**腾讯云TDSQL PostgreSQL数据迁移**

目前腾讯TDSQL PostgreSQL数据库官方不支持在线迁移, 全量数据使用TDSQL PostgreSQL自带pg\_dump导出,增量数据库通过腾讯云DTS数据订阅到kafka, 然后自定义程序从kafka消费数据到其它数据库。

## 1.1环境配置

|  |  |  |
| --- | --- | --- |
| 资源类型 | 系统版本 | 访问地址 |
| Klustron | 1.3.2 | zettatech.tpddns.cn |
| TDSQL PostgreSQ | PostgreSQL 10.0 TBase V2 | 10.0.0.3 |
| 云主机 | CentOS7.9 | 10.0.0.11 |

## 1.2全量导出

注意事项:

* 使用TDSQL PostgreSQL自带pg\_dump工具备份表结构带有DISTRIBUTE BY SHARD .\* to GROUP tbase\_def\_group信息需要手工处理
* 使用TDSQL PostgreSQL自带pg\_dump工具备份全量数据可以直接导入昆仑数据库中
* 全量数据使用TDSQL PostgreSQL自带pg\_dum工具用csv格式导出数据全量数据
* TDSQL PostgreSQL 客户端psql 版本和昆仑客户端版本不兼容

### 1.2.1操作步骤

#### 1.2.1.1创建表写入数据

CREATE TABLE "test1" (

"ID" int NOT NULL,

"Name" varchar(20) NOT NULL,

"Age" varchar(20) NOT NULL,

"Sex" varchar(20) NOT NULL,

"CardID" decimal(30,0) NOT NULL,

"Joindate" timestamp NOT NULL,

"Region" varchar(12) NOT NULL,

"Tel" varchar(12) NOT NULL,

"Email" varchar(30) NOT NULL,

"Recommend" varchar(10) DEFAULT NULL,

"Identifier" varchar(100) DEFAULT NULL,

PRIMARY KEY ("ID", "Joindate", "Region")

);

INSERT INTO "test1" ("ID", "Name", "Age", "Sex", "CardID", "Joindate", "Region", "Tel", "Email", "Recommend", "Identifier")

SELECT

gs AS "ID",

CONCAT('test\_', gs) AS "Name",

(20 + floor(random() \* 51))::varchar AS "Age", -- 20-70之间的随机年龄

CASE WHEN random() < 0.5 THEN '男' ELSE '女' END AS "Sex", -- 随机选择性别

floor(random() \* 1e18)::decimal(30, 0) AS "CardID", -- 随机18位数字

timestamp '2020-01-01 00:00:00' + (random() \* interval '5 years') AS "Joindate", -- 2020-2030年之间的随机时间

(ARRAY['BJ', 'SH', 'GZ', 'SZ'])[floor(random() \* 4 + 1)] AS "Region", -- 随机选择区域

lpad(floor(random() \* 1e10)::text, 11, '0') AS "Tel", -- 随机生成11位数字

CONCAT(

chr(97 + floor(random() \* 26)::int), -- 生成随机小写字母

substr(md5(random()::text), 2, 8), -- 生成随机字符串的后8位

'@example.com'

) AS "Email", -- 随机生成 Email

CASE WHEN random() < 0.5 THEN 'Y' ELSE 'N' END AS "Recommend", -- 随机选择推荐

CONCAT('ID', lpad(floor(random() \* 1e10)::text, 10, '0')) AS "Identifier" -- 随机生成 ID+10位数字

FROM generate\_series(1, 100000) AS gs; -- 生成十万条数据

#### 1.2.1.2全量导入数据

PGPASSWORD=xxxxxx pg\_dump -h 10.0.0.3 -p 5432 -U dbadmin -d testdb --no-owner --no-privileges --quote-all-identifiers -t test1| sed 's/DISTRIBUTE BY SHARD .\* to GROUP tbase\_def\_group//g' | ./kunlun-server-1.3.2/bin/psql postgres://xxx:xxx@zettatech.tpddns.cn:30501/testdb



## 1.3增量导出-数据订阅

注意事项:

* 当前账户TDSQL PostgreSQL支持kafka数据订阅功能需要联系腾讯客户开通(默认未开通)
* 登录已在源端实例中创建订阅账号，需要账号权限如下： LOGIN 和 REPLICATION 权限,LOGIN 和 REPLICATION 授权请[提交工单](https://console.cloud.tencent.com/workorder/category)处理
* 用户必须拥有 pg\_catalog.pgxc\_node 表的 select 权限
* DN 节点的 wal\_level 必须是 logical
* 被订阅的表如果是全复制表（建表语句中有 distribute by replication 关键字），必须拥有主键；被订阅的表如果不是全复制表，必须拥有主键或 REPLICA IDENTITY 为 FULL
* 支持订阅的 SQL 操作:只支持 INSERT、UPDATE、DELETE
* 订阅任务启动之后新建的表,删除表重新创建插入数据均不支持订阅
* 不支持通过外网连接数据订阅的 Kafka 进行消费，只支持腾讯云内网的访问，并且订阅的数据库实例所属地域与数据消费的地域相同(需要购买云服务器)

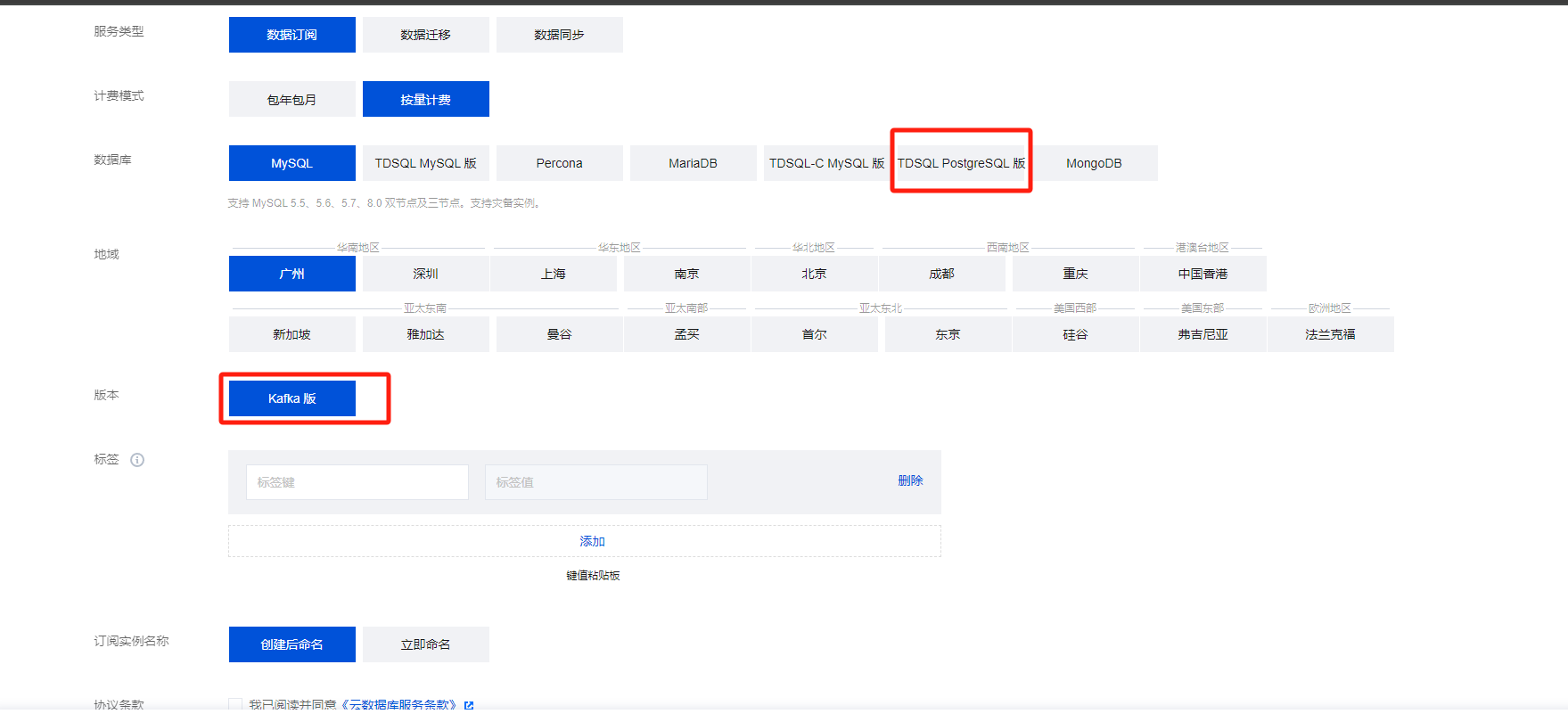
### 1.3.1操作步骤

#### 1.3.1.1创建订阅数据

* 登录 [DTS 控制台](https://console.cloud.tencent.com/dts/dss)，在左侧导航选择数据订阅，单击新建数据订阅



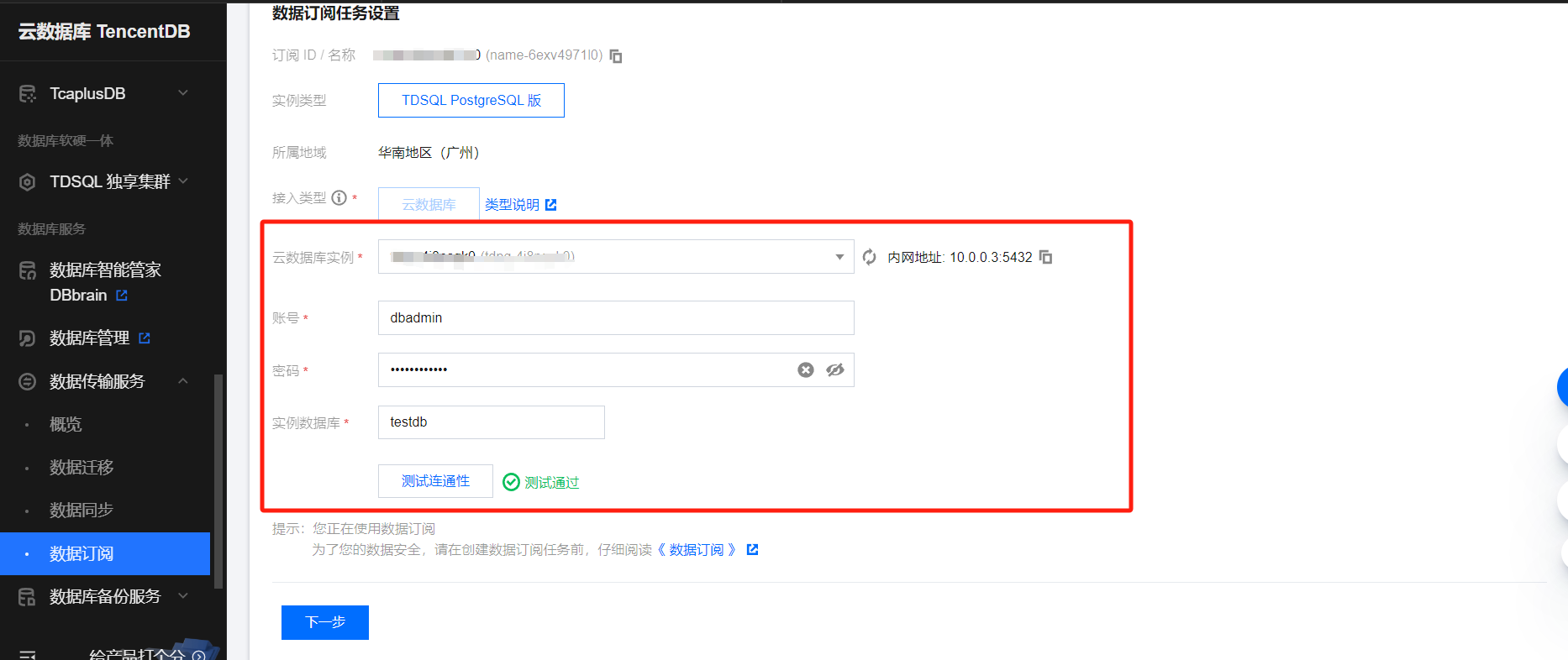
* 在新建数据订阅页，选择对应的数据库类型，单击立即购买



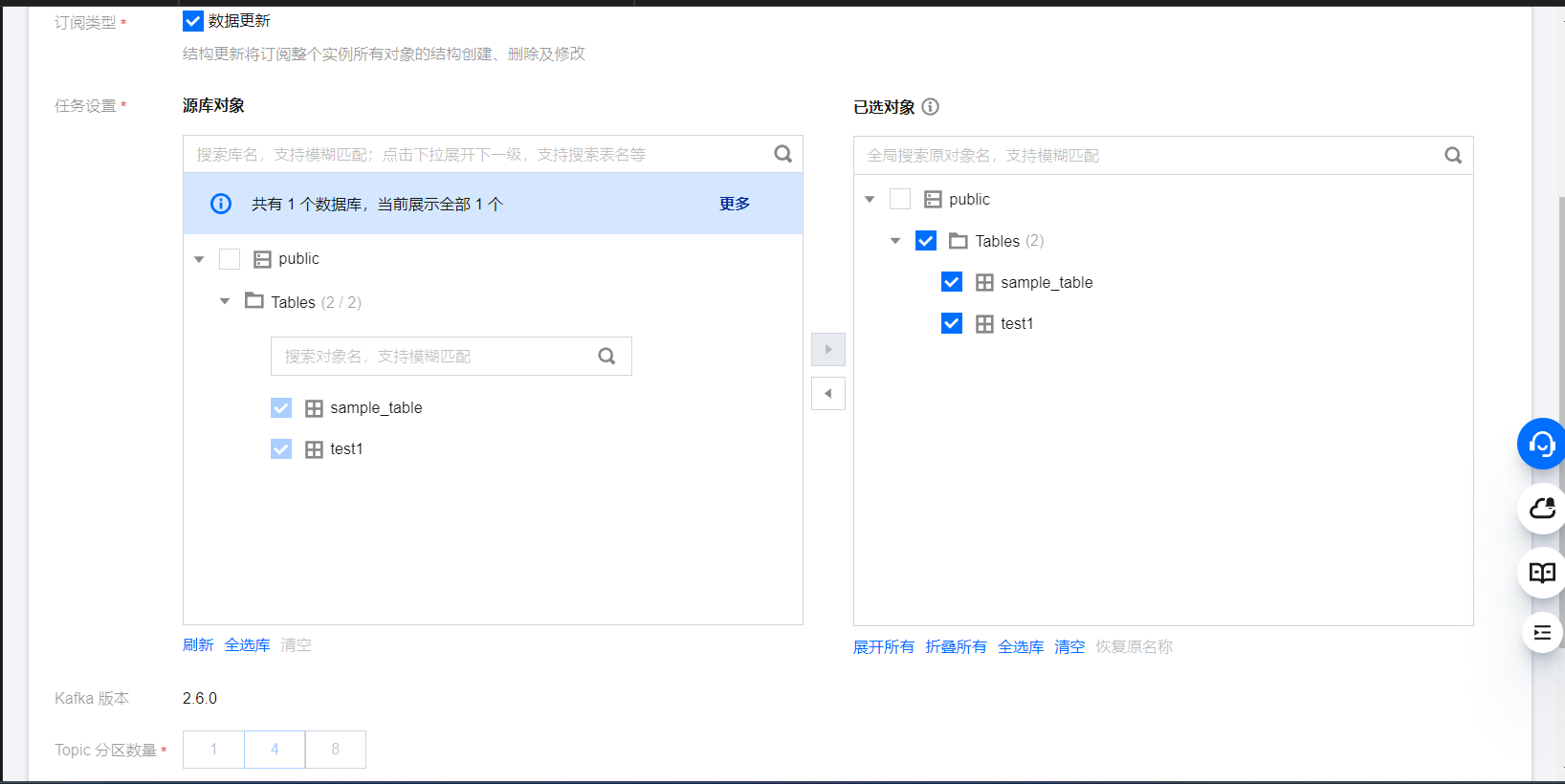
* 购买购买成功后，返回数据订阅列表，单击操作列的配置订阅对刚购买的订阅进行配置



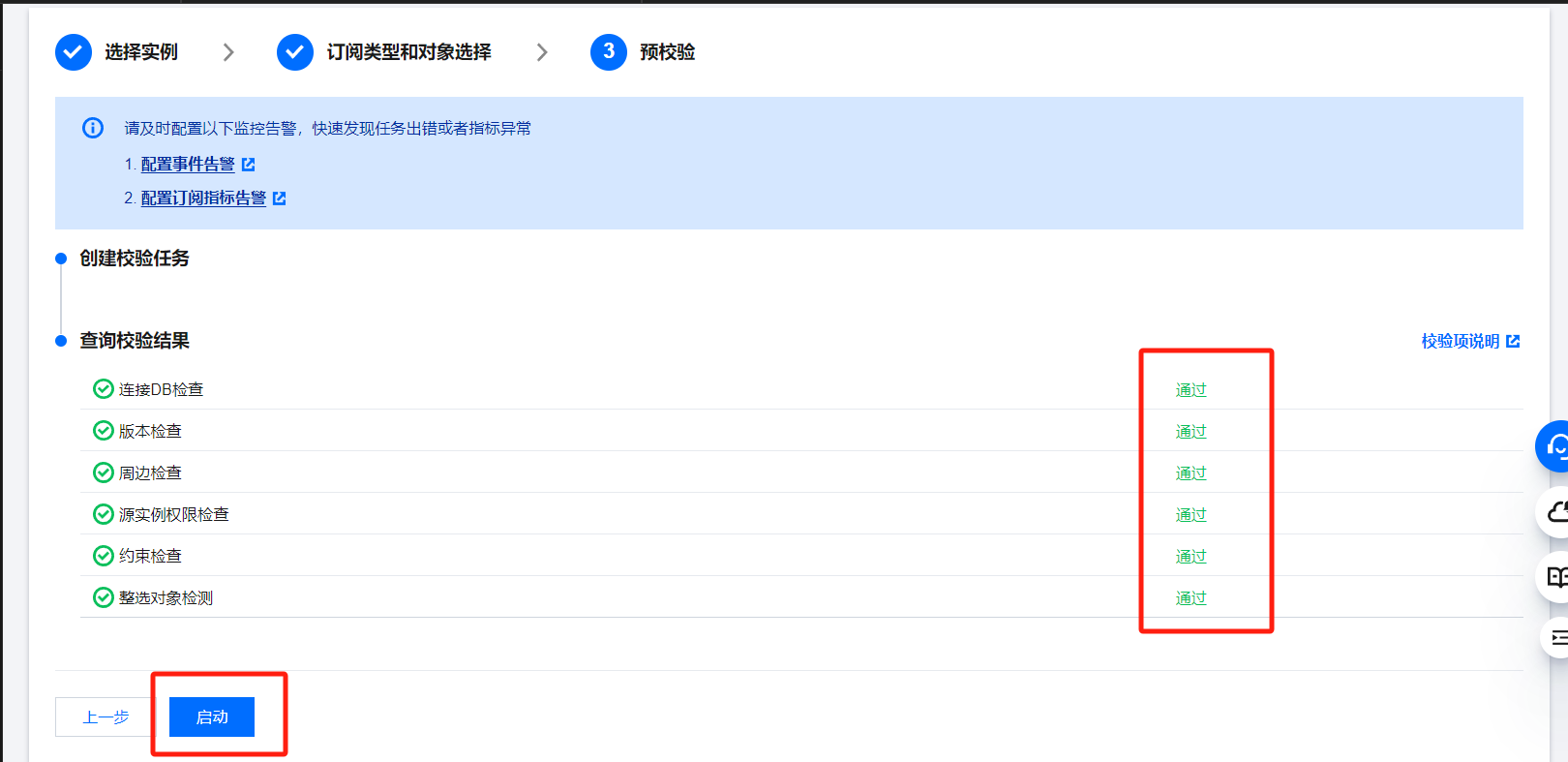
* 在配置数据库订阅页面，选择源数据库相应配置，完成后单击连通性测试，测试通过后单击下一步



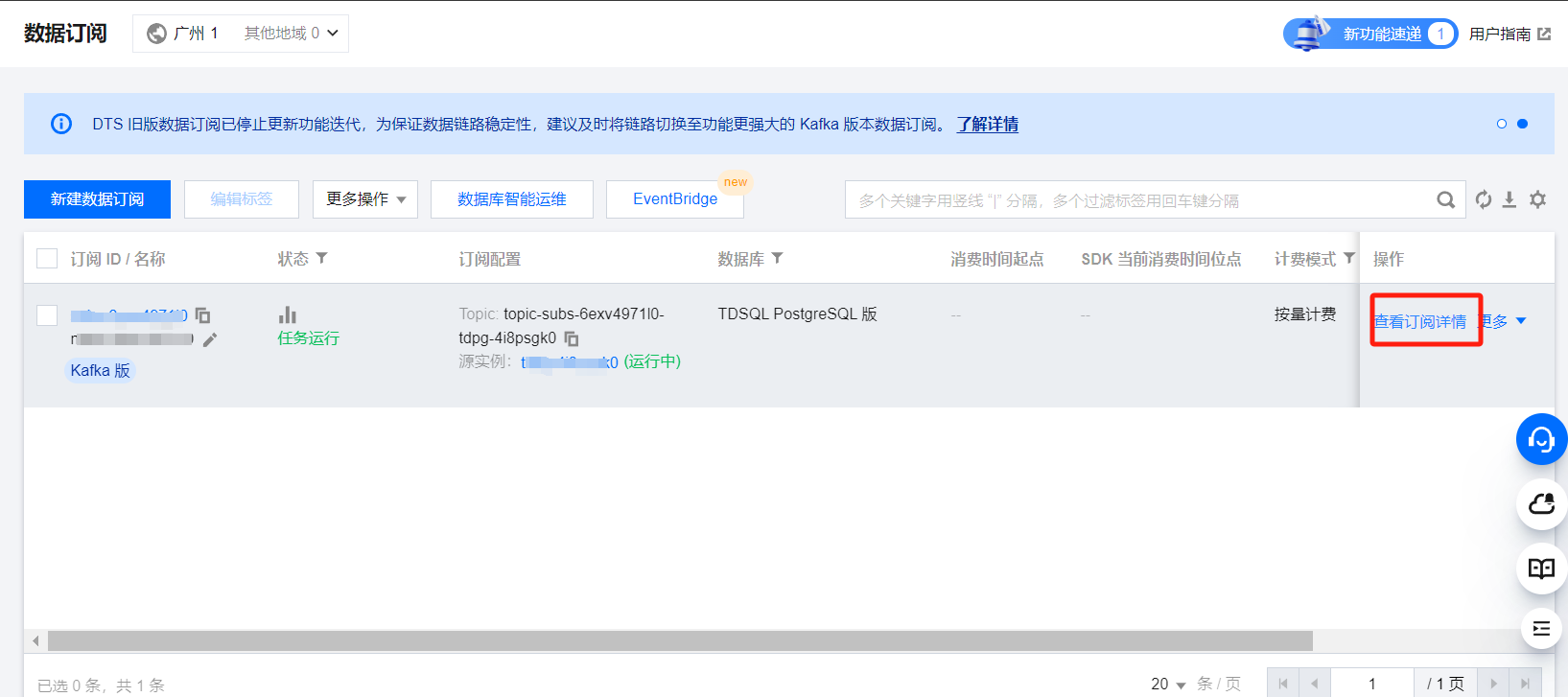
* 在订阅类型和对象选择页面，选择订阅类型，订阅类型为数据更新（订阅选择对象的数据更新，包括数据 INSERT、UPDATE、DELETE 操作）



* 购买购在预校验页面，预校验任务预计会进行校验，预校验通过后，单击启动完成数据订阅任务配置 ,如果校验失败根据提示(一般权限问题,可以找腾讯客服开通相关权限),点击启动



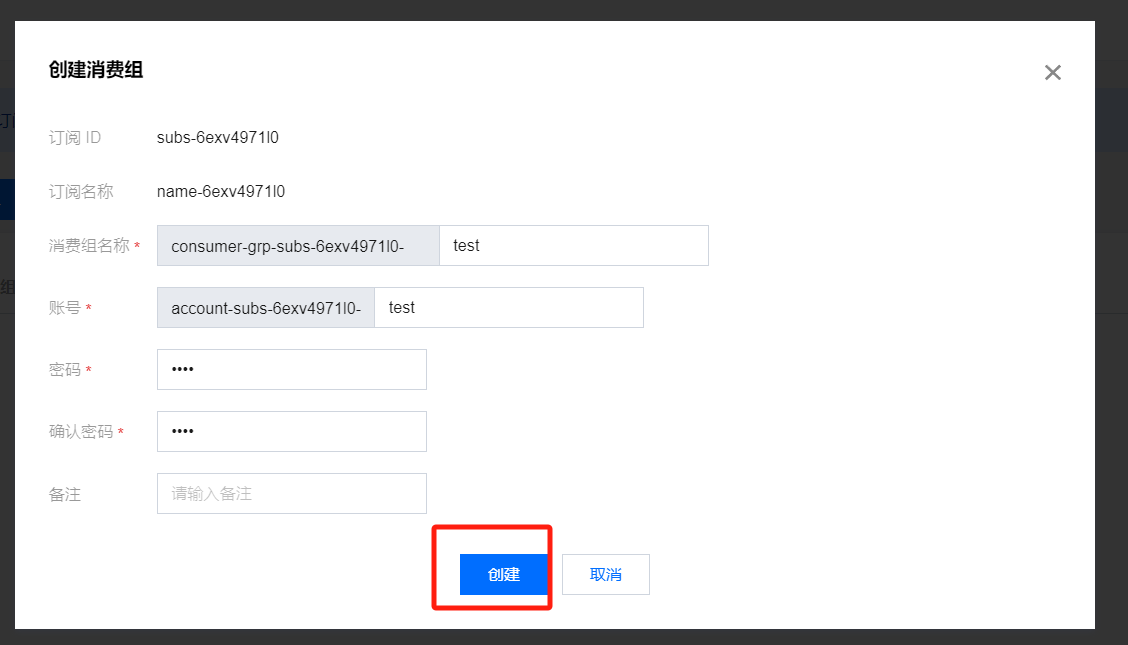
* 启动成功后，进入查看订阅详情



* 在消费管理标签下面新增消费组



* 在创建消费组页面,新增消费名称,账户和密码 ,点击创建



* 创建成功后,在控制台列表显示状态为任务运行,即可完成TDSQL PostgreSQ到kafka数据订阅创建



具体请查询官方文档说明:

https://cloud.tencent.com/document/product/571/70498

#### 1.3.1.2消费订阅数据

官方提供了Demo (语言包括:Go Java Python)并不包含消费数据的用法演示，仅对数据做了打印处理，您需要在此基础上自行编写数据处理逻辑，您也可以使用其他语言的 Kafka客户端消费并解析数据.

##### 1.3.1.2 Python3 Demo 操作步骤为例

编译运行环境：腾讯云服务器（需要与订阅实例相同地域，才能够访问到 Kafka 服务器的内网地址），安装 Python3，pip3（用于依赖包安装）

1. 下载 Python3 Demo ，然后解压该文件。

wget https://dts-subs-demos-1258344699.cos.ap-nanjing.myqcloud.com/tdsql\_postgresql\_subscribe\_python\_demo.zip

unzip tdsql\_postgresql\_subscribe\_python\_demo.zip

2 .

Cd tdsql\_postgresql\_subscribe\_python\_demo\_1.0.0

运行如下代码：python main.py --brokers=xxx --topic=xxx --group=xxx --user=xxx --password=xxx --trans2sql=1

Brokers:为数据订阅

Kafka 的内网访问地址

topic 为数据订阅任务的订阅 topic，这两个可在 订阅详情 页查看。

group、user、password 分别为消费组的名称、账号和密码，可在 消费管理 页查看。

trans2sql 表示是否转换为 SQL 语句。

##### 1.3.1.3 查询消费情况

python main.py --brokers=guangzhou-kafka-9.cdb-dts.tencentcs.com.cn:32129 \

--topic=topic-subs-6exv4971l0-tdpg-4i8psgk0 \

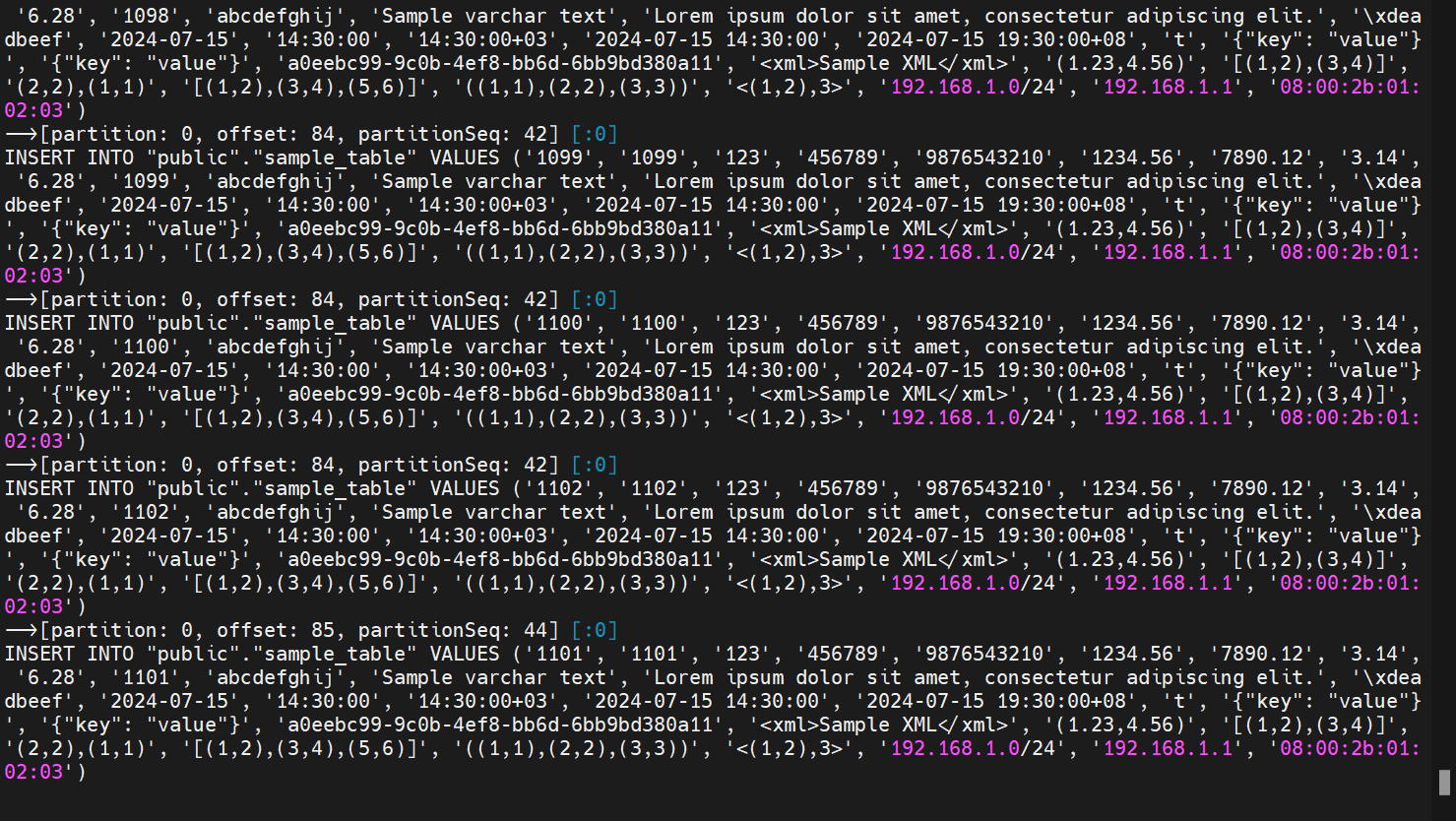
--group=consumer-grp-subs-6exv4971l0-test \

--user=account-subs-6exv4971l0-test \

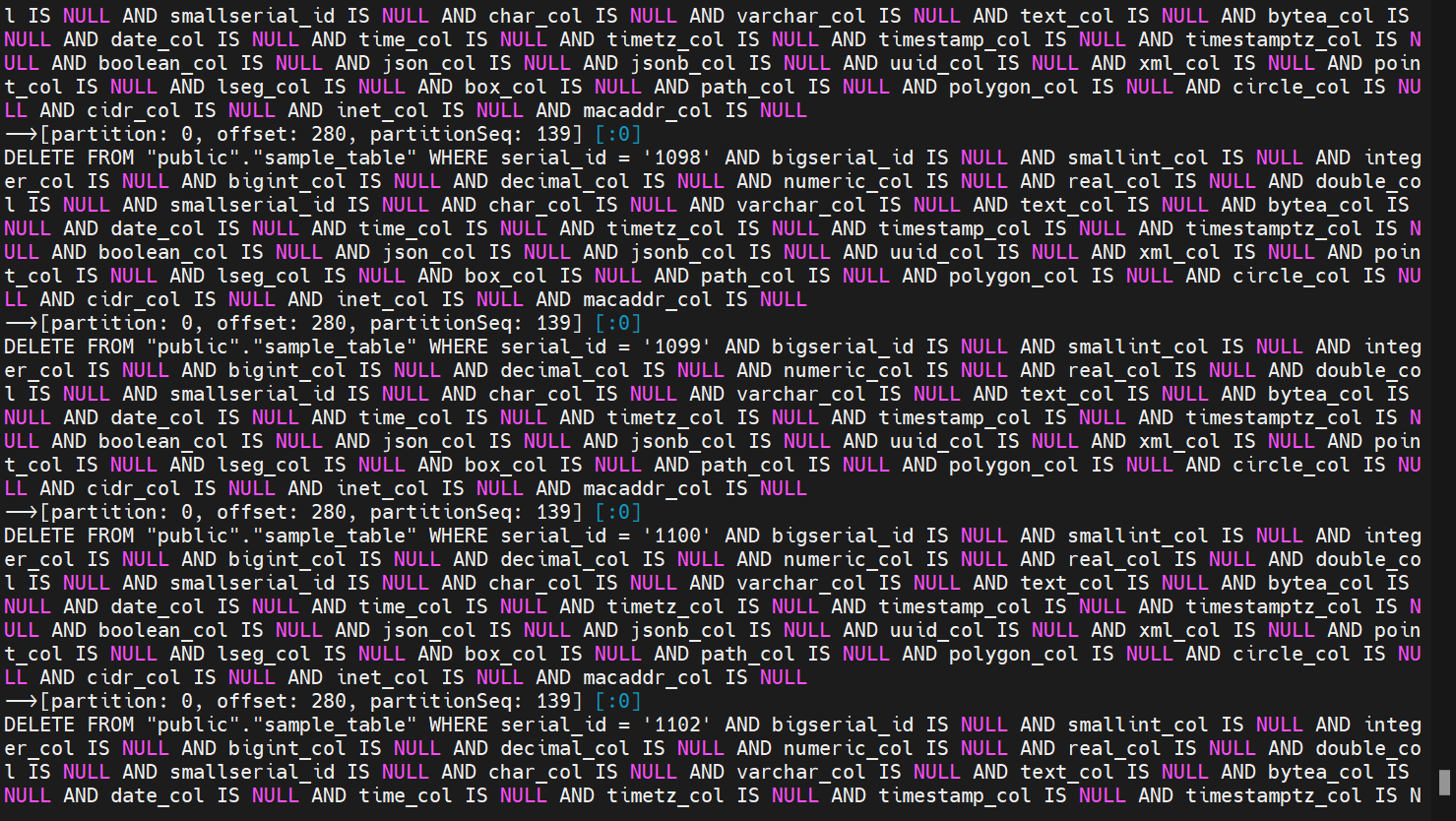
--password=test --trans2sql=1

在源TDSQL PostgreSQL库中,分别写人insert,delete,update,分别打印出对应的消费记录, 可以自定义程序把消费的数据写人其它数据库中.

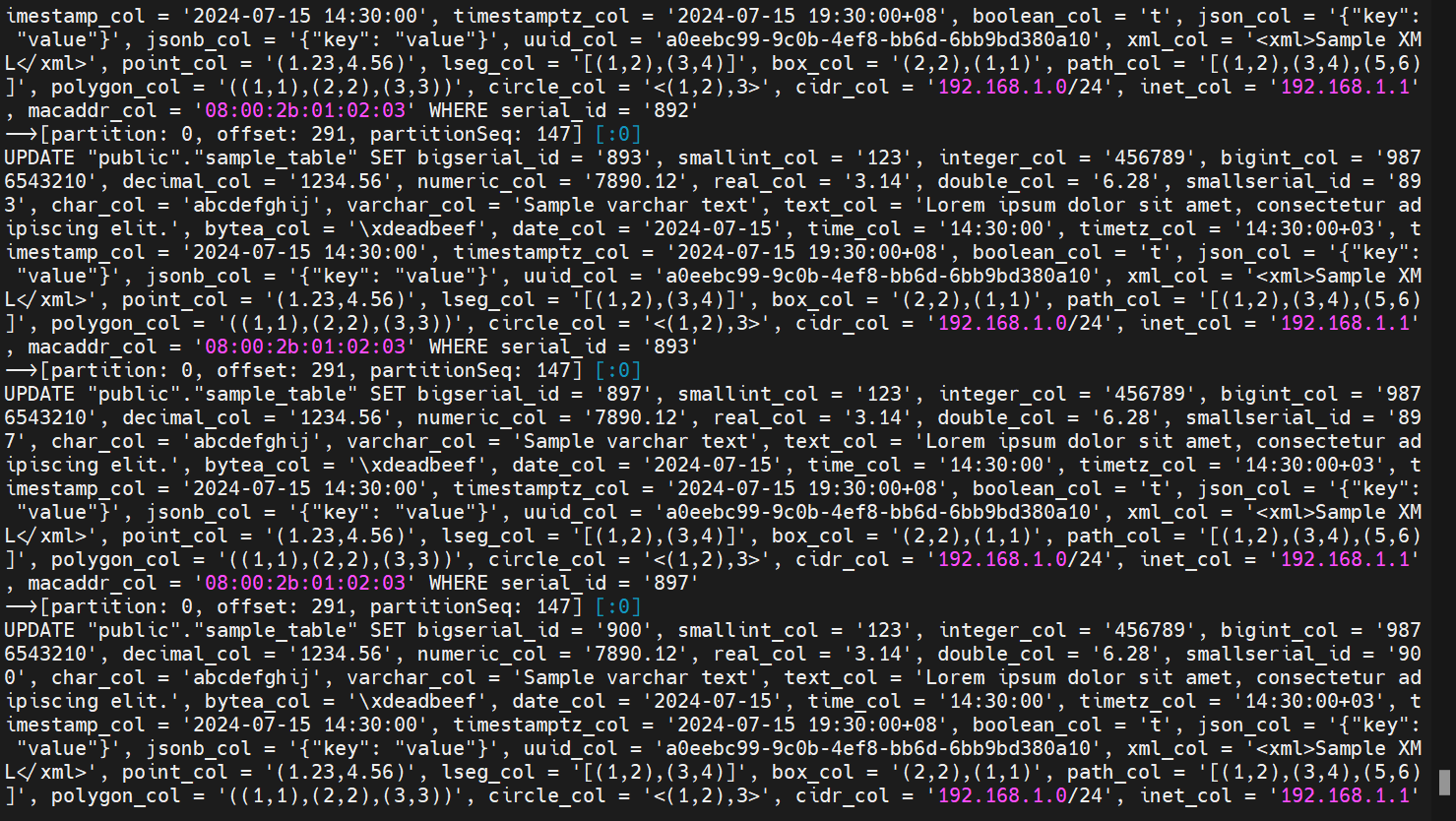
Insert:



Delete:



Update



具体请查询官方文档说明:

<https://cloud.tencent.com/document/product/571/76611>