

数据传输服务DTS

用户指南

用户指南

访问控制

DTS支持RAM主子账号

数据传输服务DTS已经支持阿里云主子账号体系。用户即可以使用阿里云主账号购买并管理DTS任务，同时也可以将权限授权给子账号，让子账号有权限购买并管理DTS任务。

本文简单介绍，DTS主子账号使用方式。

子账号定义

DTS目前只支持读写和只读两种访问子账号，尚不支持API粒度的访问授权。

DTS支持的读写和只读策略属于RAM系统授权策略，策略分别为：

- 读写策略

读写策略名称为：AliyunDTSFullAccess，其策略定义为：

```
{
  "Version": "1",
  "Statement": [
    {
      "Action": "dts:*",
      "Resource": "*",
      "Effect": "Allow"
    },
    {
      "Action": "ram:PassRole",
      "Resource": "*",
      "Effect": "Allow",
      "Condition": {
        "StringEquals": {
          "acs:Service": "dts.aliyuncs.com"
        }
      }
    }
  ]
}
```

读写策略拥有DTS所有读写权限，一旦授权某个子账号读写策略，那么这个子账号可以进行从DTS实例的购买、配置到管理的全生命周期管理。

- 只读策略

只读策略的策略名字为：AliyunDTSReadOnlyAccess, 其策略定义为：

```
{
  "Version": "1",
  "Statement": [
    {
      "Action": "dts:Describe*",
      "Resource": "*",
      "Effect": "Allow"
    }
  ]
}
```

读写策略拥有DTS所有读权限，一旦授权某个子账号只读策略，那么这个子账号可以查看主账号下所有DTS任务的任务详情配置等信息，但是不能进行变更操作。这里面的变更操作主要包括：购买、配置、启动、修改、启动、暂停、结束及释放。

子账号授权

为了安全起见，您可以将DTS的任务创建及管理的工作授权给某个子账号进行，而不将主账号提供给公司每个同学。本小节简单介绍如何创建可使用DTS服务的子账号。

(1) 创建子用户

如果尚未创建子账号，那么先创建一个子账号，具体创建流程参考RAM使用手册

(2) 用户授权

当用户创建完成后，即开始策略授权，将DTS系统策略授权给子账号。进入RAM控制台，进入用户管理界面，点击要授权子账号后面的**授权**入口，开始角色策略授权。



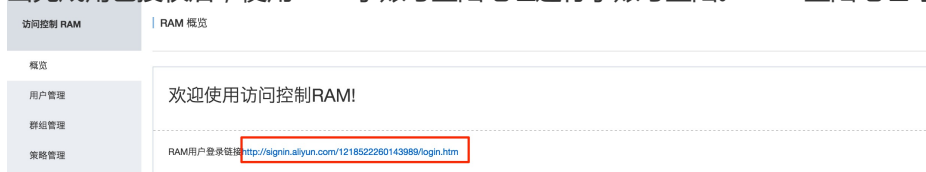
在策略选择步骤中，搜索DTS相关策略。



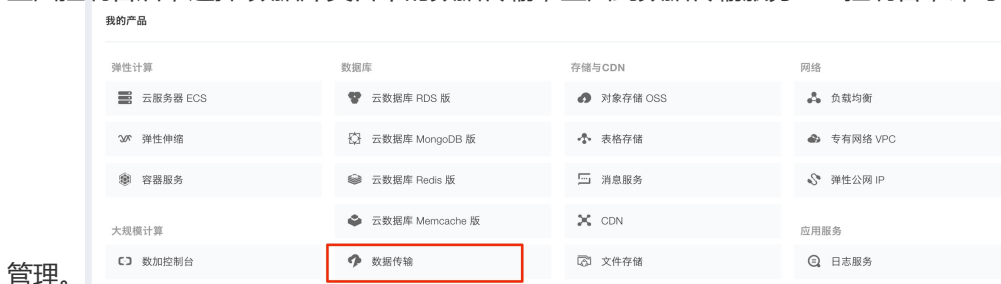
从左侧的可选授权列表中，选择要授权的DTS系统策略,添加到已选授权策略列表中，并点击确认，完成角色授权。

(3) 访问DTS服务

当完成角色授权后，使用RAM子账号登陆地址进行子账号登陆。RAM登陆地址可以在RAM概览页中查看。



登陆控制台后，选择 数据库类目中的数据传输，登陆到数据传输服务DTS控制台，即可开始DTS任务的创建及



管理。

至此，完成DTS子账号授权及DTS控制台登陆操作。

为提高安全性，用户如果在第一次使用DTS服务之前，DTS要求用户将名称为：**AliyunDTSDefaultRole** 的系统默认角色授权给DTS 服务账号。

用户将这个角色授权给DTS后，DTS才能访问用户的RDS、ECS等云资源，进行后续的任务配置，数据传输。

角色定义

AliyunDTSDefaultRole 的角色权限，主要包括访问用户RDS、ECS及Datahub相关API的权限，以便于进行后续的任务配置及数据同步。具体权限定义如下：

```
{
  "Version": "1",
  "Statement": [
    {
```



```

"Action": [
  "rds:Describe*",
  "rds:CreateDBInstance",
  "rds:CreateAccount",
  "rds:CreateDataBase",
  "rds:ModifySecurityIps",
  "rds:GrantAccountPrivilege"
],
"Resource": "*",
"Effect": "Allow"
},
{
  "Action": [
    "ecs:DescribeInstances",
    "ecs:DescribeSecurityGroup",
    "ecs:JoinSecurityGroup",
    "ecs:RevokerSecurityGroup"
  ],
  "Resource": "*",
  "Effect": "Allow"
},
{
  "Action": "dhs:*",
  "Effect": "Allow",
  "Resource": "*"
}
]
}

```

您也可以到RAM控制台 - 角色管理 中查看这个角色的权限定义。请勿修改权限定义，否则会导致任务配置异常。

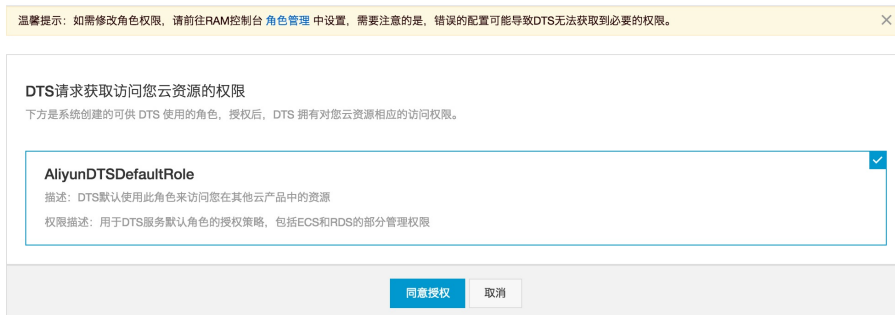
角色授权流程

当您第一次登陆DTS控制台时，并且未将服务角色AliyunDTSDefaultRole授权给DTS的服务角色时，DTS会弹出如下的授权界面，要求对DTS服务角色进行角色授权。



此时，需单击 **前往RAM角色授权** 跳转到授权页面，将服务角色授权给DTS服务账号。

云资源访问授权



单击 **同意授权**，即将服务角色授权给DTS服务账号。此时即可开始使用DTS进行传输任务的创建及管理。

数据传输DTS支持两个阿里云账号下的实例进行数据迁移同步。为了实现这个功能，源实例所属阿里云账号需将相关授权给目标实例所属阿里云账号，然后使用目标实例所属阿里云账号登录DTS控制台进行任务配置。本小节介绍，如何通过RAM进行跨账号授权，以实现不同阿里云账号下的实例间的数据迁移同步。

DTS支持的跨账号迁移同步功能

DTS能够支持的跨账号迁移同步功能包括：

数据迁移

- RDS实例->RDS实例
- RDS实例->DRDS实例
- RDS实例->PetaData实例
- RDS实例->OceanBase实例
- RDS实例->ECS自建数据库
- RDS实例->有公网IP的自建数据库

数据同步

- RDS实例->RDS实例
- RDS实例->MaxCompute(原ODPS)实例
- RDS实例->Datahub(流计算使用)实例

跨账号传输配置信息

进行跨账号数据迁移同步时，源RDS实例所属阿里云账号需要通过RAM的跨主账号授权，授权目标实例所属阿里云账号可以在DTS中访问自己的相关云资源。使用DTS配置跨账号数据迁移同步时，源实例除了实例的连接信息，还需要配置授权相关信息，具体如下：

实例类型: RDS For MySQL

实例地区: 华东 1

*RDS所属阿里云账号:

*角色名称:

*RDS实例ID:

[跨账号角色授权](#)

当前登录账号下的RDS实例

上图中，RDS所属阿里云账号为源RDS所属阿里云账号的账号ID，可以到账号管理的 安全设置 界面获取



角色名称为源RDS所属阿里云账号给目标实例所属阿里云账号进行授权角色的角色名称。下面会介绍如何创建这个授权角色。

跨账号角色授权

本小节介绍如何使用RAM进行跨阿里云账号授权，授权另外一个阿里云账号在DTS中访问自己的云资源。下面就以需要将账号A下面的RDS实例迁移到账号B下的云资源为例，介绍角色授权过程。

角色创建

1. 使用账号A登录RAM管理控制台，进入角色管理界面,点击页面右上角的 **新建角色**，开始创建跨账号授权角色。
2. 第一步的角色类型，选择用户角色
3. 受信云账号，选择 **其他云账号**，同时，**受信云账号ID** 配置目标实例所属阿里云账号即账号B的账号ID

创建角色

1: 选择角色类型 2: 填写类型信息 3: 配置角色基本信息 4: 创建成功

选择受信云账号，受信云帐号将可以使用此角色来访问您的云资源。

选择云账号 当前云账号 其他云账号

*受信云账号ID:

可以访问[账户管理->安全设置](#)获取帐号ID。

(4) 第四步，配置角色名称

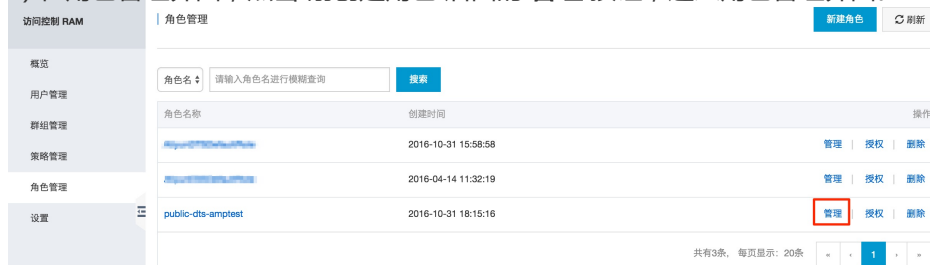
，这个名称就是DTS同步作业配置过程中，需要填写的角色名称。

角色授权

角色创建完成后，需要修改角色授权策略，包括：

(1) **受信云账号,限制受信云账号 只能在DTS 控制台访问自己的云资源。** 具体修改步骤如下：

1) 在角色管理界面，点击 刚创建角色 后面的 **管理** 按钮，进入角色管理界面。



2) 在角色管理界面，点击右

上角的 **编辑基本信息**，进入角色编辑框，在编辑框中，修改Principal,添加service定义：

```
"Service": [
  "受信阿里云账号ID@dts.aliyuncs.com"
]
```

受信云账号的账号ID,即最后配置DTS同步作业的阿里云账号ID。dts.aliyuncs.com 为DTS服务代号。假配置DTS同步作业的阿里云账号ID为：121852226014398，那么service定义为：

```
"Service": [
  "121852226014398@dts.aliyuncs.com"
]
```

所以，完整的角色定义如下：

```
{
  "Statement": [
    {
      "Action": "sts:AssumeRole",
      "Effect": "Allow",
      "Principal": {
        "RAM": [
          "acs:ram::1218522260143989:root"
        ],
        "Service": [
          "1218522260143989@dts.aliyuncs.com"
        ]
      }
    }
  ],
  "Version": "1"
}
```

(2)角色授权，授权角色能够源实例所属云账号下的资源

当受信云账号B配置完成后，需要修改角色的授权策略，将源账号A下的部分云资源的访问权限授权给角色，这

样DTS可以扮演这个角色访问A账号下的云实例。修改步骤如下：

1) 在角色管理界面，点击 刚创建角色 后面的 授权 按钮，进入角色授权界面。



2) 在精确策略搜索框中，搜索AliyunDTSRolePolicy，将这个系统策略授权给角色。



至此，完成跨账号授权角色的创建及授权。后面配置跨账号迁移同步任务时，源实例信息中的，角色名称 即为上面创建的角色名称。

数据传输服务DTS已经支持阿里云RAM主子账号体系。除了可以使用子账号进行任务的创建管理，数据订阅SDK也可以使用子账号的Access Key/Access Secret 进行数据的实时订阅。本文介绍SDK如何使用子账号的Access Key/Access Secret 进行数据的实时订阅。

子账号需要授权DTS读写策略

DTS目前只支持读写和只读两种访问子账号。读写策略名称为：AliyunDTSFullAccess，只读策略的策略名字为：AliyunDTSReadOnlyAccess。为了让子账号能够使用SDK订阅实时数据，需要给子账号授权读写策略AliyunDTSFullAccess。

如果子账号没有DTS的任何权限，或只有只读权限AliyunDTSReadOnlyAccess，那么使用子账号AccessKey/AccessSecret时，启动SDK会报如下错误：

```
SLF4J: Failed to load class "org.slf4j.impl.StaticLoggerBinder".
SLF4J: Defaulting to no-operation (NOP) logger implementation
SLF4J: See http://www.slf4j.org/codes.html#StaticLoggerBinder for further details.
Exception in thread "main" java.lang.Exception: get guid info failed
    at com.aliyun.drc.regionmanager.RegionRouter.getRegionRouterInfo(RegionRouter.java:40)
    at com.aliyun.drc.clusterclient.partition.PartitionPool.init(PartitionPool.java:94)
    at com.aliyun.drc.clusterclient.impl.DefaultClusterClientImpl.start(DefaultClusterClientImpl.java:41)
    at com.aliyun.drc.clusterclient.DefaultClusterClient.start(DefaultClusterClient.java:23)
    at com.aliyun.drc.clusterclient.MainClass.main(MainClass.java:40)
```

如上图所示，会报 get guid info failed 的错误，即DTS会认为要订阅的通道不属于SDK配置的

AccessKey/AccessSecret 对应的云账号。

子账号创建及授权

本小节简单介绍，如何在RAM控制台创建一个可以通过数据订阅SDK订阅实时增量数据的子账号。

(1) 创建子账号

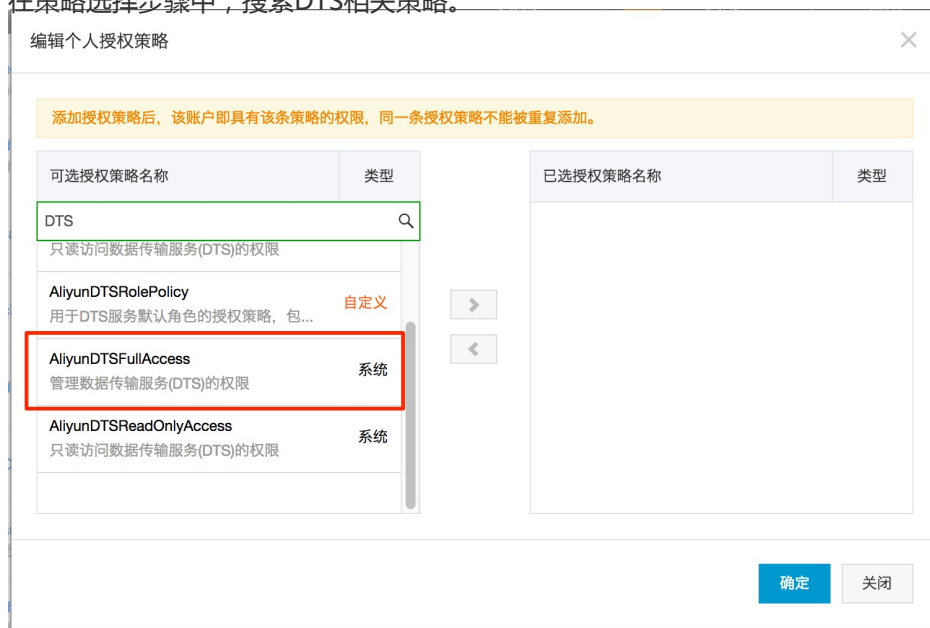
如果尚未创建子账号，那么先创建一个子账号，特别注意需要给子账号创建一个AccessKey/AccessSecret,具体创建流程参考RAM使用手册

(2) 账号授权

当用户创建完成后，即开始策略授权，将DTS系统策略AliyunDTSFullAccess授权给子账号。进入RAM控制台，进入用户管理界面，点击要授权子账号后面的 授权入口，开始角色策略授权。



在策略选择步骤中，搜索DTS相关策略。



从左侧的可选授权列表中，选择AliyunDTSFullAccess,添加到已选授权策略列表中，并点击确认，完成账号授权。

通过子账号AccessKey/AccessSecret订阅数据

当子账号创建且授权完成后，即可以使用DTS提供的SDK订阅数据，注意需要将SDK Demo中的用户AccessKey/AccessSecret修改成子账号的AccessKey/AccessSecret。

```

RegionContext context = new RegionContext();
// 运行SDK的服务器是否使用公网IP连接DTS
context.setUsePublicIp(true);
// 用户accessKey secret
context.setAccessKey("LTAIHBol44ANDLAd");
context.setSecret("11E1q1F4PRMj5LW0QC9S1WLNNGv185");
// 创建消费者
final ClusterClient client = new DefaultClusterClient(context);
// 创建订阅监听者listener
ClusterListener listener = new ClusterListener(){

```

运行SDK Demo，发现使用子账号AccessKey/AccessSecret能够正常订阅数据。

```

SLF4J: Failed to load class "org.slf4j.impl.StaticLoggerBinder".
SLF4J: Defaulting to no-operation (NOP) logger implementation
SLF4J: See http://www.slf4j.org/codes.html#StaticLoggerBinder for further details.
timestamp:1483673552
record_encoding:US-ASCII
record_type:begin
checkpoint:526494@2518
source_type:mysql
source_category:full_recorded
instance:100.114.126.165-3306
record_id:2562986

```

```

:null:BEGIN
timestamp:1483673552
db:amptest
record_encoding:utf8
record_type:insert
primary:ID
table_name:sbtest1
checkpoint:526632@2518
source_type:mysql
fields_enc:,,utf8,utf8
source_category:full_recorded
instance:100.114.126.165-3306
logevent:1
record_id:2562987

```

```

Field name: ID
Field type: 8
Field length: 1
Field value: 3

```

至此，完成数据订阅所用子账号创建及授权的说明。

数据迁移

本小节简单介绍数据传输服务提供的数据库迁移功能支持的实例类型及数据源类型。

数据库迁移支持实例类型

源实例	目的实例
本地IDC自建数据库	经典网络RDS、VPC 网络RDS实例 经典网络MongoDB实、VPC 网络MongoDB 经典网络Redis实例、VPC网络Redis实例 经典网络PetaData 经典网络OceanBase ECS自建数据库
ECS自建数据库	经典网络RDS、VPC 网络RDS实例

	经典网络MongoDB 经典网络PetaData 经典网络OceanBase ECS自建数据库
RDS实例	经典网络RDS、VPC 网络RDS实例 经典网络PetaData 经典网络OceanBase ECS自建数据库
经典网络MongoDB实例	经典网络MongoDB实例 ECS自建数据库

数据迁移支持数据源

数据迁移功能，目前支持的迁移功能如下表：

源数据库	结构迁移	全量迁移	增量迁移
Oracle->RDS For MySQL	支持	支持	支持
Oracle->RDS For PPAS	支持	支持	支持
MySQL ->RDS For MySQL	支持	支持	支持
SQLServer->RDS For SQLServer	支持	支持	支持
PostgreSQL-> RDS For PostgreSQL	支持	支持	支持
MongoDB -> MongoDB	支持	支持	支持
Redis -> Redis	支持	支持	支持
MySQL->DRDS	不支持	支持	支持
MySQL->PetaData	不支持	支持	支持
MySQL->OCeanBase	支持	支持	支持

预检查

检测数据传输服务器是否能够连通要迁移的源数据库。数据传输服务通过JDBC协议创建到源数据库的连接，当创建连接失败时，这个检查项即失败。

当源库连接性预检查失败时，可能是由于如下几种原因：

创建迁移任务时，提供的账号或密码不正确

检测方法：在任何一台网络能够跟源数据库互通的服务器上，使用创建迁移任务时使用的账号/密码，通过客户端软件连接源数据库，确认是否能够正常连接。如果连接报错，且报错信息中含有Access deny，那么说明账号或密码有误。

修复方法：在数据传输服务控制台修改迁移任务，将账号密码修改成正确的配置后重新预检查。

源库迁移账号限制了来源IP

检测方法：

在一台网络能够跟源数据库连通的服务器上，使用创建迁移任务时使用的账号/密码，通过客户端软件连接源数据库，能够正常连接。此时说明源数据库限制了来源IP，只有源库迁移账号允许的服务器可以连接，而数据传输服务器的IP不在授权列表，所以不能正常连接。

如果源数据库为MySQL数据库，那么直接使用MySQL客户端连接源数据库，执行select host from mysql.user where user='迁移账号',password='迁移账号密码',如果查询结果不为“%”，那么说明数据传输服务器IP不在授权IP列表中，所以不能正常连接。

如果是SQL Server，那么检查SQL Server安装服务器上是否设置了防火墙，或者检查源数据库中是否有endpoint或trigger限制了访问来源IP。

如果是Oracle，那么检查源数据库的配置文件sqlnet.ora，确认配置项TCP.VALIDNODE_CHECKING是否为yes，如果为yes，说明源数据库限制了访问来源IP。

修复方法:

如果源数据库为MySQL，那么在源数据库中执行grant all on . to “迁移账号” @“%” identified by “迁移密码” 给迁移账号重新授权。这个命令中的迁移账号用真实的迁移账号替换，迁移密码用真实的迁移密码替换。授权后重新进行预检查。

如果源数据库为SQL Server，那么需要先关闭掉防火墙或禁用掉trigger后，重新进行预检查。

如果源数据库为Oracle，那么先修改TCP.VALIDNODE_CHECKING=no并重启监听后，重新进行预检查。

源数据库安装服务器设置了防火墙

检测方法：如果安装源数据库的服务器为Linux，那么在shell中运行 `iptables -L` 检查服务器是否配置了防火墙。如果安装源数据库的服务器为window，那么在控制面板中查找到Windows防火墙，查看是否配置了防火墙。

修复方法：先将防火墙关闭后，重新进行预检查。

数据传输服务器与源数据库网络不通

如果不是上面三种情况中的任何一种，那么可能是由于数据传输服务器跟源数据库网络不连通，当处于这种情况时，请联系数据传输服务值班同学协助解决。

源库binlog是否开启检查

这个配置项只有当进行MySQL->MySQL增量迁移时，才会进行检查。这个检查项主要检查源数据库是否开启binlog日志。如果检查项失败，那么说明源数据库没有开启binlog。

修复方式：如果binlog没有打开，那么需要修改源数据库配置文件中的log_bin，开启二进制日志log_bin=mysql_bin。配置完成后重启源数据库后，重新进行预检查。

源库binlog模式检查

这个配置项只有当进行MySQL->MySQL增量迁移时，才会检查。这个检查项主要检查源数据库的binlog模式是否为ROW。如果检查失败，那么说明源数据库的binlog格式不为ROW。

修复方式：如果检查项失败，那么在源数据库中执行set global binlog_format=ROW后重新进行预检查。建议修改后重启MySQL，否则连接中的session可能以非ROW模式继续写入，这会导致数据丢失。

源库binlog存在性检查

这个配置项只有当进行MySQL->MySQL增量迁移时，才会检查。这个检查项主要检查源数据库存在的binlog文件是否被删除。如果检查失败，那么说明源数据库的binlog文件不完整。

修复方式：如果检查失败，那么在源数据库执行PURGE BINARY LOGS TO 未被删除的第一个binlog文件名后，重新进行预检查。

具体purge的文件名参考预检查的修复方法。

Mysql源库binlog_row_image是否为FULL

这个配置项只有当进行进行MySQL->MySQL增量迁移时，才会检查。这个检查项主要检查源数据库的binlog_row_image是否为full，即是否记录全镜像。如果检查失败，那么说明源数据库的binlog未记录全镜像。

。

修复方式：如果检查失败，那么在源库执行set global binlog_row_image=FULL命令后，重新进行预检查。

这个检查项主要检查迁移的对象列表中，有外键依赖关系的父子表是否都被迁移，防止破坏外键约束完整性。

当这个检查项失败时，失败原因为待迁移表子表名字依赖的父表名字没有迁移。

对于这种错误的修复方式为：

- 不进行约束完整性检查失败的子表的迁移，修改迁移任务配置，将这些约束完整性检查失败的子表从迁移列表中删除后，重新进行预检查。
- 进行约束完整性检查失败的子表对应的父表的迁移，修改迁移任务配置，将这些约束完整性检查失败的子表对应的父表添加到迁移列表中后，重新进行预检查。
- 删除约束完整性检查失败的子表的外键依赖，修改源数据库，将这些子表的外键依赖删除掉，重新进行预检查。

这个配置项只有当进行MySQL->MySQL增量迁移时，才会检查。这个检查项主要检查源数据库中是否存在增量迁移不支持的存储引擎。目前增量迁移不支持federated, mrg_myisam两种存储引擎。

当检查失败，且失败原因为源数据库中引擎为FEDERATED的表时，说明源数据库中要迁移的部分表的存储引擎为federated，所以导致检查失败。

当检查失败，失败原因为源数据库中引擎为MRG_MYISAM的表时，说明源数据库中要迁移的部分表的存储引擎为mrg_myisam，所以导致检查失败。

当检查失败时，对于这些存储引擎为federated或mrg_myisam的表，修改迁移任务，将这些表从迁移列表删除掉。然后单独创建一个迁移任务，对这些表进行结构迁移+全量数据迁移。

源库权限检查

这个检查项主要检查源数据库的迁移账号是否有足够的权限可以进行数据迁移。每种数据库需要的迁移权限详见《迁移手册》。

目标库权限检查

这个检查项主要检查目标数据库的迁移账号是否有足够的权限可以进行数据迁移。每种数据库需要的迁移权限详见《迁移手册》。

这个检查项主要检查目标库是否已经存在跟待迁移对象同名的对象。当这个检查项失败时，说明目标RDS实例中已经存在了跟待迁移的对象一样名字的对象，这会导致迁移失败。

当这个检查项失败时，失败原因为：目标库存在跟待迁移对象同名的结构对象。

当这个检查项失败时，修复方式为：

- 使用数据传输服务提供的库表映射功能，将待迁移对象迁移到目标库中的另外一个对象名的对象中。
- 删除或重命名目标库中跟迁移对象同名的对象。
- 修改迁移任务配置，将同名对象从迁移对象中删除掉，暂时不迁移这些同名对象。

主要检查待迁移的数据库在目标RDS实例中是否存在。如果不存在，那么数据传输服务会自动创建。但是对于如下几种情况，自动创建数据库会失败，所以直接报预检查失败。

源数据库名中含有小写字母、数字、下划线、中划线的其他特殊字符。

当遇到这种情况时，预检查的失败原因为：迁移的数据库迁移库名的名称不符合rds规范。

修复方法为：到RDS控制台的数据库管理页面中，创建满足rds规范的数据库，将新建库的读写权限赋给迁移账号。然后使用数据传输服务提供的库名映射功能将要迁移的数据库映射至新创建的数据库后，然后重新进行预检查。

源数据库的字符集不为utf8、gbk、latin1、utf8mb4中的一种。

当遇到这种情况时，预检查的失败原因为：迁移的数据库库名的字符集不满足rds的要求。

修复方法为：到RDS控制台的数据库管理页面中，创建满足rds规范的数据库，然后将新建库的读写权限赋给迁移账号。如果新创建的库名与待迁移的库名不同，那么使用数据传输服务提供的库名映射功能将要迁移的数据库映射至新创建的数据库后，重新进行预检查。

目标库迁移账号无迁移库的读写权限。

当遇到这种情况时，预检查的失败原因为：该用户没有权限使用目标库的迁移库名。

修复方法为：先到RDS控制台的数据库管理页面中账号管理页面，修改权限，将迁移库的读写权限授予迁移账号后，重新进行预检查。

这个配置项只有当进行MySQL->MySQL增量迁移时，才会进行检查。这个检查项主要检查源数据库的server-id是否设置为大于1的整数。

当检查失败时，只要在源数据库执行set global server_id=某个不为1的数，然后重新进行预检查。

目标数据库连接性检查

检测数据传输服务器是否能够连通要迁移的目标数据库。数据传输服务通过JDBC协议创建到目标数据库的连接

，当创建连接失败时，这个检查项即失败。

当目标库连接性预检查失败时，可能是由于如下的两种原因：

创建迁移任务时，提供的账号或密码不正确

检测方法：在任何一台网络能够跟目标数据库互通的服务器上，使用创建迁移任务时使用的账号/密码，通过客户端软件连接源数据库，确认是否能够正常连接。如果连接报错，且报错信息中含有Access deny，那么说明账号或密码有误。

修复方法：在数据传输服务控制台修改迁移任务，将账号密码修改成正确的配置后重新预检查。

数据传输服务器与目标数据库网络不通

如果不是迁移账号密码不正确的话，那么可能是由于数据传输服务器跟目标数据库网络不连通，当处于这种情况时，请联系数据传输服务值班同学协助解决。

这个检查项主要检查数据传输服务是否支持源数据库的版本号。数据传输服务支持的各种源数据库的版本如下：

源数据库类型	支持的版本号
MySQL	5.0、5.1、5.5、5.6 当进行增量迁移时，仅支持5.1、5.5、5.6
SQL Server	2005、2008、2008R2、2012
Oracle	10g、11g

当版本检查失败时，只能将源数据库升降级到数据传输服务支持的版本范围后，再进行预检查。

使用数据传输DTS可以将本地的Oracle的数据迁移至RDS MySQL实例。DTS可以实现结构定义、全量数据的迁移以及增量数据迁移。通过增量数据迁移，可以在不影响源Oracle的情况下，实现Oracle数据库的不停服迁移。

本小节简单介绍使用DTS进行Oracle->RDS For MySQL数据迁移的任务配置流程。

迁移步骤

对于Oracle->RDS For MySQL的迁移，支持结构定义、全量数据的迁移以及增量数据迁移。各迁移类型的限制如下：

结构迁移

DTS会将迁移对象的结构定义迁移到目标实例。目前DTS支持结构迁移的对象有：表。其他对象如视

图、同义词、触发器、存储过程、存储函数、包、自定义类型等暂不支持。

全量数据迁移

DTS会将源数据库迁移对象在的数据全部迁移到目标实例。如果迁移过程中，本地Oracle库有数据写入的话，那么这些增量数据不一定能够被迁移到MySQL中。所以，为保证迁移数据一致性，在迁移过程中，Oracle实例需停止写入。

增量数据迁移

增量迁移过程中，DTS会监听并获取源Oracle实例的变更redo log，然后将增量变更数据实时同步到目标RDS实例，通过增量数据迁移可以实现目标RDS实例同源Oracle实例的实时数据同步。

迁移权限要求

当使用DTS进行Oracle迁移时，在不同迁移类型情况下，源跟目标数据库的迁移帐号权限要求如下：

迁移类型	结构迁移	全量迁移	增量数据迁移
本地Oracle实例	schema的owner	schema的owner	SYSDBA
目的RDS MySQL实例	待迁入db的读写权限	待迁入db的读写权限	待迁入db的读写权限

迁移前置条件

- Oracle版本为10g, 11g, 12c。
- Oracle 开启supplement log，且要求 supplemental_log_data_pk，supplemental_log_data_ui 开启。
- Oracle要求开启archive log。

数据类型映射关系

由于Oracle跟MySQL的数据类型并不是一一对应的，所以DTS在进行结构迁移时，会根据两种数据库类型的数据类型定义，进行类型映射，下面是数据类型映射关系。

Oracle数据类型	MySQL数据类型	DTS是否支持
varchar2(n [char/byte])	varchar(n)	支持
nvarchar2[(n)]	national varchar[(n)]	支持
char[(n [byte/char])]	char[(n)]	支持
nchar[(n)]	national char[(n)]	支持
number[(p[,s])]	decimal[(p[,s])]	支持
float(p)	double	支持

long	longtext	支持
date	datetime	支持
binary_float	decimal(65,8)	支持
binary_double	double	支持
timestamp[(fractional_seconds_precision)]	datetime[(fractional_seconds_precision)]	支持
timestamp[(fractional_seconds_precision)]with local time zone	datetime[(fractional_seconds_precision)]	支持
timestamp[(fractional_seconds_precision)]with local time zone	datetime[(fractional_seconds_precision)]	支持
clob	longtext	支持
nclob	longtext	支持
blob	longblob	支持
raw	varbinary(2000)	支持
long raw	longblob	支持
bfile	—	不支持
interval year(year_precision) to month	—	不支持
interval day(day_precision) to second[(fractional_seconds_precision)]	—	不支持

- 对于char类型，当char(n)的定义长度n超过255时，DTS会自动将类型转换为varchar(n)。
- 由于MySQL本身不支持类似Oracle中的bfile、interval year to month、interval day to second这种数据类型，所以DTS在进行结构迁移时，这三种类型不会进行转化。迁移时如果表中含有这三种类型，会导致结构迁移失败，用户根据业务需求，将这三种类型修改成MySQL中对应的数据类型即可。
- 由于MySQL的timestamp类型不包含时区，而Oracle的timestamp with [local] time zone带时区，所以DTS在迁移这两种类型的数据时，会将其转换成UTC时区后，存入MySQL。

迁移步骤

下面详细介绍下使用DTS将本地的Oracle迁移到RDS for MySQL的任务配置流程。

创建RDS实例数据库

在数据迁移过程中，如果待迁移的数据库在目标RDS实例中不存在，那么DTS会自动创建。但是对于如下两种情况，用户需要在配置迁移任务之前，手动创建数据库。

1. 数据库名称不符合：RDS定义规范(由小写字母、数字、下划线、中划线组成，字母开头，字母或数字结尾，最长64个字符)。
2. 待迁移数据库，在Oracle跟目标RDS实例中名称不同。

对于这两种情况，用户需要在配置迁移任务之前，先在RDS控制台完成数据库创建。具体参考RDS数据库创建流程RDS使用手册。

创建迁移帐号

迁移任务配置，需要提供Oracle数据库及目标RDS实例的迁移账号。迁移账号所需权限详见上文的**迁移权限要求**。

如果您的Oracle实例的迁移账号尚未创建，那么您可以参考 [Oracle Grant语法说明](#)，创建满足要求的迁移账号。

RDS For MySQL 迁移账号的创建及授权操作详见RDS使用手册RDS使用手册。

迁移任务配置

当上面的所有前置条件都配置完成后，就可以开始迁移任务配置。下面详细介绍下具体的迁移步骤。

1. 进入数据传输DTS控制台，单击右上角的创建迁移任务，正式开始任务配置。

本地Oracle及目标RDS实例的连接信息配置。

这个步骤主要配置迁移任务名称，Oracle连接信息及目标RDS实例连接信息。其中：

任务名称

DTS为每个任务自动生成一个任务名称，任务名称没有唯一性要求。您可以根据需要修改任务名称，建议为任务配置具有业务意义的名称，便于后续的任务识别。

源实例信息

- 实例类型：选择 **有公网IP的自建数据库**
- 数据库类型：选择 **Oracle**
- 主机名或IP地址：配置Oracle访问地址，这个地址必须为公网访问方式
- 端口：Oracle实例的监听端口
- SID：Oracle实例的SID
- 账号：Oracle实例的连接账号
- 密码：上面Oracle连接账号对应的密码

目标实例信息

- 实例类型：选择 **RDS实例**
- RDS实例ID：配置迁移的目标RDS实例的实例ID。DTS支持经典网络、VPC网

络的RDS实例

- 账号：RDS实例的连接账号
- 密码：上面RDS实例连接账号对应的密码

当配置完连接信息后，单击右下角 **授权白名单并进入下一步** 进行白名单授权。这个步骤DTS会将DTS服务器的IP地址添加到目标RDS实例的白名单中，避免因为RDS实例设置了白名单，导致DTS服务器连接不上RDS实例导致迁移失败。

迁移对象及迁移类型配置。

迁移类型包括：结构迁移、全量数据迁移、增量数据迁移。默认选择 **结构迁移+全量数据迁移**。

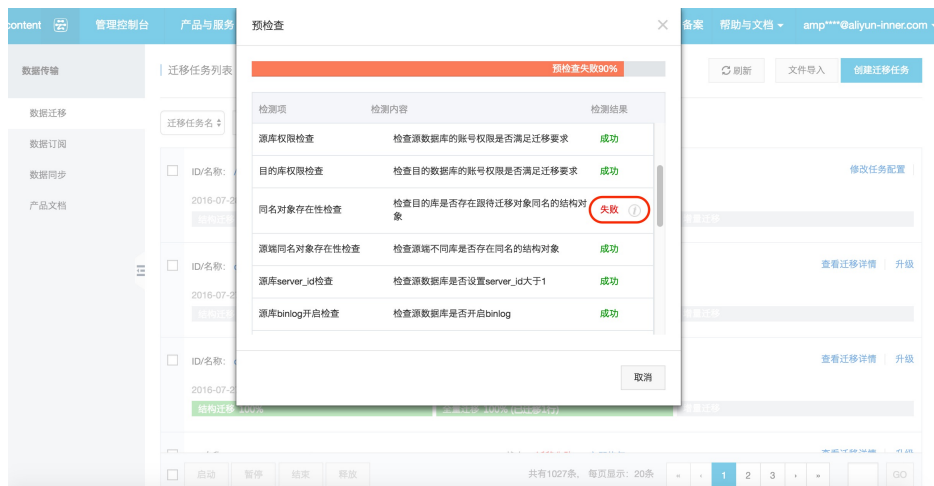
迁移对象，需要选择您要迁移的对象。迁移对象选择的粒度可以为：库、表、列三个粒度。默认情况下，对象迁移到RDS实例后，对象名跟Oracle一致。如果您迁移的对象在源实例跟目标实例上名称不同，那么需要使用DTS提供的对象名映射功能，详细使用方式可以参考库表列映射。

当配置完迁移对象及迁移类型后，即进入任务启动前的预检查步骤。

任务预检查。

在迁移任务正式启动之前，会先进行前置预检查，只有预检查通过后，才能成功启动迁移。

如果预检查失败，那么可以点击具体检查项后的按钮，查看具体的失败详情，并根据失败原因修复后，重新进行预检查。



启动迁移任务。

当预检查通过后，我们可以启动迁移任务，任务启动后，可以到任务列表中查看任务具体的迁移状态及进度。

当任务进入增量数据迁移阶段，任务不会自动停止，且一旦源Oracle实例有增量写入，增量数据就会自动同步到目标RDS实例。增量数据迁移是个动态同步的过程，建议在增量迁移达到无延迟状态时

，在目标数据库上进行业务验证。如果验证成功，那么可以停掉迁移任务，将业务切换到目标数据库。

至此，完成将本地Oracle数据库到RDS For MySQL的数据迁移任务配置。

使用数据传输DTS可以将本地MySQL实例的数据迁移到RDS for MySQL实例。DTS提供的增量数据同步功能，可以实现在本地应用不停服的情况下，平滑完成MySQL数据的迁移工作。

本小节简单介绍使用DTS进行MySQL->RDS For MySQL数据迁移的任务配置流程。

迁移步骤

对于MySQL->RDS For MySQL数据迁移，DTS支持结构迁移、全量数据迁移及增量数据迁移，这个迁移类型的功能及限制如下：

结构迁移

数据传输会将迁移对象的结构定义迁移到目标实例。目前DTS支持结构迁移的对象有：表、视图、触发器、存储过程、存储函数。

全量数据迁移

全量迁移将MySQL实例的存量数据全部迁移到目标RDS实例。

如果用户只进行全量数据迁移，那么迁移过程中本地MySQL实例新增的业务写入不会被同步到目标RDS实例。

如果用户还选择了增量数据迁移，那么全量迁移过程中，为了保证迁移数据一致性，无主键的非事务表会被锁定，锁定期间这些表无法写入，锁定时长依赖于这些表的数据量大小，在这些无主键非事务表迁移完成后，锁才会释放。

增量数据迁移

增量迁移会将迁移过程中，本地MySQL实例的增量变更数据同步到目标RDS实例，增量数据迁移可以实现本地MySQL实例在迁移过程中产生的增量数据被实时同步到目标库，最终实现本地MySQL同目标RDS实例进入动态数据同步的过程。

迁移限制

- 迁移过程中，不支持DDL操作
- 结构迁移不支持event的迁移
- 如果使用了对象名映射功能后，依赖这个对象的其他对象可能迁移失败

- 当选择增量迁移时，源MySQL实例需要开启binlog
- 当选择增量迁移时，源库的binlog_format 要为row
- 当选择增量迁移且源MySQL如果为5.6及以上版本时，它的binlog_row_image必须为full

迁移权限要求

当使用DTS进行MySQL->RDS For MySQL迁移时，在不同迁移类型，对源跟目标MySQL实例的迁移帐号权限要求如下：

迁移类型	结构迁移	全量迁移	增量迁移
本地MySQL实例	select	select	super select replication slave replication client
目的RDS实例	读写权限	读写权限	读写权限

迁移流程

DTS 在进行MySQL->RDS For Mysql数据迁移时，为了解决对象间的依赖关系，提高迁移成功率。结构对象及数据的迁移顺序如下：

1. 结构对象：表、视图的迁移
2. 全量数据迁移
3. 结构对象：存储过程、函数、触发器、外键的迁移
4. 增量数据迁移

如果任务没有选择增量数据迁移，那么当全量数据迁移完成后，任务列表中的迁移进度为：**结构迁移100%，全量迁移100%**，迁移状态为“迁移中”。此时迁移任务正在进行步骤(3)中的对象的迁移。此时，请勿手动结束任务，否则会造成迁移数据丢失。

迁移任务配置

下面详细介绍使用DTS将本地的MySQL迁移到RDS for MySQL的任务配置流程。

RDS实例数据库创建

在数据迁移过程中，如果待迁移的数据库在目标RDS实例中不存在，那么DTS会自动会创建。但是对于如下两种情况，用户需要在配置迁移任务之前，手动创建数据库。

1. 数据库名称不符合：RDS定义规范(由小写字母、数字、下划线、中划线组成，字母开头，字母或数字结尾，最长64个字符)。
2. 待迁移数据库，在Oracle跟目标RDS实例中名称不同。

对于这两种情况，用户需要在配置迁移任务之前，先在RDS控制台完成数据库创建。具体参考RDS数据库创建流程RDS使用手册。

迁移帐号创建

迁移任务配置，需要提供本地MySQL数据库及目标RDS实例的迁移账号。迁移账号所需权限详见上文的 迁移权限要求。

如果源MySQL实例迁移账号尚未创建，那么您可以参考 Grant语法说明，创建满足权限要求的迁移账号。

如果目标RDS实例迁移账号创建，那么您可以参考 RDS账号创建流程，创建对待迁移数据库有读写权限的迁移账号。

其他准备工作

如果需要进行增量迁移，那么需要确认源库的binlog是否开启，格式是否满足要求。具体确认项如下。

确认源库的binlog是否开启。

```
mysql> show global variables like "log_bin";
+-----+-----+
| Variable_name | Value |
+-----+-----+
| log_bin       | ON    |
+-----+-----+
1 row in set (0.04 sec)
```

如果这里面查询出来的log_bin=OFF,那么说明源库没有开启binlog，这个时候如果需要使用增量迁移的话，需要修改本地MySQL的binlog配置，打开binlog日志。

- i. 开启二进制日志，修改log_bin=mysql_bin。
- ii. 设置binlog模式为row，修改binlog_format=row。
- iii. 配置server_id大于1，即server_id=某一个大于1的整数。
- iv. 如果版本为5.6,修改binlog_row_image=full。
- v. 修改完成后，重启MySQL进程。

确认源库的binlog格式为row模式。

```
mysql> show global variables like "binlog_format";
+-----+-----+
| Variable_name | Value |
+-----+-----+
| binlog_format | ROW   |
+-----+-----+
1 row in set (0.04 sec)
```

使用上面的命令确认本地MySQL的binlog格式是否为row,如果查询出来的结果不为ROW的话,那么通过如下的参数设置命令将模式修改为row。

```
mysql> set global binlog_format=ROW;
Query OK, 0 rows affected (0.04 sec)
```

当修改完成后,最好将已有的老连接KILL掉重连,否则可能出现其他的连接依然使用statement格式记录binlog日志。

当本地MySQL版本大等于5.6.2时,确认源库的binlog_row_image=full。

```
mysql> show global variables like "binlog_row_image";
+-----+-----+
| Variable_name      | Value |
+-----+-----+
| binlog_row_image  | FULL  |
+-----+-----+
1 row in set (0.04 sec)
```

如果查询出来的结果不为FULL的话,那么通过如下的参数设置命令将其修改为full。

```
mysql> set global binlog_row_image=full;
Query OK, 0 rows affected (0.03 sec)
```

迁移任务配置

当数据库、迁移账号都创建完成后,就可以开始配置迁移任务了。下面详细介绍下具体的配置步骤。

1. 进入数据传输DTS控制台,点击右上角的**创建迁移任务**,开始任务配置。

本地MySQL及目标RDS实例连接信息配置。

在这个步骤中,主要配置迁移任务名称,迁移源实例及目标实例连接信息。其中:

任务名称

默认情况下，DTS为每个任务自动生成一个任务名称。任务名称没有唯一性要求，您可以修改这个名称，为任务配置一个具有业务意义的名称，便于后续的任务识别。

源实例连接信息

- 实例类型：选择 **有公网IP的自建数据库**
- 数据库类型：选择 **MySQL**
- 主机名或IP地址：配置本地MySQL数据库访问地址，这个地址必须为公网访问方式
- 端口：MySQL实例监听端口
- 账号：MySQL实例访问账号
- 密码：上面MySQL访问账号对应的密码

目标RDS实例连接信息。

- 实例类型：选择 **RDS实例**
- RDS实例ID: 配置迁移的目标RDS实例的实例ID。DTS支持经典网络、VPC网络的RDS实例
- 数据库账号：RDS实例的连接账号
- 数据库密码：上面数据库账号对应的数据库密码

创建迁移任务 [返回上级](#)

1. 源库及目标库 2. 迁移类型及列表 3. 预检查

* 任务名称: dts7x4p3le6

源库信息

* 实例类型: 有公网IP的自建数据库

* 数据库类型: MySQL

* 主机名或IP地址: 120.55.141.62

* 端口: 3306

* 数据库账号: jianglutest

* 数据库密码:

目标库信息

* 实例类型: RDS实例

* RDS实例ID: rm-bp16w2438yct639x9

* 数据库账号: jianglutest

* 数据库密码:

[取消](#) [上一步](#) [授权白名单并进入下一步](#)

迁移对象及迁移类型配置。

迁移类型

DTS支持 结构迁移、全量数据迁移、增量数据迁移。

如果需要进行不停机迁移，那么需要选择：结构迁移 + 全量数据迁移 + 增量数据迁移。

如果只进行全量迁移，那么需要选择：结构迁移 + 全量数据迁移。

迁移对象

迁移对象，需要选择您要迁移的对象。迁移对象选择的粒度可以为：库、表、列三个粒度。

默认情况下，对象迁移到RDS实例后，对象名跟本地MySQL实例一致。如果您迁移的对象在源实例跟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，详细使用方式可以参考库表列映射。

当配置完迁移对象及迁移类型后，即进入任务启动前的预检查步骤

预检查。

在迁移任务正式启动之前，会先进行前置预检查，只有预检查通过后，才能成功启动迁移。

如果预检查失败，那么可以点击具体检查项后的按钮，查看具体的失败详情，并根据失败原因修复后，重新进行预检查。



启动迁移任务。

当预检查通过后，我们可以启动迁移任务，任务启动后，可以到任务列表中查看任务具体的迁移状态及进度。

增量数据迁移是个动态同步的过程，所以建议在增量迁移达到无延迟状态时，在目标数据库上进行业务验证，如果验证成功，那么可以停掉迁移任务，然后将业务切换到目标数据库。

至此，完成将本地MySQL数据库到RDS For MySQL的数据迁移任务配置。

使用数据传输 DTS 可以将本地 SQL Server 实例的数据迁移到 RDS for SQL Server 实例。SQL Server 迁移支持增量数据同步功能，可以实现在本地应用不停服的情况下，平滑完成 SQL Server 数据库的迁移工作。

本小节简单介绍使用数据传输服务 DTS（以下简称 DTS）进行 SQL Server->RDS for SQL Server 数据迁移的任务配置流程。

迁移类型

对于 SQL Server->RDS for SQL Server 数据迁移，DTS 支持结构迁移、全量数据迁移及增量数据迁移，这个迁移类型的功能及限制如下：

结构迁移

DTS 将迁移对象的结构定义迁移到目标实例。目前 DTS 支持结构迁移的对象有：表、视图、表触发器、同义词、SQL 存储过程、SQL 函数、plan guid、自定义类型、rule、default。

全量数据迁移

全量迁移将 SQL Server 实例的存量数据全部迁移到目标 RDS 实例。如果用户只进行全量数据迁移，那么迁移过程中本地 SQL Server 实例新增的业务写入不会被同步到目标 RDS 实例。

如果用户还选择了增量数据迁移，那么迁移过程中，本地 SQL Server 实例的增量写入数据会被同步到目标 RDS 实例。

迁移限制

- 迁移过程中，不支持 DDL 操作
- 结构迁移不支持 assemblies、库级存储过程、service broker、全文索引、全文目录、分布式 schema、分布式函数、CLR 标量函数、CLR 标值函数、内部表、聚合函数、系统的迁移
- 如果使用了对象名映射功能后，依赖这个对象的其他对象可能迁移失败
- 增量迁移，只支持含有聚簇索引，且聚簇索引为唯一键的表
- 增量迁移，不支持只更新大字段的 update 语句的同步
- 增量迁移，不支持含有计算列的表
- 一个增量迁移任务，只支持一个数据库的迁移。如果同时有多个数据库需要进行增量数据迁移，那么需要创建多个迁移任务

迁移权限要求

当使用 DTS 进行 SQL Server 迁移时，不同迁移类型，对本地 SQL Server 数据库及目标 RDS for SQL Server 实例的迁移帐号权限要求如下：

迁移类型	结构迁移	全量迁移
------	------	------

本地 SQLServer 实例	select	select
目的 RDS 实例	读写权限	读写权限

迁移流程

数据传输服务在进行 SQL Server 上云迁移时，为了解决对象间的依赖，提高迁移成功率。

结构对象及数据的迁移顺序如下：

1. 进行结构对象 **表、视图、同义词、自定义类型、rule、default、plan guid** 的迁移。
2. 全量数据迁移。
3. 进行结构对象 **SQL 存储过程、SQL 函数、触发器、外键** 的迁移。
4. 增量数据迁移。

如果任务没有选择增量数据迁移，那么当全量数据迁移完成后，任务列表中的迁移进度为：**结构迁移100%，全量迁移100%**，迁移状态为“迁移中”。此时迁移任务正在进行步骤(3)中的对象的迁移。此时，请勿手动结束任务，否则会造成迁移数据丢失。

迁移任务配置

下面详细介绍使用 DTS 将本地的 SQL Server 迁移到 RDS for SQL Server 的任务配置流程。

RDS 实例数据库创建

在数据迁移过程中，如果待迁移的数据库在目标 RDS for SQL Server 实例中不存在，那么 DTS 会自动创建同名的数据库。但是对于如下两种情况，用户需要在配置迁移任务之前，手动创建数据库。

1. 数据库名称不符合：RDS 定义规范 (由小写字母、数字、下划线、中划线组成，字母开头，字母或数字结尾，最长 64 个字符)。
2. 待迁移数据库，在本地 SQL Server 跟目标 RDS 实例中名称不同。

对于这两种情况，用户需要在配置迁移任务之前，先在 RDS 控制台完成数据库创建。具体参考 [RDS 数据库创建流程](#) [RDS 使用手册](#)。

迁移帐号创建

迁移任务配置，需要提供本地 SQL Server 数据库及目标 RDS 实例的迁移账号。迁移账号所需权限详见上文的 [迁移权限要求](#)。

如果本地 SQL Server 实例迁移账号尚未创建，那么您可以参考 [SQL Server User 创建](#)，创建满足权限要求的迁移账号。

如果目标 RDS 实例迁移账号创建，那么您可以参考 [RDS 账号创建流程](#)，创建对待迁移数据库有读写权限的迁移账号。

其他准备工作

如果您需要进行不停机迁移，那么还需要设置本地 SQL Server 日志格式为 full。

如果本地 SQL Server 数据库的日志格式不为 full，那么需要通过下面两个步骤设置：

1. 在源数据库执行：`alter database database_name set recovery_model_desc=' full'` ，其中 `database_name` 为需要迁移的数据库名。
2. 为了保证开启完整日志生效，需要在源数据库进行一次日志备份，在源数据库执行：`BACKUP LOG database_name to DISK=backup_place WITH init`，其中 `database_name` 为待迁移的数据库名，`backup_place` 为备份文件存储的地址。

迁移任务配置

当数据库、迁移账号都创建完成后，就可以开始配置迁移任务了。下面详细介绍下具体的配置步骤。

1. 进入数据传输 DTS 控制台，点击右上角的创建迁移任务，开始任务配置。

本地 SQL Server 及目标 RDS 实例连接信息配置。

在这个步骤中，主要配置迁移任务名称，迁移源实例及目标实例连接信息。其中：

任务名称

默认情况下，DTS 为每个任务自动生成一个任务名称。任务名称没有唯一性要求，您可以修改这个名称，为任务配置一个具有业务意义的名称，便于后续的任务识别。

源实例连接信息

- 实例类型：选择 **有公网 IP 的自建数据库**
- 数据库类型：选择 **SQL Server**
- 主机名或IP地址：配置本地 SQL Server 数据库访问地址，这个地址必须为公网访问方式
- 端口：SQL Server 实例监听端口
- 数据库账号：SQL Server 实例访问账号
- 数据库密码：上面指定的 SQL Server 访问账号对应的密码

目标 RDS 实例连接信息

- 实例类型：选择 **RDS 实例**
- RDS 实例 ID: 配置迁移的目标 RDS 实例的实例 ID。DTS 支持经典网络、VPC 网络的 RDS 实例
- 数据库账号：RDS 实例的连接账号
- 数据库密码：上面数据库账号对应的数据库密码

创建迁移任务 返回上级

1. 源库及目标库 2. 迁移类型及列表 3. 预检查

* 任务名称: dts6wba5ljg

源库信息

* 实例类型: 有公网IP的自建数据库

* 数据库类型: SQL Server

* 主机名或IP地址: [Redacted]

* 端口: 3433

* 数据库账号: jiangluteest

* 数据库密码: [Redacted]

目标库信息

* 实例类型: RDS实例

* RDS实例ID: [Redacted]

* 数据库账号: jiangluteest

* 数据库密码: [Redacted]

取消 上云评估 授权白名单并进入下一步

迁移对象及迁移类型配置。

迁移类型

DTS 支持结构迁移、全量数据迁移、增量数据迁移。

如果需要进行不停机迁移，那么需要选择：结构迁移 + 全量数据迁移 + 增量数据迁移。

如果只进行全量迁移，那么需要选择：结构迁移 + 全量数据迁移。

迁移对象

迁移对象，需要选择您要迁移的对象。迁移对象选择的粒度可以为：库、表、列三个粒度。默认情况下，对象迁移到 RDS 实例后，对象名跟本地 SQL Server 实例一致。如果您迁移的对象在源实例跟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，详细使用方式可以参考库表列映射。

当配置完迁移对象及迁移类型后，即进入任务启动前的预检查步骤

预检查。

在迁移任务正式启动之前，会先进行前置预检查，只有预检查通过后，才能成功启动迁移。

如果预检查失败，那么可以点击具体检查项后的按钮，查看具体的失败详情，并根据失败原因修复后，重新进行预检查。



启动迁移任务。

当预检查通过后，可以启动迁移任务，任务启动后，可以到任务列表中查看任务具体的迁移状态及进度。

增量数据迁移是个动态同步的过程，所以建议在增量迁移达到无延迟状态时，在目标数据库上进行业务验证，如果验证成功，那么可以停掉迁移任务，然后将业务切换到目标数据库。

至此，完成将本地 SQL Server 数据库到 RDS for SQL Server 的数据迁移任务配置。

适用场景

DTS数据传输支持SQLServer增量数据迁移，通过增量数据迁移可以实现本地SQLServer业务不停服的情况下，完成SQLServer到RDS For SQLServer数据库的迁移。

本小节简单介绍使用DTS进行SQLServer到RDS For SQLServer不停机迁移的任务配置流程。

迁移类型

对于SQLserver到RDS For SQLserver数据迁移，DTS支持结构迁移、全量数据迁移及增量数据迁移，这个迁移类型的功能及限制如下：

结构迁移

DTS将迁移对象的结构定义迁移到目标实例。目前DTS支持结构迁移的对象有：表、视图、表触发器、同义词、SQL存储过程、SQL函数、plan guid、自定义类型、rule、default。其他对象暂不支持。

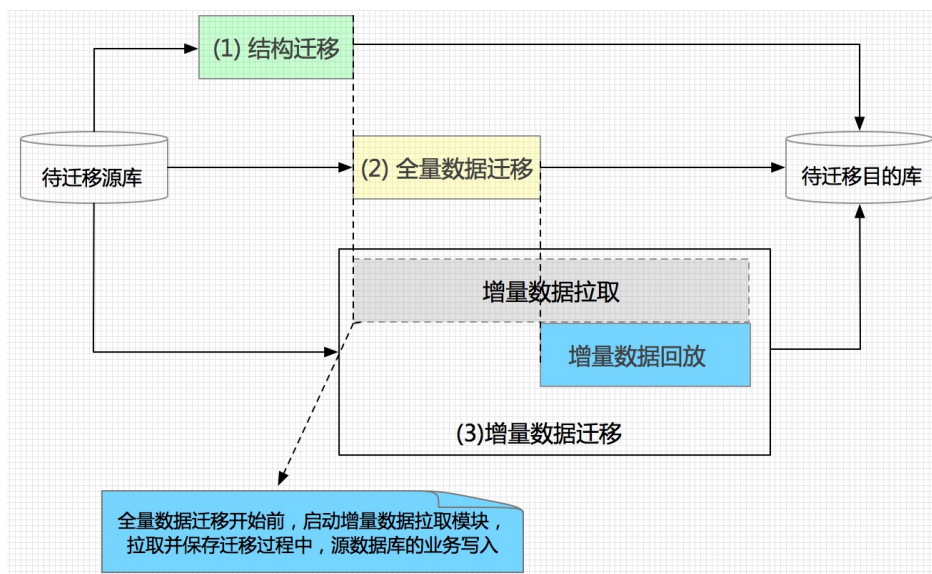
全量数据迁移

全量迁移将SQLServer实例的存量数据全部迁移到目标RDS实例。

增量数据迁移

将迁移过程中产生的增量数据迁移到目标库，通过增量数据迁移可以保证源数据库迁移过程产生的增量数据可以被同步到目标数据库，最终保证两个数据库的数据一致。

不停服迁移原理



为保证迁移过程中，源数据库业务写入的所有变更数据能够被迁移到目标库，在全量迁移之前，数据传输后台就会开启增量数据拉取模块，拉取并保存迁移过程中，源数据库的业务写入。当全量数据迁移完成后，会启动增量写入模块，开始回放这些增量数据，从而实现源跟目标实例增量数据实时同步。

当迁移进入增量迁移阶段时，增量数据迁移是一个动态的过程，除非用户手动停止迁移，否则会一直保持两边数据实时同步，所以建议当增量迁移无延迟时，可以进行源数据库和目标数据库的数据校验，并在目标数据库进行业务测试，当这些验证都通过后，可以停止迁移任务，然后将业务切换到目标数据库上。通过这种迁移方式，可实现在数据迁移过程中，应用停机时间降低到分钟级别。

迁移限制

- 如果使用了对象名映射功能后，依赖这个对象的其他对象可能迁移失败。
- 增量迁移，不支持DDL同步。
- 增量迁移，只支持含有聚簇索引，且聚簇索引为唯一键的表。
- 增量迁移，不支持只更新大字段的update语句的同步。
- 增量迁移，不支持含有计算列的表。
- 一个增量迁移任务，只支持一个数据库的迁移。如果同时有多个数据库需要进行增量数据迁移，那么需要创建多个迁移任务。

迁移权限要求

当使用数据传输进行SQLServer迁移时，在不同迁移类型情况下，源跟目标数据库的迁移账号权限要求如下：

迁移类型	结构迁移	全量数据迁移	增量数据迁移
本地SQLServer实例	select	select	sysadmin
目的RDS实例	读写权限	读写权限	读写权限

环境准备

源数据库的日志格式必须为full，如果源数据库的日志格式不为full，那么需要通过下面两个步骤设置：

在源数据库执行：

```
USE master;
GO
ALTER DATABASE database_name
SET RECOVERY FULL
WITH ROLLBACK IMMEDIATE;
GO
```

其中database_name为需要迁移的数据库名。

ii. 为了保证开启完整日志生效，需要在源数据库进行一次日志备份，在源数据库执行：

```
BACKUP LOG database_name to DISK=backup_place WITH init
```

其中database_name为待迁移的数据库名，backup_place为备份文件存储的地址。

创建RDS实例数据库

在数据迁移过程中，如果待迁移的数据库在目标RDS实例中不存在，那么DTS会自动创建。但是对于如下两种情况，用户需要在配置迁移任务之前，手动创建数据库。

- 数据库名称不符合：RDS定义规范（由小写字母、数字、下划线、中划线组成，字母开头，字母或数字结尾、最长64个字符）。

待迁移数据库，在自建数据库跟目标RDS实例中名称不同。

对于这两种情况，用户需要在配置迁移任务之前，先在RDS控制台完成数据库创建。具体参考RDS数据库创建流程。

迁移账号创建

迁移任务配置，需要提供本地SQLServer数据库及目标RDS实例的迁移账号。迁移账号所需权限详见上文的 迁移权限要求。

如果本地SQLServer实例迁移账号尚未创建，那么您可以参考SQLServer User创建，创建满足权限要求的迁移账号。

如果目标RDS实例迁移账号创建，那么您可以参考 RDS账号创建流程，创建对待迁移数据库有读写权限的迁移账号。

操作步骤

1. 登录DTS控制台

以目标RDS实例所属阿里云账号登录DTS数据传输控制台，单击右上角的**创建迁移任务**，开始任务配置。

2. 设置实例连接信息

在这个步骤中，主要配置迁移任务名称，迁移源实例（本地SQLServer）及目标RDS实例连接信息。其中：

- 任务名称

DTS为每个任务自动生成一个任务名称，任务名称没有唯一性要求。您可以根据需要修改任务名称，建议为任务配置具有业务意义的名称，便于后续的任务识别。

源实例连接信息

- 数据库类型：选择 **SQLServer**。
- 主机名或IP地址：配置本地SQLServer数据库访问地址，如果本地SQLServer实例有公网访问地址，那么配置实例的公网访问地址。如果本地SQLServer实例没有公网访问地址，但是可以通过专线或VPN网关被VPC ECS访问，那么可以在VPC ECS通过nginx进行本地SQLServer的代理转发，这个配置项就配置nginx所部署ECS的EIP。
- 端口：SQLServer实例监听端口,或nginx的代理转发端口。
- 数据库账号：SQLServer实例访问账号。
- 数据库密码：上面SQLServer访问账号对应的密码。

目标RDS实例连接信息

- 实例类型：选择 **RDS实例**。
- RDS实例ID: 配置迁移的目标RDS实例的实例ID。DTS支持经典网络、VPC网络的RDS实例。
- 数据库账号：RDS实例的连接账号。
- 数据库密码：上面数据库账号对应的数据库密码。

创建迁移任务 [返回上级](#)

1. 源库及目标库 2. 迁移类型及列表 3. 预检查

* 任务名称: dts6wba5lyg

源库信息

* 实例类型: 有公网IP的自建数据库

* 数据库类型: SQLServer

* 主机名或IP地址: [REDACTED]

* 端口: 3433

* 数据库账号: jianglunest

* 数据库密码: [REDACTED]

目标库信息

* 实例类型: RDS实例

* RDS实例ID: [REDACTED]

* 数据库账号: jianglunest

* 数据库密码: [REDACTED]

[取消](#) [上一步](#) [授权白名单并进入下一步](#)

3. 迁移对象及迁移类型配置

DTS支持迁移类型

- 结构迁移
- 全量数据迁移
- 增量数据迁移

说明：进行不停机迁移，需要选择：结构迁移 + 全量数据迁移 + 增量数据迁移。

- 迁移对象

迁移对象，需要选择您要迁移的对象。迁移对象选择的粒度可以为：库、表、列三个粒度。默认情况下，对象迁移到RDS实例后，对象名跟本地SQLServer实例一致。如果您迁移的对象在源实例跟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，详细使用方式可以参考库表列映射。

4. 预检查

在迁移任务正式启动之前，会先进行前置预检查，只有预检查通过后，才能成功启动迁移。

如果预检查失败，那么可以点击具体检查项后的按钮，查看具体的失败详情，并根据失败原因修复后，重新进行预检查。



5. 启动迁移任务

当预检查通过后，可以启动迁移任务，任务启动后，可以到任务列表中查看任务具体的迁移状态及进度。

增量数据迁移是个动态同步的过程，所以建议在增量迁移达到无延迟状态时，在目标数据库上进行业务验证，如果验证成功，那么可以停掉迁移任务，然后将业务切换到目标数据库。

6. 数据同步验证

为了验证迁移过程中，源数据库变更是否可以同步到目标库，我们在迁移过程中，往源库插入一条记录，具体如下图。



可以查看此时源库表sbtest1中新增了一条记录，此时到目标库中查询这条记录。



所以，通过这种方式可以将迁移过程中源库产生的增量数据实时同步到目标库，从而实现数据迁移过程中，应用停机时间降低到分钟级别。

至此，完成本地SQLServer数据库到RDS For SQLServer不停服迁移的任务配置。

使用数据传输 DTS 可以将用户本地 Oracle 实例中的数据迁移至 DRDS 实例。数据传输 DTS 可以实现全量数据迁移以及增量数据迁移。通过增量数据迁移，可以在不影响源 Oracle 实例的情况下，实现 Oracle 数据库的不停服迁移。

本小节简单介绍使用数据传输 DTS（以下简称 DTS）进行 Oracle->DRDS 数据迁移的任务配置流程。

迁移步骤

对于 Oracle->DRDS 的迁移，全量数据迁移及增量数据迁移的机制如下：

全量数据迁移

DTS 会将源数据库迁移对象包含的数据全部迁移到目标实例。如果迁移任务的迁移类型只选择了全量数据迁移，那么迁移过程中，本地 Oracle 库在全量迁移过程中新增的数据不一定能够被迁移到目标的 DRDS 实例中，在全量迁移过程中发生的数据修改和删除也不一定能够被同步到目标的 DRDS 实例中。所以，为保证迁移数据一致性，对于只选择全量数据迁移的任务，在迁移过程中，Oracle 实例需停止写入，处于只读状态。

增量数据迁移

增量迁移过程中，DTS 会监听并获取源 Oracle 实例由于生成的 redo log，然后将增量变更数据实时同步到目标 DRDS 实例，通过增量数据迁移可以实现目标 DRDS 实例同源 Oracle 实例的实时数据同步。

迁移权限要求

当使用 DTS 进行 Oracle->DRDS 迁移时，在不同迁移类型情况下，源跟目标数据库的迁移帐号权限要求如下：

迁移类型	结构迁移	全量迁移	增量数据迁移
本地 Oracle 实例	schema 的 owner	schema 的 owner	SYSDBA
目标 DRDS 实例	待迁入 db 的读写权限	待迁入 db 的读写权限	待迁入 db 的读写权限

迁移前置条件

- Oracle版本为 10g，11g，12c。
- Oracle 实例开启 supplement log，且要求 supplemental_log_data_pk，supplemental_log_data_ui 开启。
- Oracle 实例要求开启 archive log，处于归档模式。

下面详细介绍下使用 DTS 将本地的 Oracle 迁移到 DRDS 的任务配置流程。

创建 DRDS 实例迁移对象

在数据迁移过程中，如果待迁移的数据库、数据表在目标 DRDS 实例中不存在，用户需要在配置迁移任务之前，先手动在目标 DRDS 实例中创建好数据库和对应的表。

创建迁移帐号

迁移任务配置，需要提供 Oracle 数据库及目标 DRDS 实例的迁移账号。迁移账号所需权限详见上文的 **迁移权限要求**。

如果您的 Oracle 实例的迁移账号尚未创建，那么您可以参考 Oracle Grant 语法说明，创建满足要求的迁移账号。

DRDS 实例迁移账号的创建及授权操作详见 DRDS 使用手册。

迁移步骤

当上面的所有前置条件都配置完成后，就可以开始配置迁移任务。下面详细介绍下具体的迁移步骤。

进入数据传输 DTS 控制台，单击右上角的创建迁移任务，正式开始任务配置。

本地 Oracle 实例及目标 DRDS 实例的连接信息配置

这个步骤主要配置迁移任务名称，Oracle 连接信息及目标 DRDS 实例连接信息。其中：

任务名称

DTS 为每个任务自动生成一个任务名称，任务名称没有唯一性要求。您可以根据需要修改任务名称，建议为任务配置具有业务意义的名称，便于后续的任务识别。

源实例信息

- 实例类型：选择 **有公网 IP 的自建数据库**。
- 数据库类型：选择 **Oracle**。
- 主机名或 IP 地址：配置 Oracle 访问地址，这个地址必须为公网访问方式。
- 端口：Oracle 实例的监听端口。
- SID：Oracle 实例的 SID。
- 账号：Oracle 实例的连接账号。
- 密码：上面指定的 Oracle 连接账号对应的密码。

目标实例信息

- 实例类型：选择 **DRDS 实例**。
- DRDS 实例 ID：配置迁移的目标 DRDS 实例的实例 ID。
- DRDS 数据库：配置目标 DRDS 实例的数据库。
- 账号：DRDS 实例的连接账号。
- 密码：上面指定的 DRDS 实例连接账号对应的密码。

当配置完连接信息后，单击右下角 **授权白名单并进入下一步** 进行白名单授权。

迁移对象及迁移类型配置

迁移类型包括：全量数据迁移、增量数据迁移。默认选择 **全量数据迁移**，如果需要通过 Oracle->DRDS 的零停机迁移，需要选择**全量数据迁移+增量数据迁移**。

迁移对象，需要选择您要迁移的对象。迁移对象选择的粒度可以为：库、表、列三个粒度级别。默认情况下，对象迁移到 DRDS 实例后，对象名与源 Oracle 实例中的对象名保持一致。如果您迁移的对象在源实例跟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，详细使用方式可以参考库表列映射。

当配置完迁移对象及迁移类型后，即进入无主键表迁移方式的选择。

无主键表迁移方式选择

为了支持 Oracle 中无主键表的增量迁移，并保证迁移后数据保持一致，您需要选择在 DRDS 实例的数据库中，对于无主键表是否手工新增列：alibaba_dts_row_id。对于需要进行增量迁移的无主键表，用户需要在迁移的目标 DRDS 实例中创建对应的表时，手工添加这个列。在迁移过程中，这个列由 DTS 主动维护，该列在迁移过程中不能删除，否则会导致迁移失败。

如果选择新增列 alibaba_dts_row_id，那么可以保证无主键表的迁移不产生重复数据，保证迁移数据跟源实例数据完全一致，否则无主键表可能会出现比源实例多出部分数据的情况。

任务预检查

在迁移任务正式启动之前，会先进行前置预检查，只有预检查通过后，才能成功启动迁移。

如果预检查失败，那么可以点击具体检查项后的按钮，查看具体的失败详情，并根据失败原因修复后，重新进行预检查。



启动迁移任务

当预检查通过后，我们可以启动迁移任务，任务启动后，可以到任务列表中查看任务具体的迁移状态及进度。

当任务进入增量数据迁移阶段，任务不会自动停止，且一旦源 Oracle 实例有增量写入，增量数据就会自动同步到目标 DRDS 实例。增量数据迁移是个动态同步的过程，所以建议在增量迁移达到无延迟状态时，在目标数据库上进行业务验证，如果验证成功，那么可以停掉迁移任务，然后将业务切换到目标数据库。

至此，完成将本地 Oracle 数据库实例到 DRDS 实例的数据迁移任务配置。

适用场景

使用数据传输DTS可以将本地的Oracle的数据迁移至RDS For PostgreSQL Plus Advanced Server。同时，DTS提供的增量数据迁移功能，可以实现Oracle在提供服务的过程中，进行数据迁移，当Oracle跟PPAS两边数据达到动态同步过程，将业务切换到PPAS，使得依赖Oracle的应用停机时间最短。

本小节简单介绍使用DTS进行Oracle到RDS For PostgreSQL Plus Advanced Server不停机数据迁移的任务配置流程。

迁移类型

Oracle到RDS For PPAS的迁移，支持结构迁移、全量数据迁移及增量数据迁移。各个迁移阶段的支持范围如下：

结构迁移

DTS将迁移对象的结构定义迁移到目标实例。目前DTS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。

全量数据迁移

DTS将源数据库迁移对象的存量数据迁移到目标实例。

如果未选择增量数据迁移，那么Oracle数据库有数据写入的话，新增增量数据不一定能够被迁移到PPAS中。所以，如果要保证迁移数据一致性，那么尽量选择在业务低峰期进行停机迁移。

增量数据迁移

增量数据迁移会将迁移过程中，源Oracle实例的增量变更数据同步到目标PPAS实例，增量数据迁移可以实现Oracle实例在迁移过程中产生的增量数据被实时同步到目标库，最终Oracle同目标RDS PPAS实例进入动态数据同步的过程。

迁移限制

- 迁移过程中，不支持DDL操作。
- 不支持物化视图的迁移。
- 结构迁移时，reverse index迁移到RDS For PPAS中，存储成普通索引。
- 结构迁移时，位图索引迁移到RDS For PPAS，存储成普通索引。
- 结构迁移时，分区索引迁移到RDS For PPAS，在每个分区上创建独立的索引。
- 增量数据迁移只支持有主键，或有非空唯一索引的表。
- 增量数据迁移不支持long类型。

迁移权限要求

当使用DTS进行Oracle到RDS For PPAS迁移时，不同迁移类型，对源及目标数据库的迁移帐号权限要求如下：

迁移类型	结构迁移	全量迁移	增量数据迁移
本地Oracle实例	schema的owner	schema的owner	SYSDBA
目的RDS for PPAS实例	schema的owner	schema的owner	schema的owner

迁移前置条件

- Oracle版本为10g，11g，12c。
- Oracle 开启supplement log，且要求 supplemental_log_data_pk，supplemental_log_data_ui 开启。
- Oracle要求开启archive log。

数据类型映射关系

由于Oracle跟RDS For PPAS的数据类型不是——对应的，所以数据传输服务在进行结构迁移时，会根据两种数据库类型的数据类型定义，进行类型映射，下表为数据传输服务定义的数据类型映射关系。

Oracle数据类型	PPAS数据类型	数据传输服务是否支持
varchar2(n [char/byte])	varchar2[(n)]	支持
nvarchar2[(n)]	nvarchar2[(n)]	支持
char[(n [byte/char])]	char[(n)]	支持
nchar[(n)]	nchar[(n)]	支持
number[(p,s)]	number[(p,s)]	支持
float(p)	double precision	支持
long	long	支持
date	date	支持
binary_float	real	支持
binary_double	double precision	支持
timestamp[(fractional_seconds_precision)]	timestamp[(fractional_seconds_precision)]	支持
timestamp[(fractional_seconds_precision)]with time zone	timestamp[(fractional_seconds_precision)]with time zone	支持
timestamp[(fractional_seconds_precision)]with local time zone	timestamp[(fractional_seconds_precision)]with time zone	支持
clob	clob	支持
nclob	nclob	支持
blob	blob	支持
raw	raw(size)	支持
long raw	long raw	支持
bfile	—	不支持
interval year(year_precision) to month	interval year to month	不支持
interval day(day_precision) to second[(fractional_seconds_precision)]	interval day to second[(fractional_seconds_precision)]	不支持

由于RDS For PPAS不支持数据类型timestamp[(fractional_seconds_precision)]with local time zone，所以数据传输服务在迁移这种类型的数据时，会将其转换成UTC时区后，存入RDS For PPAS的数据类型timestamp[(fractional_seconds_precision)]with time zone中。

环境准备

创建迁移帐号

迁移任务配置，需要提供Oracle数据库及目标RDS实例的迁移账号。迁移账号所需权限详见上文的 迁移权限要求。

如果您的Oracle实例的迁移账号尚未创建，那么您可以参考 [Oracle Grant语法说明](#)，创建满足要求的迁移账号。

RDS For PPAS 迁移账号的创建及授权操作如下：

通过PostgreSQL客户端，在RDS For PPAS中创建迁移账号，语法如下所示：

```
create user username IDENTIFIDE BY password;
```

参数说明：

- username：要创建的账号。
- password：该账号的登录密码。

在RDS FOR PPAS中给迁移账号授权，语法如下所示：

```
GRANT privileges ON tablename TO username;
```

参数说明：

- privileges：该账号的操作权限，如SELECT、INSERT、UPDATE等。如果要授权账号所有权限，则使用ALL。
- table：表名。如果要授权该账号所有的表权限则使用通配符*。
- username：要授权的账号名。

操作步骤

以目标RDS实例所属阿里云账号登录DTS数据传输控制台，单击右上角的**创建迁移任务**，开始任务配置。

设置实例连接信息。这个步骤主要配置 迁移任务名称，Oracle连接信息及目标RDS实例连接信息。其中：

任务名称

DTS为每个任务自动生成一个任务名称，任务名称没有唯一性要求。您可以根据需要修改任务名称，建议为任务配置具有业务意义的名称，便于后续的任务识别。

源实例信息

- 实例类型：选择 **有公网IP的自建数据库**。
- 数据库类型：选择 **Oracle**。
- 主机名或IP地址：配置Oracle访问地址，这个地址必须为公网访问方式。
- 端口：Oracle实例的监听端口。
- SID：Oracle实例的SID。
- 数据库账号：Oracle实例的连接账号。
- 数据库密码：上面Oracle连接账号对应的密码。

目标实例信息

- 实例类型：选择 **RDS实例**。
- RDS实例ID：配置迁移的目标RDS实例的实例ID。DTS支持经典网络、VPC网络的RDS实例。
- 账号：RDS实例的连接账号。
- 密码：上面RDS实例连接账号对应的密码。

The screenshot displays the configuration interface for a DTS task, divided into three steps: 1. Source and Target (1.源库及目标库), 2. Migration Type and List (2.迁移类型及列表), and 3. Preview (3.预检查). The 'Source Instance Information' (源库信息) section is highlighted with a red box and contains the following fields: Task Name (任务名称) set to '江都测试'; Instance Type (实例类型) set to '有公网IP的自建数据库'; Database Type (数据库类型) set to 'Oracle'; Host Name or IP Address (主机名或IP地址) with a masked input; Port (端口) set to '1521'; SID set to 'orcl'; Database Account (数据库账号) with a masked input; and Database Password (数据库密码) with a masked input. The 'Target Instance Information' (目标库信息) section is also highlighted with a red box and contains: Instance Type (实例类型) set to 'RDS实例'; RDS Instance ID (RDS实例ID) with a masked input; Database Name (数据库名称) with a masked input; Database Account (数据库账号) set to 'dts'; and Database Password (数据库密码) with a masked input. At the bottom right, there are three buttons: '取消' (Cancel), '上一步' (Previous Step), and '授权白名单并进入下一步' (Authorize Whitelist and Proceed to Next Step), with the latter button highlighted by a red box.

当配置完连接信息后，单击右下角 **授权白名单并进入下一步** 进行白名单授权。这个步骤DTS会将DTS服务器的IP地址添加到目标RDS实例的白名单中，避免因RDS实例设置了白名单，导致DTS服务器连接不上RDS实例导致迁移失败。

迁移对象及迁移类型配置。

迁移类型

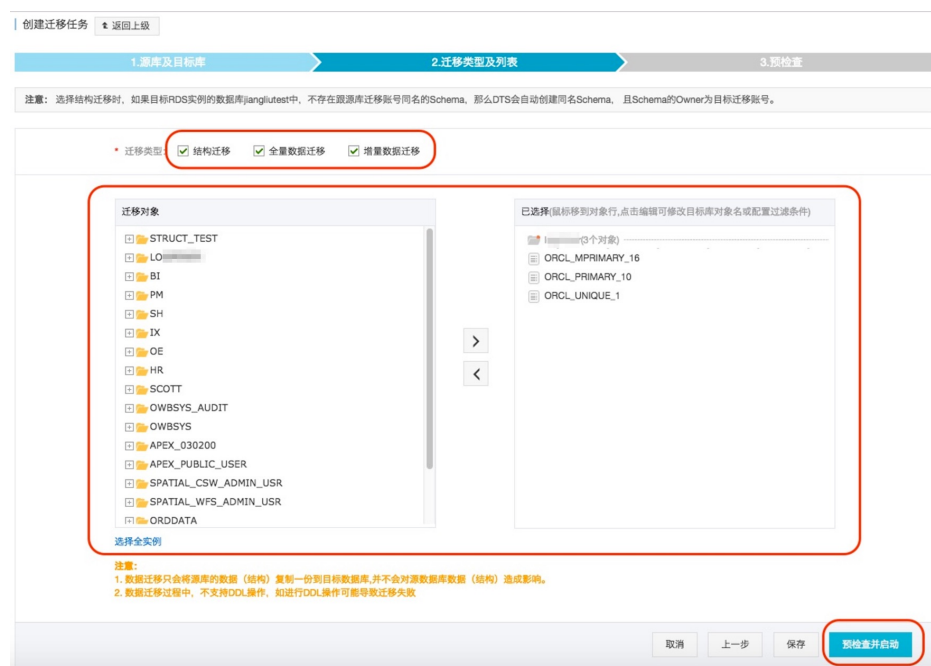
- 结构迁移
- 全量数据迁移
- 增量数据迁移

说明：

- 如果需要进行不停机迁移，那么需要选择：**结构迁移 + 全量数据迁移 + 增量数据迁移**。
- 如果只进行全量迁移，那么需要选择：**结构迁移 + 全量数据迁移**。

迁移对象

迁移对象，需要选择您要迁移的对象。迁移对象选择的粒度可以为：库、表、列三个粒度。默认情况下，对象迁移到RDS For PPAS实例后，对象名跟本地Oracle实例一致。如果您迁移的对象在源实例跟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，详细使用方式可以参考库表列映射。



当配置完迁移对象及迁移类型后，即进入任务启动前的预检查步骤。

预检查。

在迁移任务正式启动之前，会先进行前置预检查，只有预检查通过后，才能成功启动迁移。

如果预检查失败，那么可以单击具体检查项后的按钮，查看具体的失败详情，并根据失败原因修复后，重新进行预检查。



当修复完成后，可以在任务列表中，选择这个任务，重新启动预检查。

购买实例并启动任务。

当预检查通过后即进入迁移实例购买清单确认界面。请根据业务压力选择具体的规格并同意产品服务条款后，启动迁移任务。

当任务启动后，可以到任务列表中查看任务具体的迁移状态及进度。

增量数据迁移是个动态同步的过程，所以建议在增量迁移达到无延迟状态时，在目标数据库上进行业务验证，如果验证成功，那么可以停掉迁移任务，然后将业务切换到目标数据库。

至此，完成本地Oracle数据库到RDS for PPAS实例的不停机数据迁移的任务配置。

附：预检查内容

DTS在启动迁移之前，会进行前置预检查，本节简单介绍Oracle->RDS For PPAS的预检查内容：

预检查项	检查内容	备注
源库连接性检查	检查DTS服务同Oracle实例的连通性	(1) 填写信息是否有误？如果填写信息有误，请修改后重新预检查 (2) 检查Oracle是否开启监听端口
目的库连接性检查	检查DTS服务同目的RDS For PPAS实例的连通性	检查填写信息是否有误，如果有误请先修改后重新预检查
源库版本检查	检查Oracle实例的版本，DTS是否支持	DTS目前只支持10g,11g,12c三个版本
源库权限检查	检查Oracle实例访问账号的权限	如果权限不足，请参照上面的权

	是否满足要求	限要求一节授权后，重新预检查
目的库权限检查	检查RDS For PPAS实例访问账号的权限是否满足要求	如果权限不足，请参照上面的权限要求一节授权后，重新预检查
同名对象存在性检查	检查待迁移对象在目标RDS For PPAS是否已经存在	如果检查失败，请将目标库中这些已经存在的对象删除后，重新进行预检查
源端同名对象存在性检查	检查待迁移对象中，要迁入目标同一个schema的对象是否同名	如果检查失败，可以参考 库表列映射 将重名对象进行重命名
源库日志模式检查	检查源库是否开启archive log	如果未开启，请启用后，重新预检查
约束完整性检查	检查待迁移对象依赖的父对象是否迁移	如果检查失败，那么可以修改迁移对象，同时迁移依赖的父对象后，重新预检查
DBLINK存在性检查	检查源库是否存在DBLINK	如果存在，那么需要修改迁移对象，不选择DBLINK
增量拓扑冲突检查	检查同一个迁移对象是否已经存在在迁移链路	如果存在冲突链路，那么需要删除掉冲突链路后，重新预检查
字段类型检查	检查待迁移表的是否存在数据类型为long类型的字段	如果存在那么对应的表只能进行全量数据迁移，不能选择增量数据迁移
表是否存在主键或者唯一性非空索引检查	检查待迁移表是否包含主键或非空唯一键	如果存在那么对应的表只能进行全量数据迁移，不能选择增量数据迁移
补偿日志开启检查	检查是否开启 supplemental_log	如果未开启，请启用后，重新预检查

使用数据传输DTS可以将本地PostgreSQL实例的数据迁移到RDS For PostgreSQL实例。PostgreSQL迁移支持增量数据同步功能，可以实现在本地应用不停服的情况下，平滑完成PostgreSQL数据库的迁移工作。

本小节简单介绍使用DTS进行PostgreSQL->RDS for PostgreSQL数据迁移的任务配置流程。

迁移类型简介

结构迁移

数据传输DTS将迁移对象的结构定义迁移到目标实例。对于PostgreSQL,数据传输服务支持结构迁移的对象包含：**Table、trigger、view、sequence、function、user defined type、rule、domain、operation、aggregate。**

全量数据迁移

数据传输DTS将源数据库迁移对象的存量数据全部迁移到目标实例。

增量数据迁移

增量数据迁移将迁移过程中，本地PostgreSQL实例的增量更新数据同步到目标RDS实例，最终实现本地PostgreSQL同目标RDS实例进入动态数据同步的过程。使用增量数据迁移，可以实现在本地PostgreSQL正常提供服务的时候，平滑完成PostgreSQL->RDS For PostgreSQL的数据迁移。

迁移限制

- 全量迁移支持的PostgreSQL版本为：9.2、9.3、9.4、9.5；增量迁移支持的源实例PostgreSQL版本：9.4.8、9.5
- 迁移过程中，不支持DDL操作
- 不支持迁移使用C语言编写的function
- 如果使用了对象名映射功能后，依赖这个对象的其他对象可能迁移失败

迁移权限要求

当使用数据传输服务进行PostgreSQL迁移时，在不同迁移类型情况下，源跟目标数据库的迁移帐号权限要求如下：

迁移类型	结构迁移	全量迁移	增量数据迁移
本地PostgreSQL实例	pg_catalog的usage权限	迁移对象的select	superuser
目的RDS实例	迁移对象的create、usage权限	schema的owner	schema的owner

迁移顺序

数据传输服务在进行PostgreSQL->RDS For PostgreSQL数据迁移时，为了解决对象间的依赖，提高迁移成功率。结构对象及数据的迁移顺序如下：

1. 进行结构对象：Table、view、sequence、function、user defined type、rule、domain、operation、aggregate的迁移
2. 全量数据迁移
3. 进行结构对象：trigger、foreign key的迁移

全量数据迁移完成后，任务列表中的迁移进度为：结构迁移100%，全量迁移100%，迁移状态为“迁移中”，此时迁移任务正在进行步骤(3)中的对象的迁移。此时，请勿手动结束任务，否则会造成迁移数据丢失。

迁移步骤

下面详细介绍下使用DTS将本地的PostgreSQL数据库迁移到RDS for PostgreSQL的任务配置流程。

RDS实例数据库创建

在数据迁移过程中，如果待迁移的数据库在目标RDS实例中不存在，那么DTS会自动创建。但是对于如下两种情况，用户需要在配置迁移任务之前，手动创建数据库。

1. 数据库名称不符合：RDS定义规范(由小写字母、数字、下划线、中划线组成，字母开头，字母或数字结尾，最长64个字符)。
2. 待迁移数据库，在本地PostgreSQL跟目标RDS For PostgreSQL实例中存储名称不同。

对于这两种情况，用户需要在配置迁移任务之前，先在RDS控制台完成数据库创建。具体参考RDS数据库创建流程。

迁移帐号创建

迁移任务配置，需要提供Oracle数据库及目标RDS实例的迁移账号。迁移账号所需权限详见上文的 迁移权限要求。

如果您的本地PostgreSQL或RDS For PostgreSQL实例的迁移账号尚未创建，那么可以参考如下流程创建迁移账号：

通过PostgreSQL客户端，在PostgreSQL中创建迁移账号。

```
create user username password 'password'
```

如果您要使用增量迁移，那么创建的账号必须是superuser，所以创建账号语句调整为：

```
create user username with superuser password 'password'
```

参数说明：

- username：要创建的账号
- password：该账号的登录密码

给迁移账号授权，本地PostgreSQL及RDS For PostgreSQL的迁移账号权限要求详见上表。

```
GRANT privileges ON tablename TO username;
```

参数说明：

- privileges：该账号的操作权限，如 SELECT、INSERT、UPDATE 等。如果要授权该账号所有权限，则使用 ALL
- tablename：表名。如果要授权该账号所有的表权限，则使用通配符 *
- username：要授权的账号名

逻辑流复制插件安装

如果您需要使用增量数据迁移进行不停机迁移，那么在任务配置之前，需要在本地PostgreSQL中安装DTS提供的逻辑流复制插件。

插件下载。

本地PostgreSQL 9.4版本

本地PostgreSQL 9.5版本

插件安装。

- i. 解压下载的压缩包。

将ali_decoding.so文件拷贝到 PostgreSQL安装路径的lib目录下。

如果用rpm包安装，那么这个绝对路径为：/usr/pgsql-{\$version}/lib/，其中\$version为版本号，例如版本为9.5，那么绝对路径为:/usr/pgsql-9.5/lib。

将ali_decoding.contorl文件拷贝到PostgreSQL安装路径下的share/extension 目录下。

如果用rpm包安装，那么这个绝对路径为：/usr/pgsql-{\$version}/share/extension/，其中{\$version}为PostgreSQL版本号，如果版本为9.5，那么绝对路径为：/usr/pgsql-9.5/share/extension/。

测试安装是否成功。

使用superuser账号登录PostgreSQL客户端，运行如下SQL，看是否能够成功创建replication slot。如果成功创建那么说明插件安装成功。

```
SELECT * FROM pg_create_logical_replication_slot('replication_slot_test', 'ali_decoding');
```

如果输出结果如下，说明插件安装成功。

```
postgres=# SELECT * FROM pg_create_logical_replication_slot('replication_slot_test', 'ali_decoding');
 slot_name | xlog_position
-----|-----
 replication_slot_test | 0/17D3F58
(1 row)
```

测试成功后，使用如下SQL将replication slot删除掉。

```
SELECT pg_drop_replication_slot('replication_slot_test');
```

迁移任务配置

当上面的所有前置条件都配置完成后，就可以开始正式的数据迁移了。下面详细介绍迁移任务配置流程。

1. 进入数据传输DTS控制台，点击右上角的创建迁移任务，开始迁移任务配置。

本地PostgreSQL跟目标RDS For PostgreSQL实例连接信息配置。

这个步骤主要配置 迁移任务名称，PostgreSQL连接信息及目标RDS For PostgreSQL实例连接信息。其中：

任务名称

DTS为每个任务自动生成一个任务名称，任务名称没有唯一性要求。您可以根据需要修改任务名称，建议为任务配置具有业务意义的名称，便于后续的任务识别。

源实例信息

- 实例类型：选择 **有公网IP的自建数据库**
- 数据库类型：选择 **PostgreSQL**
- 主机名或IP地址：配置PostgreSQL访问地址，这个地址必须为公网访问方式
- 端口：本地PostgreSQL实例的监听端口
- 数据库名称：连接PostgreSQL的默认数据库名
- 数据库账号：本地PostgreSQL实例的连接账号
- 数据库密码：本地PostgreSQL实例连接账号对应的密码

目标实例信息

- 实例类型：选择 **RDS实例**
- RDS实例ID：配置迁移的目标RDS实例的实例ID。DTS支持经典网络、VPC网络的RDS实例
- 数据库名称：连接RDS For PostgreSQL的默认数据库名
- 数据库账号：RDS For PostgreSQL实例的连接账号
- 数据库密码：上面RDS For PostgreSQL实例连接账号对应的密码

The screenshot shows the configuration interface for DTS, divided into three steps: 1. Source and Target (1.源库及目标库), 2. Migration Type and List (2.迁移类型及列表), and 3. Pre-check (3.预检查). The current step is 1. Source and Target.

源库信息 (Source Database Information):

- 任务名称: dtsekioxwmu
- 实例类型: 有公网IP的自建数据库
- 数据库类型: PostgreSQL
- 主机名或IP地址: [Redacted]
- 端口: 5432
- 数据库名称: postgres
- 数据库账号: jiangludest
- 数据库密码: [Redacted]

目标库信息 (Target Database Information):

- 实例类型: RDS实例
- RDS实例ID: [Redacted]
- 数据库名称: postgres
- 数据库账号: dts
- 数据库密码: [Redacted]

Buttons at the bottom: 取消 (Cancel), 上一步 (Previous Step), 授权白名单并进入下一步 (Authorize Whitelist and Enter Next Step).

当配置完连接信息后，点击右下角 授权白名单并进入下一步 进行白名单授权。这个步骤DTS会将DTS服务器的IP地址添加到目标RDS实例的白名单中，避免因为RDS实例设置了白名单，导致DTS服务器连接不上RDS实例导致迁移失败。

选择迁移对象及迁移类型。

迁移类型

对于PostgreSQL->RDS For PostgreSQL，支持 结构迁移、全量数据迁移、增量数据迁移。

。

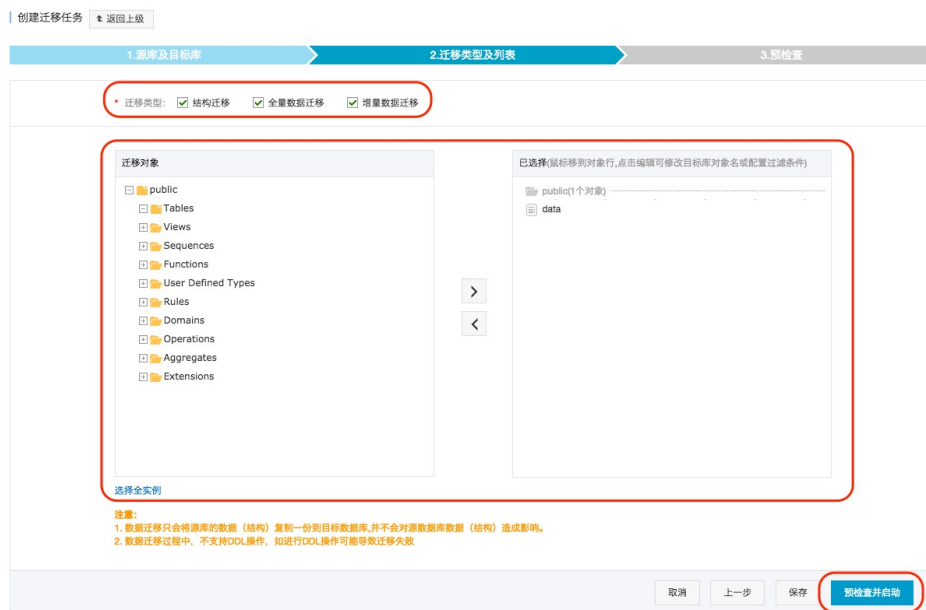
如果只需要进行全量迁移，那么迁移类型选择：结构迁移 + 全量数据迁移。

如果需要进行不停机迁移，那么迁移类型选择：结构迁移 + 全量数据迁移 + 增量数据迁移。

。

迁移对象

选择您要迁移的对象。迁移对象选择的粒度可以为：库、表、列三个粒度。默认情况下，对象迁移到RDS for PostgreSQL实例后，对象名跟本地PostgreSQL实例一致。如果您迁移的对象在源实例跟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，详细使用方式可以参考库表列映射。



预检查。

在迁移任务正式启动之前，会先进行前置预检查，只有预检查通过后，才能成功启动迁移。预检查的内容及修复方式可以参考本文末尾的 **预检查简介** 一节。

如果预检查失败，那么可以点击具体检查项后的按钮，查看具体的失败详情，并根据失败原因修复后，重新进行预检查。



启动迁移任务。

当预检查通过后，可以启动迁移任务，任务启动成功后，可以在任务列表中查看迁移的具体状态及迁移进度。

至此，完成本地PostgreSQL数据库到RDS for PostgreSQL实例的数据迁移任务配置。

预检查内容

DTS在启动迁移之前，会进行前置预检查，本小节简单介绍PostgreSQL->RDS for PostgreSQL的预检查内容：

预检查项	检查内容	备注
源库连接性检查	检查DTS服务同本地PostgreSQL实例的连通性	(1) 填写信息是否有误？如果填写信息有误，请修改后重新预检查 (2) 检查端口是否配置从其他服务器连接
目的库连接性检查	检查DTS服务同目的RDS For PostgreSQL实例的连通性	检查填写信息是否有误，如果有误请先修改后重新预检查
目的库连接性检查	检查DTS服务同目的RDS For PostgreSQL实例的连通性	检查填写信息是否有误，如果有误请先修改后重新预检查
源库版本检查	检查本地PostgreSQL版本跟目标RDS For PostgreSQL是否一致	如果版本不一致，预检查会有提醒。可以根据提醒对本地PostgreSQL进行升级或降级，也可以继续迁移
数据库可用性检查	检查待迁移数据库在目标RDS For PostgreSQL是否已存在	如果待迁移数据库命名规范不满足RDS实例要求，那么DTS在目标RDS实例创建待迁移数据库会报错失败，即数据库可用性检查会失败。此时可以参考 库表列映射 对迁移数据库进行重命名
源库权限检查	检查任务配置时，提供的本地PostgreSQL数据库账号的权限是否满足要求	如果检查失败，那么请参考本文 迁移账号创建 一节对迁移账号进行授权后，重新进行预检查
目的库权限检查	检查任务配置时，提供的目的RDS For PostgreSQL数据库账号的权限是否满足要求	如果检查失败，那么请参考本文 迁移账号创建 一节对迁移账号进行授权后，重新进行预检查
同名对象存在性检查	检查待迁移对象在目标RDS For PostgreSQL实例中是否已经存在	如果检查失败，请将目标库中这些已经存在的对象删除后，重新进行预检查
源端同名对象存在性检查	检查要迁移到同一个数据库中的多个对象是否重名	如果检查失败，可以参考 库表列映射 将重名对象进行重命名
约束完整性检查	检查待迁移对象依赖的父对象是否迁移	如果检查失败，那么可以修改迁移对象，同时迁移依赖的父对象后，重新预检查
增量拓扑冲突检查	检查同一个迁移对象是否已经存在在迁移链路	如果存在冲突链路，那么需要删除掉冲突链路后，重新预检查

LC_MONETERY 一致性检查	检查源库、目标库的Lc_monetary定义是否一致	如果检查失败，可以修改目标RDS For PostgreSQL的LC_MONETERY定义，或者继续迁移
PostgreSQL逻辑流复制插件检查	检查本地PostgreSQL实例是否安装了逻辑流复制插件	如果检查失败，可以参考下面的 逻辑流复制插件安装 一节安装逻辑流复制插件后重新预检查
PostgreSQL逻辑流slot存在性检查	检查本地PostgreSQL实例存在跟DTS创建的replication slot重名的replication slot	如果检查失败，可以删除本地PostgreSQL中已经存在的同名replication slot后，重新预检查

使用数据传输DTS可以将本地的MongoDB实例的数据迁移至RDS For MongoDB。Mongoddb迁移支持增量数据同步功能，可以实现在本地应用不停服的情况下，平滑完成MongoDB数据库的迁移工作。

本小节简单介绍使用DTS进行MongoDB->RDS For MongoDB数据迁移的任务配置流程。

迁移类型简介

对于MongoDB->云数据库 MongoDB数据迁移，DTS支持全量数据迁移及增量数据迁移，这个迁移类型的功能及限制如下。

全量数据迁移

数据传输DTS将源数据库迁移对象的存量数据全部迁移到目标实例。

增量数据迁移

增量数据迁移将迁移过程中，本地MongoDB实例的增量更新数据同步到云数据库MongoDB，最终本地MongoDB同云数据库MongoDB进入动态数据同步的过程。使用增量数据迁移，可以实现在本地MongoDB正常提供服务的时候，平滑完成MongoDB->RDS For MongoDB的数据迁移。

迁移功能

MongoDB版本支持支持MongoDB 3.0,3.2 两个版本

MongoDB全量数据迁移支持

- 支持database 迁移
- 支持collection 迁移
- 支持index 迁移

MongoDB增量数据迁移支持

- 支持document 新增、删除、更新 操作的同步
- 支持collection新建、删除 操作的同步
- 支持database 新建、删除 操作的同步
- 支持index 新建、删除 操作的同步

迁移权限要求

当使用DTS进行MongoDB-> 云数据库MongoDB迁移时，不同迁移类型，对源跟目标MySQL实例的迁移帐号权限要求如下：

迁移类型	全量数据迁移	增量数据迁移
本地MongoDB	read	待迁移库的read admin的read权限 local的read权限
目的MongoDB实例	readWrite	readWrite

迁移任务配置

下面详细介绍使用DTS将本地的MongoDB迁移到云数据库MongoDB的任务配置流程。

迁移帐号创建

迁移任务配置，需要提供本地自建MongoDB实例及目标MongoDB云数据库的迁移账号，迁移账号所需权限详见上面的 **迁移权限要求** 一节。

如果您的本地MongoDB或MongoDB云数据库的迁移账号尚未创建，那么可以参考如下流程创建迁移账号：

```
db.createUser({user:"username",pwd:"password",roles:[{role:"rolename1",db:"database_name1"},{role:"rolename2",db:"database_name2"}]})
```

参数说明：

- username: 要创建的账号
- password: 上面账号对应的密码
- rolename1/rolename2：待授权给username的角色名，例如上面的read, readWrite

database_name1/database_name2: 只将database_name1/database_name2上的角色role1/role2授权给username

关于MongoDB的角色授权也可以参考 [MongoDB Create User说明](#)

迁移任务配置

当上面的所有前置条件都配置完成后，就可以开始正式的数据迁移了。下面详细介绍迁移任务配置流程。

1. 进入数据传输DTS控制台，点击右上角的创建迁移任务，开始迁移任务配置。

本地MongoDB跟目标MongoDB实例的连接信息配置。

这个步骤主要配置 迁移任务名称，本地MongoDB连接信息及目标MongoDB云数据库连接信息。
其中：

任务名称

DTS为每个任务自动生成一个任务名称，任务名称没有唯一性要求。您可以根据需要修改任务名称，建议为任务配置具有业务意义的名称，便于后续的任务识别。

源实例信息

- 实例类型：选择 有公网IP的自建数据库
- 数据库类型：选择 MongoDB
- 主机名或IP地址：配置MongoDB访问地址，这个地址必须为公网访问方式
- 端口：本地MongoDB实例的监听端口
- 数据库名称：连接MongoDB实例的默认数据库名
- 数据库账号：本地MongoDB实例的连接账号
- 数据库密码：本地MongoDB实例连接账号对应的密码

目标实例信息

- 实例类型：选择 MongoDB实例
- MongoDB实例ID：配置迁移的目标MongoDB实例的实例ID。DTS支持经典网络的MongoDB实例。如果您的MongoDB实例为VPC网络的实例，那么需要切换到经典网络模式后，再使用DTS进行迁移
- 数据库名称：连接MongoDB的默认数据库名
- 数据库账号：连接MongoDB实例的连接账号
- 数据库密码：上面数据库账号对应的密码

当配置完连接信息后，点击右下角 授权白名单并进入下一步 进行白名单授权。这个步骤DTS会将DTS服务器的IP地址添加到目标MongoDB云数据库的白名单中，避免因为MongoDB实例设置了白名单，DTS服务器连接不上MongoDB实例导致迁移失败。

The screenshot displays the configuration page for a data migration task, divided into three steps: 1. Source and Target Databases, 2. Migration Type and List, and 3. Pre-check. The 'Task Name' field is set to 'dtslqp77cpv'. The 'Source Database Information' section includes fields for Instance Type (Public IP自建数据库), Database Type (MongoDB), Hostname/IP Address, Port (27017), Database Name (admin), Database Username (dba), and Database Password. The 'Target Database Information' section includes Instance Type (MongoDB实例), MongoDB Instance ID, Database Name (admin), Database Username (root), and Database Password. A 'Authorize Whitelist and Enter Next Step' button is highlighted at the bottom right.

选择迁移对象及迁移类型。

迁移类型

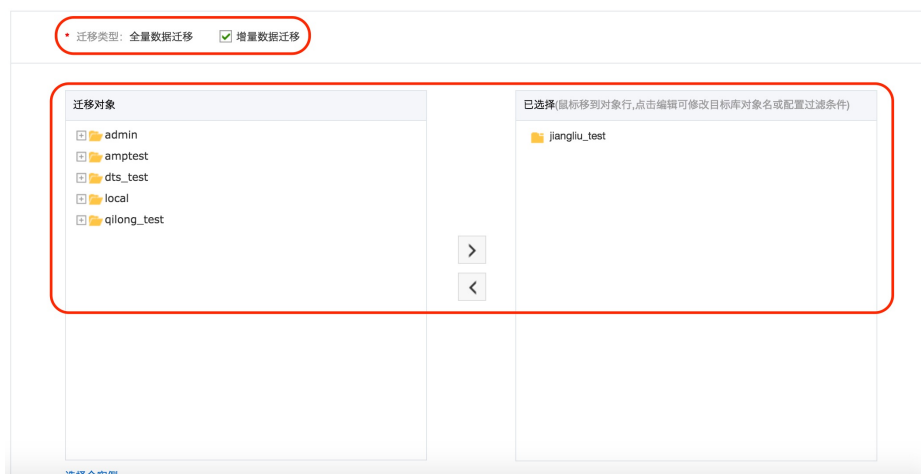
对于MongoDB，支持 全量数据迁移、增量数据迁移。

如果只需要进行全量迁移，那么迁移类型选择：全量数据迁移。

如果需要不停机迁移，那么迁移类型选择：全量数据迁移 + 增量数据迁移。

迁移对象

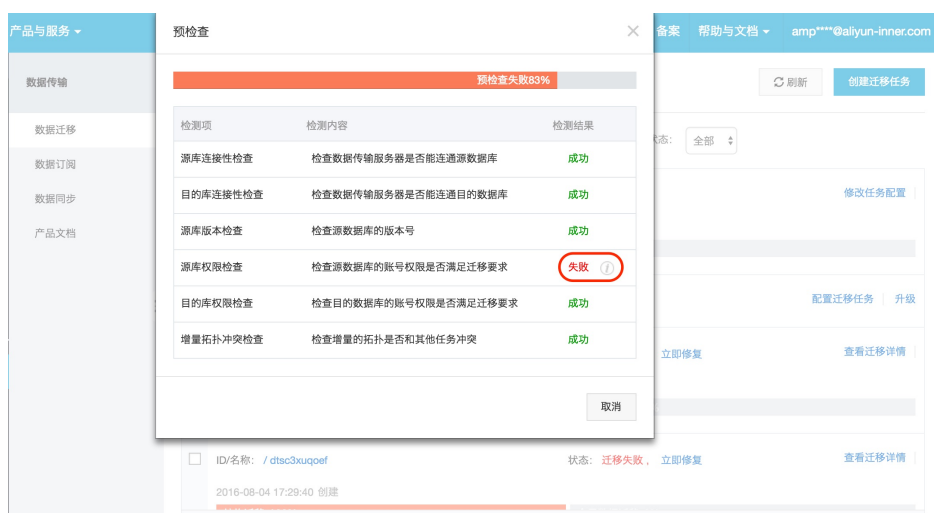
这个步骤需要选择要迁移的对象。迁移对象选择的粒度可以为：库、collection/function 两个粒度。默认情况下，对象迁移到MongoDB实例后，对象名跟本地MongoDB实例一致。如果您迁移的对象在源实例跟目标实例上名称不同，那么需要使用DTS提供的对象名映射功能，详细使用方式可以参考库表映射。



预检查。

在迁移任务正式启动之前，会先进行前置预检查，只有预检查通过后，才能成功启动迁移。预检查的内容及修复方式可以参考本文末尾的 **预检查简介** 一节。

如果预检查失败，那么可以点击具体检查项后的按钮，查看具体的失败详情，并根据失败原因修复后，重新进行预检查。



启动迁移任务。

当预检查通过后，可以启动迁移任务，任务启动成功后，可以在任务列表中查看迁移的具体状态及迁移进度。

至此，完成本地MongoDB数据库到MongoDB实例的数据迁移任务配置。

预检查

DTS在启动迁移之前，会进行前置预检查，本小节简单介绍MongoDB->RDS For MongoDB的预检查内容：

检查项	检查内容	备注
源库连接性检查	检查DTS服务器跟本地MongoDB实例的连通性	(1) 填写信息是否有误？如果填写信息有误，请修改后重新预检查 (2) 检查端口是否配置从其他服务器连接访问
目标库连接性检查	检查DTS服务器跟目标MongoDB实例的连通性	检查填写信息是否有误，如果有误请先修改后重新预检查
源库版本检查	检查本地MongoDB的版本，DTS是否支持	请先升级版本到3.2后，重新预检查
源库权限检查	检查本地MongoDB提供的迁移账号的权限是否满足需求	如果检查失败，请参照 迁移账号创建 一节授权后，重新预检查
目的库权限检查	检查目的MongoDB数据库提供的迁移账号的权限是否满足需求	如果检查失败，请参照 迁移账号创建 一节授权后，重新预检查
增量拓扑冲突检查	检查MongoDB实例上是否有其他增量迁移任务正在运行	如果检查失败，那么需要结束或删除其他的增量迁移任务后，重新预检查

使用数据传输DTS可以进行Redis实例间的数据迁移。如果源实例为自建Redis，那么Redis迁移支持增量数据同步功能，可以实现在本地应用不停写的情况下，平滑完成Redis数据迁移。

本小节简单介绍使用DTS进行Redis->云Redis实例数据迁移的迁移流程。云Redis实例间的迁移过程也可以参考这个流程

迁移类型简介

当迁移源实例为自建Redis时，可以支持全量数据迁移 + 增量数据迁移，当迁移源实例为云Redis实例时，目前只支持全量数据迁移。全量数据迁移及增量数据迁移的功能及限制如下。

全量数据迁移

数据传输DTS将自建Redis中现有的Key全部迁移到云Redis实例中。

增量数据迁移

增量数据迁移将迁移过程中，自建Redis实例的更新key同步到云数据库Redis。最终，自建Redis同云Redis实例进入动态数据复制的过程。通过增量数据迁移，可以实现在自建Redis正常提供服务的时候，平滑完成Redis->云Redis实例的数据迁移。

迁移功能

Redis增量迁移支持的命令包括：

APPEND
BITOP, BLPOP, BRPOP, BRPOPLPUSH,
DECR, DECRBY, DEL,
EVAL, EVALSHA, EXEC, EXPIRE, EXPIREAT,
FLUSHALL, FLUSHDB,
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

迁移前置条件

VPC Redis或通过专线接入阿里云VPC的自建Redis

如果待迁移的redis是通过专线接入阿里云vpc的自建redis，或是专有网络的云Redis实例，那么需要架设代理，进行数据转发。

为了让DTS能够访问专有网络Redis/通过专线接入阿里云的自建Redis，需要在VPC内选择一台有公网EIP的ECS，并在ECS上部署nginx，通过nginx进行代理转发。

对于专有网络的Redis实例，ECS所在的VPC必须同专有网络Redis在同一个VPC，对于通过专线接入阿里云的自建Redis，ECS所在的VPC必须为专线对端的阿里云VPC。

下面介绍如何使用nginx进行redis的转发配置，让DTS服务器可以访问专有网络内的redis实例。

(1) nginx部署

首先在ECS服务器上，通过如下命令部署nginx。

```
yum install nginx
```

(2) nginx转发配置

安装完nginx，修改nginx配置文件 /etc/nginx/nginx.conf，设置后端监听redis。注释掉配置文件中http的相关配置，添加tcp的配置内容。需要注释掉的http配置内容如下：

```

# http {
#     log_format main '$remote_addr - $remote_user [$time_local] "$request" '
#                     '$status $body_bytes_sent "$http_referer" '
#                     '"$http_user_agent" "$http_x_forwarded_for"';
#
#     access_log /var/log/nginx/access.log main;
#
#     sendfile        on;
#     tcp_nopush      on;
#     tcp_nodelay     on;
#     keepalive_timeout 65;
#     types_hash_max_size 2048;
#
#     include         /etc/nginx/mime.types;
#     default_type    application/octet-stream;
#
#     # Load modular configuration files from the /etc/nginx/conf.d directory.
#     # See http://nginx.org/en/docs/nginx\_core\_module.html#include
#     # for more information.
#     include /etc/nginx/conf.d/*.conf;
#}

```

需要在配置文件中添加tcp配置内容：

```

stream{
    upstream backend{
        hash $remote_addr consistent;
        #设置后端Redis连接串和端口，失败超时时间为10s，最多尝试3次。
        server r-bp1b294374634044.redis.rds.aliyuncs.com:6379 max_fails=3 fail_timeout=10s;
    }
    server{
        # nginx访问端口
        listen 3333;
        #指定nginx连接后端服务器的超时时间，指定为20s。
        proxy_connect_timeout 20s;
        #距离上一次成功访问（连接或读写）后端服务器的时间超过了5分钟就判定为超时，断开此连接。
        proxy_timeout 5m;
        #将TCP连接及数据收发转向叫“backend”的后端服务器。
        proxy_pass backend;
    }
}

```

例如，需要访问的redis的连接地址为：r-bp1b294374634044.redis.rds.aliyuncs.com:6379, nginx代理转发端口为3333, 那么tcp相关配置如下：

```

stream{
    upstream backend{
        hash $remote_addr consistent;
        server r-bp1b294374634044.redis.rds.aliyuncs.com:6379 max_fails=3 fail_timeout=10s;
    }
    server{
        listen 3333;
        proxy_connect_timeout 20s;
        proxy_timeout 5m;
        proxy_pass backend;
    }
}

```

(3) 通过转发接口访问redis

当完成上面的配置后，运行nginx即成功启动nginx代理服务。

假设nginx部署的ECS服务的EIP为：114.55.89.152,那么可以直接用redis_cli访问nginx转发端口，测试代理转发是否正常。

```
[root@iZ23bsc1llgZ ~]# redis-cli -h 114.55.89.152 -p 3333
114.55.89.152:3333> auth 
OK
114.55.89.152:3333> ping
PONG
114.55.89.152:3333> get a1
"10"
```

如上图所示，可以通过访问nginx代理转发端口来访问Redis实例。

下面配置DTS迁移任务时，对于专有网络Redis实例或对于通过专线接入阿里云的自建Redis，直接配置nginx代理的连接地址即可。

DTS支持vpc后，对于专有网络Redis实例或者通过专线接入阿里云的自建Redis无需再架设代理。DTS支持vpc的产品时间参考官网通知。

云Redis实例

由于老版本的云Redis实例(即Kvstore)不支持config get 命令，无法直接使用DTS将云Redis实例迁移到其他Redis实例。如果您需要支持云Redis实例到其他Redis实例的迁移，需提交工单，申请迁移源实例(云Redis)的版本升级。请注意：源实例的版本升级会存在实例闪断风险。升级完成后，可以直接使用DTS配置云Redis实例到其他Redis实例的数据迁移。

迁移任务配置

当上面的所有前置条件都配置完成后，就可以开始正式的数据迁移了。本小节将以 通过专线接入阿里云的自建Redis实例 到 经典网络云Redis实例的迁移为例，详细介绍迁移任务配置流程。

一、进入数据传输DTS控制台，点击右上角的创建迁移任务，开始配置迁移任务。

二、实例连接信息配置

这个步骤主要配置 迁移任务名称，自建Redis连接信息及云Redis实例连接信息。其中：

- 任务名称

DTS为每个任务自动生成一个任务名称，任务名称没有唯一性要求。您可以根据需要修改任务名称，建议为任务配置具有业务意义的名称，便于后续的任务识别。

源实例信息

实例类型：选择 有公网IP的自建数据库

实例区域：对于自建Redis，选择跟Redis实例物理距离最近的地域。选择的地域离Redis实例越近迁移性能越高。

数据库类型：选择 Redis

实例模式：默认为 单机，后续DTS将支持集群模式的Redis实例

主机名或IP地址：自建Redis实例的访问地址，如果配置了nginx转发，那么为nginx转发的访问地址

端口：自建Redis实例的监听端口。如果配置了nginx转发，那么为nginx转发端口

数据库密码：自建Redis实例访问密码，为非必填项，如果自建Redis没有设置密码，那么可以不填

目标实例信息

实例类型：Redis实例

实例区域：实例区域为云Redis实例所在区域

Redis实例ID：配置迁移的目标云Redis实例的实例ID

数据库密码：访问Redis实例的密码

当连接信息配置完成后，即可点击页面右下角的 **授权白名单并进入下一步**，开始进行迁移库的选择。

三、选择迁移对象及迁移类型

在这个步骤中，需要配置迁移类型及迁移对象。

- 迁移类型

对于Redis，DTS支持 全量数据迁移及增量数据迁移。

如果只需要进行全量迁移，那么迁移类型选择：全量数据迁移。

如果需要进行源库不停写迁移，那么迁移类型选择：全量数据迁移 + 增量数据迁移

- 迁移对象

这个步骤需要选择要迁移的库。目前Redis只支持整库迁移，所以只能选择要迁移的库，而不能选择要部分Key。



四、预检查

在迁移任务正式启动之前，会先进行前置预检查，只有预检查通过后，才能成功启动迁移。预检查的内容及修复方式可以参考本文末尾的 [预检查简介](#) 一节。

如果预检查失败，那么可以点击具体检查项后的按钮，查看具体的失败详情，并根据失败原因修复后，重新进行预检查。



五、启动迁移任务

当预检查通过后，可以启动迁移任务，任务启动成功后，可以在任务列表中查看迁移的具体状态及迁移进度。至此，完成自建Redis到云Redis实例的数据迁移任务配置。

预检查

DTS在启动迁移之前，会进行前置预检查，本小节简单介绍Redis数据迁移的预检查内容：

检查项	检查内容	备注
-----	------	----

源库连接性检查	检查DTS服务器跟自建Redis实例的连通性	(1) 填写信息是否有误？如果填写信息有误，请修改后重新预检查 (2) 检查端口是否允许从其他服务器连接访问
目标库连接性检查	检查DTS服务器跟目标Redis实例的连通性	检查填写信息是否有误，如果有误请先修改后重新预检查
库一对一检查	检查是否存在多个库迁移到一个库的情况	DTS暂不支持多个库迁移到一个库，如果出现这种情况，那么请先修改任务配置后，重新预检查
目标库是否为空	检查待迁移库在目标Redis实例中是否为空	如果检查失败，请先删除目标Redis实例中对应库的Key后，重新预检查
增量拓扑冲突检查	检查目标Redis实例上是否有其他增量迁移任务正在运行	如果检查失败，那么需要结束其他的增量迁移任务后，重新预检查

使用数据传输DTS可以实现两个RDS实例间的数据迁移。对于支持增量迁移的存储引擎，还可以使用DTS在源RDS实例不停服的情况下，将数据迁移到目标RDS实例。目前对于RDS不同存储引擎，只支持同构迁移(例如RDS For MySQL迁移到RDS For MySQL)，不支持异构迁移(例如RDS For MySQL迁移到RDS For SQLServer)。

本小节简单介绍使用DTS进行两个RDS实例之间的数据迁移的任务配置流程。

迁移权限要求

当使用DTS进行两个RDS实例间的数据迁移时，不同迁移类型，对源跟目标数据库的迁移帐号权限要求如下表：

迁移类型	结构迁移	全量迁移	增量迁移
源RDS实例	读写权限	读写权限	读写权限
目的MySQL	读写权限	读写权限	读写权限

迁移任务配置

下面详细介绍下用户如何使用DTS实现两个RDS MySQL实例间的数据迁移。源RDS跟目标RDS实例可以不同也可以相同，即用户可以使用DTS实现一个RDS实例内部的数据迁移。

RDS实例数据库创建

在数据迁移过程中，如果待迁移的数据库在目标RDS实例中不存在，那么DTS会自动会创建。但是对于如下两种情况，用户需要在配置迁移任务之前，手动创建数据库。

1. 数据库名称不符合：RDS定义规范(由小写字母、数字、下划线、中划线组成，字母开头，字母或数字结尾，最长64个字符)。
2. 待迁移数据库，在本地PostgreSQL跟目标RDS For PostgreSQL实例中存储名称不同。

对于这两种情况，用户需要在配置迁移任务之前，先在RDS控制台完成数据库创建。具体参考RDS数据库创建流程。

迁移帐号创建

迁移任务配置时，需要提供源RDS实例及目的RDS实例的迁移账号。迁移账号的相关权限详见上面的 迁移权限要求 一节。如果尚未创建迁移账号，那么可以参考 RDS实例账号创建需要先在源及目的RDS实例中创建迁移帐号，并将要迁移的库表的读写权限授权给上面创建的帐号。在RDS控制台中创建帐号，并将迁移库的读写权限授权给这个帐号。

迁移任务配置

当上面的所有前置条件都配置完成后，就可以开始正式的数据迁移了。下面详细介绍迁移任务配置流程。

1. 进入数据传输DTS控制台，点击右上角的**创建迁移任务**，开始迁移任务配置。

源及目的实例连接信息配置。

这个步骤主要配置 迁移任务名称，源RDS连接信息及目标RDS实例连接信息。其中：

任务名称

DTS为每个任务自动生成一个任务名称，任务名称没有唯一性要求。您可以根据需要修改任务名称，建议为任务配置具有业务意义的名称，便于后续的任务识别。

源实例信息

- 实例类型：选择 RDS实例
- RDS实例ID: 配置迁移的源RDS实例的实例ID。DTS支持经典网络、VPC网络的RDS实例
- 数据库名称：如果RDS实例的存储引擎为PostgreSQL或PPAS, 那么需要配置数据库名称。这个数据库名称为连接RDS实例使用的默认数据库
- 数据库账号：连接RDS实例的账号
- 数据库密码：上面数据账号对应的密码

目标实例信息

- 实例类型：选择 RDS实例
- RDS实例ID：配置迁移的目标RDS实例的实例ID。DTS支持经典网络、VPC网络的RDS实例
- 数据库名称：如果RDS实例的存储引擎为PostgreSQL或PPAS, 那么需要配置数

数据库名称。这个数据库名称为连接RDS实例使用的默认数据库

- 数据库账号：连接RDS实例的账号
- 数据库密码：上面数据账号对应的密码

创建迁移任务 返回上级

1. 源库及目标库 2. 迁移类型及列表 3. 预检查

* 任务名称: dtscl7zrbhj

源库信息

* 实例类型: RDS实例

* RDS实例ID: [Redacted]

* 数据库账号: dts

* 数据库密码: [Redacted]

其他阿里云账号下的RDS实例

目标库信息

* 实例类型: RDS实例

* RDS实例ID: [Redacted]

* 数据库账号: dts

* 数据库密码: [Redacted]

取消 上云评估 授权白名单并进入下一步

当配置完连接信息后，点击右下角 授权白名单并进入下一步 进行白名单授权。这个步骤DTS会将DTS服务器的IP地址添加到目标RDS实例的白名单中，避免因RDS实例设置了白名单，导致DTS服务器连接不上RDS实例导致迁移失败。

选择迁移对象及迁移类型。

迁移类型

DTS迁移类型支持结构迁移、全量数据迁移及增量迁移。

如果只需要进行全量迁移，那么迁移类型选择：结构迁移 + 全量数据迁移。

如果需要进行不停机迁移，那么迁移类型选择：结构迁移 + 全量数据迁移 + 增量数据迁移

。

迁移对象

这个步骤选择要迁移的对象。迁移对象的选择粒度细化为：库、表、列三个粒度。默认情况下，对象迁移到目标RDS实例后，对象名跟源RDS实例一致。如果您迁移的对象在源实例跟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，详细使用方式可以参考库表列映射。



预检查。

在迁移任务正式启动之前，会先进行前置预检查，只有预检查通过后，才能成功启动迁移。

如果预检查失败，那么可以点击具体检查项后的按钮，查看具体的失败详情，并根据失败原因修复后，重新进行预检查。



迁移任务。

当预检查通过后，可以启动迁移任务，任务启动成功后，可以在任务列表中查看迁移的具体状态及迁移进度。

如果选择了增量迁移，那么进入增量迁移阶段后，源库的更新写入都会被DTS同步到目标RDS实例。迁移任务不会自动结束。如果用户只是为了迁移，那么建议在增量迁移无延迟的状态时，源库停写几分钟，等待增量迁移再次进入无延迟状态后，停止掉迁移任务，直接将业务切换到目标RDS实例上即可。

至此，完成两个RDS实例的数据迁移任务配置。

适用场景

使用数据传输DTS可以实现ECS自建数据库到RDS/MongoDB/Redis/DRDS/PetaData/OceanBase实例的数据迁移。对于支持增量迁移的存储引擎，还可以使用DTS在ECS自建数据库不停服的情况下，将数据迁移到目标实例。

本小节简单介绍使用DTS进行ECS自建数据库迁移到RDS/MongoDB/Redis/DRDS/PetaData/OceanBase等实例的数据迁移任务的配置流程。

环境准备

创建RDS实例数据库

如果目标实例为RDS实例，如果待迁移的数据库在目标RDS实例中不存在，那么DTS会自动创建。但是对于如下两种情况，用户需要在配置迁移任务之前，手动创建数据库。

- 数据库名称不符合：RDS定义规范（由小写字母、数字、下划线、中划线组成，字母开头，字母或数字结尾，最长64个字符）。

待迁移数据库，在ECS自建数据库跟目标RDS实例中存储名称不同。

对于这两种情况，用户需要在配置迁移任务之前，先在RDS控制台完成数据库创建。具体参考RDS数据库创建流程。

创建迁移帐号

迁移任务配置时，需要提供ECS自建数据库及目的实例的迁移账号。

自建数据库的存储引擎不同，账号创建及权限授权的方式不同，具体授权方式可以参考各存储引擎的官方文档。

目的实例的迁移账号可以参考各个云产品的使用手册，进行账号创建。例如，RDS实例的迁移账号的创建流程可以参考 RDS账号创建说明。

操作步骤

以目标实例所属阿里云账号登录DTS数据传输控制台，单击右上角的**创建迁移任务**，开始任务配置。

设置实例连接信息。

这个步骤主要配置迁移任务名称、源ECS连接信息及目标实例连接信息。其中：

任务名称

DTS为每个任务自动生成一个任务名称，任务名称没有唯一性要求。您可以根据需要修改任务名称，建议为任务配置具有业务意义的名称，便于后续的任务识别。

源实例信息

- 实例类型：选择 ECS上的自建数据库
- ECS实例ID: 配置迁移的源ECS实例的实例ID。DTS支持经典网络、VPC网络的ECS实例。如果ECS实例跟目标实例在不同地域，那么ECS上面必须挂载公网EIP。
- 数据库类型：选择ECS上自建数据库的数据库类型，例如 Oracle/MySQL/SQLServer/PostgreSQL/Redis/MongoDB等。
- 数据库名称：如果自建数据库为PostgreSQL/MongoDB, 那么需要配置数据库名称。这个数据库名称为连接ECS自建数据库时使用的默认数据库。
- 数据库账号：连接自建数据库的账号。
- 数据库密码：上面数据账号对应的密码。

目标实例信息

- 实例类型：选择要迁入的实例对应的实例类型，包括 RDS/MongoDB/Redis/DRDS/PetaData/OceanBase等。
- 实例ID：配置迁移的目标实例的实例ID。
- 数据库名称：如果数据库类型为PostgreSQL/PPAS/MongoDB, 那么需要配置数据库名称。这个数据库名称为连接实例时使用的默认数据库。
- 数据库账号：连接实例的账号。
- 数据库密码：上面数据账号对应的密码。

下图以ECS上自建MongoDB到MongoDB实例的数据迁移为例，展示配置内容。

The screenshot displays the '2. 迁移类型及列表' (Migration Type and List) step of the DTS migration process. It is divided into two main sections: '源库信息' (Source Database Information) and '目标库信息' (Target Database Information). Both sections are highlighted with red boxes.

源库信息 (Source Database Information):

- 任务名称: dtsouzh34r
- 实例类型: ECS上的自建数据库
- ECS实例ID: [Redacted]
- 数据库类型: MongoDB
- 端口: 27017
- 数据库名称: admin
- 数据库账号: dba
- 数据库密码: [Redacted]

目标库信息 (Target Database Information):

- 实例类型: MongoDB实例
- MongoDB实例ID: [Redacted]
- 数据库名称: admin
- 数据库账号: root
- 数据库密码: [Redacted]

At the bottom right, there are three buttons: '取消' (Cancel), '上一步评估' (Previous Step Evaluation), and '授权白名单并进入下一步' (Authorize Whitelist and Proceed to Next Step), with the last one highlighted in blue.

迁移类型及迁移对象选择

迁移类型

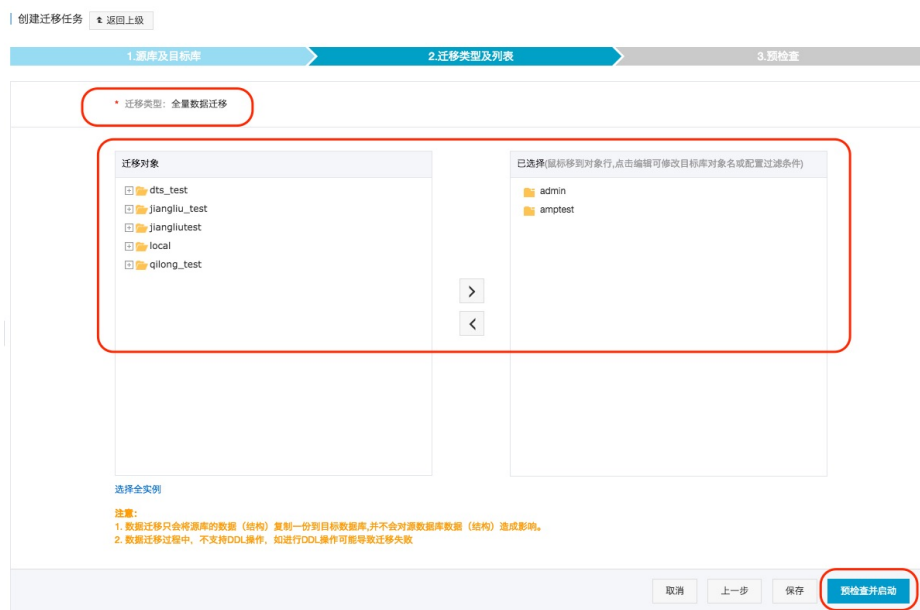
对于不同的数据库类型，DTS 支持的迁移类型不同。

如果只需要进行全量迁移，那么迁移类型选择：结构迁移 + 全量数据迁移。

对于支持增量迁移的数据库类型，如果需要进行不停机迁移，迁移类型选择：结构迁移 + 全量数据迁移 + 增量数据迁移。

迁移对象

这个步骤选择要迁移的对象。迁移对象的选择粒度细化为：库、表、列三个粒度。默认情况下，对象迁移到目标实例后，对象名跟源实例一致。如果您迁移的对象在源实例跟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，详细使用方式可以参考库表列映射。



预检查。

在迁移任务正式启动之前, 会先进行前置预检查, 只有预检查通过后, 才能成功启动迁移。

如果预检查失败, 那么可以点击具体检查项后的按钮, 查看具体的失败详情, 并根据失败原因修复后, 重新进行预检查。



启动迁移任务。

当预检查通过后, 可以启动迁移任务, 任务启动成功后, 可以在任务列表中查看迁移的具体状态及迁移进度。

如果选择了增量迁移, 那么进入增量迁移阶段后, 源库的更新写入都会被DTS同步到目标实例。迁移任务不会自动结束。如果用户只是为了迁移, 那么建议在增量迁移无延迟的状态时, 源库停写几分钟, 等待增量迁移再次进入无延迟状态后, 停止掉迁移任务, 直接将业务切换到目标实例上即可。

至此，完成ECS自建数据库到目标实例的数据迁移任务的配置。

说明：配置迁移任务时如果选择了增量迁移，那么增量迁移过程中，若用户不结束迁移，任务会一直处于迁移中。建议在业务验证通过且增量迁移无延迟的状态时，源库停写几分钟，等待增量迁移再次进入无延迟状态后，停止掉迁移任务，将业务切换到目标实例上即可。

使用数据传输DTS可以实现MySQL到DRDS的数据迁移。同时，通过增量数据迁移，可在MySQL不停服的情况下，将数据平滑迁移到目标DRDS实例。

本小节简单介绍使用DTS将MySQL数据迁移到DRDS的数据迁移任务配置流程。

数据源

DTS能够支持下面几种实例类型的MySQL到DRDS的迁移：

- 本地IDC自建MySQL->DRDS
- ECS自建MySQL->DRDS
- RDS->DRDS

迁移对象

由于DRDS只支持表，所以DTS支持的迁移对象包括：

- 单表(包括所有索引定义)
- 分区表

迁移权限要求

当使用DTS进行MySQL到DRDS实例的数据迁移时，不同迁移类型，对源跟目标数据库的迁移帐号权限要求如下。

- 本地IDC自建MySQL->DRDS实例

迁移类型	结构迁移	全量迁移	增量迁移
本地IDC自建MySQL	select	select	super select replication slave replication client
目的DRDS	读写权限	读写权限	读写权限

- ECS自建MySQL->DRDS实例

迁移类型	结构迁移	全量迁移	增量迁移
------	------	------	------

ECS上自建MySQL	select	select	super select replication slave replication client
目的DRDS	读写权限	读写权限	读写权限

- RDS实例->DRDS实例

迁移类型	结构迁移	全量迁移	增量迁移
源RDS实例	读写权限	读写权限	读写权限
目的DRDS	读写权限	读写权限	读写权限

迁移任务配置

下面以RDS到DRDS数据迁移为例，介绍DTS支持DRDS数据迁移的任务配置。自建MySQL到DRDS的迁移，除了源实例连接方式不同外，其他配置均相同。

迁移帐号创建

迁移任务配置时，需要提供源RDS实例及目的DRDS实例的迁移账号。迁移账号的相关权限详见上面的 迁移权限要求 一节。

如果RDS尚未创建迁移账号，那么可以参考 [RDS实例账号创建](#) 创建账号，并将待迁移库的读写权限授权给迁移账号。

迁移任务配置

当上面的所有前置条件都配置完成后，就可以开始正式的数据迁移了。下面详细介绍迁移任务配置流程。

1. 进入数据传输DTS控制台，点击右上角的**创建迁移任务**，开始迁移任务配置。
2. 源及目的实例连接信息配置

这个步骤主要配置 **迁移任务名称**，**源RDS连接信息**及**目标DRDS实例连接信息**。其中：

- 任务名称

DTS为每个任务自动生成一个任务名称，任务名称没有唯一性要求。您可以根据需要修改任务名称，建议为任务配置具有业务意义的名称，便于后续的任务识别。

源实例信息

- 实例类型：这个示例中选择 RDS实例；如果是本地IDC的自建MySQL,那么选择 **有公网IP的自建数据库**；如果是ECS上的自建MySQL,那么选择 **ECS上的自建数据库**。
- 实例地域：选择RDS实例所在地域；如果是本地IDC的自建MySQL，那么选择离自建MySQL物理距离最近的地域；如果是ECS上的自建MySQL,那么选择ECS所在地域。

- RDS实例ID: 配置源RDS实例的实例ID。DTS支持经典网络、VPC网络的RDS实例
- 数据库账号：连接RDS实例的访问账号
- 数据库密码：上面数据账号对应的密码

目标实例信息

- 实例类型：选择 DRDS实例
- 实例地区：选择DRDS实例所在地域
- DRDS实例ID：配置迁移的目标DRDS实例的实例ID。 DTS支持经典网络、VPC网络的DRDS实例
- 数据库名称：待迁入数据的DRDS数据库名称
- 数据库账号：连接DRDS 数据库的账号
- 数据库密码：上面数据库账号对应的密码

当配置完连接信息后，点击右下角 **授权白名单并进入下一步** 进行白名单授权。这个步骤DTS会将DTS服务器的IP地址添加到RDS及DRDS实例的白名单中，避免因为RDS或DRDS实例设置了白名单，DTS服务器连接不上实例导致迁移失败。

3. 选择迁移对象及迁移类型

- 迁移类型

DTS迁移类型支持结构迁移、全量数据迁移及增量迁移。

如果只需要进行全量迁移，那么迁移类型选择：结构迁移 + 全量数据迁移。

如果需要进行不停机迁移，那么迁移类型选择：结构迁移 + 全量数据迁移 + 增量数据迁移。

- 迁移对象

这个步骤选择要迁移的对象。迁移对象的选择粒度细化为：库、表、列三个粒度。默认情况下，对象迁移到目标DRDS实例后，对象名跟源RDS实例一致。如果您迁移的对象在源实例跟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，详细使用方式可以参考库表列映射。

4. 预检查

在迁移任务正式启动之前，会先进行前置预检查，只有预检查通过后，才能成功启动迁移。

如果预检查失败，那么可以点击具体检查项后的按钮，查看具体的失败详情，并根据失败原因修复后，重新进行预检查。



5. 启动迁移任务

当预检查通过后，可以启动迁移任务，任务启动成功后，可以在任务列表中查看迁移的具体状态及迁移进度。

如果选择了增量迁移，那么进入增量迁移阶段后，源库的更新写入都会被DTS同步到目标DRDS实例。迁移任务不会自动结束。如果用户只是为了迁移，那么建议在增量迁移无延迟的状态时，源实例停写几分钟，等待增量迁移再次进入无延迟状态后，停止掉迁移任务，直接将业务切换到目标DRDS实例上即可。

至此，完成MySQL到DRDS实例的数据迁移任务配置。

数据传输服务DTS支持通过专线接入阿里云的本地自建数据库的上云迁移。本小节以本地MySQL到RDS For MySQL实例的迁移为例，介绍如何通过DTS进行通过专线接入阿里云的本地自建数据库上云迁移任务的配置。

迁移前置条件

在迁移任务配置之前，需要在本地专线客户侧的接入设备上，添加指向DTS云端访问地址段的静态路由。

目前各个地域支持专线DB迁移的DTS IP段地址如下：

地域	DTS IP地址段
杭州	100.104.52.0/24
北京	100.104.183.0/24
青岛	100.104.72.0/24
上海	100.104.205.0/24
深圳	100.104.205.0/24
上海金融云	100.104.175.0/24
深圳金融云	100.104.72.0/24
美西	100.104.175.0/24
香港	100.104.233.0/24

迁移任务配置

下面详细介绍使用DTS将本地的MySQL迁移到RDS for MySQL的任务配置流程。

RDS实例数据库创建

在数据迁移过程中，如果待迁移的数据库在目标RDS实例中不存在，那么DTS会自动创建。但是对于如下两种情况，用户需要在配置迁移任务之前，手动创建数据库。

1. 数据库名称不符合：RDS定义规范(由小写字母、数字、下划线、中划线组成，字母开头，字母或数字结尾，最长64个字符)。
2. 待迁移数据库，在Oracle跟目标RDS实例中名称不同。

对于这两种情况，用户需要在配置迁移任务之前，先在RDS控制台完成数据库创建。具体参考RDS数据库创建流程RDS使用手册。

迁移帐号创建

迁移任务配置，需要提供本地MySQL数据库及目标RDS实例的迁移账号。迁移账号所需权限详见上文的 迁移权限要求。

如果源MySQL实例迁移账号尚未创建，那么您可以参考 Grant语法说明，创建满足权限要求的迁移账号。

如果目标RDS实例迁移账号创建，那么您可以参考 RDS账号创建流程，创建对待迁移数据库有读写权限的迁移账号。

其他准备工作

如果需要增量迁移，那么需要确认源库的binlog是否开启，格式是否满足要求。具体确认项如下。

确认源库的binlog是否开启。

```
mysql> show global variables like "log_bin";
+-----+-----+
| Variable_name | Value |
+-----+-----+
| log_bin      | ON   |
+-----+-----+
1 row in set (0.04 sec)
```

如果这里面查询出来的log_bin=OFF,那么说明源库没有开启binlog，这个时候如果需要增量迁移的话，需要修改本地MySQL的binlog配置，打开binlog日志。

- i. 开启二进制日志，修改log_bin=mysql_bin。

- ii. 设置binlog模式为row，修改binlog_format=row。
- iii. 配置server_id大于1，即server_id=某一个大于1的整数。
- iv. 如果版本为5.6,修改binlog_row_image=full。
- v. 修改完成后，重启MySQL进程。

确认源库的binlog格式为row模式。

```
mysql> show global variables like "binlog_format";
+-----+-----+
| Variable_name | Value |
+-----+-----+
| binlog_format | ROW   |
+-----+-----+
1 row in set (0.04 sec)
```

使用上面的命令确认本地MySQL的binlog格式是否为row,如果查询出来的结果不为ROW的话，那么通过如下的参数设置命令将模式修改为row。

```
mysql> set global binlog_format=ROW;
Query OK, 0 rows affected (0.04 sec)
```

当修改完成后，最好将已有的老连接KILL掉重连，否则可能出现其他的连接依然使用statement格式记录binlog日志。

当本地MySQL版本大等于5.6.2时，确认源库的binlog_row_image=full。

```
mysql> show global variables like "binlog_row_image";
+-----+-----+
| Variable_name      | Value |
+-----+-----+
| binlog_row_image   | FULL  |
+-----+-----+
1 row in set (0.04 sec)
```

如果查询出来的结果不为FULL的话，那么通过如下的参数设置命令将其修改为full。

```
mysql> set global binlog_row_image=full;
Query OK, 0 rows affected (0.03 sec)
```

迁移任务配置

当数据库、迁移账号都创建完成后，就可以开始配置迁移任务了。下面详细介绍下具体的配置步骤。

1. 进入数据传输DTS控制台，单击右上角的**创建迁移任务**，开始任务配置

本地MySQL及目标RDS实例连接信息配置

在这个步骤中，主要配置迁移任务名称，迁移源实例及目标实例连接信息。其中：

任务名称

默认情况下，DTS为每个任务自动生成一个任务名称。任务名称没有唯一性要求，您可以修改这个名称，为任务配置一个具有业务意义的名称，便于后续的任务识别。

源实例连接信息

- 实例类型：选择 **通过专线接入阿里云的本地DB**
- 实例地区：选择 **专线接入阿里云的接入点**，例如接入阿里云的北京，那么选择华北2即可。
- 对端专有网络：专线接入阿里云的VPC ID
- 数据库类型：选择 **MySQL**
- 主机名或IP地址：配置本地MySQL数据库访问地址，这个地址为本地局域网访问地址
- 端口：MySQL实例监听端口
- 数据库账号：MySQL实例访问账号
- 数据库密码：上面MySQL访问账号对应的密码

目标RDS实例连接信息。

- 实例类型：选择 **RDS实例**
- RDS实例ID: 配置迁移的目标RDS实例的实例ID。DTS支持经典网络、VPC网络的RDS实例
- 数据库账号：RDS实例的连接账号
- 数据库密码：上面数据库账号对应的数据库密码

源库信息

实例类型：通过专线接入的本地DB

实例地区：华北 2

对端专有网络：vpc-2ze3fwitarvbgpvs8lqt

数据库类型：MySQL

主机名或IP地址：192.168.100.157

端口：3306

数据库账号：jiangliu@test

数据库密码：.....

测试连接

迁移对象及迁移类型配置

迁移类型

DTS支持 结构迁移、全量数据迁移、增量数据迁移。

如果需要进行不停机迁移，那么需要选择：结构迁移 + 全量数据迁移 + 增量数据迁移。

如果只进行全量迁移，那么需要选择：结构迁移 + 全量数据迁移。

迁移对象

迁移对象，需要选择您要迁移的对象。迁移对象选择的粒度可以为：库、表、列三个粒度。

默认情况下，对象迁移到RDS实例后，对象名跟本地MySQL实例一致。如果您迁移的对象在源实例跟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，详细使用方式可以参考库表列映射。

当配置完迁移对象及迁移类型后，即进入任务启动前的预检查步骤

预检查

在迁移任务正式启动之前，会先进行前置预检查，只有预检查通过后，才能成功启动迁移。

如果预检查失败，那么可以点击具体检查项后的按钮，查看具体的失败详情，并根据失败原因修复后，重新进行预检查。



启动迁移任务

当预检查通过后，我们可以启动迁移任务，任务启动后，可以到任务列表中查看任务具体的迁移状态及进度。

增量数据迁移是个动态同步的过程，所以建议在增量迁移达到无延迟状态时，在目标数据库上进行业务验证，如果验证成功，那么可以停掉迁移任务，然后将业务切换到目标数据库。

至此，完成将本地MySQL数据库到RDS For MySQL的数据迁移任务配置。

数据传输DTS数据迁移支持对象名映射，即支持源跟目标实例的迁移对象(包括数据库、表或列)的名称不同。

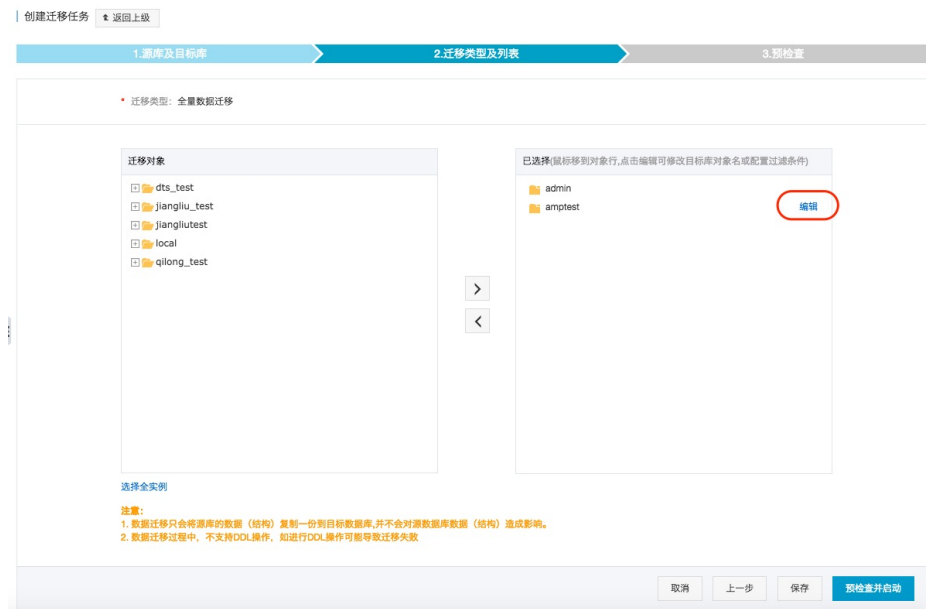
本文介绍如何在数据迁移任务配置过程中，使用对象名映射功能。

数据库名映射

如果迁移的数据库在源跟目标实例中命名不同，那么可以使用DTS提供的对象名映射功能进行数据库名映射。

数据库名映射功能在配置迁移任务的第二个步骤 **选择迁移类型及迁移对象** 中配置。下面介绍详细的配置步骤。

在**已选择**框中，将鼠标挪到需要进行数据库名映射 的数据库对象行上，右侧会浮出 **编辑** 入口。

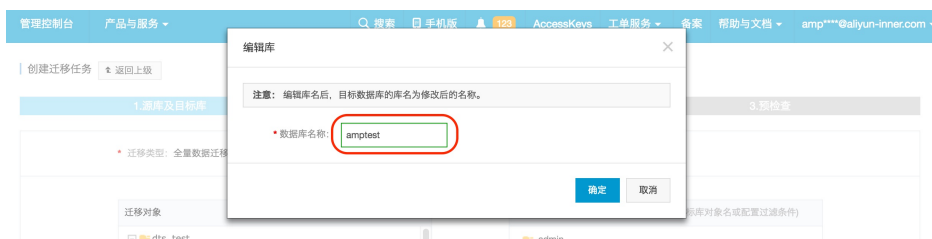


修改库名。

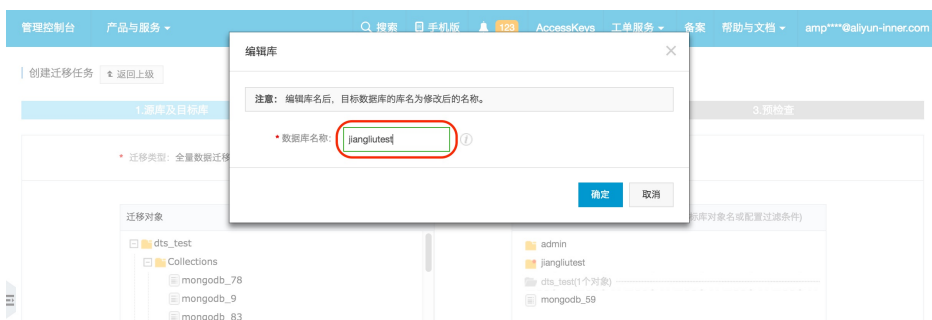
如果这个数据库迁移到目标实例中库名要修改为jiangliutest，那么点击 **编辑** 按钮进入数据库名修改页面。

在数据库名编辑页，数据库名称可以直接修改，修改后的名称即为数据库在目标实例中存储的名称。

假设修改前，数据库名为amptest,如下图:



由于数据库迁移到目标实例中库名需要修改成jiangliutest，所以直接修改界面中的 amptest 为 jiangliutest，如下图：



数据库名修改完之后，这个数据库迁移到目标实例使用的库名即为修改后的库名。

表名映射

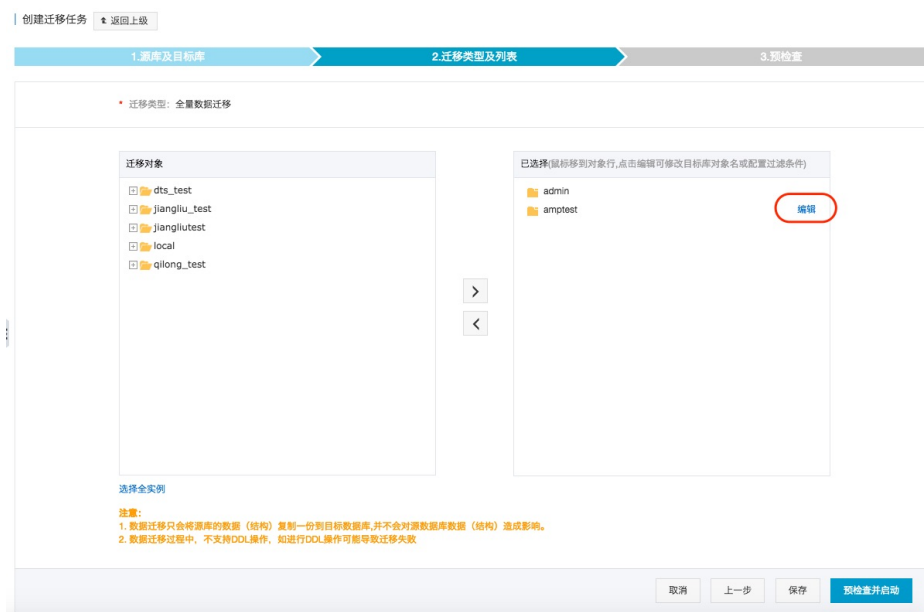
如果迁移的表在源跟目标实例中命名不同，那么可以使用DTS提供的对象名映射功能进行数据库名映射。

如果要使用表名映射功能，那么选择迁移对象时，不能直接选择整个库，必须选择具体的表

除了表，其他的结构对象（例如视图、存储过程、存储函数、同义词等），也可以进行对象名映射，使用相同的方式进行对象名映射。

表名映射功能在配置迁移任务的第二个步骤 **选择迁移类型及迁移对象** 中配置。下面介绍详细的配置步骤。

在**已选择**框中，将鼠标挪到需要进行表名映射的对象行上，右侧会浮出 **编辑** 入口。

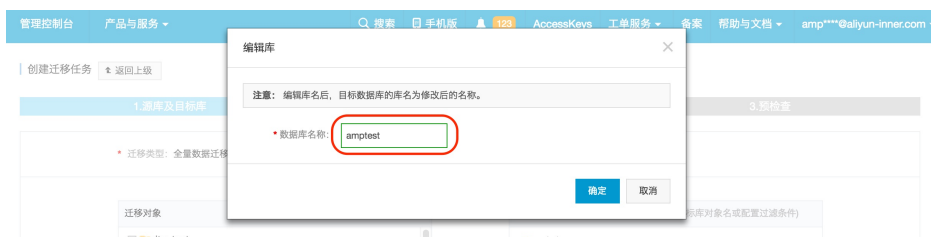


修改表名。

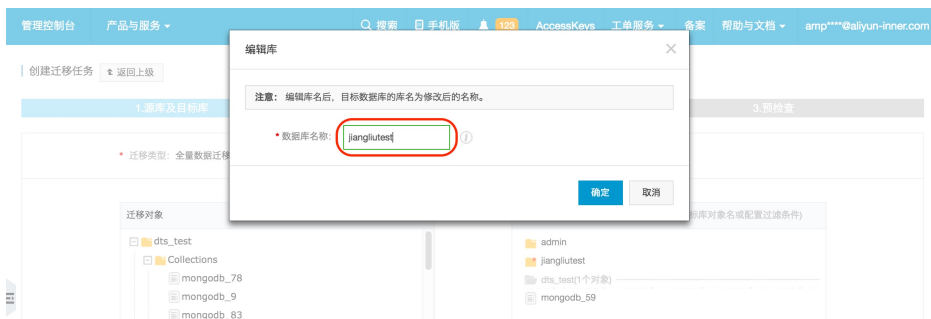
例如需将待迁移表ampptest迁移到目标实例中的jiangliu_test，那么ampptest后面的 **编辑** 按钮进入表名修改页面。

在表编辑页，表名可以直接修改，修改后的名称即为数据库在目标实例中存储的名称。

假设修改前，数据库名为ampptest,如下图:



表迁移到目标实例中库名需要修改成jiangliu_test，所以直接修改界面中的 ampptest 为 jiangliu_test，如下图：



至此，完成表名映射的相关配置。

列名映射

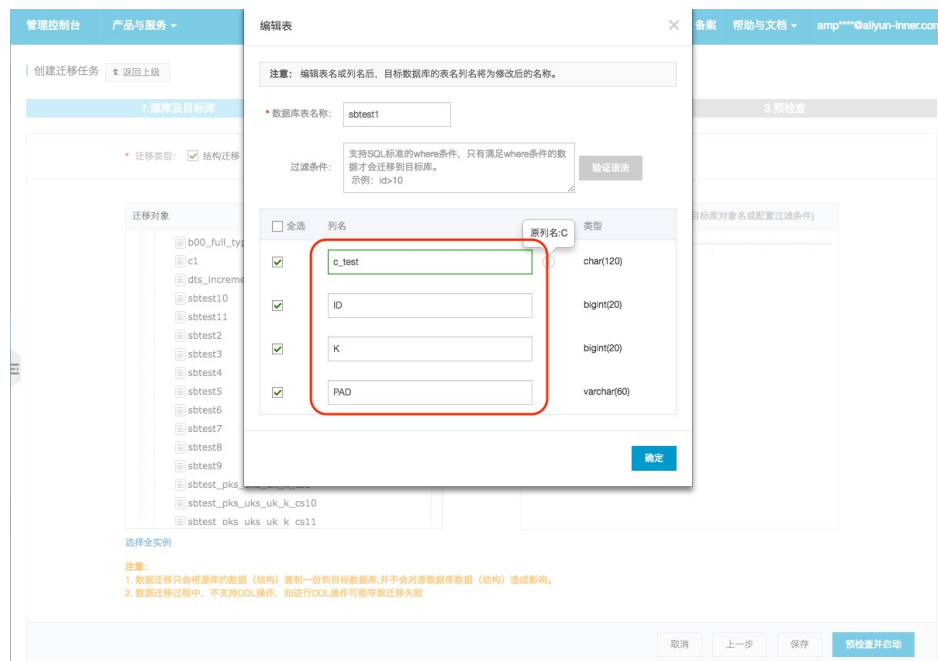
如果要迁移的源跟目标实例的某个表中对应的列名不同，那么可以使用DTS的对象名映射功能。

列名映射是在创建迁移任务的第二个步骤“选择迁移类型及迁移对象”中。如果要修改迁移列的名称，那么选择迁移对象时，必须选择这个列对应的表，不能直接选择数据库。下面简单介绍下修改迁移列对应列名的方法。

假设要修改列对应的表为sbtest1，那么将鼠标挪到界面已选择框中的sbtest1对应的行上，此时右边会出现 **编辑** 按钮。



点击编辑按钮，进入表编辑界面。



在这个步骤中，对于需要修改列名的列，直接编辑列名即可。修改后，目标库中存储的列名为修改后的列名。至此，完成列名映射的相关配置。

DTS支持通过配置SQL条件来过滤表中需要迁移的数据。这个SQL条件只能作用在配置的表上，DTS会通过这个SQL条件来过滤源库中这个表的数据，只有满足这个条件的数据才会被迁移到目标库。这个功能可以应用于定期增量迁移、拆表等多种应用场景。本小节简单介绍如何在创建迁移任务时，配置SQL条件来过滤需要的迁移数据。

功能限制

SQL过滤条件只能作用于全量数据迁移阶段，如果任务的迁移类型选择了 增量数据迁移，那么SQL过滤条件不会应用于增量数据。

配置SQL过滤条件

SQL过滤条件配置，在迁移任务配置的 **选择迁移类型及列表** 这个步骤。

如果要配置表的迁移SQL过滤条件，那么选择迁移对象时，必须选择表，不能直接选择整个数据库。下面简单介绍下配置SQL过滤条件的方法。

配置入口

对于需要配置过滤条件的表，在 **选择迁移类型及列表** 步骤的已选择 对象框中，将鼠标挪到需要配置过滤条件的表上，会浮现 **编辑** 按钮。

点击 **编辑** 按钮，即进入过滤条件配置页面。

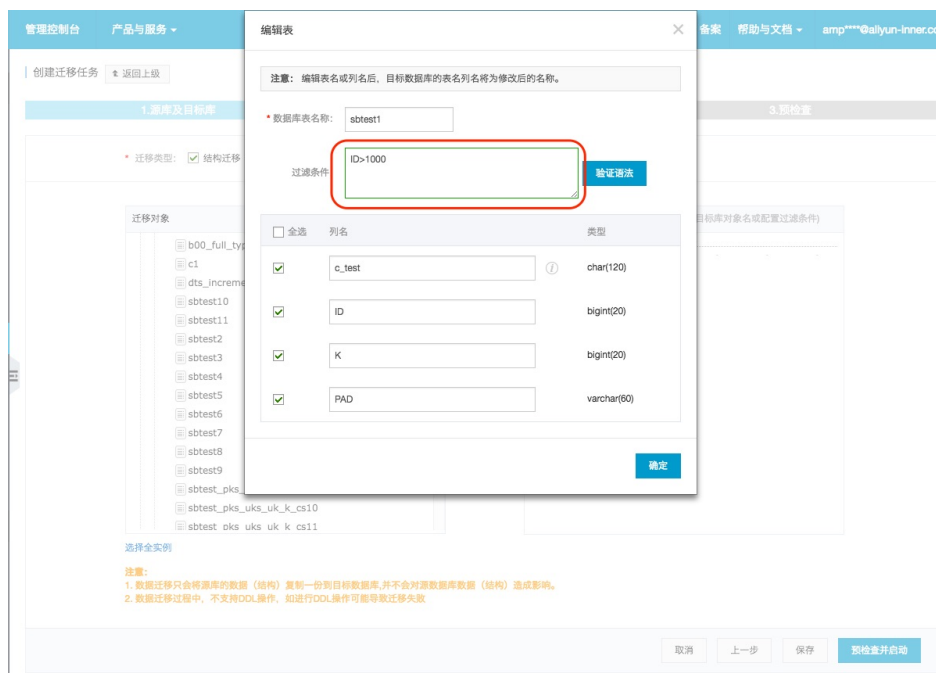


假设要修改列对应的表为sbtest1，那么将鼠标挪到界面已选择框中的sbtest1对应的行上，此时右边会出现 **编辑** 按钮。

修改SQL过滤条件

DTS的过滤条件同数据库标准的SQL WHERE条件，可以支持运算及简单的函数。

根据需求在过滤条件的编辑框内，填写SQL过滤条件，例如我们只需要将表中id>1000的记录迁移到目标实例，那么具体配置如下：



至此，完成数据过滤条件的配置流程。

DTS数据迁移功能在多个阶段都提供了失败在线修复功能。主要包括：

- 结构迁移 由于DTS支持异构数据源之间的迁移，例如对于Oracle->MySQL，对于一些不支持的数据类型，在结构迁移时，直接导入到目标实例，会导致迁移失败。
- 全量迁移 全量迁移过程中，可能因为目标RDS实例空间满、RDS实例白名单被删除等问题导致迁移失败，这些失败任务都可以通过用户手动修改后，重启任务来自助恢复任务。

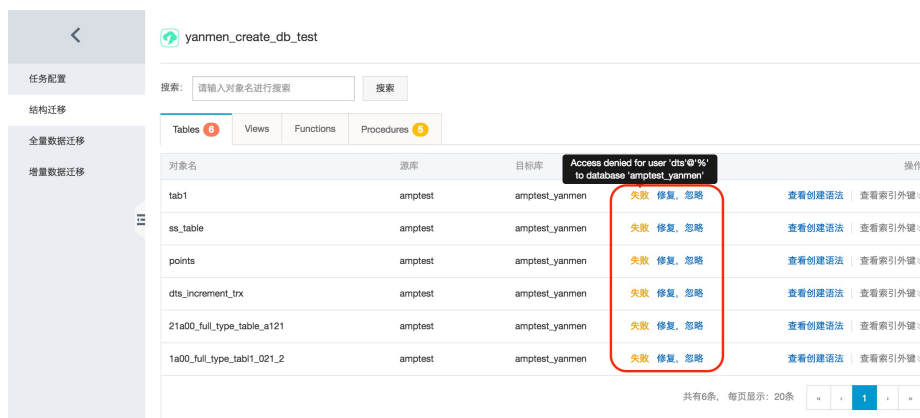
为了解决这些问题，DTS提供了在线修复功能，当用户在迁移过程遇到这些问题时，可以修复续传。下面简单介绍结构迁移和全量迁移过程中的失败修复流程。

结构迁移—失败修复

如果结构迁移失败了，任务状态会显示迁移失败，同时状态后，提供立即修复入口。



此时，点击 **立即修复**，进入结构对象修复界面。



对于每个对象，如果失败了，点击修复，进入对象修复页面。



根据失败原因提示，修改结构定义语法，修改完成后点击修复，将修改后的语法重新导入到目标实例。

如果点击修复失败原因变成修复失败，且显示修复失败原因，那么说明没有修复彻底。需要根据修复失败原因继续修复，直至修复成功。

如果修复成功，那么会回到结构迁移详情页，且刚才修复的对象的状态变成完成。

当所有的对象全部修复成功后，这个任务就会继续往下执行，例如进入全量迁移阶段。

全量迁移—失败修复

对于下列几种失败原因，DTS提供了修复重试的功能：

1. 源或目标数据库连接失败，确认连接没有问题后，重试任务。
2. 源或目标数据库连接超时，确认连接没有问题后，重试任务。
3. 目标RDS实例空间满，实例被锁定，需要升级RDS实例规格或清理实例日志空间后，重试任务。
4. 源数据库中部分Myisam被损坏，需要repair后，重试任务。

其他情况下，如果全量迁移失败，DTS只提供忽略操作，只能忽略失败对象后，进行其他正常对象的迁移。

如果全量迁移失败了，那么迁移任务的状态为 迁移失败，状态后面提供立即修复入口。



当出现失败时，点击“立即修复”进入修复页面。

如果遇到上面可以重试的错误，那么根据错误提示修复问题后，点击全量迁移详情页面的重试按钮，触发任务断点续传。

对于其他失败原因，DTS只提供“忽略”操作，即忽略掉这个对象的全量数据迁移。一旦点击忽略，那么这个对象的数据不会迁移，但是其他正常对象的数据会迁移到目标实例。

实时同步

本小节介绍如何使用数据传输服务快速创建两个RDS(MySQL)实例之间的实时同步作业，实现RDS增量数据的实时同步。

支持功能

1. 支持阿里云账号下两个RDS MySQL实例间的实时同步。
2. 支持不同阿里云账号下的RDS MySQL实例间的实时同步。

同步限制

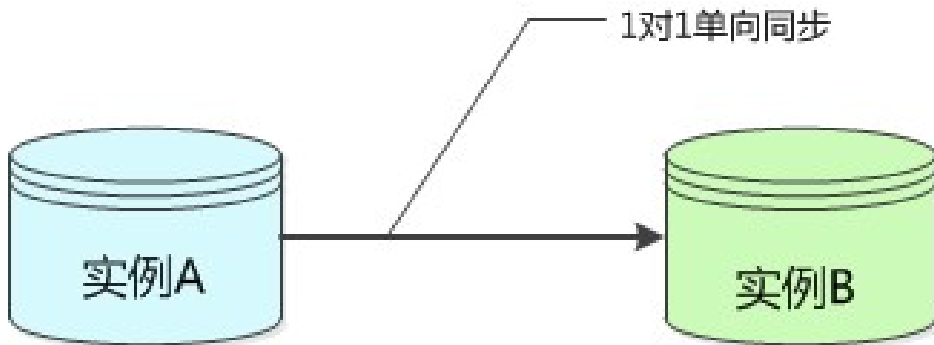
数据源

1. 目前实时同步只能支持RDS MySQL实例,暂不支持其他数据源类型。
2. 目标实例不支持访问模式为标准模式且只有外网连接地址的RDS实例。
3. 不支持香港可用区A的RDS实例的实时同步。
4. 对于rename table tbl_name to new_tbl_name、create table tbl_name like new_tbl_name、create...select...from new_tbl_name、alter table tbl_name rename to new_tbl_name 如果 new_tbl_name不在同步对象中，暂不支持。

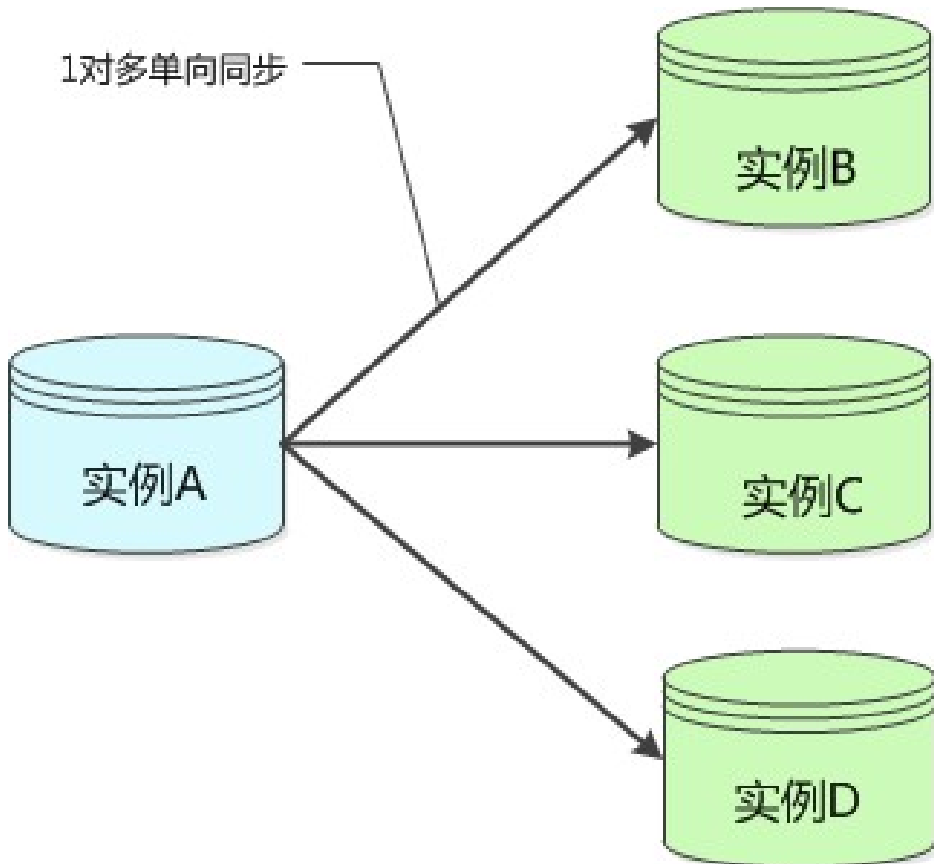
同步架构

目前数据传输服务提供的实时同步功能支持的同步架构有限，其仅能支持如下两种架构：

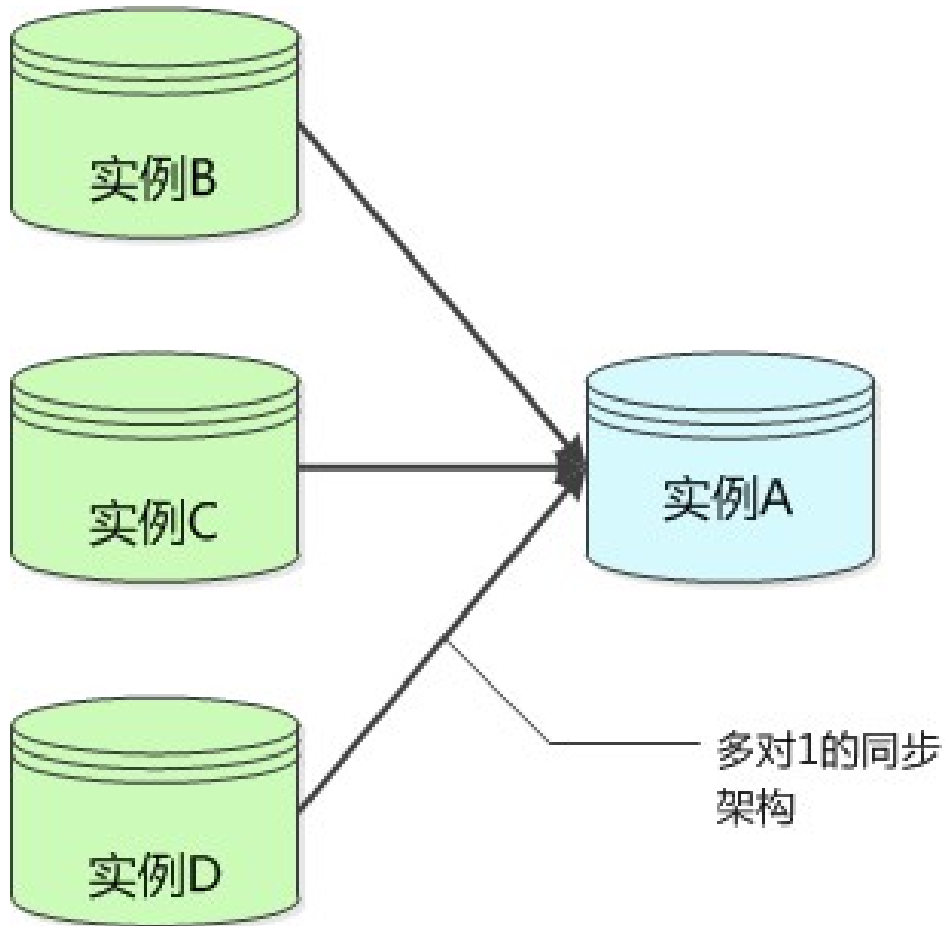
A->B 即两个实例之间的单向同步。且要求B中同步的对象必须为只读，否则可能导致同步链路异常。



A->B/C/D 即1对多的分发式同步架构,这个架构对目标RDS实例个数没有限制,但是要求目标实例中的同步对象必须为只读,否则可能导致同步链路异常。

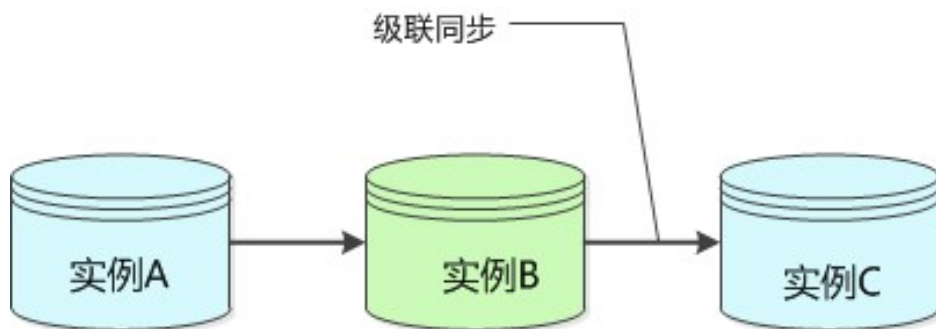


B/C/D->A 即多对1的数据汇总架构。对于这种多对1的同步架构,为了保证同步数据一致性,要求每条同步链路同步的对象不相同。

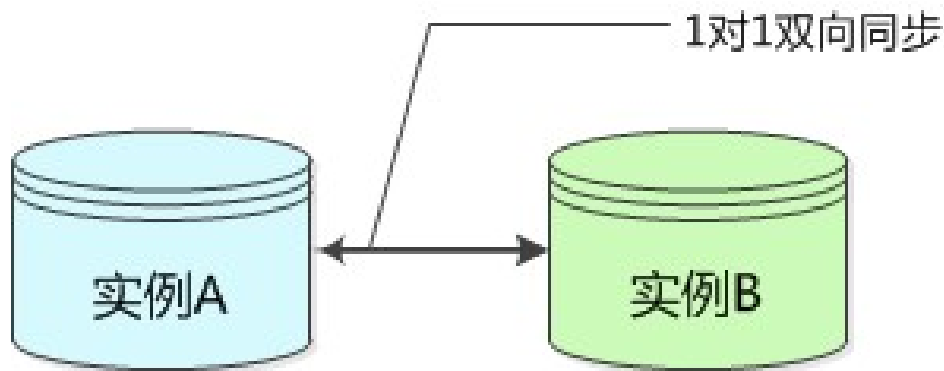


对于下面几种同步架构，暂时不支持：

A->B->C 即级联架构。



A->B->A 即实例A和实例B之间的双向同步架构。



如果用户配置同步链路过程中，配置了这些不支持的同步架构，那么预检查中的**复杂拓扑**检查项会检查失败。

功能限制

不兼容触发器

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

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

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

rename table 限制

rename table操作可能导致同步数据不一致。例如同步对象只包含表A，不包含表B，如果同步过程中源实例执行了**rename A to B**的操作，那么改名后的B表的操作不会被同步到目标库。为了解决这个问题，可以选择同步表A、B对应的整个数据库。

准备事项

在配置同步作业前，要确保同步作业的源及目标RDS实例都已经存在。如果不存在，那么请先购买RDS实例购买RDS实例。

配置步骤

下面我们详细介绍下创建任意两个RDS实例之间的同步链路的具体步骤。

购买同步链路。

进入数据传输服务控制台，进入数据同步页面，点击控制台右上角“创建同步作业”开始作业配置。

在链路配置之前需要购买一个同步链路。同步链路目前支持包年包月及按量付费两种付费模式，可以根据需要选择不同的付费模式。

在购买页面需要配置的参数包括：

源地域

源地域为同步链路源实例所在地域。

目标地域

目标地域为同步链路目标实例所在地域。

实例规格

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

数量

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

数据传输服务（包年包月）

包年包月 | 按量付费

① 如果需要购买数据迁移链路，请到[按量付费](#)页面购买

功能：
数据同步 | 数据迁移

源地域：
华东1(杭州) | 华北2(北京) | 华北1(青岛) | 华南1(深圳) | 华东2(上海) | 美国(硅谷) | 美东
亚太(新加坡) | 香港

目标地域：
华东1(杭州) | 华北2(北京) | 华北1(青岛) | 华南1(深圳) | 华东2(上海) | 美国(硅谷) | 美东
亚太(新加坡) | 香港

实例规格：
micro

订购时长：
1个月 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 1年 | 2年 | 3年

数量：
1

购买清单 0例

当前配置

功能：数据同步
源地域：华东1(杭州)
目标地域：华东1(杭州)
实例规格：micro
订购时长：1个月
数量：1
配置费用：
¥0.00
原价¥770.00 现价¥770.00
数据传输服务(包年包月DTS)_0折

立即购买 | 加入购物车

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

同步链路连接信息配置。

在这一步主要配置：

同步作业名称

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

同步链路的RDS实例ID

源跟目标RDS实例必须为两个不同的实例，选择RDS实例ID时，下拉菜单中只列出对应阿里云账号下的RDS For MySQL实例。

当这些内容配置完成后，可以点击**授权白名单并进入下一步**。

授权RDS实例白名单。

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

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

当白名单授权后，点击下一步，进入同步账号创建。

创建目标库上的同步账号。

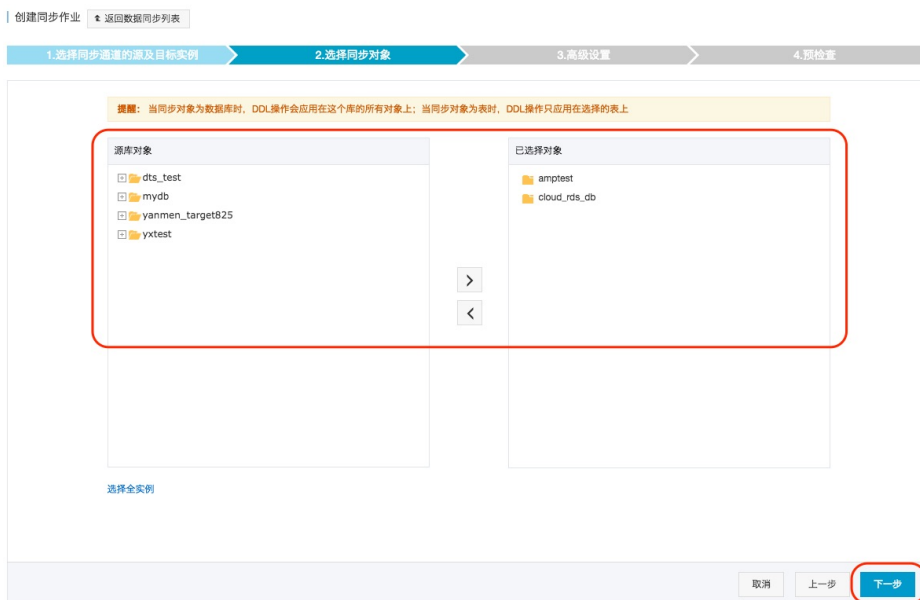
这个步骤主要是在目标RDS实例上创建一个同步账号，账号名字为:dtssyncwriter，在同步过程中，不能删除这个账号，否则会导致同步链路中断。

选择同步对象。

当创建完目标RDS实例的同步账号后，即进入同步对象的选择步骤。实时同步的同步对象的选择粒度可以支持到表级别，即用户可以选择同步某些库或是同步某几张表。

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

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



当配置完同步对象后，进入同步初始化配置。

同步初始化配置。

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

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



预检查。

当上面所有选项配置完成后，即进入启动之前的预检查。

当同步作业配置完成后，数据传输服务会进行限制预检查，当预检查通过后，可以点击 **启动** 按钮启动同步作业。

当同步作业启动之后，即进入同步作业列表。此时刚启动的作业处于**同步初始化**状态。初始化的时间

长度依赖于源实例中同步对象的数据量大小。当初始化完成后同步链路即进入**同步中**的状态，此时源跟目标实例的同步链路才真正建立完成。

本小节介绍如何使用数据传输服务快速创建RDS实例到MaxCompute实例间的实时同步作业，实现在线(RDS)到离线系统(MaxCompute)的数据实时同步，进一步为数据实时分析奠定基础。

支持功能

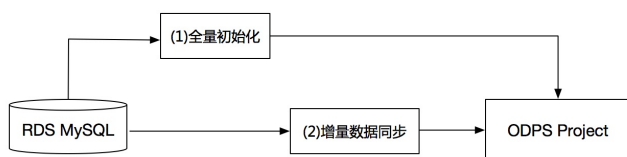
数据源

- 支持同一个阿里云账号下RDS MySQL实例到MaxCompute实例的数据实时同步。
- 支持不同阿里云账号下的RDS MySQL实例到MaxCompute实例的数据实时同步。
- 支持的RDS实例包括，经典网络和VPC网络两种网络模式。

同步对象

- 只支持表的同步，不支持其他非表对象的同步。

同步原理



如上图所示，整个同步过程分为两步：

(1) 全量初始化，这个步骤将RDS MySQL中已经存在的全量数据初始化到MaxCompute中。对于同步的每个表，全量初始化的数据都会独立存储在MaxCompute中的全量基线表中，这个表的默认格式为：源表名_base。例如表 t1，那么全量基线表在MaxCompute中存储的表名为：t1_dts_base。这个存储表名前缀可以根据需要变更，您可以在配置任务时，修改表在MaxCompute存储的名称。

(2) 增量数据同步，这个步骤将RDS MySQL产生的增量数据数据实时同步到MaxCompute中。并存储在增量日志表中，每个同步表对应一个增量日志表。增量日志表在MaxCompute中存储的表名的默认格式为：源表名_log。这个存储表名前缀可以根据需要变更，您可以在配置任务时，修改表在MaxCompute存储的名称。

增量日志表除了存储更新数据，它还会存储一些元信息，增量日志表的表结构定义如下：

record_id	operation_flag	utc_timestamp	before_flag	after_flag	col1	colN
1	I	1476258462	N	Y	1	JustInsert
2	U	1476258463	Y	N	1	JustInsert
2	U	147625	N	Y	1	JustUpd

		8463					ate
3	D	147625 8464	Y	N	1	JustUpdate

其中：

record_id: 这条增量日志的唯一标识，唯一递增。如果变更类型为update，那么增量更新会被拆分成2条，一条Insert，一条Delete。那么这两条记录的record_id相同。

operation_flag: 标示这条增量日志的操作类型。取值包括:

I: insert 操作

D: delete 操作

U: update 操作

dts_utc_timestamp: 这条增量日志的操作时间戳，为这个更新操作记录binlog的时间戳。这个时间戳为UTC时间。

before_flag: 表示这条增量日志后面带的各个column值是否更新前的值。取值包括：Y 和 N。当后面的column为更新前的值时，before_flag = Y, 当后面的column值为更新后的值时，before_flag=N。

after_flag: 表示这条增量日志后面带的各个column值是否更新后的值。取值包括：Y 和 N。当后面的column为更新前的值时，after_flag=N, 当后面的column值为更新后的值时，after_flag=Y。

对于不同的操作类型，增量日志中的before_flag和after_flag定义如下：

1) 操作类型为：insert

record_id	operation_flag	utc_timestamp	before_flag	after_flag	col1	colN
1	I	147625 8462	N	Y	1	JustInsert

当操作类型为insert时，后面的所有column值为新插入的记录值，即为更新后的值。所以before_flag = N, after_flag=Y。

2) 操作类型为：update

record_id	operation_flag	utc_timestamp	before_flag	after_flag	col1	colN
2	U	147625 8463	Y	N	1	JustInsert
2	U	147625 8463	N	Y	1	JustUpdate

当操作类型为update时，会将update操作拆为2条增量日志。这两条增量日志的record_id, operation_flag 及 dts_utc_timestamp相同。

第一条日志记录了更新前的值，所以before_flag=Y, after_flag=N。

第二条日志记录了更新后的值，所以before_flag=N, after_flag=Y。

3) 操作类型为：delete

record_id	operation	dts_utc_	before_f	after fla	col1	colN
-----------	-----------	----------	----------	-----------	------	------	------

d	on_flag	timestamp	lag	g			
3	D	1476258464	Y	N	1	JustUpdate

当操作类型为delete时，后面的所有column值为被删除的记录值，即为更新前的值。所以before_flag = Y, after_flag=N。

(3) RDS->MaxCompute数据同步，对于每个同步表，都会在MaxCompute中生成一个全量基线表 + 一个增量日志表，所以如果需要在MaxCompute中获取某个时刻某张表的全量数据，就需要merge这张表的全量基线表和增量日志表。具体实现方法后面会详细讲解。

下面详细介绍RDS到MaxCompute数据实时同步作业的配置流程。

同步作业配置流程

下面详细介绍RDS到MaxCompute数据实时同步作业的配置流程。

1.购买同步链路

进入数据传输服务控制台，进入数据同步页面，点击控制台右上角“创建同步作业”开始作业配置。

在链路配置之前需要购买一个同步链路。同步链路目前支持包年包月及按量付费两种付费模式，可以根据需要选择不同的付费模式。

在购买页面需要配置的参数包括：

- 源实例
同步作业的源实例类型，目前只支持RDS For MySQL。
- 源地域
源地域为同步实例的源RDS实例所在地域。
- 目标实例
目标实例为同步作业的目标实例类型，目前支持 RDS For MySQL, MaxCompute(原ODPS),DataHub。配置RDS->MaxCompute同步链路时，目标实例选择：**MaxCompute** 即可。
- 目标地域
由于MaxCompute目前只在上海地区售卖，所以目标地域选择上海。
- 实例规格
实例规格影响了链路的同步性能，可以根据业务性能选择合适的规则。
- 网络类型
RDS->MaxCompute支持通过公网、私网 同步数据。如果源RDS没有公网连接地址，那么网络类型只能选择 **私网**。
- 数量
数量为一次性购买的同步链路的数量，如果购买的是按量付费实例，一次最多购买99条链路。

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

2.同步链路连接信息配置。

在这一步主要配置：

同步作业名称

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

实例ID配置

在这个步骤中需要配置源RDS实例的实例ID，及目标MaxCompute实例的project。配置的MaxCompute project 必须属于登录DTS的阿里云账号的资源。

当这些内容配置完成后，可以点击**授权白名单并进入下一步**。

3.授权RDS实例白名单

这个步骤，主要是将给DTS服务账号授权MaxCompute写权限，让DTS能够将数据同步复制到MaxCompute中。

授权权限包括对project的：

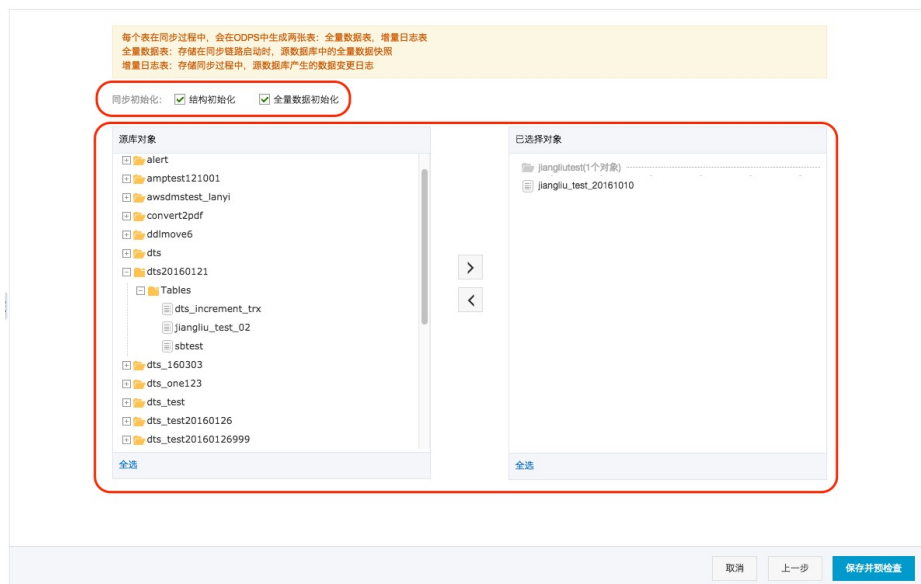
CreateTable
CreateInstance
CreateResource
CreateJob
List

为了保证同步作业的稳定性，在同步过程中，请勿将写权限回收。当白名单授权后，点击下一步，进入同步账号创建。

当授权完成后，即进入同步对象选择。

4.选择同步对象

当MaxCompute账号授权完成后，即进入同步表及同步初始化的相关配置。



在这个步骤中，需要配置同步初始化和同步表。其中：

(1)同步初始化

同步初始化选项包括：结构初始化和全量数据初始化。

结构初始化是指对于待同步的表，在MaxCompute中创建对应的表，完成表结构定义。

全量数据初始化是指对于待同步的表，将历史数据初始化到MaxCompute中。

配置任务时，建议同时选择结构初始化 + 全量数据初始化。

(2) 同步表选择

同步表只能选择某些表，不能直接选择整个库。对于同步的表，可以修改表在MaxCompute对应的全量基线表及增量日志表的表名前缀。如需修改，可以点击右边已选择对象后面的编辑按钮，进入修改界面。



当配置完同步对象后，进入同步初始化配置。

5. 预检查

当上面所有选项配置完成后，即进入启动之前的预检查。具体检查项内容详见本文最后的 **预检查内容** 一节。当同步作业配置完成后，数据传输服务会进行限制预检查，当预检查通过后，可以点击 **确定** 按钮，启动同步作业。

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

当同步任务进入 **同步中** 时，可以在MaxCompute中可以查询出jiangliu_test对应的全量基线表和增量日志表：

```
ALIYUN$dts_public_cloud@aliyun-inner.com jiangliu_test_20161010_base
ALIYUN$dts_public_cloud@aliyun-inner.com jiangliu_test_20161010_log
```

至此，完成RDS->MaxCompute数据实时同步作业的配置。

全量数据合并方案

本小节介绍，如何根据同步到MaxCompute中的全量基线表和增量日志数据得到某个时刻表的全量数据。DTS提供通过MaxCompute SQL实现全量数据合并的能力。

通过MaxCompute SQL merge 全量基线数据 和 增量日志表 得到时刻t的全量数据。MaxCompute SQL的写法如下：

```
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, col1,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< timestmap
union all
select 0 as record_id, 'I' as operation_flag, 'Y' as after_flag, base.col1, base.col2,base.colN
from table_base base) t) gt
where record_num=1
and after_flag='Y'

```

上面代码中的几个变量意义如下:

- 1) result_storage_table 表示全量merge结果集的存储表的表名
- 2) col1, col2,colN 表示同步表中列的列名
- 3) primary_key_column 表示同步表中的主键列的列名
- 4) table_log 表示增量日志表
- 5) table_base 表示全量基线表
- 6) timestmap 表示需要merge哪个时刻的全量数据

例如对于上面配置任务中的jiangliu_test表, jiangliu_test_20161010_base 为jiangliu_test对应的全量基线表, jiangliu_test_20161010_log 为jiangliu_test对应的增量日志表。

jiangliu_test的表结构定义为：

列名	类型	长度	备注	可空	主键
1 id	int	11		<input type="checkbox"/>	<input checked="" type="checkbox"/>
2 name	varchar	20		<input checked="" type="checkbox"/>	<input type="checkbox"/>

那么查询时间戳1476263486 时刻, jiangliu_test表的全量数据的MaxCompute SQL如下：

```

insert overwrite table jiangliu_test_1476263486
select id,
name
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, name
from(
select incr.record_id, incr.operation_flag, incr.after_flag, incr.id, incr.name
from jiangliu_test_20161010_log incr
where utc_timestamp< 1476263486
union all
select 0 as record_id, 'I' as operation_flag, 'Y' as after_flag, base.id, base.name
from jiangliu_test_20161010_base base) t) gt
where gt.row_number= 1
and gt.after_flag= 'Y' ;

```

您也可以通过大数据开发套件, 在后续的计算分析操作之前, 添加全量数据Merge节点, 当全量数据merge完成后, 可自动调度起后续的计算分析节点。同时可以配置调度周期, 进行周期性的数据离线分析。

至此完成RDS->MaxCompute数据同步任务配置及全量数据合并。

本小节介绍如何使用数据传输服务DTS快速配置RDS实例到DataHub实例间的数据实时同步作业，让数据可以进入流计算等大数据产品进行数据实时分析。

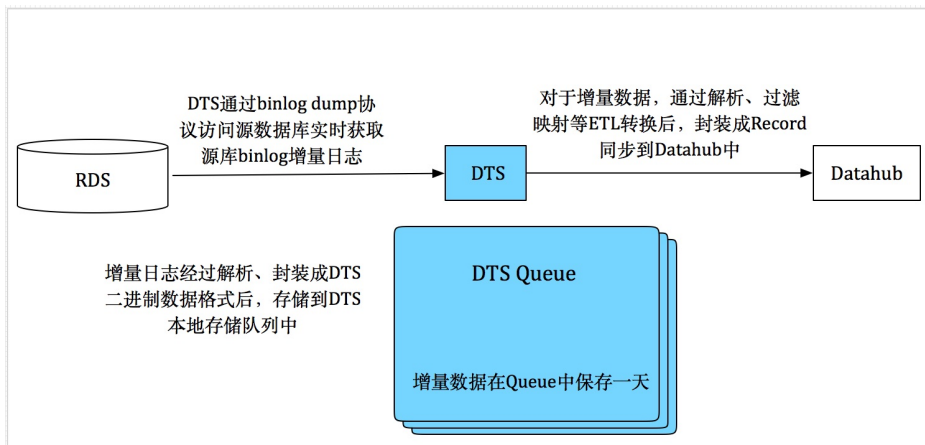
支持功能

- 支持同一个阿里云账号下RDS MySQL实例到DataHub实例的数据实时同步。
- 支持不同阿里云账号下的RDS MySQL实例到DataHub实例的数据实时同步。
- 支持的RDS实例包括，经典网络和VPC网络两种网络模式。

同步对象

- 只支持表的同步，不支持其他非表对象的同步。

同步原理



如上图所示，RDS->DataHub数据实时同步，是将RDS MySQL产生的增量数据数据实时同步到DataHub中的topic。增量日志DataHub中存储的表名默认同RDS表同名。topic名称可以根据业务需要修改，您可以在配置任务时，修改表在DataHub中对应的topic名称。

topic除了存储更新数据，它还会存储一些元信息，topic的结构定义如下：

dts_reco rd_id	dts_i nsta nce_ id	dts_ db_ nam e	dts_ tabl e_ na me	dts_ oper atio n_ fla g	dts_ utc_ time sta mp	dts_ befo re_ fla g	dts_ after_ fla g	dts_ col1	dts_ colN
1	234	db1	sbte st1	I	147 625 846 2	N	Y	1	JustI nser t
2	234	db1	sbte st1	U	147 625 846 3	Y	N	1	JustI nser t
2	234	db1	sbte st1	U	147 625	N	Y	1	Just Upd

					8463					ate
3	234	db1	sbtest1	D	1476258464	Y	N	1	Just Update

其中：

dts_record_id: 这条增量日志的唯一标识，唯一递增。如果变更类型为update，那么增量更新会被拆分成2条，一条Insert，一条Delete。那么这两条记录的record_id相同。

dts_instance_id: 这条增量日志所对应的数据库的server id。

dts_db_name: 这条增量更新日志更新的表所在的数据库库名。

dts_table_name: 这条增量更新日志更新的表。

dts_operation_flag: 标示这条增量日志的操作类型。取值包括:

I: insert 操作

D: delete 操作

U: update 操作

dts_utc_timestamp: 这条增量日志的操作时间戳，为这个更新操作记录binlog的时间戳。这个时间戳为UTC时间。

dts_before_flag: 表示这条增量日志后面带的各个column值是否更新前的值。取值包括：Y 和 N。当后面的column为更新前的值时，dts_before_flag = Y, 当后面的column值为更新后的值时，dts_before_flag=N.

dts_after_flag: 表示这条增量日志后面带的各个column值是否更新后的值。取值包括：Y 和 N。当后面的column为更新前的值时，dts_after_flag=N，当后面的column值为更新后的值时，dts_after_flag=Y.

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：

1) 操作类型为：insert

dts_record_id	dts_instance_id	dts_db_name	dts_table_name	dts_operation_flag	dts_utc_timestamp	dts_before_flag	dts_after_flag	dts_col1	dts_colN
1	234	db1	sbtest1	I	1476258462	N	Y	1	JustInsert

当操作类型为insert时，后面的所有column值为新插入的记录值，即为更新后的值。所以before_flag = N, after_flag=Y

2) 操作类型为：update

dts_record_id	dts_instance_id	dts_db_name	dts_table_name	dts_operation_flag	dts_utc_timestamp	dts_before_flag	dts_after_flag	dts_col1	dts_colN
---------------	-----------------	-------------	----------------	--------------------	-------------------	-----------------	----------------	----------	------	----------

2	234	db1	sbtest1	I	1476258463	Y	N	1	JustInsert
2	234	db1	sbtest1	I	1476258463	N	Y	1	JustUpdate

当操作类型为update时，会将update操作拆为2条增量日志。这两条增量日志的dts_record_id, dts_operation_flag 及dts_utc_timestamp相同。

第一条日志记录了更新前的值，所以dts_before_flag=Y, dts_after_flag=N

第二条日志记录了更新后的值，所以dts_before_flag=N, dts_after_flag=Y

3) 操作类型为：delete

dts_record_id	dts_instance_id	dts_db_name	dts_table_name	dts_operation_flag	dts_utc_timestamp	dts_before_flag	dts_after_flag	dts_col1	dts_colN
3	234	db1	sbtest1	D	1476258464	Y	N	1	JustUpdate

当操作类型为delete时，后面的所有column值为被删除的记录值，即为更新前的值。所以dts_before_flag = Y, dts_after_flag=N

下面详细介绍RDS到Datahub数据实时同步作业的配置流程。

同步作业配置流程

下面详细介绍RDS到Datahub数据实时同步作业的配置流程。

1.购买同步链路

进入数据传输服务控制台，进入数据同步页面，点击控制台右上角“创建同步作业”开始作业配置。

在链路配置之前需要购买一个同步链路。同步链路目前支持包年包月及按量付费两种付费模式，可以根据需要选择不同的付费模式。

在购买页面需要配置的参数包括：

- 源实例
同步作业的源实例类型，目前只支持RDS For MySQL.
- 源地域
源地域为同步实例的源RDS实例所在地域。
- 目标实例

目标实例为同步作业的目标实例类型，目前支持 RDS For MySQL, MaxCompute(原ODPS),DataHub。配置RDS->MaxCompute同步链路时，目标实例选择：**Datahub** 即可。

- 目标地域

由于Datahub目前只在上海地区售卖，所以目标地域选择杭州。

- 实例规格

实例规格影响了链路的同步性能，可以根据业务性能选择合适的规则。

- 网络类型

RDS->Datahub支持通过公网、私网 同步数据。如果源RDS没有公网连接地址，那么网络类型只能选择 **私网**。

- 数量

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

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

2.同步链路连接信息配置。

在这一步主要配置：

同步作业名称

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

实例ID配置

在这个步骤中需要配置源RDS实例的实例ID，及目标Datahub实例的project。配置的Datahub project 必须属于登录DTS的阿里云账号的资源。

当这些内容配置完成后，可以点击**授权白名单并进入下一步**。

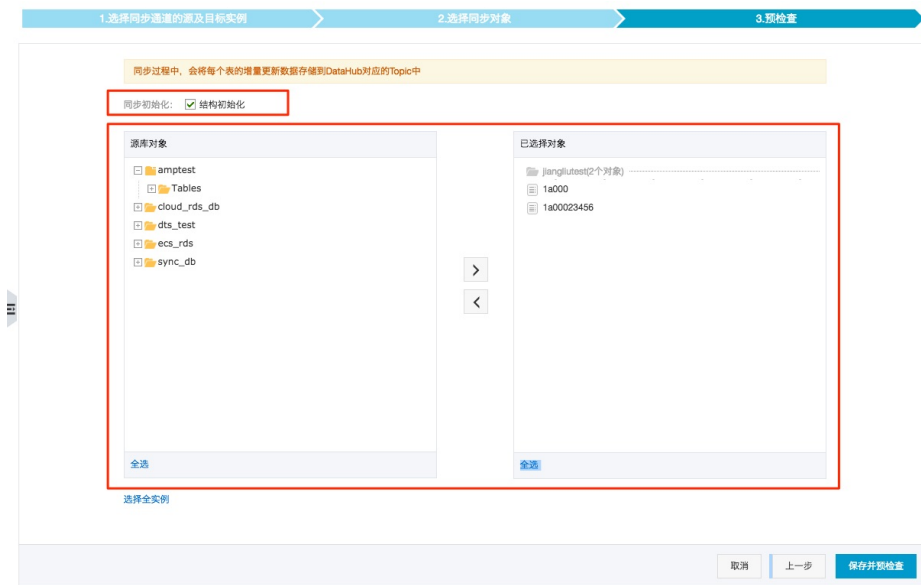
2.授权RDS实例白名单。

这个步骤，主要是将给DTS服务器IP地址添加到RDS实例白名单中，让DTS能够访问RDS实例。

当授权完成后，即进入同步对象选择。

3.选择同步对象

当RDS实例白名单添加完成后，即进入同步表及同步初始化的相关配置。



在这个步骤中，需要配置 同步初始化 和 同步表。其中：

(1)同步初始化

同步初始化选项包括：结构初始化

结构初始化是指对于待同步的表，在Datahub中创建对应的topic，完成topic schema定义。建议选择结构初始化。

(2) 同步表选择

同步表只能选择某些表，不能直接选择整个库。对于同步的表，可以修改topic名称，选择对应的shard key。如需修改，可以点击右边已选择对象后面的编辑按钮，进入修改界面。

当配置完同步对象后，进入同步启动前的预检查阶段。

5.预检查

当上面所有选项配置完成后，即进入启动之前的预检查。具体检查项内容详见本文最后的 **预检查内容** 一节。当同步作业配置完成后，数据传输服务会进行限制预检查，当预检查通过后，可以点击 **确定** 按钮，启动同步作业。

当同步作业启动之后，即进入同步作业列表。此时刚启动的作业处于**同步初始化**状态。初始化的时间依赖于同步表的数量大小。当初始化完成后同步链路即进入**同步中**的状态，此时源跟目标实例的同步链路才真正建立完成。

当同步任务进入 **同步中** 时，可以在Datahub中可以查询出同步表对应的topic：

至此，完成RDS->Datahub数据实时同步作业的配置。

当同步作业建立成功后，可以到流式计算中注册Datahub topic，同时创建流式源表，对于同步到Datahub中的数据进行后续的实时数据计算分析。

实时同步链路在启动之前，会先进行简单的预检查，本小节简单介绍预检查的内容及失败时的修复方式。

预检查项说明

源库连接性检查

检查内容

这个检查项主要检查数据传输服务器同源RDS实例的连接性。数据传输服务使用JDBC协议尝试连接源RDS实例，如果连接失败，那么这个检查项即检查失败。

失败原因

- i. 数据传输服务暂不支持源实例所在地区的RDS实例的实时同步。
- ii. 源实例账号密码有误。

修复方法

提交工单，联系阿里云技术售后，协助解决。

目标库连接性检查

检查内容

这个检查项主要检查数据传输服务器同目标RDS实例的连通性。数据传输服务使用JDBC协议尝试连接目标RDS实例，如果连接失败，那么这个检查项即检查失败。

失败原因

- i. 数据传输服务暂不支持目标实例所在地区的RDS实例的实时同步。
- ii. 目标实例账号密码有误。

修复方法

提交工单，联系阿里云技术售后，协助解决。

源库版本检查

检查内容

这个检查项主要检查：

- i. 源RDS实例的版本号，实时同步功能是否支持。
- ii. 目标RDS实例的版本号是否高等于源RDS实例的版本。

失败原因

- i. 源RDS实例的版本低于数据传输服务支持的版本，实时同步暂时只支持MySQL5.1、5.5、5.6 三个版本。
- ii. 目标RDS实例版本低于源RDS实例版本。

修复方法

- i. 如果源RDS版本过低，那么请首先到RDS控制台，将源RDS升级到5.6版本后，重新创建同步链路。
- ii. 如果是目标RDS实例版本低于源RDS实例，那么请先到RDS控制台，将目标RDS实例升级到5.6版本后，重新创建同步链路。

数据库存在性检查

检查内容

这个检查项主要检查同步DB，在目标库是否已经存在。如果同步DB在目标库不存在，那么数据传输服务会自动创建，但是在如下情况下，自动创建DB会出现失败，此时这个检查项还是会提示检查失败。

- i. 数据库名中含有小写字母、数字、下划线、中划线的其他特殊字符。
- ii. 数据库的字符集不为utf8、gbk、latin1、utf8mb4中的一种。

目标库迁移账号无同步库的读写权限。

当源为RDS实例时，这个检查项不会失败。

源库权限检查

检查内容

这个检查项主要检查源库同步账号的权限是否满足要求。当权限不满足要求时，这个检查项会失败。当源为RDS实例时，这个检查项不会出现失败。

目标库权限检查

检查内容

这个检查项主要检查目标库同步账号的权限是否满足要求，当权限不满足时，这个检查项会失败。

失败原因

- i. 数据传输服务自动在目标RDS实例创建账号失败。
- ii. 数据传输服务对目标RDS实例的同步账号进行读写权限授权失败。

修复方法

提交工单，联系技术同学协助处理。

同名对象存在性检查

检查内容

这个检查项只有选择了同步初始化的同步链路时会进行检查，它主要检查目标RDS实例中是否已经存在跟同步对象同名的对象。

失败原因

当目标RDS实例中已经存在跟同步对象相同名称的对象时，这个检查项会失败。

修复方法

- i. 先将目标库中跟同步对象相同名称的对象删除掉。
- ii. 重新创建同步链路，同步初始化选择：结构初始化+全量数据初始化。

源库server_id检查

检查内容

这个检查主要检查源数据库的server_id 是否设置为大等于2的整数。如果源为RDS实例，那么这个检查项不会失败。

源库binlog是否开启检查

检查内容

这个检查项主要检查源数据库是否开启binlog日志。如果检查项失败，那么说明源数据库没有开启binlog。当源为RDS实例时，那么这个检查项不会失败。

源库binlog模式检查

检查内容

这个检查项主要检查源数据库的binlog模式是否为ROW。如果检查失败，那么说明源数据库的binlog格式不为ROW。当源为RDS实例时，这个检查项不会失败。

约束完整性检查

检查内容

这个检查项主要检查同步的对象列表中，有外键依赖关系的父子表是否都被同步，防止破坏外键约束完整性。

失败原因

当同步的对象中，有部分对象为外键依赖的子表，但是它对应的父表没有同步，破坏了外键约束完整性。

修复方法

对于这种情况，有如下三种修复方式：

- i. 不进行约束完整性检查失败的子表的同步，重新创建同步作业，不进行这些约束完整性检查失败的子表的同步。
- ii. 进行约束完整性检查失败的子表对应的父表的同步，重新创建同步作业，将这些约束完整性检查失败的子表对应的父表添加到同步列表中。
- iii. 删除约束完整性检查失败的子表的外键依赖，修改源数据库将这些子表的外键依赖删除掉，重新创建同步作业。

存储引擎检查

检查内容

这个检查项主要检查同步的对象中，是否存在实时同步不支持的存储引擎。如federated, mrg_myisam, tokudb。

失败原因

当同步对象中，有表的存储引擎为federated, mrg_myisam或tokudb时，这个检查项会失败。

修复方法

先将不支持存储引擎的表的存储引擎修改为InnoDB后，重新创建同步作业。

字符集检查

检查内容

这个检查项主要检查同步的对象中，是否存在实时同步不支持的字符集，例如ucs2。

失败原因

当同步对象的字符集为不支持的字符集时，这个检查项会失败。

修复方法

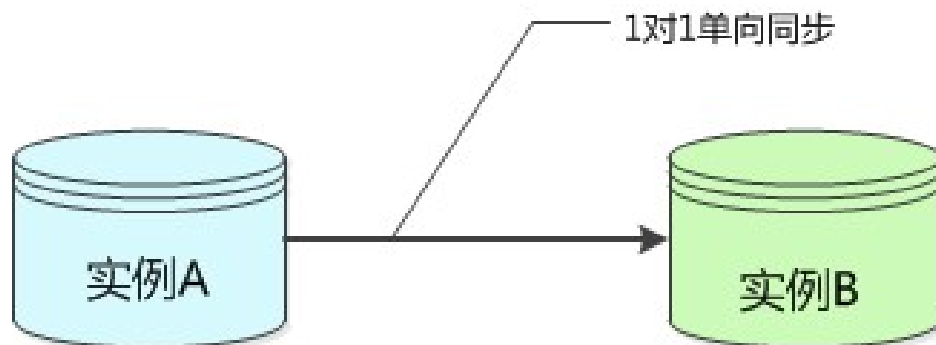
先将不支持字符集的表的字符集修改为utf8、gbk、latin1 中的一种后，重新创建同步作业。

复杂拓扑存在性检查

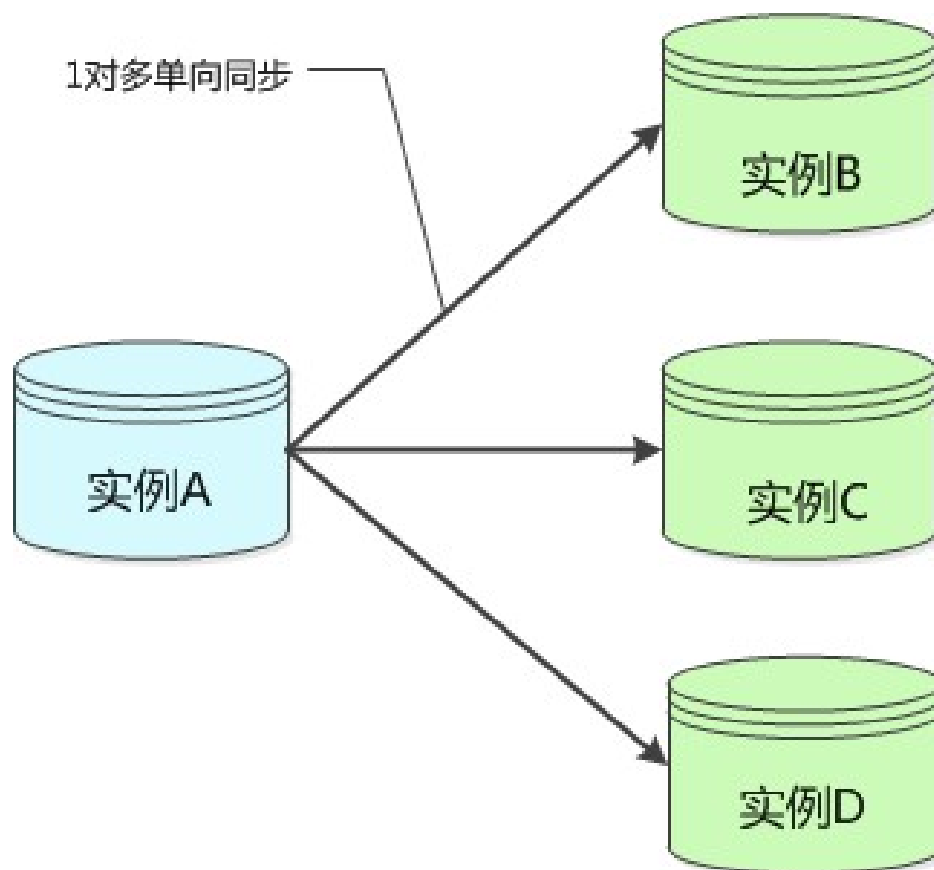
检查内容

这个检查项主要检查同步作业的源跟目标RDS实例上是否存在不支持的同步架构。目前实时同步只能支持：

1对1



1对多



两种架构，对于多对1，级联，双向等同步架构暂不支持。

失败原因

- 源RDS实例上已经存在同步作业,且其为同步作业的目标实例。
- 目标RDS实例上已经存在同步作业。
- 源跟目标RDS存在正在运行的迁移任务，且这个迁移任务迁移的对象跟要创建的同步作业重叠。

修复方法

- 如果源跟目标RDS实例同已存在同步作业相同，创建新同步作业是为了同步新的对象，那么直接修改已经存在的同步链路，修改同步对象，将要添加对象添加进去即可。
- 如果这个同步链路同已经存在的迁移任务冲突，那么必须等迁移任务完成后，才能重新创建同步作业。
- 如果新建同步作业会跟老的同步作业构成：级联架构或双向同步或多对1架构，那么暂不支持。

MySQL老密码格式检查

检查内容

这个检查项，主要检查源实例使用的密码是否为old password。当源为RDS实例时，这个检查项不会失败。

数据传输服务支持在同步过程中，动态修改同步对象。本小节简单介绍一下，如何在同步过程中增加同步对象。

修改限制

只有当同步作业处于**同步中**，**同步失败**的状态时，才可以修改同步对象。

同步起始时间

当新增同步对象时，根据同步作业是否选择同步初始化，同步起始时间不同，具体如下：

1. 如果同步任务选择了同步初始化，那么新增对象也会先进行同步初始化，然后开始增量同步。
2. 如果同步任务未选择同步初始化，那么同步对象会从当前时间点源实例产生的增量数据开始同步。

修改步骤

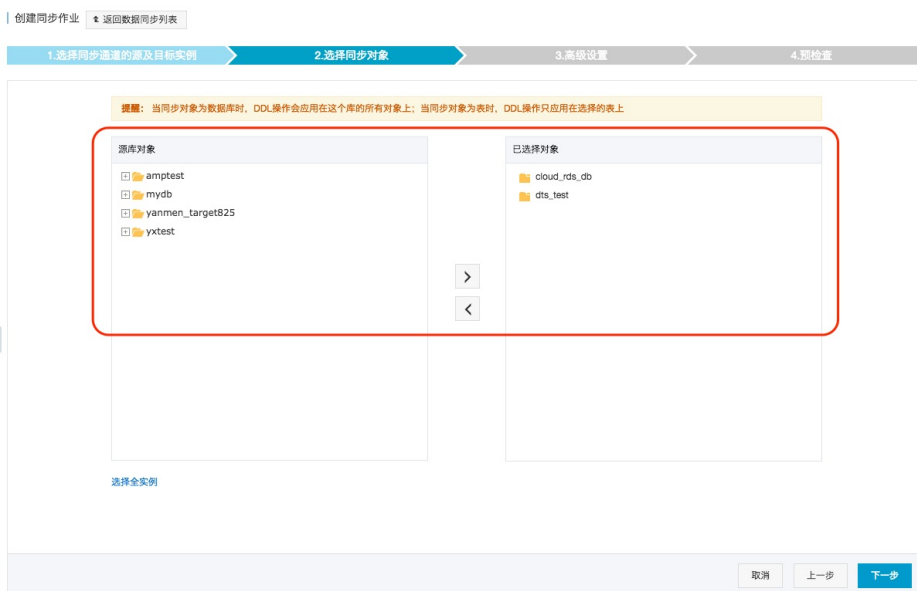
修改同步对象入口。

同步作业列表

实例ID/作业名称	状态(全部)	同步概况	付费方式	操作
dtstofn82nfrqs 杭州-杭州-micro	未配置		包年包月 2016-09-10 00:00:00 到期	配置同步链路 升级
dtscvhfyepwe0 杭州-杭州-large	未配置		包年包月 2016-10-10 00:00:00 到期	配置同步链路
dtst44wtek9v79 杭州-杭州-micro	未配置		包年包月 2016-09-10 00:00:00 到期	配置同步链路 升级
dts1hrj8tao4c483 杭州-杭州-micro	未配置		包年包月 2016-08-21 00:00:00 到期	配置同步链路 升级
dts1yhsqj03d4n 江柳测试	同步中	延时: 0 ms 速度: 0TPS(0.0MB/s)	按量付费	转包年包月 暂停同步 升级 修改同步对象 监控报警 释放同步

对于修改同步对象的同步作业，点击链路右侧**更多**里面的**修改同步对象**，即可以开始修改同步对象。

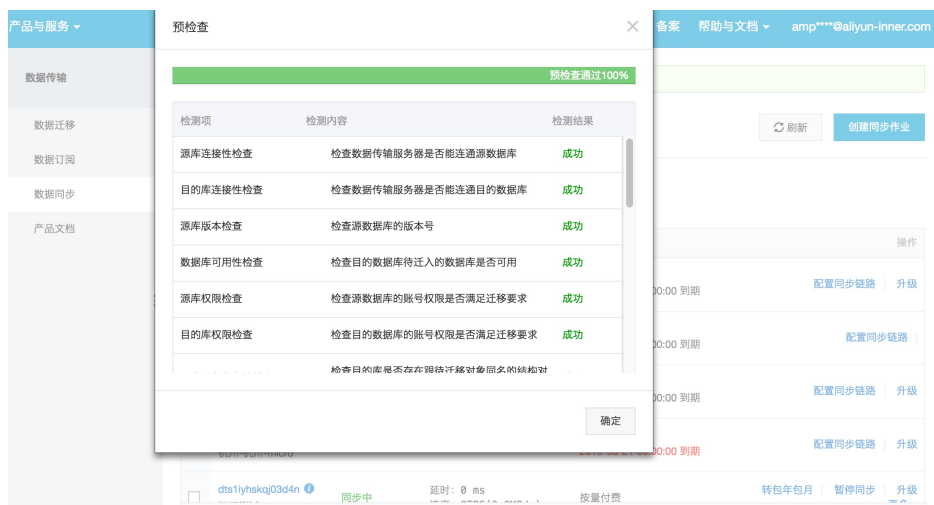
添加同步对象。



在修改同步对象选择界面中，可以根据需要添加需要同步的对象。

当对象修改完成后即进入启动前的预检查。

预检查。



当预检查通过后，即可以点击“启动”完成同步对象添加。

当完成同步对象添加后，如果同步作业需要进行初始化，那么这个任务的状态会由原来的同步中，变成同步中(新增对象初始化中 详情)。此时，后台会重新启动同步作业通道，所以同步延迟会变成-1s，当重启成功后，延迟及同步速度恢复正常。

在同步作业列表中，可以点击详情，查看新增对象初始化的具体进度详情。当新增同步对象完成初始化后，同步作业状态回到同步中。

数据传输服务支持在同步过程中，动态修改同步对象。本小节简单介绍下，如何在同步过程中减少同步对象。

修改限制

只有当同步作业处于**同步中**，**同步失败**的状态时，才可以修改同步对象。

修改步骤

修改同步对象入口

同步作业列表

实例ID/作业名称	状态(全部)	同步概况	付费方式	操作
dtstofn82nfrqs 杭州-杭州-micro	未配置		包年包月 2016-09-10 00:00:00 到期	配置同步链路 升级
dtsqvhfhypwe0 杭州-杭州-large	未配置		包年包月 2016-10-10 00:00:00 到期	配置同步链路
dts444wtek9v79 杭州-杭州-micro	未配置		包年包月 2016-09-10 00:00:00 到期	配置同步链路 升级
dts1hrj8tao4c483 杭州-杭州-micro	未配置		包年包月 2016-08-21 00:00:00 到期	配置同步链路 升级
dts1yhskq03d4n 江柳测试	同步中	延时: 0 ms 速度: 0TPS(0.0MB/s)	按量付费	转包年包月 暂停同步 升级 更多

共有5条，每页显示：20条

对于修改同步对象的同步作业，点击链路右侧 **更多** 里面的 **修改同步对象**，即可以开始修改同步对象。

减少同步对象

创建同步作业 | 返回数据同步列表

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

提醒：当同步对象为数据库时，DDL操作会应用在这个库的所有对象上；当同步对象为表时，DDL操作只应用在选择的表上

源库对象	已选择对象
<input type="checkbox"/> ampctest <input type="checkbox"/> ampctest1 <input type="checkbox"/> ampctest11 <input type="checkbox"/> ampctest121001 <input type="checkbox"/> ampctest121501 <input type="checkbox"/> ampctest2 <input type="checkbox"/> ampctest_3 <input type="checkbox"/> dts <input type="checkbox"/> dts_demo_db1 <input type="checkbox"/> jiangliutest <input type="checkbox"/> sakila123 <input type="checkbox"/> sys_info <input type="checkbox"/> yuzhi_test	<input type="checkbox"/> ampctest(3个对象) <input type="checkbox"/> 1ble_a02 <input type="checkbox"/> 1ble_a03 <input type="checkbox"/> 1ble_a04

选择全实例

取消 保存并预检查

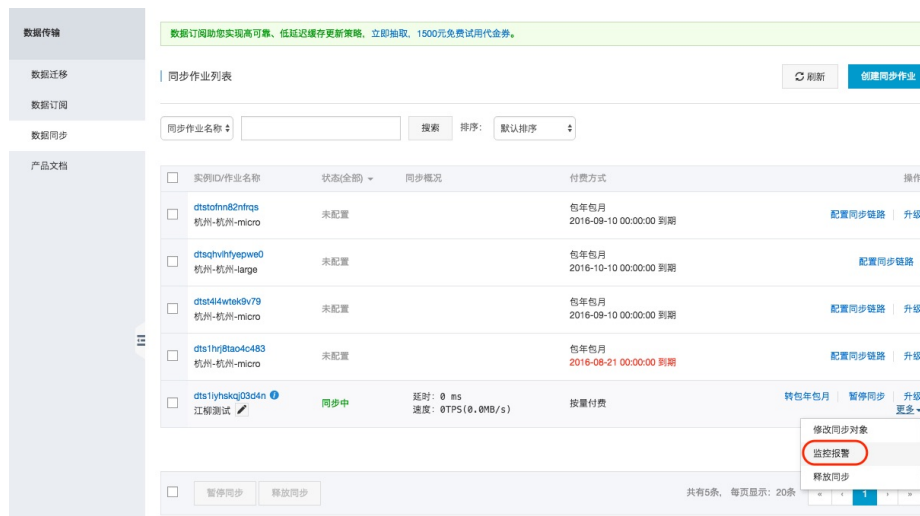
在修改同步对象选择界面中，可以根据需要删除不需要同步的对象。

至此，完成减少同步对象的整个配置流程。

数据传输服务提供了监控报警的功能，可以监控同步作业的状态及同步延迟。同步作业在任何状态下都可以配置监控报警功能。本小节简单介绍如何在数据传输服务控制台配置同步作业状态及延迟的监控报警。

配置步骤

监控报警入口



对于要配置监控报警的同步作业，点击任务右侧 更多 中的的**监控报警**，即可以开始配置。

修改报警策略及联系人入口



默认情况下，监控项都是开启的状态。如果需要修改报警触发条件及手机联系人，可以点击每个报警项后的**修改**入口，进行修改。

修改报警策略及联系人



在这个图中，可以修改同步延迟的延迟阈值及延迟报警的手机联系人。

报警规则：同步延迟 > 10s，是指当这个同步作业的延迟大于10s时，才会向报警手机联系人发送报警短信。用户可以根据业务延迟敏感度，配置延迟阈值。

实时同步支持向多个手机联系人发送报警短信，用户可以根据需要添加报警联系人。

对于同步状态，不能修改报警规则。只有在同步作业状态异常的情况下才会发送报警短信。同步作业状态异常包括：同步初始化失败，同步失败，新增对象初始化失败。

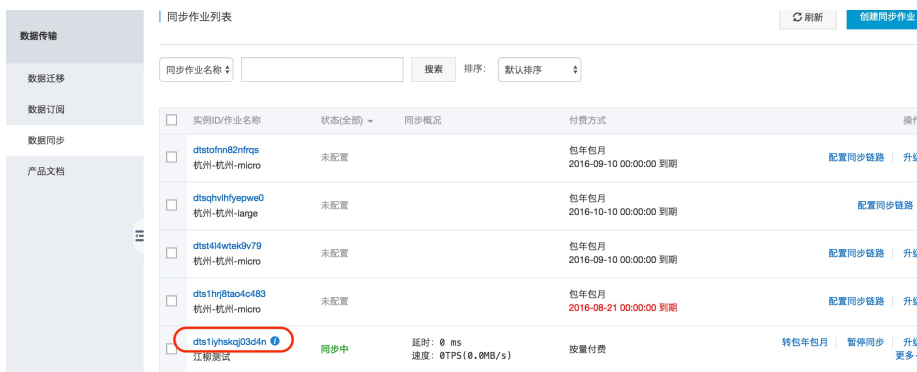


当配置完成后，一旦监控项触发报警规则，数据传输服务会自动向报警联系人发送报警短信。

至此，完成监控报警的相关配置。

数据传输服务提供了同步延迟、同步TPS、同步流量三种性能趋势图，用户可以实时查看同步作业的运行性能状态。本小节简单介绍具体从哪里查看同步性能状态。

同步性能查看入口



当需要查看某一个同步作业的性能时，只要在同步作业列表，点击这个同步作业的作业ID，即进入这个作业的详情页，在详情页中，点击左侧的同步性能，即可以查看具体的同步性能趋势。

同步性能趋势图

数据传输服务了同步延迟、同步TPS、同步流量三种性能趋势图。

- 同步流量：是指每秒钟数据传输服务中的数据写入模块从数据拉取模块中拉取的数据流量，单位为MB/s。
- 同步TPS：是指每秒钟数据传输服务同步到目标RDS实例的query数。
- 同步延迟：是指目标RDS实例同步的最新数据的时间戳跟源RDS实例当前时间戳的时间差,单位为s。

数据传输支持不同阿里云账号下的两个RDS实例之间的实时同步。本小节介绍跨阿里云账号RDS实例间数据实时同步作业的配置流程。

支持场景

1. 支持不同阿里云账号下的两个RDS MySQL实例间的实时同步。
2. 支持公共云与金融云账号下的RDS实例间的数据实时同步。

同步限制

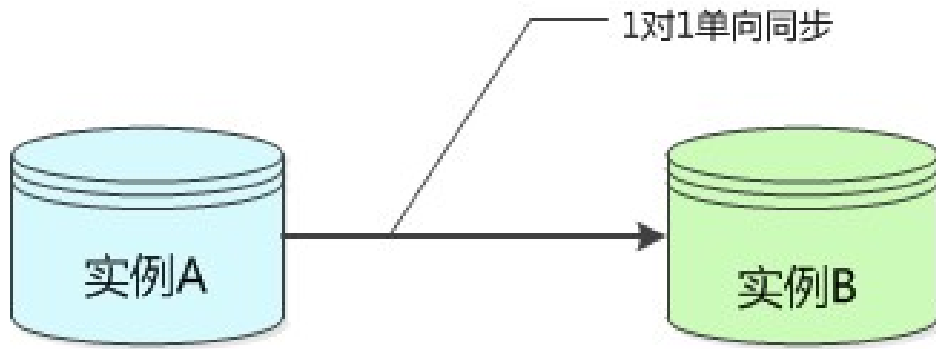
数据源

1. 目前实时同步只能支持RDS MySQL实例,暂不支持其他数据源类型。
2. 目标实例不支持访问模式为标准模式且只有外网连接地址的RDS实例。
3. 不支持香港可用区A的RDS实例的实时同步。

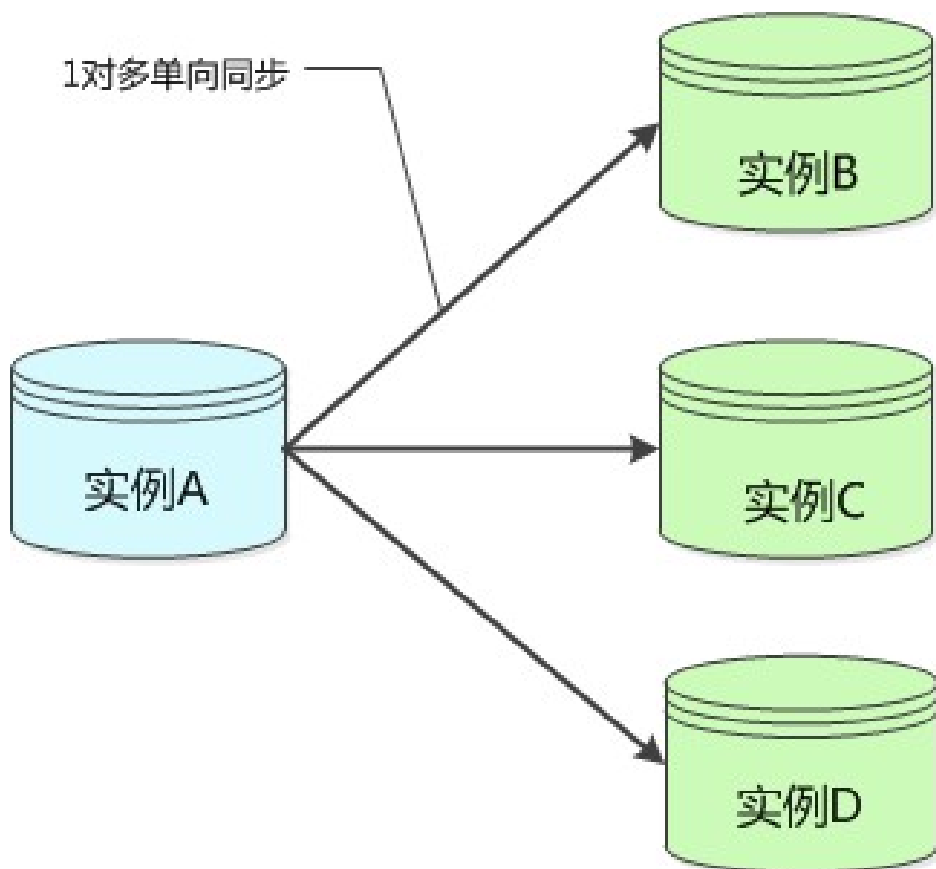
同步架构

目前数据传输服务提供的实时同步功能支持的同步架构有限，其仅能支持如下两种架构：

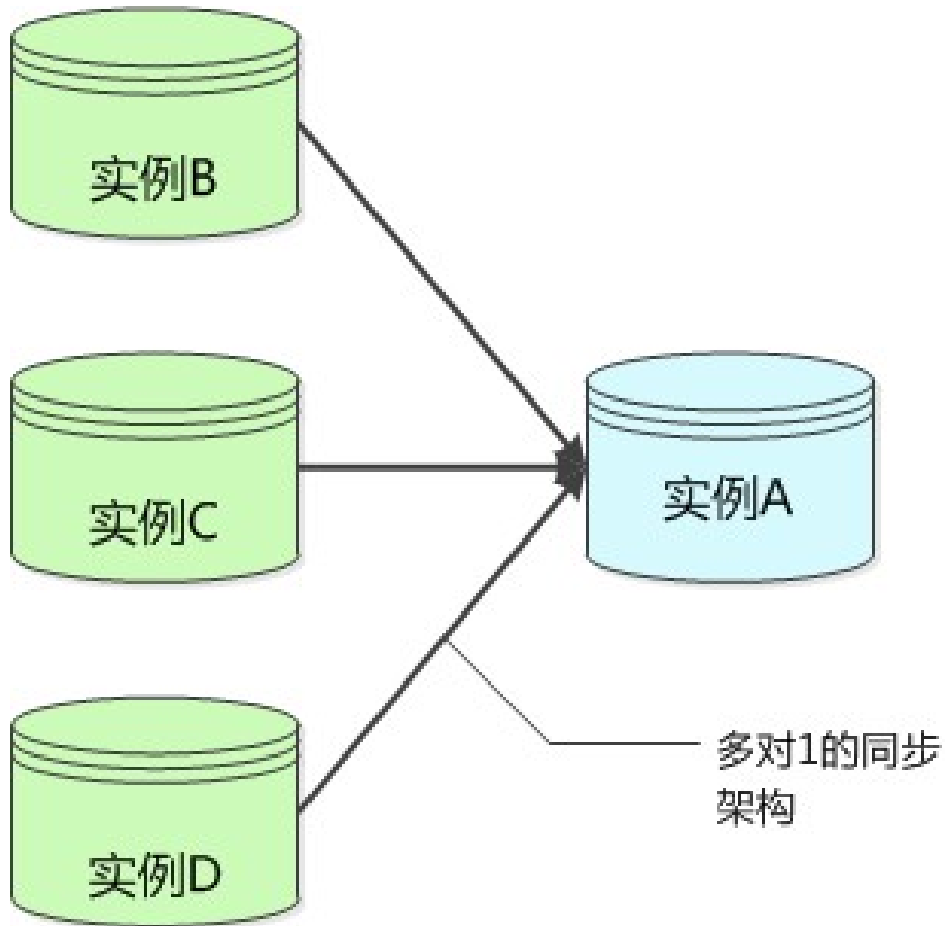
- A->B 即两个实例之间的单向同步。且要求B中同步的对象必须为只读，否则可能导致同步链路异常。



A->B/C/D 即1对多的分发式同步架构,这个架构对目标RDS实例个数没有限制,但是要求目标实例中的同步对象必须为只读,否则可能导致同步链路异常。

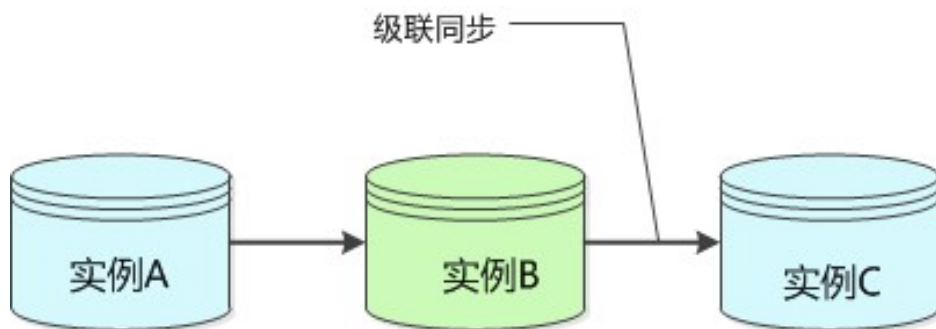


B/C/D->A 即多对1的数据汇总架构。对于这种多对1的同步架构,为了保证同步数据一致性,要求每条同步链路同步的对象不相同。

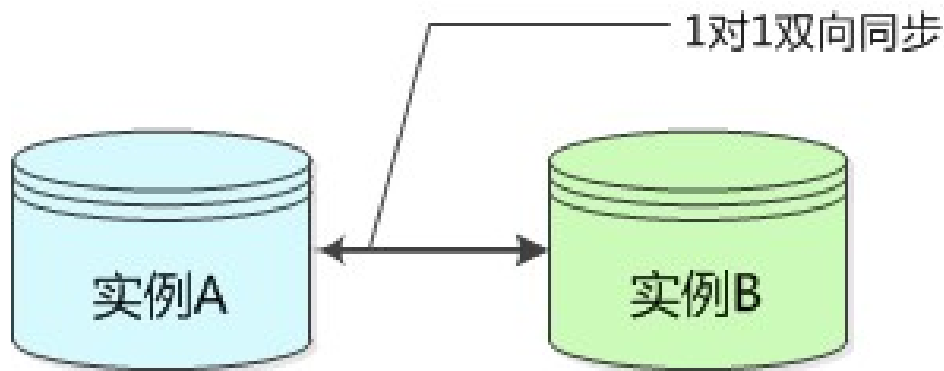


对于下面几种同步架构，暂时不支持：

A->B->C 即级联架构



A->B->A 即实例A和实例B之间的双向同步架构



如果用户配置同步链路过程中，配置了这些不支持的同步架构，那么预检查中的**复杂拓扑**检查项会检查失败。

功能限制

- 不兼容触发器

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

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

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

- rename table 限制

rename table操作可能导致同步数据不一致。例如同步对象只包含表A，不包含表B，如果同步过程中源实例执行了rename A to B的操作，那么改名后的B表的操作不会被同步到目标库。为了解决这个问题，可以选择同步表A、B对应的整个数据库。

准备事项

在配置同步作业前，要确保同步作业的源及目标RDS实例都已经存在。如果不存在，那么请先购买RDS实例。

配置步骤

下面我们详细介绍下创建同步作业的具体步骤。

1. 购买同步链路

使用目标实例对应的阿里云账号登录数据传输服务控制台，进入数据同步页面。

点击控制台右上角“创建同步作业”开始作业配置。

在链路配置之前需要购买一个同步链路。同步链路目前支持包年包月及按量付费两种付费模式，可以根据需要选择不同的付费模式。

在购买页面需要配置的参数包括：

源实例

源实例为同步作业的源实例类型，目前只支持RDS For MySQL。

源地域

源地域为同步链路源实例所在地域。

目标实例

目标实例为同步作业的目标实例类型，目前支持RDS For MySQL, MaxCompute, DataHub。如果进行RDS实例间的同步，那么选择RDS For MySQL即可。

目标地域

目标地域为同步链路目标实例所在地域。

- 实例规格

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

网络类型

对于RDS实例间的数据同步，目前只支持通过私网同步。

数量

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

包年包月
按量付费

① 如果需要购买数据迁移链路，请到[按量付费](#)页面购买

功能：
数据同步
数据订阅

数据迁移只支持按量付费模式，如需购买迁移链路，请到 [按量付费](#) 页面

源实例：
RDS For MySQL

源地域：

华东1(杭州)
华北2(北京)
华北1(青岛)
华南1(深圳)
华东2(上海)
美西(硅谷)
美东

源地域为同步链路源RDS实例所在地域，香港只支持可用区B的RDS实例，订购后不支持更换地域，请谨慎选择

目标实例：
RDS For MySQL

目标地域：

华东1(杭州)
华北2(北京)
华北1(青岛)
华南1(深圳)
华东2(上海)
美西(硅谷)
美东

目标地域为同步链路目标实例所在地域，香港只支持可用区B的RDS实例，订购后不支持更换地域，请谨慎选择

实例规格：
micro

各链路规格对应了同步链路不同的性能表现，如何选择链路规格，请参考 [数据同步规格说明](#) >>

网络类型：
私网

为跨地域传输数据使用的网络模式，当源RDS实例没有公网连接时，只能选择私网

订购时长：

1个月
2
3
4
5
6
7
8
9
1年
2年
3年

数量：

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

2.同步实例连接信息

在这一步主要配置：

- 同步作业名称

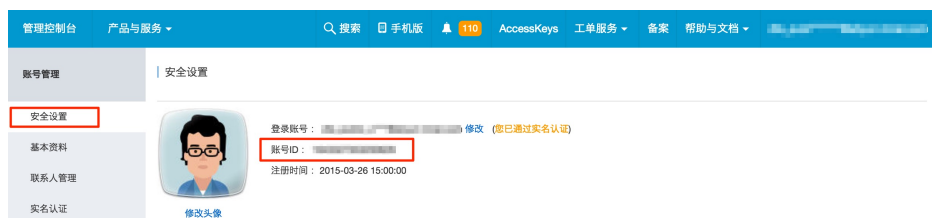
同步作业名称没有唯一性要求，主要为了更方便识别具体的作业，建议选择一个有业务意义的作业名称，方便后续的链路查找及管理。

- 同步链路的RDS实例ID

由于源实例不属于登录的阿里云账号，所以源实例配置时，点击界面右侧的 **其他阿里云账号下的RDS实例**，然后配置：

RDS实例所属阿里云账号

为源RDS实例所属阿里云账号的账号ID，在登录后，到账号管理的安全设置界面获取。



角色名称

为了提升安全性，配置跨账号RDS同步任务的用户，需要得到源RDS实例所属云账号的授权后，才能配置源RDS实例。

这里面配置的角色名称，即为RAM跨账号授权的角色名称。跨账号授权的流程如下：

- (1) 进入RAM控制台的角色管理界面，点击页面右上角的 **新建角色**，开始创建跨账号授权角色。
- (2) 第一步的角色类型，选择用户角色
- (3) 第二步的受信云账号，选择 **其他云账号**，同时，**受信云账号ID** 配置最终配置DTS同步作业的阿里云账号的账号ID



(4) 第四步，配置角

色名称，这个名称就是DTS同步作业配置过程中，需要填写的角色名称。

角色创建完成后，需要修改角色授权策略，授权 受信云账号 只能在DTS 控制台访问自己的云资源。

具体修改步骤如下：

- (1) 在角色管理界面，点击 刚创建角色 后面的 **管理** 按钮，进入角色管理界面。



(2) 在角色管理界面

，点击右上角的 **编辑基本信息**，进入角色编辑框，在编辑框中，修改Principal,添加service定义：

```
"Service": [
  "受信阿里云账号ID@dts.aliyuncs.com"
]
```

受信云账号的账号ID,即最后配置DTS同步作业的阿里云账号ID。dts.aliyuncs.com 为DTS服务代号。

假设配置DTS同步作业的阿里云账号ID为：121852226014398,那么service定义为：

```
"Service": [
  "121852226014398@dts.aliyuncs.com"
]
```

所以，完整的角色定义如下：

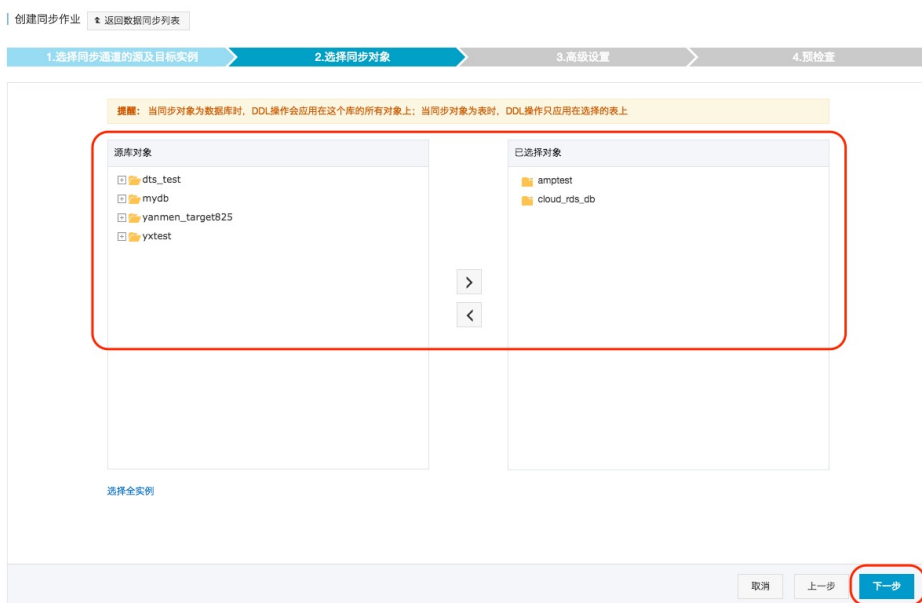
```
{
  "Statement": [
    {
      "Action": "sts:AssumeRole",
      "Effect": "Allow",
      "Principal": {
        "RAM": [
          "acs:ram::1218522260143989:root"
        ]
      },
      "Service": [
        "1218522260143989@dts.aliyuncs.com"
      ]
    }
  ],
  "Version": "1"
}
```

当配置完角色受信身份后，需要将配置DTS任务需要的相关权限授权给角色后，DTS才能扮演这个角色完成任务配置及运行。进入RAM角色管理界面，点击刚才创建的角色后面的 **授权** 按钮，进行DTS系统策略授权。进入角色授权界面后，进入 **精确授权** 界面，在搜索框中搜索 **AliyunDTSRolePolicy**，将这个系统策略授权给角色。



当配置完成后，DTS控制台中填写的角色名称，即为刚才创建的跨账号角色名称。

- RDS实例ID



当配置完同步对象后，进入同步初始化配置。

6.同步初始化配置

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

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



7.预检查

当上面所有选项配置完成后，即进入启动之前的预检查。

当同步作业配置完成后，数据传输服务会进行限制预检查，当预检查通过后，可以点击 **启动** 按钮启动同步作业。

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

本小节介绍如何使用数据传输 DTS（以下简称 DTS）快速创建两个 RDS for MySQL 实例之间的双向同步作业，以实现异地多活（单元化），数据容灾等多种应用场景。

支持功能

支持数据源

- 支持同一个阿里云账号下的两个 RDS for MySQL 实例间的双向同步，包括 VPC 网络 RDS 实例和经典网络 RDS 实例。
- 支持不同阿里云账号下的 RDS for MySQL 实例间的双向同步。

支持同步语法

双向同步支持 RDS for MySQL 所有 DML (Insert/Update/Delete) 和部分 DDL 语法的同步，对 DDL 语法同步支持的具体信息请参考 [创建 RDS 实例间数据实时同步作业](#) 中的信息。

同时，为了保证双向同步链路的稳定性，对于同一张表的 DDL 更新只能在其中一个同步方向进行同步。例如：对于 A->B, B->A 同步，用户需要配置在 **A->B 或者 B->A** 方向上进行 DDL 同步。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。

冲突检测

为了保证同步数据的一致性，用户需要保证同一个主键/业务主键 / 唯一键的记录只在双向同步的某个节点进行更新。

如果用户业务发生误操作或其他原因，对于同一个主键/业务主键 / 唯一键在双向同步的两个节点均进行了更新，那么会出现同步冲突，为了最大程度维护双向同步实例的稳定性，DTS 支持冲突检测和修复。

注意：由于数据双向同步过程中可能存在两端系统时间上的差异，以及同步存在延时等多种因素，DTS 无法完全保证其冲突检测机制能够完全防止数据的冲突，用户在使用双向同步时，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 / 唯一键的记录只在双向同步的某个节点进行更新。

下面介绍 DTS 支持的冲突类型和修复方法。

支持冲突类型

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

Insert 导致的唯一性冲突是指当同步 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前后分别）Insert 一条主键值一样的记录，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 同步冲突主要包括两种：

- (1) Update 要更新的记录在同步目标实例中不存在。
- (2) Update 要更新的记录除了主键/唯一键外，其他字段值不完全匹配。

- Delete 对应的记录不存在

Delete 同步冲突是指要 Delete 的记录在同步的目标实例中不存在。

出现这种冲突的原因可能是目标实例自身的业务对这条记录进行了主键更新或删除了这条记录，导致同步过来的 Delete 操作命中不了记录。

支持冲突修复策略

对于上面的同步冲突，DTS 提供了如下几种修复策略，用户可以在配置双向同步时，选择任意一种冲突修复策略。

- TaskFailed (遇到冲突，任务报错退出)

这种冲突修复策略是指当同步遇到上面的冲突类型时，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。这种冲突修复方法为默认的冲突修复策略

- Ignore (遇到冲突，直接使用目标实例中的冲突记录)

这种冲突修复策略是指当同步遇到上面的冲突类型时，直接跳过当前同步语句，继续往下执行，选择使用目标实例中的冲突记录。

- Overwrite (遇到冲突，直接覆盖目标实例中的冲突记录)

这种冲突修复策略是指当同步遇到上面的冲突类型时，直接覆盖目标实例中的冲突记录。

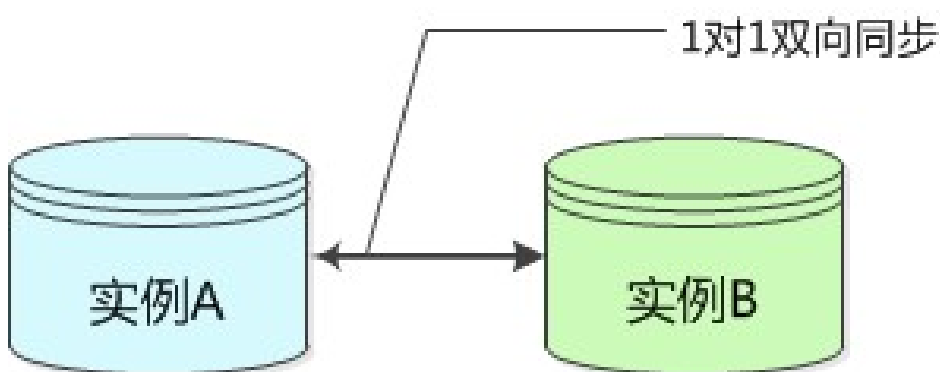
同步限制

数据源

1. 目前只有 RDS for MySQL 实例间支持双向同步,其他异构数据源间暂不支持双向同步。
2. 目标实例不支持访问模式为标准模式且只有外网连接地址的 RDS 实例。

同步架构

目前数据传输服务提供的双向同步功能只支持两个 RDS for MySQL 实例之间的双向同步，暂不支持多实例之间的双向同步，即其能支持如下架构：



功能限制

不兼容触发器

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

。

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

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

rename table 限制

rename table 操作可能导致同步数据不一致。例如同步对象只包含表A，不包含表B，如果同步过程中源实例执行了 **rename a to b** 的操作，那么改名后的表 b 的操作不会被同步到目标库。为了解决这个问题，可以选择同步表 a、b 对应的整个数据库。

DDL 语法同步方向限制

为了保证双向同步链路的稳定性，对于同一张表的 DDL 更新只能在其中一个同步方向进行同步。例如：对于 A->B，B->A 同步，用户需要配置在 **A->B 或者 B->A** 方向上进行 DDL 同步。一旦在某个同步方向配置了 DDL 同步，则在反方向上不支持 DDL 同步。

准备事项

在配置同步作业前，要确保双向同步作业的源及目标 RDS for MySQL 实例都已经存在。如果不存在，那么请先购买 RDS 实例 购买 RDS 实例。

配置步骤

下面我们详细介绍下创建任意两个RDS实例之间的双向同步实例的具体步骤。

1. 购买同步链路

进入数据传输 DTS 控制台，进入数据同步页面，点击控制台右上角“创建同步作业”开始作业配置。

在链路配置之前需要购买一个双向同步实例。双向同步实例目前支持包年包月及按量付费两种付费模式，可以根据需要选择不同的付费模式。

在购买页面需要配置参数包括：

- 源实例

源实例选择 RDS For MySQL。

- 源地域

源地域为双向同步实例的某个 RDS 实例所在的地域。

- 目标实例

目标实例选择 RDS for MySQL。

- 目标地域

目标地域为双向同步实例的另外一个 RDS 实例所在地域。

- 同步拓扑

同步拓扑选择：双向同步

- 网络类型

网络类型是指跨地域同步数据使用的网络模式，目前只支持私网。

- 实例规格

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

- 数量

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

功能： 数据同步 数据订阅
数据迁移只支持按量付费模式，如需购买迁移链路，请到 [按量付费](#) 页面

源实例：

源地域：

源地域为同步链路源RDS实例所在地域，香港只支持可用区B的RDS实例，订购后不支持更换地域，请谨慎选择

目标实例：

目标地域：
目标地域为同步链路目标实例所在地域，香港只支持可用区B的RDS实例，订购后不支持更换地域，请谨慎选择

同步拓扑： 单向同步 双向同步
双向同步可以支持两个RDS实例间的数据双向同步

网络类型：
为跨地域传输数据使用的网络模式，当源RDS实例设有公网连接串时，只能选择私网

实例规格：

当购买完同步实例，返回数据传输 DTS 控制台。此时，刚购买的同步实例处于未配置状态，它包含了两个同步任务，可以进行两个方向的同步任务的配置。

<input type="checkbox"/>	实例ID/作业名称	状态(全部) ▾	同步概况	付费方式	同步架构(全部) ▾	操作
<input type="checkbox"/>	dts1v0zipsar6ev	--	--	按量付费	双向同步	查看同步作业 ▾ 升级更多
<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"/>	hangzhou-hangzhou-micro	未配置		未配置 未配置	配置同步链路	
<input type="checkbox"/>	hangzhou-hangzhou-micro	未配置		未配置 未配置	配置同步链路	

此时，可以点击同步作业右侧的“配置同步链路”开始某个方向的同步任务的配置。下面详细介绍某个方向的

同步任务的配置流程。

2. 同步链路连接信息配置

在这一步主要配置：

- 同步作业名称

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

- 同步作业的 RDS 实例 ID

选择 RDS 实例 ID 时，下拉菜单中只列出对应阿里云账号下的 RDS for MySQL 实例。如果要配置跨阿里云账号的两个 RDS for MySQL 实例间的双向同步，同步任务的配置请参考 [跨阿里云账号下的实例迁移同步](#)



当这些内容配置完成后，可以点击**授权白名单并进入下一步**进行 RDS 实例的白名单配置。

3. 授权 RDS 实例白名单

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

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

当白名单授权后，点击下一步，进入同步账号创建。

4. 创建目标库上的同步账号

这个步骤主要是在目标 RDS 实例上创建一个同步账号，账号名字为：dtssyncwriter，在同步过程中，不能删除这个账号，否则会导致同步链路中断。

5. 同步对象及同步策略配置

当创建完目标 RDS 实例的同步账号后，即进入同步对象及同步策略的配置。

- 是否过滤 DDL

这个配置项定义这个同步方向的同步是否要同步 DDL，如果需要同步那么选择 **否**，否则选择 **是**。一旦这个同步方向选择同步 DDL，那么同一张表在另一个同步方向不支持 DDL 同步。

- DML 同步类型

这个配置项定义需要同步的DML类型，默认包括：Insert、Update、Delete，用户可以根据业务需要调整需要同步的 DML 类型。

- 冲突修复策略

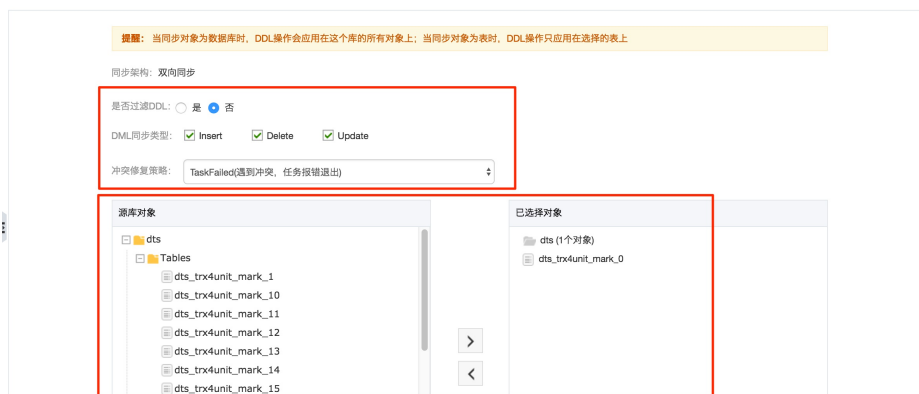
这个配置项定义同步冲突的修复策略，默认为 TaskFailed，用户可以根据业务情况选择合适的冲突修复策略。例如以 A 节点为主业务中心，B 节点为备用业务中心，那么 A 节点正常运行情况下，应该以 A 节点的业务写入为主，即 A->B 同步方向的冲突策略为 overwrite，B->A 同步方向的冲突策略为 ignore。

- 同步对象选择

实时同步的同步对象的选择粒度可以支持到表级别,即用户可以选择同步某些库或是同步某几张表。

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

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



当配置完同步对象后，进入同步初始化配置。

6. 同步初始化配置

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

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

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。



7. 预检查

当上面所有选项配置完成后，即进入启动之前的预检查。

当同步作业配置完成后，数据传输服务会进行限制预检查，当预检查通过后，可以点击 **启动** 按钮,启动同步作业。

当同步作业启动之后，即进入同步作业列表。此时刚启动的作业处于**同步初始化**状态。初始化的时间长度依赖于源实例中同步对象的数据量大小。当初始化完成后同步链路即进入**同步中**的状态，此时同步任务才真正建立完成。

当这个方向的同步任务配置完成后，另外一个方向同步任务的源及目标 RDS 实例就固定了不能更改。

作业名称	状态	同步概况	源实例/目标实例	操作
hangzhou-hangzhou-small	同步中	延时: 933 毫秒 速度: 0TPS(0.00MB/s)	rm-bp1nx2h5fm3v6q26l rm-bp185gso372vbd118	暂停同步 更多
hangzhou-hangzhou-small	未配置		rm-bp185gso372vbd118 rm-bp1nx2h5fm3v6q26l	配置同步链路

完成某一方向的同步任务的配置后，既可以开始另外一个方向的同步作业的配置，具体的配置步骤同上面的步骤 2~ 步骤 6。

数据传输服务DTS支持RDS到分析型数据库AnalyticDB之间的数据实时同步。通过DTS提供的RDS->AnalyticDB实时同步功能，可以将企业数据集中分析，实现企业内部各种BI、CRM、实时报表系统等。本小节介绍如何使用数据传输服务DTS快速创建RDS(MySQL)->AnalyticDB实例之间的实时同步作业，实现RDS增量数据到AnalyticDB的数据实时同步。

支持功能

1. 支持同一个阿里云账号下RDS(MySQL)->AnalyticDB实例的数据实时同步。
2. 支持不同阿里云账号下的RDS->AnalyticDB实例间的数据实时同步。

支持SQL操作类型

RDS->AnalyticDB实时同步支持的SQL操作类型如下：

Insert

Delete

Update

Alter table(只支持增加列，减少列/修改列名 不支持)

Truncate table

Drop table

目前暂不支持Create Table，如果同步过程中RDS实例中新建表需要同步，那么参考 **新增同步对象**，修改这个同步作业，将新建表添加到同步对象中。

配置步骤

下面我们详细介绍下创建任意两个RDS实例之间的同步链路的具体步骤。

1. 购买同步链路

进入数据传输服务控制台，进入数据同步页面，点击控制台右上角“创建同步作业”开始作业配置。

在链路配置之前需要购买一个同步链路。同步链路目前支持包年包月及按量付费两种付费模式，可以根据需要选择不同的付费模式。

在购买页面需要配置的参数包括：

- 源实例

源实例目前只支持 RDS For MySQL, 所以选择RDS For MySQL即可。

- 源地域

源地域为同步链路源实例所在地域。

- 目标实例

目标实例目前支持 RDS For MySQL、MaxCompute、Datahub 及 分析型数据库AnalyticDB。这里面选择 分析型数据库AnalyticDB。

- 目标地域

目标地域为同步AnalyticDB数据库所在地域。

- 实例规格

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

- 数量

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

数据传输服务（包年包月）

包年包月 | 按量付费

① 如果需要购买数据迁移链路，请到按量付费页面购买

功能：**数据同步** | 数据迁移

数据迁移只支持按量付费模式，如需购买迁移链路，请到按量付费页面

源地域：**华东1(杭州)** | 华北2(北京) | 华北1(青岛) | 华南1(深圳) | 华东2(上海) | 美国(硅谷) | 美东
亚太(新加坡) | 香港

源地域为同步链路源RDS实例所在地域，香港只支持可用区的RDS实例，订购后不支持更换地域，请谨慎选择

目标地域：**华东1(杭州)** | 华北2(北京) | 华北1(青岛) | 华南1(深圳) | 华东2(上海) | 美国(硅谷) | 美东
亚太(新加坡) | 香港

目标地域为同步链路目标RDS实例所在地域，香港只支持可用区的RDS实例，订购后不支持更换地域，请谨慎选择

实例规格：**micro**

各链路规格对应了同步链路不同的性能表现，如何选择链路规格，请参考 数据同步规格说明 >>

订购时长：**1年** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 1年 | 2年 | 3年

数量：**1**

购买清单 0例

当前配置

功能：数据同步
源地域：华东1(杭州)
目标地域：华东1(杭州)
实例规格：micro
订购时长：1个月
数量：1
配置费用：
¥0.00
原价¥770.00 省¥770.00
数据传输服务（包年包月DTS）_0折

立即购买 | 加入清单

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

2. 同步链路连接信息配置

在这一步主要配置：

- 同步作业名称

同步作业名称没有唯一性要求，主要为了更方便识别具体的作业，建议选择一个有业务意义的作业名称，方便

后续的链路查找及管理。

- 源实例ID

选择源RDS实例ID时，下拉菜单中只列出对应阿里云账号下的RDS For MySQL实例。

- 目标AnalyticDB 数据库

配置目标AnalyticDB对应的数据库名称

The screenshot shows the '同步作业名称' (Sync Job Name) field set to '江柳测试RDS2ADS'. Below it, the '源实例信息' (Source Instance Information) section is highlighted with a red box, showing '实例类型: RDS For MySQL' and '实例地区: 华东 1'. The '实例ID' dropdown is set to 'rsv-6pr1dvw240@yu4600n0'. The '目标实例信息' (Target Instance Information) section is also highlighted with a red box, showing '实例类型: ADS' and '实例地区: 华东 1'. The '数据库' dropdown is set to 'jiangliutest'. At the bottom right, there are buttons for '取消' (Cancel) and '授权白名单并进入下一步' (Authorize Whitelist and Proceed to Next Step).

当这些内容配置完成后，可以点击**授权白名单并进入下一步**进行RDS白名单添加及AnalyticDB账号权限授权。

3. AnalyticDB 账号授权

在这个步骤需要将AnalyticDB对应数据库的读写权限授权给DTS 同步账号。

The screenshot shows the '创建同步作业' (Create Sync Job) process at step 2, 'ADS账号授权' (ADS Account Authorization). A progress bar at the top indicates the current step. Below the progress bar, a message states: '将数据同步至ADS，需要将数据库 "jiangliutest-9c90fce7-cn-hangzhou-1.ads.aliyuncs.com" 的下述权限授予同步账号' (To sync data to ADS, you need to grant the following permissions to the sync account for the database "jiangliutest-9c90fce7-cn-hangzhou-1.ads.aliyuncs.com"). A table lists the permissions: Create, Alter, Show, Select, Insert, Delete, and Describe. At the bottom right, there are buttons for '取消' (Cancel), '上一步' (Previous Step), and '下一步' (Next Step).

4. 授权RDS实例白名单

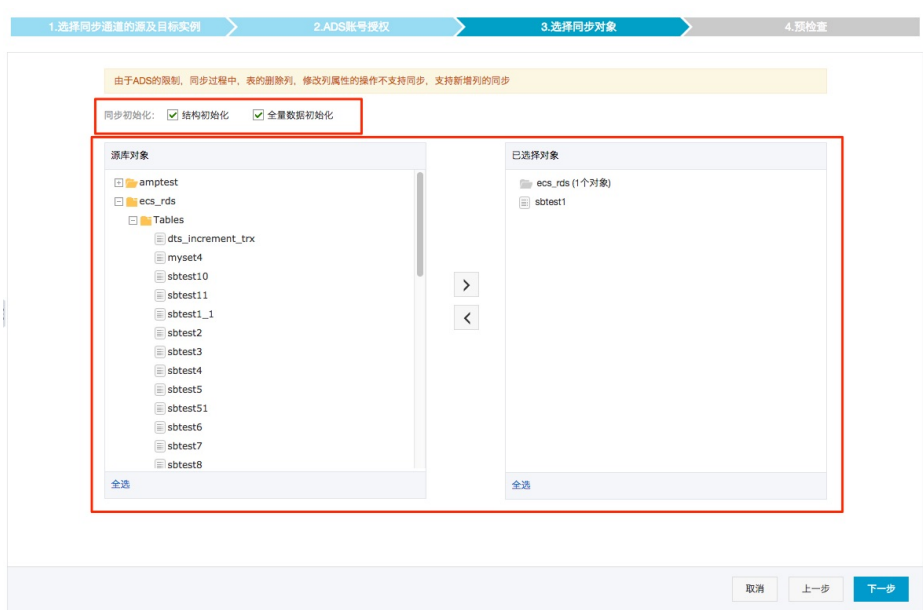
这个步骤，主要是将数据传输服务器IP添加到同步RDS实例的白名单中。避免因为RDS设置了白名单，数据传输服务器连接不上RDS导致同步作业创建失败。

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

当白名单授权后，点击下一步，进入同步账号创建。

5. 选择同步对象

当AnalyticDB同步账号授权完成后，即进入同步对象的选择步骤。实时同步的同步对象的选择粒度可以支持到表级别，即用户可以选择同步某些库或是同步某几张表。



当配置完同步对象后，进入同步初始化配置。

6. 表结构定义

如果选择了结构初始化，那么需要对每个要同步的表，定义其在AnalyticDB中的表类型，主键，分区列等。AnalyticDB的表定义。AnalyticDB的表定义可以参考 [创建和修改表](#)

7. 预检查

当上面所有选项配置完成后，即进入启动之前的预检查。

当同步作业配置完成后，数据传输服务会进行限制预检查，当预检查通过后，可以点击 **启动** 按钮,启动同步作业。

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

数据订阅

DTS提供了RDS MySQL增量数据订阅的功能。要订阅消费增量数据，需要进行两个步骤的操作：

1. 在DTS控制台创建订阅RDS实例的订阅通道。
2. 使用DTS提供的SDK连接这个订阅通道，订阅并消费增量数据。

本小节主要简单介绍在DTS控制台创建订阅通道的流程。

购买订阅实例。

登录数据传输服务控制台，进入数据订阅页面。

点击控制台右上角“创建订阅通道”开始订阅通道配置。

在链路配置之前需要购买一个订阅通道。数据订阅目前支持包年包月及按量付费两种付费模式，可以根据需要选择不同的付费模式。

在购买页面需要配置的参数包括：

源实例地域

源地域为订阅通道对应源RDS实例所在的地域。

数量

数量为一次性购买的订阅通道的数量，如果购买的是按量付费实例，一次最多购买99条链路。

数据传输服务（包年包月）

当购买完订阅通道，返回数据传输控制台，点击新购链路右侧的“配置订阅通道”开始链路配置。

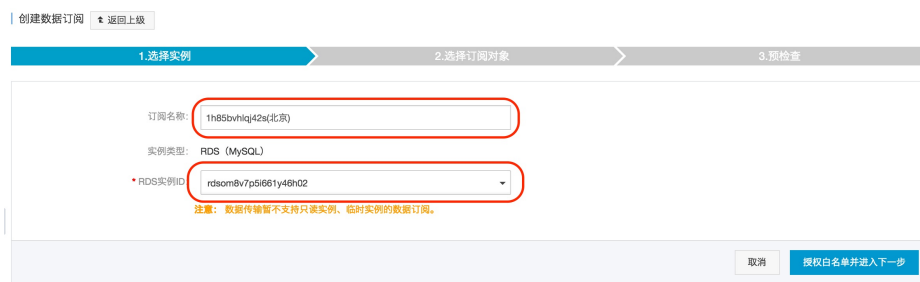
配置RDS实例ID。

订阅通道名称

订阅通道名称没有唯一性要求，主要为了方便识别具体的通道，建议选择一个有业务意义的通道名称，方便后续的链路查找及管理。

订阅通道的RDS实例ID

选择需要订阅的RDS实例，下拉菜单中只列出对应阿里云账号下的RDS For MySQL实例。

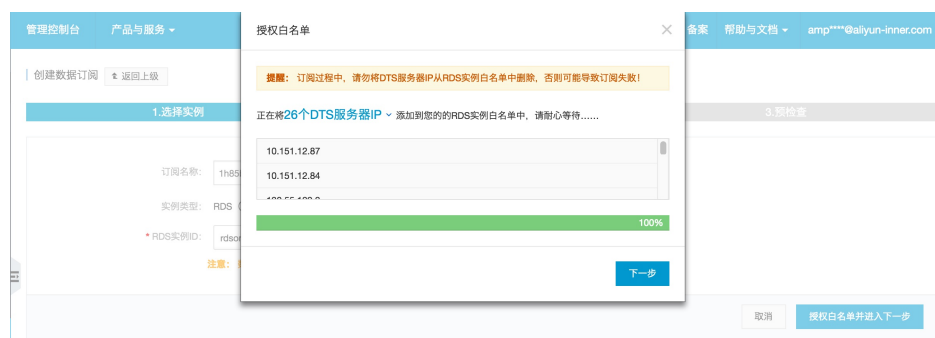


当这些内容配置完成后，可以点击**授权白名单并进入下一步**。

授权RDS实例白名单。

这个步骤，主要是将数据传输服务器IP添加到订阅RDS实例的白名单中。避免因为RDS设置了白名单，数据传输服务器连接不上RDS导致订阅通道创建失败。

为了保证订阅通道的稳定性，在订阅过程中，请勿将这些服务器IP从RDS实例的白名单中删除。



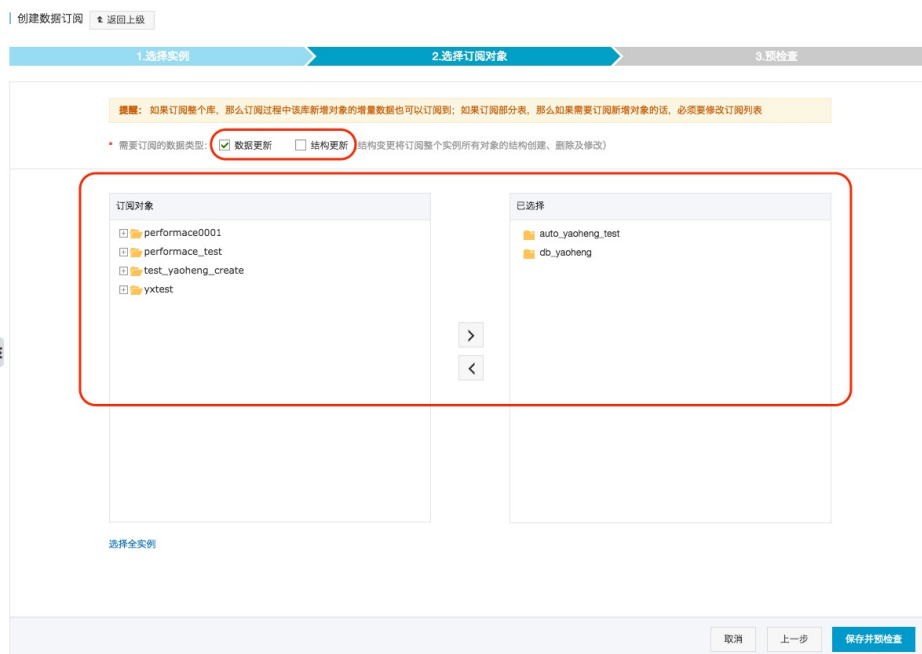
当白名单授权后，点击**下一步**，进入订阅账号创建。

选择订阅对象。

在这个步骤中，用户需要选择要订阅的对象列表。

DTS的订阅对象粒度细分为库、表。即用户可以选择订阅某些库或者是订阅某几张表。

DTS将订阅数据类型细分为数据变更、结构变更。如果只选择订阅对象及数据变更的话，那么只能订阅到insert/delete/update三种数据变更内容，如果需要订阅结构变更（DDL），那么需要选择订阅数据类型中的结构变更。一旦订阅了结构变更，那么DTS会将整个RDS实例的所有结构变更拉取出来，用户需要使用SDK过滤需要的数据。



当选择完订阅对象后，即进入订阅通道启动之前的预检查阶段。

预检查。

当订阅通道配置完成后，DTS会进行限制预检查，当预检查通过后，可以点击**开始订阅**，启动订阅通道。

启动订阅。

当订阅通道刚创建完，会进入初始化阶段，这个阶段一般会持续1分钟左右。初始化完成后，即可在线查看订阅数据，或者使用SDK订阅增量数据。

DTS支持在消费的过程中，随时修改消费时间点。一旦修改完消费时间点，那么下游SDK拉取到的增量数据从修改后的消费时间点开始。修改的消费时间点必须在订阅通道的数据范围之内。DTS目前只支持在控制台修改消费点，不支持在SDK中指定消费位点。

本小节介绍如何在控制台修改消费时间点。

操作步骤

停止SDK消费进程。

在修改消费位点之前，需要先确保连接这个订阅通道的所有SDK下游都停掉。可以在DTS控制台查看订阅通道的消费者来源(IP)是否为空来确认订阅通道的下游是否都已经停止，如果为空，表示下游都已经停止。

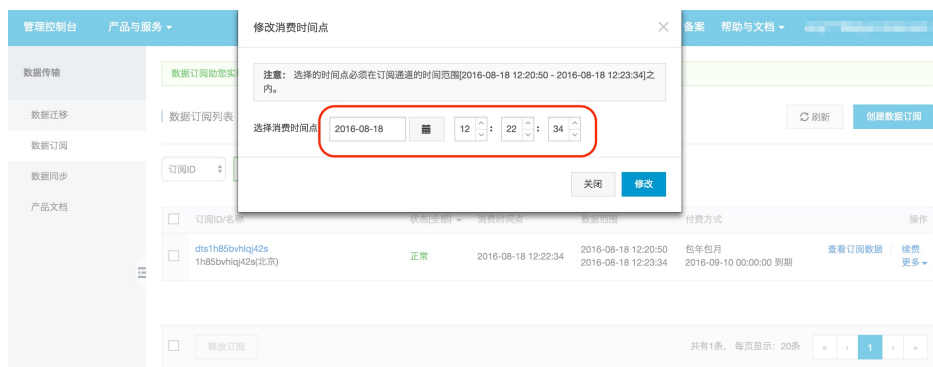
如果订阅通道的消费者来源(IP)为空，那么说明这个订阅通道的所有下游SDK都已经停止。

修改消费时间点。

DTS只支持在DTS控制台修改消费位点，修改入口如下：



当需要修改通道消费时间点时，将鼠标挪到这个通道的消费时间点上，会出现一个类似“笔”的图标，点击此进入修改页面。



这里面配置的消费时间点，必须在当前的数据通道范围内。

重启SDK消费进程。

当修改完消费位点后，即可重启本地的SDK消费进程，此时SDK会从修改的消费位点开始订阅增量数据。

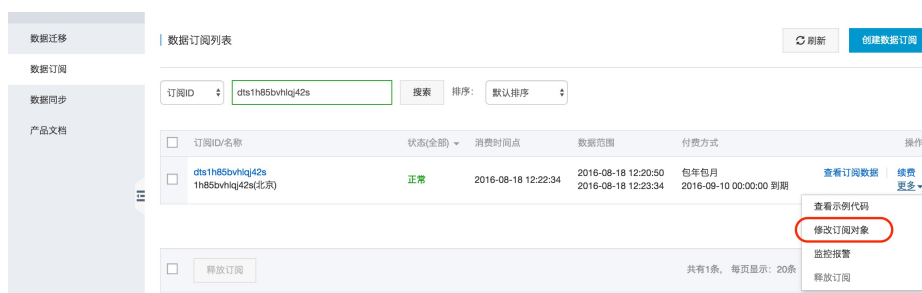
DTS支持在订阅消费的过程中，动态增加/减少订阅对象。如果增加了订阅对象，那么修改完成后，订阅通道会从当前时间拉取新增订阅对象的增量数据。如果减少订阅对象，那么修改完成后，SDK中将不能再订阅到这个对象的数据。

本小节介绍如何在控制台修改订阅对象。

操作步骤

修改订阅对象入口。

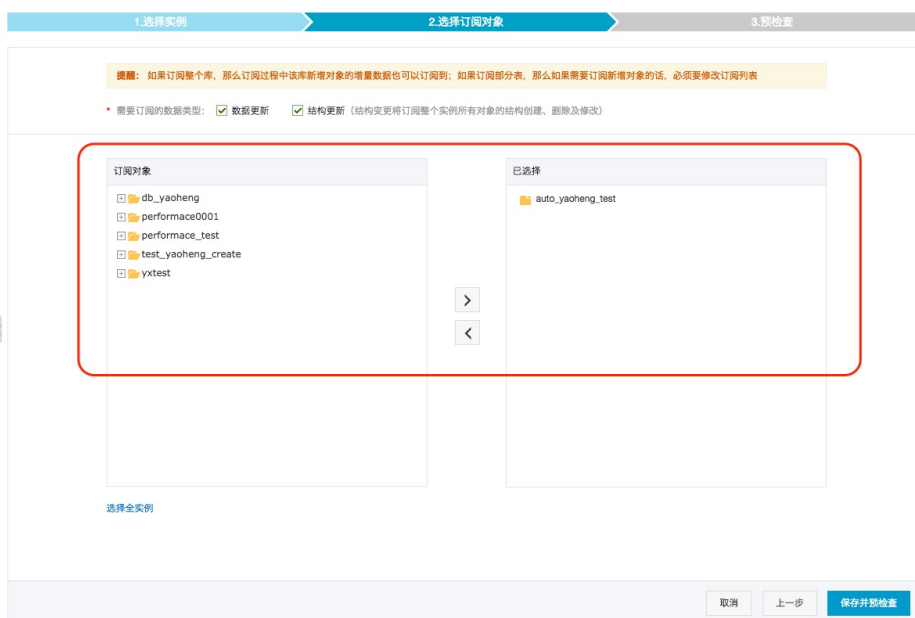
DTS只支持在控制台修改订阅对象，具体入口如下：



对于需要修改订阅对象的订阅通道，点击最右边的更多，下拉出更多的操作，其中有一个**修改订阅对象**的操作入口。

修改订阅对象。

当点击“修改订阅对象”入口之后，进入选择订阅对象的页面。



在这个页面可以增减订阅对象，也可以修改订阅的变更数据类型。当修改完之后，重新进入了预检查阶段。

当预检查通过后，可以点击开始订阅，此时订阅通道会重新进行初始化。



当初始化完成后，订阅通道进入正常状态，此时可以查看订阅数据，或使用SDK订阅增量数据。

DTS支持在线查看订阅数据，即可以在DTS控制台查看订阅通道中的增量数据，能够查看的时间范围为订阅通

道的数据范围。本小节介绍在线查看订阅数据的方式，及数据记录的各个字段的含义。

查看订阅数据入口

DTS控制台查看订阅数据的入口如下：

订阅ID/名称	状态(全部)	消费时间点	数据范围	付费方式	操作
<input type="checkbox"/> dts1h85bvhqj42s 1h85bvhqj42s(北京)	正常	2016-08-18 12:34:05	2016-08-18 12:20:50 2016-08-18 13:35:45	包年包月 2016-09-10 00:00:00 到期	查看订阅数据 续费更多

只有正常和异常状态下的订阅通道可以查看订阅数据。处于初始化、预检查中的订阅通道不能查看订阅数据。

查看订阅数据

当点击查看订阅数据时，即进入订阅数据的查看页面，如下：

订阅数据

提醒： 订阅通道aityun_bj_eca_rdsom8v7p5i661y46h02-1-0的数据范围为：2016-08-18 12:20:50 ~ 2016-08-18 13:37:43

时间：2016-08-18 12:34:05 至：2016-08-18 13:35:45 查看

服务器ID	库名	表名	变更类型	位点	时间戳	操作
			HEARTBEAT	148401@869	2016-08-18 12:34:06	
			HEARTBEAT	148401@869	2016-08-18 12:34:06	
			HEARTBEAT	148401@869	2016-08-18 12:34:06	
			HEARTBEAT	148401@869	2016-08-18 12:34:06	
			HEARTBEAT	148401@869	2016-08-18 12:34:06	
			HEARTBEAT	148401@869	2016-08-18 12:34:06	

可以根据配置具体时间范围，查看特定时间范围内的增量数据。配置的时间段必须在订阅通道的时间范围内。

订阅数据字段含义

订阅数据展示的数据格式为DTS内部的数据格式，数据记录的各个字段含义为：

- 服务ID：RDS实例的IP:PORT
- 库名：这条变更记录对应表所在的数据库的库名
- 表名：这条变更记录对应的表的表名
- 变更类型：这条变更记录的变更类型，包括：update、delete、insert、replace、begin、commit、ddl、dml、heartbeat

其中heartbeat为DTS提供的心跳记录，每秒钟会有一条记录，主要用来表示订阅通道健康状况，如果每秒有一条heartbeat，说明订阅通道正常，否则说明订阅通道异常。

- 位点：这条变更记录在binlog中的位点，格式为binlog_file_offset@binlog_file_id，其中binlog_file_id为binlog文件名的后缀，binlog_file_offset为这条binlog记录在binlog文件中的偏移量。

- 时间戳：这条变更记录写入到binlog文件中时的时间戳
- 字段名称：这条变更记录对应的每个字段的名称
- 字段类型：这条变更记录对应的每个字段的数据类型
- 字段编码：这条变更记录对应的字符集编码
- 前镜像：这条变更记录变更之前的各个字段的值。对于insert操作，没有前镜像值

服务ID	库名	表名	变更类型	位点	时间戳	操作
			HEARTBEAT	264271@869	2016-08-18 13:46:010	
100.98.198.99-3306	auto_yaoheng_test	jiangliutest	INSERT	264531@869	2016-08-18 13:46:13	查看数据详情
字段详情						
字段名称	字段类型	字段编码	前镜像	后镜像		
sn	INT64			1000		
data	STRING	utf8		jiangliutest		

- 后镜像：这条变更记录变更之前的各个字段的值。对于delete操作，没有后镜像值

服务ID	库名	表名	变更类型	位点	时间戳	操作
100.98.198.99-3306	auto_yaoheng_test	jiangliutest	DELETE	265428@869	2016-08-18 13:46:29	查看数据详情
字段详情						
字段名称	字段类型	字段编码	前镜像	后镜像		
sn	INT64		1000			
data	STRING	utf8	jiangliutest after update			

update 操作的前镜像和背景镜像都有值。

服务ID	库名	表名	变更类型	位点	时间戳	操作
100.98.198.99-3306	auto_yaoheng_test	jiangliutest	UPDATE	265124@869	2016-08-18 13:46:24	查看数据详情
字段详情						
字段名称	字段类型	字段编码	前镜像	后镜像		
sn	INT64		1000	1000		
data	STRING	utf8	jiangliutest	jiangliutest after update		

对于操作类型为DDL、DML的记录，只有字段详情，没有字段名称、字段类型、字段编码、前镜像及后镜像几个字段。如下图：

服务ID	库名	表名	变更类型	位点	时间戳	操作
100.98.198.99-3306	auto_yaoheng_test		DDL	260296@869	2016-08-18 13:43:42	查看数据详情
字段详情						
DDL定义						
<pre>create table jiangliutest(sn BIGINT AUTO_INCREMENT ,data VARCHAR(500),PRIMARY KEY (sn)) AUTO_INCREMENT = 1000</pre>						

数据传输的数据订阅功能需要通过数据传输提供的SDK进行增量数据的订阅及消费。

在使用SDK消费之前，需要现在数据传输控制台创建需要订阅的RDS实例的订阅通道。

当订阅通道创建完成后，使用SDK可以实时订阅订阅通道中的增量数据。目前：

1. 数据传输只提供JAVA版本SDK，SDK下载地址。

2. 一个订阅通道只能被一个SDK消费，如果启动多个SDK连接同一个订阅通道时，只能有一个SDK进程拉取到增量数据。如果有多个下游SDK需要订阅同一个RDS的增量数据。那么需要为每个下游SDK创建一个订阅通道。

SDK中定义了多种类对象，本小节简单介绍SDK的这些类的接口定义。

RegionContext接口定义

```
setAccessKey(accessKey)
```

设置安全凭证，参数为需要订阅数据的订阅通道对应的阿里云账号的AccessKey。

```
setSecret(AccessKeySecret)
```

设置安全凭证，参数为阿里云账号对应的AccessKeySecret。可以到AK页面创建并获取。

```
setUsePublicIp(usePublicIp)
```

配置SDK运行服务器是否使用公网订阅数据。如果通过公网订阅数据，那么参数usePublicIp参数为True，否则为False。

数据订阅可以通过内网进行订阅，但是SDK在建立订阅连接之前需要先跟数据传输管控系统通信获取订阅通道的物理连接地址，SDK跟数据传输管控系统需要通过互联网通信，所以即使通过内网订阅数据，SDK部署服务器也需要挂载公网IP。

ClusterClient接口定义

```
void addConcurrentListener(ClusterListener arg0)
```

添加下游监听者，监听者加入到一个ClusterClient中，才可以订阅订阅通道中的增量数据。

参数ClusterListener arg0 为类ClusterListener的对象。

```
void askForGUID(String arg0)
```

请求某个订阅通道的增量数据，参数String arg0 为订阅通道的ID，需要到数据传输控制台获取，如下图标识。

订阅ID/名称	状态(全部)	消费时间点	数据范围	付费方式	操作
<input type="checkbox"/> dtsus9j14cq3qk0 us9j14cq3qk0(杭州)	未配置			按量付费	配置订阅通道 转包年包月
<input type="checkbox"/> dtsrhdcqm88r7fm rhdcqm88r7fm(杭州)	正常	2016-05-18 19:50:14	2016-05-29 16:44:38 2016-05-30 17:22:06	包年包月 2016-06-27 00:00:00 到期	查看订阅数据 续费更多

```
List<ClusterListener> getConcurrentListeners()
```

获取这个ClusterClient中的监听者列表，接口返回类型为List <ClusterListener >。

```
void start()
```

启动SDK客户端，开始订阅增量数据。

```
void stop()
```

停止SDK客户端，停止订阅增量数据。由于SDK中拉取数据和回调notify的是同一个线程执行的，如果notify的消费代码中有信号不可打断的功能时，那么stop函数可能不能正常关闭掉客户端。

ClusterListener接口定义

```
void notify(List<ClusterMessage> arg0)
```

这个函数主要用于定义增量数据的消费，当SDK接收到数据时，会通过notify通知ClusterListner消费数据。例如示例demo的消费方式，就是将订阅数据打印到屏幕上。

这个函数输入参数类型为：List <ClusterMessage >，其中ClusterMessage为订阅数据存储的结构对象，具体定义详见ClusterMessage接口定义。

ClusterMessage接口定义

每个ClusterMessage保存RDS中的一个事务的数据记录，事务中的每条记录通过Record保存，本小节介绍ClusterMessage的主要接口函数。

```
Record getRecord()
```

这个接口从ClusterMessage中获取一条变更记录。这个变更记录表示RDS binlog文件中的每一条记录，例如begin ,commit,update,insert等。

```
void ackAsConsumed
```

为了简化下游SDK进程容灾，数据订阅服务端支持SDK的消费位点保存，当下游SDK异常宕机并重启后，会自动从上次异常退出的最后一个消费位点继续订阅并消费数据。

在message消费完成后，需要调用这个接口向数据传输服务端汇报一个ACK，通知服务端更新下游SDK的消费位点，保证SDK异常重启后消费数据的完整性。

Record接口定义

Record代表订阅的RDS binlog中的每条记录，例如begin, commit,update等。

String getAttribute(String key)

这个函数可以获取Record中主要的一些属性值。传入参数为属性名，返回这个属性的值。

可以调用这个函数获取属性值的属性名及对应的属性值如下表：

key	说明
record_id	这条Record的ID，这个ID在订阅过程中不保证递增
instance	这条Record对应的数据库实例的连接地址，格式为：IP:Port
source_type	这条Record对应数据库实例的引擎类型，目前取值为：mysql
source_category	这条Record的类型，目前取值为：full_recorded
timestamp	这条Record落binlog的时间，这个时间同时也是这条SQL在RDS中执行的时间
checkpoint	这条Record对应的binlog文件的位点，格式为：file_offset@file_name，file_name为binlog文件的数字后缀
record_type	这条Record对应的操作类型，主要取值包括：insert/update/delete/replace/ddl/begin/commit/heartbeat
db	这条Record更新表，对应的数据库名
table_name	这条Record更新表的表名
record_recording	这条Record对应的编码
primary	这条Record更新表的主键列名
fields_enc	这条Record每个字段值的编码，各个字段之间用逗号隔开，如果非字符类型那么取值为空

Type getOpt()

获取这条记录的变更类型，包括：

insert、delete、update、replace、ddl、begin、commit、heartbeat。

其中heartbeat为数据传输内部定义的心跳表，主要用于检查订阅通道是否健康，理论上每秒都会产生一条 heartbeat。

String getCheckpoint()

获取这条变更记录在binlog中的位点，返回的位点格式为:

binlog_offset@binlog_fid.

其中binlog_offset为变更记录在binlog文件中的偏移量，binlog_fid为binlog文件的数字后缀，例如binlog文件名为mysql-bin.0008，那么binlog_fid为8。

String gettimestamp()

获取这条变更记录在binlog中记录的运行时间戳。

String getDbname()

获取这条变更记录修改的表所对应的数据库库名。

String getTablename()

获取这条变更记录修改表对应的表名。

String getPrimaryKeys()

获取这条变更记录对应的主键列名，如果是联合主键，那么这些列名之间用逗号分隔。

DBType getDbType()

获得订阅实例的数据库类型，目前数据传输仅支持RDS MySQL，所以这个值为MySQL。

String getServerId()

获取这条变更记录对应的RDS MySQL实例运行进程的IP:PORT。

int getFieldCount()

获取这条变更记录的字段Field的个数。

List<Field> getFieldList()

这个函数的返回结果的数据类型为List <Field >。

List<Field> 包含了这条变更记录对应表的所有字段的定义及变更前后的镜像值，Field对象的定义详见**Field接口定义**。

Boolean isFirstInLogevent()

判断这条Record 是否数据库批量变更中的第一条事务日志，如果是的话返回True,否则返回False。

Field接口定义

Field类定义了每个字段的编码、类型、字段名、字段值及是否为主键等属性，本小节介绍Field类的各个接口定义。

```
String getEncoding()
```

获取这个字段值的编码格式。

```
String getFieldname()
```

获取这个字段的名称。

```
Type getType()
```

获取这个字段的数据类型，Type的定义具体参见下面的字段类型定义。

```
ByteString getValue()
```

获取这个字段的值，返回类型为ByteString,当值为空时，返回NULL。

```
Boolean isPrimary()
```

判断这个字段是否是表的主键列，如果是返回True，否则返回False。

这一节，您将学到如何用DTS Java SDK完成一些基本的操作。

初始化RegionContext

RegionContext 主要用于保存设置安全认证信息及访问网络模式设置。下面代码显示如何初始化RegionContext，设置安全认证凭证及网络访问模式。

```
import java.util.List;
import com.aliyun.drc.clusterclient.RegionContext;
import org.slf4j.Logger;
import org.slf4j.LoggerFactory;
public class MainClass
{
public static void main(String[] args) throws Exception {
// 创建一个RegionContext
RegionContext context = new RegionContext();
// 配置阿里云账号的AccessKey及AccessKeySecret
context.setAccessKey("<AccessKey>");
context.setSecret("<AccessKeySecret>");
```

```
// 运行SDK的服务器是否使用公网IP连接DTS订阅通道
context.setUsePublicIp(true);
// 下面为其他调用代码 .....
.....
}
}
```

如果要使用SDK，必须先初始化RegionContext，配置连接订阅通道的安全认证等信息。上面的接口setAccessKey设置的是阿里云账号的AccessKey

setSecret 设置的是阿里云账号的AccessKeySecret

AccessKey及AccessKeySecret是由阿里云的系统直接分配给用户的，称为ID对，用户标识用户，可到阿里云用户中心创建获取。

setUsePublicIp 是告诉DTS，您本地SDK运行服务器是否用公网IP连接订阅通道。如果设置为true，那么订阅数据流走公网，否则走内网。

初始化ClusterClient

SDK连接订阅通道，接受增量数据等操作都是通过类ClusterClient来完成的，下面代码创建了一个ClusterClient

```
import java.util.List;
import com.aliyun.drc.clusterclient.ClusterClient;
import com.aliyun.drc.clusterclient.DefaultClusterClient;
import com.aliyun.drc.clusterclient.RegionContext;
public class MainClass
{
public static void main(String[] args) throws Exception {
// 创建一个RegionContext
RegionContext context = new RegionContext();
context.setAccessKey("<AccessKey>");
context.setSecret("<AccessKeySecret>");
context.setUsePublicIp(true);

// 创建订阅消费者
final ClusterClient client = new DefaultClusterClient(context);

// 下面是一些其他调用代码
.....
}
}
```

初始化Listener

消费数据的功能通过类Listener来实现。初始化完ClusterClient，需要添加listener，Listener要定义notify函数来接受订阅数据并进行数据消费。下面的代码中实现了最简单的消费逻辑，将订阅到的数据打印到屏幕。

```
import com.aliyun.drc.clusterclient.ClusterClient;
```

```

import com.aliyun.drc.clusterclient.ClusterListener;
import com.aliyun.drc.clusterclient.DefaultClusterClient;
import com.aliyun.drc.clusterclient.RegionContext;
import com.aliyun.drc.clusterclient.message.ClusterMessage;

public class MainClass
{
public static void main(String[] args) throws Exception {
// 初始化一个RegionContext对象
.....
//初始化一个ClusterClient对象
.....
ClusterListener listener = new ClusterListener(){
@Override
public void notify(List<ClusterMessage> messages) throws Exception {
for (ClusterMessage message : messages) {
//打印订阅到的增量数据
System.out.println(message.getRecord() + ":" + message.getRecord().getTablename() + ":"
+ message.getRecord().getOpt());
//消费完数据后向DTS汇报ACK，必须调用
message.ackAsConsumed();
}
}
}
}
}

```

DTS实现了SDK的数据消费时间点保存到DTS服务端的机制，简化用户使用SDK时，实现SDK容灾的复杂度。上面示例代码中的 `askAsConsumed()`接口就是将SDK消费的最新一条数据的位点及时间戳汇报给DTS服务端。汇报了时间戳信息，如果SDK意外宕机重启后，会自动从DTS服务端获取这个消费时间点，然后从这个时间点重启，解决数据重复问题。

启动ClusterClient

```

import java.util.List;

import com.aliyun.drc.clusterclient.ClusterClient;
import com.aliyun.drc.clusterclient.ClusterListener;
import com.aliyun.drc.clusterclient.DefaultClusterClient;
import com.aliyun.drc.clusterclient.RegionContext;
import com.aliyun.drc.clusterclient.message.ClusterMessage;
import org.slf4j.Logger;
import org.slf4j.LoggerFactory;
public class MainClass
{
public static void main(String[] args) throws Exception {
//初始化RegionContext
.....
//初始化ClusterClient
.....
//初始化ClusterListener
.....
// 添加监听者

```

```

client.addConcurrentListener(listener);
// 设置请求的订阅通道ID
client.askForGUID("dts_rdsrjiei2u2afnb_DSF");
// 启动后台线程，注意这里不会阻塞，主线程不能退出
client.start();
}

```

上面代码中接口askForGUID设置这个client需要请求的订阅通道ID。这个订阅通道ID从DTS控制台上获取。一旦配置了订阅通道ID，那么这个SDK就能获取这个订阅通道中的增量数据。

在启动client之前，需要将监听者listener添加到client中，这样当client从订阅通道中拉取到增量数据时，会同步回调用listener的notify方法开始进行数据消费。

数据订阅SDK订阅到的数据为DTS自定义格式，本小节简单介绍各种类型SQL语句解析的代码实现。

DDL解析

如果一个Record 是DDL语句，那么这个Record 的操作类型为DDL。且DDL语句存储在第一列的value中。具体获取DDL语句的代码如下：

```

String ddl_string;
Record.Type type=record.getOpt();
if(type.equals(Record.Type.DDL)){

List<DataMessage.Record.Field> fields = record.getFieldList();
ddl_string = fields.get(0).getValue().toString();

}

```

Insert 解析

如果一个Record 是Insert语句，那么这个Record的操作类型为INSERT。具体获取Insert完整语句的代码如下：

```

StringBuilder insert_string=new StringBuilder();
Record.Type type=record.getOpt();
DataMessage.Record.Field field;
StringBuilder FieldName=new StringBuilder();
StringBuilder FieldValue = new StringBuilder();

if(type.equals(Record.Type.INSERT)){
int i=0;
List<DataMessage.Record.Field> fields = record.getFieldList();
for (; i < fields.size(); i++) {
field = fields.get(i); FieldName.append("`"+field.getFieldname().toLowerCase()+"`");
FieldValue.append(field.getValue());
if (i != fields.size() - 1) {
FieldName.append(',');
}
}
}

```

```

FieldValue.append(',');
}
}
insert_string.append("insert " + record.getTablename()+" "+FieldName.toString()+")
values("+FieldValue.toString()+");");

}

```

Update 解析

如果一个Record 是Update语句，那么这个Record的操作类型为UPDATE。

update更新前的字段存储在：Record.getFieldList()中索引为偶数的Field，更新后的字段值存储在：索引为奇数的Field。

下面的示例代码介绍当update的表有主键时，获取Insert完整语句的代码示例：

```

StringBuilder update_string=new StringBuilder();
Record.Type type=record.getOpt();
DataMessage.Record.Field field;
StringBuilder SetValue = new StringBuilder();
StringBuilder WhereCondition = new StringBuilder();
String ConditionStr;
boolean hasPk=false;
boolean pkMode=false;
boolean hasSet=false;

if(type.equals(Record.Type.UPDATE)){
int i=0;
DataMessage.Record.Field OldField = null;
DataMessage.Record.Field NewField = null;
List<DataMessage.Record.Field> fields = record.getFieldList();
for (; i <fields.size(); i++) {
if (i % 2 == 0) {
OldField = fields.get(i);
continue;
}

NewField = fields.get(i);
if (field.isPrimary()) {
if (hasPk) {
WhereCondition.append(" and ");
}
//where old value
ConditionStr = getFieldValue(OldField);
if(ConditionStr==null){ WhereCondition.append("`"+field.getFieldname().toLowerCase()+"`" + " " + "is null");
}else{
WhereCondition.append("`"+field.getFieldname().toLowerCase()+"`" + " = " + NewField.getValue());
}
}
hasPk = true;
}
if (hasSet) {
SetValue.append(COMMA);
}
}

```

```

SetValue.append("`"+field.getFieldname().toLowerCase()+"`" + " = " + field.getValue());
String setStr = getFieldValue(field);
hasSet = true;
}

update_string.append("Update "+record.getTablename() +" Set " + SetValue + " Where "+WhereCondition +";");
}

```

Delete 解析

如果一个Record 是Delete语句，那么这个Record的操作类型为DELETE。下面的代码示例简单介绍当delete对应的表有主键时，如何获取完整的Delete语句：

```

StringBuilder delete_string=new StringBuilder();
Record.Type type=record.getOpt();
DataMessage.Record.Field field;
StringBuilder FieldName=new StringBuilder();
StringBuilder FieldValue = new StringBuilder();
StringBuilder DeleteCondition = new StringBuilder();
boolean hasPk=false;
boolean pkMode=false;

if(type.equals(Record.Type.DELETE)){
int i=0;
List<DataMessage.Record.Field> fields = record.getFieldList();
delete_string.append("Delete From" + record.getTablename() + "where");
// 表是否有主键？
if (record.getPrimaryKeys() != null) {
pkMode = record.getPrimaryKeys().length() > 0 ? true : false;
}
for (; i < fields.size(); i++) {
if ((pkMode && !field.isPrimary())) {
continue;
}
if (hasPk) {
delete_string.append(" and ");
}
delete_string.append(field.getFieldname() + "=" + field.getValue());
hasPk = true;
}
delete_string.append(";");
}
}

```

Replace 解析

如果源库执行了Replace语句，那么这个Record的操作类型为UPDATE或INSERT。当replace设置的值不存在时，Record的操作类型为INSERT，当replace设置的值存在时，Record的操作类型为UPDATE。

Begin 解析

如果一个Record 是Begin语句，那么这个Record的操作类型为BEGIN。begin语句没有实际内容，所以不需要对Field做处理，只需要判断操作为begin即可,具体获取代码如下：

```
StringBuilder sql_string = new StringBuilder();
Record.Type type = record.getOpt();
if(type.equals(Record.Type.BEGIN)){
    sql_string.append("Begin");
}
```

Commit 解析

如果一个Record 是Commit语句，那么这个Record的操作类型为COMMIT。commit语句没有实际内容，所以不需要对Field做处理，只需要判断操作为commit即可。

```
StringBuilder sql_string = new StringBuilder();
Record.Type type = record.getOpt();
if(type.equals(Record.Type.COMMIT)){
    sql_string.append("commit");
}
```

本小节简单介绍如何运行DTS提供的多种示例代码。

使用Maven依赖形式的SDK

本小节简单介绍如何使用DTS提供的Maven依赖形式的Demo。

下载Demo

1. 首先，需要先下载Demo。
2. 使用解压工具，解压上面下载下来的Demo。

运行Demo

本小节以Eclipse为例，简单介绍如何运行SDK Demo。

1. 在Eclipse中导入Demo

在Eclipse中，点击菜单File->Import 操作入口。在弹出的窗口中，选择Maven下的Existing Maven Projects,点击Next。在Import Maven Projects窗口中，选择并导入上面解压后的Demo。

2. 运行Demo

DTS提供的demo为上面项目中文件：src/main/java/com.aliyuncs.dts.demo/Main.java, 配置Demo中的下面三个参数:accessKey, accessSecret, subscribeInstanceId。其中：
accessKey、accessSecret 为订阅实例所属阿里云账号的AccessKey及

AccessSecret,AccessKey/AccessSecret的获取方式可以参考本文最后一节。subscribeInstanceID 为需要订阅数据的数据订阅实例ID,需要到DTS控制台中获取。

<input type="checkbox"/>	dts1h85bvhlq42s 1h85bvhlq42s(北京)	正常	2016-08-18 12:34:05	2016-08-18 12:20:50 2016-08-18 13:55:07	包年包月 2016-09-10 00:00:00 到期	查看订阅数据 续费更多
--------------------------	-------------------------------------	----	---------------------	--	--------------------------------	---

上面的三个参数修改完成后,在Eclipse中右击Main.java文件 -> 运行方式 -> Java应用程序,即成功运行示例demo。

使用JAR包形式的SDK

下载并载入SDK Demo

DTS提供的JAVA SDK支持的开发环境为: J2SE Development Kit (JDK) 1.5或以上版本。

下面以eclipse工程为例,简单介绍SDK的安装步骤:

1. 下载SDK包: dts-subscribe-sdk.jar ;
2. 以Eclipse为例导入jar包的过程如下: 在Eclipse上右键点击工程 -> Properties -> Java Build Path -> Libraries->Add External JARs, 选择JAR包的存放路径,将JAR添加到工程中。
3. 选择JAR包,点击“OK”按钮,经过以上步骤,你就可以在工程中使用DTS Java SDK了。

运行Demo

DTS提供了DTS SDK的demo示例,可以在DTS控制台的“查看示例代码”功能中复制这段示例代码。下面以eclipse工程为例,简单介绍运行这段SDK demo的方式。

1. 在eclipse中,在上面新建的工程中新建类MainClass。
2. 打开生成的Java文件MainClass,删除里面的代码模板。
3. 将示例代码黏贴在文件MainClass中。
4. 修改示例代码中的AccessKeyID, AccessKeySecret,订阅通道ID三个配置项。

```

1 package com.aliyun.drc.clusterclient;
2 import java.util.List;
3 import com.aliyun.drc.clusterclient.message.ClusterMessage;
4 import org.slf4j.Logger;
5 import org.slf4j.LoggerFactory;
6 public class MainClass
7 {
8     public static void main(String[] args) throws Exception {
9         //logger.info("start");
10        // 创建一个context
11        RegionContext context = new RegionContext();
12        // 运行SDK的服务器是否使用公网IP连接DTS
13        context.setUsePublicIp(true);
14        // 用户AccessKey secret
15        context.setAccessKey("xxx");
16        context.setSecret("xxx");
17        // 创建消费者
18        final ClusterClient client = new DefaultClusterClient(context);
19        // 创建订阅监听者listener
20        ClusterListener listener = new ClusterListener(){
21            @Override
22            public void notify(List<ClusterMessage> messages) throws Exception {
23                for (ClusterMessage message : messages) {
24                    // 打印订阅到的数据
25                    System.out.println(message.getRecord() + ":" + message.getRecord().getTablename() + ":"
26                    + message.getRecord().getOpt());
27                    //消费完数据后向DTS汇报ACK, 必须调用
28                    message.ackAsConsumed();
29                }
30            }
31            @Override
32            public void noException(Exception e) {
33                e.printStackTrace();
34            }
35            // 添加监听者
36            client.addConcurrentListener(listener);
37            // 设置请求的订阅通道ID
38            client.askForGUID("订阅通道ID");
39            // 启动后台线程, 注意这里不会阻塞, 主线程不能退出
40            client.start();
41        }
42    }

```

其中，AccessKeyID/AccessKeySecret为订阅实例所属阿里云账号的AccessKeyID/AccessKeySecret，订阅通道ID为需要订阅的数据订阅实例ID，需要到DTS控制台中获取。

<input type="checkbox"/>	dts1h85bvhqj42s 1h85bvhqj42s(北京)	正常	2016-08-18 12:34:05	2016-08-18 12:20:50 2016-08-18 13:55:07	包年包月 2016-09-10 00:00:00 到期	查看订阅数据	续费 更多
--------------------------	-------------------------------------	----	---------------------	--	--------------------------------	------------------------	--------------------------

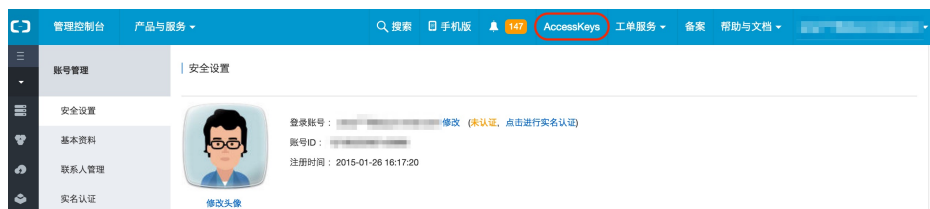
5.在Eclipse中右击demo文件 -> 运行方式 -> Java应用程序，即成功运行示例demo。

如何获取Access Key

SDK连接订阅通道时，需要通过阿里云账号的Access Key认证。所以在使用SDK之前，需要先创建Access Key。本文简单介绍如何创建并获取Access Key/Access Secret。

如果阿里云账号已经创建过Access Key，那么不需要重新创建。如果未创建，那么具体操作步骤如下：

1. 打开阿里云官网,登录您的阿里云账号。
2. 点击左上角账号，进入账号管理。
3. 点击管理控制台右上角的AccessKeys进入AccessKey管理页面。



4.创建你的Access Key。

最新版本

2017-05-15

dts-subscribe-sdk-4.6.27.12.0.jar

相较于老版订阅SDK，新版订阅SDK 进行了诸多优化，包括：

1. 修复由于服务端压力大，导致同一个SDK产生多个到服务端的连接，出现订阅数据重复的问题。
2. 修复由于服务端压力大，导致SDK提交容灾位点失败的问题。
3. 修复SDK使用client.stop()退出时，线程资源停止不干净的问题。
4. 修复SDK报keep alive error 及 152行报错java.lang.NullPointerException 的问题。

新版SDK除了提供JAR的方式，还提供了maven依赖。maven相关配置及使用方式请参考 [SDK示例代码运行简介](#)。

2017-04-11

java_sdk.jar

相较于老版SDK，新版SDK fix了诸多bug，包括Client退出异常，拉取线程跟消费线程分离等问题。

2016-04-26

java_sdk.jar