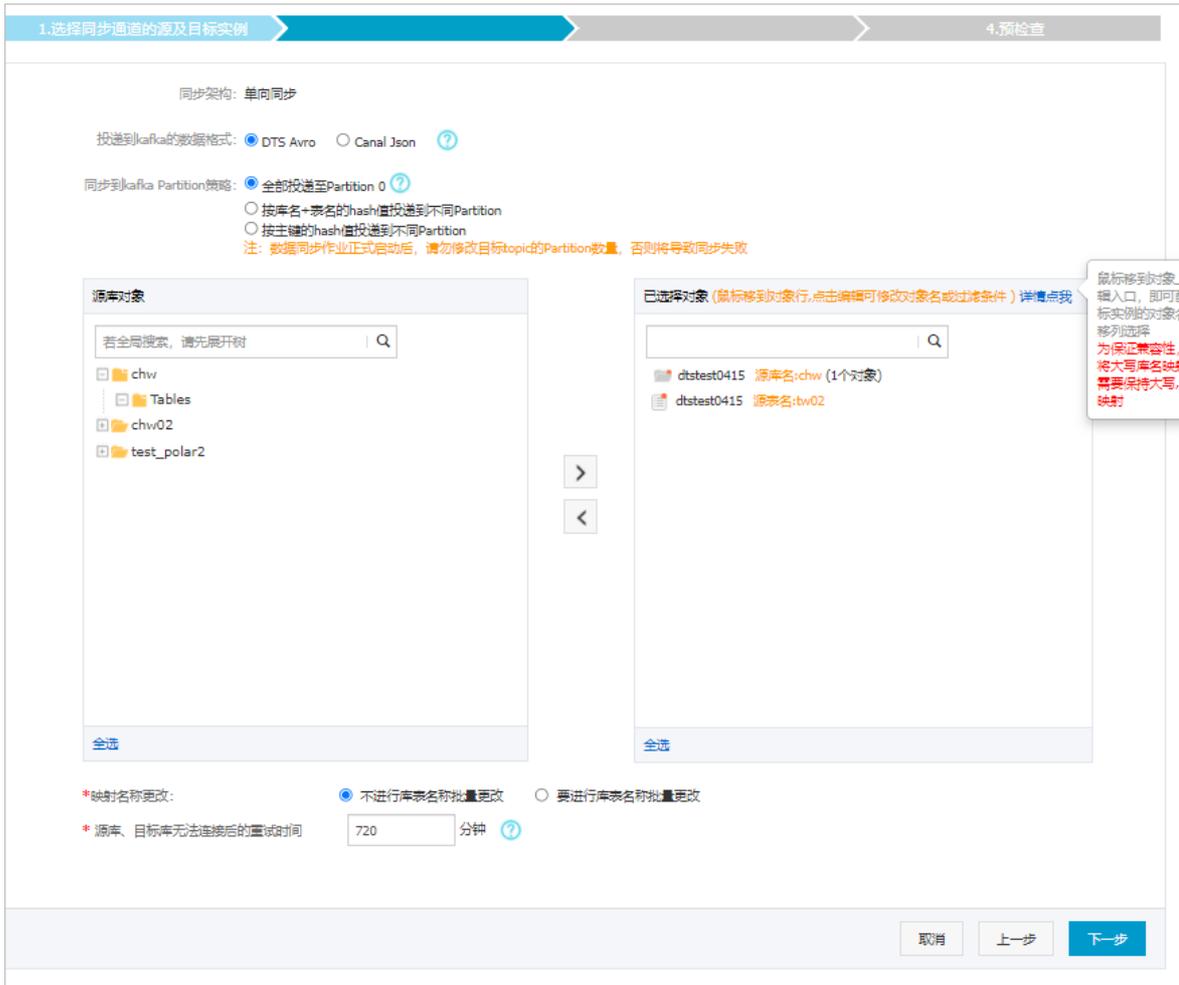
| 类别     | 配置                        | 说明  |
|--------|---------------------------|---|
| 目标实例信息 | 数据库类型                     | 固定为Kafka。   |
|        | 端口                        | 填入Kafka集群提供服务的端口，默认为9092。   |
|        | 数据库账号                     | 填入Kafka集群的用户名，如Kafka集群未开启验证可不填写。  |
|        | 数据库密码                     | 填入Kafka集群用户名的密码，如Kafka集群未开启验证可不填写。  |
|        | Kafka版本                   | 选择目标Kafka集群的版本。   |
|        | 连接方式                      | 根据业务及安全需求，选择非加密连接或SCRAM-SHA-256。  |
|        | Topic                     | 单击右侧的获取Topic列表，然后在下拉框中选择目标Topic。  |
|        | 存储DDL的Topic               | 在下拉框中选择具体的Topic，用于存储DDL信息。如果未指定，DDL信息默认存储在上方选择的Topic中。  |
|        | 是否使用Kafka Schema Registry | Kafka Schema Registry是元数据提供服务层，提供了一个RESTful接口，用于存储和检索Avro Schema。 <ul style="list-style-type: none"> <li>○ 否：不使用Kafka Schema Registry。</li> <li>○ 是：使用Kafka Schema Registry。您需要输入Avro Schema在Kafka Schema Registry注册的URL或IP。</li> </ul> |

7. 单击页面右下角的授权白名单并进入下一步。

 说明

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置同步对象信息。



| 配置                   | 说明   |
|----------------------|--|
| 投递到kafka的数据格式        | 同步到Kafka集群中的数据以avro格式或者Canal Json格式存储，定义详情请参见 <a href="#">Kafka集群的数据存储格式</a> 。   |
| 同步到Kafka Partition策略 | 根据业务需求选择同步的策略，详细介绍请参见 <a href="#">Kafka Partition同步策略说明</a> 。  |
| 同步对象                 | <p>在源库对象区域框中，选择需要同步的对象（选择的粒度为表），然后单击  图标将其移动到已选对象区域框中。</p> <p> 说明 DTS会自动将表名映射为步骤6选择的Topic名称。如果需要更换同步的目标Topic，您需要使用库表列名映射功能，详情请参见<a href="#">设置同步对象在目标实例中的名称</a>。</p> |
| 映射名称更改               | 如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见 <a href="#">库表列映射</a> 。   |
| 源、目标库无法连接重试时间        | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <p> 说明 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p>   |

- 上述配置完成后单击页面右下角的下一步。
- 配置同步初始化的高级配置信息。

1.选择同步通道的源及目标实例
4.预检查

同步初始化:  结构初始化  全量数据初始化 注: 不支持trigger的同步, 详情请参考文档

过滤选项:  忽略增量同步阶段的 DDL

取消 上一步 保存 预检查并启动

| 配置    | 说明   |
|-------|--|
| 同步初始化 | 默认选择结构初始化和全量数据初始化, DTS会在增量数据同步之前, 将源库中待同步对象的结构和存量数据, 同步到目标库。 |
| 过滤选项  | 默认选择忽略增量同步阶段的 DDL, 即增量同步阶段源库执行的DDL操作不会被DTS同步至目标库。            |

- 上述配置完成后, 单击页面右下角的预检查并启动。

? 说明

- 在同步作业正式启动之前, 会先进行预检查。只有预检查通过后, 才能成功启动同步作业。
- 如果预检查失败, 单击具体检查项后的 !, 查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项, 您也可以选择确认屏蔽、忽略告警项并重新进行预检查, 跳过告警检测项重新进行预检查。

- 在预检查对话框中显示预检查通过后, 关闭预检查对话框, 数据同步作业正式开始。您可以在数据同步页面, 查看数据同步状态。

同步作业名称
搜索
排序: 默认排序
状态: 全部

| 实例ID/作业名称  | 状态  | 同步概况                             | 付费方式 | 同步架构(全部) | 操作  |
|--|-----|----------------------------------|------|----------|---|
| <input type="checkbox"/> hangzhou-hangzhou-small | 同步中 | 耗时: 565 毫秒<br>速度: 0TPS(0.00MB/s) | 按量付费 | 单向同步     | <a href="#">暂停同步</a>   <a href="#">转包年包月</a>   <a href="#">升级更多</a> |
| <span>共有1条, 每页显示: 20条</span>                     |     |                                  |      |          |   |

## 7.14. 从ECS上的自建MySQL同步至自建Kafka集群

Kafka是应用较为广泛的分布式、高吞吐量、高可扩展性消息队列服务, 普遍用于日志收集、监控数据聚合、流式数据处理、在线和离线分析等大数据领域, 是大数据生态中不可或缺的产品之一。通过数据传输服务DTS (Data Transmission Service), 您可以将ECS上的自建MySQL数据同步至自建Kafka集群, 扩展消息处理能力。

### 前提条件

- Kafka集群的版本为0.10.1.0-2.7.0版本。
- 已创建RDS MySQL实例, 详情请参见[创建RDS MySQL实例](#)。

### 注意事项

- DTS在执行全量数据初始化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。因此您需要在执行数据同步前评估源库和目标库的性能, 同时建议您在业务低峰

期执行数据同步（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。

## 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

## 功能限制

- 同步对象仅支持数据表，不支持非数据表的对象。
- 不支持自动调整同步对象，如果对同步对象中的数据表进行重命名操作，且重命名后的名称不在同步对象中，那么这部分数据将不再同步到目标Kafka集群中。如需将修改后的数据表继续数据同步至目标Kafka集群中，您需要进行修改同步对象操作，详情请参见[新增同步对象](#)。

## 支持的同步架构

- 1对1单向同步
- 1对多单向同步
- 多对1单向同步
- 级联单向同步

## 准备工作

在正式配置数据同步作业之前，您需要为[自建MySQL创建账号并设置binlog](#)。

## 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

 说明 购买时，选择源实例为MySQL，目标实例为Kafka，选择同步拓扑为单向同步。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击[数据同步](#)。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击[配置同步链路](#)。
6. 配置同步通道的源实例及目标实例信息。

同步作业名称: MySQL\_TO\_Kafka

---

**源实例信息**

实例类型: ECS上的自建数据库

实例地区: 华东1 (杭州)

\* ECS实例ID: i-bp-\*\*\*\*\*

数据库类型: MySQL

\* 端口: 3306

\* 数据库账号: dtstest

\* 数据库密码: \*\*\*\*\*

---

**目标实例信息**

实例类型: ECS上的自建数据库

实例地区: 华东1 (杭州)

\* ECS实例ID: i-bp-\*\*\*\*\*

数据库类型: Kafka

\* 端口: 9092

数据库账号:  非必填项

数据库密码:  非必填项

\* Topic: dtstestopic

请先点击右侧按钮, 获取Topic列表后选择具体的Topic

\* Kafka版本: 0.10

\* 连接方式:  非加密连接  SCRAM-SHA-256

| 类别    | 配置      | 说明  |
|-------|---------|---|
| 无     | 同步作业名称  | DTS会自动生成一个同步作业名称, 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。  |
| 源实例信息 | 实例类型    | 选择ECS上的自建数据库。   |
|       | 实例地区    | 购买数据同步实例时选择的源实例地域, 不可变更。  |
|       | ECS实例ID | 选择作为同步数据源的ECS实例ID。  |
|       | 数据库类型   | 固定为MySQL, 不可变更。   |
|       | 端口      | 填入自建数据库服务端口, 默认为3306。   |
|       | 数据库账号   | 填入自建MySQL的数据库账号, 需要具备REPLICATION CLIENT、REPLICATION SLAVE、SHOW VIEW和所有同步对象的SELECT权限。  |
|       | 数据库密码   | 填入该数据库账号对应的密码。  |
|       | 实例类型    | 根据Kafka集群的部署位置选择, 本文以ECS上的自建数据库为例介绍配置流程。<br><br><div style="border: 1px solid #ccc; padding: 5px; background-color: #e6f2ff;"> <span style="font-size: 1.2em; color: #007bff;">?</span> <b>说明</b> 当选择为其他实例类型时, 您还需要执行相应的准备工作, 详情请参见<a href="#">准备工作概览</a>。                 </div> |
|       | 实例地区    | 购买数据同步实例时选择的目标实例地域, 不可变更。   |
|       | ECS实例ID | 选择部署了Kafka集群的ECS实例ID。   |
|       | 数据库类型   | 选择为Kafka。   |

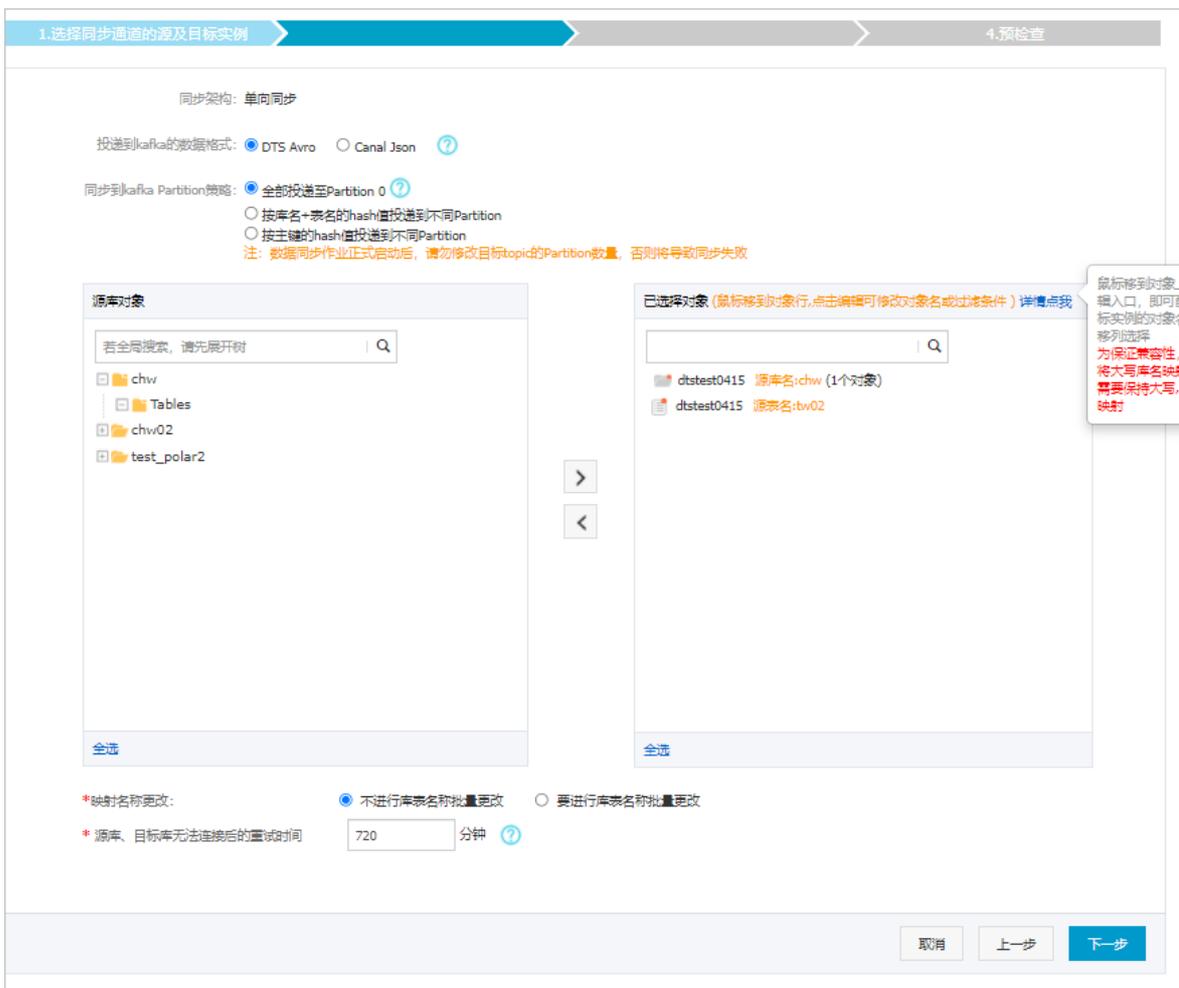
| 实例信息 | 配置      | 说明                                   |
|------|---------|--------------------------------------|
|      | 端口      | Kafka集群对外提供服务的端口，默认为9092。            |
|      | 数据库账号   | 填入Kafka集群的用户名，如Kafka集群未开启验证可不填写。     |
|      | 数据库密码   | 填入Kafka集群用户名对应的密码，如Kafka集群未开启验证可不填写。 |
|      | Topic   | 单击右侧的获取Topic列表，然后在下拉框中选择具体的Topic名称。  |
|      | Kafka版本 | 根据目标Kafka集群版本，选择对应的版本信息。             |
|      | 连接方式    | 根据业务及安全需求，选择非加密连接或SCRAM-SHA-256。     |

7. 单击页面右下角的授权白名单并进入下一步。

**说明**

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置同步对象信息。



| 配置                   | 说明  |
|----------------------|---|
| 投递到kafka的数据格式        | 同步到Kafka集群中的数据以avro格式或者Canal Json格式存储，定义详情请参见 <a href="#">Kafka集群的数据存储格式</a> 。  |
| 同步到Kafka Partition策略 | 根据业务需求选择同步的策略，详细介绍请参见 <a href="#">Kafka Partition同步策略说明</a> 。   |
| 同步对象                 | <p>在源库对象区域框中，选择需要同步的对象（选择的粒度为表），然后单击  图标将其移动到已选对象区域框中。</p> <p><b>说明</b> DTS会自动将表名映射为步骤6选择的Topic名称。如果需要更换同步的目标Topic，您需要使用库表列名映射功能，详情请参见<a href="#">设置同步对象在目标实例中的名称</a>。</p> |
| 映射名称更改               | 如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见 <a href="#">库表列映射</a> 。  |
| 源、目标库无法连接重试时间        | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <p><b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p>   |

9. 上述配置完成后单击页面右下角的下一步。

10. 配置同步初始化的高级配置信息。

1.选择同步通道的源及目标实例
4.预检查

同步初始化:  结构初始化  全量数据初始化 注: 不支持trigger的同步, 详情请参考文档

过滤选项:  忽略增量同步阶段的 DDL

取消 上一步 保存 预检查并启动

| 配置    | 说明  |
|-------|---|
| 同步初始化 | 默认选择结构初始化和全量数据初始化，DTS会在增量数据同步之前，将源库中待同步对象的结构和存量数据，同步到目标库。 |
| 过滤选项  | 默认选择忽略增量同步阶段的 DDL，即增量同步阶段源库执行的DDL操作不会被DTS同步至目标库。          |

11. 上述配置完成后，单击页面右下角的预检查并启动。

**说明**

- 在同步作业正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步作业。
- 如果预检查失败，单击具体检查项后的 ，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择**确认屏蔽**、**忽略告警项**并重新进行预检查，跳过告警检测项重新进行预检查。

12. 在预检查对话框中显示预检查通过后，关闭预检查对话框，数据同步作业正式开始。

您可以在数据同步页面，查看数据同步状态。



## 7.15. 从通过专线、VPN网关或智能接入网关接入的自建MySQL同步至自建Kafka集群

Kafka是应用较为广泛的分布式、高吞吐量、高可扩展性消息队列服务，普遍用于日志收集、监控数据聚合、流式数据处理、在线和离线分析等大数据领域，是大数据生态中不可或缺的产品之一。通过数据传输服务DTS（Data Transmission Service），您可以将通过专线、VPN网关或智能接入网关接入的自建MySQL数据同步至自建Kafka集群，扩展消息处理能力。

### 前提条件

- Kafka集群的版本为0.10.1.0-2.7.0版本。
- 自建MySQL数据库版本为5.1、5.5、5.6、5.7或8.0版本。
- 已经将自建MySQL数据库接入至阿里云专有网络。详情请参见[通过CEN实现本地IDC与DTS云服务互通](#)，本文不做详细介绍。

### 注意事项

- DTS在执行全量数据初始化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。因此您需要在执行数据同步前评估源库和目标库的性能，同时建议您在业务低峰期执行数据同步（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。

### 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

### 功能限制

- 同步对象仅支持数据表，不支持非数据表的对象。
- 不支持自动调整同步对象，如果对同步对象中的数据表进行重命名操作，且重命名后的名称不在同步对象中，那么这部分数据将不再同步到目标Kafka集群中。如需将修改后的数据表继续数据同步至目标Kafka集群中，您需要进行修改同步对象操作，详情请参见[新增同步对象](#)。

### 支持的同步架构

- 1对1单向同步
- 1对多单向同步
- 多对1单向同步
- 级联单向同步

### 准备工作

在正式配置数据同步作业之前，您需要为自建MySQL创建账号并设置binlog。

### 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

? **说明** 购买时，选择源实例为MySQL，目标实例为Kafka，选择同步拓扑为单向同步。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击[数据同步](#)。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击[配置同步链路](#)。
6. 配置同步通道的源实例及目标实例信息。

同步作业名称:

---

**源实例信息**

实例类型:

实例地区:

\* 对端专有网络:  [其他阿里云账号下的专有网络](#)

数据库类型:

\* IP地址:

\* 端口:

\* 数据库账号:

\* 数据库密码:

---

**目标实例信息**

实例类型:

实例地区:

\* ECS实例ID:

数据库类型:

\* 端口:

数据库账号:  非必填项

数据库密码:  非必填项

\* Topic:     
请先点击右侧按钮，获取Topic列表后选择具体的Topic

\* Kafka版本:

\* 连接方式:  非加密连接  SCRAM-SHA-256

| 类别    | 配置     | 说明  |
|-------|--------|---|
| 无     | 同步作业名称 | DTS会自动生成一个同步作业名称，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。                                      |
| 源实例信息 | 实例类型   | 选择通过专线/VPN网关/智能接入网关接入的自建数据库。  |
|       | 实例地区   | 购买数据同步实例时选择的源实例地域信息，不可变更。   |
|       | 对端专有网络 | 选择自建数据库接入的VPC ID。   |
|       | 数据库类型  | 固定为MySQL，不可变更。  |
|       | IP地址   | 填入自建MySQL数据库的服务器IP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数据库的账号，需要具备REPLICATION CLIENT、REPLICATION SLAVE、SHOW VIEW和所有同步对象的SELECT权限。 |

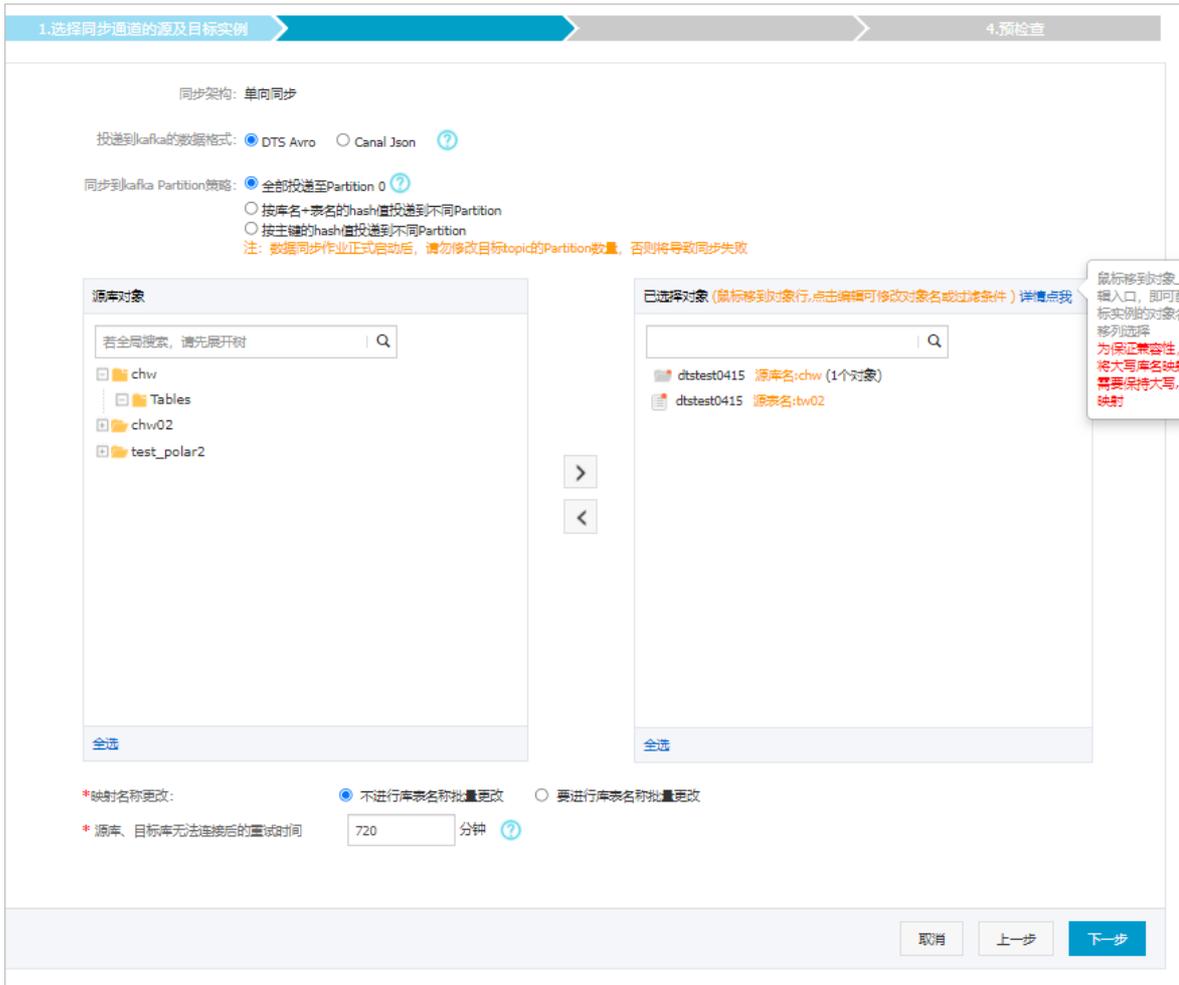
| 类别     | 配置      | 说明  |
|--------|---------|---|
|        | 数据库密码   | 填入该数据库账号的密码。  |
| 目标实例信息 | 实例类型    | 根据Kafka集群的部署位置选择，本文以ECS上的自建数据库为例介绍配置流程。<br><br><div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px;"> <p><span style="color: #00aaff;">?</span> <b>说明</b> 当选择为其他实例类型时，您还需要执行相应的准备工作，详情请参见<a href="#">准备工作概览</a>。</p> </div> |
|        | 实例地区    | 购买数据同步实例时选择的目标实例地域信息，不可变更。  |
|        | ECS实例ID | 选择部署了Kafka集群的ECS实例ID。   |
|        | 数据库类型   | 选择为Kafka。   |
|        | 端口      | Kafka集群对外提供服务的端口，默认为9092。   |
|        | 数据库账号   | 填入Kafka集群的用户名，如Kafka集群未开启验证可不填写。  |
|        | 数据库密码   | 填入Kafka集群用户名的密码，如Kafka集群未开启验证可不填写。  |
|        | Topic   | 单击右侧的 <a href="#">获取Topic列表</a> ，然后在下拉框中选择具体的Topic名称。   |
|        | Kafka版本 | 根据目标Kafka集群版本，选择对应的版本信息。  |
|        | 连接方式    | 根据业务及安全需求，选择非加密连接或SCRAM-SHA-256。  |

#### 7. 单击页面右下角的授权白名单并进入下一步。

##### ? 说明

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

#### 8. 配置同步对象信息。



| 配置                   | 说明   |
|----------------------|--|
| 投递到kafka的数据格式        | 同步到Kafka集群中的数据以avro格式或者Canal Json格式存储，定义详情请参见 <a href="#">Kafka集群的数据存储格式</a> 。   |
| 同步到Kafka Partition策略 | 根据业务需求选择同步的策略，详细介绍请参见 <a href="#">Kafka Partition同步策略说明</a> 。  |
| 同步对象                 | <p>在源库对象区域框中，选择需要同步的对象（选择的粒度为表），然后单击  图标将其移动到已选对象区域框中。</p> <p> 说明 DTS会自动将表名映射为步骤6选择的Topic名称。如果需要更换同步的目标Topic，您需要使用库表列名映射功能，详情请参见<a href="#">设置同步对象在目标实例中的名称</a>。</p> |
| 映射名称更改               | 如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见 <a href="#">库表列映射</a> 。   |
| 源、目标库无法连接重试时间        | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <p> 说明 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p>   |

- 上述配置完成后单击页面右下角的下一步。
- 配置同步初始化的高级配置信息。

1.选择同步通道的源及目标实例
4.预检查

同步初始化:  结构初始化  全量数据初始化 注: 不支持trigger的同步, 详情请参考文档

过滤选项:  忽略增量同步阶段的 DDL

取消 上一步 保存 预检查并启动

| 配置    | 说明   |
|-------|--|
| 同步初始化 | 默认选择结构初始化和全量数据初始化, DTS会在增量数据同步之前, 将源库中待同步对象的结构和存量数据, 同步到目标库。 |
| 过滤选项  | 默认选择忽略增量同步阶段的 DDL, 即增量同步阶段源库执行的DDL操作不会被DTS同步至目标库。            |

- 上述配置完成后, 单击页面右下角的预检查并启动。

? 说明

- 在同步作业正式启动之前, 会先进行预检查。只有预检查通过后, 才能成功启动同步作业。
- 如果预检查失败, 单击具体检查项后的 !, 查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项, 您也可以选择确认屏蔽、忽略告警项并重新进行预检查, 跳过告警检测项重新进行预检查。

- 在预检查对话框中显示预检查通过后, 关闭预检查对话框, 数据同步作业正式开始。您可以在数据同步页面, 查看数据同步状态。

同步作业名称  搜索 排序: 默认排序 状态: 全部

| 实例ID/作业名称  | 状态  | 同步概况                             | 计费方式 | 同步架构(全部) | 操作   |
|--|-----|----------------------------------|------|----------|--|
| <input type="checkbox"/> hangzhou-hangzhou-small                   | 同步中 | 延时: 565 毫秒<br>速度: 0TPS(0.00MB/s) | 按量付费 | 单向同步     | <a href="#">暂停同步</a>   <a href="#">转包年月</a>   <a href="#">升级更多</a> |
| <input type="checkbox"/> <a href="#">暂停同步</a> <a href="#">释放同步</a> |     |                                  |      |          |  |

共有1条, 每页显示: 20条

## 7.16. 从RDS MySQL同步至PolarDB-X

介绍云原生分布式数据库PolarDB-X（简称PolarDB-X）是阿里巴巴致力于解决单机数据库服务瓶颈而自主研发的分布式数据库产品，高度兼容MySQL协议和语法，支持自动化水平拆分、在线平滑扩容、弹性扩展、透明读写分离，具备数据库全生命周期运维管控能力。通过数据传输服务DTS（Data Transmission Service），您可以将RDS MySQL同步至PolarDB-X。

### 前提条件

- 源库中待同步的数据表必须具备主键。
- 确保目标库具备充足的存储空间。
- 已创建分布式关系型数据库服务PolarDB-X，如未创建请参见[创建实例](#)和[创建数据库](#)。

? 说明 创建实例时需选择存储类型为RDS MySQL。

### 注意事项

- DTS在执行全量数据初始化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。因此您需要在执行数据同步前评估源库和目标库的性能，同时建议您在业务低峰期执行数据同步（例如源库和目标库的CPU负载在30%以下）。
- 全量初始化过程中，并发INSERT会导致目标集群的表碎片，全量初始化完成后，目标集群的表空间比源库的表空间大。确保目标库具备充足的存储空间。
- MySQL同步至PolarDB-X暂不支持结构初始化，在配置同步作业前，您需要在目标实例中创建对应的库和表。

### 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

### 同步限制

- 同步对象仅支持数据表。
- 不支持BIT、VARBIT、GEOMETRY、ARRAY、UUID、TSQUERY、TSVECTOR、TXID\_SNAPSHOT类型的数据同步。
- 暂不支持同步前缀索引，如果源库存在前缀索引可能导致数据同步失败。
- 在数据同步时，请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。

### 支持同步的SQL操作

INSERT、UPDATE、DELETE。

### 数据库账号的权限要求

| 数据库       | 所需权限  |
|-----------|---|
| RDS MySQL | REPLICATION CLIENT、REPLICATION SLAVE、SHOW VIEW和所有同步对象的SELECT权限。 |
| PolarDB-X | 无需填写数据库账号信息，DTS会自动创建账号并授权。                                      |

### 支持的同步架构

- 1对1单向同步。
- 多对1单向同步。

### 准备工作

由于RDS MySQL同步至PolarDB-X暂不支持结构初始化，所以您需要根据源RDS MySQL实例中待同步对象的数据结构，在目标实例中创建相应的数据库和数据表，详情请参见[创建数据库](#)和[创建数据表](#)。

 **说明** 结构初始化：将源库中待同步对象的结构定义信息，同步至目标库中。

### 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

 **说明** 购买时，选择源实例为MySQL、目标实例为PolarDB-X（原DRDS升级版），并选择同步拓扑为单向同步。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击数据同步。

4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击配置同步链路。
6. 配置同步作业的源实例及目标实例信息。

| 类别     | 配置       | 说明  |
|--------|----------|---|
| 无      | 同步作业名称   | DTS会自动生成一个同步作业名称，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。  |
| 源实例信息  | 实例类型     | 选择RDS实例。  |
|        | 实例地区     | 购买数据同步实例时选择的源实例地域信息，不可变更。   |
|        | 实例ID     | 选择源RDS MySQL实例ID。   |
|        | 数据库账号    | 填入源RDS的数据库账号，权限要求请参见数据库账号的权限要求。<br><br><div style="border: 1px solid #ccc; padding: 5px; background-color: #e6f2ff;">                     ? 说明 当源RDS实例的数据库类型为MySQL 5.5或MySQL 5.6时，无需配置数据库账号和数据库密码。                 </div>       |
|        | 数据库密码    | 填入该数据库账号对应的密码。  |
| 目标实例信息 | 连接方式     | 根据需求选择非加密连接或SSL安全连接。如果设置为SSL安全连接，您需要提前开启RDS实例的SSL加密功能，详情请参见设置SSL加密。<br><br><div style="border: 1px solid #ccc; padding: 5px; background-color: #e6f2ff;">                     注意 目前仅中国内地及中国香港地域支持设置连接方式。                 </div> |
|        | 实例类型     | 固定为DRDS实例，不可变更。   |
|        | 实例地区     | 购买数据同步实例时选择的目标实例地域信息，不可变更。  |
|        | DRDS实例ID | 选择目标PolarDB-X（原DRDS升级版）实例ID。  |

7. 单击页面右下角的授权白名单并进入下一步。

说明

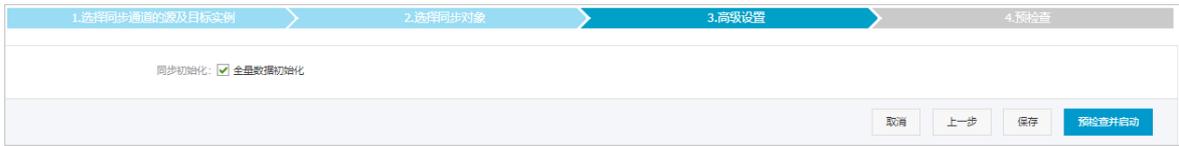
- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置同步策略及对象信息。



| 配置项目          | 配置说明  |
|---------------|---|
| 选择同步对象        | <p>在源库对象框中单击待同步的表，然后单击 <b>&gt;</b> 图标将其移动至已选择对象框。</p> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>同步对象的选择粒度为表。</li> <li>默认情况下，同步对象的名称保持不变。如果您需要改变同步对象在目标实例中的名称，那么需要使用DTS提供的对象名映射功能，详情请参见<a href="#">设置同步对象在目标实例中的名称</a>。</li> </ul> |
| 映射名称更改        | 如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见 <a href="#">库表映射</a> 。   |
| 源、目标库无法连接重试时间 | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <p><b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p>   |

- 9. 单击下一步。
- 10. 选择是否要执行全量数据初始化。



**说明** 全量数据初始化：DTS将源库中待同步表的存量数据同步至目标库中，如果不选择则不同步存量数据。

- 11. 上述配置完成后，单击页面右下角的预检查并启动。

**说明**

- 在同步作业正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步作业。
- 如果预检查失败，单击具体检查项后的 ，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择**确认屏蔽**、**忽略告警项**并重新进行预检查，跳过告警检测项重新进行预检查。

- 12. 在预检查对话框中显示预检查通过后，关闭预检查对话框，同步作业将正式开始。
- 13. 等待同步作业的链路初始化完成，直至处于同步中状态。

您可以在数据同步页面，查看数据同步作业的状态。



# 8.PolarDB MySQL数据同步

## 8.1. PolarDB MySQL集群间的双向同步

数据传输服务DTS（Data Transmission Service）支持两个PolarDB MySQL集群间的双向数据同步，适用于异地多活、异地容灾等多种应用场景。本文介绍双向数据同步的配置步骤。

### 前提条件

- 已购买源和目标PolarDB MySQL集群（简称PolarDB集群），详情请参见[创建PolarDB MySQL集群](#)。
- 源和目标PolarDB集群已开启Binlog，详情请参见[如何开启Binlog](#)。

### 注意事项

- DTS在执行全量数据初始化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。因此您需要在执行数据同步前评估源库和目标库的性能，同时建议您在业务低峰期执行数据同步（例如源库和目标库的CPU负载在30%以下）。
- 全量初始化过程中，并发INSERT会导致目标集群的表碎片，全量初始化完成后，目标集群的表空间比源集群的表空间大。
- 如果数据同步的源集群没有主键或唯一约束，且记录的全字段没有唯一性，可能会出现重复数据。
- 如双向同步任务的源实例或目标实例位于海外地域，则仅支持同地域的双向同步，不支持跨地域的双向同步。例如，支持日本地域间的双向同步，不支持日本地域与法兰克福地域间的双向同步。

### 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

### 功能限制

- 目前DTS仅支持一对一双向同步，暂不支持多个PolarDB集群间的双向同步。
- 不兼容触发器  
当同步对象为整个库，且库中的触发器（TRIGGER）会更新库内某个表时，可能导致源和目标库的数据不一致。相关解决方案请参见[源库存在触发器时如何配置同步作业](#)。
- RENAME TABLE限制  
RENAME TABLE操作可能导致同步数据不一致。例如同步对象只包含某个表，如果同步过程中源实例对该表执行了重命名操作，那么该表的数据将不会同步到目标库。为避免该问题，您可以在数据同步配置时将该表所属的整个数据库作为同步对象。
- DDL语法同步方向限制  
为保障双向同步链路的稳定性，DDL更新只能在其中一个同步方向进行同步。即一旦某个同步方向配置了DDL同步，则在反方向上不支持DDL同步，只进行DML同步。

### 支持同步的SQL操作

| 操作类型 | SQL操作语句                      |
|------|------------------------------|
| DML  | INSERT、UPDATE、DELETE、REPLACE |

| 操作类型 | SQL操作语句   |
|------|---|
| DDL  | <ul style="list-style-type: none"> <li>• ALTER TABLE、ALTER VIEW</li> <li>• CREATE FUNCTION、CREATE INDEX、CREATE PROCEDURE、CREATE TABLE、CREATE VIEW</li> <li>• DROP INDEX、DROP TABLE</li> <li>• RENAME TABLE</li> <li>• TRUNCATE TABLE</li> </ul> |

## 支持的冲突检测

为保障同步数据的一致性，您需要确保同一个主键、业务主键或唯一键的记录只在一个PolarDB集群上执行更新。如果发生了误操作，在两个PolarDB集群上均执行更新，那么将出现同步冲突。

DTS通过冲突检测和修复最大程度地维护双向同步实例的稳定性。目前DTS支持进行检测的冲突类型包括：

- **INSERT导致的唯一性冲突**  
同步INSERT语句时违背了唯一性约束，例如双向同步的两个节点同时或者在极为接近的时间新增了某个主键值相同的记录，那么同步到对端时，会因为已经存在相同主键值的记录，导致INSERT操作的同步失败。
- **UPDATE更新的记录不完全匹配**
  - UPDATE要更新的记录在同步目标集群中不存在时，DTS会自动转化为INSERT，此时可能会出现唯一键的唯一性冲突。
  - UPDATE要更新的记录出现主键或唯一键冲突。
- **DELETE对应的记录不存在**  
DELETE要删除的记录在同步的目标集群中不存在。出现这种冲突时，不论配置何种冲突修复策略，DTS都会自动忽略DELETE操作。

### 注意

- 由于数据同步两端的系统时间可能存在差异、同步存在延时等多种因素，DTS无法完全保证冲突检测机制能够完全防止数据的冲突。在使用双向同步时，您需要在业务层面配合进行相应的改造，保证同一个主键、业务主键或唯一键的记录只在双向同步的某个节点进行更新。
- 对于上述数据同步的冲突，DTS提供了修复策略，您可以在配置双向同步时选择。

## 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

 **注意** 购买时，选择源实例和目标实例均为PolarDB，并选择同步拓扑为双向同步。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击[数据同步](#)。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 配置正向同步作业。

i. 定位至已购买的数据同步实例，单击该实例下第一个同步作业操作列的配置同步链路。

**注意** 一个双向数据同步实例会包含两个同步作业，需要分别进行配置。在配置反向同步作业时，定位至第二个同步作业，单击操作列的配置同步链路。

| 同步作业名称 | 状态  | 同步概况 | 付费方式       | 同步架构(全部)               | 操作  |
|--------|-----|------|------------|------------------------|---|
| ...    | --  | --   | 按量付费       | 双向同步                   | <a href="#">转包年包月</a>   <a href="#">升级</a><br><a href="#">查看同步作业</a>   <a href="#">更多</a> |
| 作业名称   | 状态  | 同步概况 | 源实例/目标实例   | 操作                     |   |
| ...    | 未配置 |      | 未配置<br>未配置 | <a href="#">配置同步链路</a> |   |
| ...    | 未配置 |      | 未配置<br>未配置 | <a href="#">配置同步链路</a> |   |

ii. 配置同步作业的源集群及目标集群信息。

1.选择同步通道的源及目标实例
2.选择同步对象
3.高级设置
4.预检查

同步作业名称: PolarDB MySQL\_Forward

---

**源实例信息**

实例类型: PolarDB实例

实例地区: 华东1 (杭州)

\* PolarDB实例ID: pc-bp-...

\* 数据库账号: dtstest

\* 数据库密码: \*\*\*\*\*

---

**目标实例信息**

实例类型: PolarDB实例

实例地区: 华东1 (杭州)

\* PolarDB实例ID: pc-bp-...

\* 数据库账号: dtstest

\* 数据库密码: \*\*\*\*\*

取消 授权白名单并进入下一步

| 类别    | 配置          | 说明  |
|-------|-------------|---|
| 无     | 同步作业名称      | DTS会自动生成一个同步作业名称，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。                        |
| 源实例信息 | 实例类型        | 固定为PolarDB实例，不可变更。  |
|       | 实例地区        | 购买数据同步实例时选择的源实例地域信息，不可变更。   |
|       | PolarDB实例ID | 选择源PolarDB集群ID。<br><b>注意</b> 稍后配置反向同步作业时，此处需选择为正向同步作业里的目标PolarDB集群ID。 |
|       | 数据库账号       | 填入连接PolarDB集群的数据库账号，该账号需具备待同步对象读写权限。                                  |
|       | 数据库密码       | 填入数据库账号的密码。   |

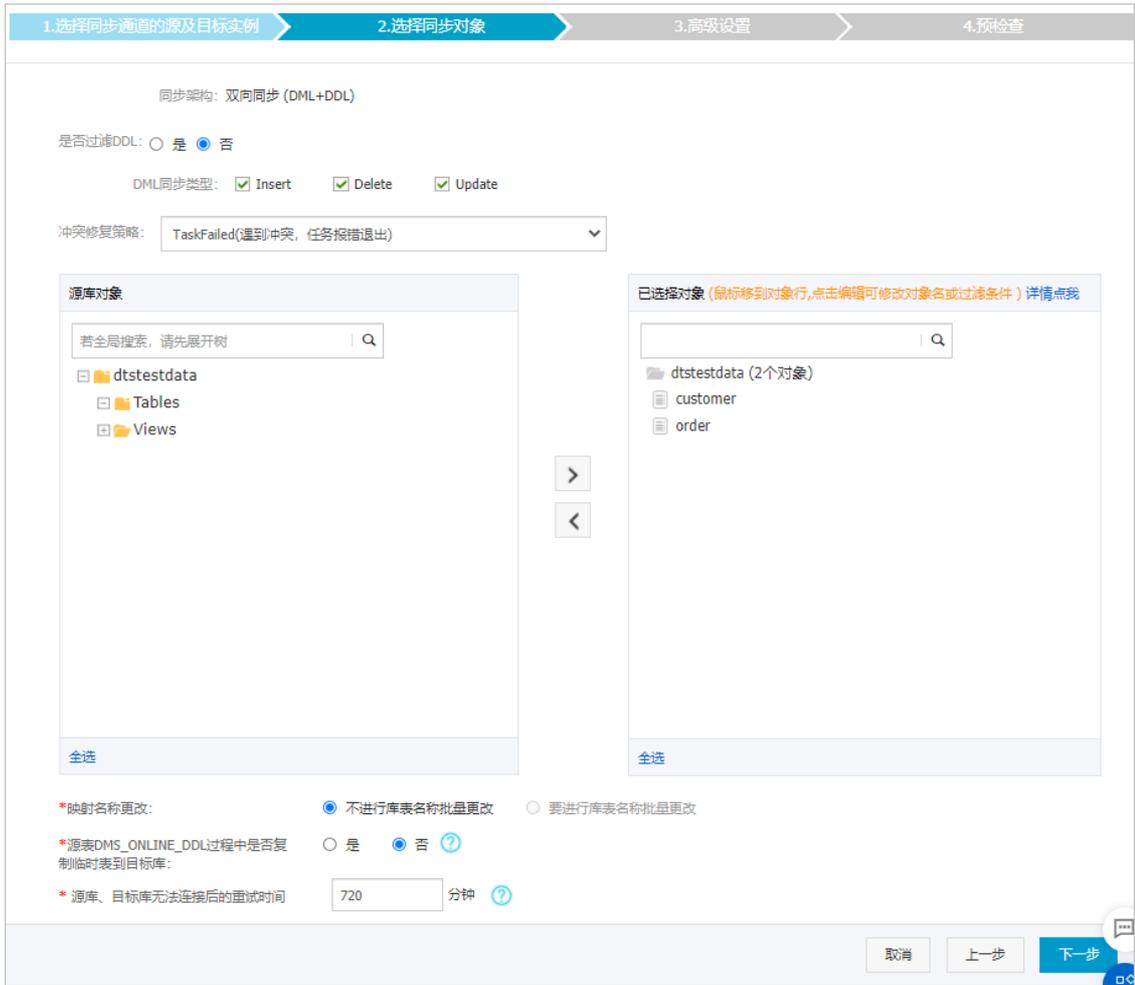
| 类别     | 配置          | 说明  |
|--------|-------------|---|
| 目标实例信息 | 实例类型        | 固定为PolarDB实例，不可变更。  |
|        | 实例地区        | 购买数据同步实例时选择的目标实例地域信息，不可变更。  |
|        | PolarDB实例ID | 选择目标PolarDB集群ID。<br> <b>注意</b> 稍后配置反向同步作业时，此处需选择为正向同步作业里的源PolarDB集群ID。 |
|        | 数据库账号       | 填入连接PolarDB集群的数据库账号，该账号需具备目标同步对象读写权限。   |
|        | 数据库密码       | 填入数据库账号的密码。   |

iii. 单击页面右下角的**授权白名单并进入下一步**。

 **说明**

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

iv. 配置同步策略及对象信息。



| 类别     | 配置      | 说明   |
|--------|---------|--|
| 同步策略配置 | 是否过滤DDL | <ul style="list-style-type: none"> <li>选择为是：不同步DDL操作。</li> <li>选择为否：同步DDL操作。</li> </ul> <p><b>注意</b> DDL语法同步方向限制。为保障双向同步链路的稳定性，只支持正向同步DDL，不支持反向同步DDL。</p>  |
|        | DML同步类型 | 选择需要同步的DML类型，默认为Insert、Update、Delete，您可以根据业务需求调整。  |
|        | 冲突修复策略  | <p>选择同步冲突的修复策略，默认为TaskFailed，您可以根据业务情况选择合适的冲突修复策略。</p> <ul style="list-style-type: none"> <li><b>TaskFailed</b>（遇到冲突，任务报错退出）<br/>默认的冲突修复策略。当数据同步遇到上述冲突类型时，同步任务直接报错并退出，同步任务进入失败状态，需要用户介入修复任务。</li> <li><b>Ignore</b>（遇到冲突，直接使用目标实例中的冲突记录）<br/>当数据同步遇到上述的冲突类型时，直接跳过当前同步语句，继续往下执行，选择使用目标库中的冲突记录。</li> <li><b>Overwrite</b>（遇到冲突，直接覆盖目标实例中的冲突记录）<br/>当数据同步遇到上述的冲突类型时，直接覆盖目标库中的冲突记录。</li> </ul> |

| 类别                              | 配置          | 说明   |
|---------------------------------|-------------|--|
|                                 | 目标已存在表的处理模式 | <ul style="list-style-type: none"> <li>■ <b>预检查并报错拦截</b>：检查目标库中是否有同名的表。如果目标库中没有同名的表，则通过该检查项目；如果目标库中有同名的表，则在预检查阶段提示错误，数据同步作业不会被启动。</li> </ul> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin: 5px 0;"> <p> <b>注意</b> 如果目标库中同名的表不方便删除或重命名，您可以<b>设置同步对象在目标实例中的名称</b>来避免表名冲突。</p> </div> <ul style="list-style-type: none"> <li>■ <b>忽略报错并继续执行</b>：跳过目标库中是否有同名表的检查项。</li> </ul> <div style="border: 1px solid #ccc; background-color: #fff9c4; padding: 5px; margin: 5px 0;"> <p> <b>警告</b> 选择为<b>忽略报错并继续执行</b>，可能导致数据不一致，给业务带来风险，例如：</p> <ul style="list-style-type: none"> <li>■ 表结构一致的情况下，如果在目标库遇到与源库主键的值相同的记录，在初始化阶段会保留目标库中的该条记录；在增量同步阶段则会覆盖目标库的该条记录。</li> <li>■ 表结构不一致的情况下，可能会导致无法初始化数据、只能同步部分列的数据或同步失败。</li> </ul> </div> |
| 选择同步对象                          | 无           | <p>在<b>源库对象框</b>中单击待同步的对象（选择粒度为库或表），然后单击  将其移动到<b>已选择对象框</b>。</p> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin: 5px 0;"> <p> <b>注意</b></p> <ul style="list-style-type: none"> <li>■ 如果选择整个库作为同步对象，该库中所有对象的结构变更操作都会同步至目标库。</li> <li>■ 默认情况下，同步对象的名称保持不变。如果您需要改变同步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<b>设置同步对象在目标实例中的名称</b>。</li> </ul> </div>   |
| 映射名称更改                          | 无           | <p>如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见<b>库表列映射</b>。</p>   |
| 源表 DMS_ONLINE_DDL过程中是否复制临时表到目标库 | 无           | <p>如源库使用<b>数据管理DMS (Data Management Service)</b> 执行Online DDL变更，您可以选择是否同步Online DDL变更产生的临时表数据。</p> <ul style="list-style-type: none"> <li>■ <b>是</b>：同步Online DDL变更产生的临时表数据。</li> </ul> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin: 5px 0;"> <p> <b>说明</b> Online DDL变更产生的临时表数据过大，可能会导致同步任务延迟。</p> </div> <ul style="list-style-type: none"> <li>■ <b>否</b>：不同步Online DDL变更产生的临时表数据，只同步源库的原始DDL数据。</li> </ul> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin: 5px 0;"> <p> <b>说明</b> 该方案会导致目标库锁表。</p> </div>   |

| 类别            | 配置 | 说明   |
|---------------|----|--|
| 源、目标库无法连接重试时间 | 无  | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p><b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p> </div> |

v. 上述配置完成后，单击页面右下角的下一步。

vi. 配置同步初始化的高级配置信息。



**注意**

- 此步骤会将源集群中待同步对象的结构及数据同步至目标集群，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，需要同时选择结构初始化和全量数据初始化。
- 在配置反向同步作业时，如果同步对象已经初始化至目标集群，则直接同步增量数据。

vii. 上述配置完成后，单击页面右下角的预检查并启动。

**说明**

- 在同步作业正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步作业。
- 如果预检查失败，单击具体检查项后的 ⓘ，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择确认屏蔽、忽略告警项并重新进行预检查，跳过告警检测项重新进行预检查。

viii. 在预检查对话框中显示预检查通过后，关闭预检查对话框，正向同步作业将正式开始。

6. 等待正向同步作业完成初始化，直至处于同步中状态。

您可以在数据同步页面，查看数据同步作业的状态。

7. 配置反向同步作业。

i. 定位至第二个同步作业，单击配置同步链路。

| 实例ID/作业名称                | 状态  | 同步概况                           | 付费方式                         | 同步架构(全部)                                   | 操作   |
|--------------------------|-----|--------------------------------|------------------------------|--|--|
| [实例ID]                   | --  | --                             | 按量付费                         | 双向同步                                       | <a href="#">转包年包月</a>   <a href="#">升级</a><br><a href="#">查看同步作业</a> |
| 作业名称                     | 状态  | 同步概况                           | 源实例/目标实例                     | 操作   |  |
| PolarDB MySQL_Forward    | 同步中 | 延时: 0 毫秒<br>速度: 0TPS(0.00MB/s) | pc-bp-...t5b<br>pc-bp-...3hf | <a href="#">暂停同步</a><br><a href="#">更多</a> |  |
| hangzhou-hangzhou-medium | 未配置 |                                | pc-bp-...3hf<br>pc-bp-...t5b | <a href="#">配置同步链路</a>                     |  |

ii. 重复步骤5中的ii到viii配置步骤，完成反向同步作业的配置。

### 执行结果

等待一段时间后，两个同步作业的链路状态均会处于同步中。

| 实例ID/作业名称             | 状态  | 同步概况                           | 付费方式                         | 同步架构(全部)                                   | 操作   |
|-----------------------|-----|--------------------------------|------------------------------|--|--|
| [实例ID]                | --  | --                             | 按量付费                         | 双向同步                                       | <a href="#">转包年包月</a>   <a href="#">升级</a><br><a href="#">查看同步作业</a> |
| 作业名称                  | 状态  | 同步概况                           | 源实例/目标实例                     | 操作   |  |
| PolarDB MySQL_Forward | 同步中 | 延时: 0 毫秒<br>速度: 0TPS(0.00MB/s) | pc-bp-...t5b<br>pc-bp-...3hf | <a href="#">暂停同步</a><br><a href="#">更多</a> |  |
| PolarDB MySQL_Reverse | 同步中 | 延时: 0 毫秒<br>速度: 0TPS(0.00MB/s) | pc-bp-...3hf<br>pc-bp-...t5b | <a href="#">暂停同步</a><br><a href="#">更多</a> |  |

共有1条， 每页显示: 20条

## 8.2. PolarDB MySQL集群间的单向同步

PolarDB是阿里巴巴自主研发的下一代关系型分布式云原生数据库，可完全兼容MySQL，具备简单易用、高性能、高可靠、高可用等优势。通过数据传输服务DTS（Data Transmission Service），您可以实现PolarDB MySQL集群间的单向数据同步。

### 前提条件

- 已购买源和目标PolarDB MySQL集群，详情请参见[创建PolarDB MySQL集群](#)。支持的数据库版本，请参见[同步方案概览](#)。
- 源PolarDB MySQL集群已开启Binlog，详情请参见[如何开启Binlog](#)。

### 注意事项

- DTS在执行全量数据初始化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。因此您需要在执行数据同步前评估源库和目标库的性能，同时建议您在业务低峰期执行数据同步（例如源库和目标库的CPU负载在30%以下）。
- 全量初始化过程中，并发INSERT会导致目标集群的表碎片，全量初始化完成后，目标集群的表空间比源库的表空间大。
- 如果数据同步的源库没有主键或唯一约束，且记录的全字段没有唯一性，可能会出现重复数据。
- 建议源和目标PolarDB MySQL集群的数据库版本保持一致，或者从低版本同步到高版本以保障兼容性。

### 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

## 支持同步的SQL操作

| 操作类型 | SQL操作语句   |
|------|---|
| DML  | INSERT、UPDATE、DELETE、REPLACE  |
| DDL  | <ul style="list-style-type: none"> <li>ALTER TABLE、ALTER VIEW</li> <li>CREATE FUNCTION、CREATE INDEX、CREATE PROCEDURE、CREATE TABLE、CREATE VIEW</li> <li>DROP INDEX、DROP TABLE</li> <li>RENAME TABLE</li> <li>TRUNCATE TABLE</li> </ul> |

## 支持的同步架构

- 一对一单向同步
- 一对多单向同步
- 级联单向同步
- 多对一单向同步

关于各类同步架构的介绍及注意事项，请参见[数据同步拓扑介绍](#)。

## 功能限制

- 不兼容触发器  
当同步对象为整个库，且库中的触发器（TRIGGER）会更新库内某个表时，可能导致源和目标库的数据不一致。相关解决方案请参见[源库存在触发器时如何配置同步作业](#)。
- RENAME TABLE限制  
RENAME TABLE操作可能导致同步数据不一致。例如同步对象只包含某个表，如果同步过程中源实例对该表执行了重命名操作，那么该表的数据将不会同步到目标库。为避免该问题，您可以在数据同步配置时将该表所属的整个数据库作为同步对象。

## 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

 说明 购买时，选择源实例和目标实例均为PolarDB，并选择同步拓扑为单向同步。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击[数据同步](#)。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击[配置同步链路](#)。
6. 配置同步通道的源实例及目标实例信息。

1.选择同步通道的源及目标实例 > 2.选择同步对象 > 3.高级设置 > 4.预检查

同步作业名称: POLARDB\_TO\_POLARDB

**源实例信息**

实例类型: PolarDB实例  
实例地区: 华东1 (杭州)  
\* PolarDB实例ID: pc-bp-  
\* 数据库账号: dtstest  
\* 数据库密码: \*\*\*\*\*

**目标实例信息**

实例类型: PolarDB  
实例地区: 华东1 (杭州)  
\* PolarDB实例ID: pc-bp-  
\* 数据库账号: dtstest  
\* 数据库密码: \*\*\*\*\*

取消 授权白名单并进入下一步

| 类别     | 配置          | 说明  |
|--------|-------------|---|
| 无      | 同步作业名称      | DTS会自动生成一个同步作业名称，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。  |
| 源实例信息  | 实例类型        | 固定为PolarDB实例，不可变更。  |
|        | 实例地区        | 购买数据同步实例时选择的源实例地域信息，不可变更。   |
|        | PolarDB实例ID | 选择源PolarDB集群ID。   |
|        | 数据库账号       | 填入连接PolarDB集群的数据库账号。<br><b>说明</b> 该账号需具备REPLICATION CLIENT、REPLICATION SLAVE、SHOW VIEW和所有同步对象的SELECT权限。 |
|        | 数据库密码       | 填入数据库账号对应的密码。   |
| 目标实例信息 | 实例类型        | 固定为PolarDB，不可变更。  |
|        | 实例地区        | 购买数据同步实例时选择的目标实例地域信息，不可变更。  |
|        | PolarDB实例ID | 选择目标PolarDB集群ID。  |
|        | 数据库账号       | 填入连接PolarDB集群的数据库账号。<br><b>说明</b> 用于数据同步的数据库账号需具备目标同步对象的ALL权限。  |
|        | 数据库密码       | 填入数据库账号对应的密码。   |

7. 单击页面右下角的授权白名单并进入下一步。

**说明**

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

**8. 配置目标已存在表的处理模式和同步对象。**

同步架构: 单向同步 (DML+DDL)

1.选择同步通道的源及目标实例 2.选择同步对象 3.高级设置 4.预检查

**源库对象**

若全局搜索, 请先展开树

- dtstestdata
  - Tables
  - Views

**已选择对象** (鼠标移到对象行, 点击编辑可修改对象名或过滤条件) 详情点我

- dtstestdata (2个对象)
  - customer
  - order

\* 映射名称更改:  不进行库表名称批量更改  要进行库表名称批量更改

\* 源表DMS\_ONLINE\_DDL过程中是否复制临时表到目标库:  是  否

\* 源库、目标库无法连接后的重试时间: 720 分钟

取消 上一步 下一步

| 配置项目 | 配置说明 |
|------|------|
|------|------|

| 配置项目                           | 配置说明  |
|--------------------------------|---|
| 目标已存在表的处理模式                    | <ul style="list-style-type: none"> <li>◦ <b>预检查并报错拦截</b>：检查目标数据库中是否有同名的表。如果目标数据库中没有同名的表，则通过该检查项目；如果目标数据库中有同名的表，则在预检查阶段提示错误，数据同步作业不会被启动。                     <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin: 5px 0;"> <p> <b>说明</b> 如果目标库中同名的表不方便删除或重命名，您可以<a href="#">设置同步对象在目标实例中的名称</a>来避免表名冲突。</p> </div> </li> <li>◦ <b>忽略报错并继续执行</b>：跳过目标数据库中是否有同名表的检查项。                     <div style="border: 1px solid #ccc; background-color: #fff9c4; padding: 5px; margin: 5px 0;"> <p> <b>警告</b> 选择为忽略报错并继续执行，可能导致数据不一致，给业务带来风险，例如：</p> <ul style="list-style-type: none"> <li>■ 表结构一致的情况下，如果在目标库遇到与源库主键的值相同的记录，在初始化阶段会保留目标库中的该条记录；在增量同步阶段则会覆盖目标库的该条记录。</li> <li>■ 表结构不一致的情况下，可能会导致无法初始化数据、只能同步部分列的数据或同步失败。</li> </ul> </div> </li> </ul> |
| 选择同步对象                         | <p>在源库对象框中单击待同步的对象，然后单击  图标将其移动至已选择对象框。</p> <p>同步对象的选择粒度为库、表。</p> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin: 5px 0;"> <p> <b>说明</b></p> <ul style="list-style-type: none"> <li>◦ 如果选择整个库作为同步对象，那么该库中所有对象的结构变更操作都会同步至目标库。</li> <li>◦ 默认情况下，同步对象的名称保持不变。如果您需要改变同步对象在目标集群中的名称，请使用对象名映射功能，详情请参见<a href="#">设置同步对象在目标实例中的名称</a>。</li> </ul> </div>  |
| 映射名称更改                         | <p>如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见<a href="#">库表列映射</a>。</p>   |
| 源表DMS_ONLINE_DDL过程中是否复制临时表到目标库 | <p>如源库使用<a href="#">数据管理DMS (Data Management Service)</a> 执行Online DDL变更，您可以选择是否同步Online DDL变更产生的临时表数据。</p> <ul style="list-style-type: none"> <li>◦ <b>是</b>：同步Online DDL变更产生的临时表数据。             <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin: 5px 0;"> <p> <b>说明</b> Online DDL变更产生的临时表数据过大，可能会导致同步任务延迟。</p> </div> </li> <li>◦ <b>否</b>：不同步Online DDL变更产生的临时表数据，只同步源库的原始DDL数据。             <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin: 5px 0;"> <p> <b>说明</b> 该方案会导致目标库锁表。</p> </div> </li> </ul>  |

| 配置项目          | 配置说明  |
|---------------|---|
| 源、目标库无法连接重试时间 | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <p><b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p> |

9. 上述配置完成后，单击页面右下角的下一步。

10. 配置同步初始化的高级配置信息。



**说明** 同步初始化类型细分为：结构初始化，全量数据初始化。选中结构初始化和全量数据初始化后，DTS会在增量数据同步之前，将源数据库中待同步对象的结构和存量数据，同步到目标数据库。

11. 上述配置完成后，单击页面右下角的预检查并启动。

**说明**

- 在同步作业正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步作业。
- 如果预检查失败，单击具体检查项后的 ，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择**确认屏蔽**、**忽略告警项**并重新进行预检查，跳过告警检测项重新进行预检查。

12. 在预检查对话框中显示预检查通过后，关闭预检查对话框，同步作业将正式开始。

13. 等待同步作业的链路初始化完成，直至处于同步中状态。

您可以在数据同步页面，查看数据同步作业的状态。



## 8.3. 从PolarDB MySQL同步至RDS MySQL

本文介绍如何使用数据传输服务DTS（Data Transmission Service），将PolarDB MySQL同步至RDS MySQL，实现增量数据的实时同步。

### 前提条件

- 已购买PolarDB MySQL集群，详情请参见[创建PolarDB MySQL集群](#)。
- PolarDB MySQL集群已开启Binlog，详情请参见[如何开启Binlog](#)。

## 注意事项

- DTS在执行全量数据初始化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。因此您需要在执行数据同步前评估源库和目标库的性能，同时建议您在业务低峰期执行数据同步（例如源库和目标库的CPU负载在30%以下）。
- 全量初始化过程中，并发INSERT会导致目标集群的表碎片，全量初始化完成后，目标集群的表空间比源库的表空间大。
- 如果数据同步的源库没有主键或唯一约束，且记录的全字段没有唯一性，可能会出现重复数据。

## 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

## 支持同步的SQL操作

| 操作类型 | SQL操作语句   |
|------|---|
| DML  | INSERT、UPDATE、DELETE、REPLACE  |
| DDL  | <ul style="list-style-type: none"> <li>• ALTER TABLE、ALTER VIEW</li> <li>• CREATE FUNCTION、CREATE INDEX、CREATE PROCEDURE、CREATE TABLE、CREATE VIEW</li> <li>• DROP INDEX、DROP TABLE</li> <li>• RENAME TABLE</li> <li>• TRUNCATE TABLE</li> </ul> |

## 支持的同步架构

- 一对一单向同步
- 一对多单向同步
- 级联单向同步
- 多对一单向同步

关于各类同步架构的介绍及注意事项，请参见[数据同步拓扑介绍](#)。

## 功能限制

- 不兼容触发器  
当同步对象为整个库，且库中的触发器（TRIGGER）会更新库内某个表时，可能导致源和目标库的数据不一致。相关解决方案请参见[源库存在触发器时如何配置同步作业](#)。
- RENAME TABLE限制  
RENAME TABLE操作可能导致同步数据不一致。例如同步对象只包含表A，如果同步过程中源实例将表A重命名为表B，那么表B将不会被同步到目标库。为避免该问题，您可以在数据同步配置时，选择同步表A和表B所在的整个数据库作为同步对象。

## 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

 说明 购买时，选择源实例为PolarDB，选择目标实例为MySQL，并选择同步拓扑为单向同步。

2. 登录[数据传输控制台](#)。

3. 在左侧导航栏，单击数据同步。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击配置同步链路。
6. 配置同步通道的源实例及目标实例信息。

1. 选择同步通道的源及目标实例
2. 选择同步对象
3. 高级设置
4. 预检查

同步作业名称:

**源实例信息**

实例类型:

实例地区: 华东1 (杭州)

\* PolarDB实例ID:

\* 数据库账号:

\* 数据库密码:

**目标实例信息**

实例类型:

实例地区: 华东1 (杭州)

\* 实例ID:

\* 数据库账号:

\* 数据库密码:

\* 连接方式:  非加密连接  SSL安全连接

| 类别     | 配置          | 说明  |
|--------|-------------|---|
| 无      | 同步作业名称      | DTS会自动生成一个同步作业名称，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。  |
| 源实例信息  | 实例类型        | 固定为PolarDB实例，不可变更。  |
|        | 实例地区        | 购买数据同步实例时选择的源实例地域信息，不可变更。   |
|        | PolarDB实例ID | 选择源PolarDB集群ID。   |
|        | 数据库账号       | 填入源PolarDB集群的数据库账号。   |
|        | 数据库密码       | 填入数据库账号对应的密码。   |
| 目标实例信息 | 实例类型        | 选择为RDS实例。   |
|        | 实例地区        | 购买数据同步实例时选择的目标实例地域信息，不可变更。  |
|        | 数据库账号       | 填入目标RDS实例的数据库账号。  |
|        | 数据库密码       | 填入数据库账号对应的密码。<br><div style="border: 1px solid #add8e6; padding: 5px; margin-top: 5px;"> <span style="font-size: 1.2em;">?</span> <b>说明</b> 当目标RDS实例的数据库类型为MySQL 5.5或MySQL 5.6时，无需配置数据库账号和数据库密码。                 </div> |

| 类别 | 配置   | 说明  |
|----|------|---|
|    | 连接方式 | <p>根据需求选择非加密连接或SSL安全连接。如果设置为SSL安全连接，您需要提前开启RDS实例的SSL加密功能，详情请参见<a href="#">设置SSL加密</a>。</p> <p><b>注意</b> 目前仅中国内地及中国香港地域支持设置连接方式。</p> |

7. 单击页面右下角的授权白名单并进入下一步。

**说明**

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置目标已存在表的处理模式和同步对象。

同步架构: 单向同步 (DML+DDL)

源库对象

若全局搜索，请先展开树

- dtstestdata
  - Tables
  - Views

已选择对象 (鼠标移到对象行, 点击编辑可修改对象名或过滤条件) 详情点我

- dtstestdata (2个对象)
  - customer
  - order

\* 映射名称更改:  不进行库表名称批量更改  要进行库表名称批量更改

\* 源表DMS\_ONLINE\_DDL过程中是否复制临时表到目标库:  是  否

\* 源库、目标库无法连接后的重试时间: 720 分钟

取消 上一步 下一步

| 配置项目 | 配置说明 |
|------|------|
|------|------|

| 配置项目                           | 配置说明   |
|--------------------------------|--|
| 目标已存在表的处理模式                    | <ul style="list-style-type: none"> <li> <b>预检查并报错拦截</b>：检查目标数据库中是否有同名的表。如果目标数据库中没有同名的表，则通过该检查项目；如果目标数据库中有同名的表，则在预检查阶段提示错误，数据同步作业不会被启动。                     </li> </ul> <p><b>说明</b> 如果目标库中同名的表不方便删除或重命名，您可以<b>设置同步对象在目标实例中的名称</b>来避免表名冲突。</p> <ul style="list-style-type: none"> <li> <b>忽略报错并继续执行</b>：跳过目标数据库中是否有同名表的检查项。                     </li> </ul> <p><b>警告</b> 选择为<b>忽略报错并继续执行</b>，可能导致数据不一致，给业务带来风险，例如：</p> <ul style="list-style-type: none"> <li>表结构一致的情况下，如果在目标库遇到与源库主键的值相同的记录，在初始化阶段会保留目标库中的该条记录；在增量同步阶段则会覆盖目标库的该条记录。</li> <li>表结构不一致的情况下，可能会导致无法初始化数据、只能同步部分列的数据或同步失败。</li> </ul> |
| 选择同步对象                         | <p>在<b>源库对象框</b>中单击待同步的对象，然后单击  图标将其移动至<b>已选择对象框</b>。</p> <p>同步对象的选择粒度为库、表。</p> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>如果选择整个库作为同步对象，那么该库中所有对象的结构变更操作都会同步至目标库。</li> <li>默认情况下，同步对象的名称保持不变。如果您需要改变同步对象在目标集群中的名称，请使用对象名映射功能，详情请参见<b>设置同步对象在目标实例中的名称</b>。</li> </ul>  |
| 映射名称更改                         | <p>如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见<b>库表列映射</b>。</p>   |
| 源表DMS_ONLINE_DDL过程中是否复制临时表到目标库 | <p>如源库使用<b>数据管理DMS (Data Management Service)</b> 执行Online DDL变更，您可以选择是否同步Online DDL变更产生的临时表数据。</p> <ul style="list-style-type: none"> <li> <b>是</b>：同步Online DDL变更产生的临时表数据。                     </li> </ul> <p><b>说明</b> Online DDL变更产生的临时表数据过大，可能会导致同步任务延迟。</p> <ul style="list-style-type: none"> <li> <b>否</b>：不同步Online DDL变更产生的临时表数据，只同步源库的原始DDL数据。                     </li> </ul> <p><b>说明</b> 该方案会导致目标库锁表。</p>   |
| 源、目标库无法连接重试时间                  | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <p><b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p>  |

9. 上述配置完成后，单击页面右下角的下一步。

10. 配置同步初始化的高级配置信息。



**说明** 同步初始化类型细分为：结构初始化，全量数据初始化。选中结构初始化和全量数据初始化后，DTS会在增量数据同步之前，将源数据库中待同步对象的结构和存量数据，同步到目标数据库。

11. 上述配置完成后，单击页面右下角的预检查并启动。

**说明**

- 在同步作业正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步作业。
- 如果预检查失败，单击具体检查项后的 ，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择确认屏蔽、忽略告警项并重新进行预检查，跳过告警检测项重新进行预检查。

12. 在预检查对话框中显示预检查通过后，关闭预检查对话框，同步作业将正式开始。

13. 等待同步作业的链路初始化完成，直至处于同步中状态。

您可以在数据同步页面，查看数据同步作业的状态。



## 8.4. 从PolarDB MySQL同步至Datahub

阿里云流式数据服务DataHub是流式数据（Streaming Data）的处理平台，提供对流式数据的发布、订阅和分发功能，让您可以轻松构建基于流式数据的分析和应用。通过数据传输服务（Data Transmission Service，简称DTS），您可以将PolarDB MySQL同步至DataHub，帮助您快速实现使用流计算等大数据产品对数据实时分析。

### 前提条件

- DataHub实例的地域为华东1、华东2、华北2或华南1。
- DataHub实例中，已创建用作接收同步数据的项目（Project），详情请参见[创建项目](#)。
- PolarDB MySQL集群已开启Binlog，详情请参见[如何开启Binlog](#)。
- PolarDB MySQL中待同步的表需具备主键或唯一约束。

### 费用说明

|               |                                 |
|---------------|---------------------------------|
| 同步类型          | 链路配置费用                          |
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

### 功能限制

- 不支持全量数据初始化，即DTS不会将源PolarDB集群中同步对象的存量数据同步至目标DataHub实例。
- 仅支持表级别的数据同步。
- 不支持新增列的数据同步，即源数据表新增了某个列，该列的数据不会同步至目标DataHub实例。
- 数据同步的过程中，请勿对源库中待同步的表执行DDL变更，否则会导致同步失败。

### 支持同步的SQL操作

INSERT、UPDATE、DELETE。

### 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

 **说明** 购买时，选择源实例为PolarDB、目标实例为DataHub，并选择同步拓扑为单向同步。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击[数据同步](#)。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击[配置同步链路](#)。
6. 配置同步作业的源实例及目标实例信息。

| 类别    | 配置          | 说明   |
|-------|-------------|--|
| 无     | 同步作业名称      | DTS会自动生成一个同步作业名称，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。 |
| 源实例信息 | 实例类型        | 固定为PolarDB实例，不可变更。                             |
|       | 实例地区        | 购买数据同步实例时选择的源实例地域信息，不可变更。                      |
|       | PolarDB实例ID | 选择源PolarDB集群ID。                                |
|       | 数据库账号       | 填入PolarDB集群的数据库账号。                             |
|       | 数据库密码       | 填入数据库账号对应的密码。                                  |

| 类别     | 配置      | 说明                         |
|--------|---------|----------------------------|
| 目标实例信息 | 实例类型    | 固定为DataHub，不可变更。           |
|        | 实例地区    | 购买数据同步实例时选择的目标实例地域信息，不可变更。 |
|        | Project | 选择DataHub实例的Project。       |

7. 单击页面右下角的授权白名单并进入下一步。

**说明**

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

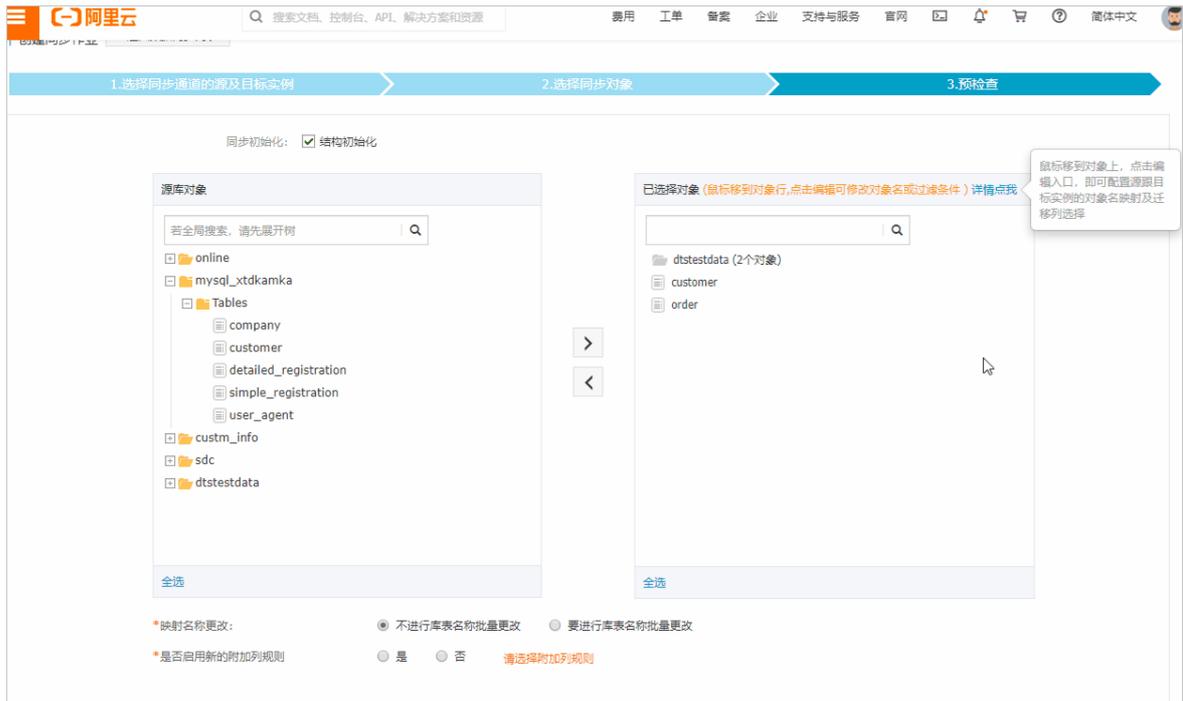
8. 配置同步策略和同步对象。



| 配置    | 说明   |
|-------|--|
| 同步初始化 | <p>勾选结构初始化。</p> <p><b>说明</b> 勾选结构初始化后，在数据同步作业的初始化阶段，DTS会将同步对象的结构信息（例如表结构）同步至目标DataHub实例。</p> |

| 配置                             | 说明  |
|--------------------------------|---|
| 选择同步对象                         | <p>在源库对象框中单击待迁移的对象，然后单击  将其移动至已选择对象框。</p> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin-top: 10px;"> <p> <b>说明</b></p> <ul style="list-style-type: none"> <li>◦ 同步对象的选择粒度为表。</li> <li>◦ 默认情况下，同步对象的名称保持不变。如果您需要改变同步对象在目标实例中的名称，需要使用对象名映射功能，详情请参见<a href="#">设置同步对象在目标实例中的名称</a>。</li> </ul> </div>  |
| 选择附加列规则                        | <p>DTS在将数据同步到DataHub时，会在同步的目标Topic中添加一些附加列。如果附加列和目标Topic中已有的列出现名称冲突将会导致数据同步失败。您需要根据业务需求选择是否启用新的附加列规则为是或否。</p> <div style="border: 1px solid #ccc; background-color: #fff9c4; padding: 5px; margin-top: 10px;"> <p> <b>警告</b> 在选择附加列规则前，您需要评估附加列和目标Topic中已有的列是否会出现名称冲突。关于附加列的规则和定义说明，请参见<a href="#">附加列名称和定义说明</a>。</p> </div>  |
| 映射名称更改                         | <p>如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见<a href="#">库表列映射</a>。</p>   |
| 源表DMS_ONLINE_DDL过程中是否复制临时表到目标库 | <p>如源库使用<a href="#">数据管理DMS (Data Management Service)</a> 执行Online DDL变更，您可以选择是否同步Online DDL变更产生的临时表数据。</p> <ul style="list-style-type: none"> <li>◦ 是：同步Online DDL变更产生的临时表数据。</li> </ul> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin-top: 10px;"> <p> <b>说明</b> Online DDL变更产生的临时表数据过大，可能会导致同步任务延迟。</p> </div> <ul style="list-style-type: none"> <li>◦ 否：不同步Online DDL变更产生的临时表数据，只同步源库的原始DDL数据。</li> </ul> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin-top: 10px;"> <p> <b>说明</b> 该方案会导致目标库锁表。</p> </div> |
| 源、目标库无法连接重试时间                  | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin-top: 10px;"> <p> <b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p> </div>   |

9. (可选) 将鼠标指针放置在已选择对象框中待同步的Topic名上，单击对象后出现的编辑，然后在弹出的对话框中设置Shardkey（即用于分区的key）。



10. 上述配置完成后，单击页面右下角的预检查并启动。

**说明**

- 在同步作业正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步作业。
- 如果预检查失败，单击具体检查项后的，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择**确认屏蔽**、**忽略告警项**并重新进行预检查，跳过告警检测项重新进行预检查。

11. 在预检查对话框中显示预检查通过后，关闭预检查对话框，同步作业将正式开始。

12. 等待同步作业的链路初始化完成，直至处于同步中状态。

您可以在数据同步页面，查看数据同步作业的状态。



### Topic结构定义说明

DTS在将数据变更同步至DataHub实例的Topic时，目标Topic中除了存储变更数据外，还会新增一些附加列用于存储元信息，示例如下。

**说明** 本案例中的业务字段为 `id`、`name`、`address`，由于在配置数据同步时选用的是旧版附加列规则，DTS会为业务字段添加 `dtst_` 的前缀。

| dts_id | dts_name | dts_address | dts_record_id       | dts_operation_flag | dts_instance_id | dts_db_name | dts_table_name | dts_utc_timestamp | dts_before_flag | dts_after_flag |
|--------|----------|-------------|---------------------|--------------------|-----------------|-------------|----------------|-------------------|-----------------|----------------|
| 10006  |          | 曲靖市         | 1574832130000000000 | U                  |                 | dtstestdata | customer       | 1574832130        | Y               | N              |
| 10006  |          | 杭州市         | 1574832130000000000 | U                  |                 | dtstestdata | customer       | 1574832130        | N               | Y              |
| 10009  |          | 马鞍山市        | 1574832919000000000 | D                  |                 | dtstestdata | customer       | 1574832919        | Y               | N              |
| 10112  |          | 北京市         | 1574832919000000000 | I                  |                 | dtstestdata | customer       | 1574832919        | N               | Y              |

结构定义说明：

| 旧版附加列名称                         | 新版附加列名称                                      | 数据类型   | 说明   |
|---------------------------------|--|--------|--|
| <code>dts_record_id</code>      | <code>new_dts_sync_dts_record_id</code>      | String | 增量日志的记录ID，为该日志唯一标识。<br><div style="border: 1px solid #add8e6; padding: 5px; margin-top: 10px;"> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>• 正常情况下是全局唯一自增的，在容灾的情况下会有回退且无法保证自增和唯一。</li> <li>• 如果增量日志的操作类型为UPDATE，那么增量更新会被拆分成两条记录（分别记录更新前和更新后的值），且 <code>dts_record_id</code> 的值相同。</li> </ul> </div> |
| <code>dts_operation_flag</code> | <code>new_dts_sync_dts_operation_flag</code> | String | 操作类型，取值： <ul style="list-style-type: none"> <li>• I: INSERT操作。</li> <li>• D: DELETE操作。</li> <li>• U: UPDATE操作。</li> </ul>  |
| <code>dts_instance_id</code>    | <code>new_dts_sync_dts_instance_id</code>    | String | 数据库的server ID。暂不支持显示实际的值，目前固定为 <code>null</code> 。   |
| <code>dts_db_name</code>        | <code>new_dts_sync_dts_database_name</code>  | String | 数据库名称。   |
| <code>dts_table_name</code>     | <code>new_dts_sync_dts_table_name</code>     | String | 表名。  |
| <code>dts_utc_timestamp</code>  | <code>new_dts_sync_dts_utc_timestamp</code>  | String | 操作时间戳，即binlog的时间戳（UTC时间）。  |
| <code>dts_before_flag</code>    | <code>new_dts_sync_dts_before_flag</code>    | String | 所有列的值是否更新前的值，取值：Y或N。   |
| <code>dts_after_flag</code>     | <code>new_dts_sync_dts_after_flag</code>     | String | 所有列的值是否更新后的值，取值：Y或N。   |

### 关于dts\_before\_flag和dts\_after\_flag的补充说明

对于不同的操作类型，增量日志中的 `dts_before_flag` 和 `dts_after_flag` 定义如下：

• INSERT

当操作类型为INSERT时，所有列的值为新插入的值，即为更新后的值，所以 `dts_before_flag` 取值为N，`dts_after_flag` 取值为Y，示例如下。

| dts_id | dts_name | dts_address | dts_record_id       | dts_operation_flag | dts_instance_id | dts_db_name | dts_table_name | dts_utc_timestamp | dts_before_flag | dts_after_flag |
|--------|----------|-------------|---------------------|--------------------|-----------------|-------------|----------------|-------------------|-----------------|----------------|
| 10112  |          | 北京市         | 1574832919000000000 | I                  |                 | dtstestdata | customer       | 1574832919        | N               | Y              |

• UPDATE

当操作类型为UPDATE时，DTS会将UPDATE操作拆为两条增量日志。这两条增量日志的 `dts_record_id`、`dts_operation_flag` 及 `dts_utc_timestamp` 对应的值相同。

第一条增量日志记录了更新前的值，所以 `dts_before_flag` 取值为Y，`dts_after_flag` 取值为N。第二条增量日志记录了更新后的值，所以 `dts_before_flag` 取值为N，`dts_after_flag` 取值为Y，示例如下。

| dts_id | dts_name | dts_address | dts_record_id       | dts_operation_flag | dts_instance_id | dts_db_name | dts_table_name | dts_utc_timestamp | dts_before_flag | dts_after_flag |
|--------|----------|-------------|---------------------|--------------------|-----------------|-------------|----------------|-------------------|-----------------|----------------|
| 10006  |          | 曲靖市         | 1574832130000000000 | U                  |                 | dtstestdata | customer       | 1574832130        | Y               | N              |
| 10006  |          | 杭州市         | 1574832130000000000 | U                  |                 | dtstestdata | customer       | 1574832130        | N               | Y              |

#### • DELETE

当操作类型为DELETE时，增量日志中所有的列值为被删除的值，即列值不变，所以 `dts_before_flag` 取值为Y，`dts_after_flag` 取值为N，示例如下。

| dts_id | dts_name | dts_address | dts_record_id       | dts_operation_flag | dts_instance_id | dts_db_name | dts_table_name | dts_utc_timestamp | dts_before_flag | dts_after_flag |
|--------|----------|-------------|---------------------|--------------------|-----------------|-------------|----------------|-------------------|-----------------|----------------|
| 10009  |          | 马鞍山市        | 1574832919000000000 | D                  |                 | dtstestdata | customer       | 1574832919        | Y               | N              |

## 后续操作

配置完数据同步作业后，您可以对同步到DataHub实例中的数据执行计算分析。更多详情，请参见[阿里云实时计算](#)。

# 8.5. 从PolarDB MySQL同步至Elasticsearch

阿里云Elasticsearch完全兼容开源Elasticsearch的功能及Security、Machine Learning、Graph、APM等商业功能，致力于数据分析、数据搜索等场景服务，支持企业级权限管控、安全监控告警、自动报表生成等功能。通过数据传输服务DTS（Data Transmission Service），您可以将PolarDB MySQL同步至Elasticsearch，帮助您快速构建数据。

## 前提条件

- 已创建目标Elasticsearch实例（5.5、5.6、6.3、6.7或7.4版本），详情请参见[开通阿里云Elasticsearch服务](#)。
- 已开启PolarDB MySQL的Binlog，详情请参见[如何开启Binlog](#)。

## 注意事项

- DTS在执行全量数据初始化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。因此您需要在执行数据同步前评估源库和目标库的性能，同时建议您在业务低峰期执行数据同步（例如源库和目标库的CPU负载在30%以下）。
- 不支持同步DDL操作，如果源库中待同步的表在同步的过程中，已经执行了DDL操作，您需要先移除同步对象，然后在Elasticsearch实例中移除该表对应的索引，最后新增同步对象。详情请参见[移除同步对象](#)和[新增同步对象](#)。
- 如果源库中待同步的表需要执行增加列的操作，您需要先在Elasticsearch实例中修改对应表的mapping，然后在PolarDB MySQL中执行相应的DDL操作，最后暂停并启动DTS同步实例。

## 费用说明

|               |                                 |
|---------------|---------------------------------|
| 同步类型          | 链路配置费用                          |
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

## 支持同步的SQL操作

INSERT、DELETE、UPDATE

## 数据类型映射关系

由于PolarDB MySQL与Elasticsearch实例支持的数据类型不同，数据类型无法完全对应。DTS在进行结构初始化时，会根据Elasticsearch实例支持的数据类型进行类型映射，详情请参见[结构初始化涉及的数据类型映射关系](#)。

## 操作步骤

（可选）

1. 购买数据同步作业，详情请参见[购买流程](#)。

**说明** 购买时，选择源实例为PolarDB、目标实例为Elasticsearch，并选择同步拓扑为单向同步。

2. 登录**数据传输控制台**。
3. 在左侧导航栏，单击**数据同步**。
4. 在**同步作业列表**页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击**配置同步链路**。
6. 配置同步作业的源实例及目标实例信息。

| 类别     | 配置            | 说明   |
|--------|---------------|--|
| 无      | 同步作业名称        | DTS会自动生成一个同步作业名称，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。           |
| 源实例信息  | 实例类型          | 固定为 <b>PolarDB实例</b> ，不可变更。                              |
|        | 实例地区          | 购买数据同步实例时选择的源实例地域信息，不可变更。                                |
|        | PolarDB实例ID   | 选择源PolarDB MySQL集群ID。                                    |
|        | 数据库账号         | 填入PolarDB MySQL的数据库账号。<br><b>说明</b> 该账号需具备同步对象所属数据库的读权限。 |
|        | 数据库密码         | 填入数据库账号对应的密码。  |
| 目标实例信息 | 实例类型          | 固定为 <b>Elasticsearch</b> ，不可变更。                          |
|        | 实例地区          | 购买数据同步实例时选择的目标实例地域信息，不可变更。                               |
|        | Elasticsearch | 选择目标Elasticsearch实例ID。                                   |
|        | 数据库账号         | 填入连接Elasticsearch实例的账号，默认账号为elastic。                     |

| 类别 | 配置    | 说明          |
|----|-------|-------------|
|    | 数据库密码 | 填入该账号对应的密码。 |

7. 单击页面右下角的授权白名单并进入下一步。

**说明**

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置索引名称、目标已存在表的处理模式和同步对象。

The screenshot shows the '2. 选择同步对象' (Select Sync Objects) step in the DTS configuration wizard. The '同步架构' (Sync Architecture) is set to '单向同步' (One-way Sync). The '索引名称' (Index Name) is set to '库名\_表名' (Database Name\_Table Name). Under '目标已存在表的处理模式' (Target Existing Table Handling Mode), '预检查并报错拦截' (Pre-check and intercept with error) is selected. The '源库对象' (Source Object) tree shows 'sys', 'dtstestdata', and 'Tables'. The '已选择对象' (Selected Objects) list shows 'dtstestdata (2个对象)' with sub-items 'dtstestdata\_customer' (源表名: customer) and 'dtstestdata\_order' (源表名: order). At the bottom, there are options for '映射名称更改' (Mapping Name Change) and '源表DMS\_ONLINE\_DDL过程中是否复制临时表到目标库' (Whether to copy temporary tables to the target database during the DMS\_ONLINE\_DDL process of the source table), along with a '源库、目标库无法连接后的重试时间' (Retry time after source/target database connection failure) set to 720 minutes.

| 配置   | 说明  |
|------|---|
| 索引名称 | <ul style="list-style-type: none"> <li><b>表名</b><br/>选择为表名后，在目标Elasticsearch实例中创建的索引名称和表名一致，在本案例中即为 customer。</li> <li><b>库名_表名</b><br/>选择为库名_表名后，在目标Elasticsearch实例中创建的索引名称为库名_表名，在本案例中即为 dtstestdata_customer。</li> </ul> |

| 配置                             | 说明  |
|--------------------------------|---|
| 目标已存在表的处理模式                    | <ul style="list-style-type: none"> <li> <b>预检查并报错拦截</b>：检查目标数据库中是否有同名的索引。如果目标数据库中没有同名的索引，则通过该检查项目；如果目标数据库中有同名的索引，则在预检查阶段提示错误，数据同步作业不会被启动。                     <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p> <b>说明</b> 如果目标库中同名的索引不方便删除或重命名，您可以<b>设置同步对象在目标实例中的名称</b>来避免表名冲突。</p> </div> </li> <li> <b>忽略报错并继续执行</b>：跳过目标数据库中是否有同名索引的检查项。                     <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p> <b>警告</b> 选择为忽略报错并继续执行，可能导致数据不一致，给业务带来风险，例如：</p> <ul style="list-style-type: none"> <li>mapping结构一致的情况下，如果在目标库遇到与源库主键的值相同的记录，在初始化阶段会保留目标库中的该条记录；在增量同步阶段则会覆盖目标库的该条记录。</li> <li>mapping结构不一致的情况下，可能会导致无法初始化数据、只能同步部分列的数据或同步失败。</li> </ul> </div> </li> </ul> |
| 选择同步对象                         | <p>在<b>源库对象框</b>中单击待同步的对象，然后单击  将其移动至<b>已选择对象框</b>。</p> <p>同步对象的选择粒度为库、表。</p>  |
| 映射名称更改                         | <p>如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见<b>库表列映射</b>。</p>  |
| 源表DMS_ONLINE_DDL过程中是否复制临时表到目标库 | <p>如源库使用<b>数据管理DMS (Data Management Service)</b> 执行Online DDL变更，您可以选择是否同步Online DDL变更产生的临时表数据。</p> <ul style="list-style-type: none"> <li> <b>是</b>：同步Online DDL变更产生的临时表数据。                             <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p> <b>说明</b> Online DDL变更产生的临时表数据过大，可能会导致同步任务延迟。</p> </div> </li> <li> <b>否</b>：不同步Online DDL变更产生的临时表数据，只同步源库的原始DDL数据。                             <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p> <b>说明</b> 该方案会导致目标库锁表。</p> </div> </li> </ul>   |
| 源、目标库无法连接重试时间                  | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p> <b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p> </div>  |

- 在**已选择对象**区域框中，将鼠标指针放置在待同步的表上，并单击表名后出现的**编辑**，设置该表在目标Elasticsearch实例中的索引名称、Type名称等信息。

**编辑表** ✕

**注意：** 编辑表名或列名后，目标数据库的表名列名将为修改后的名称。

\* 索引名称:  ⓘ

\* Type名称:  ⓘ

过滤条件: 

支持SQL标准的where条件，只有满足where条件的数据才会迁移到目标库。  
示例: id>10

设置\_routing:  是  否 ⓘ

\_id取值:

| <input checked="" type="checkbox"/> 全选 | 列名                                   | 类型            | 字段参数    | 字段参数值   |      |
|--|--------------------------------------|---------------|---------|---------|------|
| <input checked="" type="checkbox"/>    | <input type="text" value="Name"/>    | varchar(32... | index ▾ | false ▾ | 添加参数 |
| <input checked="" type="checkbox"/>    | <input type="text" value="address"/> | varchar(32... | index ▾ | false ▾ | 添加参数 |
| <input checked="" type="checkbox"/>    | <input type="text" value="id"/>      | int(11)       | index ▾ | false ▾ |      |

**确定**

| 配置         | 说明  |
|------------|---|
| 索引名称       | 详情请参见 <a href="#">Elasticsearch基本概念</a> 。   |
| Type名称     | <div style="background-color: #fff9c4; padding: 10px; border: 1px solid #ccc;"> <p><b>警告</b></p> <ul style="list-style-type: none"> <li>◦ 索引名称和Type名称支持的特殊字符仅为下划线 ( _ )。</li> <li>◦ 如需将多个相同结构的源表同步到一个目标对象，您需要重复本步骤为这些表设置相同的索引名称和Type名称。</li> </ul> </div>   |
| 过滤条件       | 您可以设置SQL过滤条件，过滤待同步的数据，只有满足过滤条件的数据才会被同步到目标实例，详情请参见 <a href="#">通过SQL条件过滤任务数据</a> 。   |
| 是否分区       | 选择是否设置分区，如果您选择为是，您还需要设置分区列和分区数量。  |
| 设置_routing | <p>设置_routing可以将文档路由存储在目标Elasticsearch实例的指定分片上，详情请参见<a href="#">_routing</a>。</p> <ul style="list-style-type: none"> <li>◦ 选择为是，您可以自定义列进行路由。</li> <li>◦ 选择为否，则用_id进行路由。</li> </ul> <div style="background-color: #e1f5fe; padding: 10px; border: 1px solid #ccc; margin-top: 10px;"> <p><b>说明</b> 创建的目标Elasticsearch实例为7.4版本时，您必须选择为否。</p> </div> |

| 配置    | 说明   |
|-------|--|
| _id取值 | <ul style="list-style-type: none"> <li>表的主键列<br/>联合主键合并为一列。</li> <li>业务主键<br/>如果选择为业务主键，那么您还需要设置对应的业务主键列。</li> </ul>   |
| 添加参数  | 选择所需的字段参数和字段参数值，字段参数及取值介绍请参见 <a href="#">Elasticsearch官方文档</a> 。<br><span style="background-color: #e0f2f1; padding: 5px; border: 1px solid #ccc;"> <span style="font-size: 1.2em; color: #0070c0;">?</span> 说明 DTS目前仅支持可以选中的参数。                 </span> |

10. 上述配置完成后，单击页面右下角的**预检查并启动**。

? 说明

- 在同步作业正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步作业。
- 如果预检查失败，单击具体检查项后的!，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择**确认屏蔽**、**忽略告警项**并重新进行预检查，跳过告警检测项重新进行预检查。

11. 在**预检查**对话框中显示**预检查通过**后，关闭**预检查**对话框，同步作业将正式开始。

12. 等待同步作业的链路初始化完成，直至处于**同步中**状态。

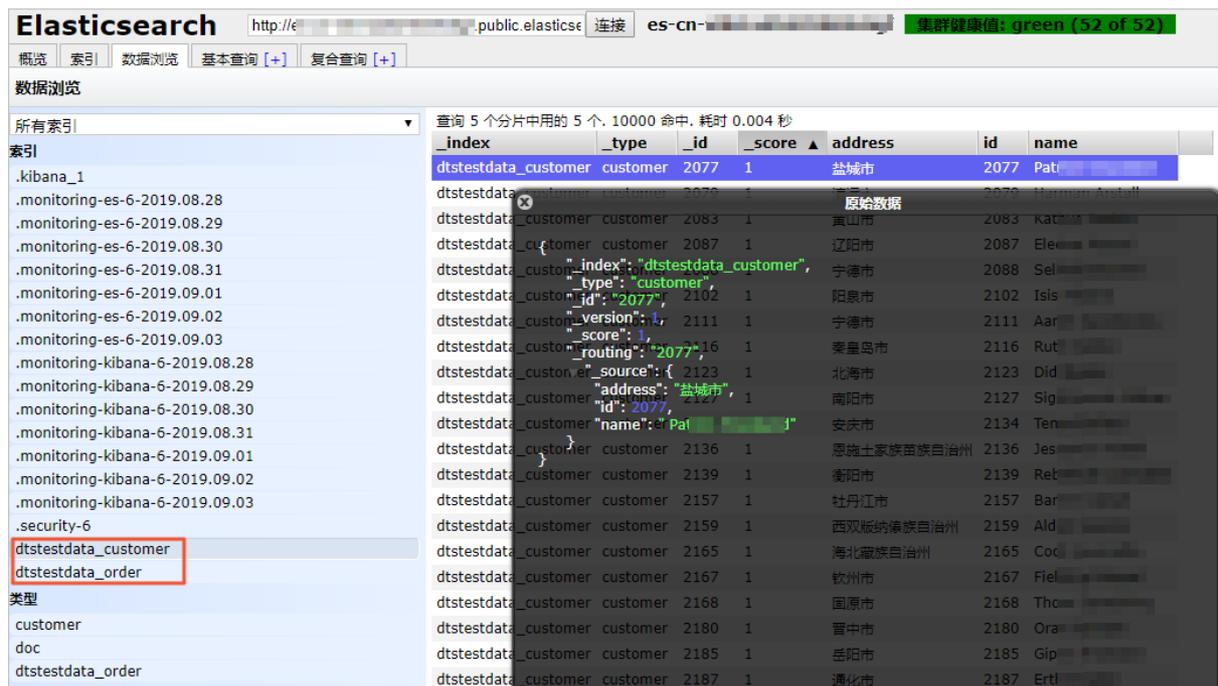
您可以在**数据同步**页面，查看数据同步作业的状态。



### 查看同步后的索引和数据

数据同步作业处于**同步中**状态后，您可以连接Elasticsearch实例（本案例使用**Head插件**进行连接），确认创建的索引和同步的数据是否符合业务的预期。

? 说明 如果不符合业务预期，您可以删除该索引及对应的数据，然后重新配置数据同步作业。



## 8.6. 从PolarDB MySQL同步到Kafka

Kafka是应用较为广泛的分布式、高吞吐量、高可扩展性消息队列服务，普遍用于日志收集、监控数据聚合、流式数据处理、在线和离线分析等大数据领域，是大数据生态中不可或缺的产品之一。通过数据传输服务DTS（Data Transmission Service），您可以将PolarDB MySQL同步至自建Kafka集群，扩展消息处理能力。

### 前提条件

- Kafka集群的版本为0.10.1.0-2.7.0版本。
- PolarDB MySQL已开启Binlog，详情请参见[如何开启Binlog](#)。

### 注意事项

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。

### 费用说明

|               |                                 |
|---------------|---------------------------------|
| 同步类型          | 链路配置费用                          |
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

### 功能限制

- 仅支持表粒度的数据同步。
- 不支持自动调整同步对象。

**说明** 如果在同步的过程中，对源库中待同步的表执行了重命名操作，且重命名后的名称不在同步对象中，那么该表将不再被同步到目标Kafka集群中。如果该表还需要同步，那么您需要[新增同步对象](#)。

### 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

**说明** 购买时，选择源实例为PolarDB、目标实例为Kafka，并选择同步拓扑为单向同步。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击[数据同步](#)。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击[配置同步链路](#)。
6. 配置源实例及目标实例信息。

同步作业名称:

---

**源实例信息**

实例类型:

实例地区:

\* PolarDB实例ID:

\* 数据库账号:

\* 数据库密码:

---

**目标实例信息**

实例类型:

实例地区:

\* ECS实例ID:

数据库类型:

\* 端口:

数据库账号:  非必填项

数据库密码:  非必填项

\* Topic:

请先点击右侧按钮，获取Topic列表后选择具体的Topic

\* Kafka版本:

\* 连接方式:  非加密连接  SCRAM-SHA-256

---

| 类别    | 配置          | 说明  |
|-------|-------------|---|
| 无     | 同步作业名称      | DTS会自动生成一个同步作业名称，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。  |
| 源实例信息 | 实例类型        | 固定为PolarDB实例，不可变更。  |
|       | 实例地区        | 购买数据同步实例时选择的源实例地域，不可变更。   |
|       | PolarDB实例ID | 选择PolarDB集群ID。  |
|       | 数据库账号       | 填入PolarDB集群的数据库账号，需要具备待同步数据库的读权限。   |
|       | 数据库密码       | 填入该数据库账号的密码。  |
|       | 实例类型        | 根据Kafka集群的部署位置选择，本文以ECS上的自建数据库为例介绍配置流程。<br><div style="border: 1px solid #add8e6; padding: 5px; margin-top: 5px;"> <span style="font-size: 1.2em;">?</span> <b>说明</b> 当选择为其他实例类型时，您还需要执行相应的准备工作，详情请参见<a href="#">准备工作概览</a>。                 </div> |
|       | 实例地区        | 购买数据同步实例时选择的目标实例地域，不可变更。  |
|       | ECS实例ID     | 选择部署了Kafka集群的ECS实例ID。   |

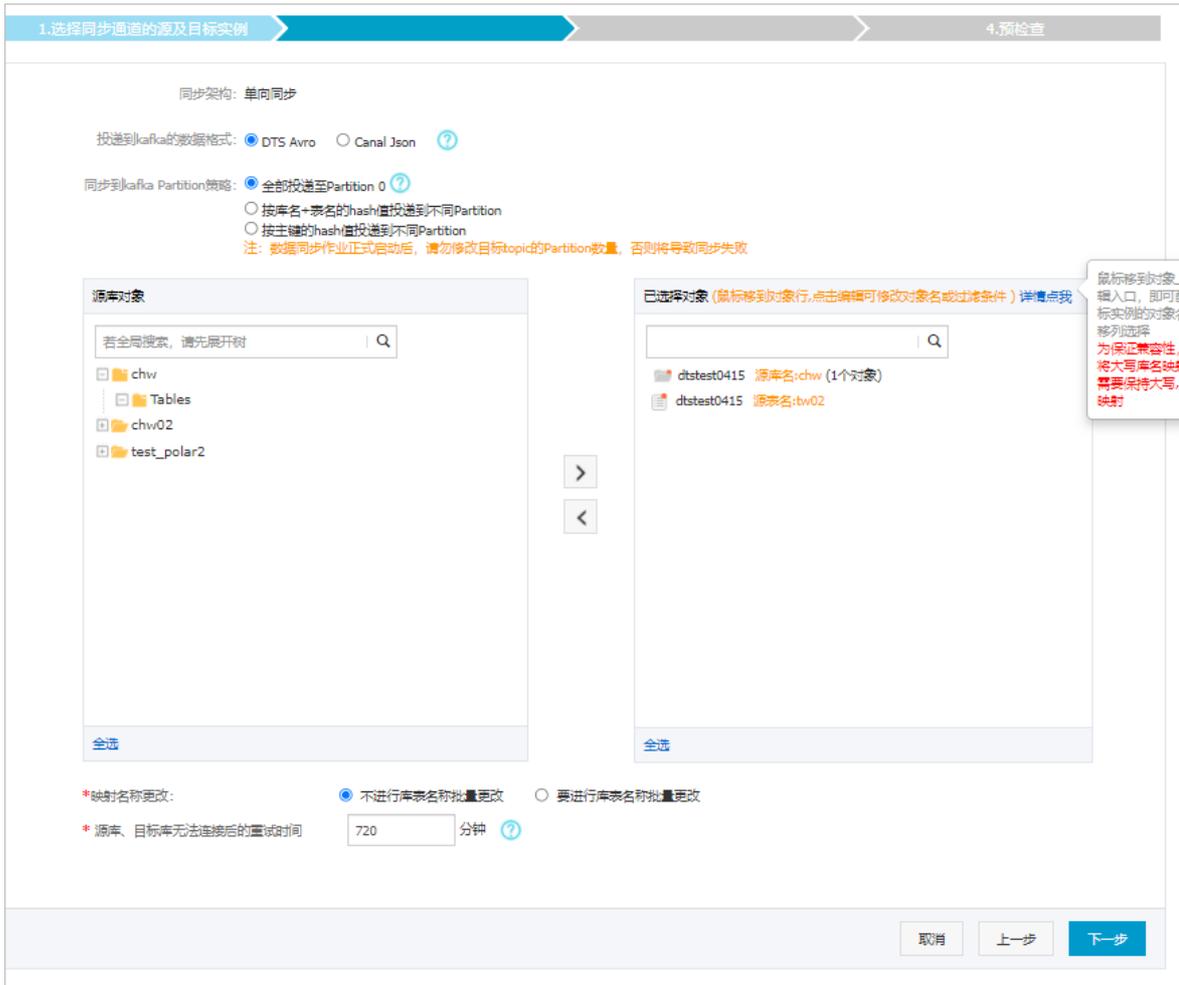
| 类别     | 配置      | 说明                                   |
|--------|---------|--------------------------------------|
| 目标实例信息 | 数据库类型   | 选择为Kafka。                            |
|        | 端口      | Kafka集群对外提供服务的端口，默认为9092。            |
|        | 数据库账号   | 填入Kafka集群的用户名，如Kafka集群未开启验证可不填写。     |
|        | 数据库密码   | 填入Kafka集群用户名对应的密码，如Kafka集群未开启验证可不填写。 |
|        | Topic   | 单击右侧的获取Topic列表，然后在下拉框中选择具体的Topic。    |
|        | Kafka版本 | 根据目标Kafka集群版本，选择对应的版本信息。             |
|        | 连接方式    | 根据业务及安全需求，选择非加密连接或SCRAM-SHA-256。     |

7. 单击页面右下角的授权白名单并进入下一步。

 说明

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置同步对象信息。



| 配置                   | 说明   |
|----------------------|--|
| 投递到kafka的数据格式        | 同步到Kafka集群中的数据以avro格式或者Canal Json格式存储，定义详情请参见 <a href="#">Kafka集群的数据存储格式</a> 。   |
| 同步到Kafka Partition策略 | 根据业务需求选择同步的策略，详细介绍请参见 <a href="#">Kafka Partition同步策略说明</a> 。  |
| 同步对象                 | <p>在源库对象区域框中，选择需要同步的对象（选择的粒度为表），然后单击  图标将其移动到已选对象区域框中。</p> <p> 说明 DTS会自动将表名映射为步骤6选择的Topic名称。如果需要更换同步的目标Topic，您需要使用库表列名映射功能，详情请参见<a href="#">设置同步对象在目标实例中的名称</a>。</p> |
| 映射名称更改               | 如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见 <a href="#">库表列映射</a> 。   |
| 源、目标库无法连接重试时间        | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <p> 说明 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p>   |

- 上述配置完成后单击页面右下角的下一步。
- 配置同步初始化的高级配置信息。

1.选择同步通道的源及目标实例
4.预检查

同步初始化:  结构初始化  全量数据初始化 注: 不支持trigger的同步, 详情请参考文档

过滤选项:  忽略增量同步阶段的 DDL

取消 上一步 保存 预检查并启动

| 配置    | 说明   |
|-------|--|
| 同步初始化 | 默认选择结构初始化和全量数据初始化, DTS会在增量数据同步之前, 将源库中待同步对象的结构和存量数据, 同步到目标库。 |
| 过滤选项  | 默认选择忽略增量同步阶段的 DDL, 即增量同步阶段源库执行的DDL操作不会被DTS同步至目标库。            |

- 上述配置完成后, 单击页面右下角的预检查并启动。

**说明**

- 在同步作业正式启动之前, 会先进行预检查。只有预检查通过后, 才能成功启动同步作业。
- 如果预检查失败, 单击具体检查项后的 , 查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项, 您也可以选择确认屏蔽、忽略告警项并重新进行预检查, 跳过告警检测项重新进行预检查。

- 在预检查对话框中显示预检查通过后, 关闭预检查对话框, 数据同步作业正式开始。您可以在数据同步页面, 查看数据同步状态。

同步作业名称

搜索
排序: 默认排序
状态: 全部

| 实例ID/作业名称  | 状态  | 同步概况                             | 付费方式 | 同步架构(全部) | 操作   |
|--|-----|----------------------------------|------|----------|--|
| <input type="checkbox"/> hangzhou-hangzhou-small                     | 同步中 | 延时: 565 毫秒<br>速度: 0TPS(0.00MB/s) | 按量付费 | 单向同步     | <a href="#">暂停同步</a>   <a href="#">转包年月</a>   <a href="#">升级更多</a> |
| <input type="checkbox"/> <a href="#">暂停同步</a>   <a href="#">释放同步</a> |     | 共有1条, 每页显示: 20条                  |      |          |  |

## 8.7. 从PolarDB MySQL同步至云原生数据仓库 AnalyticDB MySQL

云原生数据仓库AnalyticDB MySQL是阿里巴巴自主研发的海量数据实时高并发在线分析（Realt ime OLAP）云计算服务, 可以对千亿级数据进行毫秒级的即时多维分析透视和业务探索。通过数据传输服务DTS（Data Transmission Service）, 您可以将PolarDB MySQL同步到云原生数据仓库AnalyticDB MySQL, 帮助您快速构建企业内部BI、交互查询、实时报表等系统。

### 前提条件

- 已创建目标云原生数据仓库AnalyticDB MySQL集群, 详情请参见[创建云原生数据仓库AnalyticDB MySQL集群](#)。
- 确保目标云原生数据仓库AnalyticDB MySQL具备充足的存储空间。
- PolarDB MySQL已开启Binlog, 详情请参见[如何开启Binlog](#)。

### 注意事项

- DTS在执行全量数据初始化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。因此您需要在执行数据同步前评估源库和目标库的性能，同时建议您在业务低峰期执行数据同步（例如源库和目标库的CPU负载在30%以下）。

- 请勿在数据同步时，对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。
- 由于云原生数据仓库AnalyticDB MySQL本身的使用限制，当云原生数据仓库AnalyticDB MySQL集群中的节点磁盘空间使用量超过80%，该集群将被锁定。请提前根据待同步的对象预估所需空间，确保目标集群具备充足的存储空间。
- 暂不支持同步前缀索引，如果源库存在前缀索引可能导致数据同步失败。

### 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

### 支持同步的SQL操作

- DDL操作：CREATE TABLE、DROP TABLE、RENAME TABLE、TRUNCATE TABLE、ADD COLUMN、DROP COLUMN、MODIFY COLUMN
- DML操作：INSERT、UPDATE、DELETE

 **说明** 如果在数据同步的过程中变更了源表的字段类型，同步作业将报错并中断。您可以[提交工单](#)处理或手动修复，详情请参见[修复因变更字段类型导致的同步失败](#)。

### 数据库账号的权限要求

| 数据库                     | 所需权限       |
|-------------------------|------------|
| PolarDB MySQL           | 待同步对象的读权限。 |
| 云原生数据仓库AnalyticDB MySQL | 读写权限。      |

数据库账号的创建和授权方法，请参见[创建PolarDB MySQL数据库账号](#)或[创建云原生数据仓库AnalyticDB MySQL数据库账号](#)。

### 数据类型映射关系

详情请参见[结构初始化涉及的数据类型映射关系](#)。

### 操作步骤

1. [购买数据同步实例](#)。

 **说明** 购买时，选择源实例为PolarDB、目标实例为AnalyticDB MySQL，并选择同步拓扑为单向同步。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击[数据同步](#)。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击[配置同步链路](#)。
6. 配置同步通道的源实例及目标实例信息。

1.选择同步通道的源及目标实例
2.ADS账号授权
3.选择同步对象
4.预检查

同步作业名称: POLARDB\_TO\_AnalyticDB for MySQL

---

**源实例信息**

实例类型: PolarDB实例

实例地区: 华东1 (杭州)

\* PolarDB实例ID: pc-bp-xxxxxx

\* 数据库账号: dtstest

\* 数据库密码: \*\*\*\*\*

---

**目标实例信息**

实例类型: ADS

实例地区: 华东1 (杭州)

\* 版本:  2.0  3.0

\* 数据库: am-bp-xxxxxx

\* 数据库账号: dtstest

\* 数据库密码: \*\*\*\*\*

取消 授权白名单并进入下一步

| 类别     | 配置          | 说明   |
|--------|-------------|--|
| 无      | 同步作业名称      | DTS会自动生成一个同步作业名称，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。     |
| 源实例信息  | 实例类型        | 固定为PolarDB实例，不可变更。                                 |
|        | 实例地区        | 购买数据同步实例时选择的源实例地域信息，不可变更。                          |
|        | PolarDB实例ID | 选择源PolarDB集群ID。                                    |
|        | 数据库账号       | 填入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       | 填入数据库账号对应的密码。                                      |
| 目标实例信息 | 实例类型        | 固定为ADS，不可变更。                                       |
|        | 实例地区        | 购买数据同步实例时选择的目标实例地域信息，不可变更。                         |
|        | 版本          | 选择为3.0。  |
|        | 数据库         | 选择目标云原生数据仓库AnalyticDB MySQL的集群ID。                  |
|        | 数据库账号       | 填入云原生数据仓库AnalyticDB MySQL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       | 填入数据库账号对应的密码。                                      |

7. 单击页面右下角的授权白名单并进入下一步。

说明

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见DTS服务器的IP地址段。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置同步策略及对象信息。

The screenshot shows the '3. 选择同步对象' (Select Synchronization Objects) step in the DTS configuration wizard. It includes options for synchronization initialization (structure and full data), target table processing modes (pre-check and report error, or ignore error and continue), and synchronization operation types (Insert, Update, Delete, Alter Table, Truncate Table, Create Table, Drop Table). It also features a tree view for selecting source and target objects, with 'customer' and 'order' tables selected in the target view. Additional settings include mapping name changes, source table DDL replication, and a retry interval of 720 minutes.

| 配置    | 说明   |
|-------|--|
| 同步初始化 | 默认情况下，您需要同时选中结构初始化和全量数据初始化。预检查完成后，DTS会将源实例中待同步对象的结构及数据在目标集群中初始化，作为后续增量同步数据的基线数据。 |

| 配置          | 说明   |
|-------------|--|
| 目标已存在表的处理模式 | <ul style="list-style-type: none"> <li>○ <b>预检查并报错拦截</b>：检查目标数据库中是否有同名的表。如果目标数据库中没有同名的表，则通过该检查项目；如果目标数据库中有同名的表，则在预检查阶段提示错误，数据同步作业不会被启动。</li> </ul> <p><b>说明</b> 如果目标库中同名的表不方便删除或重命名，您可以更改该表在目标库中的名称，详情请参见<a href="#">设置同步对象在目标实例中的名称</a>。</p> <ul style="list-style-type: none"> <li>○ <b>忽略报错并继续执行</b>：跳过目标数据库中是否有同名表的检查项。</li> </ul> <p><b>警告</b> 选择为忽略报错并继续执行，可能导致数据不一致，给业务带来风险，例如：</p> <ul style="list-style-type: none"> <li>■ 表结构一致的情况下，在目标库遇到与源库主键的值相同的记录，则会保留目标集群中的该条记录，即源库中的该条记录不会同步至目标数据库中。</li> <li>■ 表结构不一致的情况下，可能会导致无法初始化数据、只能同步部分列的数据或同步失败。</li> </ul> |
| 多表归并        | <ul style="list-style-type: none"> <li>○ 选择为是：DTS将在每个表中增加 <code>__dts_data_source</code> 列来存储数据来源，且不再支持DDL同步。</li> <li>○ 选择为否：默认选项，支持DDL同步。</li> </ul> <p><b>说明</b> 多表归并功能基于任务级别，即不支持基于表级别执行多表归并。如果需要让部分表执行多表归并，另一部分不执行多表归并，您可以创建两个数据同步作业。</p>   |
| 同步操作类型      | <p>根据业务选中需要同步的操作类型，支持的同步操作详情请参见<a href="#">支持同步的SQL操作</a>，默认情况下都处于选中状态。</p>  |
| 选择同步对象      | <p>在源库对象框中单击待同步的对象，然后单击  图标将其移动至已选择对象框。</p> <p>同步对象的选择粒度为库、表。</p> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>○ 如果选择整个库作为同步对象，那么该库中所有对象的结构变更操作会同步至目标库。</li> <li>○ 如果选择某个表作为同步对象，那么只有这个表的ADD COLUMN操作会同步至目标库。</li> <li>○ 默认情况下，同步对象的名称保持不变。如果您需要同步对象在目标集群上名称不同，请使用对象名映射功能，详情请参见<a href="#">设置同步对象在目标实例中的名称</a>。</li> </ul>   |
| 映射名称更改      | <p>如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见<a href="#">库表列映射</a>。</p>  |

| 配置                             | 说明   |
|--------------------------------|--|
| 源表DMS_ONLINE_DDL过程中是否复制临时表到目标库 | <p>如源库使用数据管理DMS (Data Management Service) 执行Online DDL变更, 您可以选择是否同步Online DDL变更产生的临时表数据。</p> <ul style="list-style-type: none"> <li>是: 同步Online DDL变更产生的临时表数据。</li> </ul> <p><b>说明</b> Online DDL变更产生的临时表数据过大, 可能会导致同步任务延迟。</p> <ul style="list-style-type: none"> <li>否: 不同步Online DDL变更产生的临时表数据, 只同步源库的原始DDL数据。</li> </ul> <p><b>说明</b> 该方案会导致目标库锁表。</p> |
| 源、目标库无法连接重试时间                  | <p>当源、目标库无法连接时, DTS默认重试720分钟(即12小时), 您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库, 同步任务将自动恢复。否则, 同步任务将失败。</p> <p><b>说明</b> 由于连接重试期间, DTS将收取任务运行费用, 建议您根据业务需要自定义重试时间, 或者在源和目标库实例释放后尽快释放DTS实例。</p>   |

9. 上述配置完成后, 单击页面右下角的下一步。

10. 设置待同步的表在目标库中类型。

11. 上述配置完成后, 单击页面右下角的预检查并启动。

**说明**

- 在同步作业正式启动之前, 会先进行预检查。只有预检查通过后, 才能成功启动同步作业。
- 如果预检查失败, 单击具体检查项后的 , 查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项, 您也可以选择**确认屏蔽**、**忽略告警项**并重新进行预检查, 跳过告警检测项重新进行预检查。

12. 在预检查对话框中显示预检查通过后, 关闭预检查对话框, 同步作业将正式开始。

13. 等待同步作业的链路初始化完成, 直至处于同步中状态。

您可以在数据同步页面, 查看数据同步作业的状态。

## 8.8. 从PolarDB MySQL同步至云原生数据仓库AnalyticDB PostgreSQL

云原生数据仓库AnalyticDB PostgreSQL版（原分析型数据库PostgreSQL版）为您提供简单、快速、经济高效的PB级云端数据仓库解决方案。通过数据传输服务DTS（Data Transmission Service），您可以将PolarDB MySQL同步至云原生数据仓库AnalyticDB PostgreSQL，帮助您快速实现对海量数据的即席查询分析、ETL处理和可视化探索。

### 前提条件

- PolarDB MySQL集群已开启Binlog，详情请参见[如何开启Binlog](#)。
- PolarDB MySQL集群中待同步的数据表必须具备主键。
- 已创建目标云原生数据仓库AnalyticDB PostgreSQL实例，详情请参见[创建云原生数据仓库AnalyticDB PostgreSQL实例](#)。

### 注意事项

- DTS在执行全量数据初始化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。因此您需要在执行数据同步前评估源库和目标库的性能，同时建议您在业务低峰期执行数据同步（例如源库和目标库的CPU负载在30%以下）。
- 全量初始化过程中，并发INSERT会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集群的表空间大。

### 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

### 同步限制

- 同步对象仅支持数据表。
- 不支持BIT、VARBIT、GEOMETRY、ARRAY、UUID、TSQUERY、TSVECTOR、TXID\_SNAPSHOT类型的数据同步。
- 暂不支持同步前缀索引，如果源库存在前缀索引可能导致数据同步失败。
- 在数据同步时，请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。

### 支持同步的SQL操作

- DML操作：INSERT、UPDATE、DELETE。
- DDL操作：ADD COLUMN。

 **说明** 不支持CREATE TABLE操作，如果您需要将新增的表作为同步对象，则需要执行[新增同步对象](#)操作。

### 支持的同步架构

- 1对1单向同步。
- 1对多单向同步。
- 多对1单向同步。

### 术语对应关系

|               |                              |
|---------------|------------------------------|
| PolarDB MySQL | 云原生数据仓库AnalyticDB PostgreSQL |
| Database      | Schema                       |
| Table         | Table                        |

### 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

? **说明** 购买时，选择源实例为PolarDB、目标实例为AnalyticDB PostgreSQL，并选择同步拓扑为单向同步。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击[数据同步](#)。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击[配置同步链路](#)。
6. 配置同步通道的源实例及目标实例信息。

1. 选择同步通道的源及目标实例
2. 选择同步对象
3. 预检查

同步作业名称:

**源实例信息**

实例类型:

实例地区:

\* PolarDB实例ID:

\* 数据库账号:

\* 数据库密码:

**目标实例信息**

实例类型:

实例地区:

\* 实例ID:

\* 数据库名称:

\* 数据库账号:

\* 数据库密码:

| 类别 | 配置          | 说明   |
|----|-------------|--|
| 无  | 同步作业名称      | DTS会自动生成一个同步作业名称，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。 |
|    | 实例类型        | 固定为PolarDB实例。                                  |
|    | 实例地区        | 购买数据同步实例时选择的源PolarDB集群的地域信息，不可变更。              |
|    | PolarDB实例ID | 选择PolarDB集群ID。                                 |

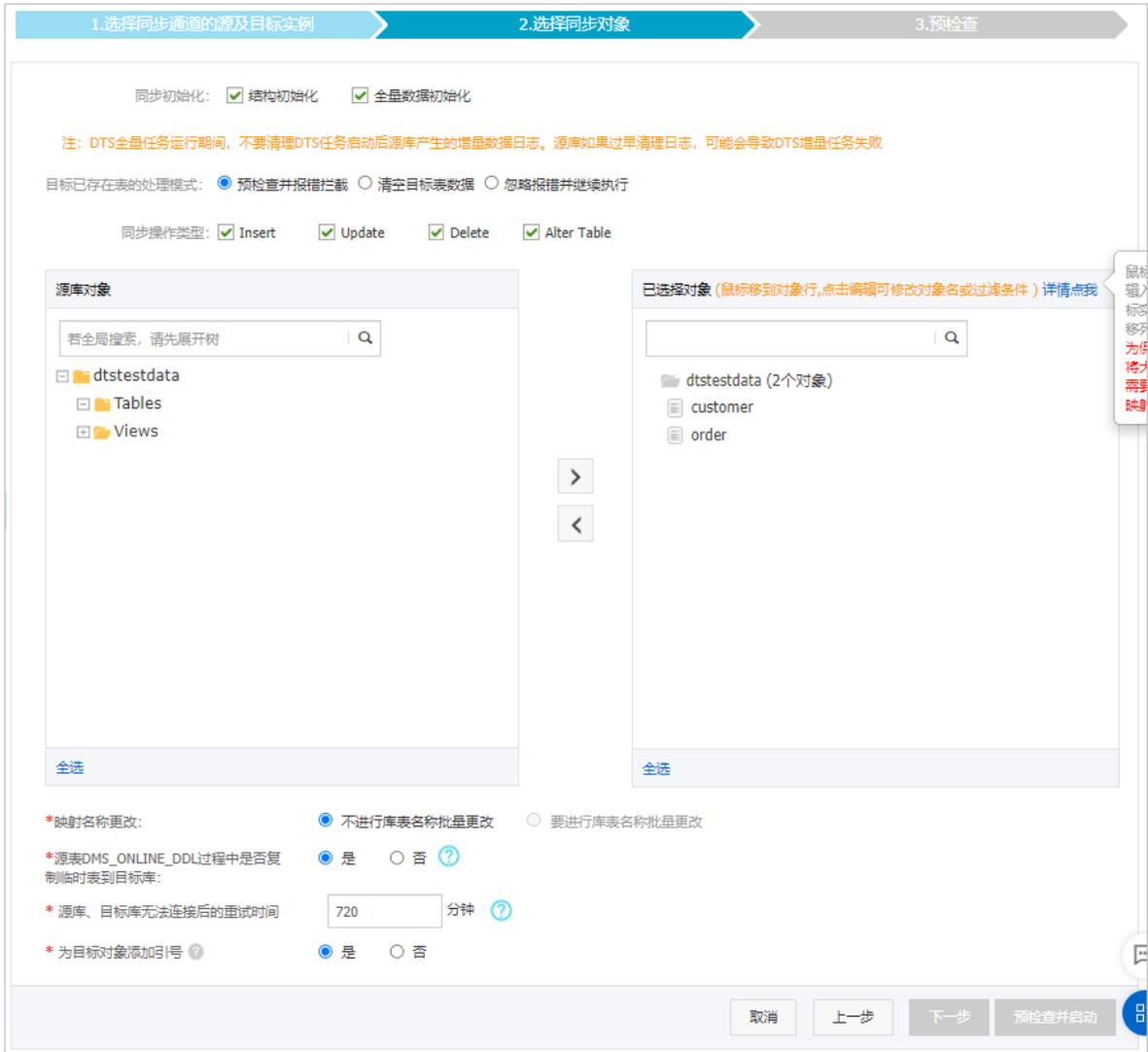
| 类别     | 配置          | 说明  |
|--------|-------------|---|
| 源实例信息  | 数据库账号       | 填入PolarDB集群的数据库账号。<br><br><b>? 说明</b> 该账号需具备待同步对象的读权限。  |
|        | 数据库密码       | 填入该数据库账号的密码。  |
| 目标实例信息 | 实例类型        | 固定为AnalyticDB for PostgreSQL，无需设置。  |
|        | 实例地区        | 购买数据同步实例时选择的目标实例地域信息，不可变更。  |
|        | 实例ID        | 选择云原生数据仓库AnalyticDB PostgreSQL实例ID。   |
|        | 数据库名称       | 填入云原生数据仓库AnalyticDB PostgreSQL实例中，待同步的目标表所属的数据库名称。  |
|        | 数据库账号       | 填入云原生数据仓库AnalyticDB PostgreSQL的 <b>初始账号</b> ，详情请参见 <b>创建数据库账号</b> 。<br><br><b>? 说明</b> 您也可以填入具备RDS_SUPERUSER权限的账号，创建方法请参见 <b>用户权限管理</b> 。 |
| 数据库密码  | 填入数据库账号的密码。 |   |

#### 7. 单击页面右下角的授权白名单并进入下一步。

##### **? 说明**

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见**DTS服务器的IP地址段**。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

#### 8. 配置同步策略及同步对象。



| 类别     | 配置          | 说明  |
|--------|-------------|---|
| 同步策略配置 | 同步初始化       | 默认情况下，您需要同时选中 <b>结构初始化</b> 和 <b>全量数据初始化</b> 。预检查完成后，DTS会将源实例中待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。  |
|        | 目标已存在表的处理模式 | <ul style="list-style-type: none"> <li><b>清空目标表的数据</b><br/>在预检查阶段跳过同名对象存在性检查的检查项目。全量初始化之前将目标表的数据清空。适用于完成同步任务测试后的正式同步场景。</li> <li><b>忽略报错并继续执行</b><br/>在预检查阶段跳过同名对象存在性检查的检查项目。全量初始化时直接追加数据。适用于多张表同步到一张表的汇总同步场景。</li> </ul> |
|        | 同步操作类型      | 根据业务需求选择需要同步的操作类型： <ul style="list-style-type: none"> <li>Insert</li> <li>Update</li> <li>Delete</li> <li>AlterTable</li> </ul>   |

| 类别                              | 配置 | 说明  |
|---------------------------------|----|---|
| 选择同步对象                          | 无  | <p>在源库对象框中单击待同步的表，然后单击  图标将其移动至已选择对象框。</p> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>同步对象的选择粒度为表。</li> <li>如果需要目标表中的列名称与源表不同，需要使用DTS的字段映射功能，详情请参见<a href="#">设置同步对象在目标实例中的名称</a>。</li> </ul>   |
| 映射名称更改                          | 无  | <p>如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见<a href="#">库表列映射</a>。</p>   |
| 源表 DMS_ONLINE_DDL过程中是否复制临时表到目标库 | 无  | <p>如源库使用<a href="#">数据管理DMS (Data Management Service)</a> 执行Online DDL变更，您可以选择是否同步Online DDL变更产生的临时表数据。</p> <ul style="list-style-type: none"> <li><b>是</b>：同步Online DDL变更产生的临时表数据。</li> </ul> <p><b>说明</b> Online DDL变更产生的临时表数据过大，可能会导致同步任务延迟。</p> <ul style="list-style-type: none"> <li><b>否</b>：不同步Online DDL变更产生的临时表数据，只同步源库的原始DDL数据。</li> </ul> <p><b>说明</b> 该方案会导致目标库锁表。</p> |
| 源、目标库无法连接重试时间                   | 无  | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <p><b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p>   |

9. 设置待同步的表在云原生数据仓库AnalyticDB PostgreSQL中的主键列和分布列信息。

1.选择同步通道的源及目标实例
2.选择同步对象
3.预检查

| Schema    | Table    | 主键列                             | 分布列                             | 定义状态(全部) |
|-----------|----------|---------------------------------|---------------------------------|----------|
| mysqltest | customer | <input type="text" value="id"/> | <input type="text" value="id"/> | 已定义      |

共有1条， 每页显示：20条

**说明** 当您在上一步中选择了结构初始化才会出现该页面。关于主键列和分布列的详细说明，请参见[表的约束定义](#)和[表分布键定义](#)。

10. 上述配置完成后，单击页面右下角的预检查并启动。

**说明**

- 在同步作业正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步作业。
- 如果预检查失败，单击具体检查项后的 ，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择**确认屏蔽**、**忽略告警项**并重新进行预检查，跳过告警检测项重新进行预检查。

11. 在预检查对话框中显示预检查通过后，关闭预检查对话框，同步作业将正式开始。

12. 等待同步作业的链路初始化完成，直至处于同步中状态。

您可以在数据同步页面，查看数据同步作业的状态。



## 8.9. 从PolarDB MySQL同步至PolarDB-X

云原生分布式数据库PolarDB-X是阿里巴巴致力于解决单机数据库服务瓶颈而自主研发的分布式数据库产品，高度兼容MySQL协议和语法，支持自动化水平拆分、在线平滑扩缩容、弹性扩展、透明读写分离，具备数据库全生命周期运维管控能力。通过数据传输服务DTS（Data Transmission Service），您可以将PolarDB MySQL同步至PolarDB-X。

### 前提条件

- PolarDB MySQL已开启Binlog，详情请参见[如何开启Binlog](#)。
- 源库中待同步的数据表必须具备主键。
- 确保目标库具备充足的存储空间。
- 已创建分布式关系型数据库服务PolarDB-X，如未创建请参见[创建实例](#)和[创建数据库](#)。

**说明** 创建实例时需选择存储类型为RDS MySQL。

### 注意事项

- DTS在执行全量数据初始化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。因此您需要在执行数据同步前评估源库和目标库的性能，同时建议您在业务低峰期执行数据同步（例如源库和目标库的CPU负载在30%以下）。
- 全量初始化过程中，并发INSERT会导致目标集群的表碎片，全量初始化完成后，目标集群的表空间比源库的表空间大。确保目标库具备充足的存储空间。
- MySQL同步至DRDS暂不支持结构初始化，在配置同步作业前，您需要在目标实例中创建对应的库和表。

### 费用说明

|               |                                 |
|---------------|---------------------------------|
| 同步类型          | 链路配置费用                          |
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

### 同步限制

- 同步对象仅支持数据表。
- 不支持BIT、VARBIT、GEOMETRY、ARRAY、UUID、TSQUERY、TSVECTOR、TXID\_SNAPSHOT类型的数据同步。
- 暂不支持同步前缀索引，如果源库存在前缀索引可能导致数据同步失败。
- 在数据同步时，请勿对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。

### 支持同步的SQL操作

INSERT、UPDATE、DELETE。

### 数据库账号的权限要求

| 数据库           | 所需权限  |
|---------------|---|
| PolarDB MySQL | REPLICATION CLIENT、REPLICATION SLAVE、SHOW VIEW和所有同步对象的SELECT权限。 |
| PolarDB-X     | 无需填写数据库账号信息，DTS会自动创建账号并授权。                                      |

### 支持的同步架构

- 1对1单向同步。
- 多对1单向同步。

### 支持的同步架构

- 1对1单向同步。
- 多对1单向同步。

### 准备工作

由于PolarDB MySQL同步至PolarDB-X暂不支持**结构初始化**，所以您需要根据源PolarDB MySQL集群中待同步对象的数据结构，在目标实例中创建相应的数据库和数据表，详情请参见[创建数据库](#)和[创建数据表](#)。

 **说明** 结构初始化：将源库中待同步对象的结构定义信息，同步至目标库中。

### 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。
2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击**数据同步**。
4. 在**同步作业列表**页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击**配置同步链路**。
6. 配置同步作业的源集群及目标实例信息。

1.选择同步通道的源及目标实例
2.选择同步对象
3.高级设置
4.预检查

同步作业名称:

**源实例信息**

实例类型:

实例地区: 华东1 (杭州)

\* PolarDB实例ID:

\* 数据库账号:

\* 数据库密码:

**目标实例信息**

实例类型: DRDS实例

实例地区: 华东1 (杭州)

\* DRDS实例ID:

| 类别     | 配置          | 说明   |
|--------|-------------|--|
| 无      | 同步作业名称      | DTS会自动生成一个同步作业名称，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。         |
| 源实例信息  | 实例类型        | 选择PolarDB实例。   |
|        | 实例地区        | 购买数据同步实例时选择的源集群地域信息，不可变更。                              |
|        | PolarDB实例ID | 选择源PolarDB集群ID。  |
|        | 数据库账号       | 填入PolarDB集群的数据库账号，权限要求请参见 <a href="#">数据库账号的权限要求</a> 。 |
|        | 数据库密码       | 填入该数据库账号对应的密码。   |
| 目标实例信息 | 实例类型        | 固定为DRDS实例，不可变更。  |
|        | 实例地区        | 购买数据同步实例时选择的目标实例地域信息，不可变更。                             |
|        | DRDS实例ID    | 选择目标PolarDB-X实例ID。                                     |

7. 单击页面右下角的授权白名单并进入下一步。

**说明**

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

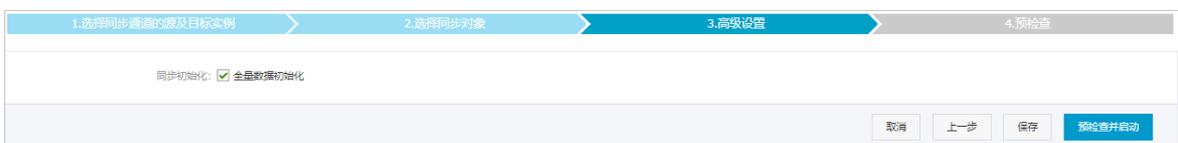
8. 配置同步策略及对象信息。



| 配置项目          | 配置说明   |
|---------------|--|
| 选择同步对象        | <p>在源库对象框中单击待同步的表，然后单击  图标将其移动至已选择对象框。</p> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>同步对象的选择粒度为表。</li> <li>默认情况下，同步对象的名称保持不变。如果您需要改变同步对象在目标实例中的名称，那么需要使用DTS提供的对象名映射功能，详情请参见<a href="#">设置同步对象在目标实例中的名称</a>。</li> </ul> |
| 映射名称更改        | 如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见 <a href="#">库表映射</a> 。  |
| 源、目标库无法连接重试时间 | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <p><b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p>  |

9. 单击下一步。

10. 选择是否要执行全量数据初始化。



**说明** 全量数据初始化：DTS将源库中待同步表的存量数据同步至目标库中，如果不选择则不同步存量数据。

11. 上述配置完成后，单击页面右下角的**预检查并启动**。

**说明**

- 在同步作业正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步作业。
- 如果预检查失败，单击具体检查项后的，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择**确认屏蔽**、**忽略告警项**并重新进行预检查，跳过告警检测项重新进行预检查。

12. 在**预检查**对话框中显示**预检查通过**后，关闭**预检查**对话框，同步作业将正式开始。

13. 等待同步作业的链路初始化完成，直至处于**同步中**状态。

您可以在**数据同步**页面，查看数据同步作业的状态。

| 实例ID/作业名称               | 状态  | 同步概况                                    | 付费方式 | 同步架构(全部) | 操作  |
|-------------------------|-----|---|------|----------|---|
| hangzhou-hangzhou-small | 同步中 | 耗时: 1376 毫秒<br>速度: 0.00RPS/(0.0000MB/s) | 按量付费 | 单向同步     | <a href="#">暂停同步</a> <a href="#">转包年包月</a> <a href="#">升级更多</a> |

共有1条，每页显示：20条

# 9.PolarDB O引擎数据同步

## 9.1. PolarDB O引擎集群间的单向同步

云原生关系型数据库PolarDB O引擎是阿里巴巴自主研发的下一代云原生关系型数据库，高度兼容Oracle语法。通过数据传输服务DTS（Data Transmission Service），您可以实现PolarDB O引擎集群间的单向数据同步。

### 前提条件

- 源PolarDB O引擎集群需为最新版本，升级方式详情请见[版本管理](#)。
- 源PolarDB O引擎集群中，待同步的表需具备主键或非空唯一索引。
- 源PolarDB O引擎集群中，wal\_level参数的值需设置为logical，即在预写式日志WAL（Write-ahead logging）中增加支持逻辑编码所需的信息。详情请参见[设置集群参数](#)。

### 注意事项

- DTS在执行全量数据初始化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。因此您需要在执行数据同步前评估源库和目标库的性能，同时建议您在业务低峰期执行数据同步（例如源库和目标库的CPU负载在30%以下）。
- 一个数据同步作业只能同步一个数据库，如有多个数据库需要同步，您需要为每个数据库配置数据同步作业。
- 为保障同步延迟时间展示的准确性，DTS会在源库中新增一个表，表名为 `dts_postgres_heartbeat`，结构及内容如下图所示。

The screenshot shows a SQL query: `select * from "dtstest"."dts_postgres_heartbeat"`. The result is a table with the following columns: SLOT\_NAME, REVICE\_TIME, REVICE\_LSN, FLUSHED\_LSN, UPDATE\_TIME, and DTS\_SERVICE\_TIME. The data row shows: w8i, 1585104490008, 0/44, null, 2020-03-25 10:48:15.037646+08, 1585104495032.

|   | SLOT_NAME | REVICE_TIME   | REVICE_LSN | FLUSHED_LSN | UPDATE_TIME                   | DTS_SERVICE_TIME |
|---|-----------|---------------|------------|-------------|-------------------------------|------------------|
| 1 | w8i       | 1585104490008 | 0/44       | null        | 2020-03-25 10:48:15.037646+08 | 1585104495032    |

- 在数据同步过程中，如果同步对象的选择粒度为Schema，在待同步的Schema中创建了新的表或使用RENAME命令重建了待同步的表，您需要在对该表写入数据前执行 `ALTER TABLE schema.table REPLICA IDENTITY FULL;` 命令。

**说明** 将上述命令中的 `schema` 和 `table` 替换成真实的Schema名和表名。

### 费用说明

|               |                                 |
|---------------|---------------------------------|
| 同步类型          | 链路配置费用                          |
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

### 支持同步的SQL操作

|      |                      |
|------|----------------------|
| 操作类型 | SQL操作语句              |
| DML  | INSERT、UPDATE、DELETE |

| 操作类型 | SQL操作语句   |
|------|---|
| DDL  | <ul style="list-style-type: none"> <li>• CREATE TABLE、DROP TABLE</li> <li>• ALTER TABLE (包括RENAME TABLE、ADD COLUMN、ADD COLUMN DEFAULT、ALTER COLUMN TYPE、DROP COLUMN、ADD CONSTRAINT、ADD CONSTRAINT CHECK、ALTER COLUMN DROP DEFAULT)</li> <li>• CREATE INDEX ON TABLE</li> </ul> <div style="background-color: #e6f2ff; padding: 10px; margin-top: 10px;"> <p><span style="color: #00aaff;">?</span> <b>说明</b> 不支持DDL同步的有以下场景：</p> <ul style="list-style-type: none"> <li>• DDL中CASCADE和RESTRICT等附加信息不会被同步。</li> <li>• 一个事务中包含DML和DDL，则该DDL不会被同步。</li> <li>• 一个事务中同时包含在同步对象的DDL和不在同步对象的DDL。</li> <li>• 通过设置SET session_replication_role = replica会话下执行的DDL不支持同步。</li> <li>• 省略schema的DDL（会使用 show search_path 默认为public）。</li> <li>• DDL语句中包含if not exist的场景。</li> </ul> </div> |

### 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

? **说明** 购买时，源实例和目标实例均选择为POLARDB，并选择同步拓扑为单向同步。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击[数据同步](#)。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击[配置同步链路](#)。
6. 配置同步通道的源实例及目标实例信息。

1.选择同步通道的源及目标实例
2.选择同步对象
3.高级设置
4.预检查

同步作业名称:

**源实例信息**

实例类型:

实例地区: 华东1 (杭州)

\* PolarDB实例ID:

\* 数据库名称:

\* 数据库账号:

\* 数据库密码:

**目标实例信息**

实例类型: PolarDB

实例地区: 华东1 (杭州)

\* PolarDB实例ID:

\* 数据库名称:

\* 数据库账号:

\* 数据库密码:

| 类别     | 配置          | 说明   |
|--------|-------------|--|
| 无      | 同步作业名称      | DTs会自动生成一个同步作业名称，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。   |
| 源实例信息  | 实例类型        | 固定为PolarDB实例，不可变更。   |
|        | 实例地区        | 购买数据同步实例时选择的源实例地域信息，不可变更。  |
|        | PolarDB实例ID | 选择源PolarDB集群ID。  |
|        | 数据库名称       | 填入待同步的数据库名称。   |
|        | 数据库账号       | 填入源PolarDB集群的高权限账号。数据库账号创建方法请参见 <a href="#">创建数据库账号</a> 。  |
|        | 数据库密码       | 填入数据库账号对应的密码。  |
| 目标实例信息 | 实例类型        | 固定为PolarDB，不可变更。   |
|        | 实例地区        | 购买数据同步实例时选择的目标实例地域信息，不可变更。   |
|        | PolarDB实例ID | 选择目标PolarDB集群ID。   |
|        | 数据库名称       | 填入同步的目标数据库名称。  |
|        | 数据库账号       | 填入目标PolarDB集群的数据库账号，该账号需具备数据库Owner权限。<br><div style="border: 1px solid #add8e6; padding: 5px; margin-top: 5px;"> <span style="font-size: 1.2em;">🔊</span> <b>注意</b> 数据库Owner在创建数据库时已指定。                 </div> |
|        | 数据库密码       | 填入数据库账号对应的密码。  |

7. 单击页面右下角的授权白名单并进入下一步。

**说明**

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置目标已存在表的处理模式和同步对象。

同步架构：单向同步

目标已存在表的处理模式： 预检查并报错拦截  忽略报错并继续执行

源库对象

若全局搜索，请先展开树

- anydata
- aq\$\_agent
- aq\$\_descriptor
- aq\$\_reg\_info
- dbms\_alert
- dbms\_application\_info
- dbms\_aq
- dbms\_aqadm
- dbms\_crypto
- dbms\_job
- dbms\_lock
- dbms\_mview
- dbms\_obfuscation\_toolkit
- dbms\_output

已选择对象 (鼠标移到对象行, 点击编辑可修改对象名或过滤条件) 详情点我

- dbms\_job

\* 映射名称更改： 不进行库表名称批量更改  要进行库表名称批量更改

\* 源库、目标库无法连接后的重试时间  分钟

取消 上一步 下一步

| 配置项目 | 配置说明 |
|------|------|
|------|------|

| 配置项目          | 配置说明   |
|---------------|--|
| 目标已存在表的处理模式   | <ul style="list-style-type: none"> <li>◦ <b>预检查并报错拦截</b>：检查目标数据库中是否有同名的表。如果目标数据库中没有同名的表，则通过该检查项目；如果目标数据库中有同名的表，则在预检查阶段提示错误，数据同步作业不会被启动。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin: 5px 0;"> <p><b>说明</b> 如果目标库中同名的表不方便删除或重命名，您可以<b>设置同步对象在目标实例中的名称</b>来避免表名冲突。</p> </div> <ul style="list-style-type: none"> <li>◦ <b>忽略报错并继续执行</b>：跳过目标数据库中是否有同名表的检查项。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin: 5px 0;"> <p><b>警告</b> 选择为<b>忽略报错并继续执行</b>，可能导致数据不一致，给业务带来风险，例如：</p> <ul style="list-style-type: none"> <li>■ 表结构一致的情况下，如果在目标库遇到与源库主键的值相同的记录，在初始化阶段会保留目标库中的该条记录；在增量同步阶段则会覆盖目标库的该条记录。</li> <li>■ 表结构不一致的情况下，可能会导致无法初始化数据、只能同步部分列的数据或同步失败。</li> </ul> </div> |
| 选择同步对象        | <p>在<b>源库对象框</b>中单击待同步的对象，然后单击  图标将其移动至<b>已选择对象框</b>。</p> <p>同步对象的选择粒度为库、表。</p> <div style="border: 1px solid #ccc; padding: 5px; margin: 5px 0;"> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>◦ 如果选择整个库作为同步对象，那么该库中所有对象的结构变更操作都会同步至目标库。</li> <li>◦ 默认情况下，同步对象的名称保持不变。如果您需要改变同步对象在目标集群中的名称，请使用对象名映射功能，详情请参见<b>设置同步对象在目标实例中的名称</b>。</li> </ul> </div>  |
| 映射名称更改        | <p>如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见<b>库表列映射</b>。</p>   |
| 源、目标库无法连接重试时间 | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <div style="border: 1px solid #ccc; padding: 5px; margin: 5px 0;"> <p><b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p> </div>  |

9. 上述配置完成后，单击页面右下角的下一步。

10. 配置同步初始化。

1.选择同步通道的源及目标实例
2.选择同步对象
3.高级设置
4.预检查

同步初始化： 结构初始化  全量数据初始化 注：不支持trigger的同步，详情请[参考文档](#)

取消 上一步 保存 预检查并启动

| 初始化类型   | 说明  |
|---------|---|
| 结构初始化   | <p>DTS将同步对象的结构定义同步到目标PolarDB集群。目前支持的对象包括：表、视图、同义词、触发器、存储过程、存储函数、包、自定义类型。</p> <p><b>注意</b> 当同步对象包含触发器时可能导致数据不一致，解决方案请参见<a href="#">源库存在触发器时如何配置同步作业</a>。</p> |
| 全量数据初始化 | <p>DTS会将源库中同步对象的存量数据，全部同步至目标PolarDB集群。</p> <p><b>注意</b> 在同步启动运行后，请勿对同步对象执行DDL操作，否则可能导致同步失败。</p>   |

11. 上述配置完成后，单击页面右下角的**预检查并启动**。

**说明**

- 在同步作业正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步作业。
- 如果预检查失败，单击具体检查项后的，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择**确认屏蔽**、**忽略告警项**并重新进行预检查，跳过告警检测项重新进行预检查。

12. 在**预检查**对话框中显示**预检查**通过后，关闭**预检查**对话框，同步作业将正式开始。

13. 等待同步作业的链路初始化完成，直至处于**同步中**状态。

您可以在**数据同步**页面，查看数据同步作业的状态。



## 9.2. 从PolarDB O引擎集群同步至自建Kafka

Kafka是应用较为广泛的分布式、高吞吐量、高可扩展性消息队列服务，普遍用于日志收集、监控数据聚合、流式数据处理、在线和离线分析等大数据领域，是大数据生态中不可或缺的产品之一。通过数据传输服务DTS（Data Transmission Service），您可以将PolarDB O引擎集群同步至自建Kafka，扩展消息处理能力。

### 前提条件

- 源PolarDB O引擎集群需为最新版本，升级方式详情请见[版本管理](#)。
- 源PolarDB O引擎集群中，待同步的表需具备主键或非空唯一索引。
- 源PolarDB O引擎集群中，wal\_level参数的值需设置为logical，即在预写式日志WAL（Write-ahead logging）中增加支持逻辑编码所需的信息。详情请参见[设置集群参数](#)。

### 注意事项

- 本场景中，DTS仅支持，不支持和。[增量数据同步结构初始化全量数据初始化](#)
- 一个数据同步作业只能同步一个数据库，如有多个数据库需要同步，您需要为每个数据库配置数据同步作业。
- 为保障同步延迟时间展示的准确性，DTS会在源库中新增一个表，表名为 `dts_postgres_heartbeat`，结构及内容如下图所示。

| 消息 执行结果1 <span style="float:right">跨数据库实例查询</span> |             |               |             |             |                               |               |
|--|-------------|---------------|-------------|-------------|-------------------------------|---------------|
| SLOT_NAME  | REVICE_TIME | REVICE_LSN    | FLUSHED_LSN | UPDATE_TIME | DTS_SERVICE_TIME              |               |
| 1  | w8i         | 1585104490008 | 0/44        | null        | 2020-03-25 10:48:15.037646+08 | 1585104495032 |

### 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

### 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

? 说明 购买时，源实例选择为POLARDB，目标实例选择为Kafka，并选择同步拓扑为单向同步。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击[数据同步](#)。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击[配置同步链路](#)。
6. 配置同步作业的源和目标实例信息。

同步作业名称: PolarDB-O\_To\_Kafka

---

**源实例信息**

实例类型: PolarDB实例

实例地区: 华东2 (上海)

\* PolarDB实例ID: pc-XXXXXXXXXX

\* 数据库名称: dtstestdata

\* 数据库账号: dtstest

\* 数据库密码: XXXXXXXXXX

---

**目标实例信息**

实例类型: ECS上的自建数据库

实例地区: 华东2 (上海)

\* ECS实例ID: i-XXXXXXXXXX

数据库类型: Kafka

\* 端口: 9092

数据库账号:  非必填项

数据库密码:  非必填项

\* Topic: polardb\_test 获取Topic列表

请先点击右侧按钮，获取Topic列表后选择具体的Topic

\* Kafka版本: 1.0

\* 连接方式:  非加密连接  SCRAM-SHA-256

---

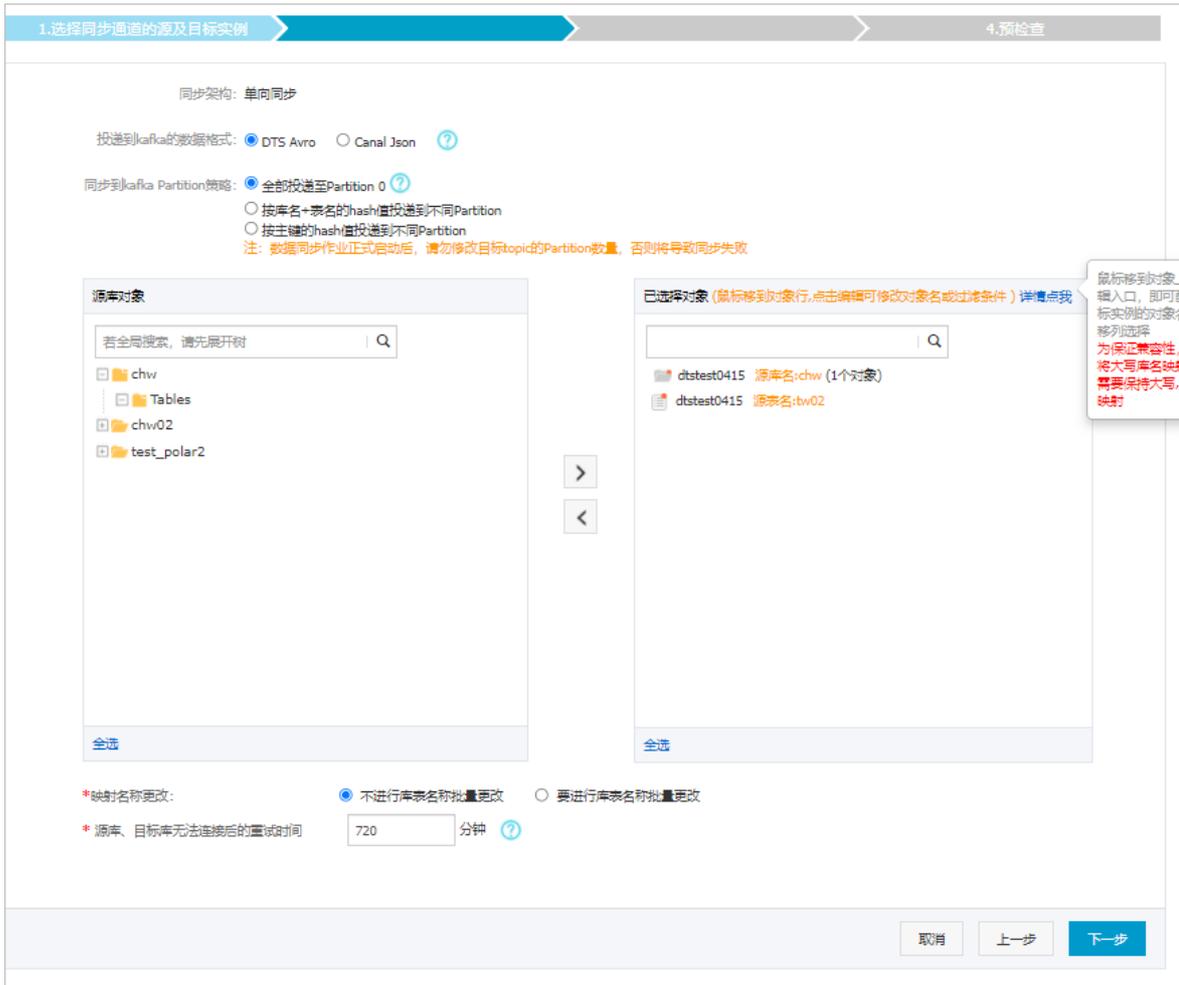
| 类别     | 配置          | 说明   |
|--------|-------------|--|
| 无      | 同步作业名称      | DTS会自动生成一个同步作业名称，建议配置具有业务意义的名称（无唯一性要求），便于后续识别。   |
| 源实例信息  | 实例类型        | 固定为PolarDB实例，不可变更。   |
|        | 实例地区        | 购买数据同步实例时选择的源实例地域信息，不可变更。  |
|        | PolarDB实例ID | 选择源PolarDB集群ID。  |
|        | 数据库名称       | 填入待同步的数据库名称。   |
|        | 数据库账号       | 填入源PolarDB集群的高权限账号。数据库账号创建方法请参见 <a href="#">创建数据库账号</a> 。  |
|        | 数据库密码       | 填入数据库账号的密码。  |
| 目标实例信息 | 实例类型        | 根据自建Kafka的部署位置进行选择，本文以ECS上的自建数据库为例介绍配置流程。<br><br><div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin-top: 5px;"> <span style="font-size: 1.2em; color: #007bff;">?</span> <b>说明</b> 当自建Kafka为其他实例类型时，您还需要执行相应的准备工作，详情请参见<a href="#">准备工作概览</a>。                 </div> |
|        | 实例地区        | 购买数据同步实例时选择的目标实例地域信息，不可变更。   |
|        | ECS实例ID     | 选择部署了Kafka集群的ECS实例ID。  |
|        | 数据库类型       | 选择为Kafka。  |
|        | 端口          | 自建Kafka对外提供服务的端口，默认为9092。  |
|        | 数据库账号       | 填入自建Kafka的用户名，如自建Kafka未开启验证可不填写。   |
|        | 数据库密码       | 填入该用户的密码，如自建Kafka未开启验证可不填写。  |
|        | Topic       | 单击右侧的获取Topic列表，然后在下拉框中选择具体的Topic。  |
|        | Kafka版本     | 根据自建Kafka的版本进行选择。  |
|        | 连接方式        | 根据业务及安全需求，选择非加密连接或SCRAM-SHA-256。   |

7. 单击页面右下角的授权白名单并进入下一步。

? **说明**

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置同步对象。



| 配置                   | 说明  |
|----------------------|---|
| 选择同步对象               | <p>在源库对象框中单击待同步的对象（仅支持表粒度的选择），然后单击 <b>&gt;</b> 将其移动至已选择对象框。</p> <p><b>说明</b> DTS会自动将表名映射为<b>步骤6</b>中选择的Topic名称。如需更换同步的目标Topic（该Topic需是Kafka实例中真实存在），您可以将鼠标指针放置在要进行名称映射的表上，并单击出现的编辑进行调整，详情请参见<b>设置同步对象在目标实例中的名称</b>。</p> |
| 投递到kafka的数据格式        | 同步到Kafka集群中的数据以avro格式或者Shareplex Json格式存储，定义详情请参见 <b>Kafka集群的数据存储格式</b> 。   |
| 同步到Kafka Partition策略 | 根据业务需求选择同步的策略，详细介绍请参见 <b>Kafka Partition同步策略说明</b> 。  |
| 映射名称更改               | 如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见 <b>库表映射</b> 。  |

| 配置            | 说明  |
|---------------|---|
| 源、目标库无法连接重试时间 | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <p><b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p> |

9. 上述配置完成后，单击页面右下角的下一步。

10. 配置同步初始化和过滤选项。

1.选择同步通道的源及目标实例
2.选择同步对象
3.高级设置
4.预检查

同步初始化:  增量数据初始化

过滤选项:  忽略增量同步阶段的 DDL

取消 上一步 保存 预检查并启动

| 配置    | 说明  |
|-------|---|
| 同步初始化 | 默认勾选 <b>增量数据初始化</b> ，即DTS会将同步过程中，源库产生的增量变更数据实时同步至目标库。   |
| 过滤选项  | <p>默认勾选<b>忽略增量同步阶段的 DDL</b>，即DTS不会将源库在增量数据同步阶段执行的DDL操作同步至目标库。</p> <p><b>说明</b> 目前还不支持过滤选项，即无论是否勾选该选项，都无法将源库在增量同步阶段执行的DDL操作同步至目标库。</p> |

11. 上述配置完成后，单击页面右下角的预检查并启动。

**说明**

- 在同步作业正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步作业。
- 如果预检查失败，单击具体检查项后的，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择**确认屏蔽、忽略告警项并重新进行预检查**，跳过告警检测项重新进行预检查。

12. 在预检查对话框中显示预检查通过后，关闭预检查对话框，同步作业将正式开始。

13. 等待同步作业的链路初始化完成，直至处于同步中状态。

您可以在**数据同步**页面，查看数据同步作业的状态。

| 同步作业名称  | 状态  | 同步概况                                   | 付费方式 | 同步架构(全部) | 操作  |
|---|-----|--|------|----------|---|
| <input type="checkbox"/> <span style="font-size: small;">实例ID/作业名称</span><br><input type="checkbox"/> hangzhou-hangzhou-small             | 同步中 | 耗时: 1376 毫秒<br>速度: 0.00RPS/(0.000MB/s) | 按量付费 | 单向同步     | <a href="#">暂停同步</a>   <a href="#">转包年包月</a>   <a href="#">升级更多</a> |
| <input type="checkbox"/> <span style="font-size: small;">暂停同步</span> <input type="checkbox"/> <span style="font-size: small;">释放同步</span> |     | 共有1条， 每页显示：20条                         |      |          |   |

# 10.PolarDB-X数据同步

## 10.1. PolarDB-X间的数据实时同步

云原生分布式数据库PolarDB-X是阿里巴巴致力于解决单机数据库服务瓶颈而自主研发的分布式数据库产品，高度兼容MySQL协议和语法，支持自动化水平拆分、在线平滑扩缩容、弹性扩展、透明读写分离，具备数据库全生命周期运维管控能力。通过数据传输服务DTS（Data Transmission Service），您可以实现PolarDB-X间的数据同步。

### 前提条件

- PolarDB-X中的数据库须基于RDS MySQL创建。
- 待同步的表须具备主键。
- 确保目标PolarDB-X中，目标数据库对应的RDS实例具备充足的存储空间。

### 注意事项

- DTS在执行全量数据初始化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。因此您需要在执行数据同步前评估源库和目标库的性能，同时建议您在业务低峰期执行数据同步（例如源库和目标库的CPU负载在30%以下）。
- 数据同步期间，请勿对PolarDB-X执行扩容、缩容、迁移热点表、变更拆分键和变更DDL等操作，否则将导致数据同步失败。
- 如果需要在数据同步期间切换PolarDB-X的网络类型，在您执行完网络类型切换操作后，请[提交工单](#)调整同步链路的网络连接信息。

### 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

### 支持同步的SQL操作

INSERT、UPDATE、DELETE。

### 准备工作

由于PolarDB-X间的数据同步不支持结构初始化，所以您需要根据源PolarDB-X中待同步对象的数据结构，在目标PolarDB-X中创建相应的数据库和数据表，详情请参见[创建数据库](#)和[创建数据表](#)。

 **说明** 结构初始化：将源库中待同步对象的结构定义信息，同步至目标库中。

### 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

 **说明** 购买时，源和目标实例均选择为PolarDB-X（原DRDS升级版），并选择同步拓扑为单向同步。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击[数据同步](#)。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击[配置同步链路](#)。
6. 配置数据同步的源实例及目标实例信息。

1.选择同步通道的源及目标实例    2.选择同步对象    3.高级设置    4.预检查

同步作业名称:

**源实例信息**

实例类型: DRDS实例  
实例地区: 华东1 (杭州)  
\* DRDS实例ID:

**目标实例信息**

实例类型: DRDS实例  
实例地区: 华东1 (杭州)  
\* DRDS实例ID:

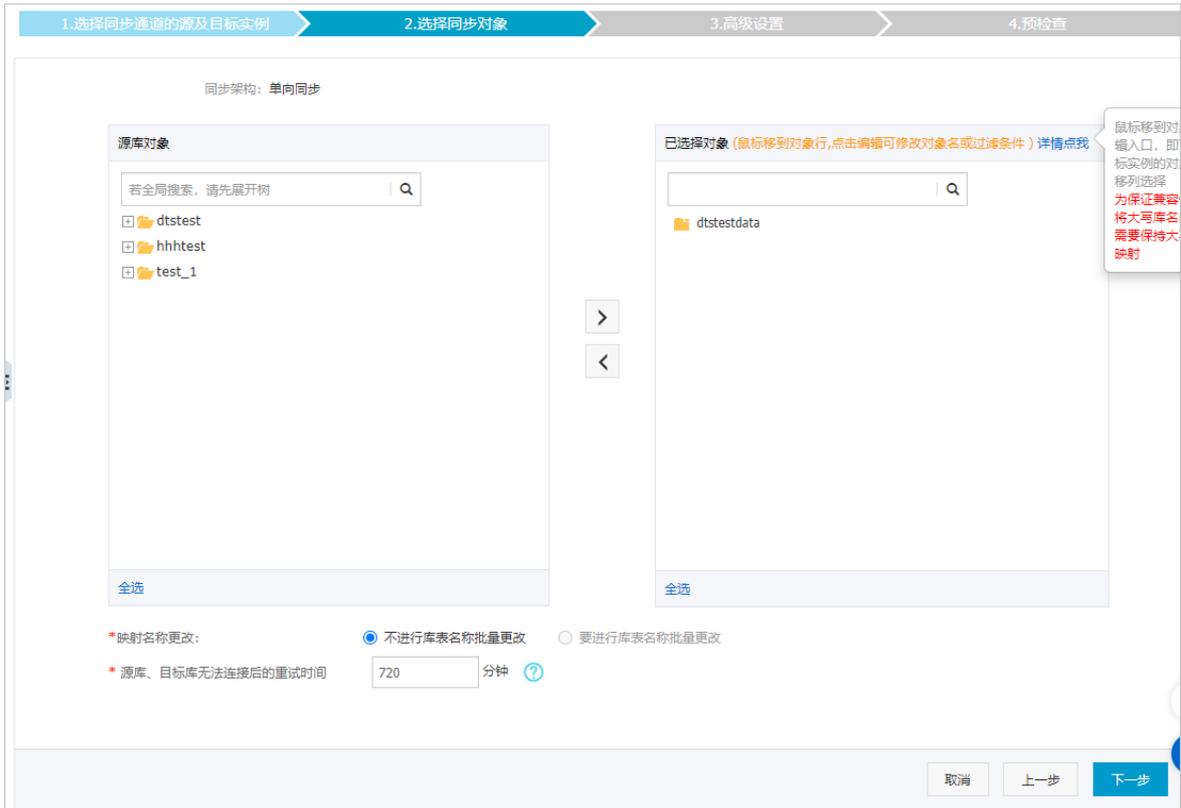
| 类别     | 配置       | 说明   |
|--------|----------|--|
| 无      | 同步作业名称   | DTS会自动生成一个同步作业名称，建议配置具有业务意义的名称（无唯一性要求），便于后续识别。 |
| 源实例信息  | 实例类型     | 固定为DRDS实例，不可变更。                                |
|        | 实例地区     | 购买数据同步实例时选择的源实例地域信息，不可变更。                      |
|        | DRDS实例ID | 选择源PolarDB-X的实例ID。                             |
| 目标实例信息 | 实例类型     | 固定为DRDS实例，不可变更。                                |
|        | 实例地区     | 购买数据同步实例时选择的目标实例地域信息，不可变更。                     |
|        | DRDS实例ID | 选择目标PolarDB-X的实例ID。                            |

7. 单击页面右下角的授权白名单并进入下一步。

**说明**

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置同步策略及对象信息。



| 配置项目        | 配置说明  |
|-------------|---|
| 目标已存在表的处理模式 | <ul style="list-style-type: none"> <li>○ <b>预检查并报错拦截</b>：检查同步的目标表是否为空。如果为空，则通过该检查项目；如果不为空，则在预检查阶段提示错误，数据同步作业不会被启动。</li> <li>○ <b>忽略报错并继续执行</b>：跳过同步的目标表是否为空的检查项。</li> </ul> <div style="border: 1px solid #ccc; background-color: #fff9c4; padding: 10px; margin-top: 10px;"> <p> <b>警告</b> 选择为忽略报错并继续执行，可能导致数据不一致，给业务带来风险，例如：</p> <ul style="list-style-type: none"> <li>■ 表结构一致的情况下，在目标库遇到与源库主键的值相同的记录，则会保留目标实例中的该条记录，即源库中的该条记录不会同步至目标数据库中。</li> <li>■ 表结构不一致的情况下，可能会导致无法初始化数据、只能同步部分列的数据或同步失败。</li> </ul> </div> |
| 选择同步对象      | <p>在源库对象框中单击待同步的表，然后单击  将其移动至已选择对象框。</p> <div style="border: 1px solid #ccc; background-color: #e1f5fe; padding: 10px; margin-top: 10px;"> <p> <b>说明</b></p> <ul style="list-style-type: none"> <li>○ 同步对象的选择粒度为表。</li> <li>○ 默认情况下，同步对象的名称保持不变。如果您需要改变同步对象在目标实例中的名称，那么需要使用DTS提供的对象名映射功能，详情请参见 <a href="#">设置同步对象在目标实例中的名称</a>。</li> </ul> </div>   |
| 映射名称更改      | <p>如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见 <a href="#">库表映射</a>。</p>   |

| 配置项目          | 配置说明  |
|---------------|---|
| 源、目标库无法连接重试时间 | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <p><b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p> |

- 单击下一步。
- 选择是否要执行全量数据初始化。

**说明** 全量数据初始化：DTS将源库中待同步表的存量数据同步至目标库中，如果不选择则不同步存量数据。

- 上述配置完成后，单击页面右下角的**预检查并启动**。

**说明**

- 在同步作业正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步作业。
- 如果预检查失败，单击具体检查项后的，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择**确认屏蔽**、**忽略告警项**并重新进行预检查，跳过告警检测项重新进行预检查。

- 在**预检查**对话框中显示**预检查通过**后，关闭**预检查**对话框，同步作业将正式开始。
- 等待同步作业的链路初始化完成，直至处于**同步中**状态。

您可以在**数据同步**页面，查看数据同步作业的状态。

| 实例ID/作业名称               | 状态  | 同步概况                                   | 付费方式 | 同步架构(全部) | 操作              |
|-------------------------|-----|--|------|----------|-----------------|
| hangzhou-hangzhou-small | 同步中 | 耗时: 1376 毫秒<br>速度: 0.00RPS/(0.000MB/s) | 按量付费 | 单向同步     | 暂停同步 转包年包月 升级更多 |

## 10.2. 从PolarDB-X同步至云原生数据仓库AnalyticDB MySQL

云原生数据仓库AnalyticDB MySQL是阿里巴巴自主研发的海量数据实时高并发在线分析（Real time OLAP）云计算服务，可以对千亿级数据进行毫秒级的即时多维分析透视和业务探索。通过数据传输服务DTS（Data Transmission Service），您可以将PolarDB-X同步到云原生数据仓库AnalyticDB MySQL，帮助您快速构建企业内部BI、交互查询、实时报表等系统。

### 前提条件

- PolarDB-X中的数据库须基于RDS MySQL创建。
- 已创建目标云原生数据仓库AnalyticDB MySQL集群，详情请参见[创建云原生数据仓库AnalyticDB MySQL集群](#)。

- 确保目标云原生数据仓库AnalyticDB MySQL具备充足的存储空间。

## 注意事项

- DTS在执行全量数据初始化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。因此您需要在执行数据同步前评估源库和目标库的性能，同时建议您在业务低峰期执行数据同步（例如源库和目标库的CPU负载在30%以下）。
- 数据同步期间，请勿对PolarDB-X执行扩容、缩容、迁移热点表、变更拆分键和变更DDL等操作，否则将导致数据同步失败。
- 如果需要在数据同步期间切换PolarDB-X的网络类型，在您执行完网络类型切换操作后，请[提交工单](#)调整同步链路的网络连接信息。
- 请勿在数据同步时，对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。
- 由于云原生数据仓库AnalyticDB MySQL本身的使用限制，当云原生数据仓库AnalyticDB MySQL集群中的节点磁盘空间使用量超过80%，该集群将被锁定。请提前根据待同步的对象预估所需空间，确保目标集群具备充足的存储空间。
- 暂不支持同步前缀索引，如果源库存在前缀索引可能导致数据同步失败。

## 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

## 支持同步的SQL操作

INSERT、UPDATE、DELETE。

## 数据库账号的权限要求

| 数据库                     | 所需权限  |
|-------------------------|---|
| PolarDB-X               | Replication slave、Replication client及待同步对象的Select权限，由DTS自动执行授权。 |
| 云原生数据仓库AnalyticDB MySQL | 读写权限。   |

## 数据类型映射关系

详情请参见[结构初始化涉及的数据类型映射关系](#)。

## 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

 **说明** 购买时，选择源实例为PolarDB-X（原DRDS升级版）、目标实例为AnalyticDB MySQL，并选择同步拓扑为单向同步。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击[数据同步](#)。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击[配置同步链路](#)。
6. 配置同步作业的源实例及目标实例信息。

1.选择同步通道的源及目标实例    2.ADS账号授权    3.选择同步对象    4.预检查

同步作业名称: DRDS\_TO\_AnalyticDB for MySQL

**源实例信息**

实例类型: DRDS实例  
实例地区: 华东1 (杭州)  
\* DRDS实例ID: drds-xxxxxx

**目标实例信息**

实例类型: ADS  
实例地区: 华东1 (杭州)  
\* 版本:  2.0  3.0  
\* 数据库: am-bp-xxxxxx  
\* 数据库账号: dtstest  
\* 数据库密码: .....

取消    授权白名单并进入下一步

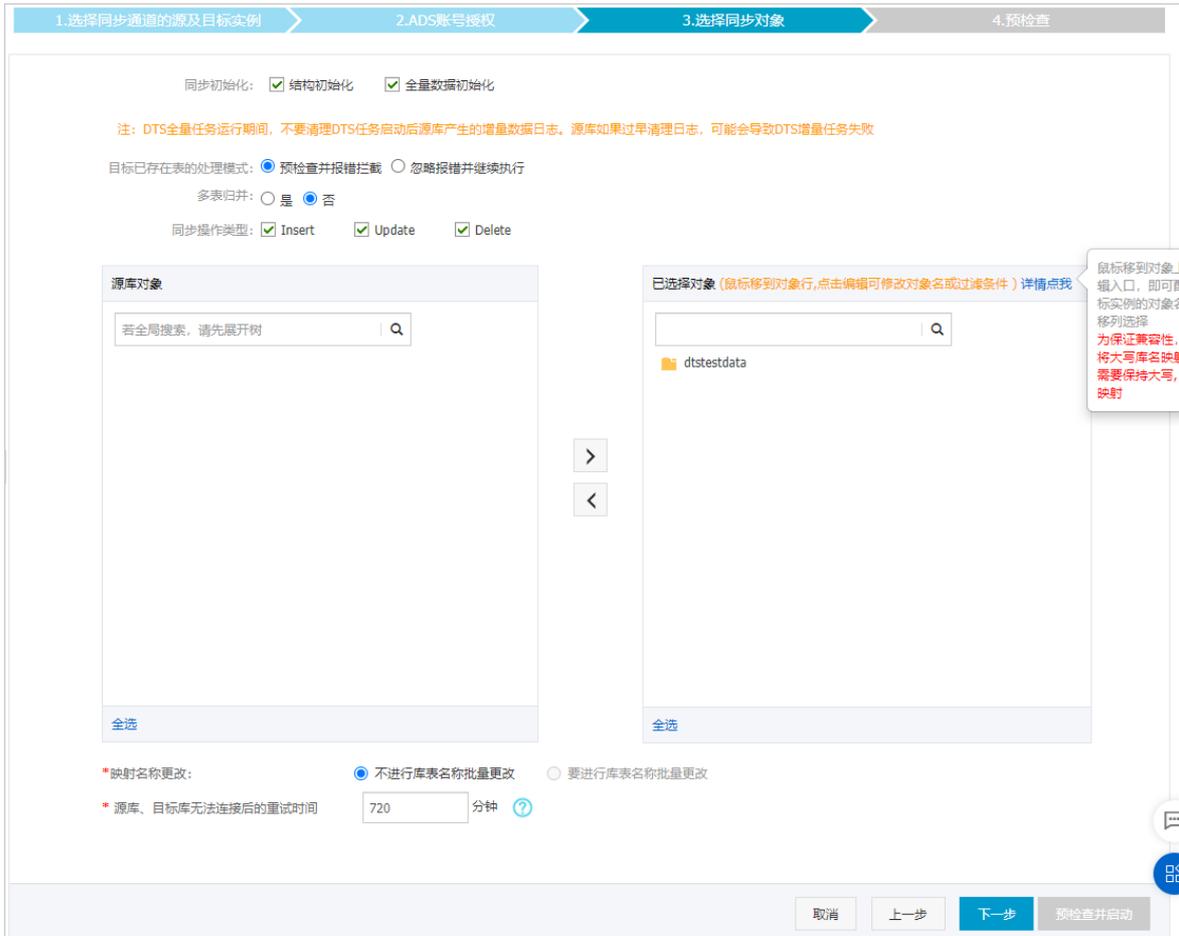
| 类别     | 配置       | 说明   |
|--------|----------|--|
| 无      | 同步作业名称   | DTS会自动生成一个同步作业名称，建议配置具有业务意义的名称（无唯一性要求），便于后续识别。     |
| 源实例信息  | 实例类型     | 固定为DRDS实例，不可变更。                                    |
|        | 实例地区     | 购买数据同步实例时选择的源实例地域信息，不可变更。                          |
|        | DRDS实例ID | 选择源PolarDB-X的实例ID。                                 |
| 目标实例信息 | 实例类型     | 固定为ADS，不可变更。                                       |
|        | 实例地区     | 购买数据同步实例时选择的目标实例地域信息，不可变更。                         |
|        | 版本       | 选择为3.0。  |
|        | 数据库      | 选择云原生数据仓库AnalyticDB MySQL实例ID。                     |
|        | 数据库账号    | 填入云原生数据仓库AnalyticDB MySQL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    | 填入该数据库账号的密码。                                       |

7. 单击页面右下角的授权白名单并进入下一步。

**说明**

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见DTS服务器的IP地址段。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置同步策略及对象信息。



| 配置          | 说明  |
|-------------|---|
| 同步初始化       | 默认情况下，您需要同时选中 <b>结构初始化</b> 和 <b>全量数据初始化</b> 。预检查完成后，DTS会将源实例中待同步对象的结构及数据在目标集群中初始化，作为后续增量同步数据的基线数据。  |
| 目标已存在表的处理模式 | <ul style="list-style-type: none"> <li><b>预检查并报错拦截</b>：检查目标数据库中是否有同名的表。如果目标数据库中没有同名的表，则通过该检查项目；如果目标数据库中有同名的表，则在预检查阶段提示错误，数据同步作业不会被启动。</li> </ul> <p><b>说明</b> 如果目标库中同名的表不方便删除或重命名，您可以更改该表在目标库中的名称，详情请参见<a href="#">设置同步对象在目标实例中的名称</a>。</p> <ul style="list-style-type: none"> <li><b>忽略报错并继续执行</b>：跳过目标数据库中是否有同名表的检查项。</li> </ul> <p><b>警告</b> 选择为<b>忽略报错并继续执行</b>，可能导致数据不一致，给业务带来风险，例如：</p> <ul style="list-style-type: none"> <li>表结构一致的情况下，在目标库遇到与源库主键的值相同的记录，则会保留目标集群中的该条记录，即源库中的该条记录不会同步至目标数据库中。</li> <li>表结构不一致的情况下，可能会导致无法初始化数据、只能同步部分列的数据或同步失败。</li> </ul> |

| 配置            | 说明  |
|---------------|---|
| 多表归并          | <ul style="list-style-type: none"> <li>选择为是：DTS将在每个表中增加 <code>__dts_data_source</code> 列来存储数据来源。</li> <li>选择为否：默认选项。</li> </ul> <p><b>说明</b> 多表归并功能基于任务级别，即不支持基于表级别执行多表归并。如果需要让部分表执行多表归并，另一部分不执行多表归并，您可以创建两个数据同步作业。</p> |
| 同步操作类型        | <p>根据业务选中需要同步的操作类型，默认情况下都处于选中状态。</p> <ul style="list-style-type: none"> <li>Insert</li> <li>Update</li> <li>Delete</li> </ul>   |
| 选择同步对象        | <p>在源库对象框中单击待同步的表，然后单击 <b>&gt;</b> 将其移动至已选择对象框。</p> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>同步对象的选择粒度为表。</li> <li>如果需要目标表中列信息与源表不同，则需要使用DTS的字段映射功能，详情请参见<a href="#">设置同步对象在目标实例中的名称</a>。</li> </ul>     |
| 映射名称更改        | <p>如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见<a href="#">库表列映射</a>。</p>   |
| 源、目标库无法连接重试时间 | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <p><b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p>   |

9. 上述配置完成后，单击页面右下角的下一步。

10. 设置待同步的表在目标库中类型。

创建同步作业 | 返回数据同步列表

1. 选择同步源和目标实例 > 2. ADS账号授权 > 3. 选择同步对象 > 4. 预检查

鼠标悬停列表表名会显示源库表名

| ADS表组 | ADS表名 | 类型(全部) | 主键列 | 分区列 | 分区数 | 定义状态(全部) |
|-------|-------|--------|-----|-----|-----|----------|
|       |       |        |     |     |     |          |

**说明** 选择了结构初始化后，您需要定义待同步的表在云原生数据仓库AnalyticDB MySQL中的类型、主键列、分区列等信息，详情请参见[CREATE TABLE操作手册](#)。

11. 上述配置完成后，单击页面右下角的预检查并启动。

**说明**

- 在同步作业正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步作业。
- 如果预检查失败，单击具体检查项后的 ，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择**确认屏蔽**、**忽略告警项**并重新进行预检查，跳过告警检测项重新进行预检查。

12. 在**预检查**对话框中显示**预检查通过**后，关闭**预检查**对话框，同步作业将正式开始。

13. 等待同步作业的链路初始化完成，直至处于**同步中**状态。

您可以在**数据同步**页面，查看数据同步作业的状态。



## 10.3. 从PolarDB-X同步至AnalyticDB for PostgreSQL

分析型数据库PostgreSQL版（原HybridDB for PostgreSQL）为您提供简单、快速、经济高效的PB级云端数据仓库解决方案。通过数据传输服务DTS（Data Transmission Service），您可以将PolarDB-X同步至AnalyticDB for PostgreSQL，帮助您快速实现对海量数据的即席查询分析、ETL处理和可视化探索。

### 前提条件

- PolarDB-X中的数据库须基于RDS MySQL创建。
- 源库中待同步的表必须具备主键。
- 已创建AnalyticDB for PostgreSQL实例，详情请参见[创建AnalyticDB for PostgreSQL实例](#)。

### 注意事项

- DTS在执行全量数据初始化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。因此您需要在执行数据同步前评估源库和目标库的性能，同时建议您在业务低峰期执行数据同步（例如源库和目标库的CPU负载在30%以下）。
- 数据同步期间，请勿对PolarDB-X执行扩容、缩容、迁移热点表、变更拆分键和变更DDL等操作，否则将导致数据同步失败。
- 如果需要在数据同步期间切换PolarDB-X的网络类型，在您执行完网络类型切换操作后，请[提交工单](#)调整同步链路的网络连接信息。
- 在成功启动数据同步作业后，如果在源库中新增了一个表并需要将其作为同步对象，您需要为数据同步作业[新增同步对象](#)。

### 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

### 功能限制

- 仅支持表级别的数据同步。
- 不支持同步JSON、GEOMETRY、CURVE、SURFACE、MULTIPOINT、MULTILINESTRING、MULTIPOLYGON、GEOMETRYCOLLECTION类型的数据。

### 支持同步的SQL操作

INSERT、UPDATE、DELETE。

### 术语/概念对应关系

|           |                           |
|-----------|---------------------------|
| PolarDB-X | AnalyticDB for PostgreSQL |
| Database  | Schema                    |
| Table     | Table                     |

### 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

? **说明** 购买时，选择源实例为PolarDB-X（原DRDS升级版）、目标实例为AnalyticDB for PostgreSQL，并选择同步拓扑为单向同步。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击数据同步。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击配置同步链路。
6. 配置同步作业的源实例及目标实例信息。

1.选择同步通道的源及目标实例
2.选择同步对象
3.预检查

同步作业名称:

**源实例信息**

实例类型: DRDS实例

实例地区: 华东1 (杭州)

\* DRDS实例ID:

**目标实例信息**

实例类型: AnalyticDB for PostgreSQL

实例地区: 华东1 (杭州)

\* 实例ID:

\* 数据库名称:

\* 数据库账号:

\* 数据库密码:

| 类别 | 配置 | 说明 |
|----|----|----|
|----|----|----|

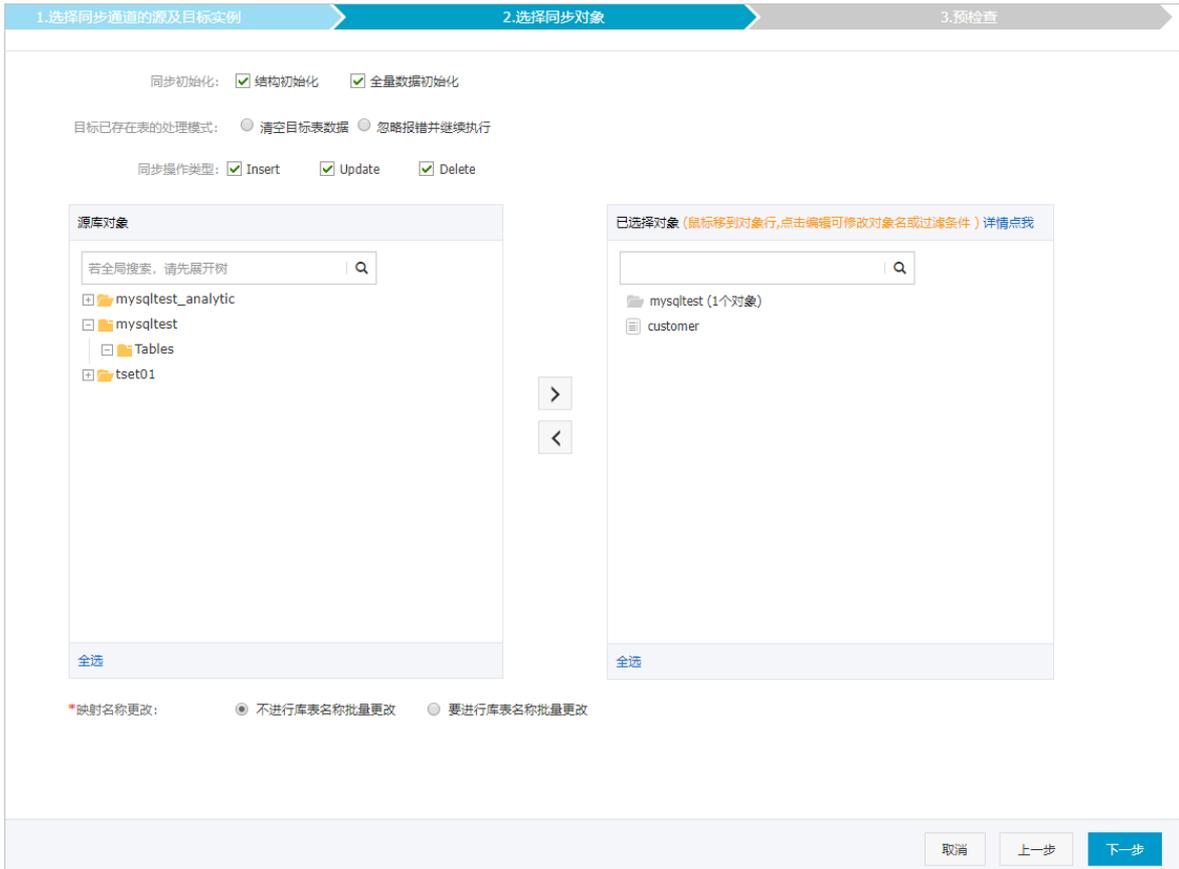
| 类别     | 配置       | 说明  |
|--------|----------|---|
| 无      | 同步作业名称   | DTS会自动生成一个同步作业名称，建议配置具有业务意义的名称（无唯一性要求），便于后续识别。  |
| 源实例信息  | 实例类型     | 固定为DRDS实例，不可变更。                                 |
|        | 实例地区     | 购买数据同步作业时选择的源实例地域信息，不可变更。                       |
|        | DRDS实例ID | 选择PolarDB-X的实例ID。                               |
| 目标实例信息 | 实例类型     | 固定为AnalyticDB for PostgreSQL，不可变更。              |
|        | 实例地区     | 购买数据同步作业时选择的目标实例地域信息，不可变更。                      |
|        | 实例ID     | 选择AnalyticDB for PostgreSQL实例ID。                |
|        | 数据库名称    | 填入目标AnalyticDB for PostgreSQL实例中，同步目标表所属的数据库名称。 |
|        | 数据库账号    | 填入AnalyticDB for PostgreSQL的数据库账号。              |
|        | 数据库密码    | 填入该数据库账号对应的密码。                                  |

7. 单击页面右下角的**授权白名单**并进入下一步。

 说明

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置同步策略及对象信息。



| 类别     | 配置          | 说明  |
|--------|-------------|---|
| 同步策略配置 | 同步初始化       | 默认情况下，您需要同时勾选 <b>结构初始化</b> 和 <b>全量数据初始化</b> 。预检查完成后，DT S会将源实例中待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。   |
|        | 目标已存在表的处理模式 | <ul style="list-style-type: none"> <li><b>清空目标表的数据</b><br/>在预检查阶段跳过<b>目标表是否为空</b>的检查项目。全量初始化之前将目标表的数据清空。适用于完成同步作业测试后的正式同步场景。</li> <li><b>忽略报错并继续执行</b><br/>在预检查阶段跳过<b>目标表是否为空</b>的检查项目。全量初始化时直接追加同步数据。适用于多张表同步到一张表的汇总同步场景。</li> </ul> |
|        | 同步操作类型      | 根据业务需求选择需要同步的操作类型： <ul style="list-style-type: none"> <li>Insert</li> <li>Update</li> <li>Delete</li> </ul>   |

| 类别            | 配置 | 说明  |
|---------------|----|---|
| 选择同步对象        | 无  | <p>在源库对象框中单击待同步的表，然后单击  将其移动至已选择对象框。</p> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>同步对象的选择粒度为表。</li> <li>如果需要目标表中列信息与源表不同，则需要使用DTS的字段映射功能，详情请参见<a href="#">设置同步对象在目标实例中的名称</a>。</li> </ul> |
| 映射名称更改        | 无  | <p>如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见<a href="#">库表列映射</a>。</p>   |
| 源、目标库无法连接重试时间 | 无  | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <p><b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p>   |

9. 上述配置完成后，单击页面右下角的下一步。

10. 设置待同步的表在AnalyticDB for PostgreSQL中的主键列和分布列信息。

**说明** 在第8步中选择了结构初始化才会出现该页面。关于主键列和分布列的详细说明，请参见[表的约束定义](#)和[表分布键定义](#)。

11. 上述配置完成后，单击页面右下角的预检查并启动。

**说明**

- 在同步作业正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步作业。
- 如果预检查失败，单击具体检查项后的 ，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择[确认屏蔽](#)、[忽略告警项](#)并重新进行预检查，跳过告警检测项重新进行预检查。

12. 在预检查对话框中显示预检查通过后，关闭预检查对话框，同步作业将正式开始。

13. 等待同步作业的链路初始化完成，直至处于同步中状态。

您可以在数据同步页面，查看数据同步作业的状态。

| 实例ID/作业名称  | 状态  | 同步概况                                  | 付费方式 | 同步架构(全部) | 操作  |
|--|-----|---------------------------------------|------|----------|---|
| <input type="checkbox"/> hangzhou-hangzhou-small | 同步中 | 延时: 1376 毫秒<br>速率: 0.00RPS(0.000MB/s) | 按量付费 | 单向同步     | <a href="#">暂停同步</a>   <a href="#">转包年包月</a>   <a href="#">升级更多</a> |

共有1条， 每页显示：20条

# 11.Redis数据同步

## 11.1. Redis实例间的单向数据同步

数据传输服务DTS (Data Transmission Service) 支持Redis数据库的单向同步, 适用于异地多活、数据异地容灾等多种应用场景。本文以Redis实例间的单向数据同步为例, 介绍数据同步作业的配置流程。

**警告** 完成数据同步作业的配置后, 请勿变更源数据库或目标数据库的**架构类型** (例如将主从架构变更为集群架构), 否则会导致数据同步失败。

### 前提条件

- 源和目标Redis实例支持的版本, 请参见[同步方案概览](#)。

**说明** 如需跨版本同步 (仅支持从低版本同步到高版本) 请提前确认兼容性。例如创建按量付费的Redis实例来测试, 测试完成后可将该实例释放或转为包年包月。

- 目标Redis实例的存储空间需大于源Redis实例已使用的存储空间。
- 源云数据库Redis为企业版 (持久内存型) 时, 需要用户开启appendonly参数。
- 云数据库Redis企业版 (容量存储型) 不支持做为源数据库, 只能做为目标数据库。
- 源Redis实例Slave和Master之间的复制超时时间参数repl-timeout, 默认为60秒, 建议使用 `config set repl-timeout 600` 命令设置为600秒。若源数据库数据量比较大, 可以适当增大repl-timeout参数的值。

### 注意事项

- DTS在执行全量数据初始化时将占用源库和目标库一定的资源, 可能会导致数据库服务器负载上升。如果数据库业务量较大或服务器规格较低, 可能会加重数据库压力, 甚至导致数据库服务不可用。建议您在执行数据同步前谨慎评估, 在业务低峰期执行数据同步。
- 如果源库中的某些Key使用了过期 (expire) 策略, 由于可能存在Key已过期但未被及时删除的情况, 所以在目标库中查看到的Key数量 (例如通过info命令查看) 会比源库的Key数量少。

**说明** 源和目标库中, 未设置过期策略或未过期的Key数量是一致的。

- 当源Redis实例的版本为2.8时不支持增量同步。
- 数据同步期间, 请勿在源库中执行 `FLUSHDB` 和 `FLUSHALL` 命令, 否则将导致源和目标的数据不一致。
- 如果目标库的数据逐出策略 ( `maxmemory-policy` ) 配置为 `noeviction` 以外的值, 可能导致目标库的数据与源库不一致。关于数据逐出策略详情, 请参见[Redis数据逐出策略介绍](#)。
- 同步期间, 如源Redis实例发生扩缩容 (如增加或者减少分片)、规格变配 (如扩大内存), 则您需重新配置任务。且为保障数据一致性, 在重新配置任务前, 建议先清空已同步至目标Redis的数据。
- 同步期间, 如源Redis实例的连接地址变化 (如迁移可用区、经典网络切换专有网络), 则您需要[提交工单](#)进行处理, 如果长时间未处理, 源Redis实例会重置AOF文件, 届时您需重新配置任务。
- 源库单机版Redis同步到目标库集群版Redis的操作限制: 由于集群cluster只允许单个命令操作单个slot, 若在源库执行多Key操作时, Key不在同一个slot或涉及多个slot, 则会出现以下报错:

```
CROSSSLOT Keys in request don't hash to the same slot
```

建议在DTS同步过程中仅执行单Key操作, 以免导致链路中断。

### 费用说明

| 同步类型          | 链路配置费用 |
|---------------|--------|
| 库表结构同步和全量数据同步 | 不收费。   |

| 同步类型   | 链路配置费用                          |
|--------|---------------------------------|
| 增量数据同步 | 收费，详情请参见 <a href="#">产品定价</a> 。 |

### 支持的同步拓扑

- 一对一单向同步
- 一对多单向同步
- 级联单向同步

关于各类同步拓扑的介绍及注意事项，请参见[数据同步拓扑介绍](#)。

### 支持的同步命令

- APPEND
- BITOP、BLPOP、BRPOP、BRPOPLPUSH
- DECR、DECRBY、DEL
- EVAL、EVALSHA、EXEC、EXPIRE、EXPIREAT
- GEOADD、GETSET
- HDEL、HINCRBY、HINCRBYFLOAT、HMSET、HSET、HSETNX
- INCR、INCRBY、INCRBYFLOAT
- LINSERT、LPOP、LPUSH、LPUSHX、LREM、LSET、LTRIM
- MOVE、MSET、MSETNX、MULTI
- PERSIST、PEXPIRE、PEXPIREAT、PFADD、PFMERGE、PSETEX
- RENAME、RENAMENX、RESTORE、RPOP、RPOPLPUSH、RPUSH、RPUSHX
- SADD、SDIFFSTORE、SELECT、SET、SETBIT、SETEX、SETNX、SETRANGE、SINTERSTORE、SMOVE、SPOP、SREM、SUNIONSTORE
- ZADD、ZINCRBY、ZINTERSTORE、ZREM、ZREMRANGEBYLEX、ZUNIONSTORE、ZREMRANGEBYRANK、ZREMRANGEBYSCORE
- SWAPDB、UNLINK（仅当源端Redis实例的版本为4.0时支持）

#### 说明

- 不支持同步PUBLISH命令。
- 对于通过EVAL或者EVALSHA调用Lua脚本，在增量数据同步时，由于目标端在执行脚本时不会明确返回执行结果，DTS无法确保该类型脚本能够执行成功。
- 对于List，由于DTS在调用sync或psync进行重传时，不会对目标端已有的数据进行清空，可能导致出现重复数据。

### 数据库账号的权限要求

| 数据库       | 权限及授权方式                                   |
|-----------|---|
| 源Redis实例  | 读权限，关于授权方式，请参见 <a href="#">创建与管理账号</a> 。  |
| 目标Redis实例 | 读写权限，关于授权方式，请参见 <a href="#">创建与管理账号</a> 。 |

### 操作步骤

1. 购买数据同步实例，详情请参见[购买数据同步实例](#)。

**说明** 购买时，源实例和目标实例均选择为Redis。

2. 登录[数据传输控制台](#)。

3. 在左侧导航栏，单击数据同步。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击配置同步链路。
6. 配置数据同步的源实例及目标实例信息。

1.选择同步通道的源及目标实例
2.选择同步对象
3.高级设置
4.预检查

同步作业名称：

**源实例信息**

实例类型：

实例地区：

\* 实例ID：

数据库密码：

**目标实例信息**

实例类型：

实例地区：

\* 实例ID：

数据库密码：

取消 授权白名单并进入下一步

| 类别     | 配置     | 说明  |
|--------|--------|---|
| 无      | 同步作业名称 | DTS会自动生成一个同步作业名称，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。  |
| 源实例信息  | 实例类型   | 选择Redis实例。  |
|        | 实例地区   | 购买数据同步实例时选择的源实例地域信息，不可变更。   |
|        | 实例ID   | 选择源Redis实例ID。   |
|        | 数据库密码  | 填入连接源Redis实例的数据库账号密码。账号权限，请参见 <a href="#">数据库账号的权限要求</a> 。<br><br><div style="border: 1px solid #add8e6; padding: 5px; background-color: #e6f2ff;"> <p><span style="font-size: 1em;">?</span> 说明 数据库密码格式为&lt;user&gt;:&lt;password&gt;。例如，Redis实例自定义的用户名为admin，密码为Rp829dlwa，则此处填入的数据库密码为admin:Rp829dlwa。</p> </div> |
| 目标实例信息 | 实例类型   | 选择Redis实例。  |
|        | 实例地区   | 购买数据同步实例时选择的目标实例地域信息，不可变更。  |
|        | 实例ID   | 选择目标Redis实例ID。  |

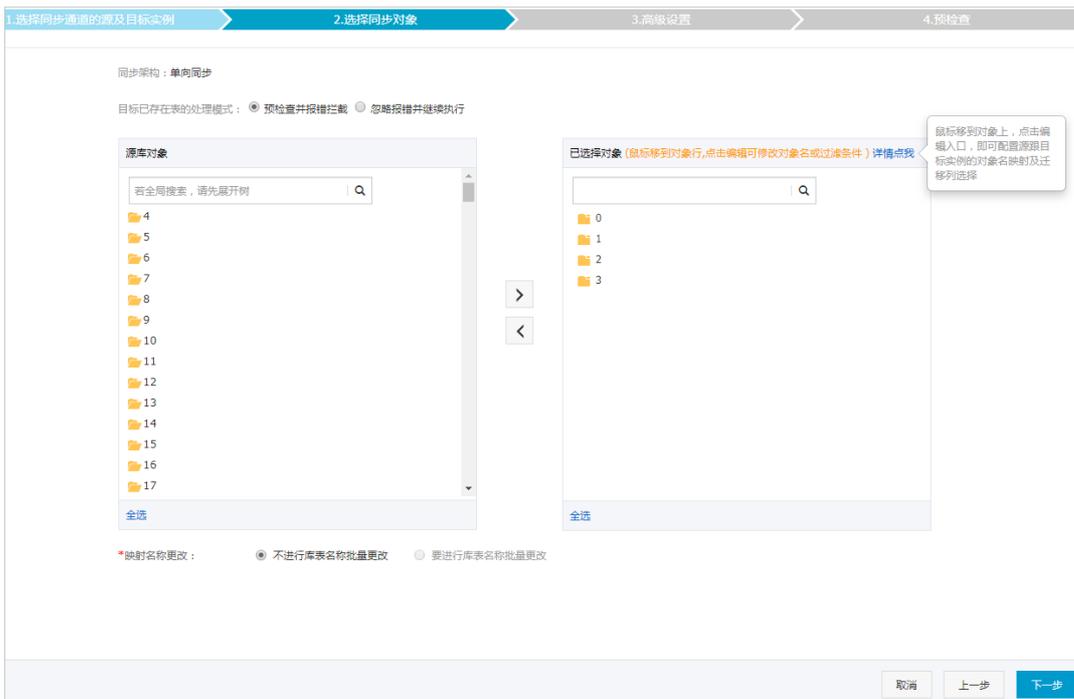
| 类别 | 配置    | 说明  |
|----|-------|---|
|    | 数据库密码 | <p>填入连接目标Redis实例的数据库账号密码。账号权限，请参见<a href="#">数据库账号的权限要求</a>。</p> <p><b>说明</b> 数据库密码格式为&lt;user&gt;:&lt;password&gt;。例如，Redis实例自定义的用户名为admin，密码为Rp829dlwa，则此处填入的数据库密码为admin:Rp829dlwa。</p> |

7. 单击页面右下角的授权白名单并进入下一步。

**说明**

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置目标已存在表的处理模式和同步对象。



| 配置          | 说明  |
|-------------|---|
| 目标已存在表的处理模式 | <ul style="list-style-type: none"> <li><b>预检查并报错拦截</b>：检查目标库是否为空。如果待同步的目标库为空，则通过该检查项目；如果不为空，则在预检查阶段提示错误，数据同步作业不会被启动。</li> <li><b>忽略报错并继续执行</b>：跳过目标库是否为空的检查项。</li> </ul> <p><b>警告</b> 选择为忽略报错并继续执行后，如果在同步过程中遇到目标库中的Key与源库中的Key相同，会将源库的数据覆盖写入目标库中，请谨慎选择。</p> |

| 配置                             | 说明  |
|--------------------------------|---|
| 同步对象                           | <ul style="list-style-type: none"> <li>在源库对象框中单击待同步的数据库，然后单击  将其移动到已选择对象框。</li> <li>同步对象的选择粒度为库，暂不支持Key粒度的选择。</li> </ul>   |
| 映射名称更改                         | 不支持更改映射名称。  |
| 源库DMS_ONLINE_DDL过程中是否复制临时表到目标库 | <p>如源库使用数据管理DMS (Data Management Service) 执行Online DDL变更，您可以选择是否同步Online DDL变更产生的临时表数据。</p> <ul style="list-style-type: none"> <li>是：同步Online DDL变更产生的临时表数据。</li> </ul> <p> 说明 Online DDL变更产生的临时表数据过大，可能会导致同步任务延迟。</p> <ul style="list-style-type: none"> <li>否：不同步Online DDL变更产生的临时表数据，只同步源库的原始DDL数据。</li> </ul> <p> 说明 该方案会导致目标库锁表。</p> |
| 源、目标库无法连接重试时间                  | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <p> 说明 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p>   |

9. 上述配置完成后单击页面右下角的下一步。

10. 配置同步初始化选项，当前固定为包含全量数据+增量数据。

1.选择同步通道的源及目标实例
2.选择同步对象
3.高级设置
4.预检查

同步初始化：包含全量数据+增量数据

取消
上一步
保存
预检查并启动

-  说明
- DTS会将源Redis实例中的存量数据同步至目标Redis实例中，并同步增量数据。
  - 如果跳出版本相关的报错提示，请您按照提示将源Redis实例升级至指定版本。关于升级版本方式，请参见[升级大版本](#)和[升级小版本](#)。

11. 上述配置完成后，单击页面右下角的预检查并启动。

-  说明
- 在同步任务正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步任务。
  - 如果预检查失败，单击具体检查项后的 ，查看失败详情。
    - 您可以根据提示修复后重新进行预检查。
    - 如无需修复告警检测项，您也可以选择确认屏蔽、忽略告警项并重新进行预检查，跳过告警检测项重新进行预检查。

- 在预检查对话框中显示预检查通过后，关闭预检查对话框，同步作业将正式开始。
- 等待同步作业的链路初始化完成，直至处于同步中状态。

| 实例ID/作业名称               | 状态  | 同步概况                                   | 付费方式 | 同步架构(全部) | 操作                  |
|-------------------------|-----|--|------|----------|---------------------|
| hangzhou-hangzhou-small | 同步中 | 耗时: 1376 毫秒<br>速度: 0.00RPS/(0.000MB/s) | 按量付费 | 单向同步     | 暂停同步   转包年包月   升级更多 |

**说明** 您可以在数据同步页面，查看数据同步作业的状态。

## 11.2. Redis企业版实例间的双向同步

数据传输服务DTS（Data Transmission Service）支持Redis企业版实例间的双向同步，适用于异地多活、数据容灾等多种应用场景，本文介绍数据同步作业的配置流程。

### 前提条件

源和目标实例为 [阿里云Redis企业版实例](#)（5.0版本）。

#### 说明

- 云数据库Redis企业版（容量存储型）不支持做为源数据库，只能做为目标数据库。
- 云数据库Redis为企业版（持久内存型）为源数据库时，需要用户开启appendonly参数。
- 架构类型不限制，即支持集群版、标准版和读写分离版。
- 源Redis实例Slave和Master之间的复制超时时间参数repl-timeout，默认为60秒，建议使用 `config set repl-timeout 600` 命令设置为600秒。若源数据库数据量比较大，可以适当增大repl-timeout参数的值。

### 注意事项

- 双向数据同步时，正向数据同步作业会执行和，反向数据同步作业仅执行增量数据同步。

**警告** 为保障数据一致性，双向数据同步作业运行期间，请勿在两端数据库同时对同一个key执行修改或写入操作。

#### 全量数据初始化增量数据同步

- DTS在执行全量数据初始化时将占用源库和目标库一定的资源，可能会导致数据库服务器负载上升。如果数据库业务量较大或服务器规格较低，可能会加重数据库压力，甚至导致数据库服务不可用。建议您在执行数据迁移前谨慎评估，在业务低峰期执行数据同步。
- 请勿在源实例中执行 `FLUSHDB` 和 `FLUSHALL` 命令，否则将导致源和目标的数据不一致。
- 如果目标库的数据逐出策略（`maxmemory-policy`）配置为 `noeviction` 以外的值，可能导致目标库的数据与源库不一致。关于数据逐出策略详情，请参见[Redis数据逐出策略介绍](#)。
- 如果源库中的某些Key使用了过期（`expire`）策略，由于可能存在Key已过期但未被及时删除的情况，所以在目标库中看到的Key数量（例如通过`info`命令查看）会比源库的Key数量少。

**说明** 源和目标库中，未设置过期策略或未过期的Key数量是一致的。

- 如果目标Redis实例没有开通直连访问，DTS将采用代理转发模式将数据写入目标实例。

**说明** 关于直连访问的开通方法，请参见[开通直连访问](#)。

- 同步期间，如源和目标Redis实例发生扩缩容（如增加或者减少分片）、规格变配（如扩大内存）、则您需重新配置任务。且为保障数据一致性，在重新配置任务前，建议先清空已同步至源和目标Redis的数据。

- 同步期间，如源和目标Redis实例的连接地址变化（如迁移可用区、经典网络切换专有网络），则需要[提交工单](#)进行处理，如果长时间未处理，源和目标Redis实例会重置AOF文件，届时您需重新配置任务。
- 如双向同步任务的源实例或目标实例位于海外地域，则仅支持同地域的双向同步，不支持跨地域的双向同步。例如，支持日本地域间的双向同步，不支持日本地域与法兰克福地域间的双向同步。
- 源库单机版Redis同步到目标库集群版Redis的操作限制：由于集群cluster只允许单个命令操作单个slot，若在源库执行多Key操作时，Key不在同一个slot或涉及多个slot，则会出现以下报错：

```
CROSSSLOT Keys in request don't hash to the same slot
```

建议在DTS同步过程中仅执行单Key操作，以免导致链路中断。

## 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

## 支持的同步命令

- APPEND
- BITOP、BLPOP、BRPOP、BRPOPLPUSH
- DECR、DECRBY、DEL
- EVAL、EVALSHA、EXEC、EXPIRE、EXPIREAT
- GEOADD、GETSET
- HDEL、HINCRBY、HINCRBYFLOAT、HMSET、HSET、HSETNX
- INCR、INCRBY、INCRBYFLOAT
- LINSERT、LPOP、LPUSH、LPUSHX、LREM、LSET、LTRIM
- MOVE、MSET、MSETNX、MULTI
- PERSIST、PEXPIRE、PEXPIREAT、PFADD、PFMERGE、PSETEX
- RENAME、RENAMENX、RPOP、RPOPLPUSH、RPUSH、RPUSHX
- SADD、SDIFFSTORE、SELECT、SET、SETBIT、SETEX、SETNX、SETRANGE、SINTERSTORE、SMOVE、SPOP、SREM、SUNIONSTORE
- UNLINK、ZADD、ZINCRBY、ZINTERSTORE、ZREM、ZREMRANGEBYLEX、ZUNIONSTORE、ZREMRANGEBYRANK、ZREMRANGEBYSCORE
- SWAPDB（当源或目标实例的架构为集群版时不支持）

### 注意

- 不支持同步PUBLISH命令。
- 对于通过EVAL或者EVALSHA调用Lua脚本，在增量数据同步时，由于目标端在执行脚本时不会明确返回执行结果，DTS无法确保该类型脚本能够执行成功。
- 对于List，由于DTS在调用sync或psync进行重传时，不会对目标端已有的数据进行清空，可能导致出现重复数据。

## 数据库账号的权限要求

| 数据库       | 权限及授权方式                                   |
|-----------|---|
| 源Redis实例  | 读写权限，关于授权方式，请参见 <a href="#">创建与管理账号</a> 。 |
| 目标Redis实例 |   |

### 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

**注意** 购买时选择源实例为Redis、目标实例为Redis，并选择同步拓扑为双向同步。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击数据同步。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 配置正向同步作业。
  - i. 定位至已购买的数据同步实例，单击该实例下第一个同步作业操作列的[配置同步链路](#)。

**注意** 一个双向数据同步实例会包含两个同步作业，需要分别进行配置。在配置反向同步作业时，定位至第二个同步作业，单击操作列的[配置同步链路](#)。

| 同步作业名称                   | 搜索        | 排序  | 默认排序 | 状态         | 全部                     |  |
|--------------------------|-----------|-----|------|------------|------------------------|--|
| <input type="checkbox"/> | 实例ID/作业名称 | 状态  | 同步概况 | 付费方式       | 同步架构(全部)               | 操作   |
| <input type="checkbox"/> | 实例ID      | --  | --   | 按量付费       | 双向同步                   | <a href="#">转包年包月</a>   <a href="#">升级</a>   <a href="#">查看同步作业</a>   <a href="#">更多</a> |
| <input type="checkbox"/> | 作业名称      | 状态  | 同步概况 | 源实例/目标实例   | 操作                     |  |
| <input type="checkbox"/> | 作业名称      | 未配置 |      | 未配置<br>未配置 | <a href="#">配置同步链路</a> |  |
| <input type="checkbox"/> | 作业名称      | 未配置 |      | 未配置<br>未配置 | <a href="#">配置同步链路</a> |  |

- ii. 配置同步作业的源实例及目标实例信息。

1.选择同步通道的源及目标实例
2.选择同步对象
3.高级设置
4.预检查

同步作业名称:

**源实例信息**

实例类型:

实例地区: 华北2 (北京)

\* 实例ID:

数据库密码:

**目标实例信息**

实例类型:

实例地区: 华北2 (北京)

\* 实例ID:

数据库密码:

| 类别 | 配置     | 说明   |
|----|--------|--|
| 无  | 同步作业名称 | DTS会自动生成一个同步作业名称，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。 |
|    | 实例类型   | 选择Redis实例。                                     |

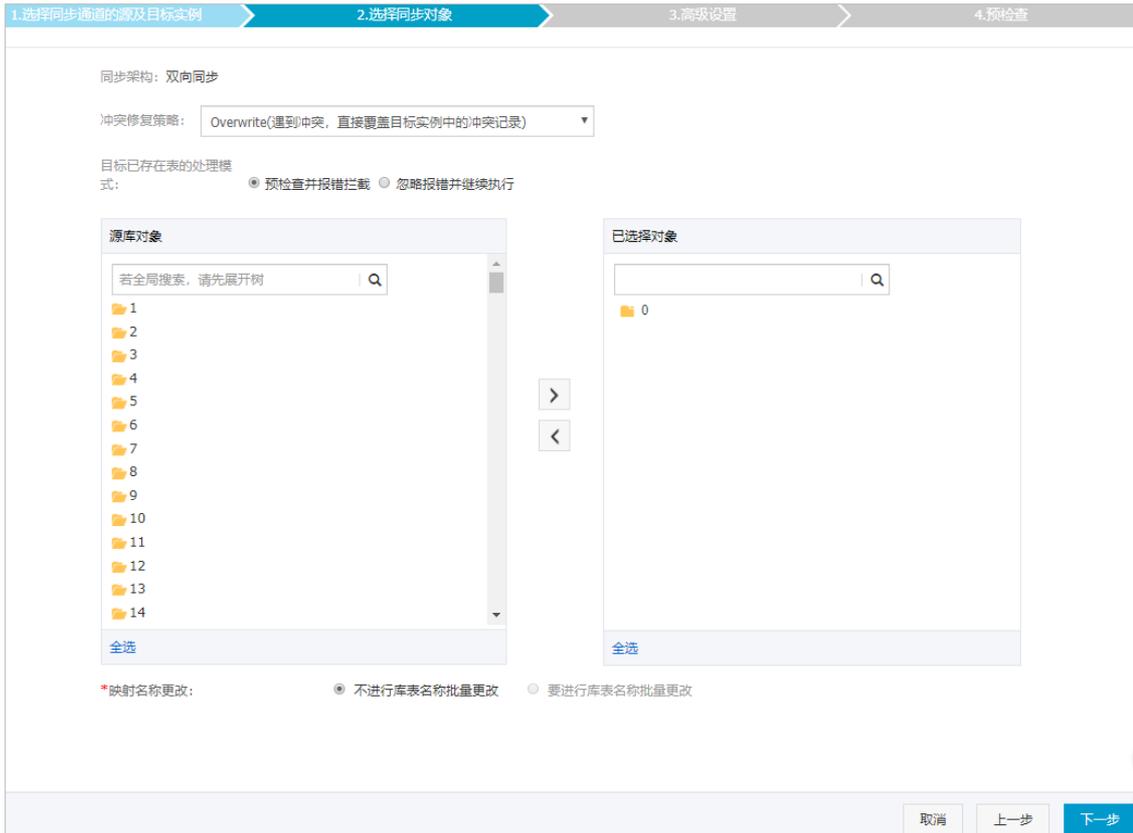
| 类别     | 配置    | 说明   |
|--------|-------|--|
| 源实例信息  | 实例地区  | 购买数据同步实例时选择的源实例地域，不可变更。  |
|        | 实例ID  | 选择源Redis实例ID。<br> <b>注意</b> 稍后配置反向同步作业时，此处需选择为正向同步作业里的目标Redis实例ID。  |
|        | 数据库密码 | 填入Redis实例的数据库账号密码。账号权限，请参见 <a href="#">数据库账号的权限要求</a> 。<br> <b>注意</b> 数据库密码格式为<user>:<password>。例如，Redis实例自定义的用户名为admin，密码为Rp829dlwa，则此处填入的数据库密码为admin:Rp829dlwa。   |
| 目标实例信息 | 实例类型  | 选择Redis实例。   |
|        | 实例地区  | 购买数据同步实例时选择的目标实例地域，不可变更。   |
|        | 实例ID  | 选择目标Redis实例ID。<br> <b>注意</b> 稍后配置反向同步作业时，此处需选择为正向同步作业里的源Redis实例ID。  |
|        | 数据库密码 | 填入Redis实例的数据库账号密码。账号权限，请参见 <a href="#">数据库账号的权限要求</a> 。<br> <b>注意</b> 数据库密码格式为<user>:<password>。例如，Redis实例自定义的用户名为admin，密码为Rp829dlwa，则此处填入的数据库密码为admin:Rp829dlwa。 |

iii. 单击页面右下角的**授权白名单并进入下一步**。

 **说明**

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

iv. 配置同步策略和同步对象。



| 类别     | 配置          | 说明   |
|--------|-------------|--|
| 同步策略   | 冲突修复策略      | <b>Overwrite</b> （遇到冲突，直接覆盖目标实例中的冲突记录）<br>当数据同步期间遇到Key相同但值不同时，Key的值更新时间晚的数据会覆盖冲突的记录。   |
|        | 目标已存在表的处理模式 | <ul style="list-style-type: none"> <li><b>预检查并报错拦截</b>：检查目标库是否为空。如果待同步的目标库为空，则通过该检查项目；如果不为空，则在预检查阶段提示错误，数据同步作业不会被启动。</li> <li><b>忽略报错并继续执行</b>：跳过目标库是否为空的检查项。</li> </ul> <div style="border: 1px solid #ccc; background-color: #fff9c4; padding: 5px; margin-top: 10px;"> <p> <b>警告</b> 选择为忽略报错并继续执行后，如果在同步初始化期间遇到目标库中的Key与源库中的Key相同，会将源库的数据覆盖写入目标库中，请谨慎选择。</p> </div> |
| 选择同步对象 | 无           | <ul style="list-style-type: none"> <li>在源库对象框中单击待同步的数据库，然后单击  将其移动到已选择对象框。</li> <li>同步对象的选择粒度为库，暂不支持Key粒度的选择。</li> </ul>   |
| 映射名称更改 | 无           | 不支持更改映射名称。   |

| 类别                                 | 配置 | 说明  |
|------------------------------------|----|---|
| 源表<br>DMS_ONLINE_DDL过程中是否复制临时表到目标库 | 无  | <p>如源库使用数据管理DMS (Data Management Service) 执行Online DDL变更，您可以选择是否同步Online DDL变更产生的临时表数据。</p> <ul style="list-style-type: none"> <li>是：同步Online DDL变更产生的临时表数据。</li> </ul> <p><b>说明</b> Online DDL变更产生的临时表数据过大，可能会导致同步任务延迟。</p> <ul style="list-style-type: none"> <li>否：不同步Online DDL变更产生的临时表数据，只同步源库的原始DDL数据。</li> </ul> <p><b>说明</b> 该方案会导致目标库锁表。</p> |
| 源、目标库无法连接重试时间                      | 无  | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <p><b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p>   |

v. 上述配置完成后，单击页面右下角的下一步。

vi. 配置同步初始化的选项。



当前固定为包含全量数据+增量数据，DTS会将源Redis实例中同步对象的存量数据初始化至目标Redis实例，并同步增量数据。

**注意**

- 在配置反向同步作业时，如果同步对象已经初始化至目标实例，则直接同步增量数据。
- 如果跳出版本相关的报错提示，请您按照提示将源Redis实例升级至指定版本。关于升级版本方式，请参见[升级大版本](#)和[升级小版本](#)。

vii. 上述配置完成后，单击页面右下角的**预检查并启动**。

**说明**

- 在同步作业正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步作业。
- 如果预检查失败，单击具体检查项后的 ，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择**确认屏蔽**、**忽略告警项**并重新进行预检查，跳过告警检测项重新进行预检查。

viii. 在**预检查**对话框中显示**预检查通过**后，关闭**预检查**对话框，正向同步作业将正式开始。

6. 等待正向同步作业完成初始化，直至处于**同步中**状态。

您可以在**数据同步**页面，查看数据同步作业的状态。

7. 配置反向同步作业。

i. 定位至第二个同步作业，单击**配置同步链路**。

| 实例ID/作业名称                              | 状态  | 同步概况                           | 付费方式                      | 同步架构(全部)                                  | 操作  |
|--|-----|--------------------------------|---------------------------|---|---|
| <input type="checkbox"/> [实例ID]        | --  | --                             | 按量付费                      | 双向同步                                      | <a href="#">转包年包月</a>   <a href="#">升级</a><br><a href="#">查看同步作业</a>   <a href="#">更多</a> |
| 作业名称                                   | 状态  | 同步概况                           | 源实例/目标实例                  | 操作  |   |
| <input type="checkbox"/> Redis_Forward | 同步中 | 延时: 0 毫秒<br>速度: 0TPS(0.00MB/s) | r-[实例ID]eo<br>r-[实例ID]ndm | <a href="#">暂停同步</a>   <a href="#">更多</a> |   |
| <input type="checkbox"/> Redis_Reverse | 未配置 |                                | r-[实例ID]ndm<br>r-[实例ID]eo | <a href="#">配置同步链路</a>                    |   |

ii. 重复**步骤5**中的ii到viii配置步骤，完成反向同步作业的配置。

### 执行结果

等待一段时间后，两个同步作业的链路状态均会处于**同步中**。

| 实例ID/作业名称                              | 状态  | 同步概况                                | 付费方式                      | 同步架构(全部)                                  | 操作  |
|--|-----|-------------------------------------|---------------------------|---|---|
| <input type="checkbox"/> dts-[实例ID]    | --  | --                                  | 按量付费                      | 双向同步                                      | <a href="#">转包年包月</a>   <a href="#">升级</a><br><a href="#">查看同步作业</a>   <a href="#">更多</a> |
| 作业名称                                   | 状态  | 同步概况                                | 源实例/目标实例                  | 操作  |   |
| <input type="checkbox"/> Redis_Forward | 同步中 | 延时: 0 毫秒<br>速度: 0.00TPS/(0.000MB/s) | r-[实例ID]eo<br>r-[实例ID]ndm | <a href="#">暂停同步</a>   <a href="#">更多</a> |   |
| <input type="checkbox"/> Redis_Reverse | 同步中 | 延时: 0 毫秒<br>速度: 0.00TPS/(0.000MB/s) | r-[实例ID]ndm<br>r-[实例ID]eo | <a href="#">暂停同步</a>   <a href="#">更多</a> |   |

## 11.3. 调用OpenAPI配置Redis企业版实例间单向或双向数据同步

您可以通过调用DTS的相关接口，为Redis企业版实例配置单向或双向数据同步。本文介绍相关注意事项及示例代码。本文展示的示例代码基于Java语言，您需要将示例代码中的 `<accessKeyId>` 和 `<accessSecret>` 分别替换为您的AccessKeyId和AccessKeySecret。您也可以在**OpenAPI Explorer**中直接运行相关接口，免去您计算签名的困扰。运行成功后，OpenAPI Explorer可以自动生成SDK代码示例。

### 前提条件

源和目标实例为**阿里云Redis企业版实例**（5.0版本）。

说明

- 单向同步时，不支持源实例为容量存储型；双向同步时，不支持源和目标实例为容量存储型。
- 源云数据库Redis为企业版（持久内存型）时，需要用户开启appendonly参数。
- 架构类型不限制，即支持集群版、标准版和读写分离版。
- 源Redis实例Slave和Master之间的复制超时时间参数repl-timeout，默认为60秒，建议使用 `config set repl-timeout 600` 命令设置为600秒。若源数据库数据量比较大，可以适当增大repl-timeout参数的值。

注意事项

- 双向数据同步时，正向数据同步作业会执行和，反向数据同步作业仅执行增量数据同步。

**警告** 为保障数据一致性，双向数据同步作业运行期间，请勿在两端数据库同时对同一个key执行修改或写入操作。

全量数据初始化增量数据同步

- DTS在执行全量数据初始化时将占用源库和目标库一定的资源，可能会导致数据库服务器负载上升。如果数据库业务量较大或服务器规格较低，可能会加重数据库压力，甚至导致数据库服务不可用。建议您在执行数据迁移前谨慎评估，在业务低峰期执行数据同步。
- 请勿在源实例中执行 `FLUSHDB` 和 `FLUSHALL` 命令，否则将导致源和目标的数据不一致。
- 如果目标库的数据逐出策略（`maxmemory-policy`）配置为 `noeviction` 以外的值，可能导致目标库的数据与源库不一致。关于数据逐出策略详情，请参见[Redis数据逐出策略介绍](#)。
- 如果源库中的某些Key使用了过期（`expire`）策略，由于可能存在Key已过期但未被及时删除的情况，所以在目标库中查看到的Key数量（例如通过`info`命令查看）会比源库的Key数量少。

**说明** 源和目标库中，未设置过期策略或未过期的Key数量是一致的。

- 如果目标Redis实例没有开通直连访问，DTS将采用代理转发模式将数据写入目标实例。

**说明** 关于直连访问的开通方法，请参见[开通直连访问](#)。

- 同步期间，如源Redis实例发生扩缩容（如增加或者减少分片）、规格变配（如扩大内存），则您需重新配置任务。且为保障数据一致性，在重新配置任务前，建议先清空已同步至目标Redis的数据。
- 同步期间，如源Redis实例的连接地址变化（如迁移可用区、经典网络切换专有网络），则您需要[提交工单](#)进行处理，如果长时间未处理，源Redis实例会重置AOF文件，届时您需重新配置任务。
- 如双向同步任务的源实例或目标实例位于海外地域，则仅支持同地域的双向同步，不支持跨地域的双向同步。例如，支持日本地域间的双向同步，不支持日本地域与法兰克福地域间的双向同步。
- 源库单机版Redis同步到目标库集群版Redis的操作限制：由于集群cluster只允许单个命令操作单个slot，若在源库执行多Key操作时，Key不在同一个slot或涉及多个slot，则会出现以下报错：

```
CROSSSLOT Keys in request don't hash to the same slot
```

建议在DTS同步过程中仅执行单Key操作，以免导致链路中断。

费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

功能限制

不支持级联单向同步，相关介绍请参见[数据同步拓扑介绍](#)。

## 支持的同步命令

- APPEND
- BITOP、BLPOP、BRPOP、BRPOPLPUSH
- DECR、DECRBY、DEL
- EVAL、EVALSHA、EXEC、EXPIRE、EXPIREAT
- GEOADD、GETSET
- HDEL、HINCRBY、HINCRBYFLOAT、HMSET、HSET、HSETNX
- INCR、INCRBY、INCRBYFLOAT
- LINSERT、LPOP、LPUSH、LPUSHX、LREM、LSET、LTRIM
- MOVE、MSET、MSETNX、MULTI
- PERSIST、PEXPIRE、PEXPIREAT、PFADD、PFMERGE、PSETEX
- RENAME、RENAMENX、RPOP、RPOPLPUSH、RPUSH、RPUSHX
- SADD、SDIFFSTORE、SELECT、SET、SETBIT、SETEX、SETNX、SETRANGE、SINTERSTORE、SMOVE、SPOP、SREM、SUNIONSTORE
- UNLINK、ZADD、ZINCRBY、ZINTERSTORE、ZREM、ZREMRANGEBYLEX、ZUNIONSTORE、ZREMRANGEBYRANK、ZREMRANGEBYSCORE
- SWAPDB（当源或目标实例的架构为集群版时不支持）

### 注意

- 不支持同步PUBLISH命令。
- 对于通过EVAL或者EVALSHA调用Lua脚本，在增量数据同步时，由于目标端在执行脚本时不会明确返回执行结果，DTS无法确保该类型脚本能够执行成功。
- 对于List，由于DTS在调用sync或psync进行重传时，不会对目标端已有的数据进行清空，可能导致出现重复数据。

## 准备工作

1. [创建AccessKey](#)。

 **说明** 为避免主账号泄露AccessKey带来的安全风险，建议您创建RAM用户，授予RAM用户访问DTS资源的权限，再使用RAM用户的AccessKey来调用SDK。详情请参见[账号访问控制](#)。

2. [下载SDK](#)。相关安装方法及介绍，请参见[DTS Java SDK使用说明](#)。

## 创建数据同步作业

本示例中，请求参数和返回参数的详细说明请参见[购买同步实例](#)。

 **说明** `Topology` 参数传入 `oneway`，则创建单向数据同步作业；传入 `bidirectiona` 则创建双向数据同步作业。

请求示例：

```
import com.aliyuncs.DefaultAcsClient;
import com.aliyuncs.IAcsClient;
import com.aliyuncs.exceptions.ClientException;
import com.aliyuncs.exceptions.ServerException;
import com.aliyuncs.profile.DefaultProfile;
import com.google.gson.Gson;
import java.util.*;
import com.aliyuncs.dts.model.v20180801.*;
public class CreateSynchronizationJob {
    public static void main(String[] args) {
        DefaultProfile profile = DefaultProfile.getProfile("cn-hangzhou", "<accessKeyId>", "<accessSecret>");
        IAcsClient client = new DefaultAcsClient(profile);
        CreateSynchronizationJobRequest request = new CreateSynchronizationJobRequest();
        request.setRegionId("cn-hangzhou");
        request.setSourceRegion("cn-beijing");
        request.setDestRegion("cn-beijing");
        request.setSynchronizationJobClass("small");
        request.setPayType("Postpaid");
        request.setTopology("oneway");
        request.setSourceEndpointInstanceType("Redis");
        request.setDestinationEndpointInstanceType("Redis");
        try {
            CreateSynchronizationJobResponse response = client.getAcsResponse(request);
            System.out.println(new Gson().toJson(response));
        } catch (ServerException e) {
            e.printStackTrace();
        } catch (ClientException e) {
            System.out.println("ErrCode:" + e.getErrCode());
            System.out.println("ErrMsg:" + e.getErrMsg());
            System.out.println("RequestId:" + e.getRequestId());
        }
    }
}
```

返回示例:

```
{
  "SynchronizationJobId": "dts*****",
  "RequestId": "158347*****-rqyjQ",
  "Success": true
}
```

## 配置数据同步作业（单向）

本示例中，源实例的数据库账号需具备读权限，目标实例的数据库账号需具备读写权限。

 **说明** 配置单向数据同步时，您也可以将自建Redis数据库或阿里云Redis社区版实例作为源库或目标库，配置方法与本案例类似（需调整对应的请求参数，例如 `MigrationReserved`）。

关于请求参数和返回参数的详细说明请参见[配置同步实例](#)。

请求示例:

```

import com.aliyuncs.DefaultAcsClient;
import com.aliyuncs.IAcsClient;
import com.aliyuncs.exceptions.ClientException;
import com.aliyuncs.exceptions.ServerException;
import com.aliyuncs.profile.DefaultProfile;
import com.google.gson.Gson;
import java.util.*;
import com.aliyuncs.dts.model.v20180801.*;
public class ConfigureSynchronizationJob {
    public static void main(String[] args) {
        DefaultProfile profile = DefaultProfile.getProfile("cn-hangzhou", "<accessKeyId>", "<accessSecret>");
        IAcsClient client = new DefaultAcsClient(profile);
        ConfigureSynchronizationJobRequest request = new ConfigureSynchronizationJobRequest();
        request.setRegionId("cn-hangzhou");
        request.setSynchronizationJobId("dts*****");
        request.setStructureInitialization(true);
        request.setDataInitialization(true);
        request.setSynchronizationObjects("[{\"DBName\": \"\", \"NewDBName\": \"\"}]");
        request.setSynchronizationJobName("apitest");
        request.setSourceEndpointInstanceId("r-2ze*****");
        request.setSourceEndpointInstanceType("redis");
        request.setSourceEndpointPassword("Test*****");
        request.setDestinationEndpointInstanceId("r-2ze*****");
        request.setDestinationEndpointInstanceType("redis");
        request.setDestinationEndpointPassword("Test*****");
        request.setMigrationReserved("{\"srcRedisType\":\"enterprise\", \"destRedisType\":\"enterprise\"}");
    }
    try {
        ConfigureSynchronizationJobResponse response = client.getAcsResponse(request);
        System.out.println(new Gson().toJson(response));
    } catch (ServerException e) {
        e.printStackTrace();
    } catch (ClientException e) {
        System.out.println("ErrCode:" + e.getErrCode());
        System.out.println("ErrMsg:" + e.getErrMsg());
        System.out.println("RequestId:" + e.getRequestId());
    }
}

```

返回示例：

```

{
  "RequestId": "1583*****-Bc05F",
  "Success": true
}

```

## 配置数据同步作业（双向）

本示例中，用于数据同步的数据库账号需具备读写权限。

关于请求参数和返回参数的详细说明请参见[配置同步实例](#)。

请求示例：

1. 配置正向数据同步作业。

```
import com.aliyuncs.DefaultAcsClient;
import com.aliyuncs.IAcsClient;
import com.aliyuncs.exceptions.ClientException;
import com.aliyuncs.exceptions.ServerException;
import com.aliyuncs.profile.DefaultProfile;
import com.google.gson.Gson;
import java.util.*;
import com.aliyuncs.dts.model.v20180801.*;
public class ConfigureSynchronizationJob {
    public static void main(String[] args) {
        DefaultProfile profile = DefaultProfile.getProfile("cn-hangzhou", "<accessKeyId>", "<accessSecret>");
        IAcsClient client = new DefaultAcsClient(profile);
        ConfigureSynchronizationJobRequest request = new ConfigureSynchronizationJobRequest();
        request.setRegionId("cn-hangzhou");
        request.setSynchronizationJobId("dts*****");
        request.setStructureInitialization(true);
        request.setDataInitialization(true);
        request.setSynchronizationObjects("[{\"DBName\": \"0\", \"NewDBName\": \"0\"}]");
        request.setSynchronizationJobName("apitest");
        request.setSynchronizationDirection("Forward");
        request.setSourceEndpointInstanceId("r-2ze*****");
        request.setSourceEndpointInstanceType("redis");
        request.setSourceEndpointPassword("Test*****");
        request.setDestinationEndpointInstanceId("r-2ze*****");
        request.setDestinationEndpointInstanceType("redis");
        request.setDestinationEndpointPassword("Test*****");
        request.setMigrationReserved("{\"srcRedisType\": \"enterprise\", \"destRedisType\": \"enterprise\"}");
        try {
            ConfigureSynchronizationJobResponse response = client.getAcsResponse(request);
            System.out.println(new Gson().toJson(response));
        } catch (ServerException e) {
            e.printStackTrace();
        } catch (ClientException e) {
            System.out.println("ErrCode:" + e.getErrCode());
            System.out.println("ErrMsg:" + e.getErrMsg());
            System.out.println("RequestId:" + e.getRequestId());
        }
    }
}
```

2. 等待正向数据同步作业初始化完成（即转变为同步中状态）。

 **说明** 您可以通过控制台或调用[查询一个数据同步实例状态](#)接口查看数据同步作业的状态信息。

3. 配置反向数据同步作业。

```
import com.aliyuncs.DefaultAcsClient;
import com.aliyuncs.IAcsClient;
import com.aliyuncs.exceptions.ClientException;
import com.aliyuncs.exceptions.ServerException;
import com.aliyuncs.profile.DefaultProfile;
import com.google.gson.Gson;
import java.util.*;
import com.aliyuncs.dts.model.v20180801.*;
public class ConfigureSynchronizationJob {
    public static void main(String[] args) {
        DefaultProfile profile = DefaultProfile.getProfile("cn-hangzhou", "<accessKeyId>", "<accessSecret>");
        IAcsClient client = new DefaultAcsClient(profile);
        ConfigureSynchronizationJobRequest request = new ConfigureSynchronizationJobRequest();
        request.setRegionId("cn-hangzhou");
        request.setSynchronizationJobId("dts*****");
        request.setStructureInitialization(true);
        request.setDataInitialization(true);
        request.setSynchronizationObjects("[{\"DBName\": \"0\", \"NewDBName\": \"0\"}]");
        request.setSynchronizationJobName("apitest");
        request.setSynchronizationDirection("Reverse");
        request.setSourceEndpointInstanceId("r-2ze*****");
        request.setSourceEndpointInstanceType("redis");
        request.setSourceEndpointPassword("Test*****");
        request.setDestinationEndpointInstanceId("r-2ze*****");
        request.setDestinationEndpointInstanceType("redis");
        request.setDestinationEndpointPassword("Test*****");
        request.setMigrationReserved("{\"srcRedisType\": \"enterprise\", \"destRedisType\": \"enterprise\"}");
        try {
            ConfigureSynchronizationJobResponse response = client.getAcsResponse(request);
            System.out.println(new Gson().toJson(response));
        } catch (ServerException e) {
            e.printStackTrace();
        } catch (ClientException e) {
            System.out.println("ErrCode:" + e.getErrCode());
            System.out.println("ErrMsg:" + e.getErrMsg());
            System.out.println("RequestId:" + e.getRequestId());
        }
    }
}
```

返回示例:

```
{
  "RequestId": "1583*****-Bc05F",
  "Success": true
}
```

## 相关文档

- [查看同步链路状态和性能](#)  
通过同步性能展示功能，您可以了解数据同步作业的运行状态，为您管理数据同步作业提供参考依据。
- [新增同步对象/移除同步对象](#)  
在数据同步过程中，您可以新增或移除同步对象。

## 11.4. 从ECS上的自建Redis同步至Redis实例

数据传输服务DTS（Data Transmission Service）支持Redis数据库的单向同步，适用于异地多活、数据异地容灾等多种应用场景。本文以ECS上的自建Redis同步至Redis实例为例，介绍数据同步作业的配置流程。

**警告** 完成数据同步作业的配置后，请勿变更源数据库或目标数据库的**架构类型**（例如将主从架构变更为集群架构），否则会导致数据同步失败。

### 前提条件

- 源Redis数据库的版本为2.8、3.0、3.2、4.0或5.0版本。

**说明** 如需跨版本同步（仅支持从低版本同步到高版本）请提前确认兼容性。例如创建按量付费的Redis实例来测试，测试完成后可将该实例释放或转为包年包月。

- 目标Redis实例的存储空间需大于源Redis数据库已使用的存储空间。
- 当源Redis数据库为集群架构时，集群的每个节点必须能够执行 `psync` 命令，且连接的密码一致。
- 源Redis实例Slave和Master之间的复制超时时间参数`repl-timeout`，默认为60秒，建议使用 `config set repl-timeout 600` 命令设置为600秒。若源数据库数据量比较大，可以适当增大`repl-timeout`参数的值。

### 注意事项

- DTS在执行全量数据初始化时将占用源库和目标库一定的资源，可能会导致数据库服务器负载上升。如果数据库业务量较大或服务器规格较低，可能会加重数据库压力，甚至导致数据库服务不可用。建议您在执行数据同步前谨慎评估，在业务低峰期执行数据同步。
- 如果源库中的某些Key使用了过期（expire）策略，由于可能存在Key已过期但未被及时删除的情况，所以在目标库中查看到的Key数量（例如通过`info`命令查看）会比源库的Key数量少。

**说明** 源和目标库中，未设置过期策略或未过期的Key数量是一致的。

- 如果在源数据库的配置文件`redis.conf`中配置了 `bind` 参数，请将该参数的值设置为ECS的内网IP地址以保障DTS可以正常连接源数据库。
- 为保障同步链路稳定性，建议将源Redis数据库的配置文件`redis.conf`中 `repl-backlog-size` 参数的值适当调大。
- 为保障同步质量，DTS会在源Redis数据库中插入一个key：`DTS_REDIS_TIMESTAMP_HEARTBEAT`，用于记录更新时间点。
- 如果配置Redis集群间的数据同步，请勿在源集群中执行 `FLUSHDB` 和 `FLUSHALL` 命令，否则将导致源和目标的数据不一致。
- 如果目标库的数据逐出策略（`maxmemory-policy`）配置为 `noeviction` 以外的值，可能导致目标库的数据与源库不一致。关于数据逐出策略详情，请参见[Redis数据逐出策略介绍](#)。
- 同步期间，如自建Redis发生扩缩容（如增加或者减少分片）、规格变配（如扩大内存），则您需重新配置任务。且为保障数据一致性，建议重新配置任务前，先清空已同步至目标Redis的数据。
- 同步期间，如自建Redis连接地址变化，则需要[提交工单](#)进行处理，如果长时间未处理，自建Redis会重置AOF文件，届时您需重新配置任务。
- 源库单机版Redis同步到目标库集群版Redis的操作限制：由于集群cluster只允许单个命令操作单个slot，若在源库执行多Key操作时，Key不在同一个slot或涉及多个slot，则会出现以下报错：

```
CROSSSLOT Keys in request don't hash to the same slot
```

建议在DTS同步过程中仅执行单Key操作，以免导致链路中断。

### 费用说明

| 同步类型          | 链路配置费用 |
|---------------|--------|
| 库表结构同步和全量数据同步 | 不收费。   |

| 同步类型   | 链路配置费用                           |
|--------|----------------------------------|
| 增量数据同步 | 收费, 详情请参见 <a href="#">产品定价</a> 。 |

## 支持的同步拓扑

- 一对一单向同步
- 一对多单向同步
- 级联单向同步

关于各类同步拓扑的介绍及注意事项, 请参见[数据同步拓扑介绍](#)。

## 支持的同步命令

- APPEND
- BITOP、BLPOP、BRPOP、BRPOPLPUSH
- DECR、DECRBY、DEL
- EVAL、EVALSHA、EXEC、EXPIRE、EXPIREAT
- GEOADD、GETSET
- HDEL、HINCRBY、HINCRBYFLOAT、HMSET、HSET、HSETNX
- INCR、INCRBY、INCRBYFLOAT
- LINSERT、LPOP、LPUSH、LPUSHX、LREM、LSET、LTRIM
- MOVE、MSET、MSETNX、MULTI
- PERSIST、PEXPIRE、PEXPIREAT、PFADD、PFMERGE、PSETEX
- RENAME、RENAMENX、RESTORE、RPOP、RPOPLPUSH、RPUSH、RPUSHX
- SADD、SDIFFSTORE、SELECT、SET、SETBIT、SETEX、SETNX、SETRANGE、SINTERSTORE、SMOVE、SPOP、SREM、SUNIONSTORE
- ZADD、ZINCRBY、ZINTERSTORE、ZREM、ZREMRANGEBYLEX、ZUNIONSTORE、ZREMRANGEBYRANK、ZREMRANGEBYSCORE
- SWAPDB、UNLINK (仅当源端Redis实例的版本为4.0时支持)

### 说明

- 不支持同步PUBLISH命令。
- 对于通过EVAL或者EVALSHA调用Lua脚本, 在增量数据同步时, 由于目标端在执行脚本时不会明确返回执行结果, DTS无法确保该类型脚本能够执行成功。
- 对于List, 由于DTS在调用sync或psync进行重传时, 不会对目标端已有的数据进行清空, 可能导致出现重复数据。

## 操作步骤

1. 购买数据同步实例, 详情请参见[购买数据同步实例](#)。

**说明** 购买时, 源实例和目标实例均选择为Redis。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏, 单击[数据同步](#)。
4. 在同步作业列表页面顶部, 选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例, 单击[配置同步链路](#)。
6. 配置数据同步的源实例及目标实例信息。

1.选择同步通道的源及目标实例
2.选择同步对象
3.高级设置
4.预检查

同步作业名称：

**源实例信息**

实例类型：

实例地区：

ECS实例ID：

数据库类型：

实例模式： 单机版  集群版

端口：

数据库密码：

**目标实例信息**

实例类型：

实例地区：

实例ID：

数据库密码：

取消 授权白名单并进入下一步

| 类别    | 配置      | 说明  |
|-------|---------|---|
| 无     | 同步作业名称  | DTS会自动生成一个同步作业名称，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。  |
| 源实例信息 | 实例类型    | 选择ECS上的自建数据库。   |
|       | 实例地区    | 购买数据同步实例时选择的源实例地域信息，不可变更。   |
|       | ECS实例ID | 选择作为同步数据源的ECS实例ID。<br><br><span style="background-color: #e1f5fe; padding: 5px; border: 1px solid #cfe2f3;">? 说明 当源Redis数据库为集群架构时，选择任一节点的Master所在的ECS实例ID。</span>             |
|       | 数据库类型   | 固定为Redis。   |
|       | 实例模式    | 根据源Redis数据库的架构选择单机版或集群版。  |
|       | 端口      | 填入源Redis数据库的服务端口，默认为6379，本案例填入7000。<br><br><span style="background-color: #e1f5fe; padding: 5px; border: 1px solid #cfe2f3;">? 说明 当源Redis数据库为集群架构时，填入任一节点的Master的服务端口。</span> |
|       | 数据库密码   | 填入连接源Redis数据库的密码。<br><br><span style="background-color: #e1f5fe; padding: 5px; border: 1px solid #cfe2f3;">? 说明 非必填项，如果没有设置密码可以不填。</span>                                     |

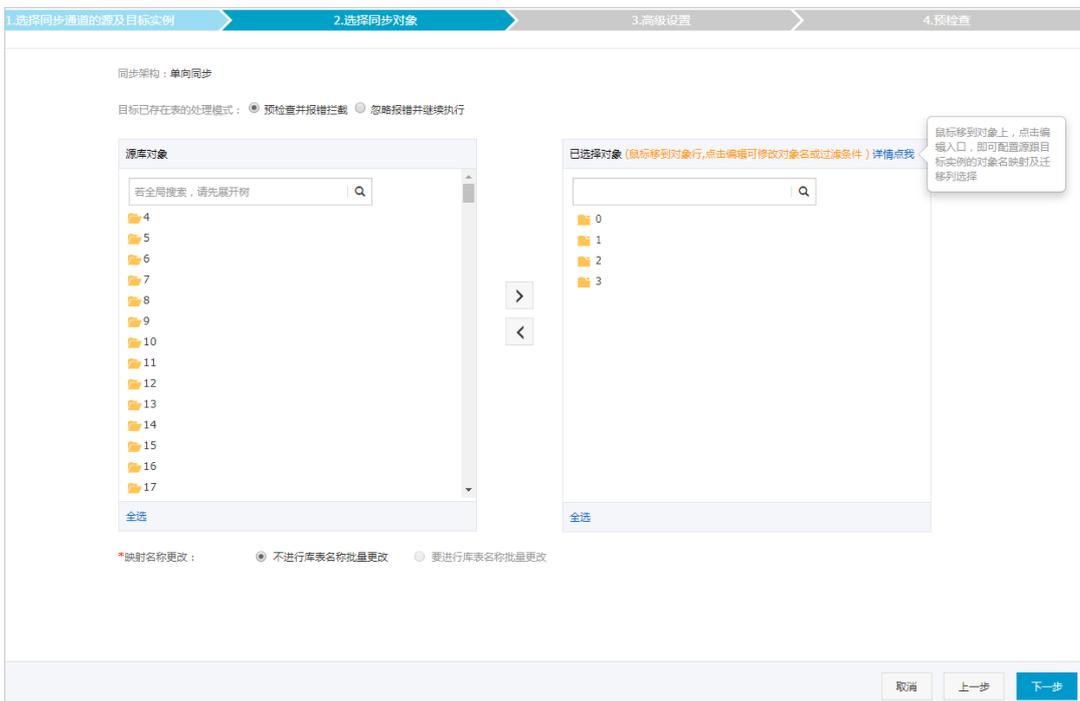
| 类别     | 配置    | 说明  |
|--------|-------|---|
| 目标实例信息 | 实例类型  | 选择Redis实例。  |
|        | 实例地区  | 购买数据同步实例时选择的目标实例地域信息，不可变更。  |
|        | 实例ID  | 选择目标Redis实例ID。  |
|        | 数据库密码 | 填入连接目标Redis实例的密码。<br><br><div style="border: 1px solid #ccc; padding: 5px; background-color: #e6f2ff;"> <p><span style="color: #00aaff;">?</span> 说明 数据库密码格式为&lt;user&gt;:&lt;password&gt;。例如，Redis实例自定义的用户名为admin，密码为Rp829dlwa，则此处填入的数据库密码为admin:Rp829dlwa。</p> </div> |

7. 单击页面右下角的授权白名单并进入下一步。

? 说明

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置目标已存在表的处理模式和同步对象。



| 配置 | 说明 |
|----|----|
|----|----|

| 配置                             | 说明   |
|--------------------------------|--|
| 目标已存在表的处理模式                    | <ul style="list-style-type: none"> <li>预检查并报错拦截：检查目标库是否为空。如果待同步的目标库为空，则通过该检查项目；如果不为空，则在预检查阶段提示错误，数据同步作业不会被启动。</li> <li>忽略报错并继续执行：跳过目标库是否为空的检查项。</li> </ul> <p> <b>警告</b> 选择为忽略报错并继续执行后，如果在同步过程中遇到目标库中的Key与源库中的Key相同，会将源库的数据覆盖写入目标库中，请谨慎选择。</p>  |
| 同步对象                           | <ul style="list-style-type: none"> <li>在源库对象框中单击待同步的数据库，然后单击  将其移动到已选择对象框。</li> <li>同步对象的选择粒度为库，暂不支持Key粒度的选择。</li> </ul>  |
| 映射名称更改                         | 不支持更改映射名称。   |
| 源表DMS_ONLINE_DDL过程中是否复制临时表到目标库 | <p>如源库使用数据管理DMS (Data Management Service) 执行Online DDL变更，您可以选择是否同步Online DDL变更产生的临时表数据。</p> <ul style="list-style-type: none"> <li>是：同步Online DDL变更产生的临时表数据。</li> </ul> <p> <b>说明</b> Online DDL变更产生的临时表数据过大，可能会导致同步任务延迟。</p> <ul style="list-style-type: none"> <li>否：不同步Online DDL变更产生的临时表数据，只同步源库的原始DDL数据。</li> </ul> <p> <b>说明</b> 该方案会导致目标库锁表。</p> |
| 源、目标库无法连接重试时间                  | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <p> <b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p>  |

9. 上述配置完成后单击页面右下角的下一步。

10. 配置同步初始化选项，当前固定为包含全量数据+增量数据。

1.选择同步通道的源及目标实例
2.选择同步对象
3.高级设置
4.预检查

同步初始化：包含全量数据+增量数据

取消 上一步 保存 预检查并启动

 **说明**

- DTS会将源Redis实例中的存量数据同步至目标Redis实例中，并同步增量数据。
- 如果跳出版本相关的报错提示，请您按照提示将源Redis实例升级至指定版本。关于升级版本方式，请参见[升级大版本](#)和[升级小版本](#)。

11. 上述配置完成后，单击页面右下角的**预检查并启动**。

**说明**

- 在同步任务正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步任务。
- 如果预检查失败，单击具体检查项后的 ，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择**确认屏蔽**、**忽略告警项**并重新进行预检查，跳过告警检测项重新进行预检查。

12. 在**预检查**对话框中显示**预检查通过**后，关闭**预检查**对话框，同步作业将正式开始。

13. 等待同步作业的链路初始化完成，直至处于**同步中**状态。



| 实例ID/作业名称               | 状态   | 同步概况                                   | 付费方式 | 同步架构(全部) | 操作              |
|-------------------------|------|--|------|----------|-----------------|
| hangzhou-hangzhou-small | 同步中  | 延时: 1376 毫秒<br>速度: 0.00RPS/(0.000MB/s) | 按量付费 | 单向同步     | 暂停同步 转包年包月 升级更多 |
| 暂停同步                    | 释放同步 |  |      |          |                 |

共有1条，每页显示：20条

**说明** 您可以在数据同步页面，查看数据同步作业的状态。

## 11.5. 从通过专线、VPN网关或智能接入网关接入的自建Redis同步至ECS上的自建Redis

数据传输服务DTS（Data Transmission Service）支持Redis数据库的单向同步，适用于异地多活、数据异地容灾等多种应用场景。本文以通过专线、VPN网关或智能接入网关接入的自建Redis同步至ECS上的自建Redis为例，介绍数据同步作业的配置流程。

**警告** 完成数据同步作业的配置后，请勿变更源数据库或目标数据库的**架构类型**（例如将主从架构变更为集群架构），否则会导致数据同步失败。

### 前提条件

- 源Redis数据库的版本为2.8、3.0、3.2、4.0或5.0版本。

**说明** 目标Redis数据库支持的版本为2.8、3.0、3.2、4.0或5.0版本，如需跨版本同步（仅支持从低版本同步到高版本）请提前确认兼容性。

- 目标Redis数据库的存储空间需大于源Redis数据库已使用的存储空间。
- 当源Redis数据库为集群架构时，集群的每个节点必须能够执行 `psync` 命令，且连接的密码一致。
- 源Redis实例Slave和Master之间的复制超时时间参数`repl-timeout`，默认为60秒，建议使用 `config set repl-timeout 600` 命令设置为600秒。若源数据库数据量比较大，可以适当增大`repl-timeout`参数的值。

### 注意事项

- DTS在执行全量数据初始化时将占用源库和目标库一定的资源，可能会导致数据库服务器负载上升。如果数据库业务量较大或服务器规格较低，可能会加重数据库压力，甚至导致数据库服务不可用。建议您在执行数据同步前谨慎评估，在业务低峰期执行数据同步。
- 如果源库中的某些Key使用了过期（expire）策略，由于可能存在Key已过期但未被及时删除的情况，所以在目标库中看到的Key数量（例如通过`info`命令查看）会比源库的Key数量少。

 **说明** 源和目标库中，未设置过期策略或未过期的Key数量是一致的。

- 如果在源数据库的配置文件`redis.conf`中配置了 `bind` 参数，请将该参数的值设置为ECS的内网IP地址以保障DTS可以正常连接源数据库。
- 为保障同步链路稳定性，建议将源Redis数据库的配置文件`redis.conf`中 `repl-backlog-size` 参数的值适当调大。
- 为保障同步质量，DTS会在源Redis数据库中插入一个key：`DTS_REDIS_TIMESTAMP_HEARTBEAT`，用于记录更新时间点。
- 如果配置Redis集群间的数据同步，请勿在源集群中执行 `FLUSHDB` 和 `FLUSHALL` 命令，否则将导致源和目标的数据不一致。
- 如果目标库的数据逐出策略（`maxmemory-policy`）配置为 `noeviction` 以外的值，可能导致目标库的数据与源库不一致。关于数据逐出策略详情，请参见[Redis数据逐出策略介绍](#)。
- 同步期间，如自建Redis发生扩缩容（如增加或者减少分片）、规格变配（如扩大内存），则您需重新配置任务。且为保障数据一致性，建议重新配置任务前，先清空已同步至目标Redis的数据。
- 同步期间，如自建Redis连接地址变化，则您需要[提交工单](#)进行处理，如果长时间未处理，自建Redis会重置AOF文件，届时您需重新配置任务。
- 源库单机版Redis同步到目标库集群版Redis的操作限制：由于集群cluster只允许单个命令操作单个slot，若在源库执行多Key操作时，Key不在同一个slot或涉及多个slot，则会出现以下报错：

```
CROSSSLOT Keys in request don't hash to the same slot
```

建议在DTS同步过程中仅执行单Key操作，以免导致链路中断。

## 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

## 支持的同步拓扑

- 一对一单向同步
- 一对多单向同步
- 级联单向同步

关于各类同步拓扑的介绍及注意事项，请参见[数据同步拓扑介绍](#)。

## 支持的同步命令

- APPEND
- BITOP、BLPOP、BRPOP、BRPOPLPUSH
- DECR、DECRBY、DEL
- EVAL、EVALSHA、EXEC、EXPIRE、EXPIREAT
- GEOADD、GETSET
- HDEL、HINCRBY、HINCRBYFLOAT、HMSET、HSET、HSETNX
- INCR、INCRBY、INCRBYFLOAT
- LINSERT、LPOP、LPUSH、LPUSHX、LREM、LSET、LTRIM
- MOVE、MSET、MSETNX、MULTI
- PERSIST、PEXPIRE、PEXPIREAT、PFADD、PFMERGE、PSETEX
- RENAME、RENAMENX、RESTORE、RPOP、RPOPLPUSH、RPUSH、RPUSHX
- SADD、SDIFFSTORE、SELECT、SET、SETBIT、SETEX、SETNX、SETRANGE、SINTERSTORE、SMOVE、SPOP、SREM、SUNIONSTORE

- ZADD、ZINCRBY、ZINTERSTORE、ZREM、ZREMRANGEBYLEX、ZUNIONSTORE、ZREMRANGEBYRANK、ZREMRANGEBYSCORE
- SWAPDB、UNLINK（仅当源端Redis实例的版本为4.0时支持）

说明

- 不支持同步PUBLISH命令。
- 对于通过EVAL或者EVALSHA调用Lua脚本，在增量数据同步时，由于目标端在执行脚本时不会明确返回执行结果，DTS无法确保该类型脚本能够执行成功。
- 对于List，由于DTS在调用sync或psync进行重传时，不会对目标端已有的数据进行清空，可能导致出现重复数据。

操作步骤

1. 购买数据同步实例，详情请参见[购买数据同步实例](#)。

说明 购买时，源实例和目标实例均选择为Redis。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击数据同步。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击配置同步链路。
6. 配置数据同步的源实例及目标实例信息。

| 类别 | 配置     | 说明   |
|----|--------|--|
| 无  | 同步作业名称 | DTS会自动生成一个同步作业名称，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。 |

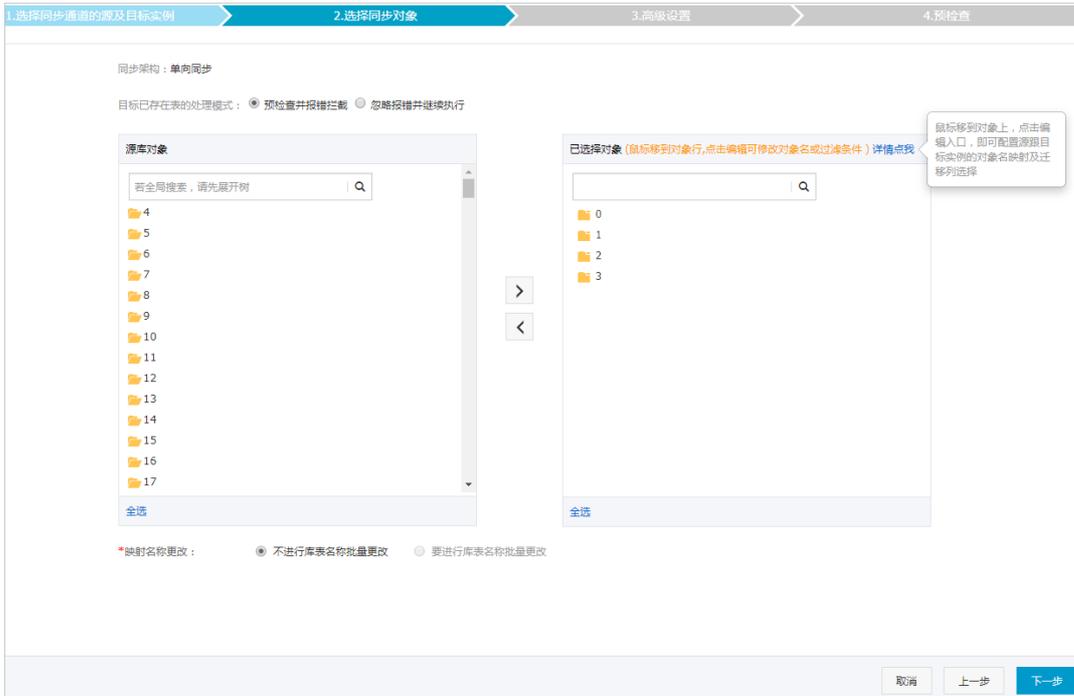
| 类别     | 配置  | 说明   |
|--------|---|--|
| 源实例信息  | 实例类型  | 选择通过专线/VPN网关/智能接入网关接入的自建数据库。   |
|        | 实例地区  | 购买数据同步实例时选择的源实例地域信息，不可变更。  |
|        | 对端专有网络  | 选择自建数据库接入的VPC ID。  |
|        | 数据库类型   | 固定为Redis。  |
|        | 实例模式  | 根据源Redis数据库的架构选择 <b>单机版</b> 或 <b>集群版</b> 。   |
|        | IP地址  | 填入源Redis数据库的IP地址。<br><br><span>❓ 说明</span> 当源Redis数据库为集群架构时，填入任一节点的Master所属服务器的IP地址。               |
|        | 端口  | 填入源Redis数据库的服务端口，默认为 <b>6379</b> 。<br><br><span>❓ 说明</span> 当源Redis数据库为集群架构时，填入任一节点的Master的服务端口。   |
| 数据库密码  | 填入连接源Redis数据库的密码。<br><br><span>❓ 说明</span> 非必填项，如果没有设置密码可以不填。 |  |
| 目标实例信息 | 实例类型  | 选择ECS上的 <b>自建数据库</b> 。   |
|        | 实例地区  | 购买数据同步实例时选择的目标实例地域信息，不可变更。   |
|        | 实例ID  | 选择作为同步目标的ECS实例ID。<br><br><span>❓ 说明</span> 当目标Redis数据库为集群架构时，选择任一节点的Master所在的ECS实例ID。              |
|        | 实例模式  | 根据目标Redis数据库的架构选择 <b>单机版</b> 或 <b>集群版</b> 。  |
|        | 端口  | 填入目标Redis数据库的服务端口，默认为 <b>6379</b> 。<br><br><span>❓ 说明</span> 当目标Redis数据库为集群架构时，填入任一节点的Master的服务端口。 |
|        | 数据库密码   | 填入连接目标Redis数据库的密码。<br><br><span>❓ 说明</span> 非必填项，如果没有设置密码可以不填。                                     |

7. 单击页面右下角的**授权白名单**并进入下一步。

说明

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置目标已存在表的处理模式和同步对象。



| 配置          | 说明  |
|-------------|---|
| 目标已存在表的处理模式 | <ul style="list-style-type: none"> <li><b>预检查并报错拦截</b>：检查目标库是否为空。如果待同步的目标库为空，则通过该检查项目；如果不为空，则在预检查阶段提示错误，数据同步作业不会被启动。</li> <li><b>忽略报错并继续执行</b>：跳过目标库是否为空的检查项。</li> </ul> <p><b>警告</b> 选择为忽略报错并继续执行后，如果在同步过程中遇到目标库中的Key与源库中的Key相同，会将源库的数据覆盖写入目标库中，请谨慎选择。</p> |
| 同步对象        | <ul style="list-style-type: none"> <li>在源库对象框中单击待同步的数据库，然后单击 <b>&gt;</b> 将其移动到已选择对象框。</li> <li>同步对象的选择粒度为库，暂不支持Key粒度的选择。</li> </ul>   |
| 映射名称更改      | 不支持更改映射名称。  |

| 配置                             | 说明   |
|--------------------------------|--|
| 源表DMS_ONLINE_DDL过程中是否复制临时表到目标库 | <p>如源库使用<a href="#">数据管理DMS (Data Management Service)</a> 执行Online DDL变更, 您可以选择是否同步Online DDL变更产生的临时表数据。</p> <ul style="list-style-type: none"> <li>是: 同步Online DDL变更产生的临时表数据。</li> </ul> <p><b>说明</b> Online DDL变更产生的临时表数据过大, 可能会导致同步任务延迟。</p> <ul style="list-style-type: none"> <li>否: 不同步Online DDL变更产生的临时表数据, 只同步源库的原始DDL数据。</li> </ul> <p><b>说明</b> 该方案会导致目标库锁表。</p> |
| 源、目标库无法连接重试时间                  | <p>当源、目标库无法连接时, DTS默认重试720分钟(即12小时), 您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库, 同步任务将自动恢复。否则, 同步任务将失败。</p> <p><b>说明</b> 由于连接重试期间, DTS将收取任务运行费用, 建议您根据业务需要自定义重试时间, 或者在源和目标库实例释放后尽快释放DTS实例。</p>   |

- 上述配置完成后单击页面右下角的下一步。
- 配置同步初始化选项, 当前固定为包含**全量数据+增量数据**。

1.选择同步通道的源及目标实例
2.选择同步对象
3.高级设置
4.预检查

同步初始化: 包含全量数据+增量数据

取消 上一步 保存 预检查并启动

**说明**

- DTS会将源Redis实例中的存量数据同步至目标Redis实例中, 并同步增量数据。
- 如果跳出版本相关的报错提示, 请您按照提示将源Redis实例升级至指定版本。关于升级版本方式, 请参见[升级大版本](#)和[升级小版本](#)。

- 上述配置完成后, 单击页面右下角的**预检查并启动**。

**说明**

- 在同步任务正式启动之前, 会先进行预检查。只有预检查通过后, 才能成功启动同步任务。
- 如果预检查失败, 单击具体检查项后的, 查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项, 您也可以选择**确认屏蔽**、**忽略告警项**并重新进行预检查, 跳过告警检测项重新进行预检查。

- 在**预检查**对话框中显示**预检查通过**后, 关闭**预检查**对话框, 同步作业将正式开始。
- 等待同步作业的链路初始化完成, 直至处于**同步中**状态。

| 实例ID/作业名称               | 状态  | 同步概况                                   | 付费方式 | 同步架构(全部) | 操作              |
|-------------------------|-----|--|------|----------|-----------------|
| hangzhou-hangzhou-small | 同步中 | 耗时: 1376 毫秒<br>速度: 0.00RPS/(0.000MB/s) | 按量付费 | 单向同步     | 暂停同步 转包年包月 升级更多 |

共有1条, 每页显示: 20条

**说明** 您可以在数据同步页面, 查看数据同步作业的状态。

## 11.6. 从自建Redis集群同步至Redis集群实例

数据传输服务DTS (Data Transmission Service) 支持Redis集群间的单向同步, 适用于数据迁移、异地多活、数据异地容灾等多种应用场景。本文以自建Redis集群同步至阿里云Redis集群实例为例, 介绍数据同步作业的配置流程。

从阿里云Redis集群实例同步至自建Redis集群的操作步骤与本文类似, 您需要根据实际的业务场景配置同步的源和目标实例信息。

**警告** 完成数据同步作业的配置后, 请勿变更源数据库或目标数据库的**架构类型** (例如将主从架构变更为集群架构), 否则会导致数据同步失败。

### 前提条件

- 自建Redis数据库的版本为2.8、3.0、3.2、4.0或5.0版本。

**说明** 目标阿里云Redis集群实例支持的版本为2.8、4.0或5.0版本, 如需跨版本同步 (仅支持从低版本同步到高版本), 请提前确认兼容性。例如创建按量付费的Redis集群实例来测试, 测试完成后可将该实例释放或转为包年包月。

- 目标阿里云Redis集群实例的存储空间需大于源Redis数据库已使用的存储空间。
- 源Redis集群的每个节点必须能够执行 `psync` 命令, 且连接的密码一致。
- 源Redis实例Slave和Master之间的复制超时时间参数`repl-timeout`, 默认为60秒, 建议使用 `config set repl-timeout 600` 命令设置为600秒。若源数据库数据量比较大, 可以适当增大`repl-timeout`参数的值。

### 注意事项

- DTS在执行全量数据初始化时将占用源库和目标库一定的资源, 可能会导致数据库服务器负载上升。如果数据库业务量较大或服务器规格较低, 可能会加重数据库压力, 甚至导致数据库服务不可用。建议您在执行数据同步前谨慎评估, 在业务低峰期执行数据同步。
- 为保障同步链路稳定性, 建议将配置文件`redis.conf`中 `repl-backlog-size` 参数的值适当调大。
- 为保障同步质量, DTS会在源Redis数据库中插入一个key: `DTS_REDIS_TIMESTAMP_HEARTBEAT`, 用于记录更新时间点。
- 请勿在源集群中执行 `FLUSHDB` 和 `FLUSHALL` 命令, 否则将导致源和目标的数据不一致。
- 如果目标库的数据逐出策略 (`maxmemory-policy`) 配置为 `noeviction` 以外的值, 可能导致目标库的数据与源库不一致。关于数据逐出策略详情, 请参见[Redis数据逐出策略介绍](#)。
- 如果源库中的某些Key使用了过期 (`expire`) 策略, 由于可能存在Key已过期但未被及时删除的情况, 所以在目标库中看到的Key数量 (例如通过`info`命令查看) 会比源库的Key数量少。

**说明** 源和目标库中, 未设置过期策略或未过期的Key数量是一致的。

- 同步期间, 如自建Redis发生扩缩容 (如增加或者减少分片)、规格变配 (如扩大内存), 则您需重新配置任务。且为保障数据一致性, 建议重新配置任务前, 先清空已同步至目标Redis的数据。
- 同步期间, 如自建Redis连接地址变化, 则您需要[提交工单](#)进行处理, 如果长时间未处理, 自建Redis会重置AOF文件, 届时您需重新配置任务。
- 源库单机版Redis同步到目标库集群版Redis的操作限制: 由于集群cluster只允许单个命令操作单个slot, 若在源库执行

多Key操作时，Key不在同一个slot或涉及多个slot，则会出现以下报错：

```
CROSSSLOT Keys in request don't hash to the same slot
```

建议在DTS同步过程中仅执行单Key操作，以免导致链路中断。

## 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

## 支持的同步拓扑

- 一对一单向同步
- 一对多单向同步
- 级联单向同步

关于各类同步拓扑的介绍及注意事项，请参见[数据同步拓扑介绍](#)。

## 支持的同步命令

- APPEND
- BITOP、BLPOP、BRPOP、BRPOPLPUSH
- DECR、DECRBY、DEL
- EVAL、EVALSHA、EXEC、EXPIRE、EXPIREAT
- GEOADD、GETSET
- HDEL、HINCRBY、HINCRBYFLOAT、HMSET、HSET、HSETNX
- INCR、INCRBY、INCRBYFLOAT
- LINSERT、LPOP、LPUSH、LPUSHX、LREM、LSET、LTRIM
- MOVE、MSET、MSETNX、MULTI
- PERSIST、PEXPIRE、PEXPIREAT、PFADD、PFMERGE、PSETEX
- RENAME、RENAMENX、RESTORE、RPOP、RPOPLPUSH、RPUSH、RPUSHX
- SADD、SDIFFSTORE、SELECT、SET、SETBIT、SETEX、SETNX、SETRANGE、SINTERSTORE、SMOVE、SPOP、SREM、SUNIONSTORE
- ZADD、ZINCRBY、ZINTERSTORE、ZREM、ZREMRANGEBYLEX、ZUNIONSTORE、ZREMRANGEBYRANK、ZREMRANGEBYSCORE
- SWAPDB、UNLINK（仅当源端Redis实例的版本为4.0时支持）

### 说明

- 不支持同步PUBLISH命令。
- 对于通过EVAL或者EVALSHA调用Lua脚本，在增量数据同步时，由于目标端在执行脚本时不会明确返回执行结果，DTS无法确保该类型脚本能够执行成功。
- 对于List，由于DTS在调用sync或psync进行重传时，不会对目标端已有的数据进行清空，可能导致出现重复数据。

## 操作步骤

1. [购买数据同步实例](#)。

**说明** 购买时，源实例和目标实例均选择为Redis。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击[数据同步](#)。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击[配置同步链路](#)。
6. 配置数据同步的源实例及目标实例信息。

1. 选择同步通道的源及目标实例
2. 选择同步对象
3. 高级设置
4. 预检查

同步作业名称：

**源实例信息**

实例类型：

实例地区：

\* ECS实例ID：

数据库类型：

实例模式： 单机版  集群版

\* 端口：

数据库密码：

**目标实例信息**

实例类型：

实例地区：

\* 实例ID：

数据库密码：

取消
授权白名单并进入下一步

| 类别    | 配置      | 说明   |
|-------|---------|--|
| 无     | 同步作业名称  | DTS会自动生成一个同步作业名称，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。   |
| 源实例信息 | 实例类型    | 选择ECS上的自建数据库。您可以根据源库部署位置，选择ECS上的自建数据库或通过专线/VPN网关/智能网关接入的自建数据库。本文以ECS上的自建数据库为例介绍配置流程，当自建Redis数据库为其他实例类型时，配置流程与该案例类似。  |
|       | 实例地区    | 购买数据同步实例时选择的源实例地域，不可变更。  |
|       | ECS实例ID | 选择自建Redis集群中任一节点的Master所在的ECS实例ID。   |
|       | 数据库类型   | 固定为Redis。  |
|       | 实例模式    | 选择集群版。   |
|       | 端口      | 填入自建Redis集群中任一节点的Master的服务端口，本案例填入7000。  |
|       | 数据库密码   | 填入连接自建Redis数据库的密码。<br><div style="background-color: #e1f5fe; padding: 5px; border: 1px solid #cfe2f3; margin-top: 5px;"> <span style="font-size: 1.2em; color: #0070c0;">?</span> 说明 非必填项，如果没有设置密码可以不填。                 </div> |
|       | 实例类型    | 选择Redis实例。   |

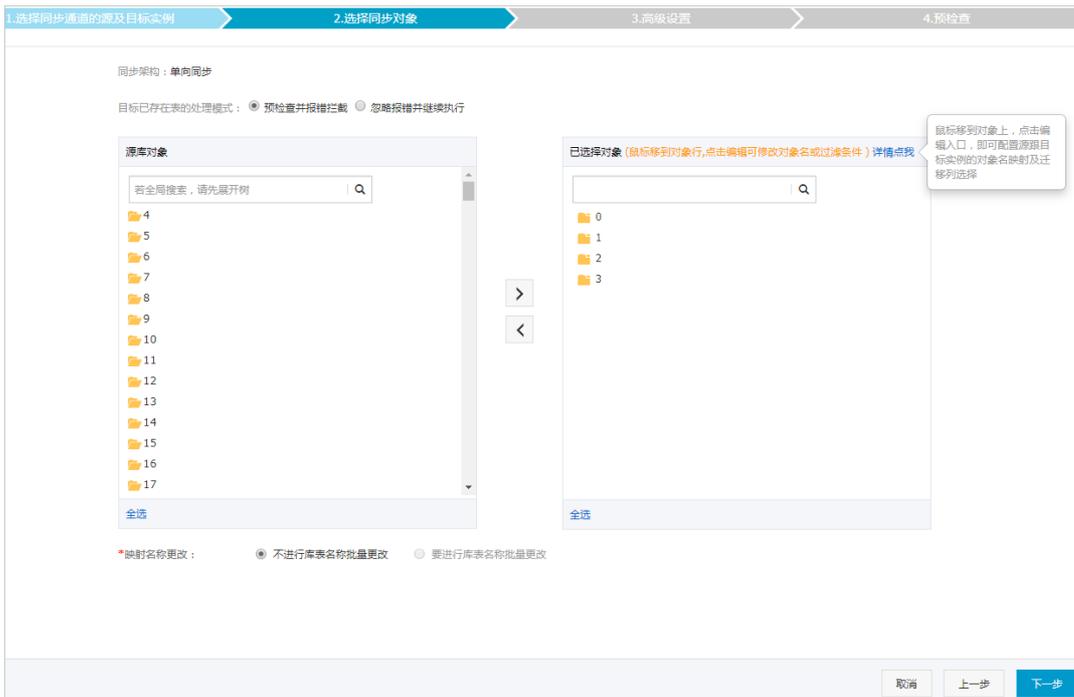
| 类别     | 配置    | 说明   |
|--------|-------|--|
| 目标实例信息 | 实例地区  | 购买数据同步实例时选择的目标实例地域，不可变更。   |
|        | 实例ID  | 选择目标阿里云Redis集群实例ID。  |
|        | 数据库密码 | 填入目标阿里云Redis集群实例的密码。<br><br><div style="border: 1px solid #ccc; padding: 5px; background-color: #e6f2ff;"> <p><b>说明</b> 数据库密码格式为&lt;user&gt;:&lt;password&gt;。例如，Redis实例自定义的用户名为admin，密码为Rp829dlwa，则此处填入的数据库密码为admin:Rp829dlwa。</p> </div> |

7. 单击页面右下角的授权白名单并进入下一步。

**说明**

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置目标已存在表的处理模式和同步对象。



| 配置 | 说明 |
|----|----|
|----|----|

| 配置                             | 说明   |
|--------------------------------|--|
| 目标已存在表的处理模式                    | <ul style="list-style-type: none"> <li>◦ <b>预检查并报错拦截</b>：检查目标库是否为空。如果待同步的目标库为空，则通过该检查项目；如果不为空，则在预检查阶段提示错误，数据同步作业不会被启动。</li> <li>◦ <b>忽略报错并继续执行</b>：跳过目标库是否为空的检查项。</li> </ul> <div style="border: 1px solid #ccc; background-color: #fff9c4; padding: 5px; margin-top: 10px;"> <p> <b>警告</b> 选择为忽略报错并继续执行后，如果在同步过程中遇到目标库中的Key与源库中的Key相同，会将源库的数据覆盖写入目标库中，请谨慎选择。</p> </div>   |
| 同步对象                           | <ul style="list-style-type: none"> <li>◦ 在源库对象框中单击待同步的数据库，然后单击  将其移动到已选择对象框。</li> <li>◦ 同步对象的选择粒度为库，暂不支持Key粒度的选择。</li> </ul>   |
| 映射名称更改                         | 不支持更改映射名称。   |
| 源表DMS_ONLINE_DDL过程中是否复制临时表到目标库 | <p>如源库使用<b>数据管理DMS (Data Management Service)</b> 执行Online DDL变更，您可以选择是否同步Online DDL变更产生的临时表数据。</p> <ul style="list-style-type: none"> <li>◦ <b>是</b>：同步Online DDL变更产生的临时表数据。</li> </ul> <div style="border: 1px solid #ccc; background-color: #e1f5fe; padding: 5px; margin-top: 10px;"> <p> <b>说明</b> Online DDL变更产生的临时表数据过大，可能会导致同步任务延迟。</p> </div> <ul style="list-style-type: none"> <li>◦ <b>否</b>：不同步Online DDL变更产生的临时表数据，只同步源库的原始DDL数据。</li> </ul> <div style="border: 1px solid #ccc; background-color: #e1f5fe; padding: 5px; margin-top: 10px;"> <p> <b>说明</b> 该方案会导致目标库锁表。</p> </div> |
| 源、目标库无法连接重试时间                  | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <div style="border: 1px solid #ccc; background-color: #e1f5fe; padding: 5px; margin-top: 10px;"> <p> <b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p> </div>   |

- 上述配置完成后单击页面右下角的下一步。
- 配置同步初始化选项，当前固定为**包含全量数据+增量数据**。

1.选择同步通道的源及目标实例
2.选择同步对象
3.高级设置
4.预检查

同步初始化：包含全量数据+增量数据

取消 上一步 保存 预检查并启动

**说明**

- DTS会将源Redis实例中的存量数据同步至目标Redis实例中，并同步增量数据。
- 如果跳出版本相关的报错提示，请您按照提示将源Redis实例升级至指定版本。关于升级版本方式，请参见[升级大版本](#)和[升级小版本](#)。

- 上述配置完成后，单击页面右下角的**预检查并启动**。

### 说明

- 在同步任务正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步任务。
- 如果预检查失败，单击具体检查项后的，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择**确认屏蔽**、**忽略告警项**并重新进行预检查，跳过告警检测项重新进行预检查。

12. 在预检查对话框中显示预检查通过后，关闭预检查对话框，同步作业将正式开始。

13. 等待同步作业的链路初始化完成，直至处于同步中状态。



| 实例ID/作业名称               | 状态  | 同步概况                                   | 付费方式 | 同步架构(全部) | 操作                  |
|-------------------------|-----|--|------|----------|---------------------|
| hangzhou-hangzhou-small | 同步中 | 延迟: 1376 毫秒<br>速度: 0.00RPS/(0.000MB/s) | 按量付费 | 单向同步     | 暂停同步   转包年包月   升级更多 |

共有1条，每页显示：20条

说明 您可以在数据同步页面，查看数据同步作业的状态。

## 11.7. 从ECS上的Codis集群同步至Redis实例

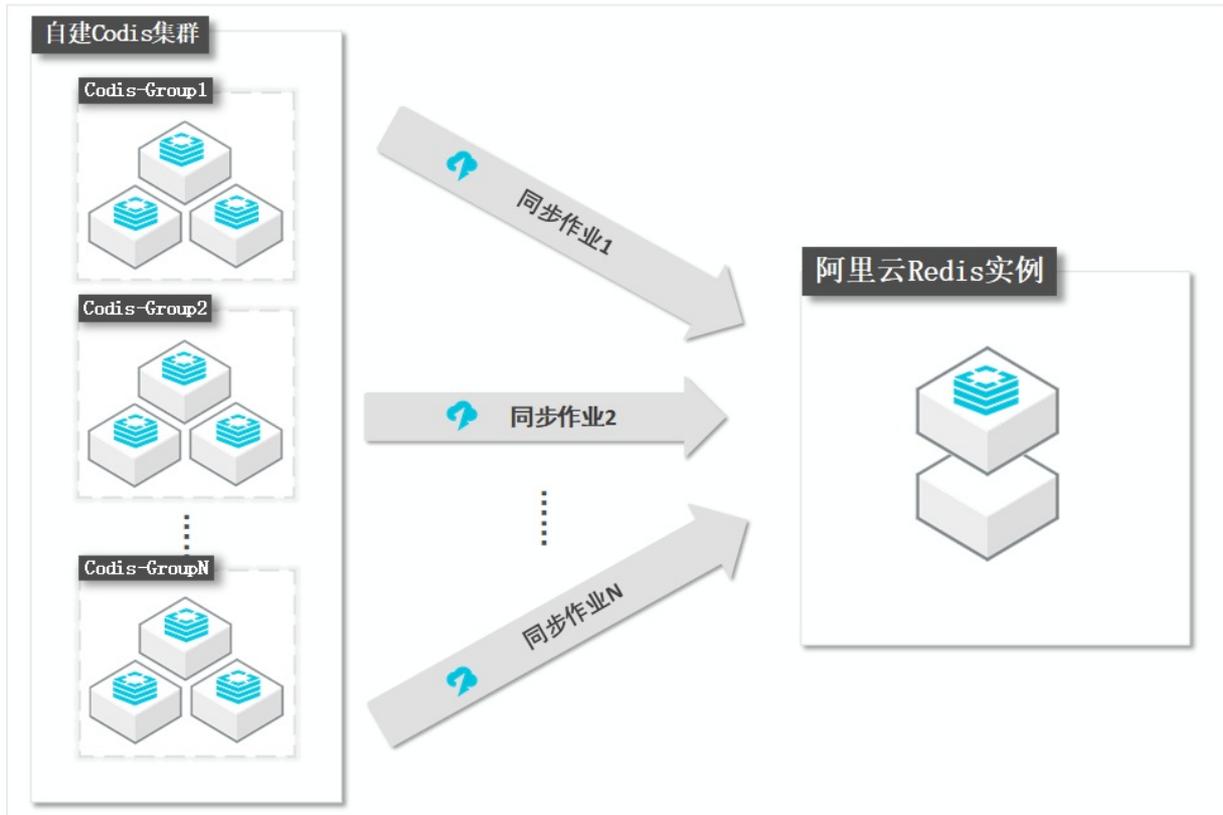
阿里云数据库Redis版是兼容开源Redis协议标准、提供内存加硬盘混合存储的数据库服务，基于高可靠双机热备架构及可平滑扩展的集群架构，可充分满足高吞吐、低延迟及弹性变配的业务需求。通过数据传输服务DTS（Data Transmission Service），您可以将Codis集群同步至阿里云Redis实例。

### 前提条件

- 已创建目标阿里云Redis实例（2.8、4.0或5.0版本），详情请参见[创建Redis实例](#)。
- 目标阿里云Redis实例的空闲存储空间需大于源Codis数据库已使用的存储空间。
- 源Codis集群的每个Master节点必须能够执行 `psync` 命令。
- 源Redis实例Slave和Master之间的复制超时时间参数`repl-timeout`，默认为60秒，建议使用 `config set repl-timeout 600` 命令设置为600秒。若源数据库数据量比较大，可以适当增大`repl-timeout`参数的值。

### 同步原理介绍

DTS通过同步Codis集群中的每个Codis-Group来实现集群的整体同步，您需要为每个Codis-Group创建一个对应的数据同步作业。



### Codis集群环境介绍

在本案例中，Codis集群具备两个Codis-Group，每个Codis-Group采用一主一备的架构，详情请参见下图。

| GROUPS: SYNC ALL |                  | REPLICA(S): ENABLE ALL |                   | REPLICA(S): DISABLE ALL  |        |  |
|------------------|------------------|------------------------|-------------------|--------------------------|--------|--|
| 1                | Server           | Data Center            | Master            |                          | Memory | Keys   |
| SYNC             | S 127.0.0.1:6379 |                        | NO:ONE            | <input type="checkbox"/> | syncd  | 38.05 MB / INF. db0:keys=31583,expires=0,avg_ttl=0 |
| PROMOTE          | S 127.0.0.1:6380 |                        | 127.0.0.1:6379:up | <input type="checkbox"/> | syncd  | 36.61 MB / INF. db0:keys=31583,expires=0,avg_ttl=0 |
| 2                | Server           | Data Center            | Master            |                          | Memory | Keys   |
| SYNC             | S 127.0.0.1:6389 |                        | NO:ONE            | <input type="checkbox"/> | syncd  | 38.10 MB / INF. db0:keys=31636,expires=0,avg_ttl=0 |
| PROMOTE          | S 127.0.0.1:6390 |                        | 127.0.0.1:6389:up | <input type="checkbox"/> | syncd  | 36.67 MB / INF. db0:keys=31636,expires=0,avg_ttl=0 |

### 注意事项

- DTS在执行全量数据初始化时将占用源库和目标库一定的资源，可能会导致数据库服务器负载上升。如果数据库业务量较大或服务器规格较低，可能会加重数据库压力，甚至导致数据库服务不可用。建议您在执行数据同步前谨慎评估，在业务低峰期执行数据同步。
- 如果在源数据库的配置文件redis.conf中配置了 bind 参数，请将该参数的值设置为ECS的内网IP地址以保障DTS可以正常连接源数据库。
- 为保障同步链路稳定性，建议将源集群的配置文件redis.conf中 repl-backlog-size 参数的值适当调大。
- 为保障同步质量，DTS会在源集群中插入一个key： DTS\_REDIS\_TIMESTAMP\_HEARTBEAT ，用于记录更新时间点。
- 请勿在源集群中执行 FLUSHDB 和 FLUSHALL 命令，否则将导致源和目标的数据不一致。
- 如果目标库的数据逐出策略（ maxmemory-policy ）配置为 noeviction 以外的值，可能导致目标库的数据与源库不一致。关于数据逐出策略详情，请参见Redis数据逐出策略介绍。
- 如果源库中的某些Key使用了过期（expire）策略，由于可能存在Key已过期但未被及时删除的情况，所以在目标库中查看到的Key数量（例如通过info命令查看）会比源库的Key数量少。

 **说明** 源和目标库中，未设置过期策略或未过期的Key数量是一致的。

- 目标阿里云Redis实例支持的版本为2.8、4.0或5.0版本，如需跨版本同步（仅支持从低版本同步到高版本）请提前确认兼容性。例如创建按量付费的Redis实例来测试，测试完成后可将该实例释放或转为包年包月。

## 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

## 支持的同步拓扑

- 一对一单向同步
- 一对多单向同步
- 级联单向同步

关于各类同步拓扑的介绍及注意事项，请参见[数据同步拓扑介绍](#)。

## 支持的同步命令

- APPEND
- BITOP、BLPOP、BRPOP、BRPOPLPUSH
- DECR、DECRBY、DEL
- EVAL、EVALSHA、EXEC、EXPIRE、EXPIREAT
- GEOADD、GETSET
- HDEL、HINCRBY、HINCRBYFLOAT、HMSET、HSET、HSETNX
- INCR、INCRBY、INCRBYFLOAT
- LINSERT、LPOP、LPUSH、LPUSHX、LREM、LSET、LTRIM
- MOVE、MSET、MSETNX、MULTI
- PERSIST、PEXPIRE、PEXPIREAT、PFADD、PFMERGE、PSETEX、PUBLISH
- RENAME、RENAMENX、RESTORE、RPOP、RPOPLPUSH、RPUSH、RPUSHX
- SADD、SDIFFSTORE、SELECT、SET、SETBIT、SETEX、SETNX、SETRANGE、SINTERSTORE、SMOVE、SPOP、SREM、SUNIONSTORE
- ZADD、ZINCRBY、ZINTERSTORE、ZREM、ZREMRANGEBYLEX、ZUNIONSTORE、ZREMRANGEBYRANK、ZREMRANGEBYSCORE

### 说明

- 不支持同步PUBLISH命令。
- 对于通过EVAL或者EVALSHA调用Lua脚本，在增量数据同步时，由于目标端在执行脚本时不会明确返回执行结果，DTS无法确保该类型脚本能够执行成功。
- 对于List，由于DTS在调用sync或psync进行重传时，不会对目标端已有的数据进行清空，可能导致出现重复数据。

## 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

 **说明** 购买时选择源实例为Redis、目标实例为Redis，并选择同步拓扑为单向同步。

2. 登录[数据传输控制台](#)。

3. 在左侧导航栏，单击数据同步。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击配置同步链路。
6. 配置数据同步的源实例及目标实例信息。

1. 选择同步通道的源及目标实例
2. 选择同步对象
3. 高级设置
4. 预检查

同步作业名称:

**源实例信息**

实例类型:

实例地区:

\* ECS实例ID:

数据库类型:

实例模式:  单机版  集群版

\* 端口:

数据库密码:

**目标实例信息**

实例类型:

实例地区:

\* 实例ID:

数据库密码:

| 类别    | 配置      | 说明  |
|-------|---------|---|
| 无     | 同步作业名称  | DTS会自动生成一个同步作业名称，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。  |
| 源实例信息 | 实例类型    | 选择ECS上的自建数据库。   |
|       | 实例地区    | 购买数据同步实例时选择的源实例地域，不可变更。   |
|       | ECS实例ID | 选择Codis-Group中Master节点所属的ECS实例ID。<br><div style="background-color: #e6f2ff; padding: 5px; margin-top: 5px;"> <span style="font-size: 1.2em; color: #007bff;">?</span> <b>说明</b> DTS通过同步Codis集群中的每个Codis-Group来实现整体的数据同步，此处先填入第一个Codis-Group中Master节点所属的ECS实例ID；稍后配置第二个数据同步作业时，此处填入第二个Codis-Group中Master节点所属的ECS实例ID；以此类推，直至为所有Codis-Group配置同步作业。                 </div> |
|       | 数据库类型   | 固定为Redis。   |

| 类别     | 配置    | 说明   |
|--------|-------|--|
|        | 实例模式  | <p>选择为<b>单机版</b>。</p> <p><b>?</b> 说明 由于Codis集群架构的特殊性，无法直接同步Codis集群，DTS通过同步Codis集群中的每个Codis-Group来实现整体的数据同步，所以此处需选择为<b>单机版</b>。</p>                 |
|        | 端口    | 填入该Codis-Group中Master节点的服务端口。  |
|        | 数据库密码 | <p>填该Master节点的数据库密码。</p> <p><b>?</b> 说明 非必填项，如果没有设置密码可以不填。</p>   |
| 目标实例信息 | 实例类型  | 选择 <b>Redis实例</b> 。  |
|        | 实例地区  | 购买数据同步实例时选择的目标实例地域，不可变更。   |
|        | 实例ID  | 选择目标阿里云Redis实例ID。  |
|        | 数据库密码 | <p>填入Redis实例的数据库密码。</p> <p><b>?</b> 说明 数据库密码格式为&lt;user&gt;:&lt;password&gt;。例如，Redis实例自定义的用户名为admin，密码为Rp829dlwa，则此处填入的数据库密码为admin:Rp829dlwa。</p> |

7. 单击页面右下角的**授权白名单**并进入下一步。

**?** 说明

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置目标已存在表的处理模式和同步对象。



| 配置          | 说明   |
|-------------|--|
| 目标已存在表的处理模式 | <p>DTS通过依次同步Codis集群中的每个Codis-Group来实现整体的数据同步。为第1个Codis-Group配置数据同步时，如果Redis实例暂无数据，请选择<b>预检查并报错拦截</b>。为第2到N个Codis-Group配置数据同步时，必须选择<b>忽略报错并继续执行</b>，否则将无法同步数据。</p> <div style="border: 1px solid #ccc; padding: 5px; background-color: #e6f2ff;"> <p><b>说明</b></p> <ul style="list-style-type: none"> <li><b>预检查并报错拦截</b>：检查目标库是否为空。如果待同步的目标库为空，则通过该检查项目；如果不为空，则在预检查阶段提示错误，数据同步作业不会被启动。</li> <li><b>忽略报错并继续执行</b>：忽略预检查阶段中检测到目标库不为空的报错，继续执行数据同步。如果在同步过程中遇到目标库中的key与源库中的key相同，会将源库的数据覆盖写入目标库中。</li> </ul> </div> |
| 同步对象        | <ul style="list-style-type: none"> <li>在<b>源库对象</b>框中单击待同步的数据库，然后单击  图标将其移动到<b>已选择对象</b>框。</li> <li>同步对象的选择粒度为库，暂不支持Key粒度的选择。</li> </ul>   |
| 映射名称更改      | 不支持更改映射名称。   |

| 配置                             | 说明   |
|--------------------------------|--|
| 源表DMS_ONLINE_DDL过程中是否复制临时表到目标库 | <p>如源库使用数据管理DMS (Data Management Service) 执行Online DDL变更, 您可以选择是否同步Online DDL变更产生的临时表数据。</p> <ul style="list-style-type: none"> <li>是: 同步Online DDL变更产生的临时表数据。</li> </ul> <p><b>说明</b> Online DDL变更产生的临时表数据过大, 可能会导致同步任务延迟。</p> <ul style="list-style-type: none"> <li>否: 不同步Online DDL变更产生的临时表数据, 只同步源库的原始DDL数据。</li> </ul> <p><b>说明</b> 该方案会导致目标库锁表。</p> |
| 源、目标库无法连接重试时间                  | <p>当源、目标库无法连接时, DTS默认重试720分钟(即12小时), 您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库, 同步任务将自动恢复。否则, 同步任务将失败。</p> <p><b>说明</b> 由于连接重试期间, DTS将收取任务运行费用, 建议您根据业务需要自定义重试时间, 或者在源和目标库实例释放后尽快释放DTS实例。</p>   |

9. 上述配置完成后, 单击页面右下角的下一步。

10. 配置同步初始化的选项。

1.选择同步通道的源及目标实例
2.选择同步对象
3.高级设置
4.预检查

同步初始化: 包含全量数据+增量数据

取消
上一步
保存
预检查并启动

**说明** 当前固定为包含全量数据+增量数据, 即DTS会将源Codis中的存量数据同步至目标Redis数据库中, 并同步增量数据。

11. 上述配置完成后, 单击页面右下角的预检查并启动。

**说明**

- 在同步作业正式启动之前, 会先进行预检查。只有预检查通过后, 才能成功启动同步作业。
- 如果预检查失败, 单击具体检查项后的 查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项, 您也可以选择**确认屏蔽**、**忽略告警项**并重新进行预检查, 跳过告警检测项重新进行预检查。

12. 在预检查对话框中显示预检查通过后, 关闭预检查对话框, 同步作业将正式开始。

13. 等待同步作业的链路初始化完成, 直至处于同步中状态。

| 实例ID/作业名称                             | 状态  | 同步概况                           | 付费方式 | 同步架构(全部) | 操作   |
|---------------------------------------|-----|--------------------------------|------|----------|--|
| <input type="checkbox"/> Codis-Group1 | 同步中 | 延时: 2 毫秒<br>速度: 0TPS(0.00MB/s) | 按量付费 | 单向同步     | <input type="checkbox"/> 暂停同步   <input type="checkbox"/> 释放同步   <a href="#">转包年包月</a>   <a href="#">升级更多</a> |

共有1条, 每页显示: 20条

说明 您可以在数据同步页面，查看数据同步作业的状态。

14. 重复第1步到第13步的操作，为剩余的Codis-Group创建数据同步作业。

### 执行结果

本案例的Codis集群具备两个Codis-Group，所以创建两个数据同步作业。如下图所示，这两个数据同步作业完成同步初始化后，已经都处于同步中状态。

| 实例ID/作业名称    | 状态  | 同步概况                           | 付费方式 | 同步架构(全部) | 操作                  |
|--------------|-----|--------------------------------|------|----------|---------------------|
| Codis-Group2 | 同步中 | 延时: 0 毫秒<br>速度: 0TPS(0.00MB/s) | 按量付费 | 单向同步     | 暂停同步   转包年包月   升级更多 |
| Codis-Group1 | 同步中 | 延时: 2 毫秒<br>速度: 0TPS(0.00MB/s) | 按量付费 | 单向同步     | 暂停同步   转包年包月   升级更多 |

本案例同步的数据库为DB0和DB1，通过DMS登录Redis实例后与源Codis集群进行对比，Key总数源Codis集群一致。

#### 阿里云Redis实例

#### 源Codis集群

| Group | Server                   | Data Center | Master            | Memory          | Keys  |
|-------|--------------------------|-------------|-------------------|-----------------|---|
| 1     | SYNC S 127.0.0.1:6379    |             | NO:ONE            | 38.03 MB / INF. | db0:keys=31584,expires=0,avg_ttl=0<br>db1:keys=1,expires=0,avg_ttl=0<br>db15:keys=1,expires=0,avg_ttl=0 |
|       | PROMOTE S 127.0.0.1:6380 |             | 127.0.0.1:6379:up | 36.61 MB / INF. | db0:keys=31584,expires=0,avg_ttl=0<br>db1:keys=1,expires=0,avg_ttl=0<br>db15:keys=1,expires=0,avg_ttl=0 |
| 2     | SYNC S 127.0.0.1:6389    |             | NO:ONE            | 38.06 MB / INF. | db0:keys=31636,expires=0,avg_ttl=0<br>db1:keys=1,expires=0,avg_ttl=0<br>db15:keys=1,expires=0,avg_ttl=0 |
|       | PROMOTE S 127.0.0.1:6390 |             | 127.0.0.1:6389:up | 36.67 MB / INF. | db0:keys=31636,expires=0,avg_ttl=0<br>db1:keys=1,expires=0,avg_ttl=0<br>db15:keys=1,expires=0,avg_ttl=0 |

## 11.8. 从ECS上的Twemproxy Redis集群同步至Redis实例

阿里云数据库Redis版是兼容开源Redis协议标准、提供内存加硬盘混合存储的数据库服务，基于高可靠双机热备架构及可平滑扩展的集群架构，可充分满足高吞吐、低延迟及弹性变配的业务需求。通过数据传输服务DTS（Data Transmission Service），您可以将Twemproxy Redis同步至阿里云Redis实例。

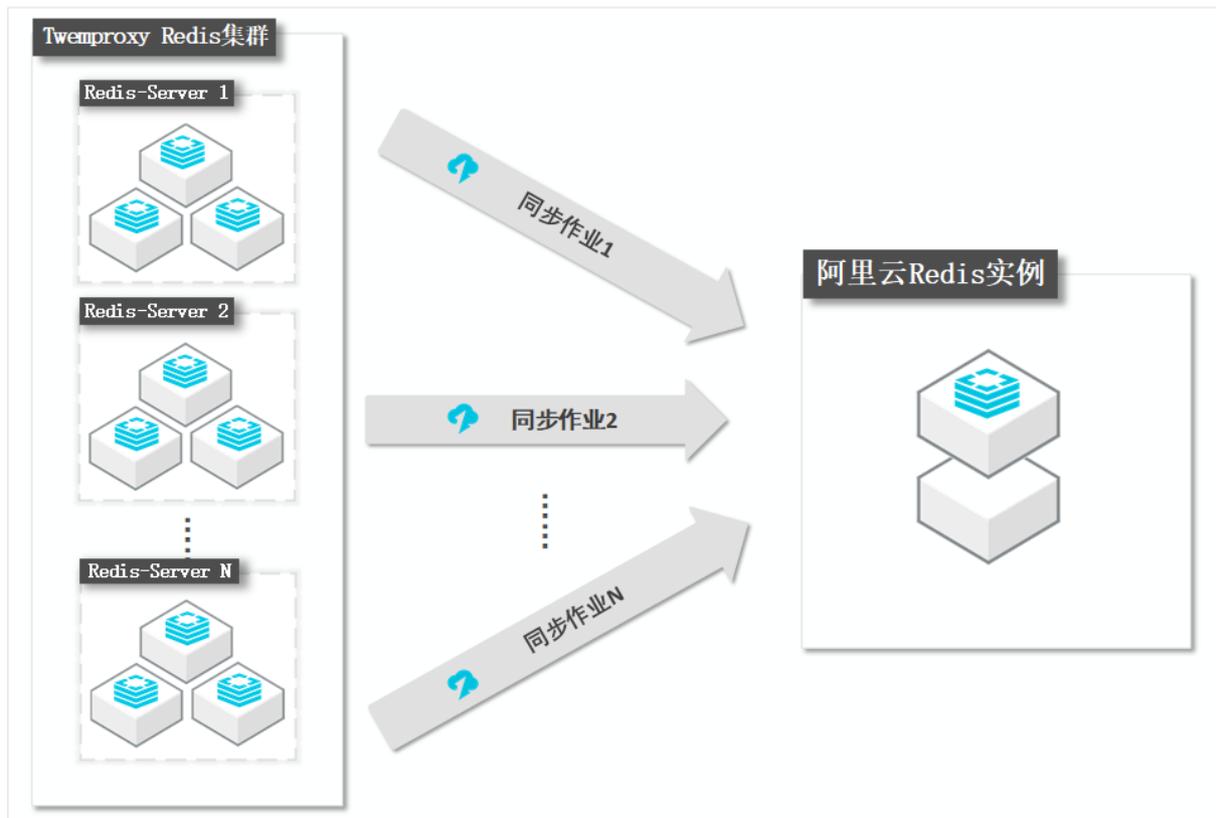
### 前提条件

- 已创建目标阿里云Redis实例，详情请参见创建Redis实例。
- 阿里云Redis实例的存储空间需大于源Twemproxy Redis数据库已使用的存储空间。
- 源Twemproxy Redis集群中的每个Master节点必须能够执行 `psync` 命令。

- 源Redis实例Slave和Master之间的复制超时时间参数repl-timeout，默认为60秒, 建议使用 `config set repl-timeout 600` 命令设置为600秒。若源数据库数据量比较大，可以适当增大repl-timeout参数的值。

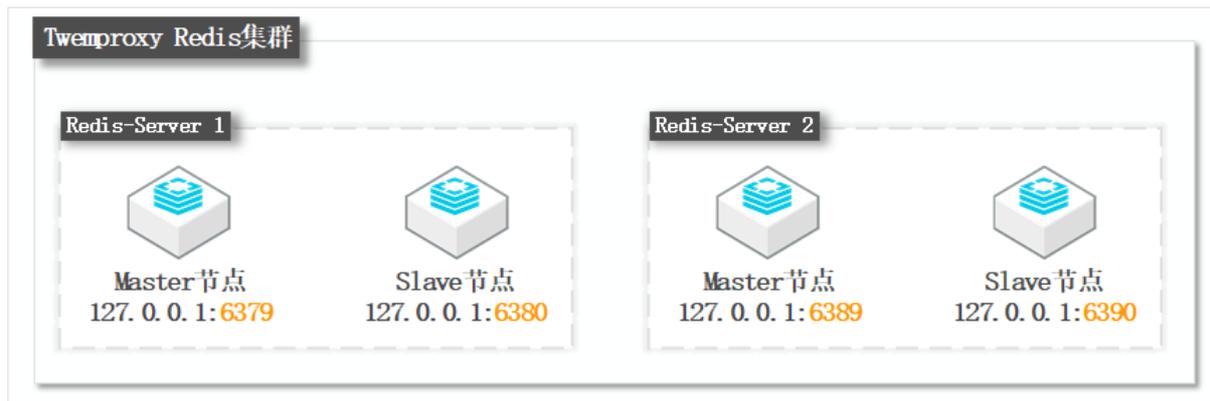
### 同步原理介绍

DTS通过同步Twemproxy Redis集群中的每个Redis-Server来实现集群的整体同步，您需要为每个Redis-Server创建一个对应的数据同步作业。



### Twemproxy Redis集群环境介绍

在本案例中，Twemproxy Redis集群具备两个Redis-Server，每个Redis-Server采用一主一从的架构，详情请参见下图。



### 注意事项

- DTS在执行全量数据初始化时将占用源库和目标库一定的资源，可能会导致数据库服务器负载上升。如果数据库业务量较大或服务器规格较低，可能会加重数据库压力，甚至导致数据库服务不可用。建议您在执行数据同步前谨慎评估，在业务低峰期执行数据同步。
- 如果在源数据库的配置文件redis.conf中配置了 `bind` 参数，请将该参数的值设置为ECS的内网IP地址以保障DTS可以正常连接源数据库。

- 为保障同步链路稳定性，建议将源集群的配置文件 `redis.conf` 中 `repl-backlog-size` 参数的值适当调大。
- 为保障同步质量，DTS会在源集群中插入一个key: `DTS_REDIS_TIMESTAMP_HEARTBEAT`，用于记录更新时间点。
- 请勿在源集群中执行 `FLUSHDB` 和 `FLUSHALL` 命令，否则将导致源和目标的数据不一致。
- 如果目标库的数据逐出策略（`maxmemory-policy`）配置为 `noeviction` 以外的值，可能导致目标库的数据与源库不一致。关于数据逐出策略详情，请参见[Redis数据逐出策略介绍](#)。
- 如果源库中的某些Key使用了过期（`expire`）策略，由于可能存在Key已过期但未被及时删除的情况，所以在目标库中看到的Key数量（例如通过`info`命令查看）会比源库的Key数量少。

 **说明** 源和目标库中，未设置过期策略或未过期的Key数量是一致的。

- 目标阿里云Redis实例支持的版本为2.8、4.0或5.0版本，如需跨版本同步（仅支持从低版本同步到高版本）请提前确认兼容性。例如创建按量付费的Redis实例来测试，测试完成后可将该实例释放或转为包年包月。

### 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

### 支持的同步拓扑

- 一对一单向同步
- 一对多单向同步
- 级联单向同步

关于各类同步拓扑的介绍及注意事项，请参见[数据同步拓扑介绍](#)。

### 支持的同步命令

- APPEND
- BITOP、BLPOP、BRPOP、BRPOPLPUSH
- DECR、DECRBY、DEL
- EVAL、EVALSHA、EXEC、EXPIRE、EXPIREAT
- GEOADD、GETSET
- HDEL、HINCRBY、HINCRBYFLOAT、HMSET、HSET、HSETNX
- INCR、INCRBY、INCRBYFLOAT
- LINSERT、LPOP、LPUSH、LPUSHX、LREM、LSET、LTRIM
- MOVE、MSET、MSETNX、MULTI
- PERSIST、PEXPIRE、PEXPIREAT、PFADD、PFMERGE、PSETEX、PUBLISH
- RENAME、RENAMENX、RESTORE、RPOP、RPOPLPUSH、RPUSH、RPUSHX
- SADD、SDIFFSTORE、SELECT、SET、SETBIT、SETEX、SETNX、SETRANGE、SINTERSTORE、SMOVE、SPOP、SREM、SUNIONSTORE
- ZADD、ZINCRBY、ZINTERSTORE、ZREM、ZREMRANGEBYLEX、ZUNIONSTORE、ZREMRANGEBYRANK、ZREMRANGEBYSCORE

说明

- 不支持同步PUBLISH命令。
- 对于通过EVAL或者EVALSHA调用Lua脚本，在增量数据同步时，由于目标端在执行脚本时不会明确返回执行结果，DTS无法确保该类型脚本能够执行成功。
- 对于List，由于DTS在调用sync或psync进行重传时，不会对目标端已有的数据进行清空，可能导致出现重复数据。

操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

说明 购买时选择源实例为Redis、目标实例为Redis，并选择同步拓扑为单向同步。

- 登录[数据传输控制台](#)。
- 在左侧导航栏，单击[数据同步](#)。
- 在同步作业列表页面顶部，选择同步的目标实例所属地域。
- 定位至已购买的数据同步实例，单击[配置同步链路](#)。
- 配置数据同步的源实例及目标实例信息。

1. 选择同步通道的源及目标实例
2. 选择同步对象
3. 高级设置
4. 预检查

同步作业名称:

**源实例信息**

实例类型:

实例地区:

\* ECS实例ID:

数据库类型:

实例模式:  单机版  集群版

\* 端口:

数据库密码:

**目标实例信息**

实例类型:

实例地区:

\* 实例ID:

数据库密码:

| 类别 | 配置     | 说明   |
|----|--------|--|
| 无  | 同步作业名称 | DTS会自动生成一个同步作业名称，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。 |
|    | 实例类型   | 选择ECS上的自建数据库。                                  |
|    | 实例地区   | 购买数据同步实例时选择的源实例地域，不可变更。                        |

| 类别     | 配置      | 说明   |
|--------|---------|--|
| 源实例信息  | ECS实例ID | <p>选择源集群Redis-Server中Master节点所属的ECS实例ID。</p> <p><b>?</b> 说明 DTS通过同步Tmemproxy Redis集群中的每个Redis-Server来实现整体的数据同步，此处先填入第一个Redis-Server中Master节点所属的ECS实例ID；稍后配置第二个数据同步作业时，此处填入第二个Redis-Server中Master节点所属的ECS实例ID；以此类推，直至为所有Redis-Server配置同步作业。</p> |
|        | 数据库类型   | 固定为Redis。  |
|        | 实例模式    | <p>选择为<b>单机版</b>。</p> <p><b>?</b> 说明 由于Tmemproxy Redis集群架构的特殊性，无法直接同步Tmemproxy Redis集群，DTS需要通过同步Tmemproxy Redis集群中的每个Redis-Server来实现整体的数据同步，所以此处需选择为<b>单机版</b>。</p>  |
|        | 端口      | 填入该Redis-Server中Master节点的服务端口。   |
|        | 数据库密码   | <p>填该Master节点的数据库密码。</p> <p><b>?</b> 说明 非必填项，如果没有设置密码可以不填。</p>   |
| 目标实例信息 | 实例类型    | 选择 <b>Redis实例</b> 。  |
|        | 实例地区    | 购买数据同步实例时选择的目标实例地域，不可变更。   |
|        | 实例ID    | 选择目标阿里云Redis实例ID。  |
|        | 数据库密码   | <p>填入Redis实例的数据库密码。</p> <p><b>?</b> 说明 数据库密码格式为&lt;user&gt;:&lt;password&gt;。例如，Redis实例自定义的用户名为admin，密码为Rp829dlwa，则此处填入的数据库密码为admin:Rp829dlwa。</p>   |

7. 单击页面右下角的授权白名单并进入下一步。

**?** 说明

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置目标已存在表的处理模式和同步对象。



| 配置          | 说明   |
|-------------|--|
| 目标已存在表的处理模式 | <p>DTS通过同步Twemproxy Redis集群中的每个Redis-Server来实现整体的数据同步。为第1个Redis-Server配置数据同步时，如果Redis实例暂无数据，请选择<b>预检查并报错拦截</b>。为第2到N个Redis-Server配置数据同步时，必须选择为<b>忽略报错并继续执行</b>，否则将无法同步数据。</p> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>预检查并报错拦截：检查目标库是否为空。如果待同步的目标库为空，则通过该检查项目；如果不为空，则在预检查阶段提示错误，数据同步作业不会被启动。</li> <li>忽略报错并继续执行：忽略预检查阶段中检测到目标库不为空的报错，继续执行数据同步。如果在同步过程中遇到目标库中的key与源库中的key相同，会将源库的数据覆盖写入目标库中。</li> </ul> |
| 同步对象        | <ul style="list-style-type: none"> <li>在源库对象框中单击待同步的数据库，然后单击  图标将其移动到已选择对象框。</li> <li>同步对象的选择粒度为库，暂不支持Key粒度的选择。</li> </ul>  |
| 映射名称更改      | 不支持更改映射名称。   |

| 配置                             | 说明   |
|--------------------------------|--|
| 源表DMS_ONLINE_DDL过程中是否复制临时表到目标库 | <p>如源库使用数据管理DMS (Data Management Service) 执行Online DDL变更, 您可以选择是否同步Online DDL变更产生的临时表数据。</p> <ul style="list-style-type: none"> <li>是: 同步Online DDL变更产生的临时表数据。</li> </ul> <p><b>说明</b> Online DDL变更产生的临时表数据过大, 可能会导致同步任务延迟。</p> <ul style="list-style-type: none"> <li>否: 不同步Online DDL变更产生的临时表数据, 只同步源库的原始DDL数据。</li> </ul> <p><b>说明</b> 该方案会导致目标库锁表。</p> |
| 源、目标库无法连接重试时间                  | <p>当源、目标库无法连接时, DTS默认重试720分钟(即12小时), 您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库, 同步任务将自动恢复。否则, 同步任务将失败。</p> <p><b>说明</b> 由于连接重试期间, DTS将收取任务运行费用, 建议您根据业务需要自定义重试时间, 或者在源和目标库实例释放后尽快释放DTS实例。</p>   |

9. 上述配置完成后, 单击页面右下角的下一步。

10. 配置同步初始化的选项。

1.选择同步通道的源及目标实例
2.选择同步对象
3.高级设置
4.预检查

同步初始化: 包含全量数据+增量数据

取消
上一步
保存
预检查并启动

**说明** 当前固定为包含全量数据+增量数据, 即DTS会将源Twemproxy Redis集群中的存量数据同步至目标Redis数据库中, 并同步增量数据。

11. 上述配置完成后, 单击页面右下角的预检查并启动。

**说明**

- 在同步作业正式启动之前, 会先进行预检查。只有预检查通过后, 才能成功启动同步作业。
- 如果预检查失败, 单击具体检查项后的 , 查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项, 您也可以选择确认屏蔽、忽略告警项并重新进行预检查, 跳过告警检测项重新进行预检查。

12. 在预检查对话框中显示预检查通过后, 关闭预检查对话框, 同步作业将正式开始。

13. 等待同步作业的链路初始化完成, 直至处于同步中状态。

| 实例ID/作业名称   | 状态  | 同步概况                           | 付费方式 | 同步架构(全部) | 操作   |
|---|-----|--------------------------------|------|----------|--|
| <input type="checkbox"/> twemproxy-node1                    | 同步中 | 延时: 0 毫秒<br>速度: 0TPS(0.00MB/s) | 按量付费 | 单向同步     | <input type="checkbox"/> 暂停同步   <input type="checkbox"/> 转包年包月   <input type="checkbox"/> 升级更多                                 |
| <input type="checkbox"/> 暂停同步 <input type="checkbox"/> 释放同步 |     |                                |      |          | 共有1条, 每页显示: 20条 <span style="float: right;">                         &lt;&lt; &lt; 1 &gt; &gt;&gt;                     </span> |

**说明** 您可以在数据同步页面，查看数据同步作业的状态。

14. 重复第1步到第13步的操作，为剩余的Redis-Server创建数据同步作业。

### 执行结果

本案例的Twemproxy Redis集群具备两个Redis-Server，所以创建两个数据同步作业。如下图所示，这两个数据同步作业完成同步初始化后，都已经处于同步中状态。

| 实例ID/作业名称       | 状态  | 同步概况                           | 付费方式 | 同步架构(全部) | 操作                  |
|-----------------|-----|--------------------------------|------|----------|---------------------|
| twemproxy-node2 | 同步中 | 延时: 0 毫秒<br>速度: 0TPS(0.00MB/s) | 按量付费 | 单向同步     | 暂停同步   转包年包月   升级更多 |
| twemproxy-node1 | 同步中 | 延时: 0 毫秒<br>速度: 0TPS(0.00MB/s) | 按量付费 | 单向同步     | 暂停同步   转包年包月   升级更多 |

共有2条，每页显示：20条

本案例同步的数据库为DB0，通过DMS登录Redis集群后与源Twemproxy Redis集群进行对比，Key总数源Twemproxy Redis集群一致。

#### 阿里云Redis实例

#### 源Twemproxy Redis集群

```
root@iZbp1ib0ezn1xo15wbfsadZ:~# redis-cli -p 6379 info|grep db0
db0:keys=29421,expires=0,avg_ttl=0
root@iZbp1ib0ezn1xo15wbfsadZ:~# redis-cli -p 6389 info|grep db0
db0:keys=33779,expires=0,avg_ttl=0
root@iZbp1ib0ezn1xo15wbfsadZ:~#
```

# 12.SQL Server数据同步

## 12.1. 从RDS SQL Server同步至云原生数据仓库AnalyticDB PostgreSQL

数据传输服务DTS（Data Transmission Service）支持将SQL Server同步至云原生数据仓库AnalyticDB PostgreSQL，帮助您轻松实现数据的流转，集中分析企业数据。

### 前提条件

- [创建RDS SQL Server实例](#)，支持的版本，请参见[同步方案概览](#)。

 **注意** 目前支持源RDS SQL Server版本为2012、2014、2016、2017或2019版本。

- [创建云原生数据仓库AnalyticDB PostgreSQL实例](#)。
- RDS SQL Server实例中待同步的表需具备主键。
- 云原生数据仓库AnalyticDB PostgreSQL实例中同步的目标表需具备主键或唯一索引。

### 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。因此您需要在执行数据迁移前评估源库和目标库的性能，同时建议您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）。
- 为保障数据同步延迟显示的准确性，DTS会在源库中新增一张心跳表（名称为 `dts_log_heart_beat`）。
- 此场景中，DTS支持初始化的结构为Schema、Table、View、Function和Procedure。

 **警告** 由于此场景属于异构数据库间的数据同步，数据类型无法一一对应，请谨慎评估数据类型的映射关系对业务的影响，详情请参见[结构初始化涉及的数据类型映射关系](#)。

- 不支持同步数据类型为TIMESTAMP、CURSOR、ROWVERSION、HIERACHYID、SQL\_VARIANT、SPATIAL GEOMETRY、SPATIAL GEOGRAPHY、TABLE的数据。

### 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

### 支持同步的SQL操作

- DDL操作：ADD COLUMN
- DML操作：INSERT、UPDATE、DELETE

### 数据库账号的权限要求

| 数据库              | 所需权限          | 授权方法                   |
|------------------|---------------|------------------------|
| RDS SQL Server实例 | 待同步数据库的所有者权限。 | <a href="#">修改账号权限</a> |

| 数据库                            | 所需权限  | 授权方法                                  |
|--------------------------------|---|---------------------------------------|
| 云原生数据仓库AnalyticDB PostgreSQL实例 | <ul style="list-style-type: none"> <li>• LOGIN权限。</li> <li>• 目标表的SELECT、CREATE、INSERT、UPDATE、DELETE权限。</li> <li>• 目标库的CONNECT、CREATE权限。</li> <li>• 目标Schema的CREATE权限。</li> <li>• Copy权限（基于内存batch copy）。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p><span style="color: #00aaff;">?</span> <b>说明</b> 您也可以使用云原生数据仓库AnalyticDB PostgreSQL实例的初始账号。</p> </div> | <p style="color: #e67e22;">用户权限管理</p> |

### 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

? **说明** 购买时，选择源实例为SQLServer，目标实例为AnalyticDB for PostgreSQL，并选择同步拓扑为单向同步。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击[数据同步](#)。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击[配置同步链路](#)。
6. 配置同步作业的源实例及目标实例信息。

1.选择同步通道的源及目标实例
2.选择同步对象
3.预检查

同步作业名称:

**源实例信息**

实例类型:

实例地区:

ECS实例ID:

数据库类型:

\* 端口:

\* 数据库账号:

\* 数据库密码:

\* 连接方式:  非加密连接  SSL安全连接

**目标实例信息**

实例类型:

实例地区:

\* 实例ID:

\* 数据库名称:

\* 数据库账号:

\* 数据库密码:

| 类别     | 配置     | 说明  |
|--------|--------|---|
| 无      | 同步作业名称 | DTS会自动生成一个同步作业名称，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。  |
| 源实例信息  | 实例类型   | 选择RDS实例。  |
|        | 实例地区   | 购买数据同步实例时选择的源实例地域信息，不可变更。   |
|        | 实例ID   | 选择源RDS SQL Server实例ID。  |
|        | 数据库账号  | 填入RDS SQL Server的数据库账号。权限要求请参见 <a href="#">数据库账号的权限要求</a> 。   |
|        | 数据库密码  | 填入该数据库账号的密码。  |
|        | 连接方式   | 根据需求选择非加密连接或SSL安全连接。如果设置为SSL安全连接，您需要提前开启RDS实例的SSL加密功能，详情请参见 <a href="#">设置SSL加密</a> 。<br><span style="background-color: #e6f2ff; padding: 5px; border: 1px solid #d9e1f2; border-radius: 5px;"> <span style="font-size: 1em;">?</span> 说明 目前仅中国内地及中国香港地域支持设置连接方式。                     </span> |
| 目标实例信息 | 实例类型   | 固定为AnalyticDB for PostgreSQL，无需设置。  |
|        | 实例地区   | 购买数据同步实例时选择的目标实例地域信息，不可变更。  |
|        | 实例ID   | 选择目标云原生数据仓库AnalyticDB PostgreSQL实例ID。   |
|        | 数据库名称  | 填入同步目标表所属的数据库名称。  |
|        | 数据库账号  | 填入云原生数据仓库AnalyticDB PostgreSQL的数据库账号。权限要求请参见 <a href="#">数据库账号的权限要求</a> 。   |
|        | 数据库密码  | 填入该数据库账号对应的密码。  |

7. 单击页面右下角的授权白名单并进入下一步。

? 说明

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置同步策略和同步对象。

1.选择同步通道的源及目标实例2.选择同步对象3.预检查

同步初始化:  结构初始化  全量数据初始化  增量数据初始化

注: DTS全量任务运行期间, 不要清理DTS任务启动后源库产生的增量数据日志。源库如果过早清理日志, 可能会导致DTS增量任务失败

目标已存在表的处理模式:  预检查并报错拦截  忽略报错并继续执行

多表归并:  是  否

同步操作类型:  Insert  Update  Delete

**源库对象**

若全局搜索, 请先展开树 | Q

- test123
- testdb

全选

**已选择对象** (鼠标移到对象行, 点击编辑可修改对象名或过滤条件) 详情点我

| Q

- dtstest

全选

>  
<

\* 映射名称更改:  不进行库表名称批量更改  要进行库表名称批量更改

\* 源库、目标库无法连接后的重试时间:  分钟 ?

\* 为目标对象添加引号 ?  是  否

取消上一步下一步预检查并启动

| 配置    | 说明  |
|-------|---|
| 同步初始化 | 默认选中结构初始化、全量数据初始化和增量数据初始化。预检查完成后, DTS会将源实例中待同步对象的结构和存量数据同步至目标在目标库, 作为后续增量同步数据的基线数据。 |

| 配置          | 说明  |
|-------------|---|
| 目标已存在表的处理模式 | <ul style="list-style-type: none"> <li>○ <b>预检查并报错拦截</b>：检查目标数据库中是否有同名的表。如果目标数据库中没有同名的表，则通过该检查项目；如果目标数据库中有同名的表，则在预检查阶段提示错误，数据同步作业不会被启动。                             <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin-top: 10px;">                                 ? <b>说明</b> 如果目标库中同名的表不能删除或重命名，您可以更改该表在目标库中的名称，详情请参见<a href="#">设置同步对象在目标实例中的名称</a>。                             </div> </li> <li>○ <b>忽略报错并继续执行</b>：跳过目标数据库中是否有同名表的检查项。                             <div style="border: 1px solid #ccc; background-color: #fff9c4; padding: 5px; margin-top: 10px;">                                 ! <b>警告</b> 选择为忽略报错并继续执行，可能导致数据不一致，给业务带来风险，例如：                                 <ul style="list-style-type: none"> <li>■ 表结构一致的情况下，在目标库遇到与源库主键的值相同的记录，则会保留目标集群中的该条记录，即源库中的该条记录不会同步至目标数据库中。</li> <li>■ 表结构不一致的情况下，可能会导致无法初始化数据、只能同步部分列的数据或同步失败。</li> </ul> </div> </li> </ul>  |
| 多表归并        | <ul style="list-style-type: none"> <li>○ <b>选择为是</b>：通常在OLTP场景中，为提高业务表响应速度，通常会做分库分表处理。而在云原生数据仓库AnalyticDB PostgreSQL中单个数据表可存储海量数据，使用单表查询更加便捷。此类场景中，您可以借助DTS的多表归并功能将源库中多个表结构相同的表（即各分表）同步至云原生数据仓库AnalyticDB PostgreSQL中的同一个表中。                             <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin-top: 10px;">                                 ? <b>说明</b> <ul style="list-style-type: none"> <li>■ 选择源库的多个表后，您需要通过对对象名映射功能，将其改为云原生数据仓库AnalyticDB PostgreSQL中的同一个表名。关于对象名映射功能的介绍，请参见<a href="#">设置同步对象在目标实例中的名称</a>。</li> <li>■ 您需要在云原生数据仓库AnalyticDB PostgreSQL的同步目标表中增加 <code>__dts_data_source</code> 列（类型为text）来存储数据来源。DTS将以 <code>&lt;dts数据同步实例ID&gt;:&lt;源数据库名&gt;.&lt;源Schema名&gt;.&lt;源表名&gt;</code> 的格式写入列值用于区分表的来源，例如 <code>dts*****:dtstestdata.testschema.customer1</code>。</li> <li>■ 多表归并功能基于任务级别，即不支持基于表级别执行多表归并。如果需要让部分表执行多表归并，另一部分不执行多表归并，您需要创建两个数据同步作业。</li> </ul> </div> </li> <li>○ <b>选择为否</b>：默认选项。</li> </ul> |
| 同步操作类型      | 根据业务选中需要同步的操作类型，默认情况下都处于选中状态。   |

| 配置            | 说明   |
|---------------|--|
| 选择同步对象        | <p>在源库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。</p> <p>本场景为异构数据库间同步，因此同步对象选择的粒度为表，且其他对象（如视图、触发器、存储过程）不会被同步至目标库。</p> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>默认情况下，同步对象的名称保持不变。如果您需要同步对象在目标实例上名称不同，请使用对象名映射功能，详情请参见<a href="#">设置同步对象在目标实例中的名称</a>。</li> <li>如果配置多表归并为是，在选择源库的多个表后，您需要通过对对象名映射功能，将其改为云原生数据仓库AnalyticDB PostgreSQL中的同一个表名。</li> </ul> </div>                        |
| 为目标对象添加引号     | <p>选择是否需要为目标对象名添加引号。如果选择为是，且存在下述情况，DTS在结构初始化阶段和增量数据迁移阶段会为目标对象添加单引号或双引号：</p> <ul style="list-style-type: none"> <li>源库所属的业务环境对大小写敏感且大小写混用。</li> <li>源表名不是以字母开头，且包含字母、数字或特殊字符以外的字符。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p><b>说明</b> 特殊字符仅支持下划线（_），井号（#）和美元符号（\$）。</p> </div> <ul style="list-style-type: none"> <li>待迁移的Schema、表或列名称是目标库的关键字、保留字或非法字符。</li> </ul> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p><b>说明</b> 如果选择添加引号，在数据同步完成后，您需使用带引号的目标对象名进行查询。</p> </div> |
| 映射名称更改        | <p>如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见<a href="#">库表列映射</a>。</p>  |
| 源、目标库无法连接重试时间 | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p><b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p> </div>   |

9. 设置待同步的表在云原生数据仓库AnalyticDB PostgreSQL中表类型、主键列和分布键信息。

1.选择同步通道的源及目标实例
2.选择同步对象
3.预检查

| 源库名         | 源表名      | 类型(全部) | 主键列       | 分布键       | 定义状态(全部) |
|-------------|----------|--------|-----------|-----------|----------|
| dtstestdata | customer | 维度表    | id        | --        | 已定义      |
| dtstestdata | Student  | 分区表    | StudentID | StudentID | 已定义      |

共有2条, 每页显示: 20 条

**说明** 关于主键列和分布键的详细说明，请参见[表的约束定义](#)和[表分布键定义](#)。

10. 上述配置完成后，单击页面右下角的预检查并启动。

**说明**

- 在同步作业正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步作业。
- 如果预检查失败，单击具体检查项后的 ，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择**确认屏蔽**、**忽略告警项**并重新进行预检查，跳过告警检测项重新进行预检查。

11. 在预检查对话框中显示预检查通过后，关闭预检查对话框，同步作业将正式开始。

12. 等待同步作业的链路初始化完成，直至处于同步中状态。

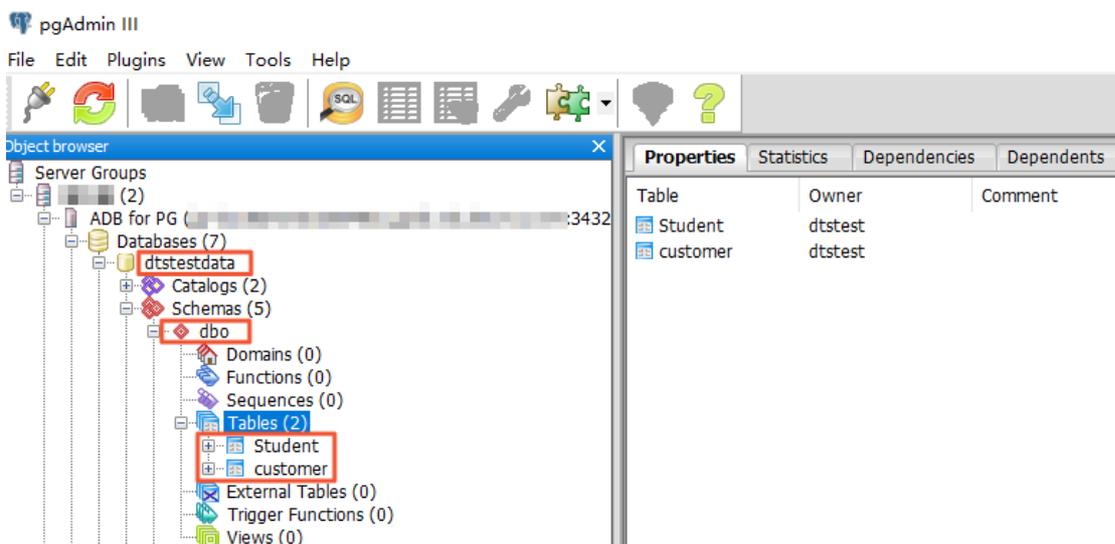
您可以在数据同步页面，查看数据同步作业的状态。



### 常见问题

Q: 如何在云原生数据仓库AnalyticDB PostgreSQL找到同步的目标表?

A: DTS的结构初始化会遵循源库的结构将其同步至目标库。本案例中，您可以在目标实例的 dtstestdata 数据库的 dbo Schema中，找到 customer 表和 Student 表，如下图所示。



## 12.2. 从ECS上的自建SQL Server同步至云原生数据仓库AnalyticDB PostgreSQL

数据传输服务DTS (Data Transmission Service) 支持将自建SQL Server同步至云原生数据仓库AnalyticDB PostgreSQL, 实现增量数据的实时同步。

### 前提条件

- 自建SQL Server数据库版本为2008、2008 R2、2012、2014、2016、2017或2019版本。

 **说明** 如果是SQL Server AlwaysOn High Availability Group，需要使用强同步模式。

- 自建SQL Server数据库中待同步的表需具备主键或者唯一性非空索引以确保同步过程中的幂等性。
- 云原生数据仓库AnalyticDB PostgreSQL实例的存储空间须大于自建SQL Server数据库占用的存储空间。

## 注意事项

- DTS在执行全量数据初始化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。因此您需要在执行数据同步前评估源库和目标库的性能，同时建议您在业务低峰期执行数据同步（例如源库和目标库的CPU负载在30%以下）。
- 为保障数据同步的正常运行，请勿对源库频繁执行备份，同时建议日志保留3天以上，避免日志截断后无法获取日志。
- 为保证数据同步延迟显示的准确性，DTS会在自建SQL Server数据库中新增一张心跳表，表名格式为 `待同步表名_dts_mysql_heartbeat`。

## 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

## 功能限制

- 不支持assemblies、service broker、全文索引、全文目录、分布式schema、分布式函数、CLR存储过程、CLR标量函数、CLR表值函数、内部表、系统、聚合函数的结构同步。
- 不支持同步数据类型为TIMESTAMP、CURSOR、ROWVERSION、HIERACHYID、SQL\_VARIANT、SPATIAL GEOMETRY、SPATIAL GEOGRAPHY、TABLE的数据。
- 不支持同步含有计算列的表。

## 支持同步的SQL操作

- DML操作：INSERT、UPDATE、DELETE
- DDL操作：ADD COLUMN

## 数据库账号的权限要求

| 数据库                            | 所需权限  | 账号创建及授权方法   |
|--------------------------------|---|---|
| 自建SQL Server实例                 | sysadmin  | <ul style="list-style-type: none"> <li>• <a href="#">CREATE USER</a></li> <li>• <a href="#">用户权限管理</a></li> </ul> |
| 云原生数据仓库AnalyticDB PostgreSQL实例 | <ul style="list-style-type: none"> <li>• LOGIN权限。</li> <li>• 目标表的SELECT、CREATE、INSERT、UPDATE、DELETE权限。</li> <li>• 目标库的CONNECT、CREATE权限。</li> <li>• 目标Schema的CREATE权限。</li> <li>• Copy权限（基于内存batch copy）。</li> </ul> <p> <b>说明</b> 您也可以使用云原生数据仓库AnalyticDB PostgreSQL实例的初始账号。</p> | <ul style="list-style-type: none"> <li>• <a href="#">创建数据库账号</a></li> <li>• <a href="#">用户权限管理</a></li> </ul>     |

## 准备工作

在正式配置数据同步任务之前，需要在自建SQL Server数据库上进行日志配置并创建聚集索引。

1. 在自建SQL Server数据库中执行如下命令，将待同步的数据库恢复模式修改为完整模式。也可通过SSMS客户端修改，具体请参见[修改数据库的recovery mode](#)。

```
use master;
GO
ALTER DATABASE <database_name> SET RECOVERY FULL WITH ROLLBACK IMMEDIATE;
GO
```

参数说明：

<database\_name>：待同步的数据库名。

示例：

```
use master;
GO
ALTER DATABASE mytestdata SET RECOVERY FULL WITH ROLLBACK IMMEDIATE;
GO
```

2. 执行如下命令，将待同步的数据库进行逻辑备份。如您已进行过逻辑备份，可跳过本步骤。

```
BACKUP DATABASE <database_name> TO DISK='<physical_backup_device_name>';
GO
```

参数说明：

- o <database\_name>：待同步的数据库名。
- o <physical\_backup\_device\_name>：指定备份文件存储的路径和文件名。

示例：

```
BACKUP DATABASE mytestdata TO DISK='D:\backup\dbdata.bak';
GO
```

3. 执行如下命令，将待同步的数据库日志进行备份。

```
BACKUP LOG <database_name> to DISK='<physical_backup_device_name>' WITH init;
GO
```

参数说明：

- o <database\_name>：待同步的数据库名。
- o <physical\_backup\_device\_name>：指定备份文件存储的路径和文件名。

示例：

```
BACKUP LOG mytestdata TO DISK='D:\backup\dblog.bak' WITH init;
GO
```

4. 为待同步的表创建聚集索引，详情请参见[创建聚集索引](#)。

## 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

 **说明** 购买时，选择源实例为SQLServer，目标实例为AnalyticDB for PostgreSQL，并选择同步拓扑为单向同步。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击[数据同步](#)。
4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
5. 定位至已购买的数据同步实例，单击[配置同步链路](#)。
6. 配置同步作业的源实例及目标实例信息。

1.选择同步通道的源及目标实例
2.选择同步对象
3.预检查

同步作业名称:

---

**源实例信息**

实例类型:

实例地区:

\* ECS实例ID:

数据库类型:

\* 端口:

\* 数据库账号:

\* 数据库密码:

\* 连接方式:  非加密连接  SSL安全连接

---

**目标实例信息**

实例类型:

实例地区:

\* 实例ID:

\* 数据库名称:

\* 数据库账号:

\* 数据库密码:

| 类别    | 配置      | 说明   |
|-------|---------|--|
| 无     | 同步作业名称  | DTS会自动生成一个同步作业名称，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。   |
| 源实例信息 | 实例类型    | 根据源库的部署位置进行选择，本文以ECS上的自建数据库为例介绍配置流程。<br><span style="background-color: #e6f2ff; padding: 5px; border: 1px solid #d9e1f2;"> <span style="font-size: 1.2em; color: #007bff;">?</span> 说明 当自建数据库为其他实例类型时，您还需要执行相应的准备工作，详情请参见<a href="#">准备工作概览</a>。                 </span> |
|       | 实例地区    | 购买数据同步实例时选择的源实例地域信息，不可变更。  |
|       | ECS实例ID | 选择源库所属的ECS实例ID。  |
|       | 数据库类型   | 固定为SQLServer。  |
|       | 端口      | 填入自建数据库的服务端口，默认为3306。  |
|       | 数据库账号   | 填入源库的数据库账号。权限要求请参见 <a href="#">数据库账号的权限要求</a> 。  |
|       | 数据库密码   | 填入该数据库账号的密码。   |
|       | 连接方式    | 根据需求选择非加密连接或SSL安全连接。<br><span style="background-color: #e6f2ff; padding: 5px; border: 1px solid #d9e1f2;"> <span style="font-size: 1.2em; color: #007bff;">?</span> 说明 目前仅中国内地及中国香港地域支持设置连接方式。                 </span>   |

| 类别     | 配置    | 说明  |
|--------|-------|---|
| 目标实例信息 | 实例类型  | 固定为AnalyticDB for PostgreSQL，无需设置。  |
|        | 实例地区  | 购买数据同步实例时选择的目标实例地域信息，不可变更。  |
|        | 实例ID  | 选择目标云原生数据仓库AnalyticDB PostgreSQL实例ID。                                     |
|        | 数据库名称 | 填入同步目标表所属的数据库名称。  |
|        | 数据库账号 | 填入云原生数据仓库AnalyticDB PostgreSQL的数据库账号。权限要求请参见 <a href="#">数据库账号的权限要求</a> 。 |
|        | 数据库密码 | 填入该数据库账号对应的密码。  |

7. 单击页面右下角的授权白名单并进入下一步。

 说明

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置同步策略和同步对象。

1.选择同步通道的源及目标实例
2.选择同步对象
3.预检查

同步初始化:  结构初始化  全量数据初始化  增量数据初始化

注: DTS全量任务运行期间, 不要清理DTS任务启动后源库产生的增量数据日志。源库如果过早清理日志, 可能会导致DTS增量任务失败

目标已存在表的处理模式:  预检查并报错拦截  忽略报错并继续执行

多表旧并:  是  否

同步操作类型:  Insert  Update  Delete

源库对象

若全局搜索, 请先展开树

| Q

- test123
- testdb

全选

已选择对象 (鼠标移到对象行, 点击编辑可修改对象名或过滤条件) 详情点我

| Q

- dtstest

全选

>
<

\* 映射名称更改:  不进行库表名称批量更改  要进行库表名称批量更改

\* 源库、目标库无法连接后的重试时间:  分钟 ?

\* 为目标对象添加引号 ?:  是  否

取消
上一步
下一步
预检查并启动

| 配置          | 说明   |
|-------------|--|
| 同步初始化       | 默认选中结构初始化、全量数据初始化和增量数据初始化。预检查完成后, DTS会将源实例中待同步对象的结构和存量数据同步至目标在目标库, 作为后续增量同步数据的基线数据。  |
| 目标已存在表的处理模式 | <ul style="list-style-type: none"> <li>○ <b>预检查并报错拦截</b>: 检查目标数据库中是否有同名的表。如果目标数据库中没有同名的表, 则通过该检查项目; 如果目标数据库中有同名的表, 则在预检查阶段提示错误, 数据同步作业不会被启动。</li> </ul> <div style="border: 1px solid #0070c0; background-color: #e6f2ff; padding: 5px; margin: 5px 0;"> <p><span style="font-size: x-small; color: #0070c0;">?</span> <b>说明</b> 如果目标库中同名的表不能删除或重命名, 您可以更改该表在目标库中的名称, 详情请参见<a href="#">设置同步对象在目标实例中的名称</a>。</p> </div> <ul style="list-style-type: none"> <li>○ <b>忽略报错并继续执行</b>: 跳过目标数据库中是否有同名表的检查项。</li> </ul> <div style="border: 1px solid #c00000; background-color: #fff9e6; padding: 5px; margin: 5px 0;"> <p><span style="font-size: x-small; color: #c00000;">🔔</span> <b>警告</b> 选择为忽略报错并继续执行, 可能导致数据不一致, 给业务带来风险, 例如:</p> <ul style="list-style-type: none"> <li>■ 表结构一致的情况下, 在目标库遇到与源库主键的值相同的记录, 则会保留目标集群中的该条记录, 即源库中的该条记录不会同步至目标数据库中。</li> <li>■ 表结构不一致的情况下, 可能会导致无法初始化数据、只能同步部分列的数据或同步失败。</li> </ul> </div> |

| 配置        | 说明   |
|-----------|--|
| 多表归并      | <p>选择为是：通常在OLTP场景中，为提高业务表响应速度，通常会做分库分表处理。而在云原生数据库AnalyticDB PostgreSQL中单个数据表可存储海量数据，使用单表查询更加便捷。此类场景中，您可以借助DTS的多表归并功能将源库中多个表结构相同的表（即各分表）同步至云原生数据库AnalyticDB PostgreSQL中的同一个表中。</p> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 10px; margin: 10px 0;"> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>选择源库的多个表后，您需要通过对象名映射功能，将其改为云原生数据库AnalyticDB PostgreSQL中的同一个表名。关于对象名映射功能的介绍，请参见<a href="#">设置同步对象在目标实例中的名称</a>。</li> <li>您需要在云原生数据库AnalyticDB PostgreSQL的同步目标表中增加 <code>__dts_data_source</code> 列（类型为text）来存储数据来源。DTS将以 <code>&lt;dts数据同步实例ID&gt;:&lt;源数据库名&gt;.&lt;源Schema名&gt;.&lt;源表名&gt;</code> 的格式写入列值用于区分表的来源，例如 <code>dts*****:dtstestdata.testschema.customer1</code>。</li> <li>多表归并功能基于任务级别，即不支持基于表级别执行多表归并。如果需要让部分表执行多表归并，另一部分不执行多表归并，您需要创建两个数据同步作业。</li> </ul> </div> <p>选择为否：默认选项。</p> |
| 同步操作类型    | 根据业务选中需要同步的操作类型，默认情况下都处于选中状态。  |
| 选择同步对象    | <p>在源库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。</p> <p>本场景为异构数据库间同步，因此同步对象选择的粒度为表，且其他对象（如视图、触发器、存储过程）不会被同步至目标库。</p> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 10px; margin: 10px 0;"> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>默认情况下，同步对象的名称保持不变。如果您需要同步对象在目标实例上名称不同，请使用对象名映射功能，详情请参见<a href="#">设置同步对象在目标实例中的名称</a>。</li> <li>如果配置多表归并为是，在选择源库的多个表后，您需要通过对象名映射功能，将其改为云原生数据库AnalyticDB PostgreSQL中的同一个表名。</li> </ul> </div>  |
| 为目标对象添加引号 | <p>选择是否需要为目标对象名添加引号。如果选择为是，且存在下述情况，DTS在结构初始化阶段和增量数据迁移阶段会为目标对象添加单引号或双引号：</p> <ul style="list-style-type: none"> <li>源库所属的业务环境对大小写敏感且大小写混用。</li> <li>源表名不是以字母开头，且包含字母、数字或特殊字符以外的字符。</li> </ul> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 10px; margin: 10px 0;"> <p><b>说明</b> 特殊字符仅支持下划线（_），井号（#）和美元符号（\$）。</p> </div> <ul style="list-style-type: none"> <li>待迁移的Schema、表或列名称是目标库的关键字、保留字或非法字符。</li> </ul> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 10px; margin: 10px 0;"> <p><b>说明</b> 如果选择添加引号，在数据同步完成后，您需使用带引号的目标对象名进行查询。</p> </div>   |
| 映射名称更改    | 如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见 <a href="#">库表列映射</a> 。   |

| 配置            | 说明  |
|---------------|---|
| 源、目标库无法连接重试时间 | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <p><b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p> |

9. 设置待同步的表在云原生数据仓库AnalyticDB PostgreSQL中表类型、主键列和分布键信息。

1.选择同步通道的源及目标实例
2.选择同步对象
3.预检查

| 源库名         | 源表名      | 类型(全部) | 主键列       | 分布键       | 定义状态(全部) |
|-------------|----------|--------|-----------|-----------|----------|
| dtstestdata | customer | 维度表    | id        | --        | 已定义      |
| dtstestdata | Student  | 分区表    | StudentID | StudentID | 已定义      |

共有2条, 每页显示: 20 条
« < 1 > »

**说明** 关于主键列和分布键的详细说明，请参见表的约束定义和表分布键定义。

10. 上述配置完成后，单击页面右下角的预检查并启动。

**说明**

- 在同步作业正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步作业。
- 如果预检查失败，单击具体检查项后的，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择确认屏蔽、忽略告警项并重新进行预检查，跳过告警检测项重新进行预检查。

11. 在预检查对话框中显示预检查通过后，关闭预检查对话框，同步作业将正式开始。

12. 等待同步作业的链路初始化完成，直至处于同步中状态。

您可以在数据同步页面，查看数据同步作业的状态。

排序: 默认排序
状态: 全部

| 实例ID/作业名称  | 状态  | 同步概况                                   | 付费方式 | 同步架构(全部) | 操作   |
|--|-----|--|------|----------|--|
| <input type="checkbox"/> hangzhou-hangzhou-small | 同步中 | 耗时: 1376 毫秒<br>速度: 0.00RPS/(0.000MB/s) | 按量付费 | 单向同步     | <input type="button" value="暂停同步"/> <input type="button" value="转包年包月"/> <input type="button" value="升级更多"/> |

共有1条, 每页显示: 20条
« < 1 > »

# 13. PostgreSQL数据同步

## 13.1. 从RDS PostgreSQL同步至云原生数据仓库AnalyticDB PostgreSQL

数据传输服务DTS（Data Transmission Service）支持将RDS PostgreSQL同步至云原生数据仓库AnalyticDB PostgreSQL。通过DTS提供的数据同步功能，可以轻松实现数据的流转，将企业数据集中分析。

### 前提条件

- RDS PostgreSQL中待同步的数据表必须具备主键。
- 已创建目标云原生数据仓库AnalyticDB PostgreSQL实例，如未创建请参见[创建云原生数据仓库AnalyticDB PostgreSQL实例](#)。

### 注意事项

- 一个数据同步作业只能同步一个数据库，如果有多个数据库需要同步，则需要为每个数据库创建数据同步作业。
- 在数据同步的过程中，如果要将在源库中创建的新表作为同步对象，您需要对该表执行如下操作以保障该表数据同步的一致性。

```
ALTER TABLE schema.table REPLICA IDENTITY FULL;
```

- 为保障同步任务的正常进行，目前仅支持RDS PostgreSQL 11进行主备切换，且需设置参数 `rds_failover_slot_mode` 为 `sync`，设置方式，请参见[逻辑复制槽故障转移（Logical Replication Slot Failover）](#)。

 **警告** 自建PostgreSQL和其他版本的RDS PostgreSQL进行主备切换，会导致同步中断。

### 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

### 同步限制

- 不支持结构初始化，即不支持将源库中待同步对象的结构定义（例如表结构）同步至目标库中。
- 同步对象仅支持数据表。
- 不支持BIT、VARBIT、GEOMETRY、UUID、TSQUERY、TSVECTOR、TXID\_SNAPSHOT类型的数据同步。
- 同步过程中，如果对源库中的同步对象执行了DDL操作，需要手动在目标库中执行对应的DDL操作，然后重启数据同步作业。

### 支持的同步语法

仅支持INSERT、UPDATE、DELETE。

### 准备工作

1. 调整源RDS实例的 `wal_level` 参数设置。

 **警告** 修改 `wal_level` 参数后需要重启实例才能生效，请评估对业务的影响，在业务低峰期进行修改。

- i. 登录[RDS管理控制台](#)。
- ii. 在页面左上角，选择实例所在地域。
- iii. 找到目标实例，单击实例ID。

- iv. 在左侧导航栏，单击参数设置。
  - v. 在参数设置页面找到 wal\_level 参数，将参数值改为 logical 。
2. 根据源RDS实例中待同步对象的结构，在目标云原生数据仓库AnalyticDB PostgreSQL中创建相应的数据库、Schema、表等结构信息，详情请参见SQL语法。

### 操作步骤

- 1. 购买数据同步作业，详情请参见[购买流程](#)。

? **说明** 购买时，选择源实例为PostgreSQL、目标实例为AnalyticDB for PostgreSQL，并选择同步拓扑为单向同步。

- 2. 登录[数据传输控制台](#)。
- 3. 在左侧导航栏，单击数据同步。
- 4. 在同步作业列表页面顶部，选择同步的目标实例所属地域。
- 5. 定位至已购买的数据同步实例，单击配置同步链路。
- 6. 配置同步作业的源实例及目标实例信息。

1.选择同步通道的源及目标实例
2.选择同步对象
3.预检查

同步作业名称:

**源实例信息**

实例类型:

实例地区: 华东1 (杭州)

\* 实例ID:  [其他阿里云账号下的RDS实例](#)

\* 数据库名称:

\* 数据库账号:

\* 数据库密码:

**目标实例信息**

实例类型: AnalyticDB for PostgreSQL

实例地区: 华东1 (杭州)

\* 实例ID:

\* 数据库名称:

\* 数据库账号:

\* 数据库密码:

| 类别 | 配置     | 说明   |
|----|--------|--|
| 无  | 同步作业名称 | DTS会自动生成一个同步作业名称，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。 |
|    | 实例类型   | 选择RDS实例。                                       |
|    | 实例地区   | 购买数据同步实例时选择的源实例地域信息，不可变更。                      |
|    | 实例ID   | 选择RDS PostgreSQL实例ID。                          |
|    | 数据库名称  | 填入待同步的表所属的数据库名称。                               |

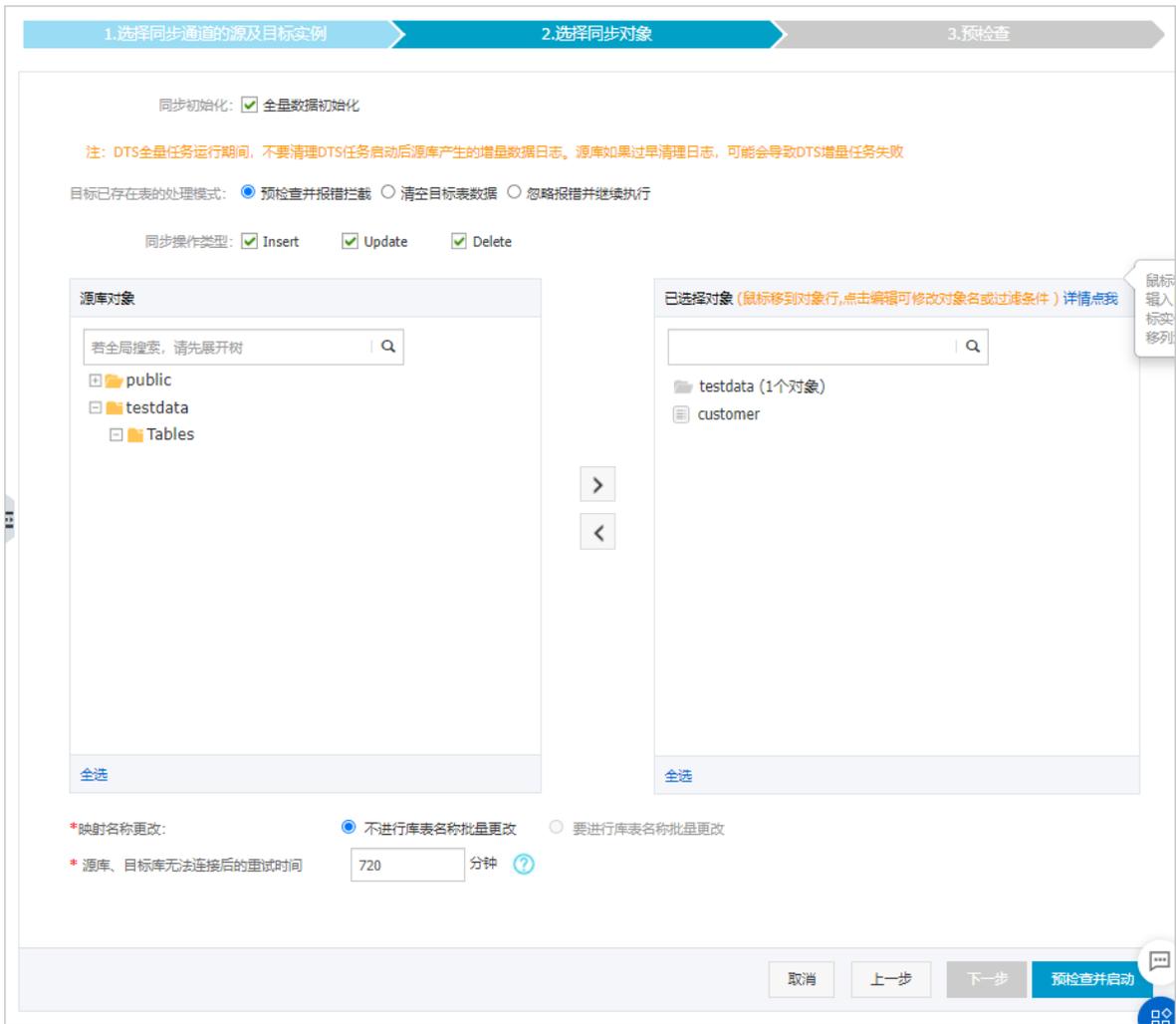
| 源实例信息  | 配置    | 说明  |
|--------|-------|---|
|        | 数据库账号 | 填入RDS PostgreSQL的数据库账号，需为高权限账号。<br><br><span style="background-color: #e6f2ff; padding: 5px;">? 说明 当源库为RDS PostgreSQL 9.4，且仅需同步DML操作时，账号具备replication权限即可。</span>                                       |
|        | 数据库密码 | 填入该数据库账号对应的密码。  |
| 目标实例信息 | 实例类型  | 固定为AnalyticDB for PostgreSQL，无需设置。  |
|        | 实例地区  | 购买数据同步实例时选择的目标实例地域信息，不可变更。  |
|        | 实例ID  | 选择云原生数据仓库AnalyticDB PostgreSQL实例ID。   |
|        | 数据库名称 | 填入同步目标表所属的数据库名称。<br><br><span style="background-color: #e6f2ff; padding: 5px;">? 说明 该库须在云原生数据仓库AnalyticDB PostgreSQL中存在，如不存在请<a href="#">创建数据库</a>。</span>  |
|        | 数据库账号 | 填入云原生数据仓库AnalyticDB PostgreSQL的初始账号，详情请参见 <a href="#">创建数据库账号</a> 。<br><br><span style="background-color: #e6f2ff; padding: 5px;">? 说明 您也可以填入具备RDS_SUPERUSER权限的账号，创建方法请参见<a href="#">用户权限管理</a>。</span> |
|        | 数据库密码 | 填入该数据库账号对应的密码。  |

7. 单击页面右下角的授权白名单并进入下一步。

? 说明

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置同步策略及对象信息。



| 类别     | 配置          | 说明  |
|--------|-------------|---|
| 同步策略配置 | 同步初始化       | 默认情况下，您需要勾选 <b>全量数据初始化</b> 。预检查完成后，DTS会将源实例中待同步对象的存量数据同步至目标实例，作为后续增量同步数据的基线数据。  |
|        | 目标已存在表的处理模式 | <ul style="list-style-type: none"> <li><b>清空目标表的数据</b><br/>在预检查阶段跳过目标表是否为空的检查项目。全量初始化之前将目标表的数据清空。适用于完成同步任务测试后的正式同步场景。</li> <li><b>忽略报错并继续执行</b><br/>在预检查阶段跳过目标表是否为空的检查项目。全量初始化时直接追加迁移数据。适用于多张表同步到一张表的汇总同步场景。</li> </ul>   |
|        | 同步操作类型      | 根据业务需求选择需要同步的操作类型： <div style="border: 1px solid #ccc; padding: 5px; margin: 5px 0;">                         ? 说明 不支持AlterTable。                     </div> <ul style="list-style-type: none"> <li>Insert</li> <li>Update</li> <li>Delete</li> <li>AlterTable</li> </ul> |

| 类别            | 配置 | 说明   |
|---------------|----|--|
| 选择同步对象        | 无  | <p>在源库对象框中单击待同步的表，然后单击  将其移动至已选择对象框。</p> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>同步对象的选择粒度为表。</li> <li>如果需要目标表中的列名称与源表不同，则需要使用DTS的字段映射功能，详情请参见<a href="#">设置同步对象在目标实例中的名称</a>。</li> </ul> |
| 映射名称更改        | 无  | 如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见 <a href="#">库表列映射</a> 。   |
| 源、目标库无法连接重试时间 | 无  | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <p><b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p>  |

9. 上述配置完成后，单击页面右下角的**预检查并启动**。

**说明**

- 在同步作业正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步作业。
- 如果预检查失败，单击具体检查项后的 ，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择**确认屏蔽**、**忽略告警项**并重新进行预检查，跳过告警检测项重新进行预检查。

10. 在**预检查**对话框中显示**预检查通过**后，关闭**预检查**对话框，同步作业将正式开始。

11. 等待同步作业的链路初始化完成，直至处于**同步中**状态。

您可以在**数据同步**页面，查看数据同步作业的状态。



## 13.2. 自建PostgreSQL同步到云原生数据仓库AnalyticDB PostgreSQL

数据传输服务DTS（Data Transmission Service）支持将自建PostgreSQL同步至云原生数据仓库AnalyticDB PostgreSQL。通过DTS提供的数据库同步功能，可以轻松实现数据的流转，将企业数据集中分析。

## 前提条件

- 自建PostgreSQL中待同步的数据表必须具备主键。
- 已创建目标云原生数据仓库AnalyticDB PostgreSQL实例，如未创建请参见[创建云原生数据仓库AnalyticDB PostgreSQL实例](#)。

## 注意事项

- 一个数据库同步作业只能同步一个数据库，如果有多个数据库需要同步，则需要为每个数据库创建数据库同步作业。
- 在数据库同步的过程中，如果要将源库中创建的新表作为同步对象，您需要对该表执行如下操作以保障该表数据同步的一致性。

```
ALTER TABLE schema.table REPLICA IDENTITY FULL;
```

- 为保障同步任务的正常进行，目前仅支持RDS PostgreSQL 11进行主备切换，且需设置参数 `rds_failover_slot_mode` 为 `sync`，设置方式，请参见[逻辑复制槽故障转移（Logical Replication Slot Failover）](#)。

 **警告** 自建PostgreSQL和其他版本的RDS PostgreSQL进行主备切换，会导致同步中断。

## 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

## 同步限制

- 不支持结构初始化，即不支持将源库中待同步对象的结构定义（例如表结构）同步至目标库中。
- 同步对象仅支持数据表。
- 不支持BIT、VARBIT、GEOMETRY、UUID、TSQUERY、TSVECTOR、TXID\_SNAPSHOT类型的数据同步。
- 同步过程中，如果对源库中的同步对象执行了DDL操作，需要手动在目标库中执行对应的DDL操作，然后重启数据库同步作业。

## 支持同步的SQL操作

INSERT、UPDATE、DELETE。

## 准备工作

根据自建PostgreSQL中待同步对象的结构，在目标云原生数据仓库AnalyticDB PostgreSQL中创建相应的数据库、Schema、表等结构信息，详情请参见[SQL语法](#)。

## 操作步骤

1. 购买数据库同步作业，详情请参见[购买流程](#)。

 **说明** 购买时，选择源实例为PostgreSQL、目标实例为AnalyticDB for PostgreSQL，并选择同步拓扑为单向同步。

- 2.
- 3.
4. 在同步作业列表页面顶部，选择数据库同步实例所属地域。
- 5.

6. 配置同步通道的源实例及目标实例信息。

1.选择同步通道的源及目标实例
2.选择同步对象
3.预检查

同步作业名称:

---

**源实例信息**

实例类型:

实例地区:

\* ECS实例ID:

数据库类型:

\* 端口:

\* 数据库名称:

\* 数据库账号:

\* 数据库密码:

---

**目标实例信息**

实例类型:

实例地区:

\* 实例ID:

\* 数据库名称:

\* 数据库账号:

\* 数据库密码:

| 类别    | 配置      | 说明  |
|-------|---------|---|
| 无     | 同步作业名称  | DTS会自动生成一个同步作业名称，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。  |
| 源实例信息 | 实例类型    | 根据源库的部署位置进行选择，本文以ECS上的自建数据库为例介绍配置流程。<br><div style="border: 1px solid #add8e6; padding: 5px; margin: 5px 0;"> <span style="color: #00aaff;">?</span> 说明 当自建数据库为其他实例类型时，您还需要执行相应的准备工作，详情请参见<a href="#">准备工作概览</a>。                     </div> 本文以ECS上的自建数据库为例介绍配置流程。 |
|       | 实例地区    | 购买数据同步实例时选择的源实例地域信息，不可变更。   |
|       | ECS实例ID | 选择自建PostgreSQL数据库所属的ECS实例ID。  |
|       | 数据库类型   | 固定为PostgreSQL。  |
|       | 端口      | 填入自建PostgreSQL数据库的服务端口。   |
|       | 数据库名称   | 填入待同步的表所属的数据库名称。  |
|       | 数据库账号   | 填入自建PostgreSQL的数据库账号。<br><div style="border: 1px solid #add8e6; padding: 5px; margin: 5px 0;"> <span style="color: #00aaff;">?</span> 说明 数据库账号须具备schema的owner权限。                     </div>   |
|       | 数据库密码   | 填入该数据库账号对应的密码。  |

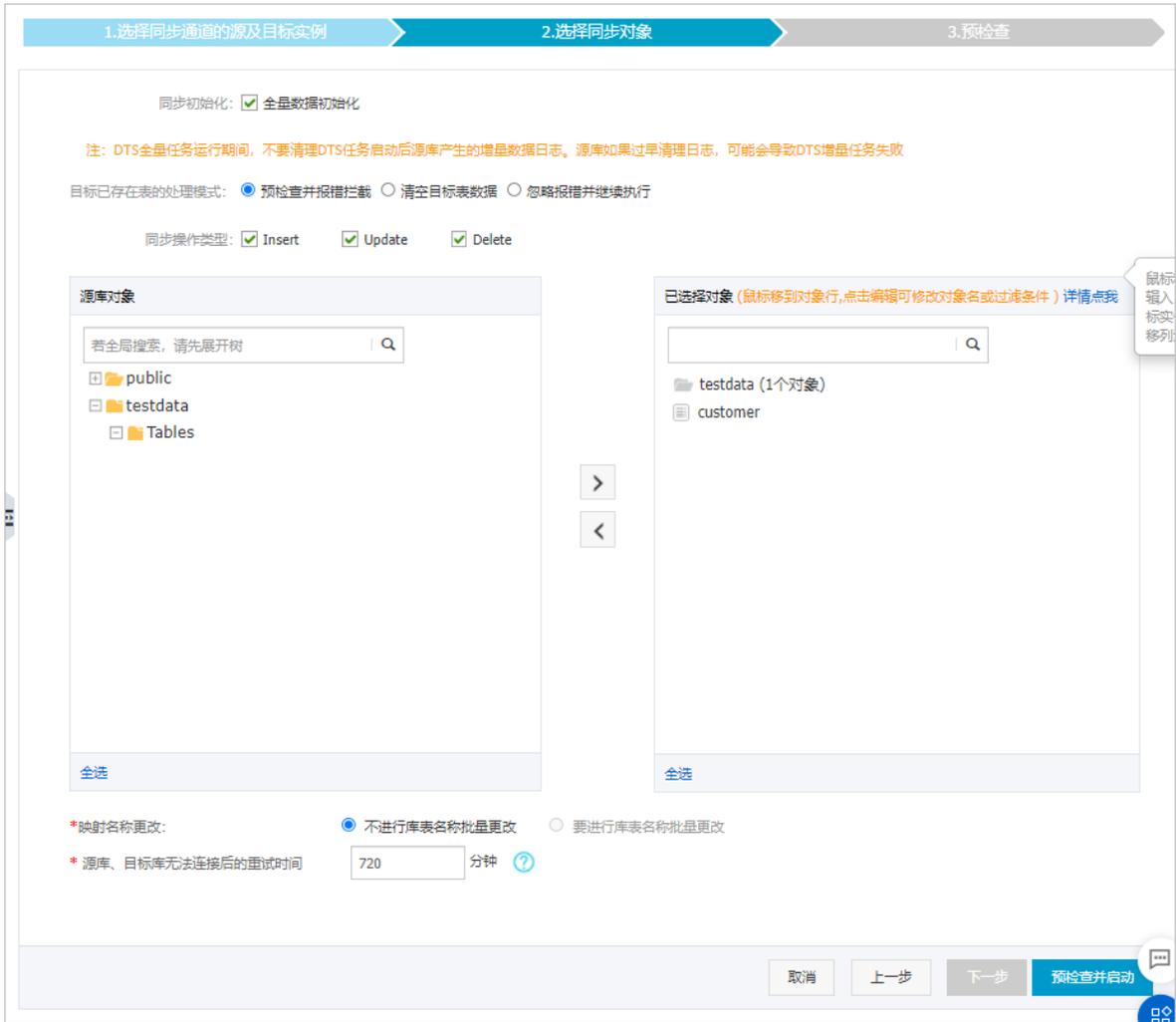
| 类别     | 配置    | 说明   |
|--------|-------|--|
| 目标实例信息 | 实例类型  | 固定为AnalyticDB for PostgreSQL，无需设置。   |
|        | 实例地区  | 购买数据同步实例时选择的目标实例地域信息，不可变更。   |
|        | 实例ID  | 选择云原生数据仓库AnalyticDB PostgreSQL实例ID。  |
|        | 数据库名称 | 填入同步目标表所属的数据库名称。   |
|        | 数据库账号 | 填入云原生数据仓库AnalyticDB PostgreSQL的初始账号，详情请参见 <a href="#">创建数据库账号</a> 。<br><br> <b>说明</b> 您也可以填入具备RDS_SUPERUSER权限的账号，创建方法请参见 <a href="#">用户权限管理</a> 。 |
|        | 数据库密码 | 填入该数据库账号对应的密码。   |

7. 单击页面右下角的授权白名单并进入下一步。

 **说明**

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置同步策略及对象信息。



| 类别     | 配置          | 说明  |
|--------|-------------|---|
| 同步策略配置 | 同步初始化       | 默认情况下，您需要勾选 <b>全量数据初始化</b> 。预检查完成后，DTS会将源实例中待同步对象的存量数据同步至目标实例，作为后续增量同步数据的基线数据。  |
|        | 目标已存在表的处理模式 | <ul style="list-style-type: none"> <li>清空目标表的数据<br/>在预检查阶段跳过目标表是否为空的检查项目。全量初始化之前将目标表的数据清空。适用于完成同步任务测试后的正式同步场景。</li> <li>忽略报错并继续执行<br/>在预检查阶段跳过目标表是否为空的检查项目。全量初始化时直接追加迁移数据。适用于多张表同步到一张表的汇总同步场景。</li> </ul>   |
|        | 同步操作类型      | 根据业务需求选择需要同步的操作类型： <div style="border: 1px solid #ccc; padding: 5px; margin: 5px 0;">             ? 说明 不支持AlterTable。           </div> <ul style="list-style-type: none"> <li>Insert</li> <li>Update</li> <li>Delete</li> <li>AlterTable</li> </ul> |

| 类别            | 配置 | 说明   |
|---------------|----|--|
| 选择同步对象        | 无  | <p>在源库对象框中单击待同步的表，然后单击  将其移动至已选择对象框。</p> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>同步对象的选择粒度为表。</li> <li>如果需要目标表中的列名称与源表不同，则需要使用DTS的字段映射功能，详情请参见<a href="#">设置同步对象在目标实例中的名称</a>。</li> </ul> |
| 映射名称更改        | 无  | 如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见 <a href="#">库表列映射</a> 。   |
| 源、目标库无法连接重试时间 | 无  | <p>当源、目标库无法连接时，DTS默认重试720分钟（即12小时），您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库，同步任务将自动恢复。否则，同步任务将失败。</p> <p><b>说明</b> 由于连接重试期间，DTS将收取任务运行费用，建议您根据业务需要自定义重试时间，或者在源和目标库实例释放后尽快释放DTS实例。</p>  |

9. 上述配置完成后，单击页面右下角的**预检查并启动**。

**说明**

- 在同步作业正式启动之前，会先进行预检查。只有预检查通过后，才能成功启动同步作业。
- 如果预检查失败，单击具体检查项后的 ，查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项，您也可以选择**确认屏蔽**、**忽略告警项**并重新进行预检查，跳过告警检测项重新进行预检查。

10. 在**预检查**对话框中显示**预检查通过**后，关闭**预检查**对话框，同步作业将正式开始。

11. 等待同步作业的链路初始化完成，直至处于**同步中**状态。

您可以在**数据同步**页面，查看数据同步作业的状态。



# 14.TiDB数据同步

## 14.1. 从自建TiDB同步至云原生数据仓库

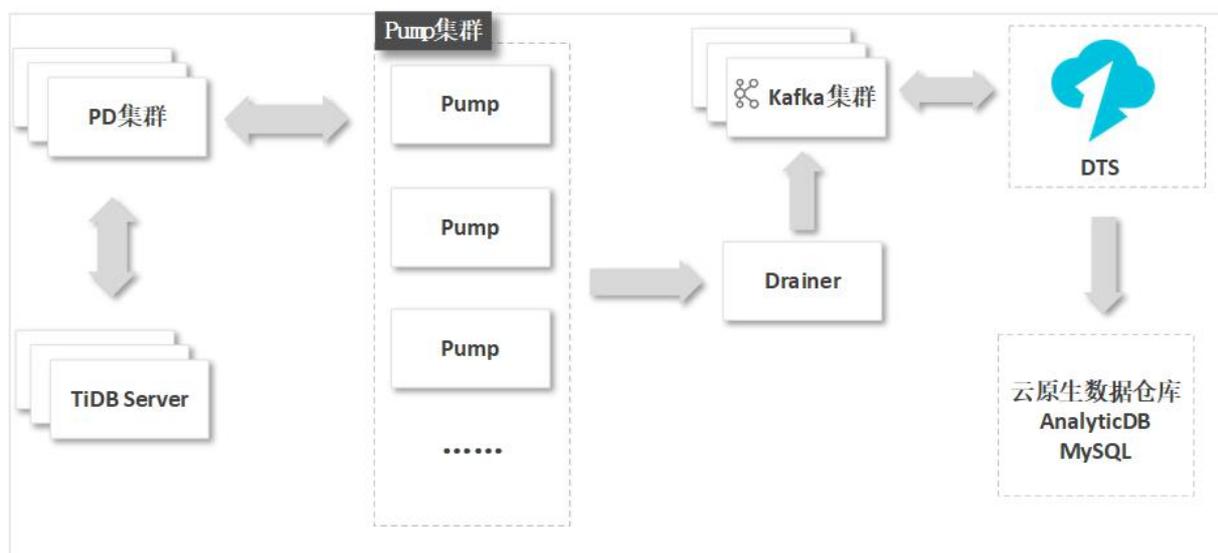
### AnalyticDB MySQL

本文介绍如何使用数据传输服务DTS（Data Transmission Service），结合Kafka集群与TiDB数据库的Pump、Drainer组件，将自建TiDB数据库同步至云原生数据仓库AnalyticDB MySQL。

#### 前提条件

- 已创建云原生数据仓库AnalyticDB MySQL，详情请参见[创建云原生数据仓库AnalyticDB MySQL](#)。
- 确保目标云原生数据仓库AnalyticDB MySQL具备充足的存储空间。

#### 背景信息



由于TiDB的Binlog格式和实现机制与MySQL数据库存在一定区别，为实现增量数据同步，同时减少对源数据库的改动，您需要部署Kafka集群以及TiDB数据库的Pump和Drainer组件。

由Pump组件实时记录TiDB产生的Binlog并提供给Drainer组件，然后由Drainer组件将获取到的Binlog写入到下游的Kafka集群。DTS在执行增量数据同步时将从Kafka集群中获取对应的数据并实时同步至目标数据库（例如云原生数据仓库AnalyticDB MySQL）。

#### 注意事项

- DTS在执行全量数据初始化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。因此您需要在执行数据同步前评估源库和目标库的性能，同时建议您在业务低峰期执行数据同步（例如源库和目标库的CPU负载在30%以下）。
- 请勿在数据同步时，对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具执行在线DDL变更，否则会导致同步失败。
- 由于云原生数据仓库AnalyticDB MySQL本身的使用限制，当云原生数据仓库AnalyticDB MySQL集群中的节点磁盘空间使用量超过80%，该集群将被锁定。请提前根据待同步的对象预估所需空间，确保目标集群具备充足的存储空间。
- 暂不支持同步前缀索引，如果源库存在前缀索引可能导致数据同步失败。

#### 费用说明

| 同步类型          | 链路配置费用                          |
|---------------|---------------------------------|
| 库表结构同步和全量数据同步 | 不收费。                            |
| 增量数据同步        | 收费，详情请参见 <a href="#">产品定价</a> 。 |

## 支持同步的SQL操作

- DDL操作：CREATE TABLE、DROP TABLE、RENAME TABLE、TRUNCATE TABLE、ADD COLUMN、DROP COLUMN
- DML操作：INSERT、UPDATE、DELETE

 **说明** 如果在数据同步的过程中变更了源表的字段类型，同步作业将报错并中断。您可以[提交工单](#)处理或手动修复，详情请参见[修复因变更字段类型导致的同步失败](#)。

## 准备工作

 **说明** 为减少网络延迟对数据同步的影响，Pump组件、Drainer组件和Kafka集群所部署的服务器需要与源库所属的服务器在同一内网中。

1. 部署Pump和Drainer组件，详情请参见[TiDB Binlog集群部署](#)。
2. 修改Drainer组件的配置文件，设置输出为Kafka，详情请参见[Kafka自定义开发](#)。
3. 选择下述方法准备Kafka集群：
  - 部署自建Kafka集群，详情请参见[Apache Kafka官网](#)。

 **警告** 为保障Kafka集群可正常接收到TiDB产生的较大的Binlog数据，请适当将Broker组件中的 `message.max.bytes`、`replica.fetch.max.bytes` 参数以及Consumer组件中的 `fetch.message.max.bytes` 参数对应的值调大，详细说明请参见[Kafka配置说明](#)。

- 使用[阿里云消息队列Kafka版（MQ for Apache Kafka）](#)，详情请参见[阿里云消息队列Kafka版快速入门](#)。

 **说明** 为保障正常通信和减少网络延迟对增量数据同步的影响，部署阿里云消息队列Kafka实例时，需配置和源库服务器相同的专有网络。

4. 在自建Kafka集群或阿里云消息队列Kafka实例中创建Topic。
5. 将DTS服务器的IP地址段加入至TiDB数据库的白名单安全设置中，具体IP地址段信息请参见[迁移、同步或订阅本地数据库时需添加的IP白名单](#)。

## 操作步骤

1. 购买数据同步作业，详情请参见[购买流程](#)。

 **说明** 购买时，选择源实例为TiDB，目标实例为AnalyticDB MySQL。

2. 登录[数据传输控制台](#)。
3. 在左侧导航栏，单击[数据同步](#)。
4. 在同步作业列表页面顶部，选择数据同步实例所属地域。
5. 定位至已购买的数据同步实例，单击[配置同步链路](#)。
6. 配置同步通道的源实例及目标实例信息。
  - i. 配置同步任务的名称和源实例信息。

同步作业名称:

---

**源实例信息**

实例类型:

实例地区:

\* ECS实例ID:

数据库类型:

\* 端口:

\* 数据库账号:

\* 数据库密码:

\* Kafka集群类型:

\* 实例地区:

\* ECS实例ID:

\* Kafka端口:

Kafka集群账号:  非必填项

Kafka集群密码:  非必填项

\* Topic:

请先点击右侧按钮, 获取Topic列表后选择具体的Topic

\* Kafka版本:

\* Kafka集群连接方式:  非加密连接  SCRAM-SHA-256

| 配置     | 说明  |
|--------|---|
| 同步任务名称 | 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。  |
| 实例类型   | <p>根据源库的部署位置进行选择, 本文以ECS上的自建数据库为例介绍配置流程。</p> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin-top: 5px;"> <p><span style="color: blue;">?</span> 说明 当自建数据库为其他实例类型时, 您还需要执行相应的准备工作, 详情请参见<a href="#">准备工作概览</a>。</p> </div> |
| 实例地区   | 购买数据同步实例时选择的源实例地域信息, 不可变更。  |
| 数据库类型  | 固定为TiDB。  |
| 端口     | 填入TiDB数据库的服务端口, 默认为4000。  |
| 数据库账号  | 填入TiDB数据库账号, 需具备SHOW VIEW和待同步对象的SELECT权限, 授权方法请参见 <a href="#">TiDB权限管理</a> 。  |
| 数据库密码  | 填入该数据库账号的密码。  |

| 配置          | 说明  |
|-------------|---|
| Kafka集群类型   | <p>根据Kafka的部署位置进行选择，本文以ECS上的自建数据库为例介绍配置流程。当自建Kafka为其他实例类型时，您还需要执行相应的准备工作，详情请参见<a href="#">准备工作概览</a>。</p> <div style="border: 1px solid #ccc; background-color: #e6f2ff; padding: 5px; margin-top: 10px;"> <p> <b>说明</b> 由于DTS暂时不支持直接选择阿里云消息队列Kafka版，如果您使用的是阿里云消息队列Kafka实例，此处需将其作为自建Kafka来配置，即选择为通过专线/VPN网关/智能接入网关接入的自建数据库，然后选择阿里云消息队列Kafka实例所属的专有网络。</p> </div> |
| 实例地区        | 和源库的实例地区保持一致，不可变更。  |
| ECS实例ID     | 选择自建Kafka所属的ECS实例ID。  |
| Kafka端口     | 自建Kafka的服务端口，默认为9092。   |
| Kafka集群账号   | 填入自建Kafka的用户名，如自建Kafka未开启验证可不填写。  |
| Kafka集群密码   | 填入该用户的密码，如自建Kafka未开启验证可不填写。   |
| Topic       | 单击右侧的获取Topic列表，然后在下拉框中选择具体的Topic。   |
| Kafka版本     | 根据自建Kafka的版本进行选择。   |
| Kafka集群连接方式 | 根据业务及安全需求，选择非加密连接或SCRAM-SHA-256。  |

ii. 配置同步任务的目标实例信息。

**目标实例信息**

实例类型: ADS

实例地区: 华东1 (杭州)

\* 版本:  3.0

\* 数据库:

\* 数据库账号:

\* 数据库密码:

| 配置    | 说明  |
|-------|---|
| 实例类型  | 固定为ADS。   |
| 实例地区  | 购买数据同步实例时选择的目标实例地域信息，不可变更。  |
| 版本    | 固定为3.0。   |
| 数据库   | 选择目标云原生数据仓库AnalyticDB MySQL的集群ID。   |
| 数据库账号 | 填入云原生数据仓库AnalyticDB MySQL的数据库账号，需具备目标库的读写权限，授权方法请参见 <a href="#">创建数据库账号</a> 。 |
| 数据库密码 | 填入该数据库账号对应的密码。  |

7. 单击页面右下角的授权白名单并进入下一步。

**说明**

- 如果源或目标数据库是阿里云数据库实例（例如RDS MySQL、云数据库MongoDB版等）或ECS上的自建数据库，DTS会自动将对应地区DTS服务的IP地址添加到阿里云数据库实例的白名单或ECS的安全规则中，您无需手动添加，请参见[DTS服务器的IP地址段](#)。
- DTS任务完成或释放后，建议您手动删除添加的DTS服务器IP地址段。

8. 配置同步策略及对象信息。

| 配置    | 说明   |
|-------|--|
| 同步初始化 | 默认情况下，您需要同时选中 <b>结构初始化</b> 和 <b>全量数据初始化</b> 。预检查完成后，DTS会将源实例中待同步对象的结构及数据在目标集群中初始化，作为后续增量同步数据的基线数据。 |

| 配置          | 说明   |
|-------------|--|
| 目标已存在表的处理模式 | <ul style="list-style-type: none"> <li>◦ <b>预检查并报错拦截</b>：检查目标数据库中是否有同名的表。如果目标数据库中没有同名的表，则通过该检查项目；如果目标数据库中有同名的表，则在预检查阶段提示错误，数据同步作业不会被启动。</li> </ul> <p><b>说明</b> 如果目标库中同名的表不方便删除或重命名，您可以更改该表在目标库中的名称，详情请参见<a href="#">设置同步对象在目标实例中的名称</a>。</p> <ul style="list-style-type: none"> <li>◦ <b>忽略报错并继续执行</b>：跳过目标数据库中是否有同名表的检查项。</li> </ul> <p><b>警告</b> 选择为忽略报错并继续执行，可能导致数据不一致，给业务带来风险，例如：</p> <ul style="list-style-type: none"> <li>■ 表结构一致的情况下，在目标库遇到与源库主键的值相同的记录，则会保留目标集群中的该条记录，即源库中的该条记录不会同步至目标数据库中。</li> <li>■ 表结构不一致的情况下，可能会导致无法初始化数据、只能同步部分列的数据或同步失败。</li> </ul> |
| 多表归并        | <ul style="list-style-type: none"> <li>◦ 选择为是：DTS将在每个表中增加 <code>__dts_data_source</code> 列来存储数据来源，且不再支持DDL同步。</li> <li>◦ 选择为否：默认选项，支持DDL同步。</li> </ul> <p><b>说明</b> 多表归并功能基于任务级别，即不支持基于表级别执行多表归并。如果需要让部分表执行多表归并，另一部分不执行多表归并，您可以创建两个数据同步作业。</p>   |
| 同步操作类型      | <p>根据业务选中需要同步的操作类型，支持的同步操作详情请参见<a href="#">支持同步的SQL操作</a>，默认情况下都处于选中状态。</p>  |
| 选择同步对象      | <p>在源库对象框中单击待同步的对象，然后单击  图标将其移动至已选择对象框。</p> <p>同步对象的选择粒度为库、表。</p> <p><b>说明</b></p> <ul style="list-style-type: none"> <li>◦ 如果选择整个库作为同步对象，那么该库中所有对象的结构变更操作会同步至目标库。</li> <li>◦ 如果选择某个表作为同步对象，那么只有这个表的ADD COLUMN操作会同步至目标库。</li> <li>◦ 默认情况下，同步对象的名称保持不变。如果您需要同步对象在目标集群上名称不同，请使用对象名映射功能，详情请参见<a href="#">设置同步对象在目标实例中的名称</a>。</li> </ul>   |
| 映射名称更改      | <p>如需更改同步对象在目标实例中的名称，请使用对象名映射功能，详情请参见<a href="#">库表映射</a>。</p>   |

| 配置                             | 说明  |
|--------------------------------|---|
| 源表DMS_ONLINE_DDL过程中是否复制临时表到目标库 | <p>如源库使用<a href="#">数据管理DMS (Data Management Service)</a> 执行Online DDL变更, 您可以选择是否同步Online DDL变更产生的临时表数据。</p> <ul style="list-style-type: none"> <li>是: 同步Online DDL变更产生的临时表数据。                     <div style="border: 1px solid #ccc; padding: 5px; margin: 5px 0;"> <p><span>?</span> 说明 Online DDL变更产生的临时表数据过大, 可能会导致同步任务延迟。</p> </div> </li> <li>否: 不同步Online DDL变更产生的临时表数据, 只同步源库的原始DDL数据。                     <div style="border: 1px solid #ccc; padding: 5px; margin: 5px 0;"> <p><span>?</span> 说明 该方案会导致目标库锁表。</p> </div> </li> </ul> |
| 源、目标库无法连接重试时间                  | <p>当源、目标库无法连接时, DTS默认重试720分钟(即12小时), 您也可以自定义重试时间。如果DTS在设置的时间内重新连接上源、目标库, 同步任务将自动恢复。否则, 同步任务将失败。</p> <div style="border: 1px solid #ccc; padding: 5px; margin: 5px 0;"> <p><span>?</span> 说明 由于连接重试期间, DTS将收取任务运行费用, 建议您根据业务需要自定义重试时间, 或者在源和目标库实例释放后尽快释放DTS实例。</p> </div>  |

9. 上述配置完成后, 单击页面右下角的下一步。

10. 设置待同步的表在目标库中类型。

1.选择同步通道的源及目标实例
2.ADS账号授权
3.选择同步对象
4.预检查

| ADB库名        | ADB表名    | 类型(全部) ▾ | 主键列     | 分布列     | 定义状态(全部) ▾ |
|--------------|----------|----------|---------|---------|------------|
| tidbtestdata | customer | 分区表 ▾    | id      | id      | 已定义        |
| tidbtestdata | order    | 分区表 ▾    | orderid | orderid | 已定义        |

全部设置为分区表 全部设置为维度表

共有2条, 每页显示: 20 条
<< < 1 > >>

? 说明 选择了结构初始化后, 您需要定义待同步的表在云原生数据仓库AnalyticDB MySQL中的类型、主键列、分区列等信息, 详情请参见[CREATE TABLE操作手册](#)。

11. 上述配置完成后, 单击页面右下角的预检查并启动。

? 说明

- 在同步作业正式启动之前, 会先进行预检查。只有预检查通过后, 才能成功启动同步作业。
- 如果预检查失败, 单击具体检查项后的 ?, 查看失败详情。
  - 您可以根据提示修复后重新进行预检查。
  - 如无需修复告警检测项, 您也可以选择**确认屏蔽**、**忽略告警项**并重新进行预检查, 跳过告警检测项重新进行预检查。

12. 在预检查对话框中显示预检查通过后, 关闭预检查对话框, 同步作业将正式开始。

13. 等待同步作业的链路初始化完成, 直至处于同步中状态。

您可以在数据同步页面，查看数据同步作业的状态。

| 实例ID/作业名称  | 状态  | 同步概况                                  | 付费方式 | 同步架构(全部) | 操作  |
|--|-----|---------------------------------------|------|----------|---|
| <input type="checkbox"/> hangzhou-hangzhou-small | 同步中 | 延时: 1376 毫秒<br>速率: 0.00RPS(0.000MB/s) | 按量付费 | 单向同步     | <a href="#">暂停同步</a>   <a href="#">转包年包月</a>   <a href="#">升级更多</a> |

[暂停同步](#) [释放同步](#) 共有1条， 每页显示：20条