阿里云
大数据计算服务

最佳实践

文档版本：20190917

法律声明

阿里云提醒您在阅读或使用本文档之前仔细阅读、充分理解本法律声明各条款的内容。如果您阅读或使用本文档，您的阅读或使用行为将被视为对本声明全部内容的认可。

1. 您应当通过阿里云网站或阿里云提供的其他授权通道下载、获取本文档，且仅能用于自身的合法合规的业务活动。本文档的内容视为阿里云的保密信息，您应当严格遵守保密义务；未经阿里云事先书面同意，您不得向任何第三方披露本手册内容或提供给任何第三方使用。
2. 未经阿里云事先书面许可，任何单位、公司或个人不得擅自摘抄、翻译、复制本文档内容的部分或全部，不得以任何方式或途径进行传播和宣传。
3. 由于产品版本升级、调整或其他原因，本文档内容有可能变更。阿里云保留在没有任何通知或者提示下对本文档的内容进行修改的权利，并在阿里云授权通道中不时发布更新后的用户文档。您应当实时关注用户文档的版本变更并通过阿里云授权渠道下载、获取最新版的用户文档。
4. 本文档仅作为用户使用阿里云产品及服务的参考性指引，阿里云以产品及服务的”现状“、“有缺陷”和“当前功能”的状态提供本文档。阿里云在现有技术的基础上尽最大努力提供相应的介绍及操作指引，但阿里云在此明确声明对本文档内容的准确性、完整性、适用性、可靠性等不作任何明示或暗示的保证。任何单位、公司或个人因为下载、使用或信赖本文档而发生任何差错或经济损失的，阿里云不承担任何法律责任。在任何情况下，阿里云均不对任何间接性、后果性、惩戒性、偶然性、特殊性或刑罚性的损害，包括用户使用或信赖本文档而遭受的利润损失，承担责任（即使阿里云已被告知该等损失的可能性）。
5. 阿里云网站上所有内容，包括但不限于著作、产品、图片、档案、资讯、资料、网站架构、网站画面的安排、网页设计，均由阿里云和/或其关联公司依法拥有其知识产权，包括但不限于商标权、专利权、著作权、商业秘密等。非经阿里云和/或其关联公司书面同意，任何人不得擅自使用、修改、复制、公开传播、改变、散布、发行或公开发表阿里云网站、产品程序或内容。此外，未经阿里云事先书面同意，任何人不得为了任何营销、广告、促销或其他目的使用、公布或复制阿里云的名称（包括但不限于单独为或以组合形式包含”阿里云”、Aliyun”、“万网”等阿里云和/或其关联公司品牌，上述品牌的附属标志及图案或任何类似公司名称、商号、商标、产品或服务名称、域名、图案标示、标志、标识或通过特定描述使第三方能够识别阿里云和/或其关联公司）。
6. 如若发现本文档存在任何错误，请与阿里云取得直接联系。

通用约定

格式	说明	样例
	该类警示信息将导致系统重大变更甚至故障，或者导致人身伤害等结果。	 **禁止:** 重置操作将丢失用户配置数据。
	该类警示信息可能导致系统重大变更甚至故障，或者导致人身伤害等结果。	 **警告:** 重启操作将导致业务中断，恢复业务所需时间约10分钟。
	用于补充说明、最佳实践、窍门等，不是用户必须了解的内容。	 **说明:** 您也可以通过按Ctrl + A选中全部文件。
>	多级菜单递进。	设置 > 网络 > 设置网络类型
粗体	表示按键、菜单、页面名称等UI元素。	单击 **确定**。
`courier` 字体	命令。	执行 `cd /d C:/windows` 命令，进入Windows系统文件夹。
*## *	表示参数、变量。	`bae log list --instanceid` *`Instance_ID`*
[]或者[a|b]	表示可选项，至多选择一个。	`ipconfig` *`[-all|-t]`*
{}或者{a|b}	表示必选项，至多选择一个。	`swich` *`{stand | slave}`*

目录

1 SQL

1.1 与标准SQL的主要区别及解决方法

本文将从习惯使用关系型数据库SQL用户的实践角度出发，列举用户在使用 MaxCompute SQL时比较容易遇见的问题。

MaxCompute SQL基本区别

主要区别	问题现象	解决办法
应用场景	不支持事务（没有 commit 和 rollback，不推荐使用 Insert Into）	建议代码具有等幂性支持重跑，推荐 Insert Overwrite 写入数据。
	不支持索引和主外键约束	-
	不支持自增字段和默认值	如果有默认值，请在数据写入时自行赋值。
表分区	单表支持 6 万个分区	-
	一次查询输入的分区不能超过1万，否则执行会报错；另外如果是 2 级分区且查询时只根据 2 级分区进行过滤，总的分区数大于 1 万也可能导致报错	一次查询输入的分区数不能大于 1 万
		一次查询输出的分区数不能大于2048
精度	DOUBLE类型存在精度问题	不建议在关联时候进行直接等号关联两个DOUBLE字段，推荐的做法是把两个数做减法，如果差距小于一个预设的值就认为是相同，比如 abs(a1- a2) < 0.000000001。
	目前产品上已经支持高精度的类型DECIMAL	如果有更高精度要求的，可以先把数据存为 STRING类型，然后使用 UDF来实现对应的计算。
数据类型转换	各种预期外的错误，代码维护问题。	如果有2个不同的字段类型需要做Join，建议您先把类型转好后再Join。

主要区别	问题现象	解决办法
	日期型和字符串的隐式转换	在需要传入日期型的函数里如果传入一个字符串，字符串和日期类型的转换根据yyyy-mm-dd hh:mi:ss格式进行转换。
	其他格式转换	日期函数 > TO_DATE

DDL与DML的区别及解法

主要区别	问题现象	解决办法
表结构	不能修改分区列列名，只能修改分区列对应的值。	分区和分区列的区别
	支持增加列，但是不支持删除列以及修改列的数据类型。	SQL常见问题
INSERT	语法上最直观的区别是：Insert into/overwrite 后面有个关键字Table。	-
	数据插入表的字段映射不是根据Select的别名做的，而是根据Select的字段的顺序和表里的字段的顺序。	-
UPDATE/DELETE	目前不支持Update/Delete语句。	更新和删除数据
SELECT	输入表的数量不能超过16张	-
	一个非分组列同一个Group By Key中的数据有多条，不使用聚合函数的话就没办法展示	Group by查询中的Select字段，应是Group By的分组字段，或者需要使用聚合函数。
子查询	子查询必须要有别名	建议查询不要带别名
IN/NOT IN	In/Not In,Exist/Not Exist，后面的子查询数据量不能超过 1000 条	如何使用Not In
		如果业务上已经保证了子查询返回结果的唯一性，可以考虑去掉Distinct，从而提升查询性能。
SQL返回10000条	MaxCompute限制了单独执行Select语句时返回的数据条数	其他操作
	需要查询的结果数据条数很多	如何获取所有数据
MAPJOIN	Join不支持笛卡尔积	Join必须要用on设置关联条件

主要区别	问题现象	解决办法
		如果有一些小表需要做广播表，需要用 Mapjoin Hint
		如何解决Join报错
ORDER BY	Order By后面需要配合Limit n使用	如果希望做很大的数据量的排序，甚至需要做全表排序，可以把这个N设置的很大
		MaxCompute 查询数据的排序
UNION ALL	参与UNION ALL运算的所有列的属性不同，抛异常	参与UNION ALL运算的所有列的数据类型、列个数、列名称必须完全一致
	UNION ALL查询外面需要再嵌套一层子查询	-

1.2 快速掌握SQL写法

本文通过课程实践的方式，为您介绍MaxCompute SQL，让您快速掌握SQL的写法，并清楚MaxCompute SQL和标准SQL的区别，请结合 MaxCompute SQL 基础文档 进行阅读。

数据集准备

这里选择大家比较熟悉的Emp/Dept表做为数据集。为方便大家操作，特提供相关的MaxCompute建表语句和数据文件（emp表数据文件，dept表数据文件），您可自行在MaxCompute项目上创建表并上传数据。

创建emp表的DDL语句，如下所示：

```
CREATE TABLE IF NOT EXISTS emp (
  EMPNO string ,
  ENAME string ,
  JOB string ,
  MGR bigint ,
  HIREDATE datetime ,
  SAL double ,
  COMM double ,
  DEPTNO bigint );
```

创建 dept 表的 DDL 语句，如下所示：

```
CREATE TABLE IF NOT EXISTS dept (
  DEPTNO bigint ,
  DNAME string ,
```

```
  LOC string);
```

SQL操作

初学SQL常遇到的问题点

- 使用Group by，那么Select的部分要么是分组项，要么就得是聚合函数。
- Order by后面必须加Limit n。
- Select表达式中不能用子查询，可以改写为Join。
- Join不支持笛卡尔积，以及MapJoin的用法和使用场景。
- Union all需要改成子查询的格式。
- In/Not in语句对应的子查询只能有一列，而且返回的行数不能超过1000，否则也需要改成Join。

编写SQL进行解题

题目一：列出至少有一个员工的所有部门。

为了避免数据量太大的情况下导致 常遇问题点 中的第6点，您需要使用Join 进行改写。如下所示：

```
SELECT d.*
FROM dept d
JOIN (
    SELECT DISTINCT deptno AS no
    FROM emp
) e
ON d.deptno = e.no;
```

题目二：列出薪金比SMITH多的所有员工。

MapJoin的典型场景，如下所示：

```
SELECT /*+ MapJoin(a) */ e.empno
    , e.ename
    , e.sal
FROM emp e
JOIN (
    SELECT MAX(sal) AS sal
    FROM `emp`
    WHERE `ENAME` = 'SMITH'
) a
ON e.sal > a.sal;
```

题目三：列出所有员工的姓名及其直接上级的姓名。

非等值连接，如下所示：

```
SELECT a.ename
    , b.ename
FROM emp a
```

```
LEFT OUTER JOIN emp b
ON b.empno = a.mgr;
```

题目四：列出最低薪金大于1500的各种工作。

Having 的用法，如下所示：

```
SELECT emp.`JOB`
    , MIN(emp.sal) AS sal
FROM `emp`
GROUP BY emp.`JOB`
HAVING MIN(emp.sal) > 1500;
```

题目五：列出在每个部门工作的员工数量、平均工资和平均服务期限。

时间处理上有很多好用的内建函数，如下所示：

```
SELECT COUNT(empno) AS cnt_emp
    , ROUND(AVG(sal), 2) AS avg_sal
    , ROUND(AVG(datediff(getdate(), hiredate, 'dd')), 2) AS avg_hire
FROM `emp`
GROUP BY `DEPTNO`;
```

题目六： 列出每个部门的薪水前3名的人员的姓名以及他们的名次（Top n的需求非常常见）。

SQL 语句如下所示：

```
SELECT *
FROM (
  SELECT deptno
    , ename
    , sal
    , ROW_NUMBER() OVER (PARTITION BY deptno ORDER BY sal DESC) AS
nums
  FROM emp
) emp1
WHERE emp1.nums < 4;
```

题目七： 用一个SQL写出每个部门的人数、CLERK（办事员）的人数占该部门总人数占比。

SQL语句如下所示：

```
SELECT deptno
    , COUNT(empno) AS cnt
    , ROUND(SUM(CASE
      WHEN job = 'CLERK' THEN 1
      ELSE 0
    END) / COUNT(empno), 2) AS rate
FROM `EMP`
```

```
GROUP BY deptno;
```

1.3 修改不兼容SQL实战

MaxCompute 开发团队近期已经完成了 MaxCompute2.0 灰度升级。新升级的版本完全拥抱开源生态，支持更多的语言功能，带来更快的运行速度，同时新版本会执行更严格的语法检测，以致于一些在老编译器下正常执行的不严谨的语法 case 在 MaxCompute2.0 下会报错。

为了使 MaxCompute2.0 灰度升级更加平滑，MaxCompute 框架支持回退机制，如果 MaxCompute2.0 任务失败，会回退到 MaxCompute1.0 执行。回退本身会增加任务 E2E 时延。鼓励大家提交作业之前，手动关闭回退`set odps.sql.planner.mode=lot;`以避免 MaxCompute 框架回退策略修改对大家造成影响。

MaxCompute 团队会根据线上回退情况，邮件或者钉钉等通知有问题任务的 Owner，请大家尽快完成 SQL 任务修改，否则会导致任务失败。烦请大家仔细 check 以下报错情况，进行自检，以免通知遗漏造成任务失败。

下面列举常见的一些会报错的语法：

group.by.with.star

SELECT * …GROUP BY… 的问题。

旧版 MaxCompute 中，即使 * 中覆盖的列不在 group by key 内，也支持 `select * from group by key` 的语法，但 MaxCompute2.0 和 Hive 兼容，并不允许这种写法，除非 group by 列表是所有源表中的列。示例如下：

场景一：group by key 不包含所有列

错误写法：

```
SELECT * FROM t GROUP BY key;
```

报错信息：

```
FAILED: ODPS-0130071:[1,8] Semantic analysis exception - column
reference t.value should appear in GROUP BY key
```

正确改法：

```
SELECT DISTINCT key FROM t;
```

场景二：group by key 包含所有列

不推荐写法：

```
SELECT * FROM t GROUP BY key, value; -- t has columns key and value
```

虽然 MaxCompute2.0 不会报错，但推荐改为：

```
SELECT DISTINCT key, value FROM t;
```

bad.escape

错误的 escape 序列问题。

按照 MaxCompute 文档的规定，在 string literal 中应该用反斜线加三位8进制数字表示从 0 到 127 的 ASCII 字符，例如：使用 \001， \002 表示 0，1 等。但目前\01，\0001 也被当作 \001 处理了。

这种行为会给新用户带来困扰，比如需要用 “\0001” 表示 “\000” + “1”，便没有办法实现。同时对于从其他系统迁移而来的用户而言，会导致正确性错误。

说明：

`\000`后面在加数字，如`\0001 - \0009`或`\00001`的写法可能会返回错误。

MaxCompute2.0 会解决此问题，需要 script 作者将错误的序列进行修改，示例如下：

错误写法：

```
SELECT split(key, "\01"), value like "\0001" FROM t;
```

报错信息：

```
FAILED: ODPS-0130161:[1,19] Parse exception - unexpected escape
sequence: 01
ODPS-0130161:[1,38] Parse exception - unexpected escape sequence: 0001
```

正确改法：

```
SELECT split(key, "\001"), value like "\001" FROM t;
```

column.repeated.in.creation

create table 时列名重复的问题。

如果 create table 时列名重复，MaxCompute2.0 将会报错，示例如下：

错误写法：

```
CREATE TABLE t (a BIGINT, b BIGINT, a BIGINT);
```

报错信息：

```
FAILED: ODPS-0130071:[1,37] Semantic analysis exception - column
repeated in creation: a
```

正确改法：

```
CREATE TABLE t (a BIGINT, b BIGINT);
```

string.join.double

写 JOIN 条件时，等号的左右两边分别是 String 和 Double 类型。

出现上述情况，旧版 MaxCompute 会把两边都转成 Bigint，但会导致严重的精度损失问题，例如：1.1 = “1” 在连接条件中会被认为是相等的。但 MaxCompute2.0 会与 Hive 兼容转为 Double。

不推荐写法：

```
SELECT * FROM t1 JOIN t2 ON t1.double_value = t2.string_value;
```

warning 信息：

```
WARNING:[1,48]  implicit conversion from STRING to DOUBLE, potential
data loss, use CAST function to suppress
```

推荐改法：

```
select * from t1 join t2 on t.double_value = cast(t2.string_value as
double);
```

除以上改法外，也可使用用户期望的其他转换方式。

window.ref.prev.window.alias

Window Function 引用同级 Select List 中的其他 Window Function Alias 的问题。

示例如下：

如果 rn 在 t1 中不存在，错误写法如下：

```

SELECT row_number() OVER (PARTITION BY c1 ORDER BY c1) rn,
row_number() OVER (PARTITION by c1 ORDER BY rn) rn2
```

```
FROM t1;
```

报错信息：

```
FAILED: ODPS-0130071:[2,45] Semantic analysis exception - column rn
cannot be resolved
```

正确改法：

```
SELECT row_number() OVER (PARTITION BY c1 ORDER BY rn) rn2
FROM
(
SELECT c1, row_number() OVER (PARTITION BY c1 ORDER BY c1) rn
FROM t1
) tmp;
```

select.invalid.token.after.star

select * 后面接 alias 的问题。

Select 列表里面允许用户使用 * 代表选择某张表的全部列，但 * 后面不允许加 alias（即使 * 展开之后只有一列也不允许），新一代编译器将会对类似语法进行报错，示例如下：

错误写法：

```
select * as alias from dual;
```

报错信息：

```
FAILED: ODPS-0130161:[1,10] Parse exception - invalid token 'as'
```

正确改法：

```
select * from dual;
```

agg.having.ref.prev.agg.alias

有 Having 的情况下，Select List 可以出现前面 Aggregate Function Alias 的问题。示例如下：

错误写法：

```
SELECT count(c1) cnt,
sum(c1) / cnt avg
FROM t1
GROUP BY c2
HAVING cnt > 1;
```

报错信息：

```
```

```
FAILED: ODPS-0130071:[2,11] Semantic analysis exception - column cnt
cannot be resolved
ODPS-0130071:[2,11] Semantic analysis exception - column reference cnt
 should appear in GROUP BY key
```

其中 s、cnt 在源表 t1 中都不存在，但因为有 HAVING，旧版 MaxCompute 并未报错，MaxCompute2.0 则会提示 column cannot be resolve，并报错。

正确改法：

```
SELECT cnt, s, s/cnt avg
FROM
(
SELECT count(c1) cnt,
sum(c1) s
FROM t1
GROUP BY c2
HAVING count(c1) > 1
) tmp;
```

order.by.no.limit

ORDER BY 后没有 LIMIT 语句的问题。

MaxCompute 默认 order by 后需要增加 limit 限制数量，因为 order by 是全量排序，没有 limit 时执行性能较低。示例如下：

错误写法：

```
select * from (select *
from (select cast(login_user_cnt as int) as uv, '3' as shuzi
from test_login_cnt where type = 'device' and type_name = 'mobile') v
order by v.uv desc) v
order by v.shuzi limit 20;
```

报错信息：

```
FAILED: ODPS-0130071:[4,1] Semantic analysis exception - ORDER BY must
 be used with a LIMIT clause
```

正确改法：

在子查询 order by v.uv desc 中增加 limit。

另外，MaxCompute1.0 对于 view 的检查不够严格。比如在一个不需要检查 LIMIT 的 Projec（odps.sql.validate.orderby.limit=false）中，创建了一个 View：

```
CREATE VIEW dual_view AS SELECT id FROM dual ORDER BY id;
```

若访问此 View：

```
SELECT * FROM dual_view;
```

MaxCompute1.0 不会报错，而 MaxCompute2.0 会报如下错误信息：

```
FAILED: ODPS-0130071:[1,15] Semantic analysis exception - while
resolving view xdj.xdj_view_limit - ORDER BY must be used with a LIMIT
 clause
```

generated.column.name.multi.window

使用自动生成的 alias 的问题。

旧版 MaxCompute 会为 Select 语句中的每个表达式自动生成一个 alias，这个 alias 会最后显示在 console 上。但是，它并不承诺这个 alias 的生成规则，也不承诺这个 alias 的生成规则会保持不变，所以不建议用户使用自动生成的 alias。

MaxCompute2.0 会对使用自动生成 alias 的情况给予警告，由于牵涉面较广，暂时无法直接给予禁止。

对于某些情况，MaxCompute 的不同版本间生成的 alias 规则存在已知的变动，但因为已有一些线上作业依赖于此类 alias，这些查询在 MaxCompute 版本升级或者回滚时可能会失败，存在此问题的用户，请修改您的查询，对于感兴趣的列，显式地指定列的别名。示例如下：

不推荐写法：

```
SELECT _c0 FROM (SELECT count(*) FROM dual) t;
```

建议改法：

```
SELECT c FROM (SELECT count(*) c FROM dual) t;
```

non.boolean.filter

使用了非 boolean 过滤条件的问题。

MaxCompute 不允许布尔类型与其他类型之间的隐式转换，但旧版 MaxCompute 会允许用户在某些情况下使用 Bigint 作为过滤条件。MaxCompute2.0 将不再允许，如果您的脚本中存在这样的过滤条件，请及时修改。示例如下：

错误写法：

```
select id, count(*) from dual group by id having id;
```

报错信息：

```
FAILED: ODPS-0130071:[1,50] Semantic analysis exception - expect a
BOOLEAN expression
```

正确改法：

```
select id, count(*) from dual group by id having id <> 0;
```

post.select.ambiguous

在 order by、 cluster by、 distribute by、sort by 等语句中，引用了名字冲突的列的问题。

旧版 MaxCompute 中，系统会默认选取 Select 列表中的后一列作为操作对象，MaxCompute2.0 将会进行报错，请及时修改。示例如下：

错误写法：

```
select a, b as a from t order by a limit 10;
```

报错信息：

```
FAILED: ODPS-0130071:[1,34] Semantic analysis exception - a is
ambiguous, can be both t.a or null.a
```

正确改法：

```
select a as c, b as a from t order by a limit 10;
```

本次推送修改会包括名字虽然冲突但语义一样的情况，虽然不会出现歧义，但是考虑到这种情况容易导致错误，作为一个警告，希望用户进行修改。

duplicated.partition.column

在 query 中指定了同名的 partition 的问题。

旧版 MaxCompute 在用户指定同名 partition key 时并未报错， 而是后一个的值直接覆盖了前一个，容易产生混乱。MaxCompute2.0 将会对此情况进行报错，示例如下：

错误写法一：

```
insert overwrite table partition (ds = '1', ds = '2') select ... ;
```

实际上，在运行时 ds = ‘1’ 被忽略。

正确改法：

```
insert overwrite table partition (ds = '2') select ... ;
```

错误写法二：

```
create table t (a bigint, ds string) partitioned by (ds string);
```

正确改法：

```
create table t (a bigint) partitioned by (ds string);
```

order.by.col.ambiguous

Select list 中 alias 重复，之后的 Order by 子句引用到重复的 alias 的问题。

错误写法：

```
SELECT id, id
FROM dual
ORDER BY id;
```

正确改法：

```
SELECT id, id id2
FROM dual
ORDER BY id;
```

需要去掉重复的 alias，Order by 子句再进行引用。

in.subquery.without.result

colx in subquery 没有返回任何结果，则 colx 在源表中不存在的问题。

错误写法：

```
SELECT * FROM dual
WHERE not_exist_col IN (SELECT id FROM dual LIMIT 0);
```

报错信息：

```
FAILED: ODPS-0130071:[2,7] Semantic analysis exception - column
not_exist_col cannot be resolved
```

ctas.if.not.exists

目标表语法错误问题。

如果目标表已经存在，旧版 MaxCompute 不会做任何语法检查，MaxCompute2.0 则会做正常的语法检查，这种情况会出现很多错误信息，示例如下：

错误写法：

```
CREATE TABLE IF NOT EXISTS dual
AS
SELECT * FROM not_exist_table;
```

报错信息：

```
FAILED: ODPS-0130131:[1,50] Table not found - table meta_dev.
not_exist_table cannot be resolved
```

worker.restart.instance.timeout

旧版 MaxCompute UDF 每输出一条记录，便会触发一次对分布式文件系统的写操作，同时会向 Fuxi 发送心跳，如果 UDF 10 分钟没有输出任何结果，会得到如下错误提示：

```
FAILED: ODPS-0123144: Fuxi job failed - WorkerRestart errCode:252,
errMsg:kInstanceMonitorTimeout, usually caused by bad udf performance.
```

MaxCompute2.0 的 Runtime 框架支持向量化，一次会处理某一列的多行来提升执行效率。但向量化可能导致原来不会报错的语句（2 条记录的输出时间间隔不超过 10 分钟），因为一次处理多行，没有及时向 Fuxi 发送心跳而导致 timeout。

遇到这个错误，建议首先检查 UDF 是否有性能问题，每条记录需要数秒的处理时间。如果无法优化 UDF 性能，可以尝试手动设置 batch row 大小来绕开（默认为1024）：

```
set odps.sql.executionengine.batch.rowcount=16;
```

divide.nan.or.overflow

旧版 MaxCompute 不会做除法常量折叠的问题。

比如如下语句，旧版 MaxCompute 对应的物理执行计划如下：

```
EXPLAIN
SELECT IF(FALSE, 0/0, 1.0)
FROM dual;
In Task M1_Stg1:
    Data source: meta_dev.dual
    TS: alias: dual
      SEL: If(False, Divide(UDFToDouble(0), UDFToDouble(0)), 1.0)
        FS: output: None
```

由此可以看出，IF 和 Divide 函数仍然被保留，运行时因为 IF 第一个参数为 false，第二个参数 Divide 的表达式不需要求值，所以不会出现除零异常。

而 MaxCompute2.0 则支持除法常量折叠，所以会报错。如下所示：

错误写法：

```
SELECT IF(FALSE, 0/0, 1.0)
FROM dual;
```

报错信息：

```
FAILED: ODPS-0130071:[1,19] Semantic analysis exception - encounter
 runtime exception while evaluating function /, detailed message:
DIVIDE func result NaN, two params are 0.000000 and 0.000000
```

除了上述的 nan，还可能遇到 overflow 错误，比如：

错误写法：

```
SELECT IF(FALSE, 1/0, 1.0)
FROM dual;
```

报错信息：

```
FAILED: ODPS-0130071:[1,19] Semantic analysis exception - encounter
 runtime exception while evaluating function /, detailed message:
DIVIDE func result overflow, two params are 1.000000 and 0.000000
```

正确改法：

建议去掉 /0 的用法，换成合法常量。

CASE WHEN 常量折叠也有类似问题，比如：CASE WHEN TRUE THEN 0 ELSE 0/0，MaxCompute2.0 常量折叠时所有子表达式都会求值，导致除0错误。

CASE WHEN 可能涉及更复杂的优化场景，比如：

```
SELECT CASE WHEN key = 0 THEN 0 ELSE 1/key END
FROM (
SELECT 0 AS key FROM src
UNION ALL
SELECT key FROM src) r;
```

优化器会将除法下推到子查询中，转换类似于：

```
M (
SELECT CASE WHEN 0 = 0 THEN 0 ELSE 1/0 END c1 FROM src
UNION ALL
```

```
SELECT CASE WHEN key = 0 THEN 0 ELSE 1/key END c1 FROM src) r;
```

报错信息：

```
FAILED: ODPS-0130071:[0,0] Semantic analysis exception - physical plan
 generation failed: java.lang.ArithmeticException: DIVIDE func result
overflow, two params are 1.000000 and 0.000000
```

其中 UNION ALL 第一个子句常量折叠报错，建议将 SQL 中的 CASE WHEN 挪到子查询中，并去掉无用的 CASE WHEN 和去掉/0用法：

```
SELECT c1 END
FROM (
SELECT 0 c1 END FROM src
UNION ALL
SELECT CASE WHEN key = 0 THEN 0 ELSE 1/key END) r;
```

small.table.exceeds.mem.limit

旧版 MaxCompute 支持 Multi-way Join 优化，多个 Join 如果有相同 Join Key，会合并到一个 Fuxi Task 中执行，比如下面例子中的 J4_1_2_3_Stg1：

```
EXPLAIN
SELECT t1.*
FROM t1 JOIN t2 ON t1.c1 = t2.c1
JOIN t3 ON t1.c1 = t3.c1;
```

旧版 MaxCompute 物理执行计划：

```
In Job job0:
root Tasks: M1_Stg1, M2_Stg1, M3_Stg1
J4_1_2_3_Stg1 depends on: M1_Stg1, M2_Stg1, M3_Stg1

In Task M1_Stg1:
    Data source: meta_dev.t1

In Task M2_Stg1:
    Data source: meta_dev.t2

In Task M3_Stg1:
    Data source: meta_dev.t3

In Task J4_1_2_3_Stg1:
    JOIN: t1 INNER JOIN unknown INNER JOIN unknown
        SEL: t1._col0, t1._col1, t1._col2
            FS: output: None
```

如果增加 MapJoin hint，旧版 MaxCompute 物理执行计划不会改变。也就是说对于旧版 MaxCompute 优先应用 Multi-way Join 优化，并且可以忽略用户指定 MapJoin hint。

```
EXPLAIN
SELECT /*+mapjoin(t1)*/ t1.*
```

```
FROM t1 JOIN t2 ON t1.c1 = t2.c1
JOIN t3 ON t1.c1 = t3.c1;
```

旧版 MaxCompute 物理执行计划同上。

MaxCompute2.0 Optimizer 会优先使用用户指定的 MapJoin hint，对于上述例子，如果 t1 比较大的话，会遇到类似错误：

```
FAILED: ODPS-0010000:System internal error - SQL Runtime Internal
 Error: Hash Join Cursor HashJoin_REL… small table exceeds, memory
 limit(MB) 640, fixed memory used …, variable memory used …
```

对于这种情况，如果 MapJoin 不是期望行为，建议去掉 MapJoin hint。

sigkill.oom

同 small.table.exceeds.mem.limit，如果用户指定了 MapJoin hint，并且用户本身所指定的小表比较大。在旧版 MaxCompute 下有可能被优化成 Multi-way Join 从而成功。但在 MaxCompute2.0 下，用户可能通过设定 odps.sql.mapjoin.memory.max 来避免小表超限的错误，但每个 MaxCompute worker 有固定的内存限制，如果小表本身过大，则 MaxCompute worker 会由于内存超限而被杀掉，错误类似于：

```
Fuxi job failed - WorkerRestart errCode:9,errMsg:SigKill(OOM), usually
 caused by OOM(outof memory).
```

这里建议您去掉 MapJoin hint，使用 Multi-way Join。

wm_concat.first.argument.const

聚合函数 中关于 WM_CONCAT 的说明，一直要求 WM_CONCAT 第一个参数为常量，旧版 MaxCompute 检查不严格，比如源表没有数据，就算 WM_CONCAT 第一个参数为 ColumnReference，也不会报错。

```
函数声明:
string wm_concat(string separator, string str)
参数说明:
separator: String类型常量，分隔符。其他类型或非常量将引发异常。
```

MaxCompute2.0，会在 plan 阶段便检查参数的合法性，假如 WM_CONCAT 的第一个参数不是常量，会立即报错。示例如下：

错误写法：

```
SELECT wm_concat(value, ',') FROM src GROUP BY value;
```

报错信息：

```
FAILED: ODPS-0130071:[0,0] Semantic analysis
 exception - physical plan generation failed:
```

```
 com.aliyun.odps.lot.cbo.validator.AggregateCallValidator
$AggregateCallValidationException: Invalid argument type - The first
 argument of WM_CONCAT must be constant string.
```

pt.implicit.convertion.failed

srcpt 是一个分区表，并有两个分区：

```
CREATE TABLE srcpt(key STRING, value STRING) PARTITIONED BY (pt STRING
);
ALTER TABLE srcpt ADD PARTITION (pt='pt1');
ALTER TABLE srcpt ADD PARTITION (pt='pt2');
```

对于以上 SQL，String 类型 pt 列 IN INT 类型常量，都会转为 Double 进行比较。即使 Project 设置了 odps.sql.udf.strict.mode=true，旧版 MaxCompute 不会报错，所有 pt 都会过滤掉，而 MaxCompute2.0 会直接报错。示例如下：

错误写法：

```
SELECT key FROM srcpt WHERE pt IN (1, 2);
```

报错信息：

```
FAILED: ODPS-0130071:[0,0] Semantic analysis exception - physical
plan generation failed: java.lang.NumberFormatException: ODPS-0123091
:Illegal type cast - In function cast, value 'pt1' cannot be casted
from String to Double.
```

建议避免 String 分区列和 INT 类型常量比较，将 INT 类型常量改成 String 类型。

having.use.select.alias

SQL 规范定义 Group by + Having 子句是 Select 子句之前阶段，所以 Having 中不应该使用 Select 子句生成的 Column alias，示例如下：

错误写法：

```
SELECT id id2 FROM DUAL GROUP BY id HAVING id2 > 0;
```

报错信息：

```
FAILED: ODPS-0130071:[1,44] Semantic analysis exception - column id2
 cannot be resolvedODPS-0130071:[1,44] Semantic analysis exception -
column reference id2 should appear in GROUP BY key
```

其中 id2 为 Select 子句中新生成的 Column alias，不应该在 Having 子句中使用。

dynamic.pt.to.static

MaxCompute2.0 动态分区某些情况会被优化器转换成静态分区处理，示例如下：

```
INSERT OVERWRITE TABLE srcpt PARTITION(pt) SELECT id, 'pt1' FROM dual;
```

会被转化成

```
INSERT OVERWRITE TABLE srcpt PARTITION(pt='pt1') SELECT id FROM dual;
```

如果用户指定的分区值不合法，比如错误的使用了’${bizdate}’，MaxCompute2.0 语法检查阶段便会报错。详情请参见#unique_9 。

错误写法：

```
INSERT OVERWRITE TABLE srcpt PARTITION(pt) SELECT id, '${bizdate}'
FROM dual LIMIT 0;
```

报错信息：

```
FAILED: ODPS-0130071:[1,24] Semantic analysis exception - wrong
columns count 2 in data source, requires 3 columns (includes dynamic
partitions if any)
```

旧版 MaxCompute 因为 LIMIT 0，SQL 最终没有输出任何数据，动态分区不会创建，所以最终不报错。

lot.not.in.subquery

In subquery 中 null 值的处理问题。

在标准 SQL 的 IN 运算中，如果后面的值列表中出现 null，则返回值不会出现 false，只可能是 null 或者 true。如 1 in (null, 1, 2, 3) 为 true，而 1 in (null, 2, 3) 为 null，null in (null, 1, 2, 3) 为 null。同理 not in 操作在列表中有 null 的情况下，只会返回 false 或者 null，不会出现 true。

MaxCompute2.0 会用标准的行为进行处理，收到此提醒的用户请注意检查您的查询，IN 操作中的子查询中是否会出现空值，出现空值时行为是否与您预期相符，如果不符合预期请做相应的修改。示例如下：

```
select * from t where c not in (select accepted from c_list);
```

若 accepted 中不会出现 null 值，则此问题可忽略。若出现空值，则 c not in (select accepted from c_list) 原先返回 true，则新版本返回 null。

正确改法：

```
select * from t where c not in (select accepted from c_list where
accepted is not null)
```

1.4 导出SQL的运行结果

本文将通过示例，为您介绍几种下载MaxCompute SQL计算结果的方法。

说明：

本文中所有SDK部分仅举例介绍Java的例子。

您可以通过以下几种方法导出SQL的运行结果：

- 如果数据比较少，可以直接用SQL Task得到全部的查询结果。
- 如果只是想导出某个表或者分区，可以用 Tunnel直接导出数据。
- 如果SQL比较复杂，需要Tunnel和SQL相互配合才行。
- DataWorks 可以方便地帮您运行SQL，同步数据，并有定时调度，配置任务依赖的功能。
- 开源工具 DataX 可帮助您方便地把 MaxCompute 中的数据导出到目标数据源，详情请参见 DataX 概述。

SQLTask方式导出

SQLTask 是 SDK直接调用MaxCompute SQL的接口，能很方便地运行SQL并获得其返回结果。

从文档可以看到，`SQLTask.getResult(i)`返回的是一个 List，可以循环迭代这个List，获得完整的SQL计算返回结果。不过该方法有一个缺陷，详情请参见#unique_13中的`SetProject READ_TABLE_MAX_ROW`功能。

目前Select语句返回给客户端的数据条数最大可以调整到 1万。也即如果在客户端上（包括SQLTask）直接进行Select操作，相当于查询结果上最后加上了Limit N参数（如果使用`CREATE TABLE XX AS SELECT`或者用`INSERT INTO/OVERWRITE TABLE`把结果固化到具体的表里则没有影响）。

说明：

`SQLTask.getResult(i)`用于导出Select查询结果，不适用于导出`show tables;`等其他MaxCompute命令操作结果。

Tunnel 方式导出

如果您需要导出的查询结果是某张表的全部内容（或者是具体的某个分区的全部内容），可以通过Tunnel来实现，详情请参见 命令行工具 和基于SDK编写的 Tunnel SDK。

此处提供一个Tunnel命令行导出数据的简单示例，Tunnel SDK的编写适用于Tunnel命令行无法支持的场景，详情请参见 批量数据通道概述。

```
tunnel d wc_out c:\wc_out.dat;
2016-12-16 19:32:08 - new session: 201612161932082d3c9b0a012f68e7
total lines: 3
2016-12-16 19:32:08 - file [0]: [0, 3), c:\wc_out.dat
downloading 3 records into 1 file
2016-12-16 19:32:08 - file [0] start
2016-12-16 19:32:08 - file [0] OK. total: 21 bytes
download OK
```

SQLTask + Tunnel方式导出

从前面SQL Task方式导出的介绍可以看到，SQL Task不能处理超过1万条记录，而 Tunnel可以，两者可以互补。所以可以基于两者实现数据的导出。

代码实现的示例如下：

```
private static final String accessId = "userAccessId";
    private static final String accessKey = "userAccessKey";
    private static final String endPoint = "http://service.odps.aliyun
.com/api";
    private static final String project = "userProject";
    private static final String sql = "userSQL";
    private static final String table = "Tmp_" + UUID.randomUUID().
toString().replace("-", "_");//其实也就是随便找了个随机字符串作为临时表的名字
    private static final Odps odps = getOdps();

    public static void main(String[] args) {
        System.out.println(table);
        runSql();
        tunnel();
    }

    /*
     * 把SQLTask的结果下载过来
     * */
    private static void tunnel() {
        TableTunnel tunnel = new TableTunnel(odps);
        try {
            DownloadSession downloadSession = tunnel.createDown
loadSession(
                    project, table);
            System.out.println("Session Status is : "
                    + downloadSession.getStatus().toString());
            long count = downloadSession.getRecordCount();
            System.out.println("RecordCount is: " + count);
            RecordReader recordReader = downloadSession.openRecord
Reader(0,
                    count);
            Record record;
            while ((record = recordReader.read()) != null) {
                consumeRecord(record, downloadSession.getSchema());
            }
            recordReader.close();
        } catch (TunnelException e) {
            e.printStackTrace();
        } catch (IOException e1) {
            e1.printStackTrace();
```

```
        }
    }

    /*
     * 保存这条数据
     * 数据量少的话直接打印后拷贝走也是一种取巧的方法。实际场景可以用Java.io写到
本地文件，或者写到远端数据等各种目标保存起来。
     * */
    private static void consumeRecord(Record record, TableSchema
schema) {
        System.out.println(record.getString("username")+","+record.
getBigint("cnt"));
    }

    /*
     * 运行SQL，把查询结果保存成临时表，方便后面用Tunnel下载
     * 这里保存数据的lifecycle为1天，所以哪怕删除步骤出了问题，也不会太浪费存储空
间
     * */
    private static void runSql() {
        Instance i;
        StringBuilder sb = new StringBuilder("Create Table ").append(
table)
                .append(" lifecycle 1 as ").append(sql);
        try {
            System.out.println(sb.toString());
            i = SQLTask.run(getOdps(), sb.toString());
            i.waitForSuccess();

        } catch (OdpsException e) {
            e.printStackTrace();
        }
    }

    /*
     * 初始化MaxCompute(原ODPS)的连接信息
     * */
    private static Odps getOdps() {
        Account account = new AliyunAccount(accessId, accessKey);
        Odps odps = new Odps(account);
        odps.setEndpoint(endPoint);
        odps.setDefaultProject(project);
        return odps;
    }
```

大数据开发套件的数据同步方式导出

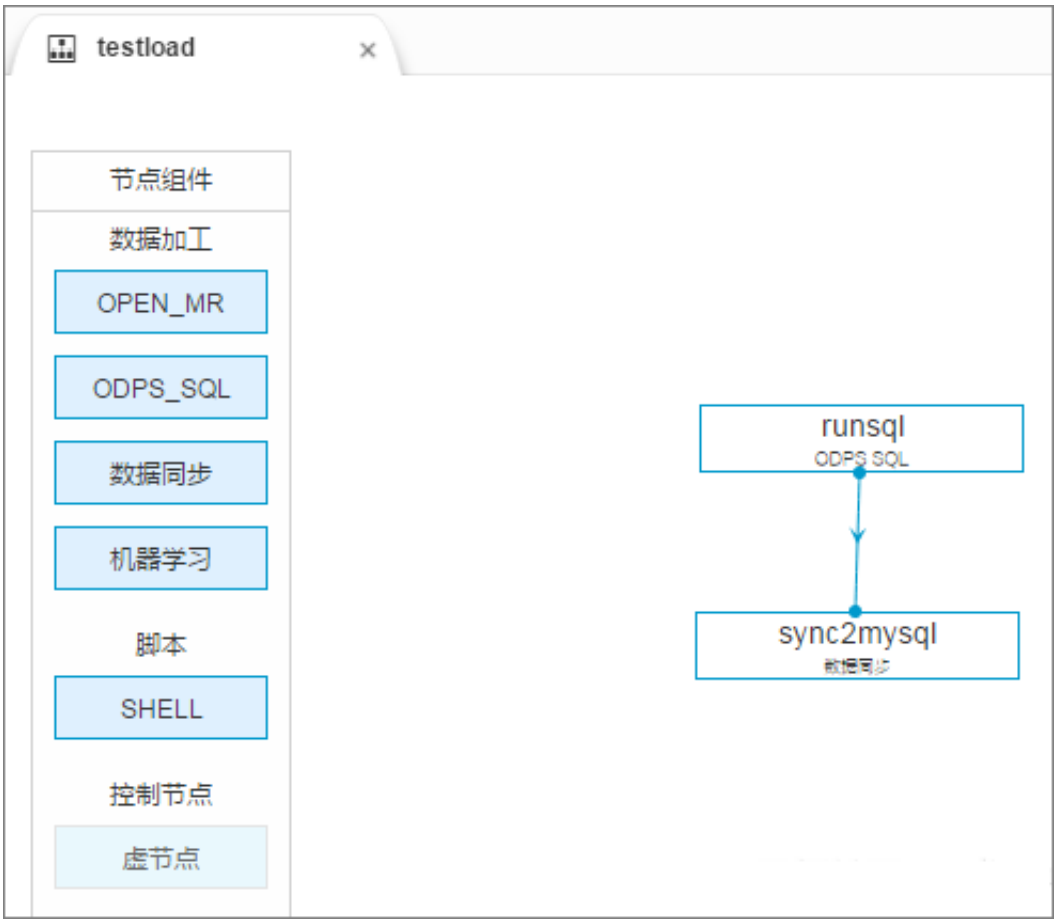
前面介绍的方式解决了数据下载后保存的问题，但是没解决数据的生成以及两个步骤之间的调度依赖的问题。

数加DataWorks 可以运行SQL、配置数据同步任务，还可以设置自动 周期性运行 和 多任务之间依赖，彻底解决了前面的烦恼。

接下来将用一个简单示例，为您介绍如何通过大数据开发套件运行SQL并配置数据同步任务，以完成数据生成和导出需求。

操作步骤

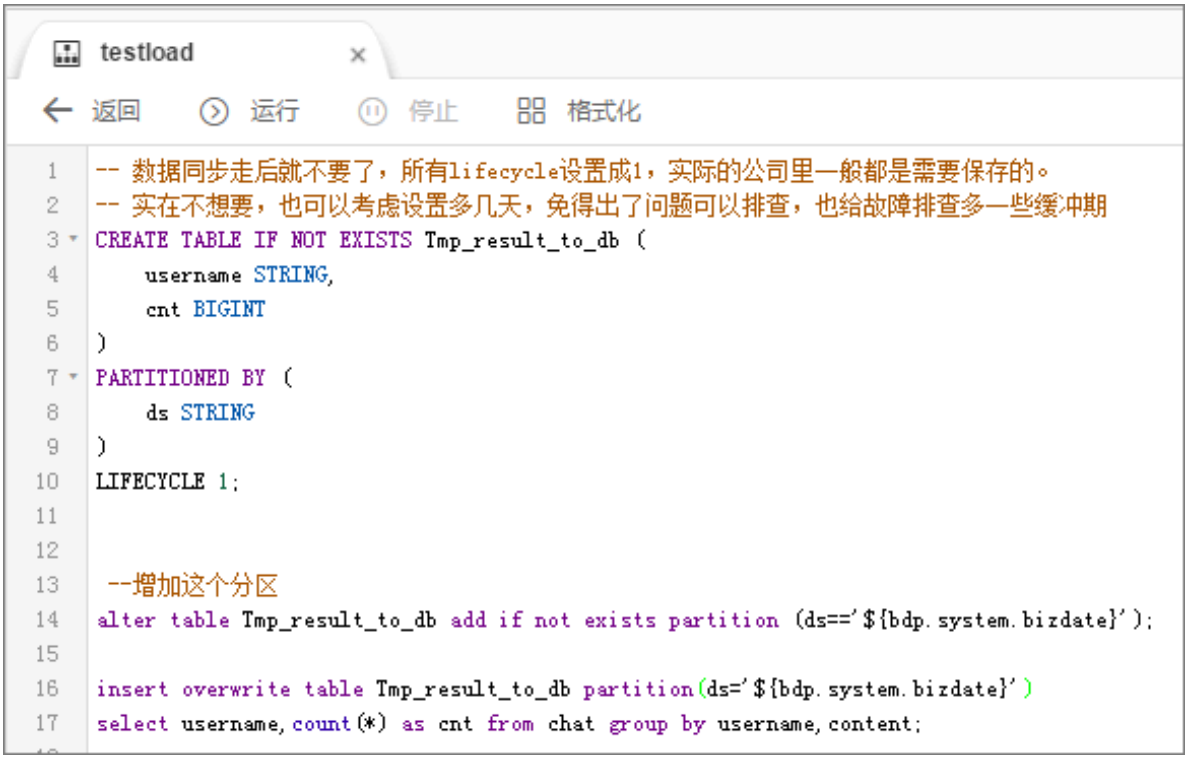
1. 创建一个工作流，工作流里创建一个SQL节点和一个数据同步节点，并将两个节点连线配置成依赖关系，SQL节点作为数据产出的节点，数据同步节点作为数据导出节点。



2. 配置SQL节点。

说明:

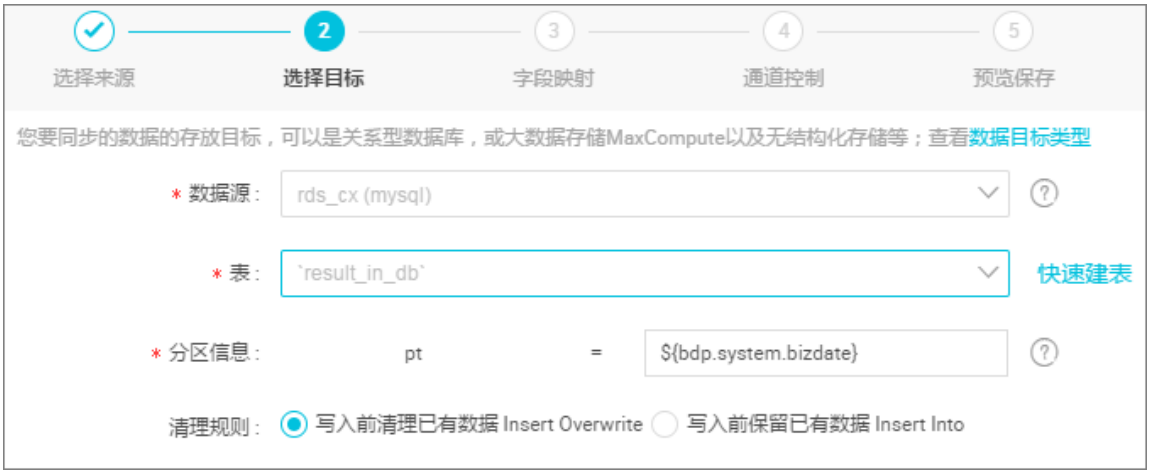
SQL这里的创建表要先执行一次再去配置同步（否则表都没有，同步任务没办法配置）。



3. 配置数据同步任务。

a. 选择来源。

b. 选择目标。



c. 字段映射。

d. 通道控制。

e. 预览保存。

4. 工作流调度配置完成后（可以直接使用默认配置），保存并提交工作流，然后单击 测试运行。

查看数据同步的运行日志，如下所示：

```
2016-12-17 23:43:46.394 [job-15598025] INFO JobContainer -
任务启动时刻 : 2016-12-17 23:43:34
任务结束时刻 : 2016-12-17 23:43:46
任务总计耗时 : 11s
任务平均流量 : 31.36KB/s
记录写入速度 : 1668rec/s
读出记录总数 : 16689
```

```
读写失败总数 ： 0
```

5. 输入SQL语句查看数据同步的结果，如下图所示：

1.5 分区剪裁合理性评估

背景及目的

MaxCompute的 分区表 是指在创建表时指定分区空间，即指定表内的某几个字段作为分区列。使用数据时，如果指定了需要访问的分区名称，则只会读取相应的分区，避免全表扫描，提高处理效率，降低费用。

分区剪裁是指对分区列指定过滤条件，使得SQL执行时只用读取表的部分分区数据，避免全表扫描引起的数据错误及资源浪费。但是分区失效的情况会经常发生，本文将通过示例为您介绍一些常见问题的解决方案。

问题示例

测试表test_part_cut的分区。

```
odps@ show partitions>show partitions test_part_cut;

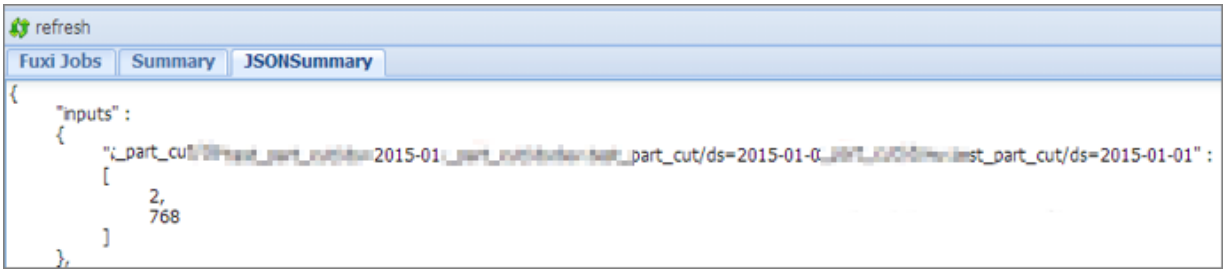
ds=2015-01-01
ds=2015-01-02
ds=2015-01-03
```

执行以下SQL代码。

```
select count(*)
from test_part_cut
where ds= bi_week_dim('20150102');
--其中为bi_week_dim自定义函数:返回格式为（年,第几周）。
--如果是正常日期，判断日期是所传入参数中年份所属周，以周四为一周的起始日期，如果碰到20140101因为属于周三所以算在2013年最后一周返回2013,52。而20150101则返回是2015,1。
```

```
--如果是类似20151231是周四又恰逢与20160101在同一周，则返回2016,1。
```

bi_week_dim('20150102')的返回结果是2015,1，不符合表test_part_cut的分区值，通常会认为上面的SQL不会读任何分区，而实际情况却是该SQL读了表test_part_cut的所有分区，LogView如下。



从上图可见，该SQL在执行时读取了表test_part_cut的所有分区。

由上述示例可见，分区剪裁使用尽管简单，但也容易出错。因此，本文将从以下两方面进行介绍：

- 判断SQL中分区剪裁是否生效。
- 了解常见的导致分区剪裁失效的场景。

判断分区剪裁是否生效

通过explain命令查看SQL的执行计划，用于发现SQL中的分区剪裁是否生效。

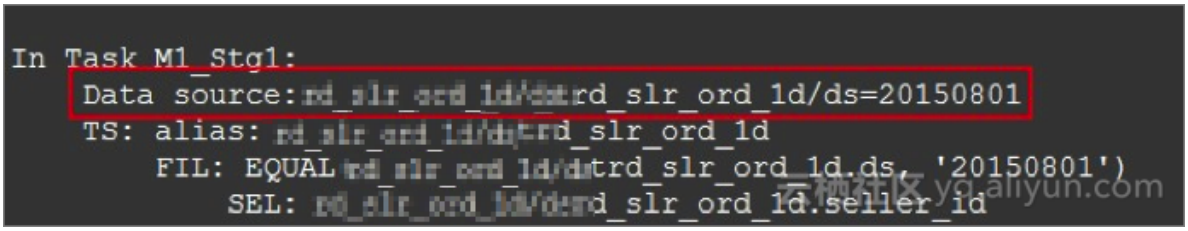
- 分区剪裁未生效

```
explain
select seller_id
from xxxxx_trd_slr_ord_1d
where ds=rand();
```

由上图可见，SQL读取了表xxxxx_trd_slr_ord_1d的1344个分区，即该表的所有分区。

- 分区剪裁生效

```
explain
select seller_id
from xxxxx_trd_slr_ord_1d
where ds='20150801';
```



由上图可见，SQL只读取了表xxxxx_trd_slr_ord_1d的20150801的分区。

分区剪裁失效的场景分析

分区剪裁在使用自定义函数或者部分系统函数的时候会失效，在join关联时的where条件中也有可能会失效。下面针对这两种场景分别举例说明：

- 自定义函数导致分区剪裁失效

  当分区剪裁的条件中使用了用户自定义函数，则分区剪裁会失效，即使是使用系统函数也可能会导致分区剪裁失效。所以，对于分区值的限定，如果使用了非常规函数需要用explain命令通过查看执行计划，确定分区剪裁是否已经生效。

```
explain
select ...
from xxxxx_base2_brd_ind_cw
where ds = concat(SPLIT_PART(bi_week_dim(' ${bdp.system.bizdate}'),
 ',', 1), SPLIT_PART(bi_week_dim(' ${bdp.system.bizdate}'), ',', 2))
```

  由上图可见，SQL因为分区剪裁使用了用户自定义的函数导致全表扫描。

  说明：
  UDF已支持分区裁剪，详情请参见#unique_17/unique_17_Connect_42_section_lwx_cv2_ggb文中对应的说明。

- 使用join时分区剪裁失效

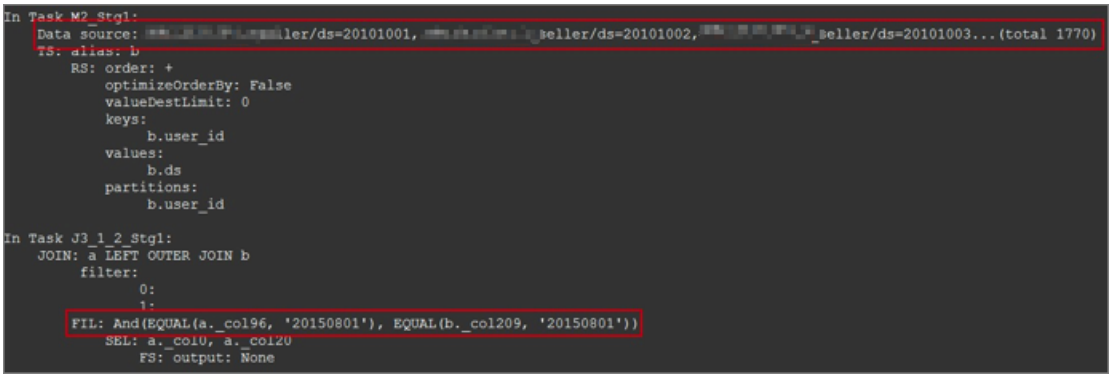
  在SQL语句中，使用join进行关联时，如果分区剪裁条件放在where中，则分区剪裁会生效，如果放在on条件中，从表的分区剪裁会生效，主表则不会生效。下面针对三种join具体说明：

  - LEFT OUTER JOIN

    ■ 分区剪裁条件均放在on中

```
set odps.sql.allow.fullscan=true;
explain
select a.seller_id
    ,a.pay_ord_pbt_1d_001
from xxxxx_trd_slr_ord_1d a
left outer join
     xxxxx_seller b
on a.seller_id=b.user_id
and a.ds='20150801'
```

```
and b.ds='20150801';
```



由上图可见，左表进行了全表扫描，只有右表的分区裁剪有效果。

■ **分区剪裁条件均放在`where`中**

```
set odps.sql.allow.fullscan=true;
explain
select a.seller_id
    ,a.pay_ord_pbt_1d_001
from xxxxx_trd_slr_ord_1d a
left outer join
    xxxxx_seller b
on a.seller_id=b.user_id
where a.ds='20150801'
```

```
and b.ds='20150801';
```

```
In Task M2_Stg1:
    Data source: seller/ds=seller/ds=20150801
    TS: alias: b
        FIL: EQUAL(b.ds, '20150801')
            RS: order: +
                optimizeOrderBy: False
                valueDestLimit: 0
                keys:
                     b.user_id
                values:
                partitions:
                     b.user_id

In Task J3_1_2_Stg1:
    JOIN: a LEFT OUTER JOIN unknown
          filter:
                0:
                1:
        SEL: a._col0, a._col20
            FS: output: None

In Task M1_Stg1:
    Data source: seller/ds=trd_slr_ord_1d/ds=20150801
    TS: alias: a
```

由上图可见，两张表的分区裁剪都有效果。

- RIGHT OUTER JOIN

  与`left outer join`类似，分区剪裁条件如果放在`on`中则只有`right outer join`的左表生效，如果放在`where`中，则两张表都会生效。

- FULL OUTER JOIN

  分区剪裁条件只有都放在`where`中才会生效，放在`on`中则都不会生效。

影响及思考

- 分区剪裁如果失效会影响比较大，且用户不容易发现。因此，分区剪裁失效最好在代码提交的时候发现比较合适。
- 对于用户自定义函数不能用于分区剪裁的问题，需要再深入思考解决方法。

1.6 分组取出每组数据的前N条

本文将为您介绍如何对数据进行分组，取出每组数据的前 N 条数据。

示例数据

目前的数据，如下表所示：

empno	ename	job	sal
7369	SMITH	CLERK	800.0
7876	SMITH	CLERK	1100.0
7900	JAMES	CLERK	950.0
7934	MILLER	CLERK	1300.0
7499	ALLEN	SALESMAN	1600.0
7654	MARTIN	SALESMAN	1250.0
7844	TURNER	SALESMAN	1500.0
7521	WARD	SALESMAN	1250.0

实现方法

您可以通过以下两种方法实现：

- 取出每条数据的行号，再用 where 语句进行过滤。

```
SELECT * FROM (
  SELECT empno
  , ename
  , sal
  , job
  , ROW_NUMBER() OVER (PARTITION BY job ORDER BY sal) AS rn
  FROM emp
) tmp
WHERE rn < 10;
```

- 使用 UDTF 实现 Split 函数。

  详情请参见 此文 中最后的示例。这个例子可以更迅速地判断当前的序号，如果是已经超过预定的条数（比如 10 条），便不做处理了，从而提高计算效率。

1.7 SQL实现多行数据转一条

本文将为您介绍，如何使用 SQL 实现多条数据压缩为一条。

场景示例

以下表数据为例：

class	gender	name
1	M	LiLei
1	F	HanMM
1	M	Jim

class	gender	name
2	F	Kate
2	M	Peter

场景一

根据需求，常见场景如下：

class	names
1	LiLei,HanMM,Jim
2	Kate,Peter

类似这样使用某个分隔符做字符串拼接，可以使用如下语句：

```
SELECT class, wm_concat(',', name) FROM students GROUP BY class;
```

场景二

另外一种常见需求，如下所示：

class	cnt_m	cnt_f
1	2	1
2	1	1

类似这样转多列的需求，可以使用如下语句：

```
SELECT
class
,SUM(CASE WHEN gender = 'M' THEN 1 ELSE 0 END) AS cnt_m
,SUM(CASE WHEN gender = 'F' THEN 1 ELSE 0 END) AS cnt_f
FROM students
GROUP BY class;
```

1.8 MaxCompute(ODPS) SQL中的JOIN ON条件

MaxCompute(ODPS) SQL中，很常用的一个操作就是关联(Join)。

概述

目前MaxCompute提供了以下几种Join类型：

类型	含义
Inner Join	输出符合关联条件的数据

类型	含义
Left Join	输出左表的所有记录，对于右表符合关联的数据，输出右表，没有符合的，右表补null。
Right Join	输出右表的所有记录，对于左表符合关联的数据，输出左表，没有符合的，左表补null。
Full Join	输出左表和右表的所有记录，对于没有关联上的数据，未关联的另一侧补null。
Left Semi Join	对于左表中的一条数据，如果右表存在符合关联条件的行，则输出左表。
Left Anti Join	对于左表中的一条数据，如果对于右表所有的行，不存在符合关联条件的数据，则输出左表。

 说明:

User Defined Join 指定两个输入流，您可以自己实现Join的逻辑，这里不展开讨论。

根据不同的场景，用户可以使用不同的Join类型来实现对应的关联操作。但是在实际使用过程中，存在这样的错误示例：

```
A (LEFT/RIGHT/FULL/LEFT SEMI/LEFT ANTI) JOIN B
ON a.key = b.key and A.ds='20180101' and B.ds='20180101';
```

这里用户的本意是希望在A和B中获取某一个分区的数据进行JOIN操作，也就是：

```
(SELECT * FROM A WHERE ds='20180101') A
(LEFT/RIGHT/FULL/LEFT SEMI/LEFT ANTI)  JOIN
(SELECT * FROM B WHERE ds='20180101') B
ON a.key = b.key
```

然而针对不同的Join类型，两者可能并不等价，不仅无法将分区条件下推，导致全表扫描，而且会导致正确性问题。这里简要辨析一下过滤条件分别在：

1. 子查询的WHERE条件。
2. JOIN ON条件。
3. JOIN ON后的WHERE条件。

原理

这里先说明一个JOIN和WHERE条件的计算顺序，对于：

```
(SELECT * FROM A WHERE {subquery_where_condition} A) A
JOIN
(SELECT * FROM B WHERE {subquery_where_condition} B) B
ON {on_condition}
```

```
WHERE {where_condition}
```

来说，计算顺序为：

1. 子查询中的`{subquery_where_condition}`
2. JOIN的`{on_condition}`的条件
3. JOIN结果集合`{where_condition}`的计算

对于不同的JOIN类型，过滤语句放在`{subquery_where_condition}`、`{on_condition}`和`{where_condition}`中，有时结果是一致的，有时候结果又是不一致的。下面分情况进行讨论。

实验

1. 准备

   首先构造表A：

   ```
   CREATE TABLE A AS SELECT * FROM VALUES (1, 20180101),(2, 20180101),(2, 20180102) t (key, ds);
   ```

key	ds
1	20180101
2	20180101
2	20180102

   表B：

   ```
   CREATE TABLE B AS SELECT * FROM VALUES (1, 20180101),(3, 20180101),(2, 20180102) t (key, ds);
   ```

key	ds
1	20180101
3	20180101
2	20180102

   则他们的笛卡尔乘积为：

a.key	a.ds	b.key	b.ds
1	20180101	1	20180101
1	20180101	3	20180101
1	20180101	2	20180102
2	20180101	1	20180101

a.key	a.ds	b.key	b.ds
2	20180101	3	20180101
2	20180101	2	20180102
2	20180102	1	20180101
2	20180102	3	20180101
2	20180102	2	20180102

2. Inner Join

   结论：过滤条件在`{subquery_where_condition}`、`{on_condition}`和`{where_condition}`中都是等价的。

   Inner Join的处理逻辑是将左右表进行笛卡尔乘积，然后选择满足ON表达式的行进行输出。

   a. 第一种情况，子查询中过滤：

   ```
   SELECT A.*, B.*
   FROM
   (SELECT * FROM A WHERE ds='20180101') A
   JOIN
   (SELECT * FROM B WHERE ds='20180101') B
   ON a.key = b.key;
   ```

   非常简单，结果只有一条：

a.key	a.ds	b.key	b.ds
1	20180101	1	20180101

   b. 第二种情况，JOIN 条件中过滤：

   ```
   SELECT A.*, B.*
   FROM A JOIN B
   ON a.key = b.key and A.ds='20180101' and B.ds='20180101';
   ```

   笛卡尔积的结果有9条，满足ON条件的结果同样只有1条：

a.key	a.ds	b.key	b.ds
1	20180101	1	20180101

   c. 第三种情况，JOIN后的WHERE条件过滤：

   ```
   SELECT A.*, B.*
   FROM A JOIN B
   ON a.key = b.key
   ```

```
WHERE A.ds='20180101' and B.ds='20180101';
```

来说，笛卡尔积的结果有9条，满足ON条件`a.key = b.key`的结果有3条，分别是：

a.key	a.ds	b.key	b.ds
1	20180101	1	20180101
2	20180102	2	20180102
2	20180101	2	20180102

此时对于这个结果再进行过滤`A.ds='20180101' and B.ds='20180101'`，结果只有1条：

a.key	a.ds	b.key	b.ds
1	20180101	1	20180101

可以看到，将过滤条件放在三个不同的地方，得到了三种不同的结果。

3. Left Join

结论：过滤条件在`{subquery_where_condition}`、`{on_condition}`和`{where_condition}`不一定等价。

对于左表的过滤条件，放在`{subquery_where_condition}`和`{where_condition}`是等价的。

对于右表的过滤条件，放在`{subquery_where_condition}`和`{on_condition}`中是等价的。

Left Join的处理逻辑是将左右表进行笛卡尔乘积，然后对于满足ON表达式的行进行输出，对于左表中不满足ON表达式的行，输出左表，右表补NULL。

a. 第一种情况，子查询中过滤：

```
SELECT A.*, B.*
FROM
(SELECT * FROM A WHERE ds='20180101') A
LEFT JOIN
(SELECT * FROM B WHERE ds='20180101') B
ON a.key = b.key;
```

过滤后，左右侧有两条，右侧有一条，结果有两条：

a.key	a.ds	b.key	b.ds
1	20180101	1	20180101

a.key	a.ds	b.key	b.ds
2	20180101	NULL	NULL

b. 第二种情况，JOIN 条件中过滤：

```
SELECT A.*, B.*
FROM A JOIN B
ON a.key = b.key and A.ds='20180101' and B.ds='20180101';
```

笛卡尔积的结果有9条，满足ON条件的结果同样只有1条，则对于左表剩余的两条输出左表，右表补NULL：

a.key	a.ds	b.key	b.ds
1	20180101	1	20180101
2	20180101	NULL	NULL
2	20180102	NULL	NULL

c. 第三种情况，JOIN后的WHERE条件过滤：

```
SELECT A.*, B.*
FROM A JOIN B
ON a.key = b.key
WHERE A.ds='20180101' and B.ds='20180101';
```

来说，笛卡尔积的结果有9条，满足ON条件`a.key = b.key`的结果有3条，分别是：

a.key	a.ds	b.key	b.ds
1	20180101	1	20180101
2	20180101	2	20180102
2	20180102	2	20180102

此时对于这个结果再进行过滤`A.ds='20180101' and B.ds='20180101'`，结果只有1条：

a.key	a.ds	b.key	b.ds
1	20180101	1	20180101

可以看到，将过滤条件放在三个不同的地方，得到了三种不同的结果。

4. Right Join

Right Join和Left Join是类似的，只是左右表的区别。结论：过滤条件在`{subquery_where_condition}`、`{on_condition}`和`{where_condition}`不一定等价。对于右表的过

滤条件，放在{subquery_where_condition}和{where_condition}是等价的。对于左表的过滤条件，放在{subquery_where_condition}和{on_condition}中是等价的。

5. Full Join

结论：过滤条件写在{subquery_where_condition}、{on_condition}和{where_condition}均不等价。

FULL Join的处理逻辑是将左右表进行笛卡尔乘积，然后对于满足ON表达式的行进行输出，对于两侧表中不满足ON表达式的行，输出有数据的表，另一侧补NULL。

a. 第一种情况，子查询中过滤：

```
SELECT A.*, B.*
FROM
(SELECT * FROM A WHERE ds='20180101') A
FULL JOIN
(SELECT * FROM B WHERE ds='20180101') B
ON a.key = b.key;
```

过滤后，左右侧有两条，右侧有两条，结果有三条：

a.key	a.ds	b.key	b.ds
1	20180101	1	20180101
2	20180101	NULL	NULL
NULL	NULL	3	20180101

b. 第二种情况，JOIN 条件中过滤：

```
SELECT A.*, B.*
FROM A FULL JOIN B
ON a.key = b.key and A.ds='20180101' and B.ds='20180101';
```

笛卡尔积的结果有9条，满足ON条件的结果同样只有1条，则对于左表剩余的两条输出左表，右表补NULL。右表剩余的两条输出右表，左表补NULL：

a.key	a.ds	b.key	b.ds
1	20180101	1	20180101
2	20180101	NULL	NULL
2	20180102	NULL	NULL
NULL	NULL	3	20180101
NULL	NULL	2	20180102

c. 第三种情况，JOIN后的WHERE条件过滤：

```
SELECT A.*, B.*
```

```
FROM A FULL JOIN B
ON a.key = b.key
WHERE A.ds='20180101' and B.ds='20180101';
```

笛卡尔积的结果有9条，满足ON条件`a.key = b.key`的结果有3条，分别是：

a.key	a.ds	b.key	b.ds
1	20180101	1	20180101
2	20180101	2	20180102
2	20180102	2	20180102

再对没有JOIN上的数据进行输出，另一侧补NULL，得到结果：

a.key	a.ds	b.key	b.ds
1	20180101	1	20180101
2	20180101	2	20180102
2	20180102	2	20180102
NULL	NULL	3	20180101

此时对于这个结果再进行过滤`A.ds='20180101' and B.ds='20180101'`，结果只有1条：

a.key	a.ds	b.key	b.ds
1	20180101	1	20180101

可以看到，和LEFT JOIN类似，得到了三种不同的结果。

6. Left Semi Join

结论：过滤条件写在`{subquery_where_condition}`、`{on_condition}`和`{where_condition}`是等价的。

LEFT SEMI Join的处理逻辑是对于左表的每一条记录，都去和右表进行匹配，如果匹配成功，则输出左表。这里需要注意的是由于只输出左表，所以JOIN后的Where条件中不能写右侧的过滤条件。LEFT SEMI JOIN常用来实现exists的语义：

a. 第一种情况，子查询中过滤：

```
SELECT A.*
FROM
(SELECT * FROM A WHERE ds='20180101') A
LEFT SEMI JOIN
(SELECT * FROM B WHERE ds='20180101') B
```

```
ON a.key = b.key;
```

过滤后，左右侧有两条，最终符合`a.key = b.key`的只有一条：

a.key	a.ds
1	20180101

b. 第二种情况，JOIN 条件中过滤：

```
SELECT A.*
FROM A LEFT SEMI JOIN B
ON a.key = b.key and A.ds='20180101' and B.ds='20180101';
```

对于左侧的三条记录，满足ON条件的结果同样只有1条：

a.key	a.ds
1	20180101

c. 第三种情况，JOIN后的WHERE条件过滤：

```
SELECT A.*
FROM A LEFT SEMI JOIN
(SELECT * FROM B WHERE ds='20180101') B
ON a.key = b.key
WHERE A.ds='20180101';
```

左侧能符合ON条件的有一条：

a.key	a.ds
1	20180101

此时对于这个结果再进行过滤`A.ds='20180101'`，结果仍然保持1条：

a.key	a.ds
1	20180101

可以看到，LEFT SEMI JOIN和INNER JOIN类似，无论过滤条件放在哪里，结果都是一致的。

7. Left Anti Join

结论：过滤条件写在{subquery_where_condition}、{on_condition}和{where_condition}不一定等价。

对于左表的过滤条件，放在{subquery_where_condition}和{where_condition}是等价的。

对于右表的过滤条件，放在{subquery_where_condition}和{on_condition}中是等价的，右表表达式不能放在{where_condition}中。

LEFT ANTI Join的处理逻辑是对于左表的每一条记录，都去和右表进行匹配，如果右表所有的记录都没有匹配成功，则输出左表。同样由于只输出左表，所以JOIN后的Where条件中不能写右侧的过滤条件。LEFT SEMI JOIN常常用来实现not exists的语义。

a. 第一种情况，子查询中过滤：

```
SELECT A.*
FROM
(SELECT * FROM A WHERE ds='20180101') A
LEFT ANTI JOIN
(SELECT * FROM B WHERE ds='20180101') B
ON a.key = b.key;
```

过滤后，左侧有两条，右侧有两条，结果有1条：

a.key	a.ds
2	20180101

b. 第二种情况，JOIN 条件中过滤：

```
SELECT A.*
FROM A LEFT ANTI JOIN B
ON a.key = b.key and A.ds='20180101' and B.ds='20180101';
```

对于左侧的三条记录，只有第一条有满足ON条件的结果，所以输出剩余的两条记录：

a.key	a.ds
2	20180101
2	20180102

c. 第三种情况，JOIN后的WHERE条件过滤：

```
SELECT A.*
FROM A LEFT ANTI JOIN
(SELECT * FROM B WHERE ds='20180101') B
ON a.key = b.key
```

```
WHERE A.ds='20180101';
```

左侧能通过ON条件的有两条：

a.key	a.ds
2	20180101
2	20180102

此时对于这个结果再进行过滤`A.ds='20180101'`，结果为1条：

a.key	a.ds
2	20180101

可以看到，LEFT ANTI JOIN中，过滤条件WHERE语句分别放在JOIN ON条件中、条件前和条件后，得到的结果是不相同的。

以上内容只是针一个常用场景测试的几种不同的写法，没有具体的推导过程，对于涉及到不等值表达式的场景会更加复杂，如果您有兴趣可以尝试推导一下。

总结

过滤条件放在不同的位置语义可能大不相同，对于用户而言，如果只是进行过滤数据后再JOIN的操作，可以简要记住以下几点。

1. INNER JOIN/LEFT SEMI JOIN 对于两侧的表达式可以随便写。
2. LEFT JOIN/LEFT ANTI JOIN 左表的过滤条件要放到`{subquery_where_condition}`或者`{where_condition}`，右表的过滤条件要放到`{subquery_where_condition}`或者`{on_condition}`中。
3. RIGHT JOIN和LEFT JOIN相反，右表的过滤条件要放到`{subquery_where_condition}`或者`{where_condition}`，左表的过滤条件要放到`{subquery_where_condition}`或者`{on_condition}`。
4. FULL OUTER JOIN 只能放到`{subquery_where_condition}`中。

当然如果还是觉得规则比较复杂的话，最好的方法就是每次都把过滤条件写到子查询中。

2 数据迁移

2.1 数据迁移

本文整理总结将数据移迁移到MaxCompute的最佳实践相关文档。

当前很多用户的数据存放在传统的关系型数据库（RDS，做业务读写操作）中，当业务数据量庞大的时候，传统关系型数据库会显得有些吃力，此时经常会将数据迁移到大数据计算服务MaxCompute上。MaxCompute为您提供了完善的数据导入方案以及多种经典的分布式计算模型，能够更快速的解决海量数据存储和计算问题，有效降低企业成本。作为MaxCompute开发套件的DataWorks为MaxCompute提供一站式的数据同步、工作流开发、数据管理和数据运维等功能。#unique_23为您介绍阿里集团对外提供的稳定高效、弹性伸缩的数据同步平台。

最佳实践合集

- 通过使用DataWorks数据同步功能，将Hadoop数据迁移到阿里云MaxCompute大数据计算服务上，请参见#unique_24。详细的视频介绍，请参见Hadoop数据迁移到MaxCompute最佳实践（视频）。自建Hadoop迁移阿里云MaxCompute实践定期整理一些数据迁移和脚本迁移遇到的问题及解决方案，帮助企业快速拥有阿里巴巴同款数据仓库，构建自己的数据平台，并开展数据业务。
- 使用DataWorks数据集成同步功能，自动创建分区，动态的将RDS中的数据迁移到MaxCompute大数据计算服务上。请参见#unique_25。
- 利用DataWorks数据集成，将JSON数据从OSS迁移到MaxCompute，并使用MaxCompute内置字符串函数GET_JSON_OBJECT提取JSON信息。详细描述请参见#unique_26。
- 利用DataWorks数据集成直接从MongoDB提取JSON字段到MaxCompute，请参见#unique_27。

2.2 MaxCompute跨项目迁移

本文为您介绍同Region下不同的MaxCompute项目如何实现配置与数据的迁移。

前提条件

请您首先完成教程《搭建互联网在线运营分析平台》的全部步骤，详情请参见#unique_29。

背景信息

本文使用的被迁移的原始项目为教程《搭建互联网在线运营分析平台》中的`bigdata_DOC`项目，您需要再创建一个迁移目标项目，用于存放原始项目的表、资源、配置及数据。

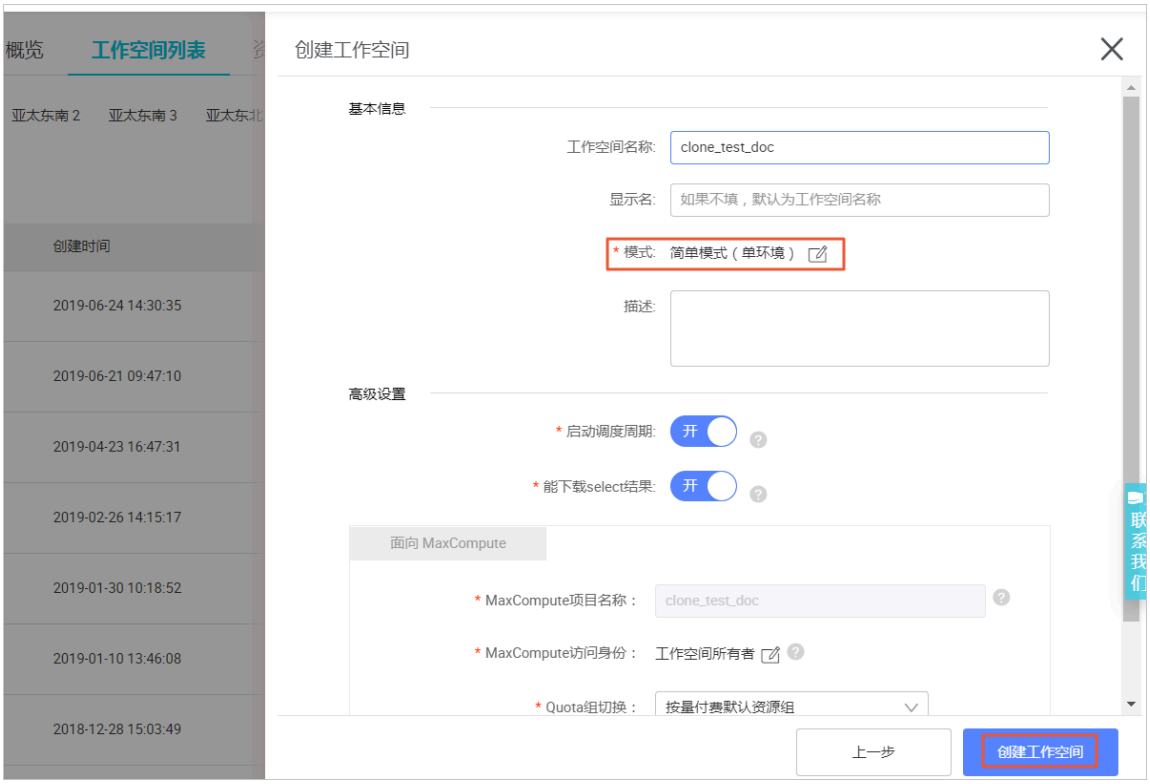
操作步骤

1. **创建迁移目标项目**

   本文中的MaxCompute项目即DataWorks的工作空间。

   a) 进入DataWorks工作空间列表，选择区域为华东1，单击创建工作空间。

   b) 选择计算引擎服务为MaxCompute、按量付费。由于原始项目bigdata_DOC为简单模式，为方便起见，本文中DataWorks工作空间模式也为简单模式（单环境）。在简单模式下，DataWorks工作空间与MaxCompute项目一一对应，详情请参见#unique_30。



说明:

工作空间名称全局唯一，建议您使用易于区分的名称，本例中使用的名称为clone_test_doc。

2. 跨项目克隆

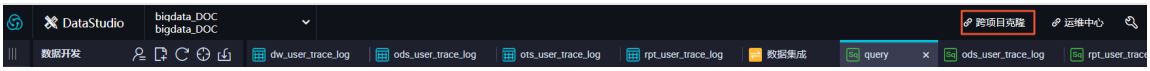
您可以通过跨项目克隆功能将原始项目bigdata_DOC的节点配置和资源复制到当前项目，详情请参见#unique_31。

说明:

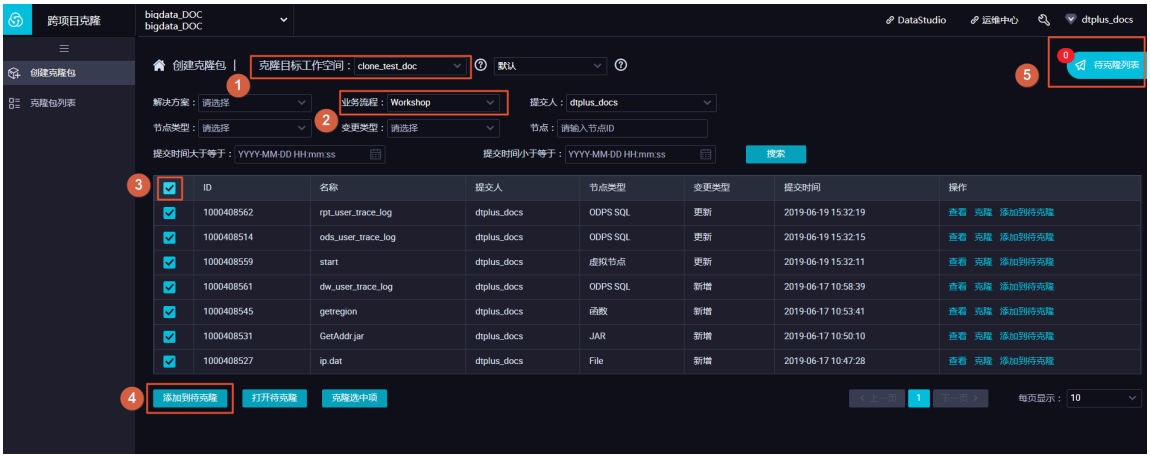
- 跨项目克隆无法复制表结构与数据。

· 跨项目克隆无法复制组合节点，需要您手动创建。

a) 在单原始的项目bigdata_DOC中击右上角的跨项目克隆，跳转至相应的克隆页面。

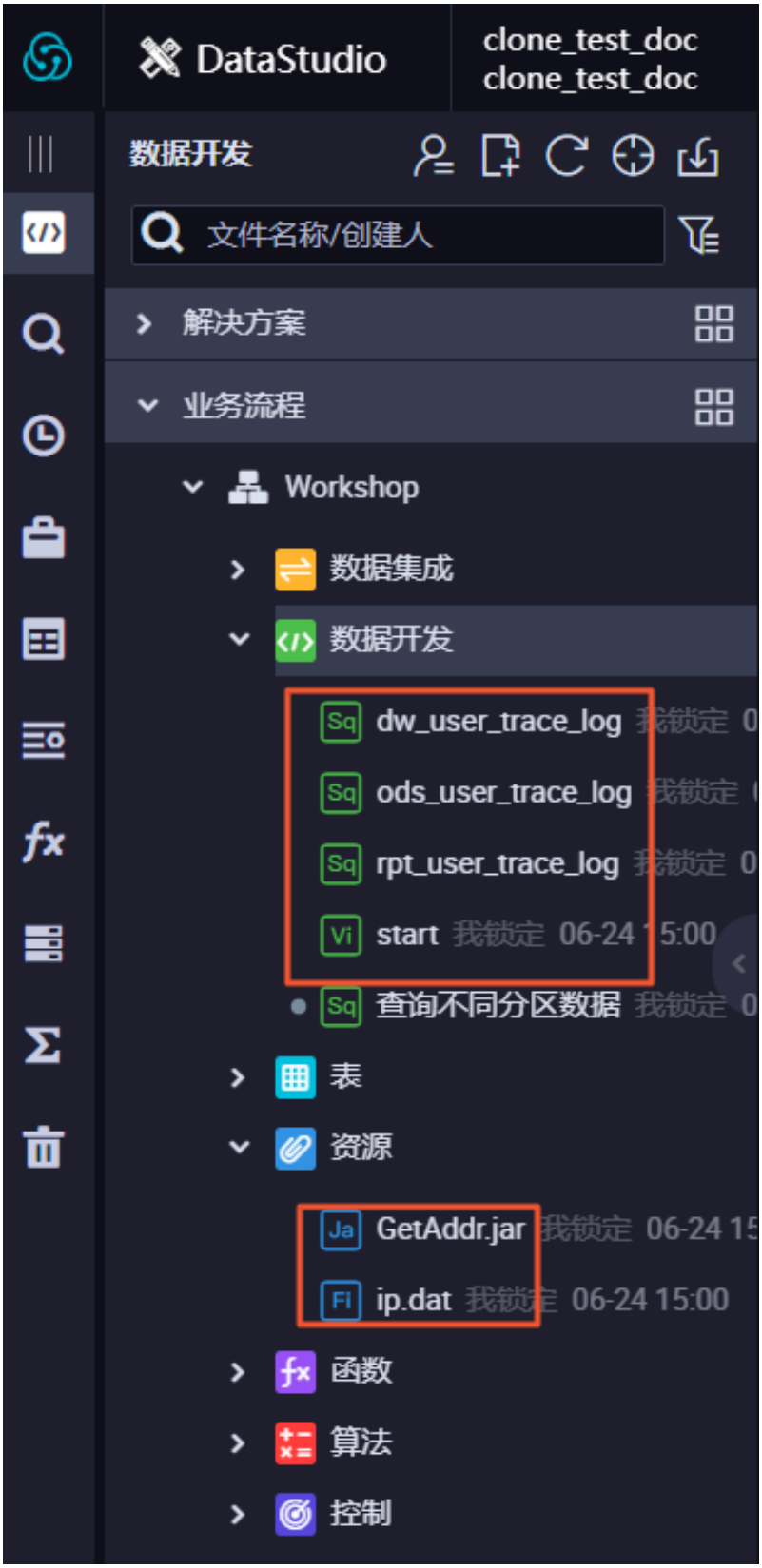


b) 选择克隆目标工作空间为clone_test_doc，业务流程为您需要克隆的业务流程Workshop，勾选所有节点，单击添加到待克隆后单击右侧的待克隆列表。



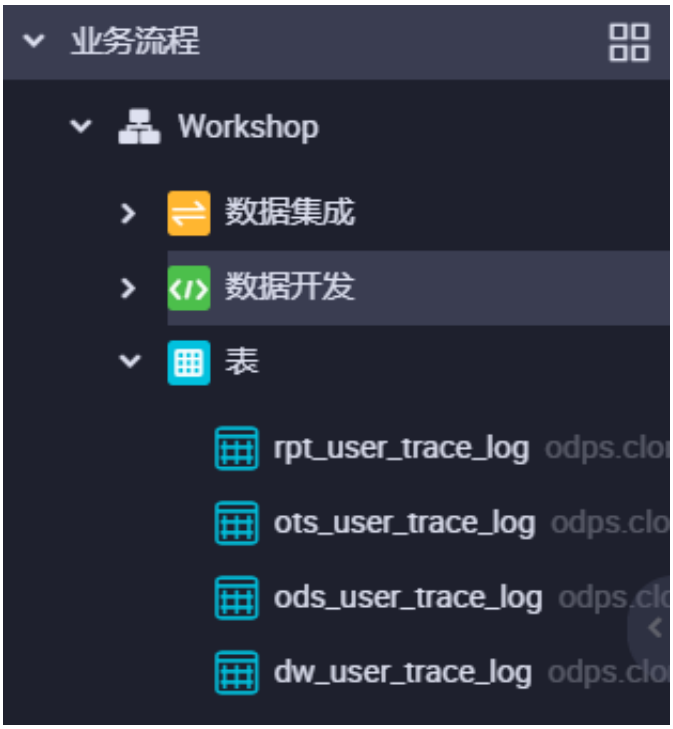
c) 单击全部克隆，将选中节点克隆至工作空间clone_test_DOC。

d) 切换到您新建的项目，检查节点是否已完成克隆。


DataStudio
clone_test_doc
clone_test_doc
数据开发
文件名称/创建人
解决方案
业务流程
Workshop
数据集成
数据开发
dw_user_trace_log
ods_user_trace_log
rpt_user_trace_log
start
查询不同分区数据
表
资源
GetAddr.jar
ip.dat
06-24 15:00
函数
算法
控制

3. 新建数据表

跨项目克隆功能无法克隆您的表结构，因此您需要手动新建表。您可以参见#unique_32完成表的创建。



说明:

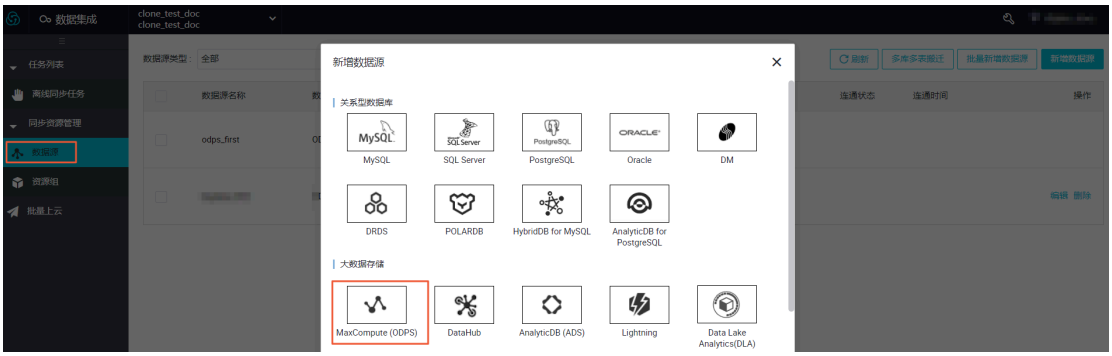
新建表后请将表提交到生产环境。

4. 数据同步

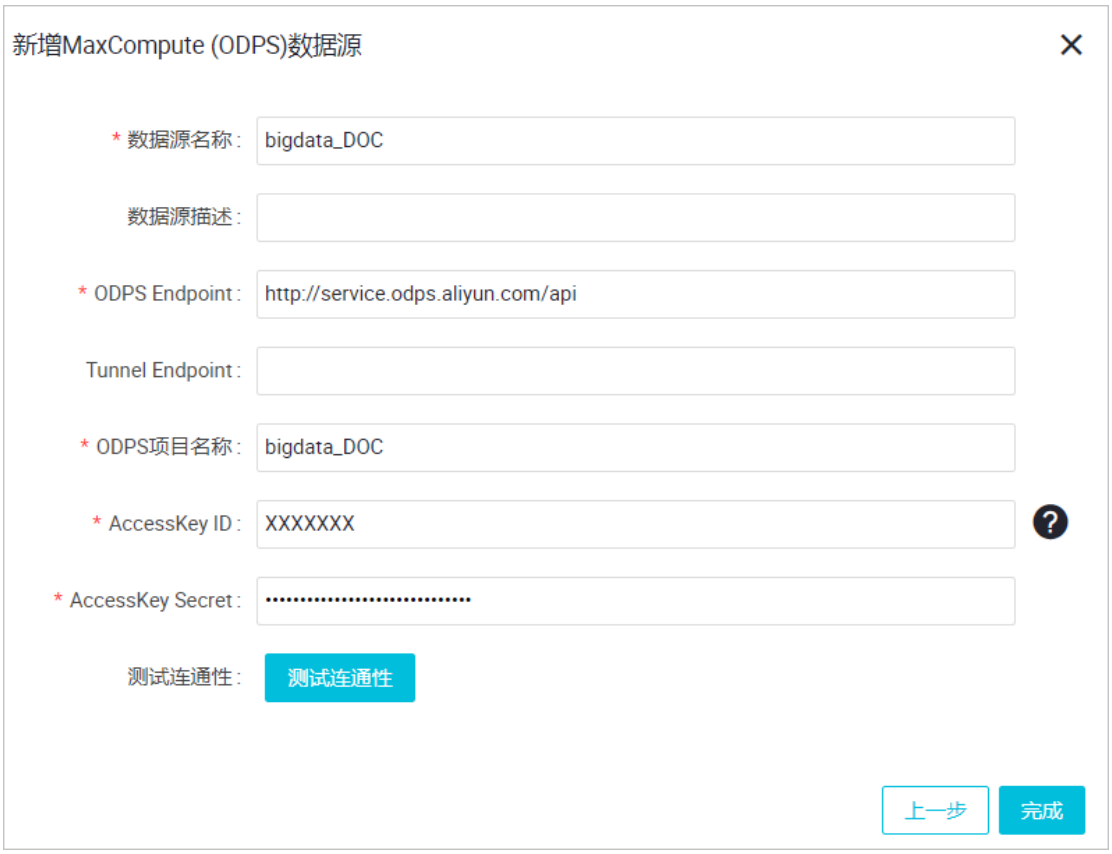
跨项目克隆功能无法复制原始项目的数据到新项目，因此您需要手动同步数据，本文中仅同步表rpt_user_trace_log的数据。

a) 新建数据源。

A. 在数据集成页面单击同步资源管理 > 数据源，选择MaxCompute。



B. 填写您的数据源名称、ODPS项目名称、AccessKey等信息，单击完成，详情请参见#unique_33。

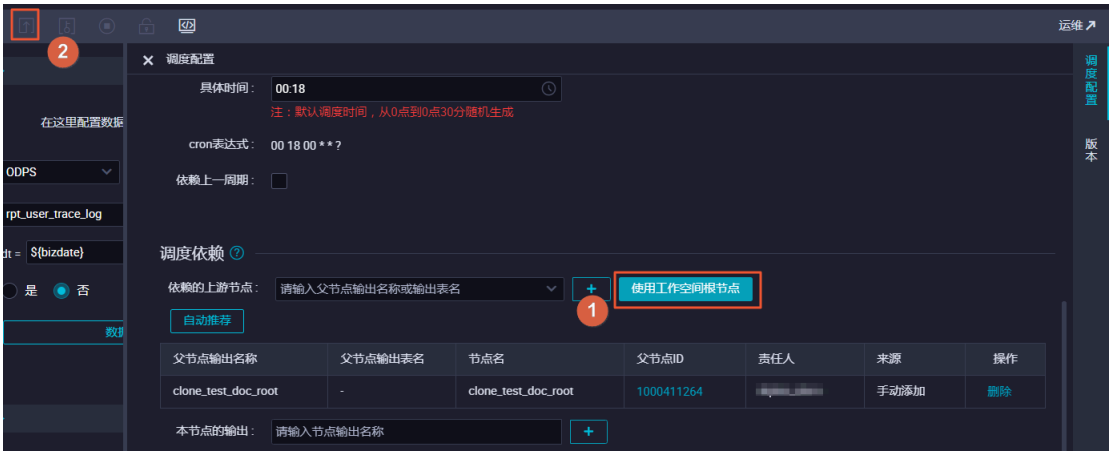


b) 创建数据同步任务。

A. 在您的数据开发页面右键您克隆的业务流程Workshop下的数据集成，单击新建数据集成节点 > 数据同步。

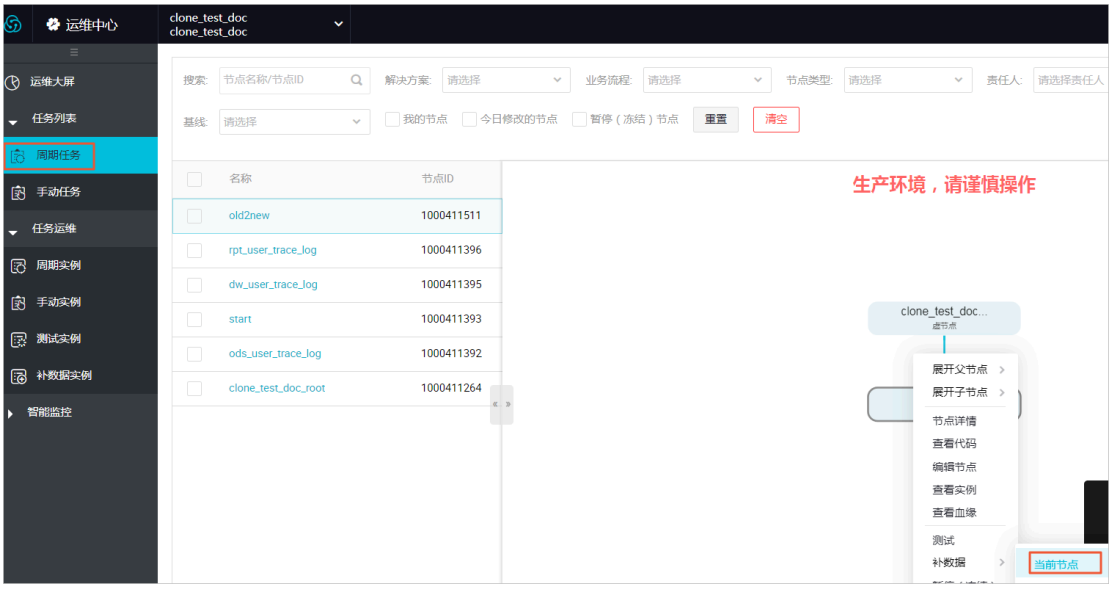
B. 编辑您新建的数据同步任务节点，填写参数如下图所示。其中数据源bigdata_DOC是您的原始项目，数据源odps_first代表您当前的新建项目，表名是您需要同步数据的表rpt_user_trace_log。完成后单击调度配置。

C. 单击使用工作空间根节点后，提交数据同步任务。

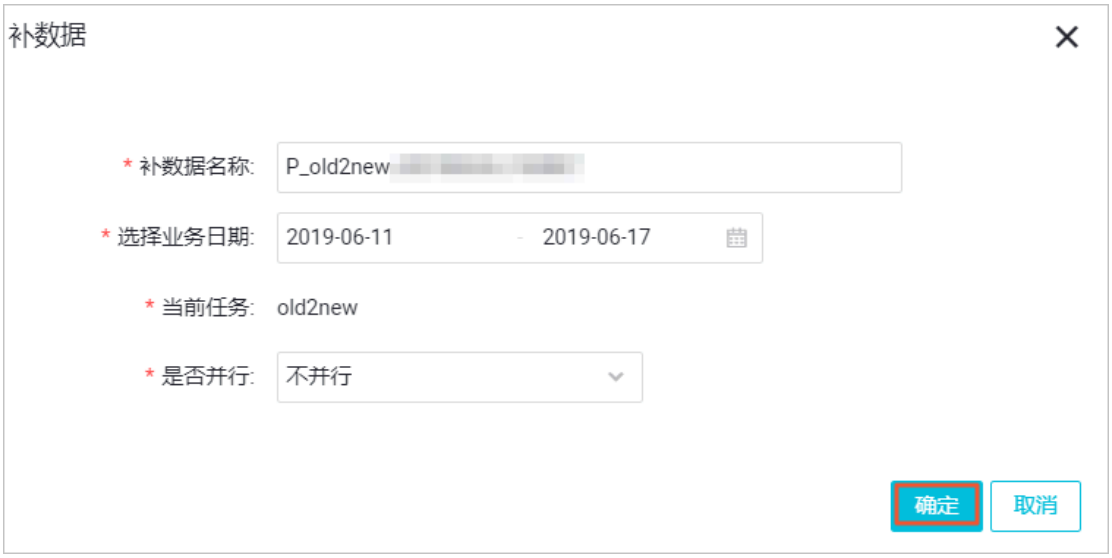


c) 补数据

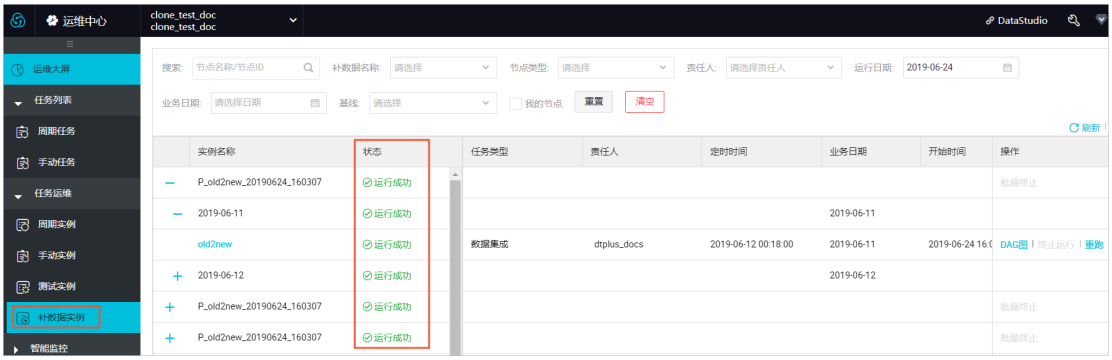
A. 切换到运维中心后，单击周期任务。右键您的数据同步任务，单击补数据 > 当前节点。



B. 本例中，需要补数据的日期分区为2019年6月11日到17日，您可以直接选择业务日期，进行多个分区的数据同步。完成设置后，单击确定。

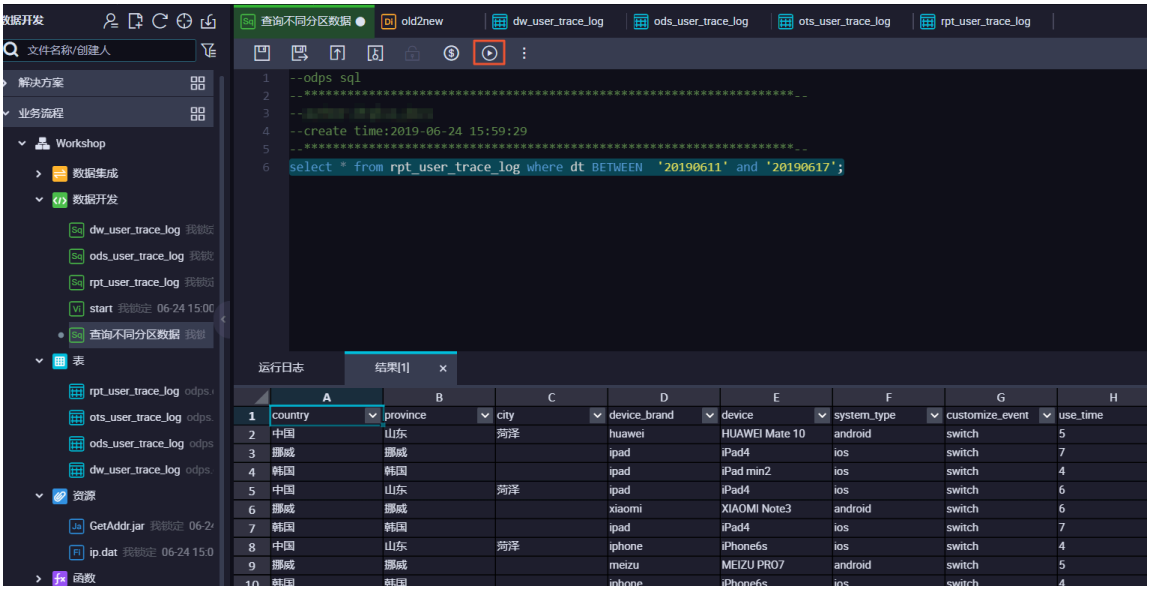


C. 在任务运维 > 补数据实例页面，您可以查看您的补数据实例任务运行状态，待显示运行成功则说明完成数据同步。



d) 结果验证

您可以在业务流程 > 数据开发中新建ODPS SQL类型节点，通过运行`select * from rpt_user_trace_log where dt BETWEEN '20190611' and '20190617';`语句查看数据是否完成同步。



2.3 Hadoop数据迁移MaxCompute最佳实践

本文将为您介绍如何通过DataWorks数据同步功能，将HDFS上的数据迁移至MaxCompute，或从MaxCompute将数据迁移至HDFS。无论您使用Hadoop还是Spark，均可以与MaxCompute之间的数据进行双向同步。

前提条件

1. Hadoop集群搭建

   进行数据迁移前，您需要保证自己的Hadoop集群环境正常。本文使用阿里云EMR服务自动化搭建Hadoop集群，详细过程请参见创建集群。

   本文使用的EMR Hadoop版本信息如下：

   EMR版本: EMR-3.11.0

   集群类型: HADOOP

   软件信息: HDFS2.7.2 / YARN2.7.2 / Hive2.3.3 / Ganglia3.7.2 / Spark2.2.1 / HUE4.1.0 / Zeppelin0.7.3 / Tez0.9.1 / Sqoop1.4.6 / Pig0.14.0 / ApacheDS2.0.0 / Knox0.13.0

   Hadoop集群使用经典网络，区域为华东1（杭州），主实例组ECS计算资源配置公网及内网IP，高可用选择为否（非HA模式）。

2. MaxCompute

   #unique_36并创建好项目。本文以在华东1（杭州）区域创建项目bigdata_DOC为例，同时启动DataWorks相关服务。

操作步骤

1. 数据准备

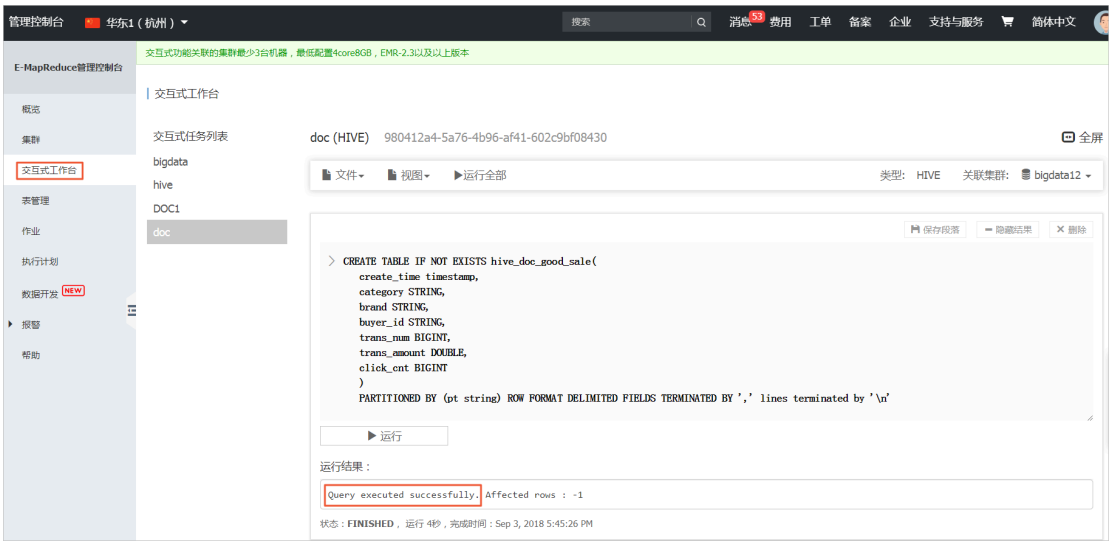
   a. Hadoop集群创建测试数据

      A. 进入EMR Hadoop集群控制台界面，使用交互式工作台，新建交互式任务doc。本例中HIVE建表语句如下。

```
CREATE TABLE IF NOT
EXISTS hive_doc_good_sale(
   create_time timestamp,
```

```
    category STRING,
    brand STRING,
    buyer_id STRING,
    trans_num BIGINT,
    trans_amount DOUBLE,
    click_cnt BIGINT
    )
    PARTITIONED BY (pt string) ROW FORMAT
DELIMITED FIELDS TERMINATED BY ',' lines terminated by '\n'
```

B. 选择运行，出现`Query executed successfully`提示，则说明成功在EMR Hadoop集群上创建了表hive_doc_good_sale。

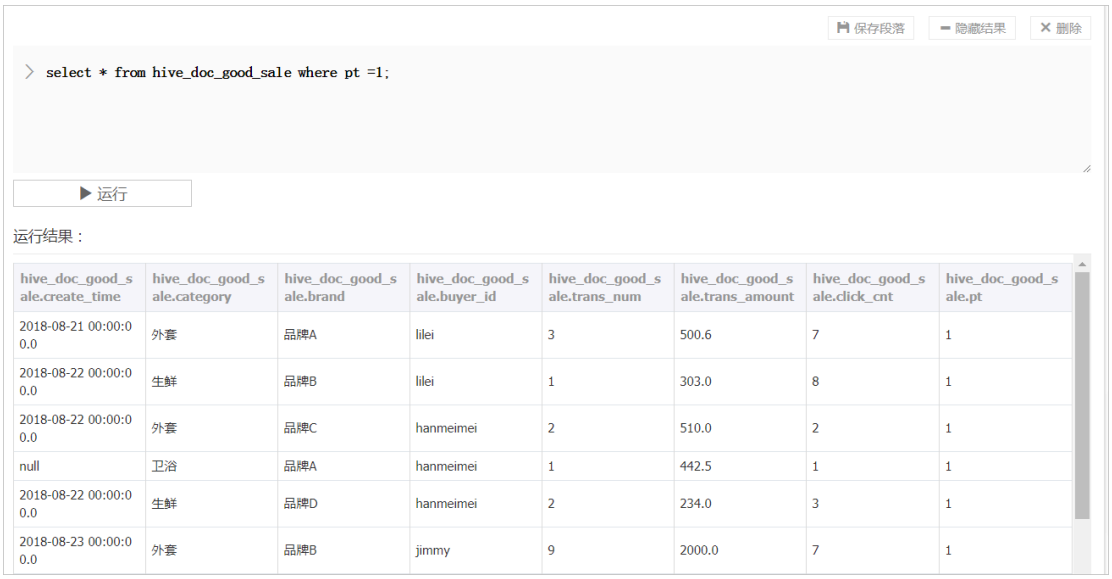


C. 插入测试数据。您可以选择从OSS或其他数据源导入测试数据，也可以手动插入少量的测试数据。本文中手动插入数据如下。

```
insert into
hive_doc_good_sale PARTITION(pt =1 ) values('2018-08-21','外
套','品牌A','lilei',3,500.6,7),('2018-08-22','生鲜','品牌B','lilei
',1,303,8),('2018-08-22','外套','品牌C','hanmeimei',2,510,2),(
2018-08-22,'卫浴','品牌A','hanmeimei',1,442.5,1),('2018-08-22
','生鲜','品牌D','hanmeimei',2,234,3),('2018-08-23','外套','品牌B
','jimmy',9,2000,7),('2018-08-23','生鲜','品牌A','jimmy',5,45.1
,5),('2018-08-23','外套','品牌E','jimmy',5,100.2,4),('2018-08-24
','生鲜','品牌G','peiqi',10,5560,7),('2018-08-24','卫浴','品牌F','
peiqi',1,445.6,2),('2018-08-24','外套','品牌A','ray',3,777,3),('
```

```
2018-08-24','卫浴','品牌G','ray',3,122,3),('2018-08-24','外套','品牌C','ray',1,62,7) ;
```

D. 完成插入数据后，您可以执行`select * from hive_doc_good_sale where pt =1 ;`语句，检查Hadoop集群表中是否已存在数据可以用于迁移。

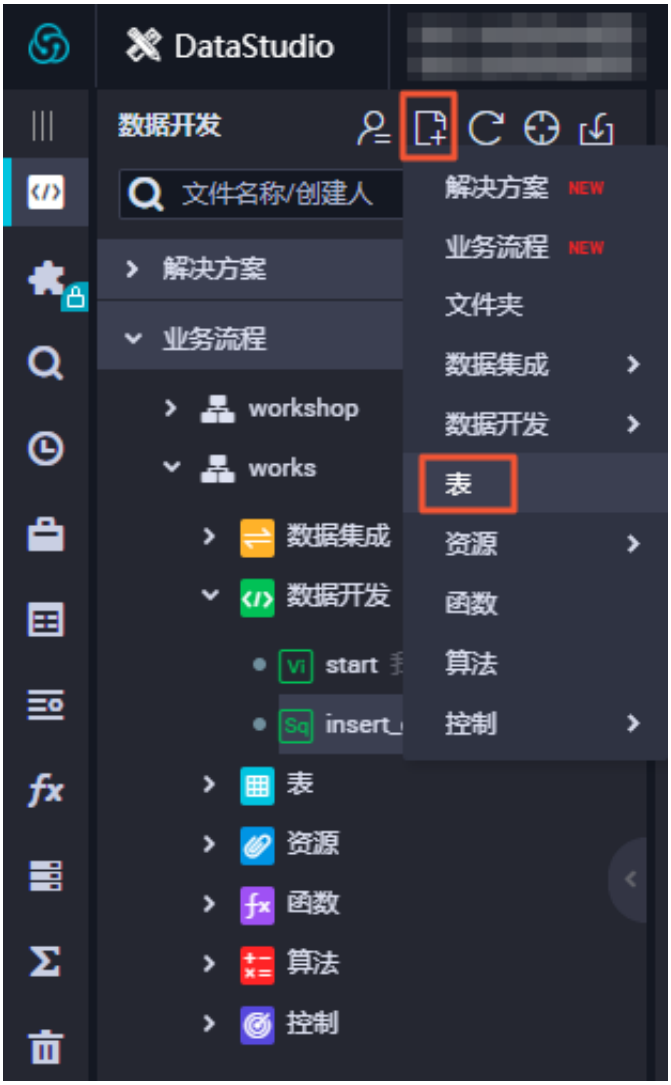


hive_doc_good_sale.create_time	hive_doc_good_sale.category	hive_doc_good_sale.brand	hive_doc_good_sale.buyer_id	hive_doc_good_sale.trans_num	hive_doc_good_sale.trans_amount	hive_doc_good_sale.click_cnt	hive_doc_good_sale.pt
2018-08-21 00:00:00.0	外套	品牌A	lilei	3	500.6	7	1
2018-08-22 00:00:00.0	生鲜	品牌B	lilei	1	303.0	8	1
2018-08-22 00:00:00.0	外套	品牌C	hanmeimei	2	510.0	2	1
null	卫浴	品牌A	hanmeimei	1	442.5	1	1
2018-08-22 00:00:00.0	生鲜	品牌D	hanmeimei	2	234.0	3	1
2018-08-23 00:00:00.0	外套	品牌B	jimmy	9	2000.0	7	1

b. 利用DataWorks新建目标表

A. 登录DataWorks控制台，单击相应工作空间操作栏下的进入数据开发。

B. 进入DataStudio（数据开发）页面，选择新建 > 表。



C. 在新建表对话框中，填写表名，并单击提交。

D. 进入新建表页面，选择DDL模式。

E. 在DDL模式对话框中输入建表语句，单击生成表结构，并确认操作。本示例的建表语句如下所示：

```
CREATE TABLE IF NOT EXISTS hive_doc_good_sale(
   create_time string,
   category STRING,
   brand STRING,
   buyer_id STRING,
   trans_num BIGINT,
   trans_amount DOUBLE,
   click_cnt BIGINT
   )
```

```
    PARTITIONED BY (pt string) ;
```

在建表过程中，需要考虑HIVE数据类型与MaxCompute数据类型的映射，当前数据映射关系请参见#unique_37。

由于本文使用DataWorks进行数据迁移，而DataWorks数据同步功能暂不支持TIMESTAMP类型数据。因此在DataWorks建表语句中，将create_time设置为STRING类型。上述步骤同样可通过odpscmd命令行工具完成，命令行工具安装和配置请参见#unique_38。

```
odps@ bigdata_DOC>CREATE TABLE IF NOT EXISTS hive_doc_good_sale(
   create_time timestamp,
   category STRING,
   brand STRING,
   buyer_id STRING,
   trans_num BIGINT,
   trans_amount DOUBLE,
   click_cnt BIGINT
   )
   PARTITIONED BY (pt string) ;
                >               >               >               >
        >               >               >               >
 >
ID = 20180906110540873gev1bpim
OK
odps@ bigdata_DOC>drop table hive_doc_good_sale;
Confirm to "drop table hive_doc_good_sale;" (yes/no)? yes

ID = 20180906110825180gxh66292
```

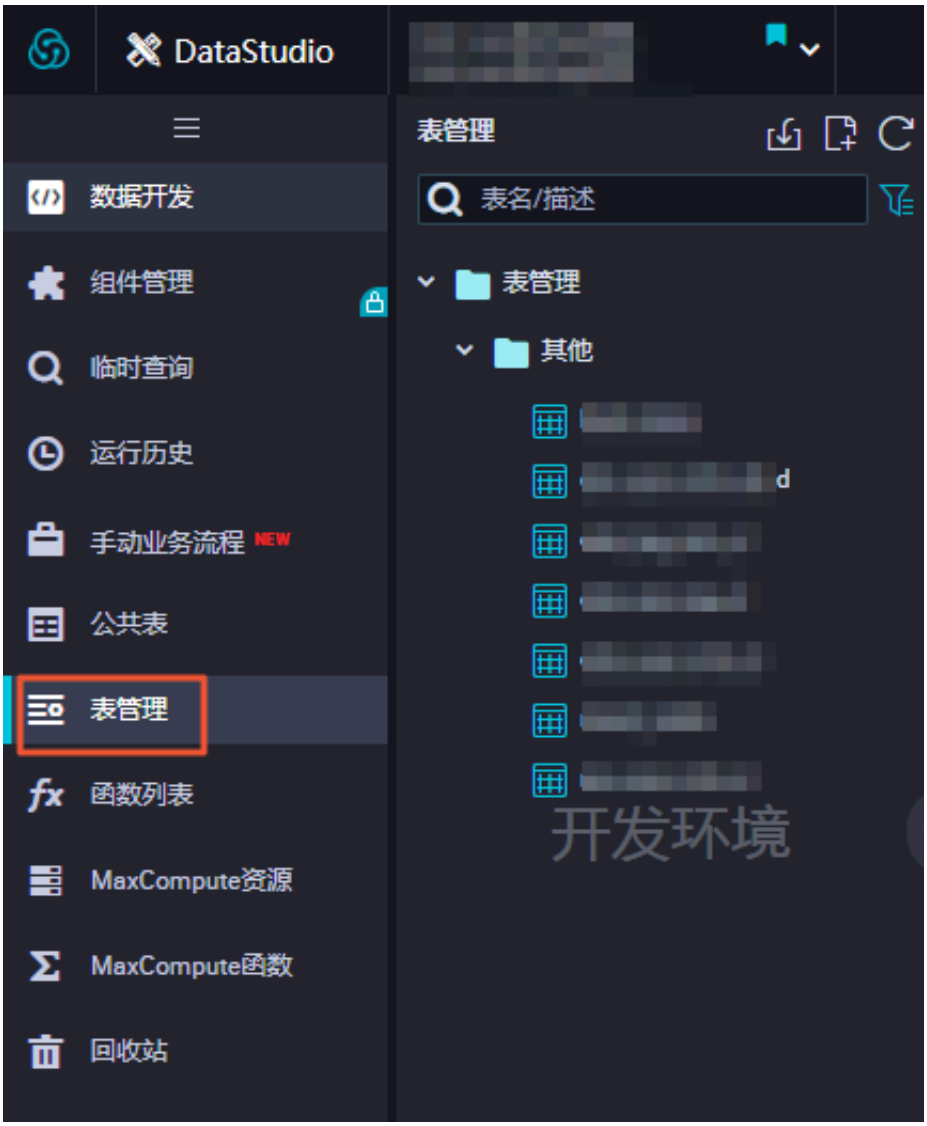
说明:

考虑到部分HIVE与MaxCompute数据类型的兼容问题，建议在odpscmd客户端上执行以下命令。

```
set odps.sql.type.system.odps2=true;
```

```
set odps.sql.hive.compatible=true;
```

F. 完成建表后，单击左侧导航栏中的表管理，即可查看当前创建的MaxCompute表。



2. 数据同步

a. 新建自定义资源组

由于MaxCompute项目所处的网络环境与Hadoop集群中的数据节点（data node）网络通常不可达，您可以通过自定义资源组的方式，将DataWorks的同步任务运行在Hadoop集群的Master节点上（Hadoop集群内Master节点和数据节点通常可达）。

A. 查看Hadoop集群data node

进入EMR控制台，选择首页 > 集群管理 > 集群 > 主机列表。

您也可以通过单击上图中Master节点的ECS ID，进入ECS实例详情页。然后单击远程连接进入ECS，执行`hadoop dfsadmin -report`命令查看data node。

```
DFS Used%: 0.05%
Under replicated blocks: 0
Blocks with corrupt replicas: 0
Missing blocks: 0
Missing blocks (with replication factor 1): 0

-------------------------------------------------
Live datanodes (2):

Name: 10.31.122.189:50010 (emr-worker-1.cluster-74503)
Hostname: emr-worker-1.cluster-74503
Decommission Status : Normal
Configured Capacity: 333373341696 (310.48 GB)
DFS Used: 155725824 (148.51 MB)
Non DFS Used: 325541888 (310.46 MB)
DFS Remaining: 332892073984 (310.03 GB)
DFS Used%: 0.05%
DFS Remaining%: 99.86%
Configured Cache Capacity: 0 (0 B)
Cache Used: 0 (0 B)
Cache Remaining: 0 (0 B)
Cache Used%: 100.00%
Cache Remaining%: 0.00%
Xceivers: 1
Last contact: Thu Sep 06 19:41:01 CST 2018


Name: 10.81.78.209:50010 (emr-worker-2.cluster-74503)
Hostname: emr-worker-2.cluster-74503
Decommission Status : Normal
Configured Capacity: 333373341696 (310.48 GB)
DFS Used: 155725824 (148.51 MB)
Non DFS Used: 325451776 (310.38 MB)
DFS Remaining: 332892164096 (310.03 GB)
DFS Used%: 0.05%
DFS Remaining%: 99.86%
Configured Cache Capacity: 0 (0 B)
Cache Used: 0 (0 B)
Cache Remaining: 0 (0 B)
Cache Used%: 100.00%
Cache Remaining%: 0.00%
Xceivers: 1
Last contact: Thu Sep 06 19:41:02 CST 2018
```

本示例的data node只具有内网地址，很难与DataWorks默认资源组互通，所以需要设置自定义资源组，将master node设置为执行DataWorks数据同步任务的节点。

B. 新建任务资源组

A. 进入数据集成 > 资源组页面，单击右上角的新增任务资源。关于自定义资源组的详细信息请参见。

说明:

目前仅专业版及以上版本方可使用此入口。

B. 添加服务器时，需要输入ECS UUID和机器IP等信息（对于经典网络类型，需要输入服务器名称。对于专有网络类型，需要输入服务器UUID）。目前仅DataWorks V2.0华

东2区支持经典网络类型的调度资源添加，对于其他区域，无论您使用的是经典网络还是专有网络类型，在添加调度资源组时都请选择专有网络类型。

机器IP需要填写master node公网IP（内网IP有可能不可达）。ECS的UUID需要进入master node管理终端，通过命令dmidecode | grep UUID获取（如果您的hadoop集群并非搭建在EMR环境上，也可以通过该命令获取）。

```
[root@emr-header-1 logs]# dmidecode | grep UUID
        UUID: F631D86C-
```

C. 添加服务器后，需要保证master node与DataWorks网络可达。如果您使用的是ECS服务器，需要设置服务器安全组。

- 如果您使用的内网IP互通，请参见添加安全组。
- 如果您使用的是公网IP，可以直接设置安全组公网出入方向规则。本文中设置公网入方向放通所有端口（实际应用场景中，为了您的数据安全，强烈建议设置详细的放通规则）。

D. 完成上述步骤后，按照提示安装自定义资源组agent。当前状态显示为可用时，则新增自定义资源组成功。

如果状态为不可用，您可以登录master node，执行`tail -f/home/admin/alisatasknode/logs/heartbeat.log`命令查看DataWorks与master node之间心跳报文是否超时。

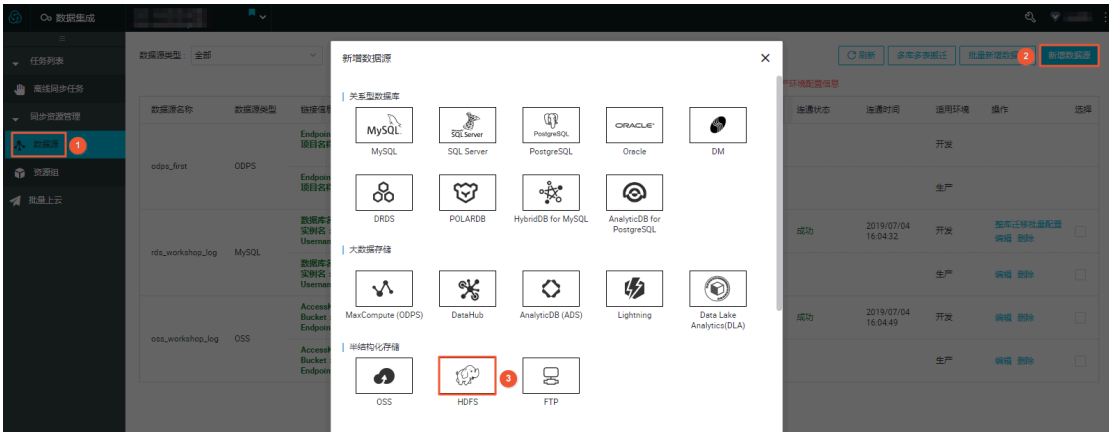
```
[root@emr-header-1 logs]# hdfs dfs -ls /user/hive/warehouse/hive_doc_good_sale/
Found 1 items
drwxr-x--x   - hive hadoop          0 2018-09-03 17:46 /user/hive/warehouse/hive_doc_good_sale/pt=1
[root@emr-header-1 logs]# tail -f /home/admin/alisatasknode/logs/heartbeat.log
2018-09-06 21:47:34,440 INFO  [pool-6-thread-1] [HeartbeatReporter.java:104] [] - heartbeat start, current status:2
2018-09-06 21:47:34,465 INFO  [pool-6-thread-1] [HeartbeatReporter.java:133] [] - heartbeat end■ cost time:0.025s
2018-09-06 21:47:39,465 INFO  [pool-6-thread-1] [HeartbeatReporter.java:104] [] - heartbeat start, current status:2
2018-09-06 21:47:39,491 INFO  [pool-6-thread-1] [HeartbeatReporter.java:133] [] - heartbeat end■ cost time:0.026s
2018-09-06 21:47:44,491 INFO  [pool-6-thread-1] [HeartbeatReporter.java:104] [] - heartbeat start, current status:2
2018-09-06 21:47:44,515 INFO  [pool-6-thread-1] [HeartbeatReporter.java:133] [] - heartbeat end■ cost time:0.024s
2018-09-06 21:47:49,516 INFO  [pool-6-thread-1] [HeartbeatReporter.java:104] [] - heartbeat start, current status:2
2018-09-06 21:47:49,538 INFO  [pool-6-thread-1] [HeartbeatReporter.java:133] [] - heartbeat end■ cost time:0.022s
2018-09-06 21:47:54,539 INFO  [pool-6-thread-1] [HeartbeatReporter.java:104] [] - heartbeat start, current status:2
2018-09-06 21:47:54,555 INFO  [pool-6-thread-1] [HeartbeatReporter.java:133] [] - heartbeat end■ cost time:0.016s
^C
```

b. 新建数据源

DataWorks新建工作空间后，默认设置自己为数据源odps_first。因此只需要添加Hadoop集群数据源。更多详情请参见#unique_40。

A. 进入数据集成页面，选择同步资源管理 > 数据源，单击新增数据源。

B. 在新增数据源弹出框中，选择数据源类型为HDFS。



C. 填写HDFS数据源的各配置项。

配置	说明
数据源名称	数据源名称必须以字母、数字、下划线组合，且不能以数字和下划线开头。
数据源描述	对数据源进行简单描述，不得超过80个字符。
适用环境	可以选择开发或生产环境。 说明： 仅标准模式工作空间会显示此配置。

配置	说明
DefaultFS	对于EMR Hadoop集群而言，如果Hadoop集群为HA集群，则此处地址为`hdfs://emr-header-1的IP:8020`。如果Hadoop集群为非HA集群，则此处地址为`hdfs://emr-header-1的IP:9000`。 本实验中的emr-header-1与DataWorks通过公网连接，因此此处填写公网IP并放通安全组。

D. 完成配置后，单击测试连通性。

E. 测试连通性通过后，单击完成。

说明:

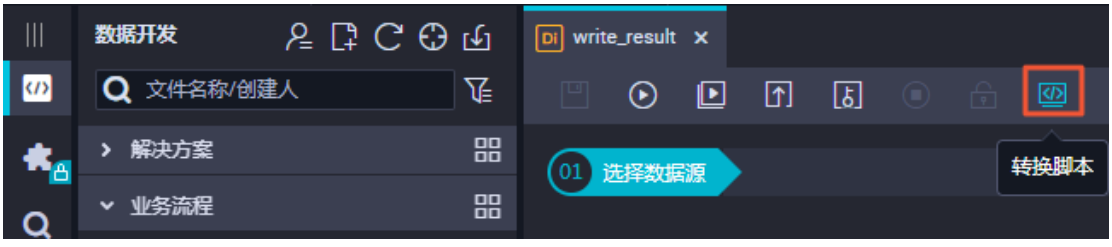
如果EMR Hadoop集群设置网络类型为专有网络，则不支持连通性测试。

c. 配置数据同步任务

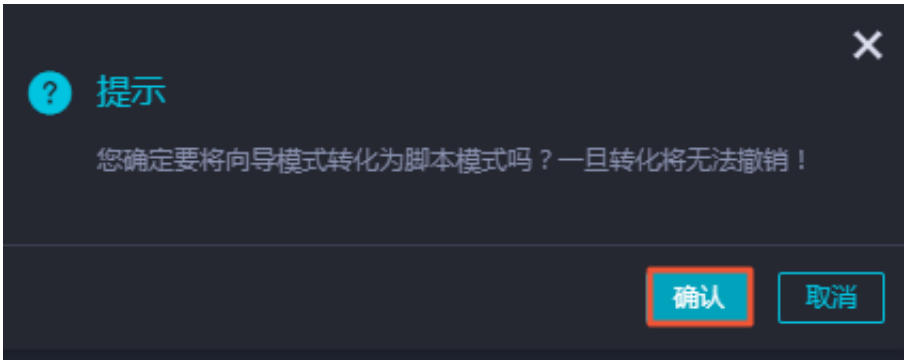
A. 进入数据开发页面，选择新建 > 数据集成 > 数据同步。

B. 在新建节点对话框中，输入节点名称，单击提交。

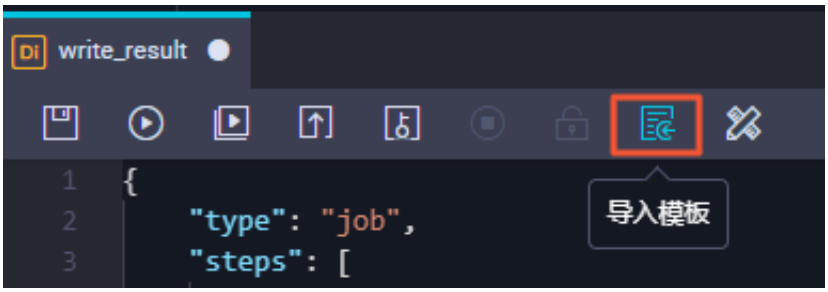
C. 成功创建数据同步节点后，单击工具栏中的转换脚本按钮。



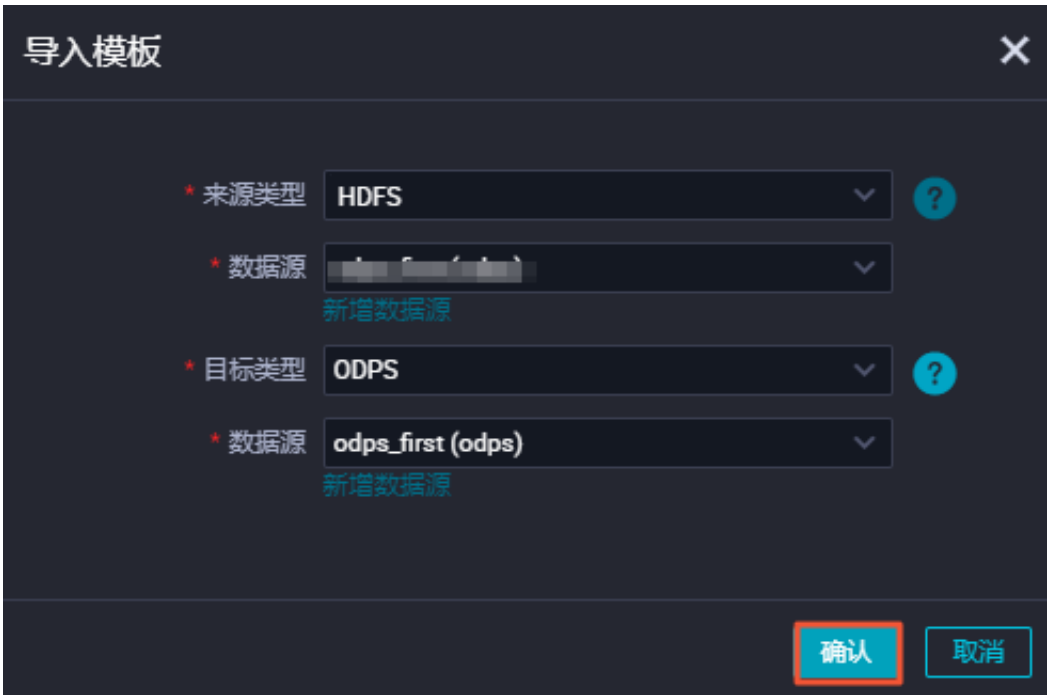
D. 单击提示对话框中的确认，即可进入脚本模式进行开发。



E. 单击工具栏中的导入模板按钮。



F. 在导入模板对话框中，选择来源类型、数据源、目标类型及数据源，单击确认。



G. 新建同步任务完成后，通过导入模板已生成了基本的读取端配置。此时您可以继续手动配置数据同步任务的读取端数据源，以及需要同步的表信息等。本示例的代码如下所示，更多详情请参见#unique_41。

```
{
  "configuration": {
    "reader": {
      "plugin": "hdfs",
      "parameter": {
        "path": "/user/hive/warehouse/hive_doc_good_sale/",
```

```
        "datasource": "HDFS1",
        "column": [
          {
            "index": 0,
            "type": "string"
          },
          {
            "index": 1,
            "type": "string"
          },
          {
            "index": 2,
            "type": "string"
          },
          {
            "index": 3,
            "type": "string"
          },
          {
            "index": 4,
            "type": "long"
          },
          {
            "index": 5,
            "type": "double"
          },
          {
            "index": 6,
            "type": "long"
          }
        ],
        "defaultFS": "hdfs://121.199.11.138:9000",
        "fieldDelimiter": ",",
        "encoding": "UTF-8",
        "fileType": "text"
      }
    },
    "writer": {
      "plugin": "odps",
      "parameter": {
        "partition": "pt=1",
        "truncate": false,
        "datasource": "odps_first",
        "column": [
          "create_time",
          "category",
          "brand",
          "buyer_id",
          "trans_num",
          "trans_amount",
          "click_cnt"
        ],
        "table": "hive_doc_good_sale"
      }
    },
    "setting": {
      "errorLimit": {
        "record": "1000"
      },
      "speed": {
        "throttle": false,
        "concurrent": 1,
        "mbps": "1",
      }
```

```
    }
  },
  "type": "job",
  "version": "1.0"
}
```

其中，path参数为数据在Hadoop集群中存放的位置.您可以在登录master node后，执行hdfs dfs -ls /user/hive/warehouse/hive_doc_good_sale命令确认。对于分区表，您可以不指定分区，DataWorks数据同步会自动递归到分区路径。

```
[root@emr-header-1 logs]# hdfs dfs -ls /user/hive/warehouse/hive_doc_good_sale/
Found 1 items
drwxr-x--x   - hive hadoop          0 2018-09-03 17:46 /user/hive/warehouse/hive_doc_good_sale/pt=1
```

H. 完成配置后，单击运行。如果提示任务运行成功，则说明同步任务已完成。如果运行失败，可以通过日志进行排查。

预期结果

1. 单击左侧导航栏中的临时查询。
2. 选择新建 > ODPS SQL。

3. 编写并执行SQL语句，查看导入hive_doc_good_sale的数据。SQL语句如下所示：

```
--查看是否成功写入MaxCompute
select * from hive_doc_good_sale where pt=1;
```

您也可以在odpscmd命令行工具中输入select * FROM hive_doc_good_sale where pt =1;，查询表结果。

如果您想实现MaxCompute数据迁移至Hadoop，步骤与上述步骤类似，不同的是同步脚本内的reader和writer对象需要对调，具体实现脚本如下。

```
{
  "configuration": {
    "reader": {
      "plugin": "odps",
      "parameter": {
      "partition": "pt=1",
      "isCompress": false,
      "datasource": "odps_first",
```

```
      "column": [
        "create_time",
        "category",
        "brand",
      "buyer_id",
      "trans_num",
      "trans_amount",
      "click_cnt"
    ],
    "table": "hive_doc_good_sale"
    }
  },
  "writer": {
    "plugin": "hdfs",
    "parameter": {
    "path": "/user/hive/warehouse/hive_doc_good_sale",
    "fileName": "pt=1",
    "datasource": "HDFS_data_source",
    "column": [
      {
        "name": "create_time",
        "type": "string"
      },
      {
        "name": "category",
        "type": "string"
      },
      {
        "name": "brand",
        "type": "string"
      },
      {
        "name": "buyer_id",
        "type": "string"
      },
      {
        "name": "trans_num",
        "type": "BIGINT"
      },
      {
        "name": "trans_amount",
        "type": "DOUBLE"
      },
      {
        "name": "click_cnt",
        "type": "BIGINT"
      }
    ],
    "defaultFS": "hdfs://47.99.162.100:9000",
    "writeMode": "append",
    "fieldDelimiter": ",",
    "encoding": "UTF-8",
    "fileType": "text"
    }
  },
  "setting": {
    "errorLimit": {
      "record": "1000"
  },
  "speed": {
    "throttle": false,
    "concurrent": 1,
    "mbps": "1",
  }
```

```
    }
},
"type": "job",
"version": "1.0"
}
```

您需要参见#unique_42，在运行上述同步任务前，对Hadoop集群进行设置。在运行同步任务后，手动复制同步过去的文件。

2.4 Kafka数据迁移MaxCompute最佳实践

本文将为您介绍如何使用DataWorks数据同步功能，将Kafka集群上的数据迁移至阿里云大数据计算服务MaxCompute。

前提条件

- 搭建Kafka集群

  进行数据迁移前，您需要保证自己的Kafka集群环境正常。本文使用阿里云EMR服务自动化搭建Kafka集群，详细过程请参见Kafka快速入门。

  本文使用的EMR Kafka版本信息如下：

  - EMR版本: EMR-3.12.1
  - 集群类型: Kafka
  - 软件信息: Ganglia 3.7.2 ZooKeeper 3.4.12 Kafka 2.11-1.0.1 Kafka-Manager 1.3.3.16

  Kafka集群使用专有网络，区域为华东1（杭州），主实例组ECS计算资源配置公网及内网IP，具体配置如下图所示。

- 创建MaxCompute项目

  开通MaxCompute服务并创建好项目，本文中在华东1（杭州）区域创建项目bigdata_DOC，同时启动DataWorks相关服务，如下图所示。详情请参见#unique_36。

背景信息

Kafka是一款分布式发布与订阅的消息中间件，具有高性能、高吞量的特点被广泛使用，每秒能处理上百万的消息。Kafka适用于流式数据处理，主要应用于用户行为跟踪、日志收集等场景。

一个典型的Kafka集群包含若干个生产者（Producer）、Broker、消费者（Consumer）以及一个Zookeeper集群。Kafka集群通过Zookeeper管理自身集群的配置并进行服务协同。

Topic是Kafka集群上最常用的消息的集合，是一个消息存储逻辑概念。物理磁盘不存储Topic，而是将Topic中具体的消息按分区（Partition）存储在集群中各个节点的磁盘上。每个Topic可以有多个生产者向它发送消息，也可以有多个消费者向它拉取（消费）消息。

每个消息被添加到分区时，会分配一个offset（偏移量，从0开始编号），是消息在一个分区中的唯一编号。

操作步骤

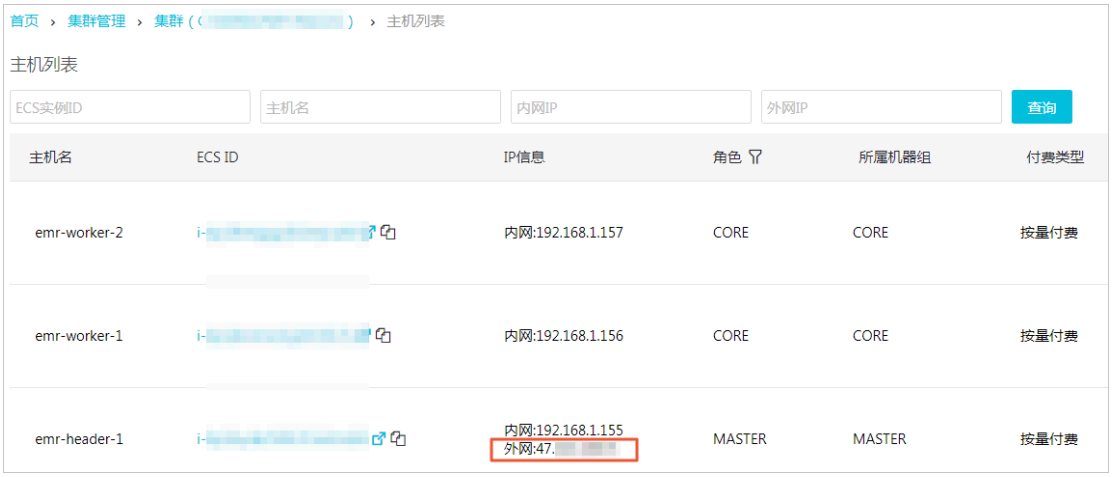
1. 准备测试表与数据

a) Kafka集群创建测试数据

为保证您可以顺利登录EMR集群Header主机及MaxCompute和DataWorks可以顺利和EMR集群Header主机通信，请您首先配置EMR集群Header主机安全组，放行TCP 22及TCP 9092端口。

A. 登录EMR集群Header主机地址

进入EMR Hadoop控制台集群管理 > 主机列表页面，确认EMR集群Header主机地址，并通过SSH连接远程登录。



B. 创建测试Topic

执行如下命令创建测试所使用的Topic testkafka。

```
[root@emr-header-1 ~]# kafka-topics.sh --zookeeper emr-header
-1:2181/kafka-1.0.1 --partitions 10 --replication-factor 3 --
topic testkafka  --create
Created topic "testkafka".
```

执行如下命令查看已创建的Topic。

```
[root@emr-header-1 ~]# kafka-topics.sh  --list --zookeeper emr-
header-1:2181/kafka-1.0.1
__consumer_offsets
_emr-client-metrics
_schemas
connect-configs
connect-offsets
connect-status
```

```
testkafka
```

C. 写入测试数据

您可以执行如下命令，模拟生产者向Topic testkafka中写入数据。由于Kafka用于处理流式数据，您可以持续不断的向其中写入数据。为保证测试结果，建议您写入10条以上的数据。

```
[root@emr-header-1 ~]# kafka-console-producer.sh --broker-list
emr-header-1:9092 --topic testkafka
>123
>abc
>
```

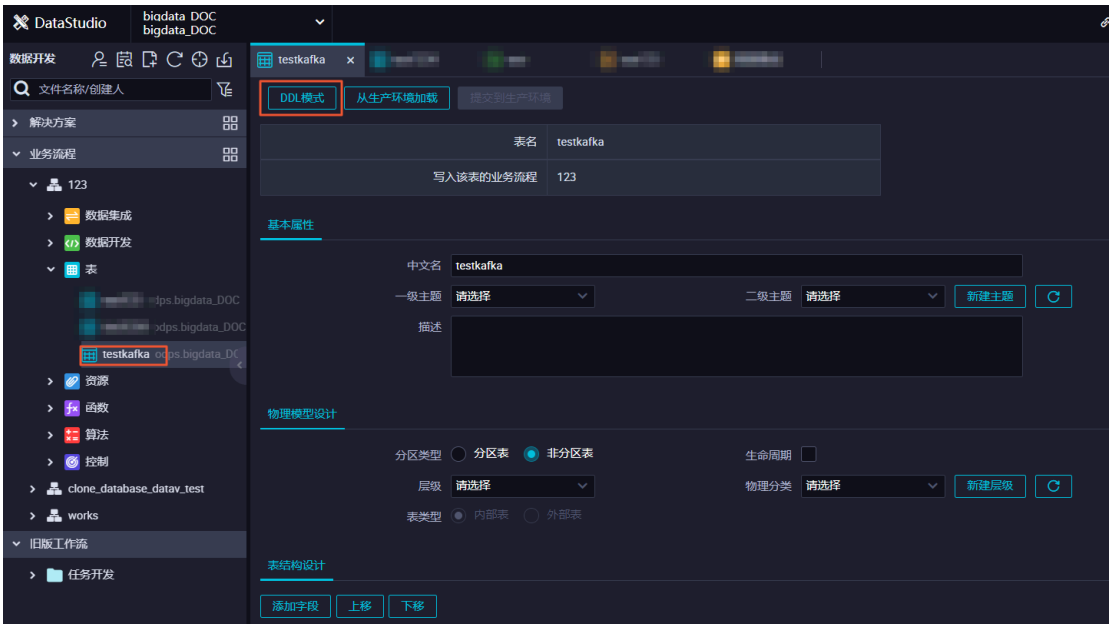
为验证写入数据生效，您可以同时再打开一个SSH窗口，执行如下命令，模拟消费者验证数据是否已成功写入Kafka。当数据写入成功时，您可以看到已写入的数据。

```
[root@emr-header-1 ~]# kafka-console-consumer.sh --bootstrap-
server emr-header-1:9092 --topic testkafka --from-beginning
123
abc
```

b) 创建MaxCompute表

为保证MaxCompute可以顺利接收Kafka数据，请您首先在MaxCompute上创建表。本例中为测试便利，使用非分区表。

A. 登录DataWorks创建表，详情请参见#unique_45。



您可以单击DDL模式进行建表，建表语句如下。

```
CREATE TABLE `testkafka` (
    `key` string,
    `value` string,
    `partition1` string,
    `timestamp1` string,
```

```
    `offset` string,
    `t123` string,
    `event_id` string,
    `tag` string
) ;
```

其中的每一列，对应于DataWorks数据集成Kafka Reader的默认列，您可以自主命名。详情请参见#unique_46：

A. __key__表示消息的key。

B. __value__表示消息的完整内容 。

C. __partition__表示当前消息所在分区。

D. __headers__表示当前消息headers信息。

E. __offset__表示当前消息的偏移量。

F. __timestamp__表示当前消息的时间戳。

2. 数据同步

a) 新建自定义资源组

由于当前DataWorks的默认资源组无法完美支持Kafka插件，您需要使用自定义资源组完成数据同步。自定义资源组详情请参见#unique_47。

在本文中，为节省资源，直接使用EMR集群Header主机作为自定义资源组。完成后，请等待服务器状态变为可用。

b) 新建并运行同步任务

A. 在您的业务流程中右键单击数据集成，选择新建数据集成节点 > 数据同步。

B. 新建数据同步节点后，您需要选择数据来源的数据源为Kafka，数据去向的数据源为ODPS，并且使用默认数据源odps_first。选择数据去向表为您新建的testkafka。完成上述配置后，请单击下图框中的按钮，转换为脚本模式。

C. 脚本配置如下，代码释义请参见#unique_48。

```
{
    "type": "job",
    "steps": [
        {
            "stepType": "kafka",
            "parameter": {
                "server": "47.xxx.xxx.xxx:9092",
                "kafkaConfig": {
                    "group.id": "console-consumer-83505"
                },
                "valueType": "ByteArray",
                "column": [
                    "__key__",
                    "__value__",
                    "__partition__",
                    "__timestamp__",
                    "__offset__",
                    "'123'",
                    "event_id",
                    "tag.desc"
                ],
                "topic": "testkafka",
                "keyType": "ByteArray",
                "waitTime": "10",
                "beginOffset": "0",
                "endOffset": "3"
            },
            "name": "Reader",
            "category": "reader"
        },
        {
            "stepType": "odps",
            "parameter": {
                "partition": "",
                "truncate": true,
                "compress": false,
                "datasource": "odps_first",
                "column": [
                    "key",
                    "value",
                    "partition1",
                    "timestamp1",
                    "offset",
                    "t123",
                    "event_id",
                    "tag"
                ],
```

```
                "emptyAsNull": false,
                "table": "testkafka"
            },
            "name": "Writer",
            "category": "writer"
        }
    ],
    "version": "2.0",
    "order": {
        "hops": [
            {
                "from": "Reader",
                "to": "Writer"
            }
        ]
    },
    "setting": {
        "errorLimit": {
            "record": ""
        },
        "speed": {
            "throttle": false,
            "concurrent": 1,
        }
    }
}
```

您可以通过在Header主机上使用kafka-consumer-groups.sh --bootstrap-server emr-header-1:9092 --list**命令查看**group.id**参数，及消费者的Group名称。**

```
[root@emr-header-1 ~]#  kafka-consumer-groups.sh  --bootstrap-
server emr-header-1:9092  --list
Note: This will not show information about old Zookeeper-based
consumers.

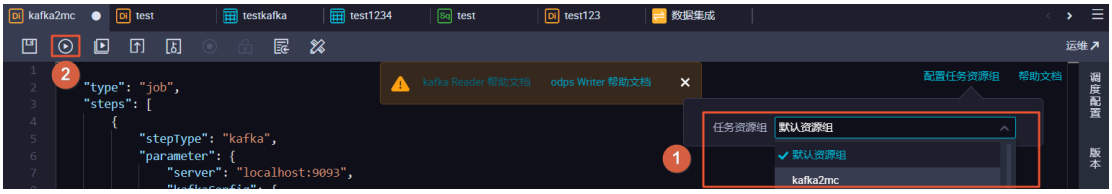
_emr-client-metrics-handler-group
console-consumer-69493
console-consumer-83505
console-consumer-21030
console-consumer-45322
console-consumer-14773
```

以console-consumer-83505**为例，您可以根据该参数在Header主机上使用**kafka-consumer-groups.sh --bootstrap-server emr-header-1:9092 --describe --group console-consumer-83505**命令确认**beginOffset**及**endOffset**参数。**

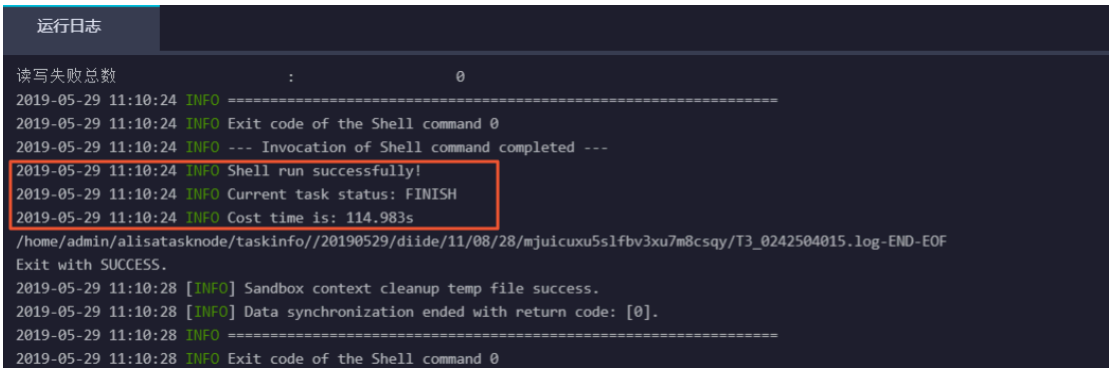
```
[root@emr-header-1 ~]# kafka-consumer-groups.sh --bootstrap-
server emr-header-1:9092 --describe --group console-consumer-
83505
Note: This will not show information about old Zookeeper-based
consumers.
Consumer group 'console-consumer-83505' has no active members.
TOPIC                          PARTITION  CURRENT-OFFSET  LOG-
END-OFFSET  LAG        CONSUMER-ID
          HOST                           CLIENT-ID
```

```
testkafka                      6          0          0
          0          -                              -
test                           6          3          3
          0          -                              -
testkafka                      0          0          0
          0          -                              -
testkafka                      1          1          1
          0          -                              -
testkafka                      5          0          0
          0          -                              -
```

完成脚本配置后，请首先切换任务资源组为您刚创建的资源组，然后单击运行。

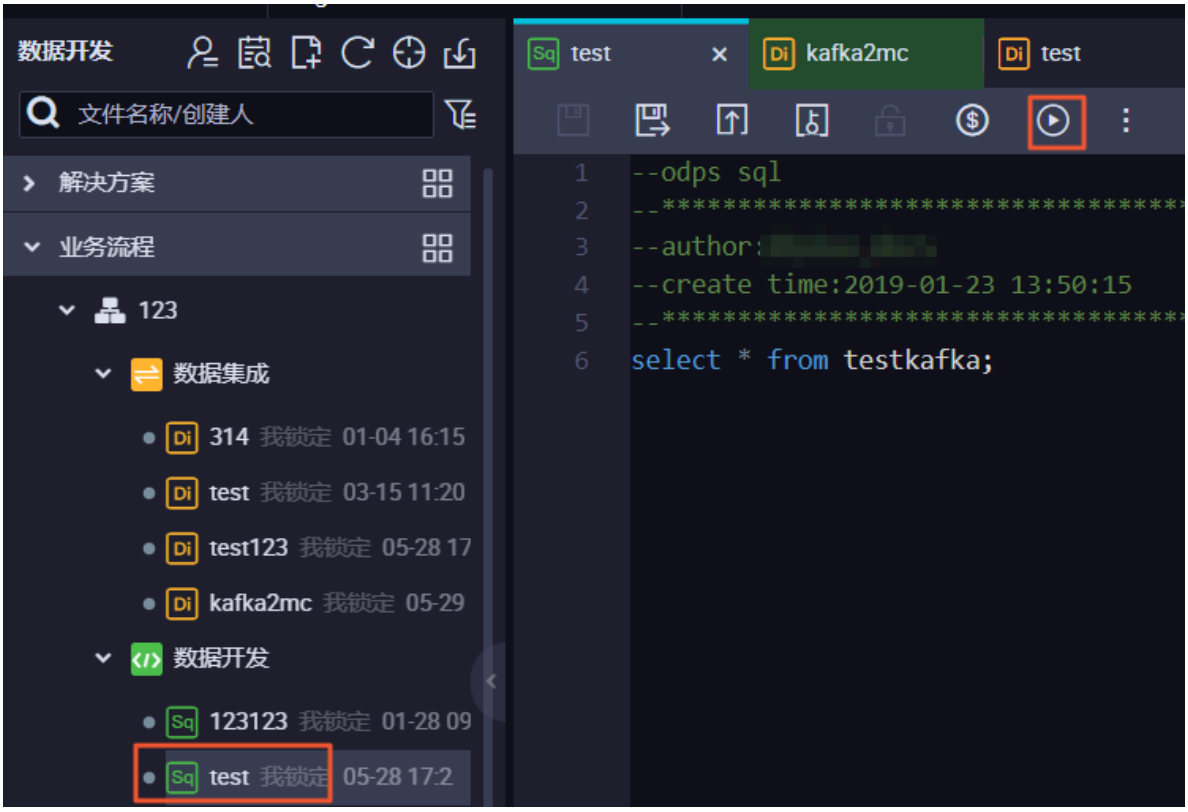


D. 完成运行后，您可以在运行日志中查看运行结果，如下为成功运行的日志。

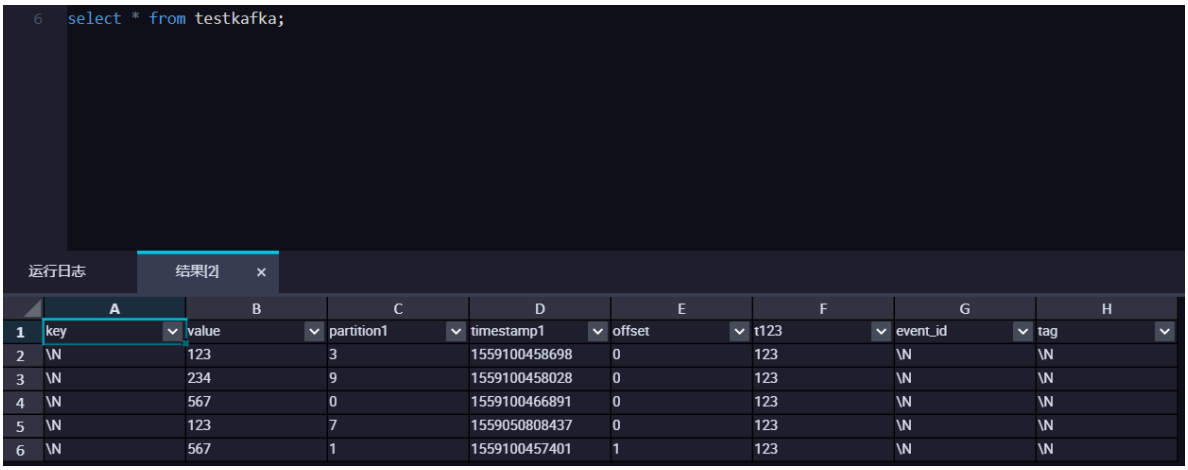


3. 结果验证

您可以通过新建一个数据开发任务运行SQL语句，查看当前表中是否已存在从Kafka同步过来的数据。本例中使用`select * from testkafka;`语句，完成后单击运行即可。



执行结果如下，本例中为保证结果，在testkafka Topic中输入了多条数据，您可以查验是否和您输入的数据一致。



	A	B	C	D	E	F	G	H
1	key	value	partition1	timestamp1	offset	t123	event_id	tag
2	\N	123	3	1559100458698	0	123	\N	\N
3	\N	234	9	1559100458028	0	123	\N	\N
4	\N	567	0	1559100466891	0	123	\N	\N
5	\N	123	7	1559050808437	0	123	\N	\N
6	\N	567	1	1559100457401	1	123	\N	\N

2.5 Elasticsearch数据迁移至MaxCompute

本文将为您介绍如何通过DataWorks数据同步功能，将阿里云Elasticsearch集群上的数据迁移至MaxCompute。

前提条件

- 搭建阿里云Elasticsearch集群

  进行数据迁移前，您需要保证自己的阿里云Elasticsearch集群环境正常。搭建阿里云Elasticsearch集群的详细过程请参见Elasticsearch快速入门。

  本示例中阿里云Elasticsearch的具体配置如下：

  - 地域：华东2（上海）
  - 可用区：上海可用区B
  - 版本：5.5.3 with Commercial Feature

- 创建MaxCompute项目

  开通MaxCompute服务并创建好项目，详情请参见#unique_36。本示例中在华东1（杭州）区域创建项目bigdata_DOC，同时启动DataWorks相关服务。

背景信息

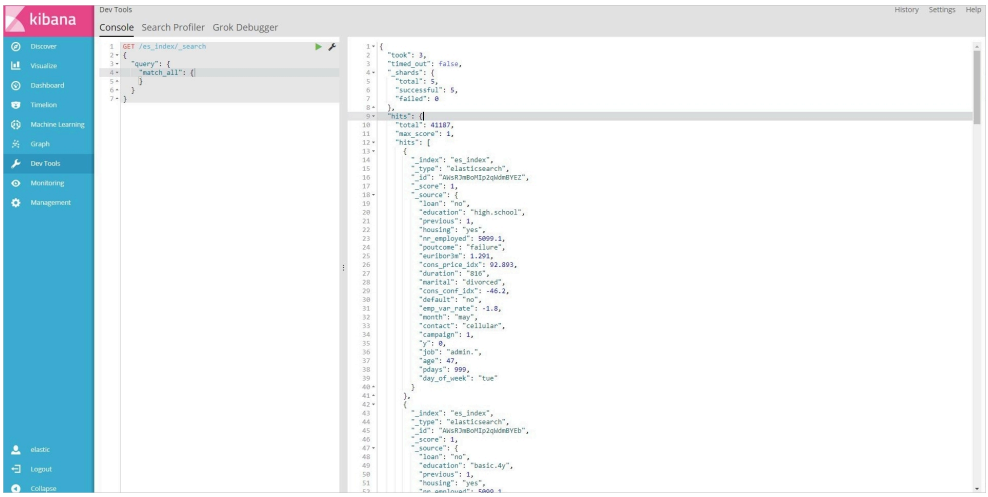
Elasticsearch是一个基于Lucene的搜索服务器，它提供了一个分布式多用户功能的全文搜索引擎。Elasticsearch是遵从Apache开源条款的一款开源产品，是当前主流的企业级搜索引擎。

阿里云Elasticsearch提供Elasticsearch 5.5.3 with Commercial Feature、6.3.2 with Commercial Feature、6.7.0 with Commercial Feature及商业插件X-pack服务，致力于数据分析、数据搜索等场景服务。在开源Elasticsearch基础上提供企业级权限管控、安全监控告警、自动报表生成等功能。

操作步骤

1. 创建测试表

   将阿里云上的数据导入至阿里云Elasticsearch（离线），具体操作请参见云上数据导入。

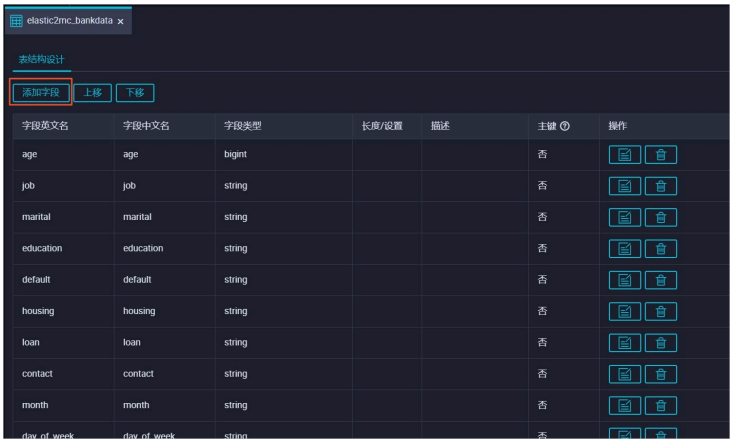


2. 创建接收表

   为保证MaxCompute可以顺利接收Elasticsearch数据，您需要在MaxCompute上创建表。本示例使用非分区表。

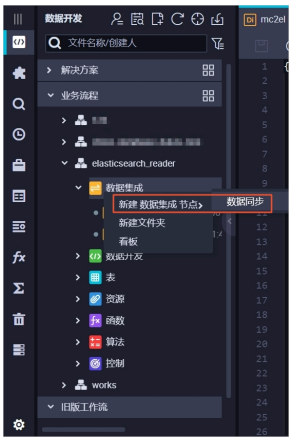
   a) 登录DataWorks控制台进行创建表的操作，详情请参见#unique_45。

   b) 为表添加字段。请确保添加的字段与需要被同步的Elasticsearch表字段相对应。



3. 同步数据

a) 在业务流程中，右键单击数据集成，选择新建数据集成节点 > 数据同步。



b) 单击下图框中的按钮，转换为脚本模式。

c) 脚本配置如下所示。代码释义请参见#unique_50。

```
{
    "type": "job",
    "steps": [
        {
            "stepType": "elasticsearch",
            "parameter": {
                "retryCount": 3,
                "column": [
                    "age",
                    "job",
                    "marital",
                    "education",
                    "default",
                    "housing",
                    "loan",
                    "contact",
                    "month",
                    "day_of_week",
                    "duration",
                    "campaign",
                    "pdays",
                    "previous",
                    "poutcome",
                    "emp_var_rate",
                    "cons_price_idx",
                    "cons_conf_idx",
                    "euribor3m",
                    "nr_employed",
                    "y"
```

```
                ],
                "scroll": "1m",
                "index": "es_index",
                "pageSize": 1,
                "sort": {
                    "age": "asc"
},
                "type": "elasticsearch",
                "connTimeOut": 1000,
                "retrySleepTime": 1000,
                "endpoint": "http://es-cn-xxxx.xxxx.xxxx.xxxx.com:
9200",
                "password": "xxxx",
                "search": {
                    "match_all": {}
                },
                "readTimeOut": 5000,
                "username": "xxxx"
            },
            "name": "Reader",
            "category": "reader"
        },
        {
            "stepType": "odps",
            "parameter": {
                "partition": "",
                "truncate": true,
                "compress": false,
                "datasource": "odps_first",
                "column": [
                    "age",
                    "job",
                    "marital",
                    "education",
                    "default",
                    "housing",
                    "loan",
                    "contact",
                    "month",
                    "day_of_week",
                    "duration",
                    "campaign",
                    "pdays",
                    "previous",
                    "poutcome",
                    "emp_var_rate",
                    "cons_price_idx",
                    "cons_conf_idx",
                    "euribor3m",
                    "nr_employed",
                    "y"
                ],
                "emptyAsNull": false,
                "table": "elastic2mc_bankdata"
            },
            "name": "Writer",
            "category": "writer"
        }
    ],
    "version": "2.0",
    "order": {
        "hops": [
            {
                "from": "Reader",
```

```
                "to": "Writer"
            }
        ]
    },
    "setting": {
        "errorLimit": {
            "record": "0"
        },
        "speed": {
            "throttle": false,
            "concurrent": 1,
            "dmu": 1
        }
    }
}
```

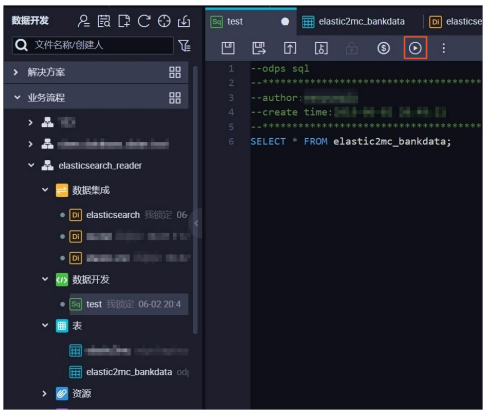
您可以在创建的阿里云Elasticsearch集群的基本信息中，查看公网地址和公网端口，并将这些信息填入上面的代码中。

d) 完成脚本配置后，单击运行。

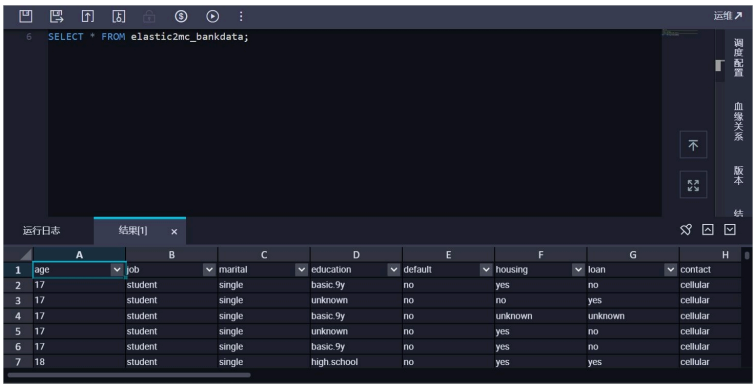
e) 在运行日志中查看运行结果。

4. 验证结果

a) 新建一个数据开发任务运行SQL语句，查看当前表中是否已存在从Elasticsearch同步过来的数据。



b) 执行结果如下所示。



2.6 RDS迁移至MaxCompute实现动态分区

本文为您介绍如何使用DataWorks数据集成同步功能自动创建分区，动态地将RDS中的数据迁移至MaxCompute大数据计算服务。

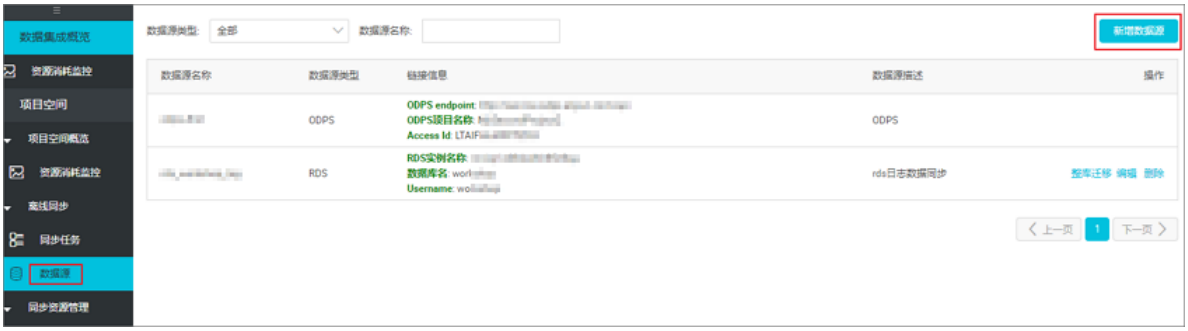
准备工作

1. #unique_36服务并创建工作空间。本文选择的区域是华北2（北京）。

> **说明:**
> 如果您是第一次使用DataWorks，请确认已经根据准备工作，准备好账号和项目角色、项目空间等。

2. 新增数据源。

新增RDS数据源作为数据来源；同时需要新增ODPS数据源作为目标数据源接收RDS数据。配置完成后，在数据集成控制台单击数据源可以看到新增的数据源。



自动创建分区

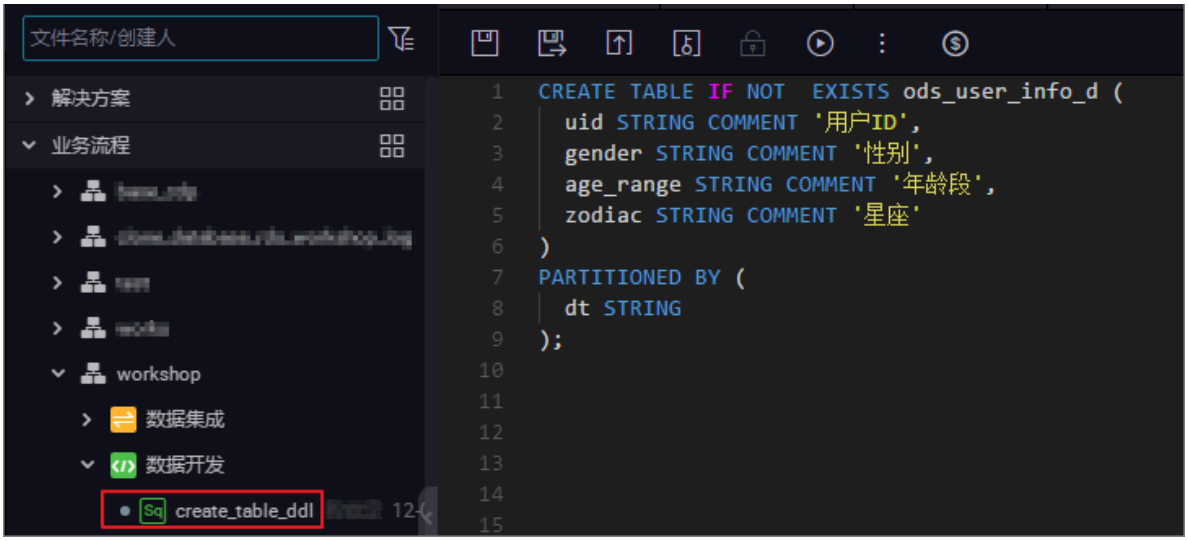
准备工作完成后，需要将RDS中的数据定时每天同步到MaxCompute中，自动创建按天日期的分区。详细的数据同步任务的操作和配置请参见DataWorks数据开发和运维。

1. 创建目标表。

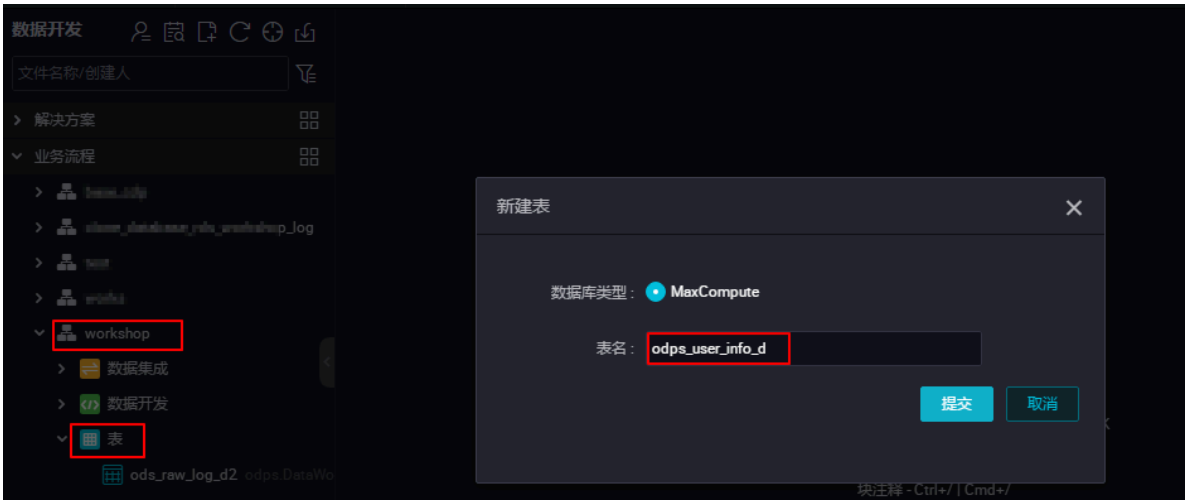
在ODPS数据库中创建RDS对应的目标表ods_user_info_d。在控制台数据开发选项下右键单击新建ODPS SQL节点创建create_table_ddl表，输入建表语句，如下所示。

```
CREATE TABLE IF NOT EXISTS ods_user_info_d (
uid STRING COMMENT '用户ID',
gender STRING COMMENT '性别',
age_range STRING COMMENT '年龄段',
zodiac STRING COMMENT '星座'
)
PARTITIONED BY (
dt STRING
```

```
);
```



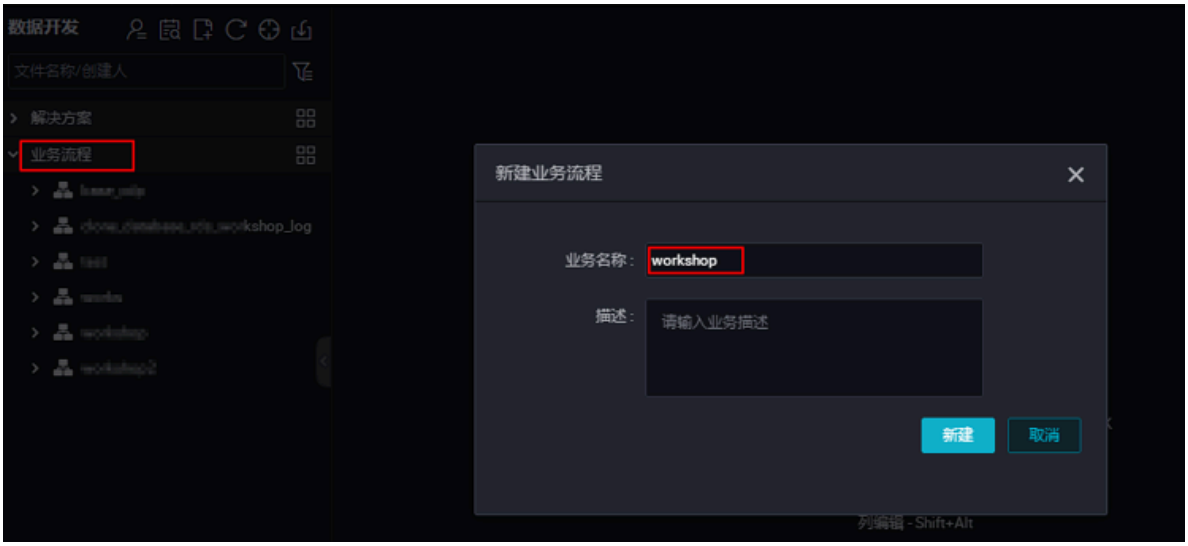
您也可以在业务流程 > 表下选择新建表。



详细的信息请参见建表并上传数据。

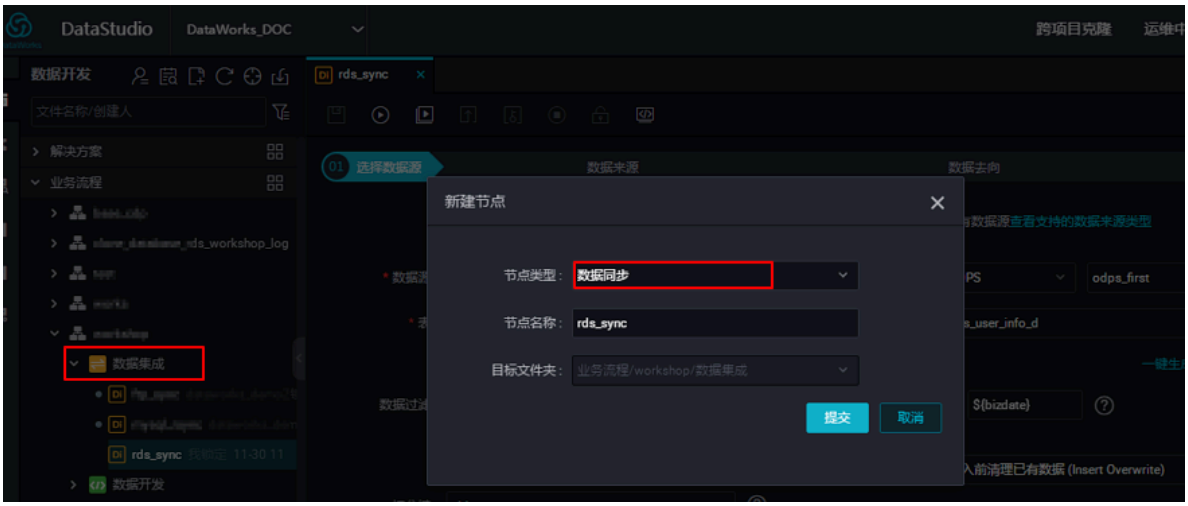
2. 新建业务流程。

   进入DataWorks管理控制台，单击对应项目后的进入数据开发，开始数据开发操作。选择数据开发 > 业务流程下的新建业务流程workshop。

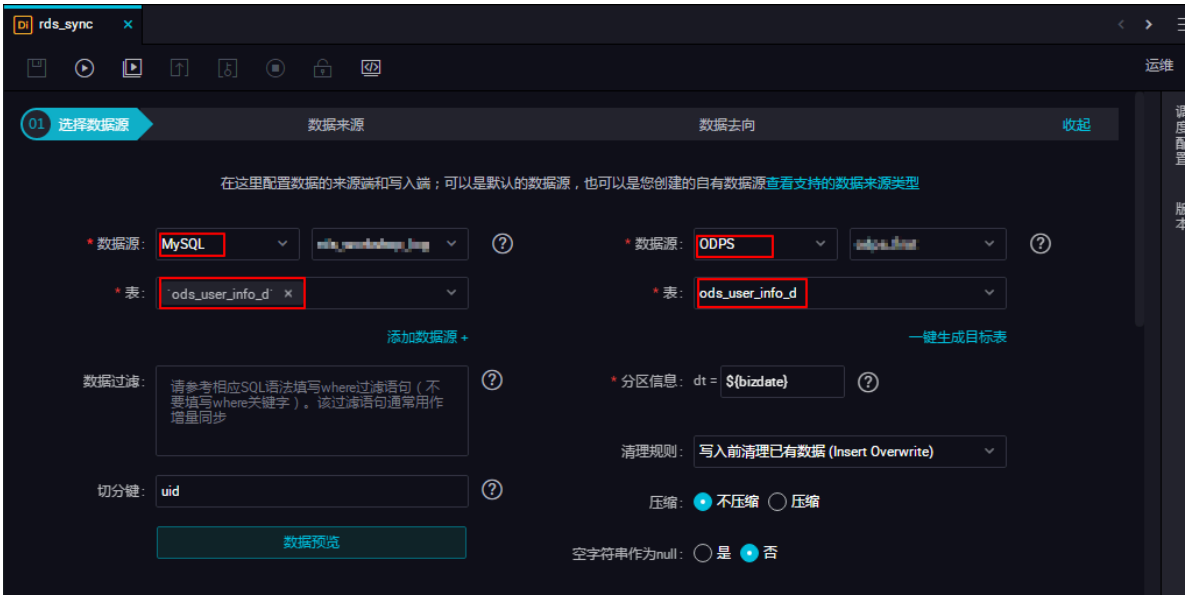


3. 新建并配置同步任务节点。

   在新创建的业务流程workshop下，新建同步节点rds_sync。

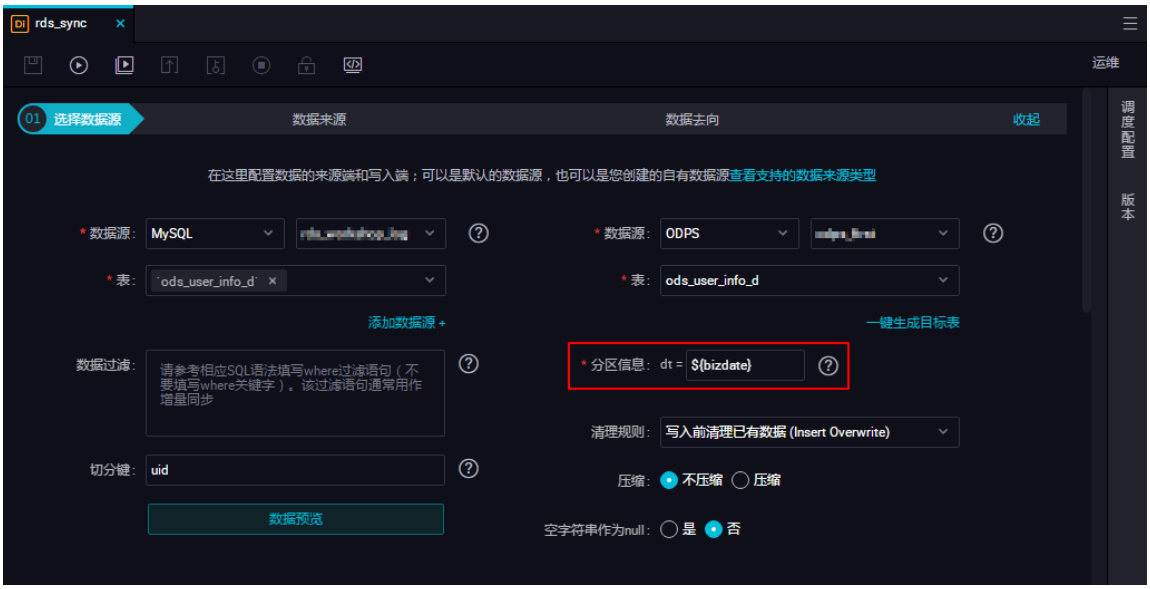


4. 分别选择数据来源和数据去向。

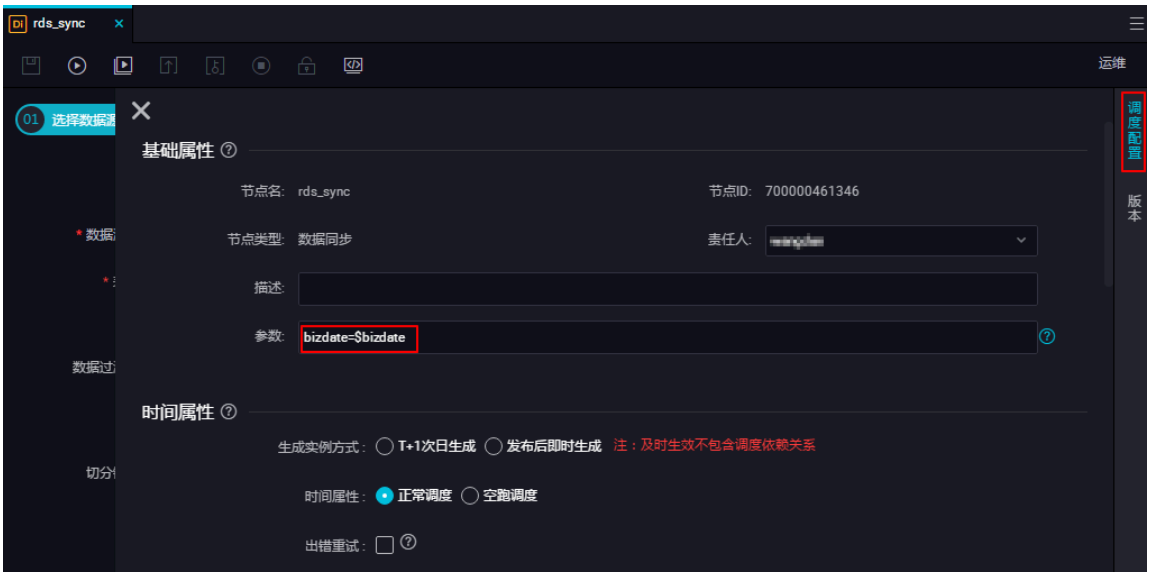


5. 配置分区参数。

a. 单击右侧调度配置弹出基础属性页面。



b. 参数值默认为系统自带的时间参数：`${bizdate}`，格式为yyyymmdd。



说明:

默认参数值与数据去向中的分区信息值对应。调度执行迁移任务时，目标表的分区值会被自动替换为任务执行日期的前一天，默认情况下，您会在当前执行前一天的业务数据，这个日期也叫做业务日期。如果您需要使用当天任务运行的日期作为分区值，则需自定义参数值。

自定义参数设置：用户可以自主选择某一天和格式配置，如下所示。

- 后N年：`$[add_months(yyyymmdd,12*N)]`
- 前N年：`$[add_months(yyyymmdd,-12*N)]`
- 前N月：`$[add_months(yyyymmdd,-N)]`
- 后N周：`$[yyyymmdd+7*N]`
- 后N月：`$[add_months(yyyymmdd,N)]`
- 前N周：`$[yyyymmdd-7*N]`
- 后N天：`$[yyyymmdd+N]`
- 前N天：`$[yyyymmdd-N]`
- 后N小时：`$[hh24miss+N/24]`
- 前N小时：`$[hh24miss-N/24]`
- 后N分钟：`$[hh24miss+N/24/60]`
- 前N分钟：`$[hh24miss-N/24/60]`

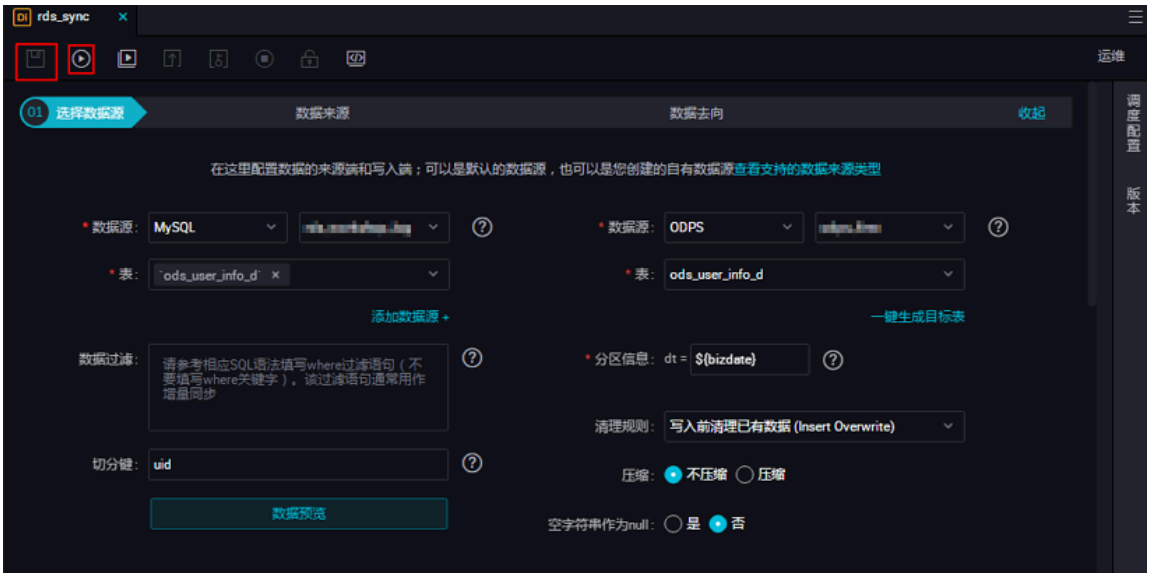
说明：

- 请以中括号[]编辑自定义变量参数的取值计算公式，例如 `key1=$[yyyy-mm-dd]`。
- 默认情况下，自定义变量参数的计算单位为天。例如 `$[hh24miss-N/24/60]` 表示 `(yyyymmddhh24miss-(N/24/60 * 1天))` 的计算结果，然后按 hh24miss 的格式取时分秒。
- 使用add_months的计算单位为月。例如 `$[add_months(yyyymmdd,12 N)-M/24/60]` 表示 `(yyyymmddhh24miss-(12 * N * 1月))-(M/24/60 * 1天)` 的结果，然后按 `yyyymmdd` 的格式取年月日。

详细的参数设置请参见参数配置。

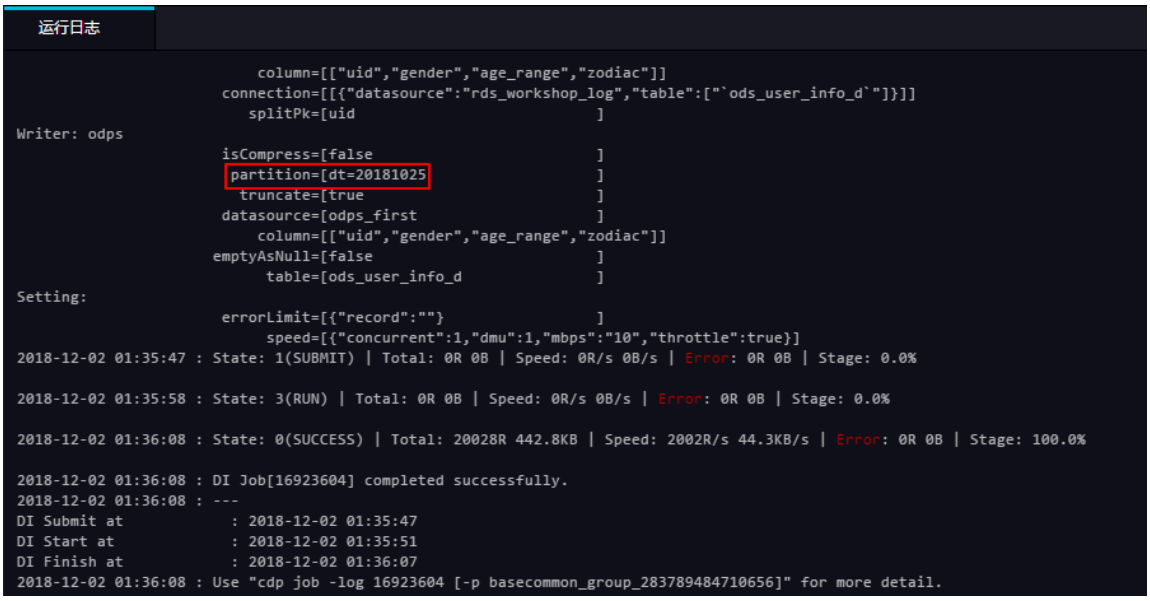
6. 测试运行。

a. 保存所有配置，单击运行。



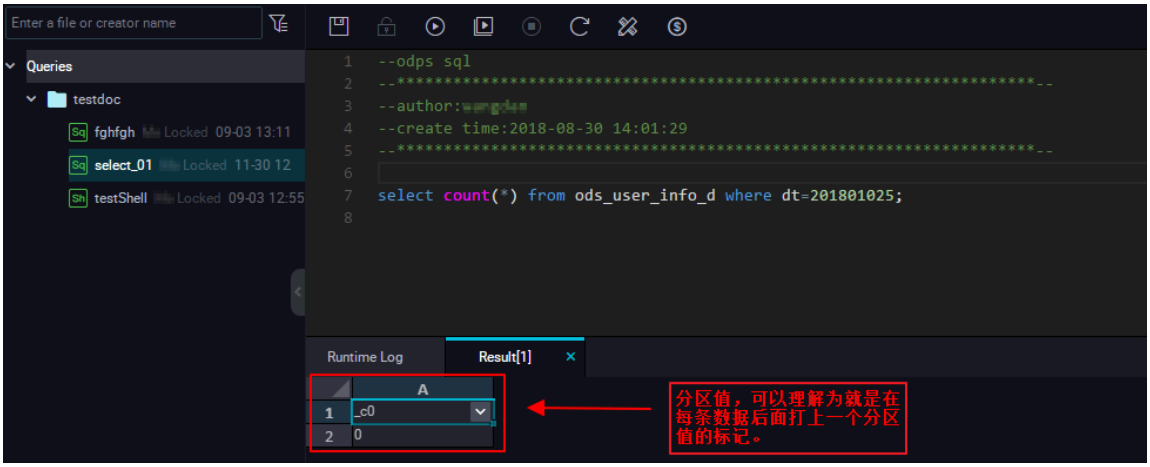
b. 查看运行日志。

日志中显示partition分区值dt=20181025已自动替换成功。



c. 验证实际的数据是否已转移到ODPS表中。

此时可看到数据已经迁移到ODPS表中，并且成功创建了一个分区值。这个任务在执行定时调度时，会将RDS中的数据每天同步至MaxCompute中的按照日期自动创建的分区里。



说明:

在MaxCompute2.0中分区表查询需要添加分区筛选，不支持全量查询，SQL语句如下。select命令详情请参见Select操作。

```
--查看是否成功写入MaxCompute。
select count(*) from ods_user_info_d where dt=业务日期;
```

补数据实验

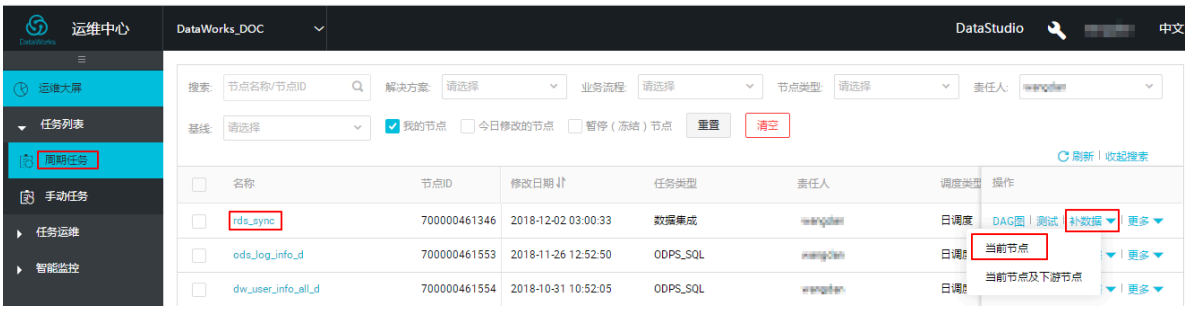
如果您的数据中存在大量运行日期之前的历史数据，想要实现自动同步和自动分区，您可以进入DataWorks的运维中心，选择当前的同步数据节点，选择补数据功能来实现。

1. 在RDS端按照日期筛选出历史数据。

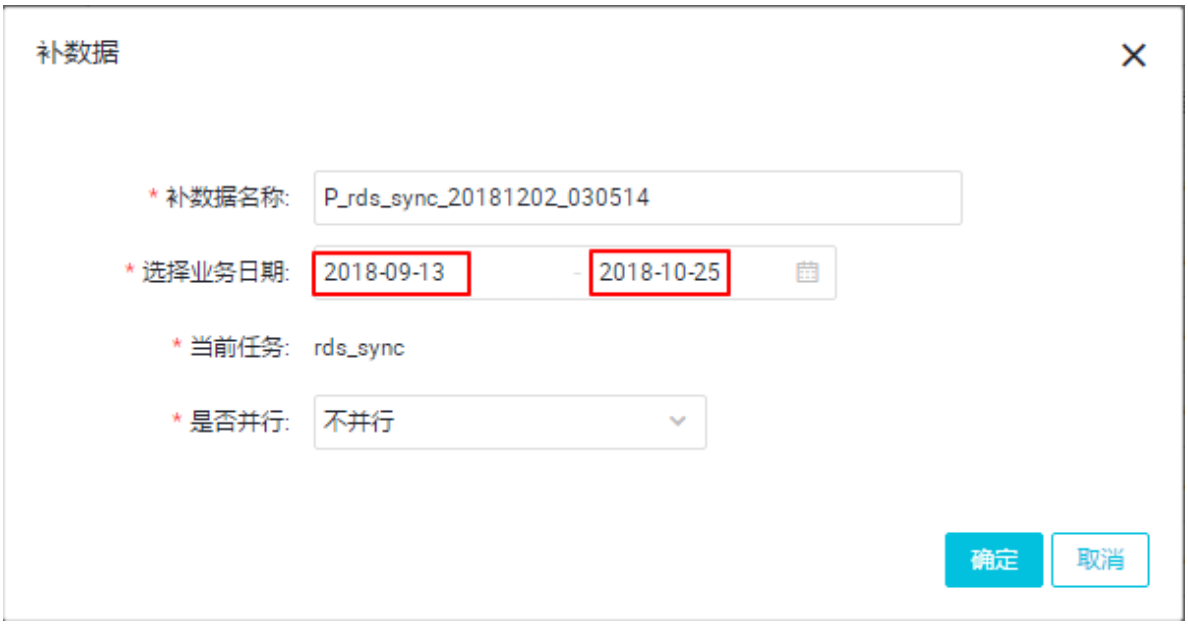
   您可以在迁移阶段设置where过滤条件。示例为将2018-09-13日的历史数据自动同步到MaxCompute的20181025的分区中。

2. 补数据操作。

   单击保存 > 提交。提交后到运维中心 > 任务列表 > 周期任务中的rds_sync节点，单击补数据 > 当前节点。



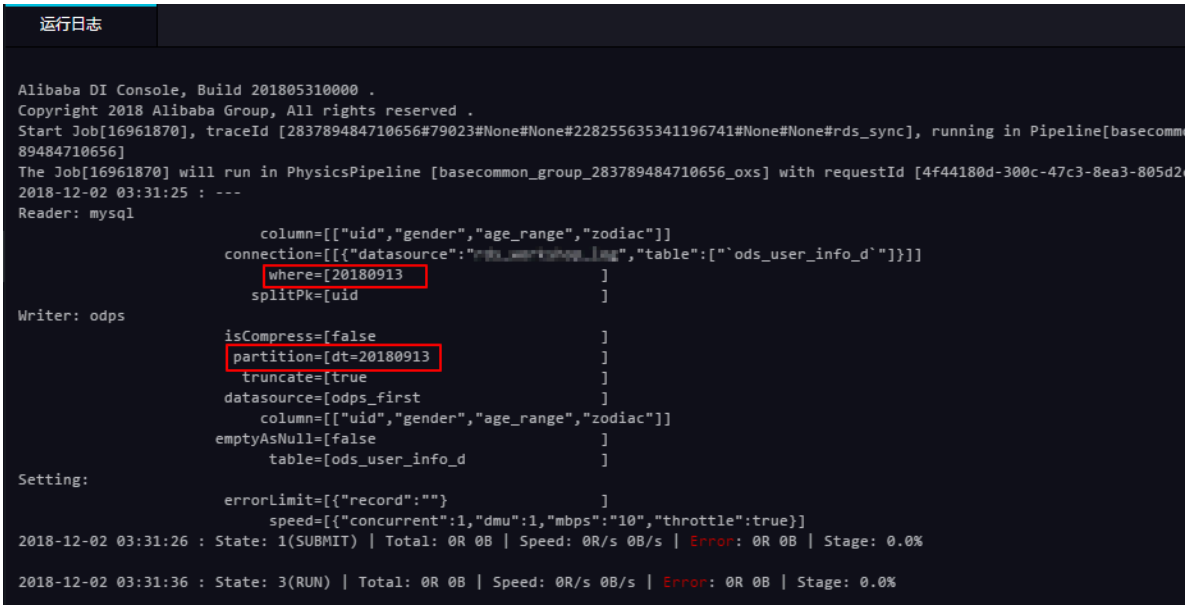
3. 跳转至补数据节点页面，选择日期区间，单击确定。



4. 此时会同时生成多个同步的任务实例，将按顺序执行。

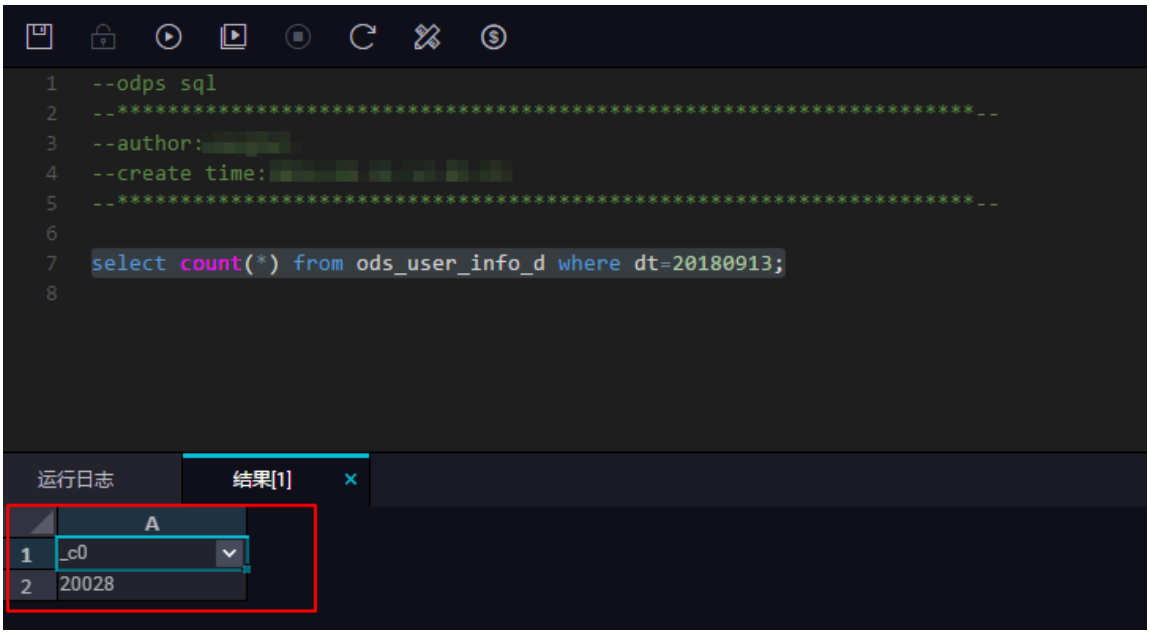
5. 在运行的日志中查看对RDS数据的抽取结果。

   可以看到MaxCompute已自动创建分区。



6. 查看运行结果。

a. 查看数据写入的情况，是否自动创建了分区，数据是否已同步到分区表中。



b. 查询对应分区信息。

说明:

在MaxCompute2.0中分区表查询需要添加分区筛选，SQL语句如下。其中分区列需要更新为业务日期，例如任务运行的日期为20180717，业务日期将为20180716。

```
--查看是否成功写入MaxCompute。
select count(*) from ods_user_info_d where dt=业务日期;
```

Hash实现非日期字段分区

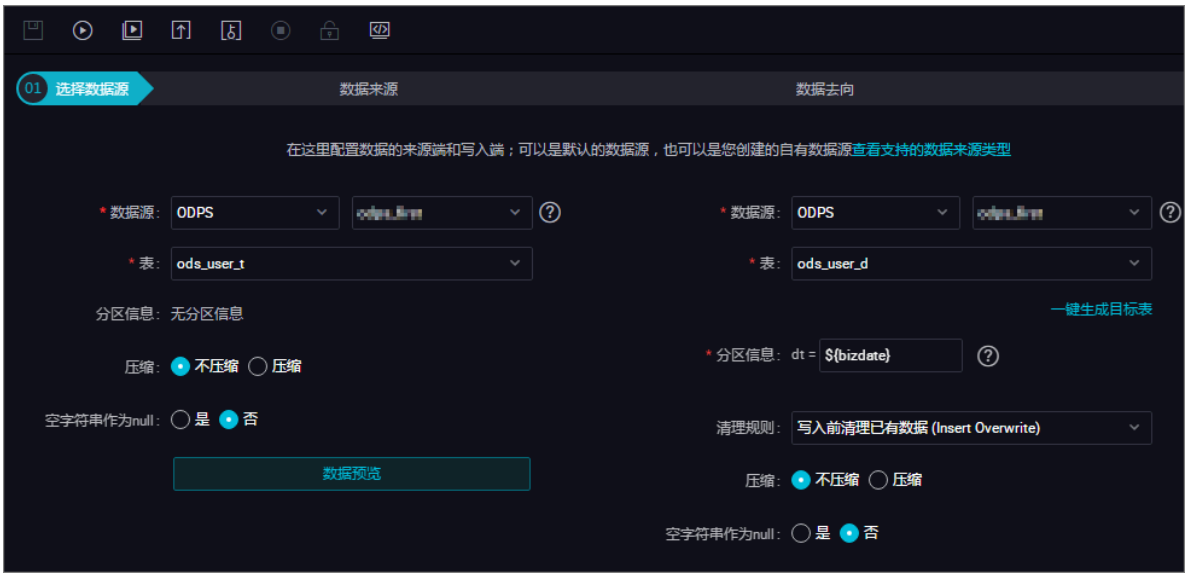
如果用户数据量较大，或是没有按照日期字段对第一次全量的数据进行分区，而是按照省份等非日期字段分区，则此时数据集成操作将不能实现自动分区。这种情况下，您可以按照RDS中某个字段进行Hash，将相同的字段值自动存放到这个字段对应值的MaxCompute分区中。

步骤如下：

1. 将数据全量同步到MaxCompute的一个临时表，创建一个SQL脚本节点。单击运行 > 保存 > 提交。 SQL命令如下：

```
drop table if exists ods_user_t;
CREATE TABLE ods_user_t (
        dt STRING,
        uid STRING,
        gender STRING,
        age_range STRING,
        zodiac STRING);
insert overwrite table ods_user_t select dt,uid,gender,age_range,
zodiac from ods_user_info_d; --将ODPS表中的数据存入临时表。
```

2. 创建同步任务的节点mysql_to_odps，即简单的同步任务。将RDS数据全量同步到MaxCompute，无需设置分区。



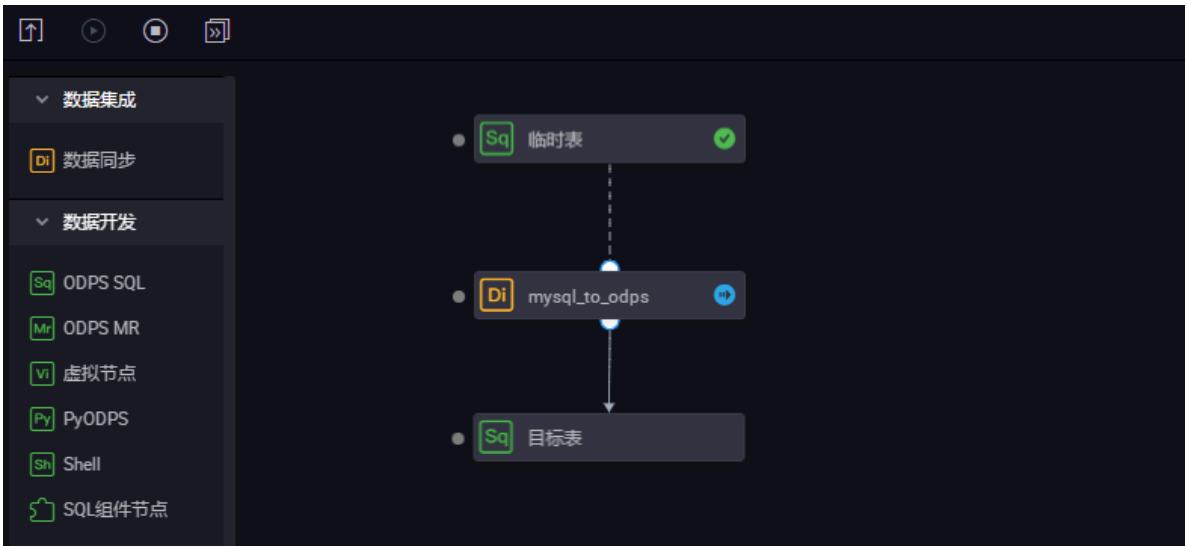
3. 使用SQL语句进行动态分区到目标表，命令如下。

```
drop table if exists ods_user_d;
//创建一个ODPS分区表（最终目的表）。
    CREATE TABLE ods_user_d (
    uid STRING,
        gender STRING,
        age_range STRING,
        zodiac STRING
)
PARTITIONED BY (
    dt STRING
);
//执行动态分区SQL，按照临时表的字段dt自动分区，dt字段中相同的数据值，会按照这个数据值自动创建一个分区值。
//例如dt中有些数据是20180913，会自动在ODPS分区表中创建一个分区，dt=20181025。
//动态分区SQL如下。
//可以注意到SQL中select的字段多写了一个dt，就是指定按照这个字段自动创建分区。
insert overwrite table ods_user_d partition(dt)select dt,uid,gender,
age_range,zodiac from ods_user_t;
```

```
//导入完成后，可以把临时表删除，节约存储成本。
drop table if exists ods_user_t;
```

在MaxCompute中您可以通过SQL语句完成数据同步。详细的SQL语句介绍请参见阿里云大数据利器MaxCompute学习之--分区表的使用。

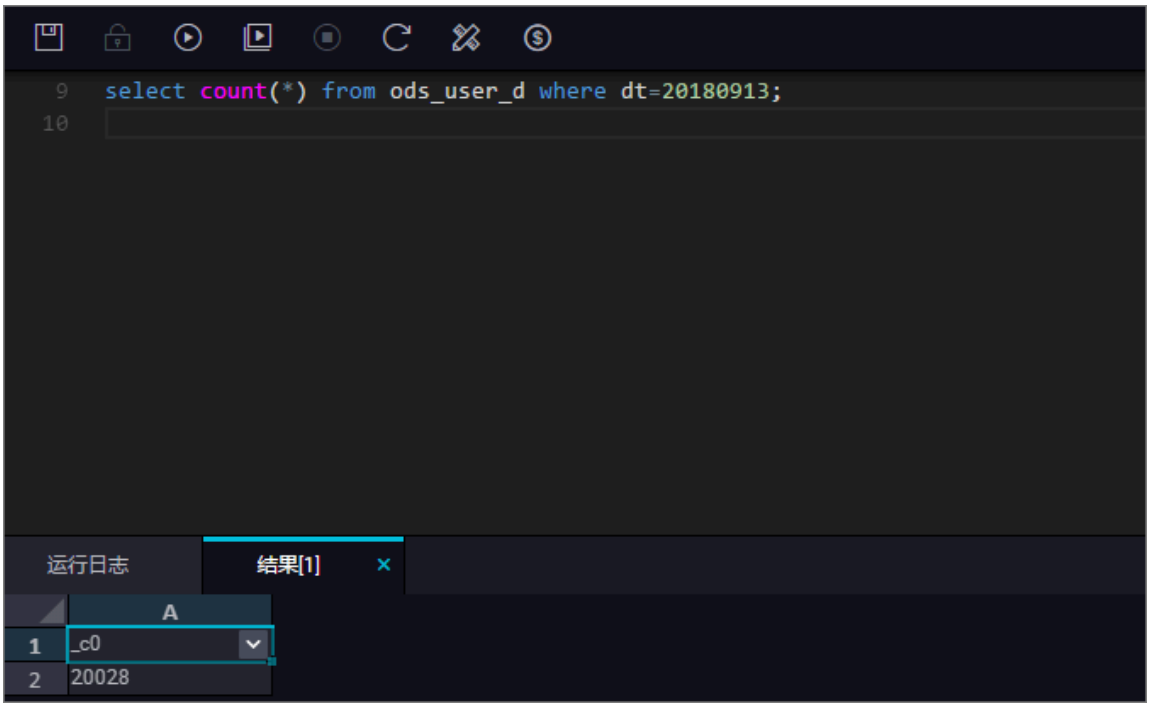
4. 将三个节点配置成一个工作流，按顺序执行。



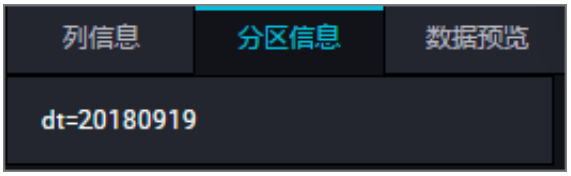
5. 查看执行过程。您可以重点观察最后一个节点的动态分区过程。

6. 查看运行结果。

a. 查看数据写入的情况，是否相同的日期数据已被迁移至同一个分区中。



b. 查询对应分区信息。



DataWorks数据同步功能可以完成大部分自动化作业，尤其是数据的同步迁移，调度等，了解更多的调度配置请参见调度配置时间属性。

2.7 JSON数据从MongoDB迁移至MaxCompute

本文将为您介绍如何通过DataWorks的数据集成功能，将从MongoDB提取的JSON字段迁移至MaxCompute。

准备工作

1. 账号准备

在数据库内新建用户，用于DataWorks添加数据源。本示例执行如下命令。

```
db.createUser({user:"bookuser",pwd:"123456",roles:["root"]})
```

新建用户名为bookuser，密码为123456，权限为root。

2. 数据准备

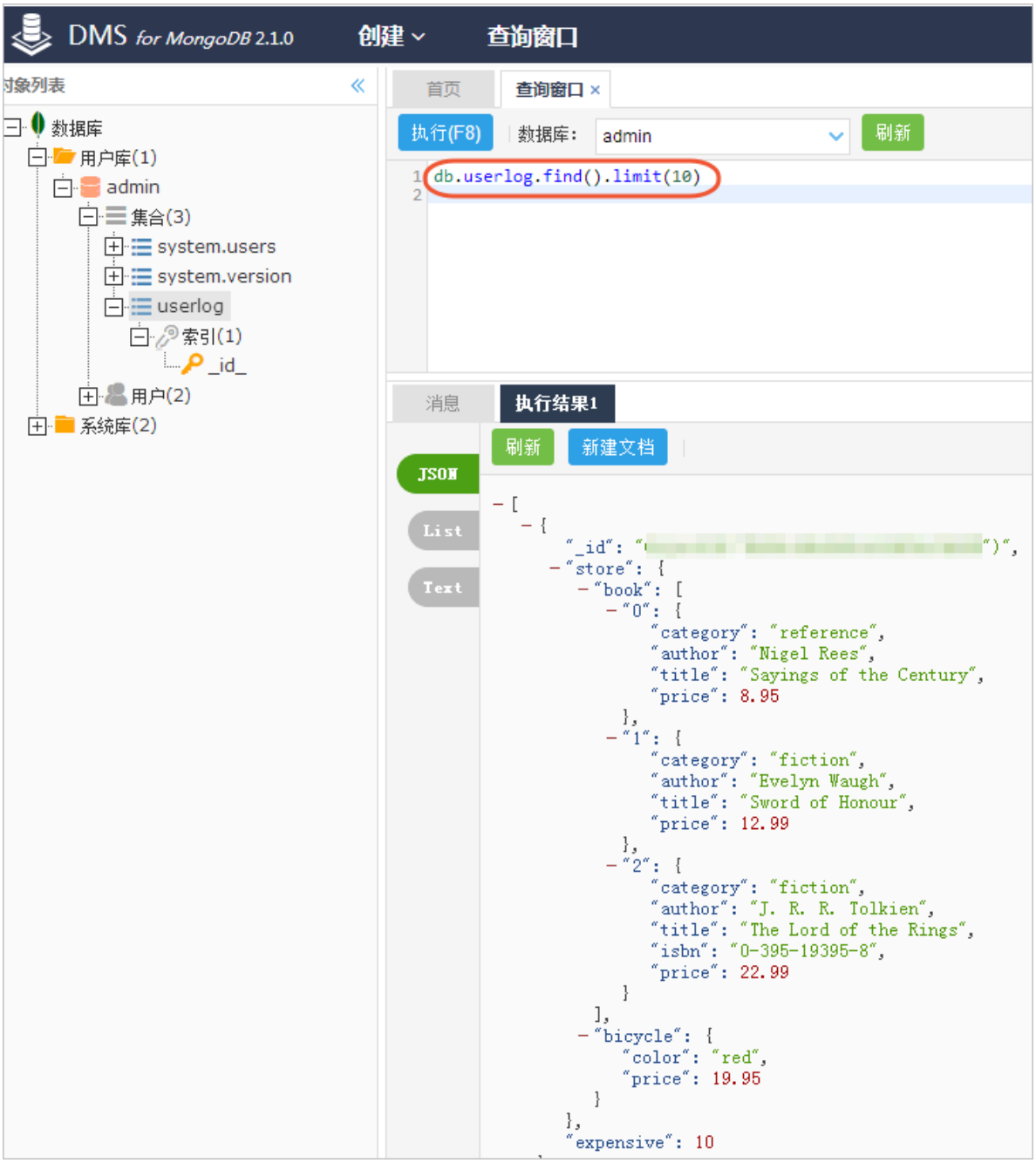
首先您需要将数据上传至您的MongoDB数据库。本示例使用阿里云的云数据库MongoDB版，网络类型为VPC（需申请公网地址，否则无法与DataWorks默认资源组互通），测试数据如下。

```
{
                        "store": {
                        "book": [
                        {
                        "category": "reference",
                        "author": "Nigel Rees",
                        "title": "Sayings of the Century",
                        "price": 8.95
                        },
                        {
                        "category": "fiction",
                        "author": "Evelyn Waugh",
                        "title": "Sword of Honour",
                        "price": 12.99
                        },
                        {
                        "category": "fiction",
                        "author": "J. R. R. Tolkien",
                        "title": "The Lord of the Rings",
                        "isbn": "0-395-19395-8",
                        "price": 22.99
                        }
                        ],
                        "bicycle": {
                        "color": "red",
                        "price": 19.95
                        }
                        },
                        "expensive": 10
```

```
                                }
```

登录MongoDB的DMS控制台，本示例使用的数据库为admin，集合为userlog。您可以在查询窗口执行如下命令，查看已上传的数据。

```
db.userlog.find().limit(10)
```

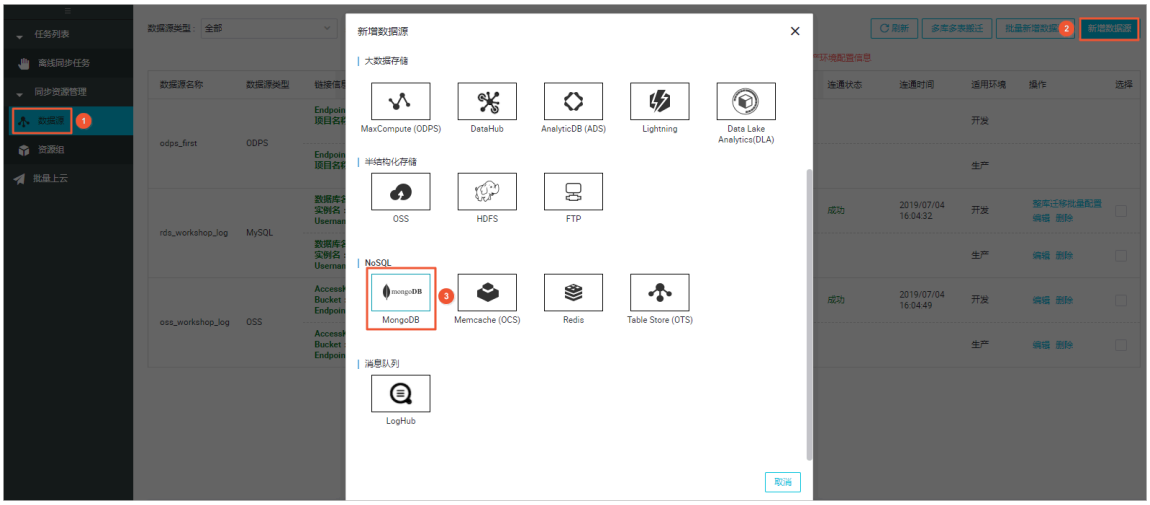


通过DataWorks将JSON数据从MongoDB迁移至MaxCompute

1. 新增MongoDB数据源

   a. 以项目管理员身份进入DataWorks控制台，单击对应工作空间操作栏中的进入数据集成。

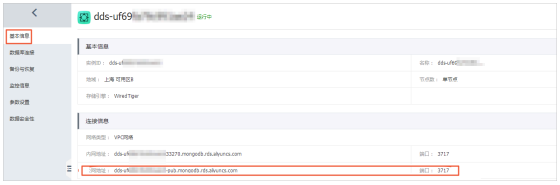
   b. 选择同步资源管理 > 数据源，单击新增数据源。



   c. 在新增数据源弹出框中，选择数据源类型为MongoDB。

   d. 填写MongoDB数据源的各配置项。

配置	说明
数据源类型	由于本文中MongoDB处于VPC环境下，因此数据源类型需选择连接串模式（数据集成网络可直接连通）。
数据源名称	数据源名称必须以字母、数字、下划线组合，且不能以数字和下划线开头。
数据源描述	对数据源进行简单描述，不得超过80个字符。
适用环境	可以选择开发或生产环境。 **说明:** 仅标准模式工作空间会显示此配置。

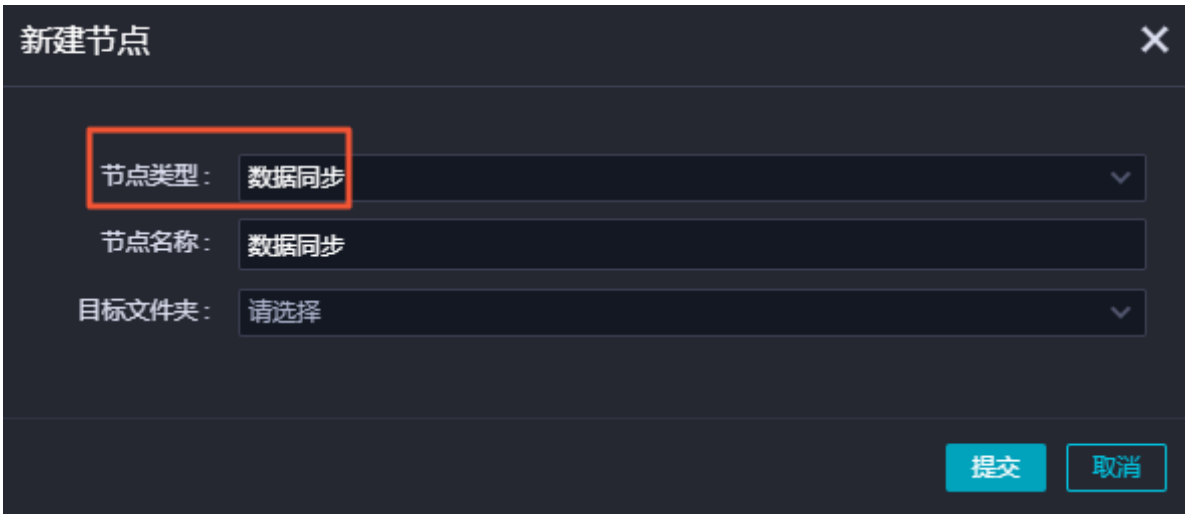
配置	说明
访问地址	访问地址及端口号可以通过单击MongoDB控制台中的实例名称获取。 
数据库名	该数据源对应的数据库名称。
用户名/密码	数据库对应的用户名和密码。

e. 单击测试连通性。

f. 测试连通性通过后，单击完成。

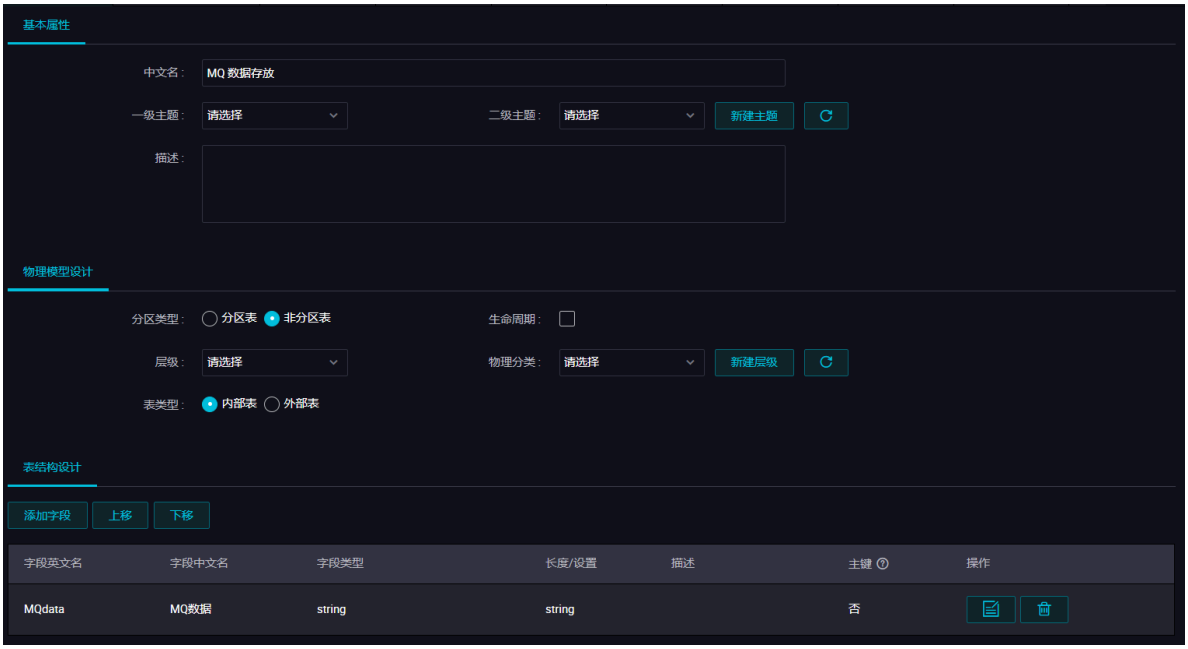
2. 新建数据同步任务

在DataWorks上新建数据同步节点，详情请参见#unique_59。



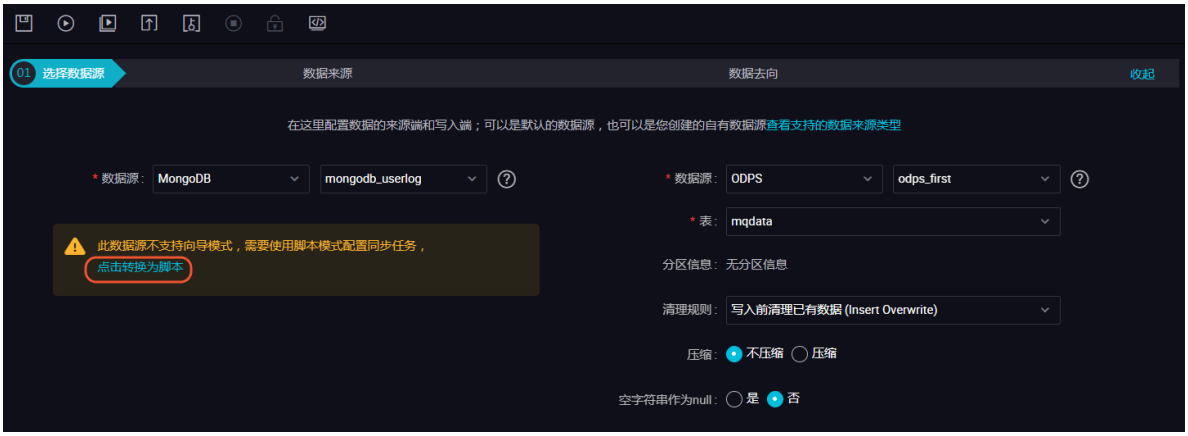
新建的同时，在DataWorks新建一个建表任务，用于存放JSON数据，本示例新建表名为mqdata。

表参数可以通过图形化界面完成。本例中mqdata表仅有一列，类型为string，列名为MQ data。



3. 配置同步任务参数

完成上述新建后，您可以在图形化界面配置数据同步任务参数，如下图所示。选择数据来源类型为MongoDB，来源表为mongodb_userlog。目标数据源名称为odps_first，目标表为刚新建的mqdata。

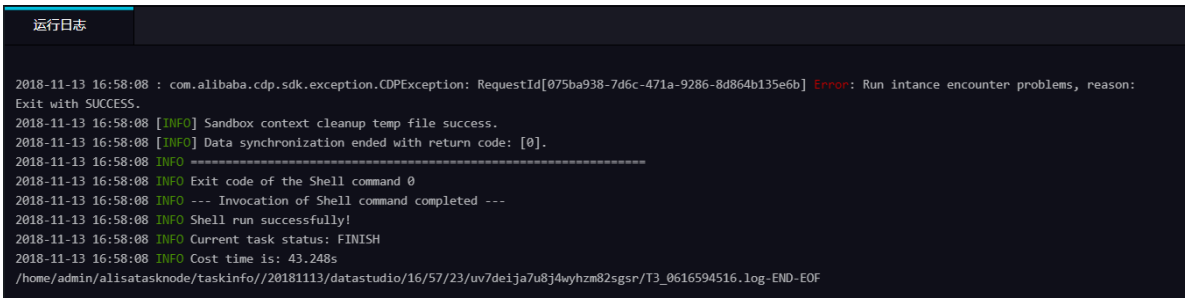


由于MongoDB数据源不支持向导模式开发，您直接点击转换为脚本，即可跳转至脚本模式进行配置。

```
{
    "type": "job",
    "steps": [
    {
        "stepType": "mongodb",
        "parameter": {
            "datasource": "mongodb_userlog",//数据源名称。
            "column": [
                {
                "name": "store.bicycle.color", //JSON字段路径，本例中提取color值。
```

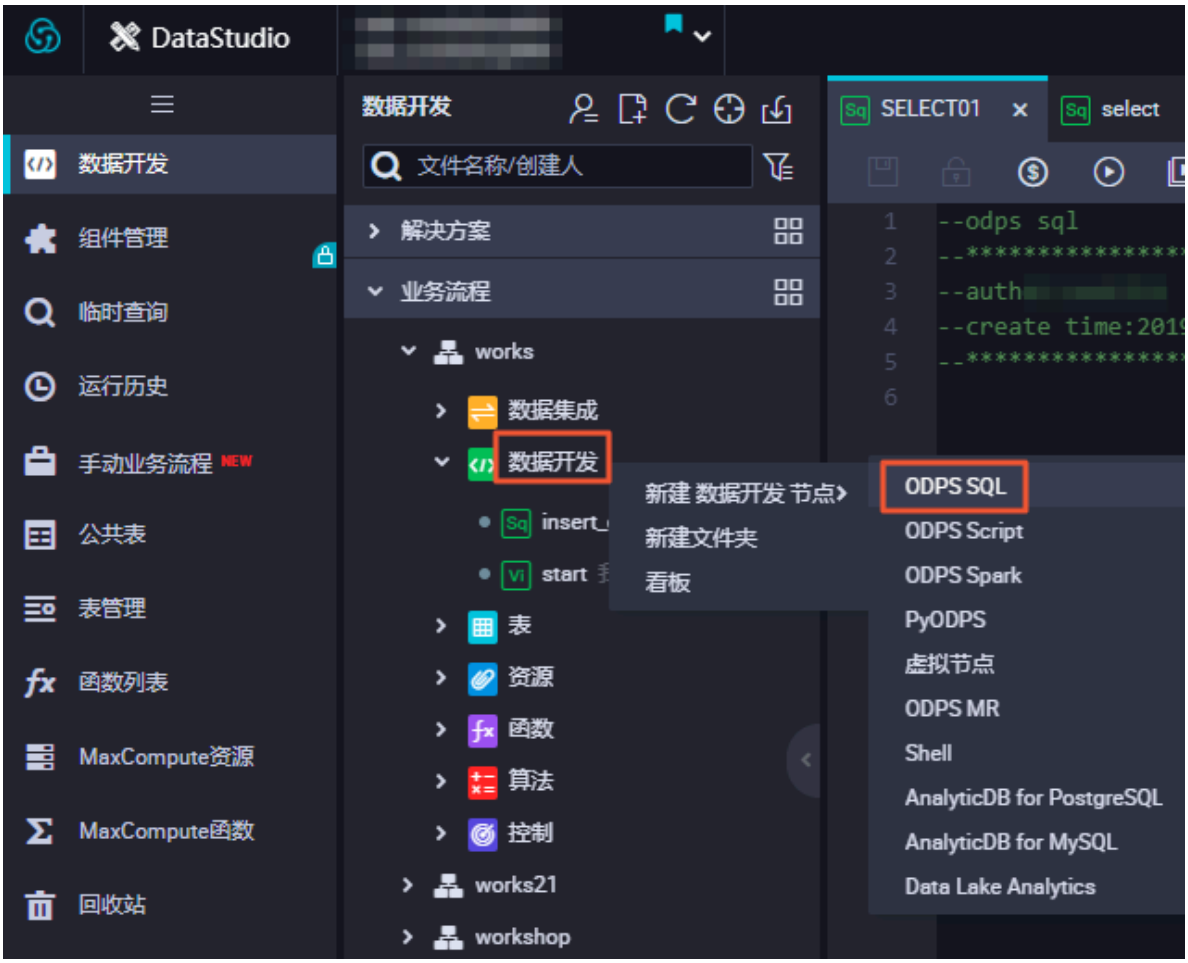
```
                    "type": "document.String" //非一层子属性以最终获取的类型为准。假如您选取的JSON字段为一级字段，如本例中的expensive，则直接填写string即可。
                    }
                ],
            "collectionName //集合名称": "userlog"
            },
        "name": "Reader",
        "category": "reader"
        },
        {
            "stepType": "odps",
            "parameter": {
            "partition": "",
            "isCompress": false,
            "truncate": true,
            "datasource": "odps_first",
            "column": [
            "mqdata"  //MaxCompute表列名。
            ],
            "emptyAsNull": false,
            "table": "mqdata"
            },
            "name": "Writer",
            "category": "writer"
            }
            ],
            "version": "2.0",
            "order": {
            "hops": [
            {
            "from": "Reader",
            "to": "Writer"
            }
            ]
            },
            "setting": {
            "errorLimit": {
            "record": ""
            },
            "speed": {
            "concurrent": 2,
            "throttle": false,
            }
            }
        }
```

完成上述配置后，单击运行即可。运行成功日志示例如下所示。

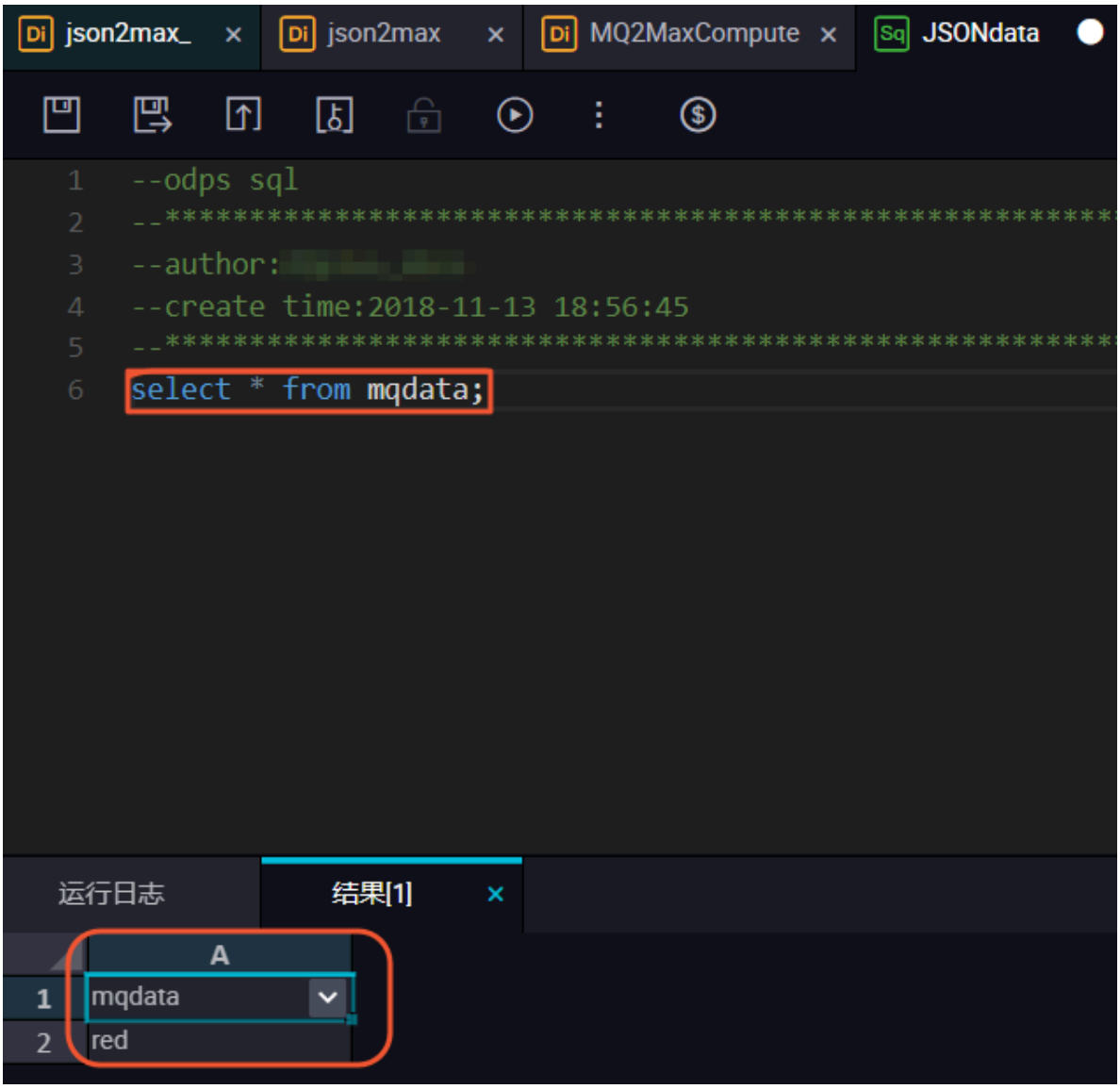


JSON数据从MongoDB迁移至MaxCompute结果验证

1. 在您的业务流程中新建一个ODPS SQL节点。



2. 输入`SELECT * from mqdata;`语句，查看当前mqdata表中数据。



说明:

您也可以直接在MaxCompute客户端中输入命令运行。

2.8 JSON数据从OSS迁移至MaxCompute

本文将为您介绍如何通过DataWorks的数据集成功能，将JSON数据从OSS迁移至MaxCompute，并使用MaxCompute内置字符串函数GET_JSON_OBJECT提取JSON信息的最佳实践。

准备工作

开始将JSON数据从OSS迁移至MaxCompute的操作前，您需要首先将JSON文件重命名为后缀是txt的文件，并上传至OSS。

本文中使用的JSON文件为*applog.txt*，将其上传至OSS，本文中OSS Bucket位于华东2区。


```
{
    "store": {
        "book": [
            {
               "category": "reference",
               "author": "Nigel Rees",
               "title": "Sayings of the Century",
               "price": 8.95
            },
            {
               "category": "fiction",
               "author": "Evelyn Waugh",
               "title": "Sword of Honour",
               "price": 12.99
            },
            {
                "category": "fiction",
                "author": "J. R. R. Tolkien",
                "title": "The Lord of the Rings",
                "isbn": "0-395-19395-8",
                "price": 22.99
            }
        ],
        "bicycle": {
            "color": "red",
            "price": 19.95
        }
    },
    "expensive": 10
```

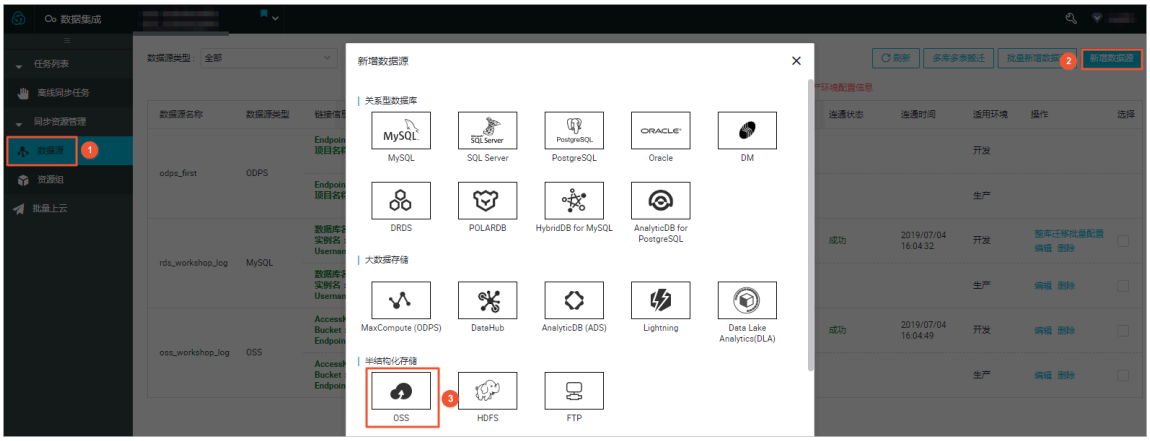
```
}
```

通过DataWorks将JSON数据从OSS迁移至MaxCompute

1. 新增OSS数据源

   a. 以项目管理员身份进入DataWorks控制台，单击对应工作空间操作栏中的进入数据集成。

   b. 选择同步资源管理 > 数据源，单击新增数据源。



   c. 在新增数据源弹出框中，选择数据源类型为OSS。

   d. 填写OSS数据源的各配置项。

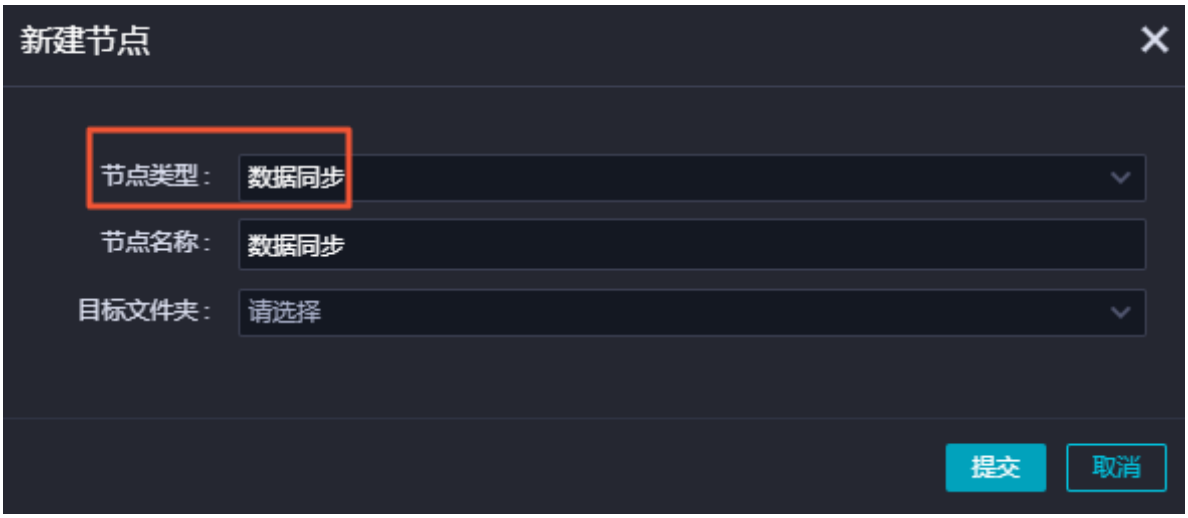
配置	说明
数据源名称	数据源名称必须以字母、数字、下划线组合，且不能以数字和下划线开头。
数据源描述	对数据源进行简单描述，不得超过80个字符。
适用环境	可以选择开发或生产环境。  说明: 仅标准模式工作空间会显示此配置。
Endpoint	OSS Endpoint信息，本示例为*http://oss-cn-shanghai.aliyuncs.com*或*http://oss-cn-shanghai-internal.aliyuncs.com*。OSS各地域的外网、内网地址请参见#unique_64/unique_64_Connect_42_section_plb_2vy_5db。  说明: 由于本文中OSS和DataWorks项目处于同一个区域中，所以本文选用后者，通过内网连接。
Bucket	相应的OSS Bucket信息，指存储空间，是用于存储对象的容器。 您可以创建一个或多个存储空间，每个存储空间可添加一个或多个文件。 您可以在数据同步任务中查找此处填写的存储空间中相应的文件，没有添加的存储空间，则不能查找其中的文件。
AccessKey ID/ AceessKey Secret	访问秘匙（AccessKeyID和AccessKeySecret），相当于登录密码。

e. 单击测试连通性。

f. 测试连通性通过后，单击完成。

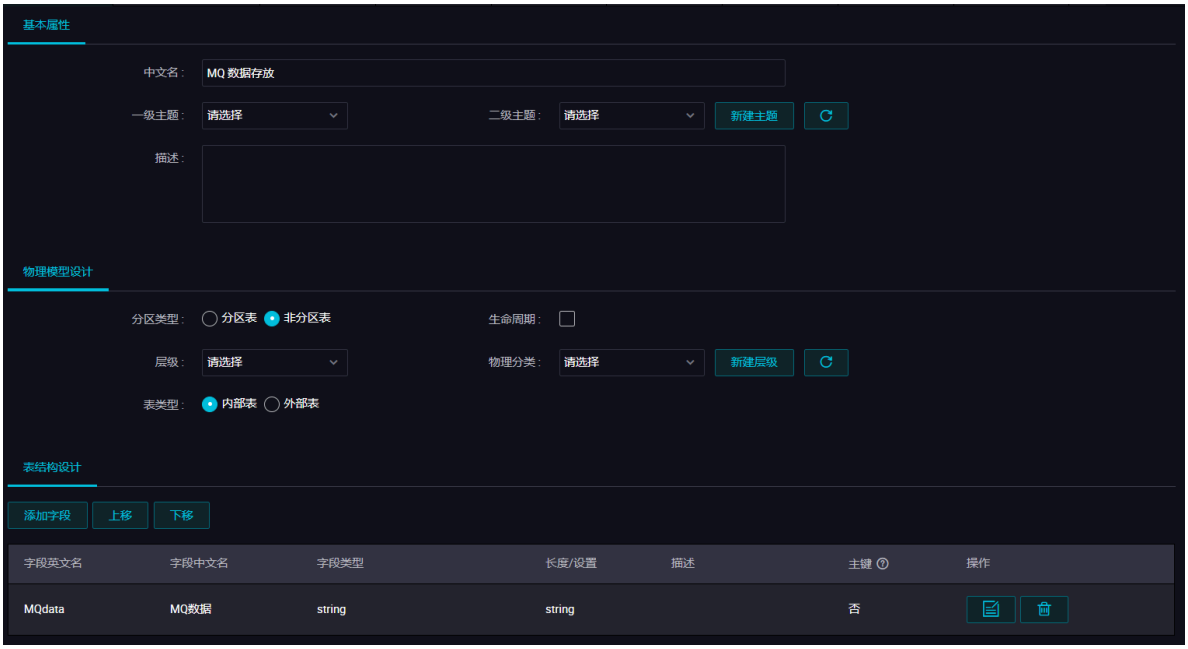
2. 新建数据同步任务

在DataWorks上新建数据同步节点，详情请参见#unique_59。



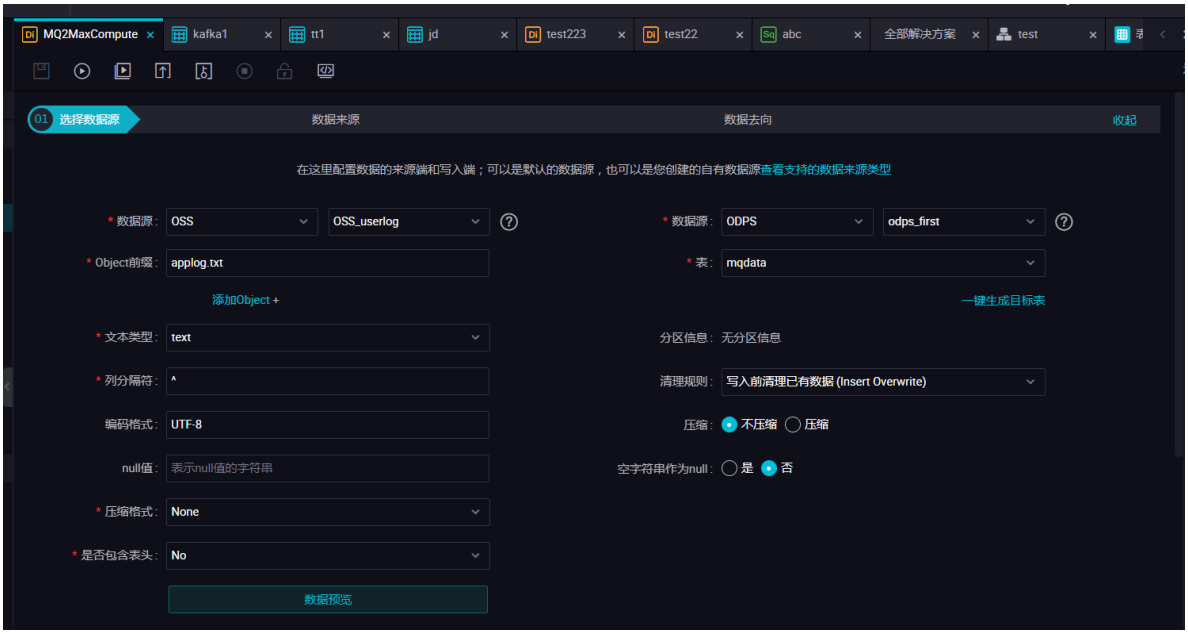
新建的同时，在DataWorks新建一个建表任务，用于存放JSON数据，本示例新建表名为mqdata。

表参数可以通过图形化界面完成。本例中mqdata表仅有一列，类型为string，列名为MQ data。



3. 配置同步任务参数

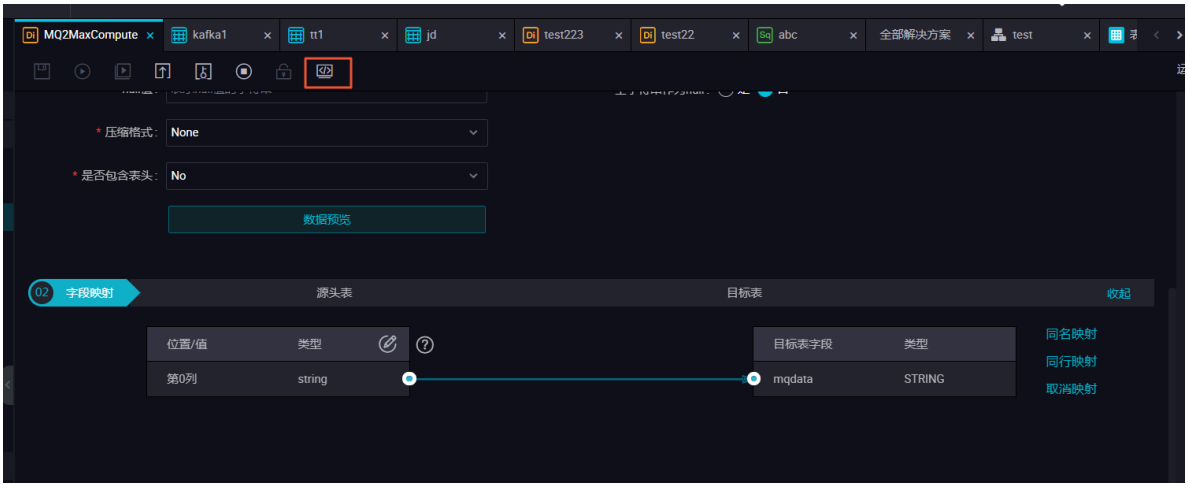
完成上述新建后，您可以在图形化界面配置数据同步任务参数，如下图所示。选择目标数据源名称为odps_first，选择目标表为刚建立的mqdata。数据来源类型为OSS，Object前缀可填写文件路径及名称。



说明：

列分隔符使用TXT文件中不存在的字符即可，本文使用（^）。对于OSS中的TXT格式数据源，Dataworks支持多字符分隔符，您可以使用（%&%#^$$^%）这种很难出现的字符串作为列分隔符。

映射方式选择默认的同行映射即可。



单击左上方的切换脚本按钮，切换为脚本模式。修改fileFormat参数为`"fileFormat":"binary"`。脚本模式代码示例如下。

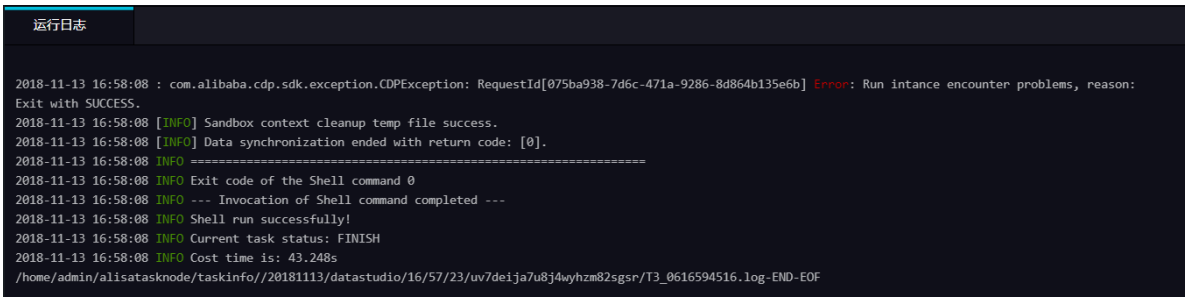
```
{
    "type": "job",
    "steps": [
        {
            "stepType": "oss",
            "parameter": {
                "fieldDelimiterOrigin": "^",
                "nullFormat": "",
                "compress": "",
                "datasource": "OSS_userlog",
                "column": [
                    {
                        "name": 0,
                        "type": "string",
                        "index": 0
                    }
                ],
                "skipHeader": "false",
                "encoding": "UTF-8",
                "fieldDelimiter": "^",
                "fileFormat": "binary",
                "object": [
                    "applog.txt"
                ]
            },
            "name": "Reader",
            "category": "reader"
        },
        {
            "stepType": "odps",
            "parameter": {
                "partition": "",
                "isCompress": false,
```

```
                "truncate": true,
                "datasource": "odps_first",
                "column": [
                    "mqdata"
                ],
                "emptyAsNull": false,
                "table": "mqdata"
            },
            "name": "Writer",
            "category": "writer"
        }
    ],
    "version": "2.0",
    "order": {
        "hops": [
            {
                "from": "Reader",
                "to": "Writer"
            }
        ]
    },
    "setting": {
        "errorLimit": {
            "record": ""
        },
        "speed": {
            "concurrent": 2,
            "throttle": false,
        }
    }
}
```

说明:

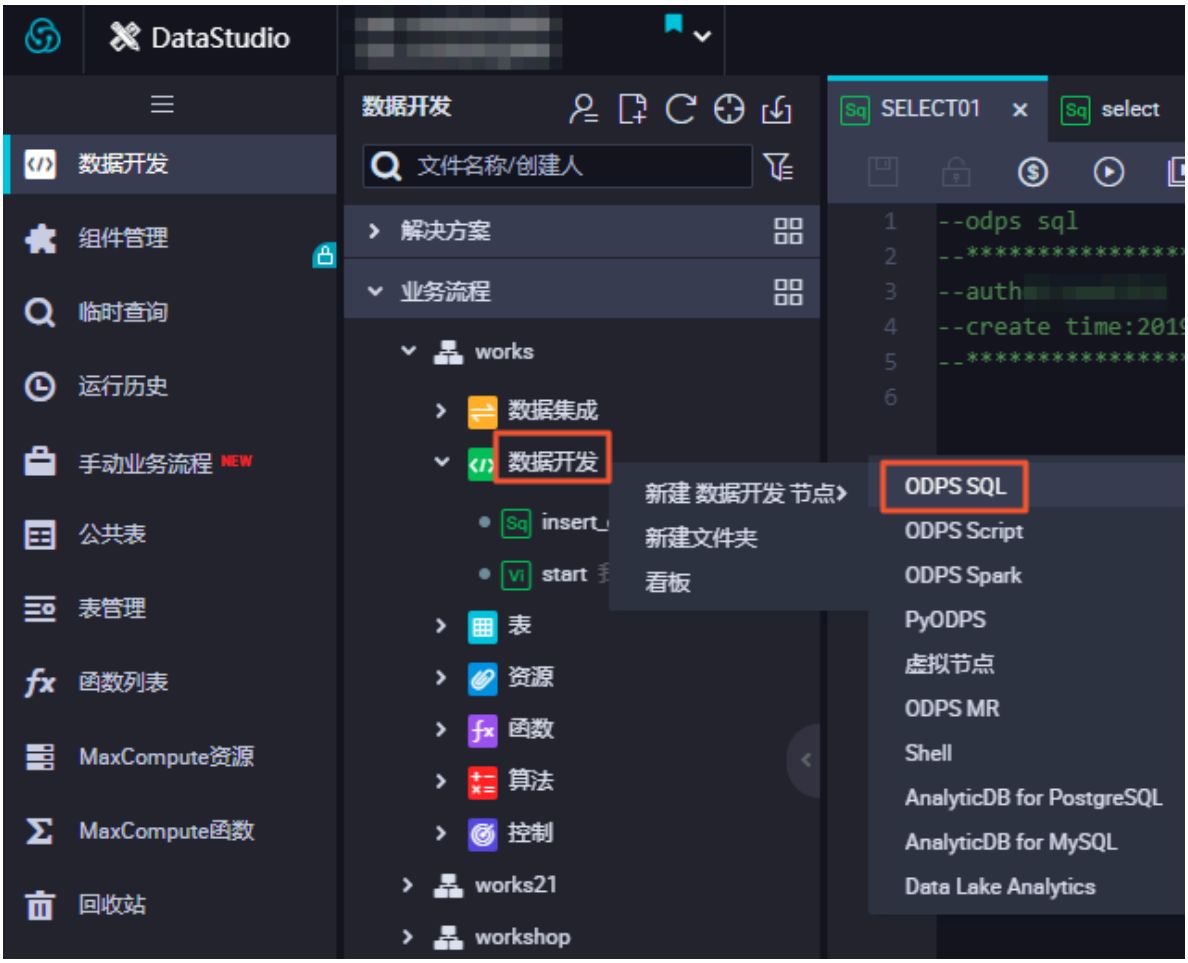
该步骤可以保证OSS中的JSON文件同步到MaxCompute之后存在同一行数据中，即为一个字段，其他参数保持不变。

完成上述配置后，单击运行即可。运行成功日志示例如下所示。

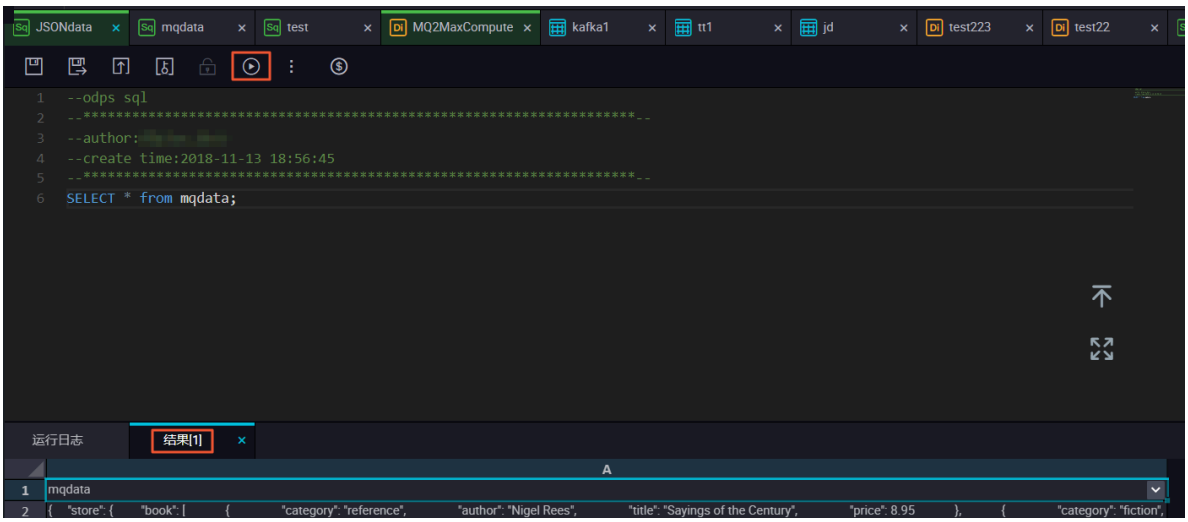


JSON数据从OSS迁移至MaxCompute结果验证

1. 在您的业务流程中新建一个ODPS SQL节点。

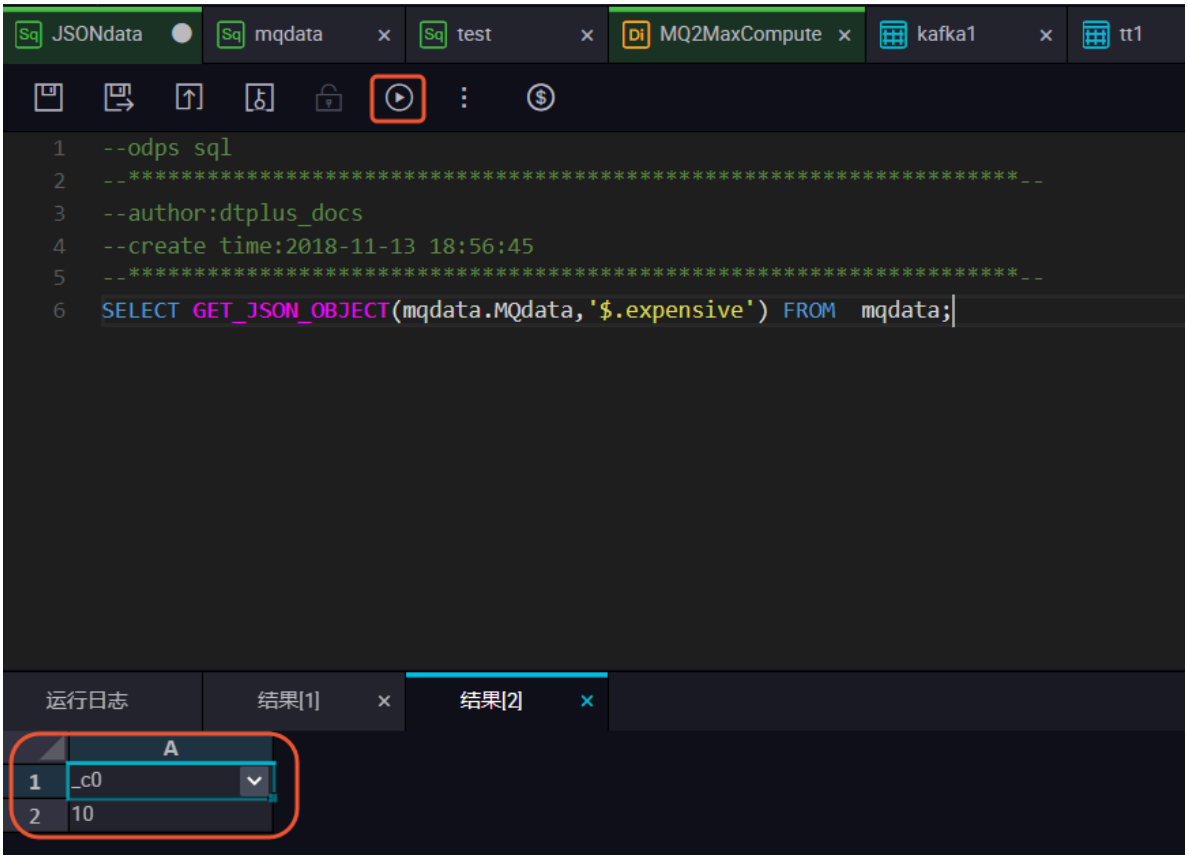


2. 查看当前mqdata表中数据，输入`SELECT * from mqdata;`语句。



> **说明:**
> 这一步及后续步骤，也可以直接在MaxCompute客户端中输入命令运行。

3. 确认导入表中的数据结果无误后，使用`SELECT GET_JSON_OBJECT(mqdata.MQdata,'$.expensive') FROM mqdata;`获取JSON文件中的`expensive`值。



更多信息

在进行迁移后结果验证时，您可以使用MaxCompute内建字符串函数GET_JSON_OBJECT获取您想要的JSON数据。

2.9 MaxCompute数据迁移至OTS

本文为您介绍如何将MaxCompute数据迁移至Table Store（OTS）。

前提条件

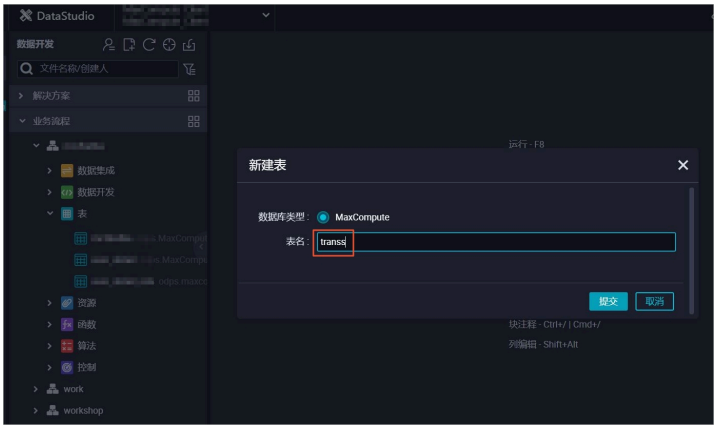
请提前开通MaxCompute和DataWorks服务。本例使用DataWorks简单模式。

操作步骤

1. 在DataWorks控制台中创建表。

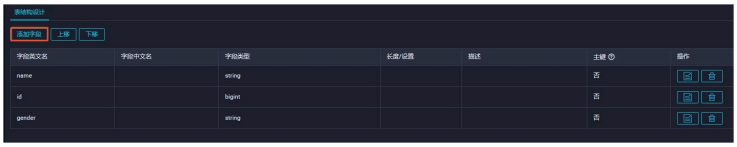
   a) 新建业务流程，并将业务流程命名为mc2ots。

   b) 新建表transs。



   c) 输入表的基本属性。

   d) 单击添加字段，创建表字段name、id、gender。



   e) 单击提交到生产环境。

f) 导入表数据。

表数据如下。

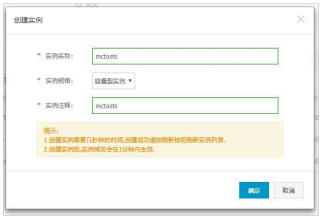
```
qwe,145,F
asd,256,F
xzc,345,M
rgth,234,F
ert,456,F
dfg,12,M
tyj,4,M
bfg,245,M
nrtjeryj,15,F
rwh,2344,M
trh,387,F
srjeyj,67,M
```

```
saerh,567,M
```

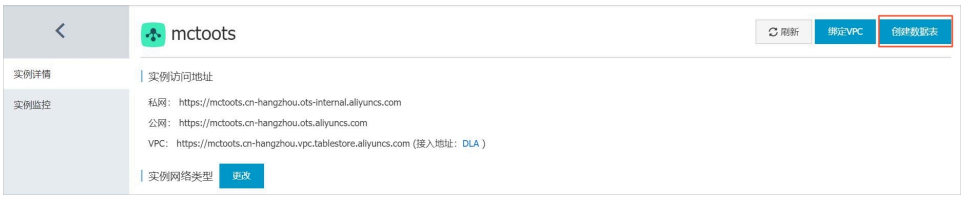
2. 在表格存储Table Store控制台中创建表。

   a) 登录表格存储Table Store控制台，单击创建实例。

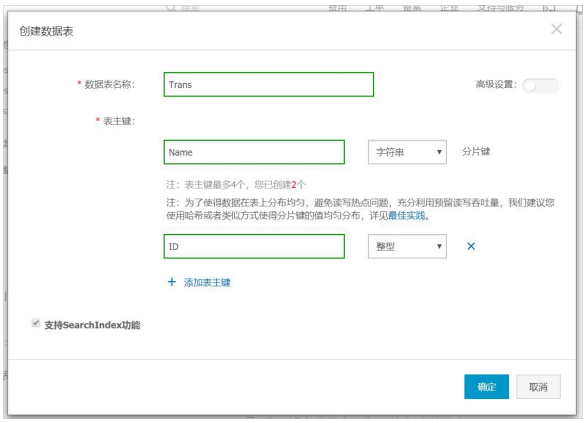
   b) 创建实例mctoots。



   c) 单击创建数据表。

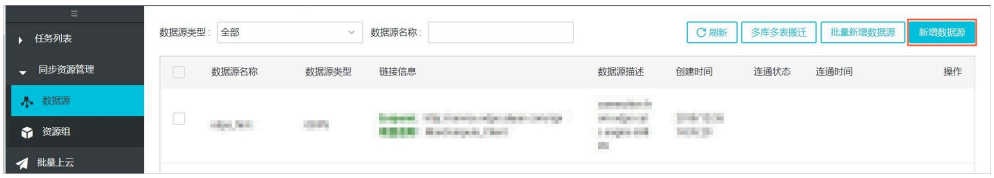


   d) 创建数据表Trans。

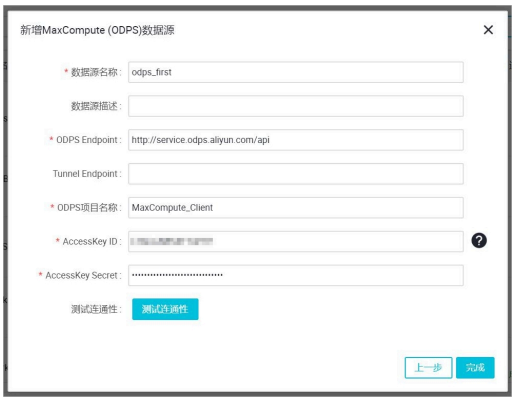


3. 新增数据源。

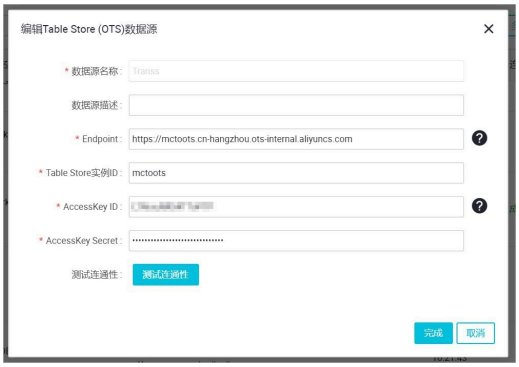
a) 单击新增数据源。



b) 新增MaxCompute（ODPS）数据源，并将数据源命名为odps_first。



c) 新增Table Store（OTS）数据源，并将数据源命名为Transs。

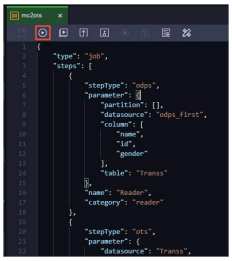


4. 配置MaxCompute（ODPS）Reader和Table Store（OTS）Writer。

a) 新建数据同步节点mc2ots。

b) 在脚本模式中，导入模板。

c) 修改JSON代码后，单击运行按钮。

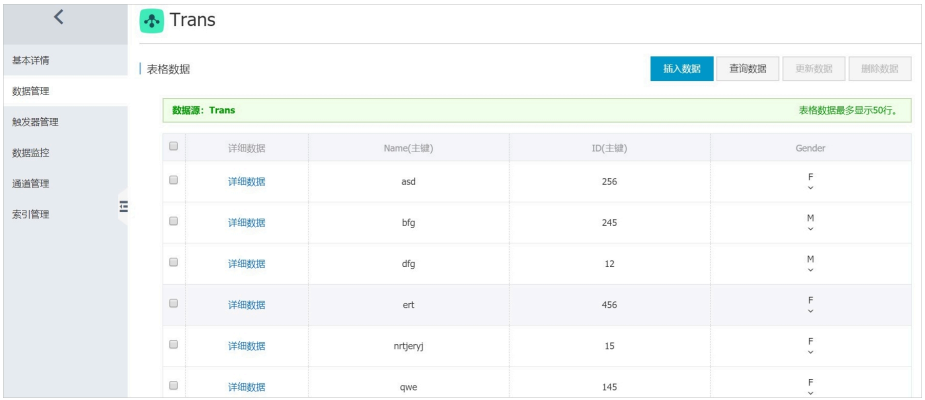


代码如下。

```
{
    "type": "job",
    "steps": [
        {
            "stepType": "odps",
            "parameter": {
                "partition": [],
                "datasource": "odps_first",
                "column": [
                    "name",
                    "id",
                    "gender"
                ],
                "table": "Transs"
            },
            "name": "Reader",
            "category": "reader"
        },
        {
            "stepType": "ots",
            "parameter": {
                "datasource": "Transs",
                "column": [
```

```
                    {
                        "name": "Gender",
                        "type": "STRING"
                    }
                ],
                "writeMode": "UpdateRow",
                "table": "Trans",
                "primaryKey": [
                    {
                        "name": "Name",
                        "type": "STRING"
                    },
                    {
                        "name": "ID",
                        "type": "INT"
                    }
                ]
            },
            "name": "Writer",
            "category": "writer"
        }
    ],
    "version": "2.0",
    "order": {
        "hops": [
            {
                "from": "Reader",
                "to": "Writer"
            }
        ]
    },
    "setting": {
        "errorLimit": {
            "record": "0"
        },
        "speed": {
            "throttle": false,
            "concurrent": 1,
            "dmu": 1
        }
    }
}
```

5. 在表格存储Table Store（OTS）控制台中查看新增的表数据。



2.10 MaxCompute数据迁移至OSS

本文为您介绍如何将MaxCompute数据迁移至Object Storage Service（OSS）。

前提条件

请提前开通MaxCompute和DataWorks服务。本例使用DataWorks简单模式。

操作步骤

1. 在DataWorks控制台中创建表。

   a) 新建业务流程，并将业务流程命名为mc2oss。

   b) 新建表transition。

   c) 单击添加字段，创建表字段name、id、gender。

   d) 单击提交到生产环境。

   e) 单击导入数据。

   f) 导入表数据。

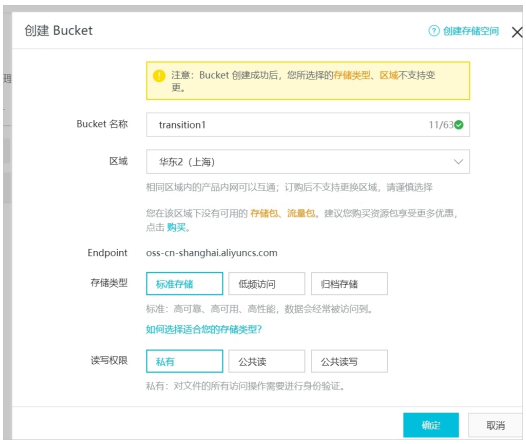
表数据如下。

```
qwe,145,F
asd,256,F
xzc,345,M
rgth,234,F
ert,456,F
dfg,12,M
tyj,4,M
bfg,245,M
nrtjeryj,15,F
rwh,2344,M
trh,387,F
srjeyj,67,M
```

```
saerh,567,M
```

2. 在对象存储（OSS）控制台中创建表。

   a) 登录对象存储（OSS）控制台，创建Bucket。



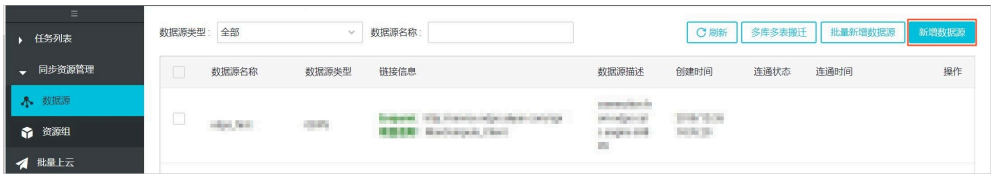
   b) 单击文件管理 > 上传文件。

   c) 在上传文件页签中，传入文件*qwee.csv*。

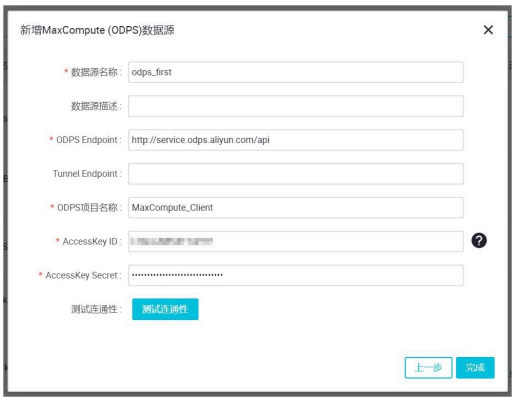
   请确保*qwee.csv*文件中的字段与表transition的字段完全一致。

3. 新增数据源。

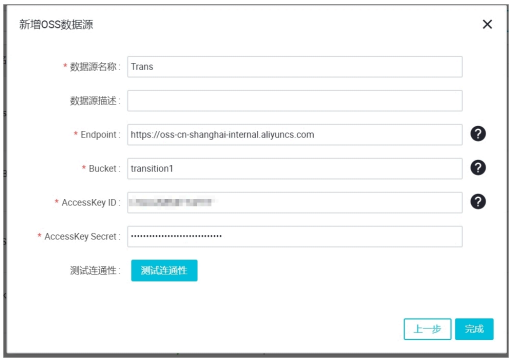
a) 单击新增数据源。



b) 新增MaxCompute（ODPS）数据源，并将数据源命名为odps_first。



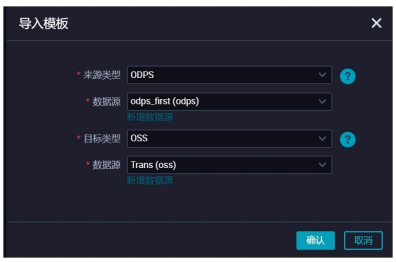
c) 新增表格存储（OSS）数据源，并将数据源命名为Trans。



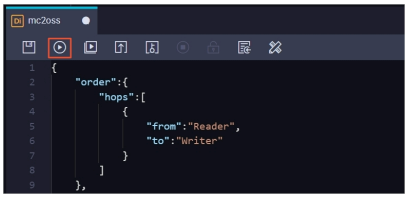
4. 配置MaxCompute（ODPS）Reader和对象存储（OSS）Writer。

a) 新建数据同步节点mc2oss。

b) 在脚本模式中，导入模板。



c) 修改JSON代码后，单击运行按钮。



代码如下。

```
{
    "order":{
        "hops":[
            {
                "from":"Reader",
                "to":"Writer"
            }
        ]
    },
    "setting":{
        "errorLimit":{
            "record":"0"
        },
        "speed":{
            "concurrent":1,
            "dmu":1,
            "throttle":false
        }
    },
    "steps":[
        {
            "category":"reader",
            "name":"Reader",
            "parameter":{
                "column":[
                    "name",
                    "id",
                    "gender"
```

```
                ],
                "datasource":"odps_first",
                "partition":[],
                "table":"Transs"
            },
            "stepType":"odps"
        },
        {
            "category":"writer",
            "name":"Writer",
            "parameter":{
                "datasource":"Trans",
                "dateFormat":"yyyy-MM-dd HH:mm:ss",
                "encoding":"UTF-8",
                "fieldDelimiter":",",
                "fileFormat":"csv",
                "nullFormat":"null",
                "object":"qweee.csv",
                "writeMode":"truncate"
            },
            "stepType":"oss"
        }
    ],
    "type":"job",
    "version":"2.0"
}
```

5. 在对象存储（OSS）控制台中查看新增的表数据。

3 数据开发

3.1 Eclipse Java UDF开发最佳实践

本文将为您介绍如何使用Eclipse开发工具，配合ODPS插件进行Java UDF开发的全流程操作。

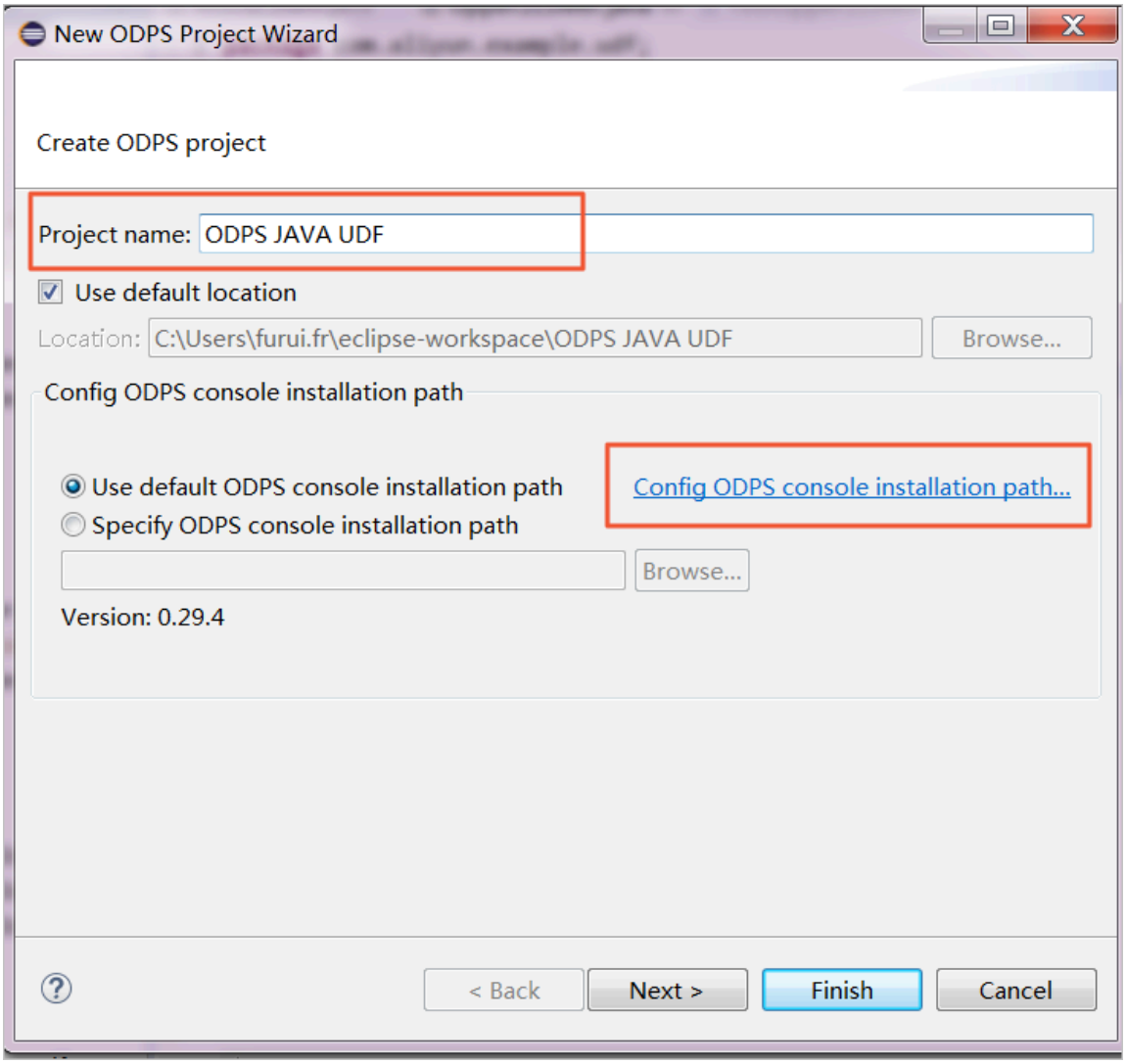
准备工作

开始使用Eclipse进行Java UDF开发前，您需要进行如下准备工作：

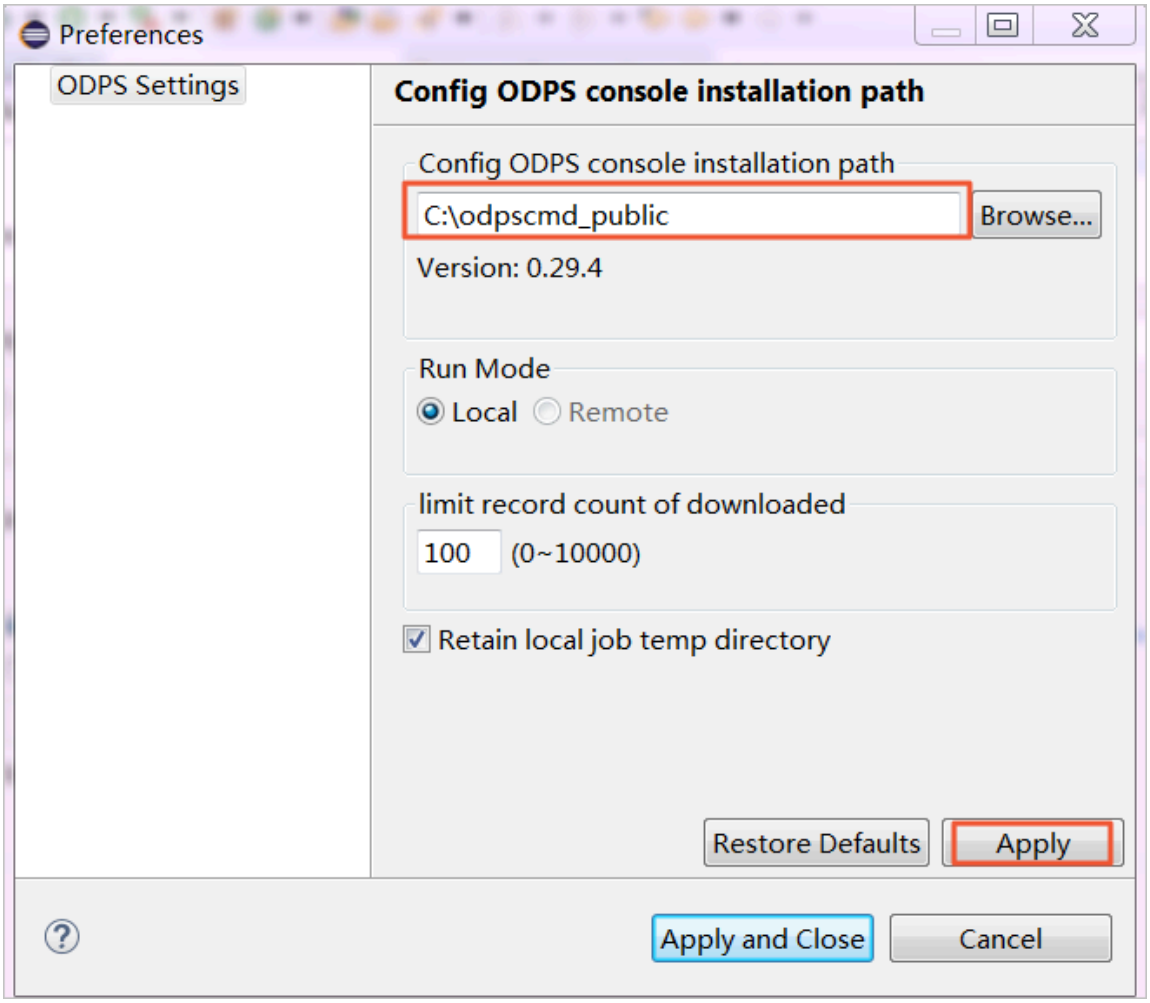
1. 使用Eclipse安装ODPS插件。

2. 创建ODPS Project。

   a. 在Eclipse中选择File > New > ODPS Project，输入项目名称，单击Config ODPS console installation path，配置odpscmd客户端安装路径。



   b. 输入客户端整体安装包的路径后，单击Apply，ODPS插件会为您自动解析出客户端版本。

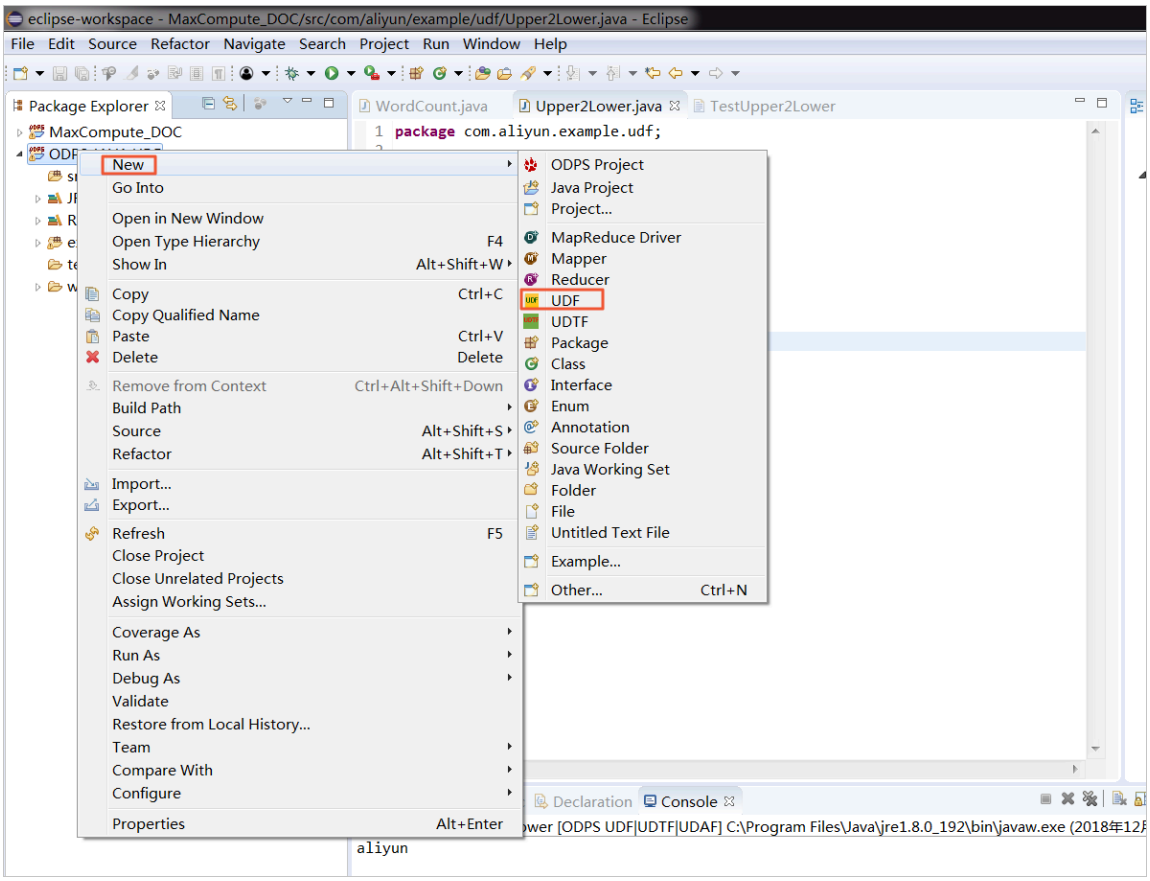


c. 单击Finish，即可完成项目的创建。

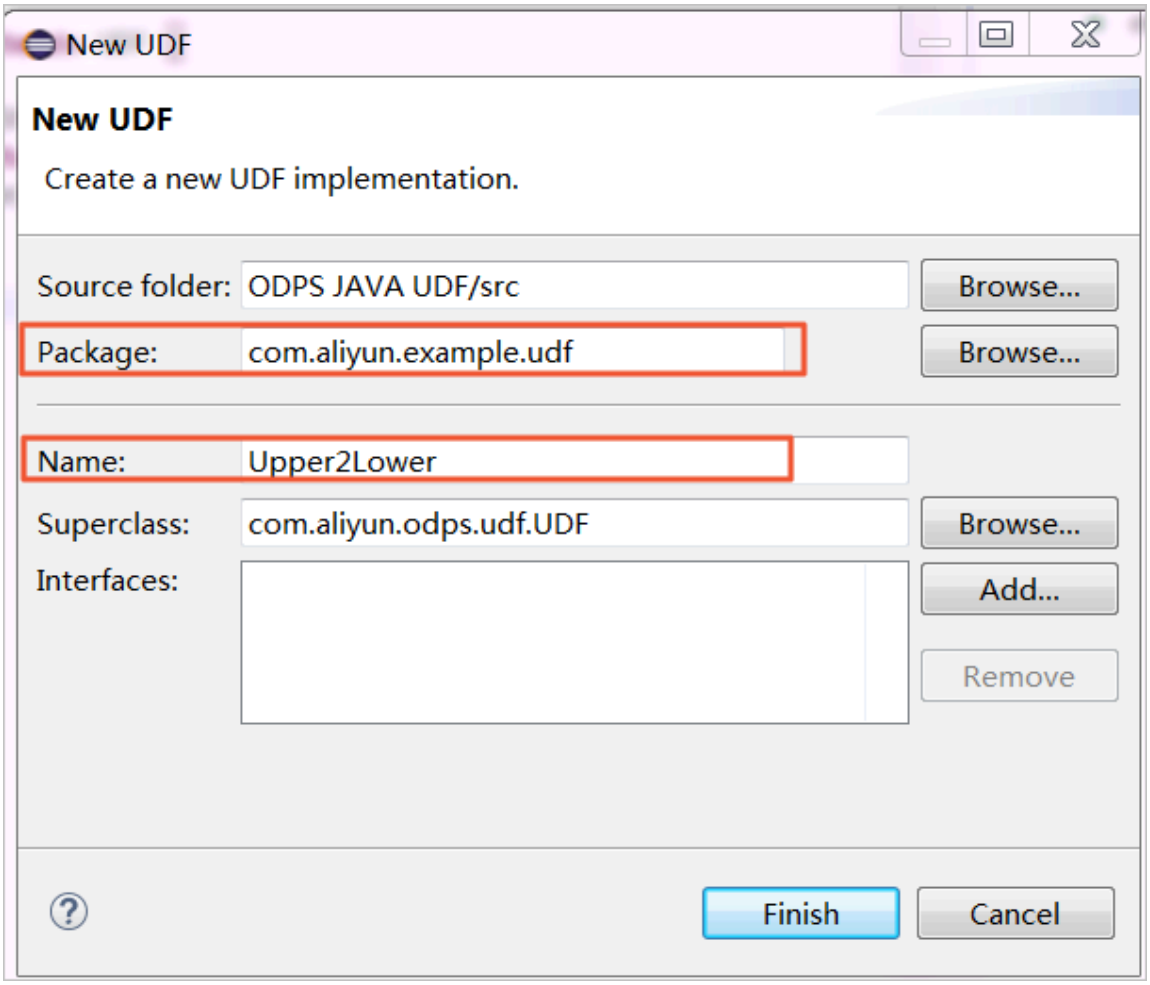
开发步骤

1. 在ODPS Project中创建Java UDF。

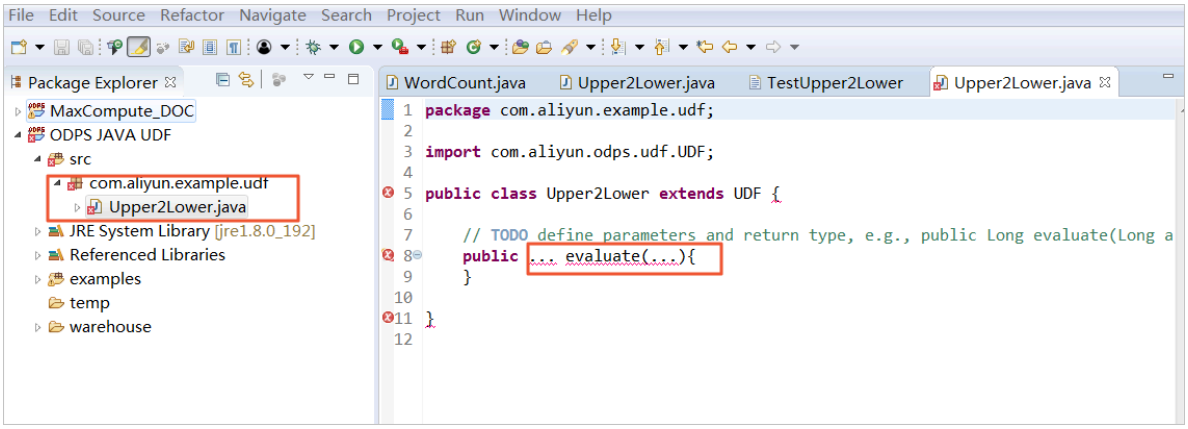
   a. 在左侧Package Exploer中右键单击新建的ODPS Java UDF项目，选择New > UDF。



   b. 输入UDF的Package名称（本例中为com.aliyun.example.udf）和Name（本例中为Upper2Lower），单击Finish，即可完成UDF的创建。

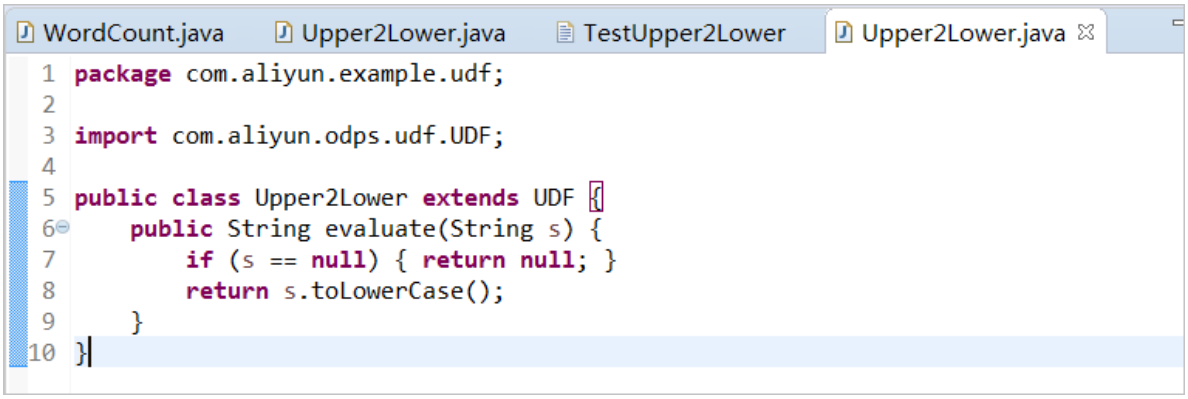


完成UDF创建后，您即可看到生成的默认Java代码，请注意不要改变evaluate()方法的名称。



2. 实现UDF类文件中的evaluate方法。

将您想要实现的功能代码写到evaluate方法中，且不要改变evaluate()方法的名称。此处以实现大写字母转化为小写字母为例。




```
package com.aliyun.example.udf;

import com.aliyun.odps.udf.UDF;

public class Upper2Lower extends UDF {
    public String evaluate(String s) {
        if (s == null) { return null; }
        return s.toLowerCase();
    }
}
```

代码编写完成后，请及时保存。

测试Java UDF

您可以先在MaxCompute上存放一些大写字母作为输入数据，以便测试Java UDF代码。

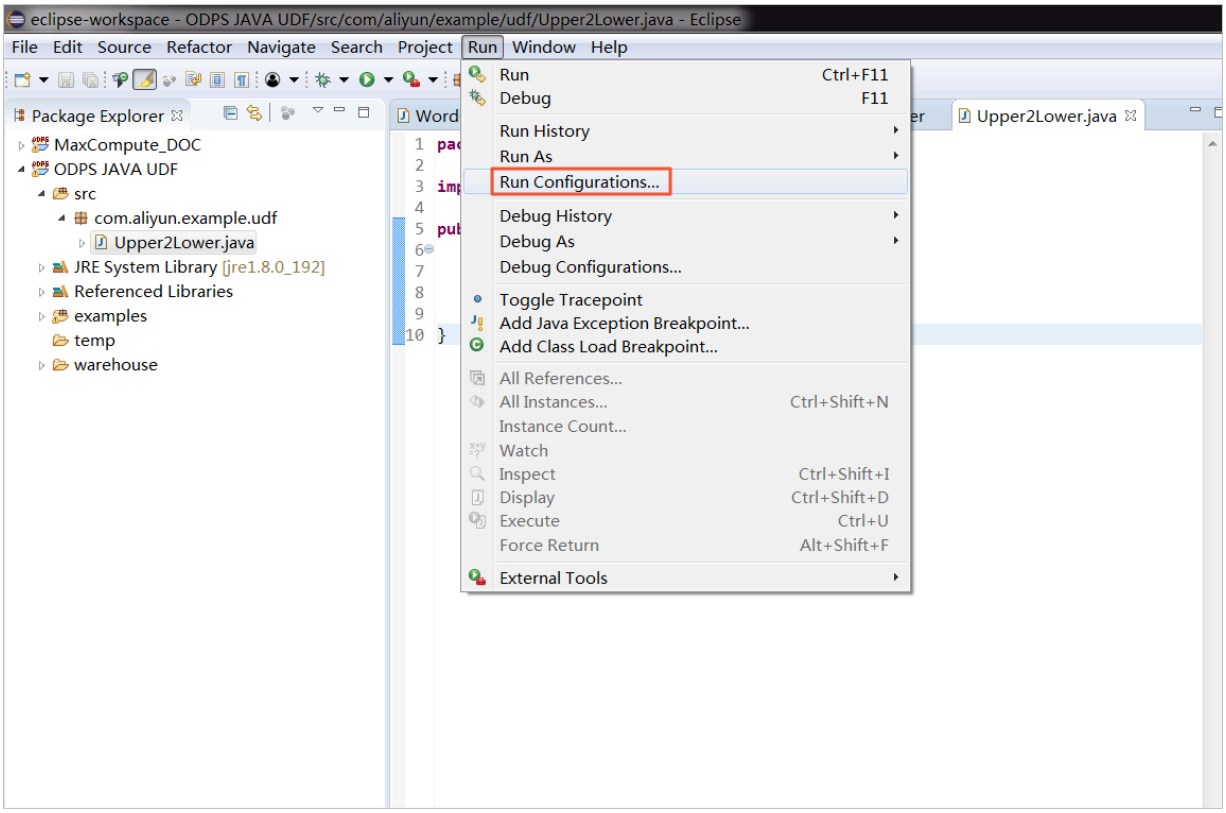
在odpscmd客户端使用SQL语句`create table upperABC(upper string);`，新建一个名为upperABC的测试表格。

```
odps@ MaxCompute_DOC>create table upperABC(upper string) ;

ID = 20181214094323883gb23j292
OK
```

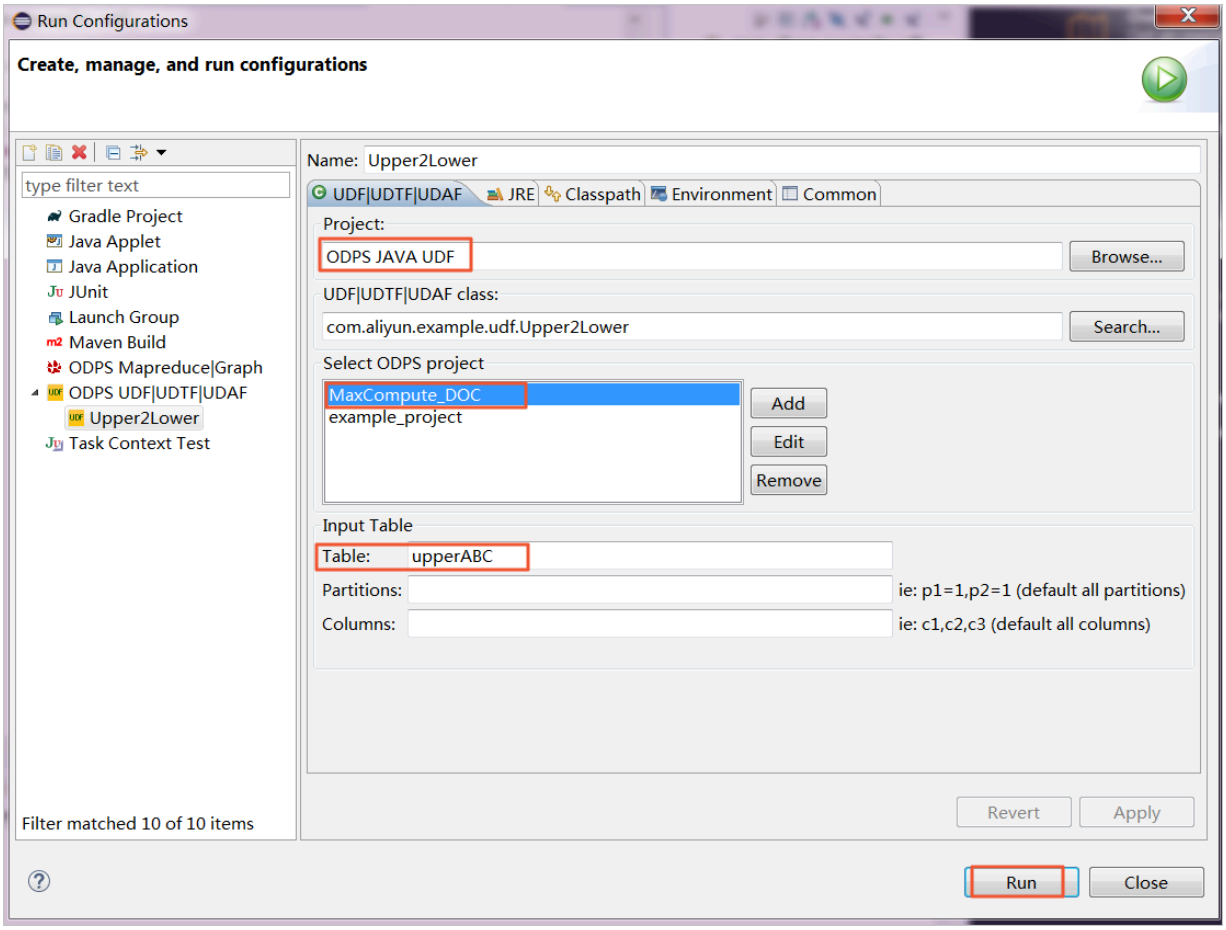
使用SQL语句`insert into upperABC values('ALIYUN');`，在表格中插入测试用的大写字母ALIYUN。

完成测试数据准备后，单击Run，选择Run Configurations，配置测试参数。

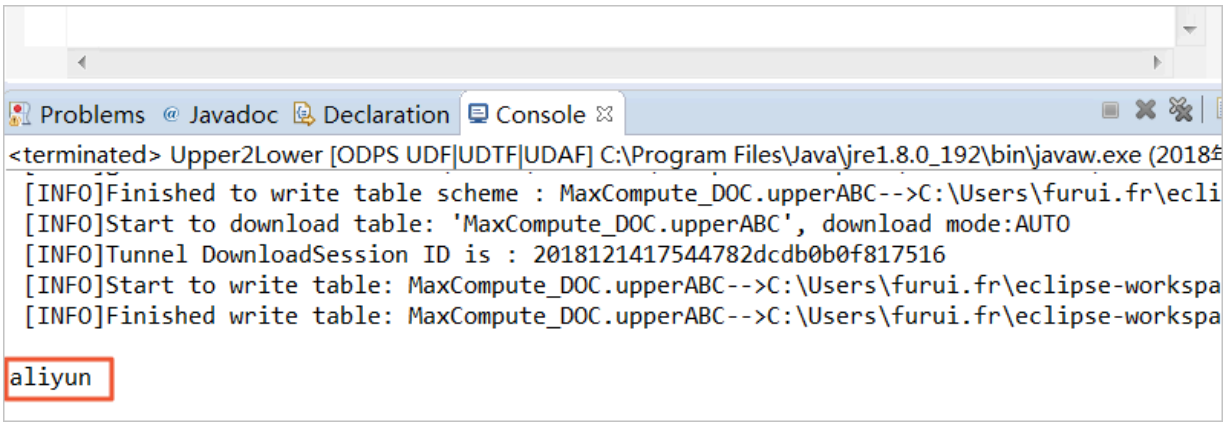


参数	说明
Project	填写创建的Java ODPS Project名称
Select ODPS project	填写您的MaxCompute项目名称（请注意与） **说明:** 此处填写的MaxCompute项目名称与odpscmd客户端当前连接的MaxCompute项目保持一致。
Table	填写刚才创建的测试表格名称

完成配置后，单击Run进行测试。



您可以在Console中查看测试结果。



说明:

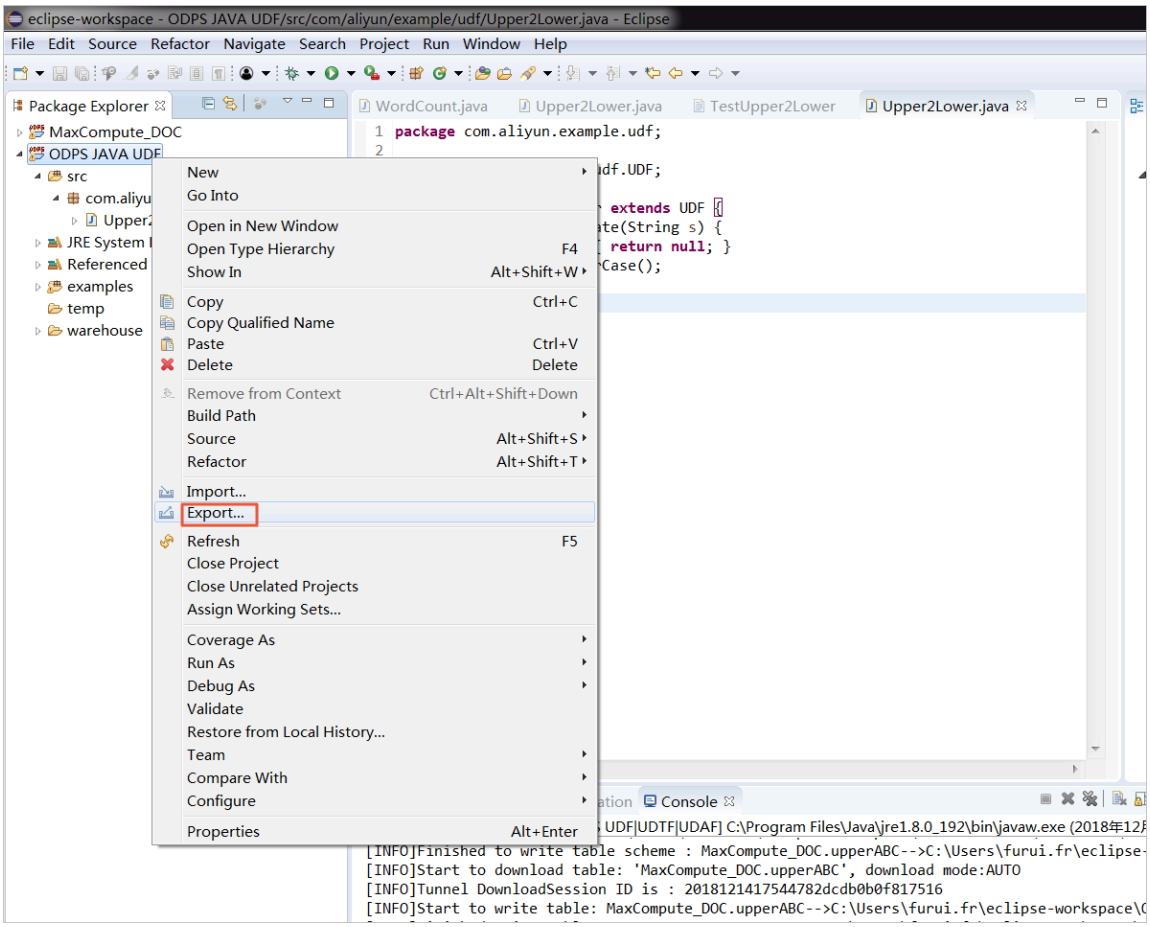
测试结果只是Eclipse获取表格中的数据后在本地转换的结果，并不代表MaxCompute中的数据已经转换为小写的aliyun了。

使用Java UDF

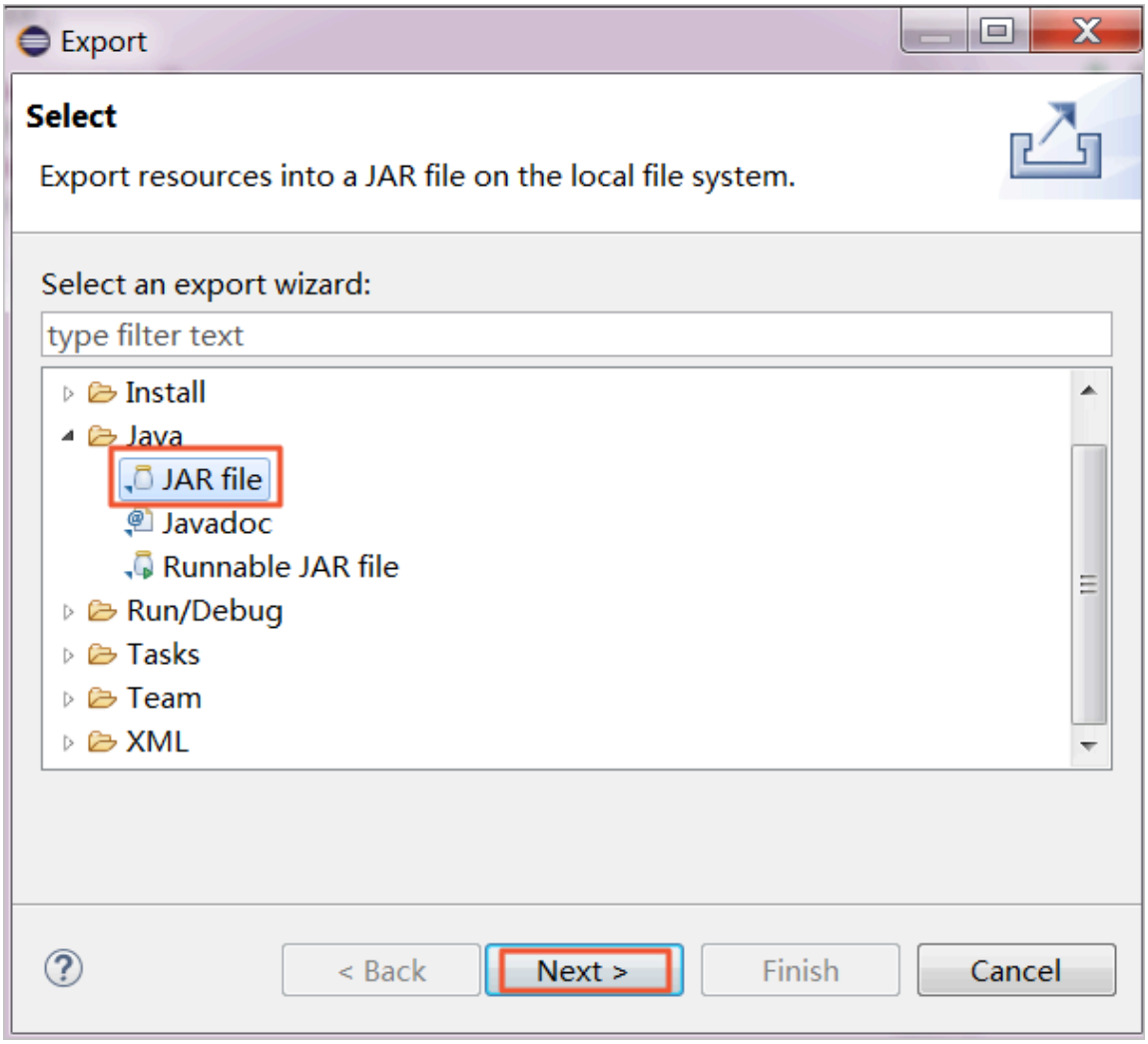
确定测试结果正确后，即可开始使用Java UDF，操作步骤如下：

1. 导出Jar包

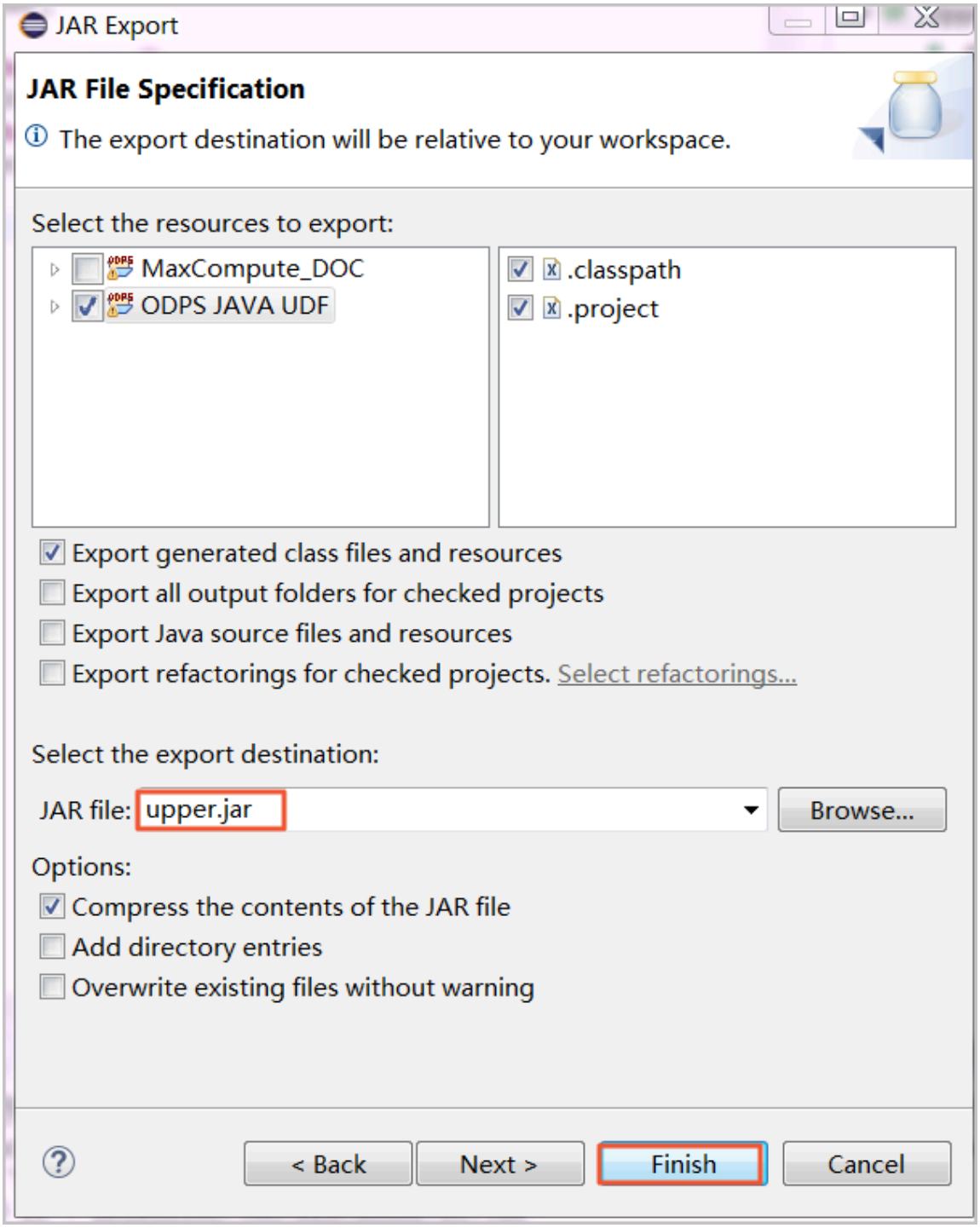
a. 右键单击左侧新建的ODPS Project，选择Export。



b. 在弹框中选择JAR file，单击Next。



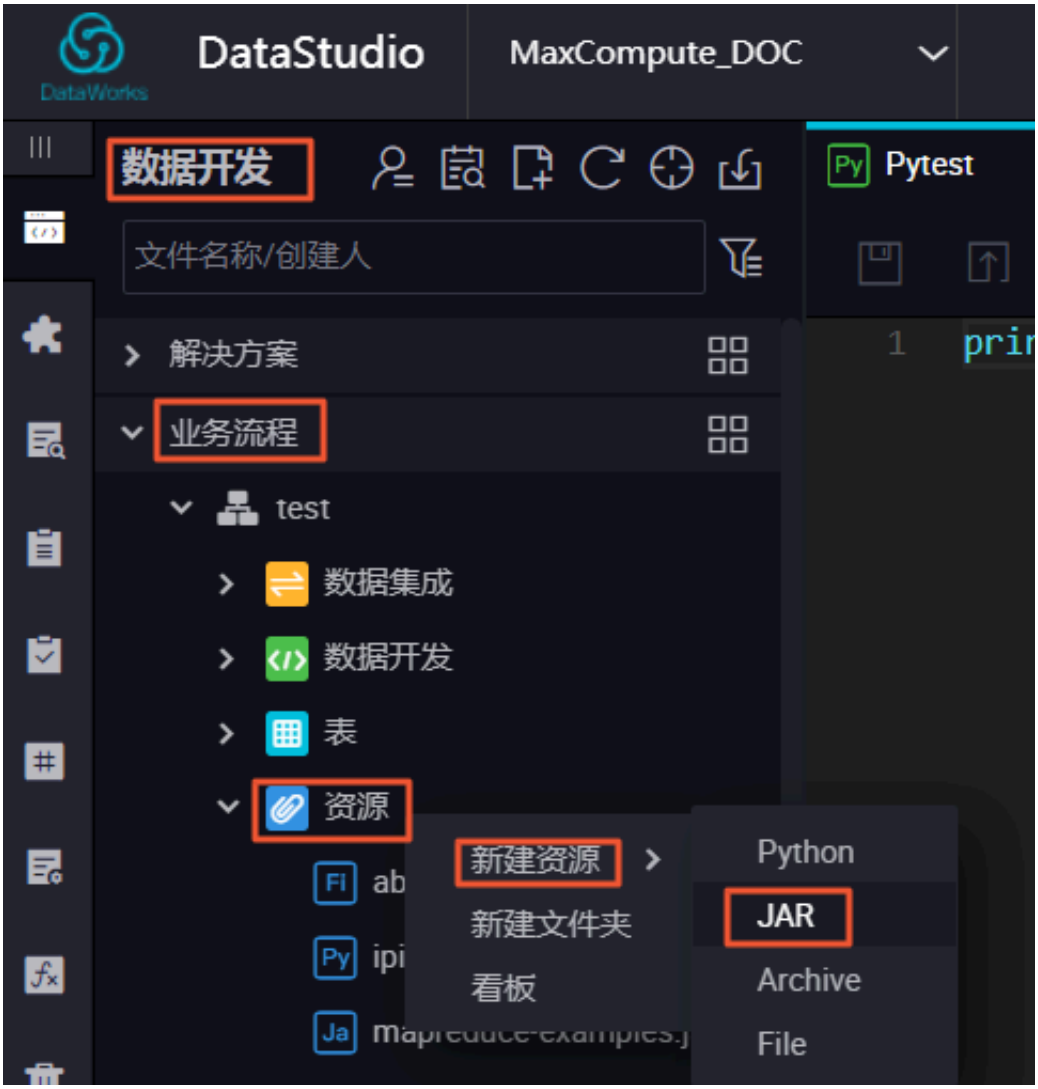
c. 在对话框中的JAR file处填写Jar包名称，单击Finish，即可导出至当前workspace目录下。


JAR Export
JAR File Specification
The export destination will be relative to your workspace.
Select the resources to export:
MaxCompute_DOC
ODPS JAVA UDF
.classpath
.project
Export generated class files and resources
Export all output folders for checked projects
Export Java source files and resources
Export refactorings for checked projects. Select refactorings...
Select the export destination:
JAR file: upper.jar
Browse...
Options:
Compress the contents of the JAR file
Add directory entries
Overwrite existing files without warning
< Back
Next >
Finish
Cancel

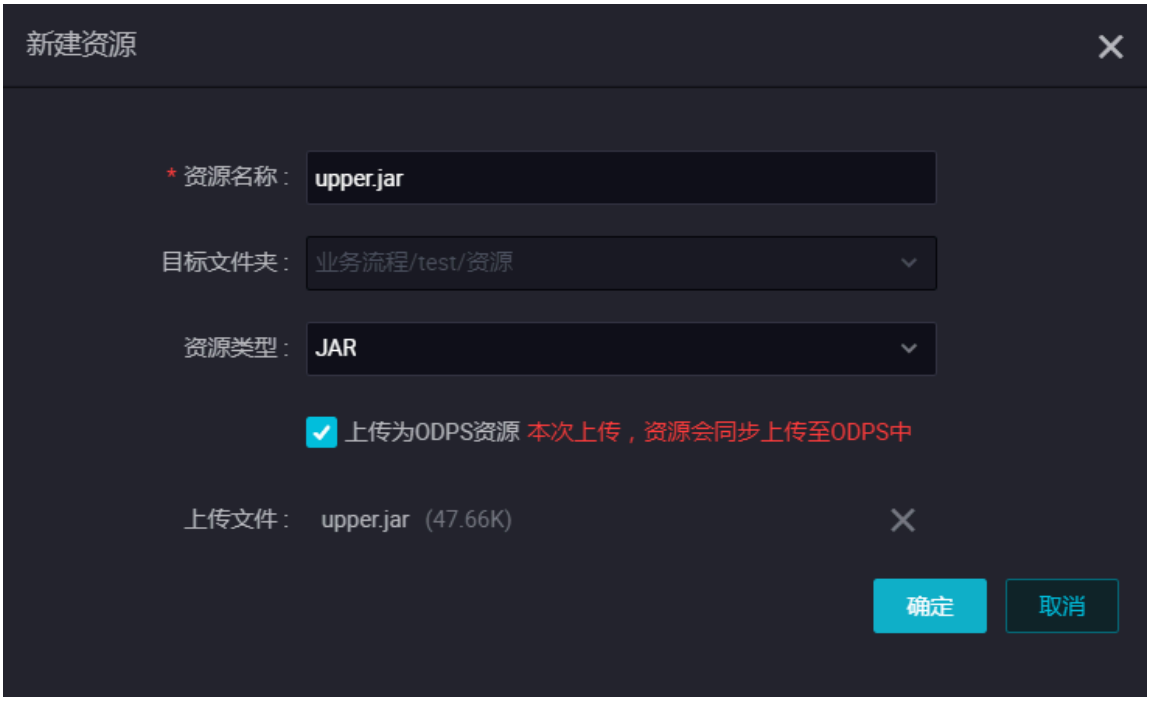
2. 使用DataWorks引用Jar包

a. 登录DataWorks控制台，进入同一个项目（本例中为项目MaxCompute_DOC）的数据开发页面。

b. 选择业务流程 > 资源 > 新建资源 > JAR，新建一个JAR类型的资源。详情请参见资源。



c. 在弹窗中上传您刚导出的JAR资源至DataWorks。



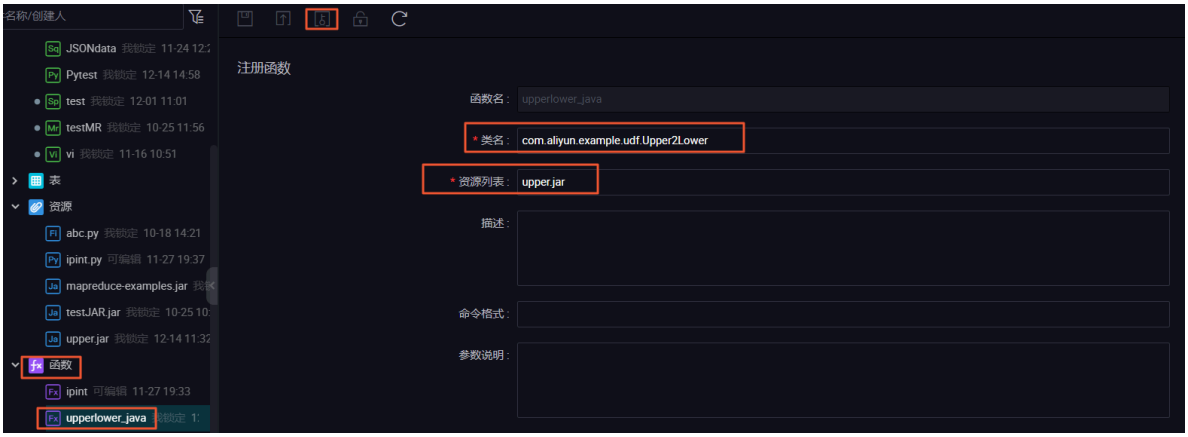
d. 单击确定，进入已上传的JAR资源页面。

e. 单击提交并解锁（提交）按钮，即可将资源上传至MaxCompute。

f. 完成上传后，您可以在odpscmd客户端使用`list resources`命令查看您上传的JAR资源。

3. 创建资源函数

现在JAR资源已经存在于您的MaxCompute项目中，接下来需要打开相应的业务流程，右键单击函数，选择新建函数，新建一个与Jar资源对应的函数，本示例的函数名称为upperlower_Java。详情请参见函数。



创建完成后，依次单击保存和提交并解锁（提交）。

完成提交后，您可以在odpscmd客户端使用`list functions`命令，查看已注册的函数。至此，您使用Eclipse开发工具注册的Java UDF函数upperlower_Java已经可用。

使用Java UDF结果验证

打开您的odpscmd命令行界面，运行`select upperlower_Java('ABCD') from dual;`命令，即可观察到该Java UDF成功转换字母的大小写，函数运行正常。

更多信息

更多Java UDF开发示例请参见Java UDF。

如果您要使用IntelliJ IDEA开发工具完成完整的Java UDF开发过程，请参见#unique_75。

3.2 IntelliJ IDEA Java UDF开发最佳实践

IntelliJ IDEA是Java语言的集成开发环境，可以帮助我们快速的开发Java程序。本文为您详细介绍如何使用IntelliJ IDEA进行Java UDF开发。

前提条件

在开始UDF开发实践之前，您需要做如下准备工作：

1. 准备IntelliJ IDEA开发工具，请参见安装Studio。
2. 通过IntelliJ IDEA MaxCompute Studio创建MaxCompute项目连接。
3. 连接MaxCompute项目成功后，您需要创建MaxCompute Java Module。

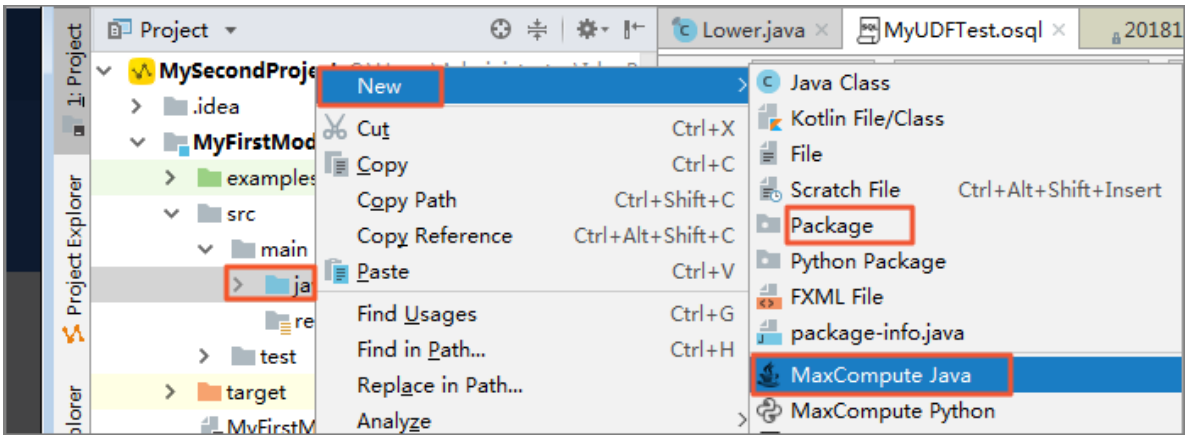
开发环境准备完成后即可开发UDF，下面将为您介绍一个字符小写转换功能的UDF实现示例。

> 说明：
> 更多UDF开发的相关资料请参见Java UDF。

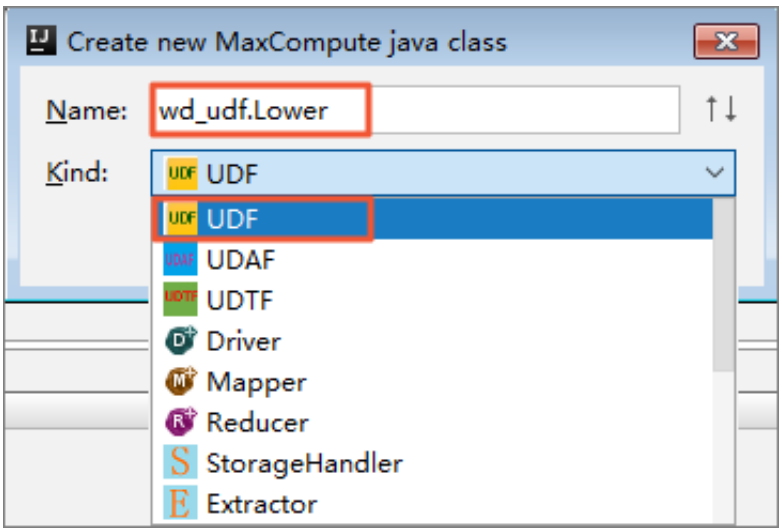
操作步骤

1. 创建Java UDF Project

   首先在您的IntelliJ IDEA中展开已创建的MaxCompute Java Module目录，导航至src > main > java > New，单击MaxCompute Java，如下图所示。



   填写Name，输入package名称.文件名，Kind选择UDF，单击OK，如下图所示。

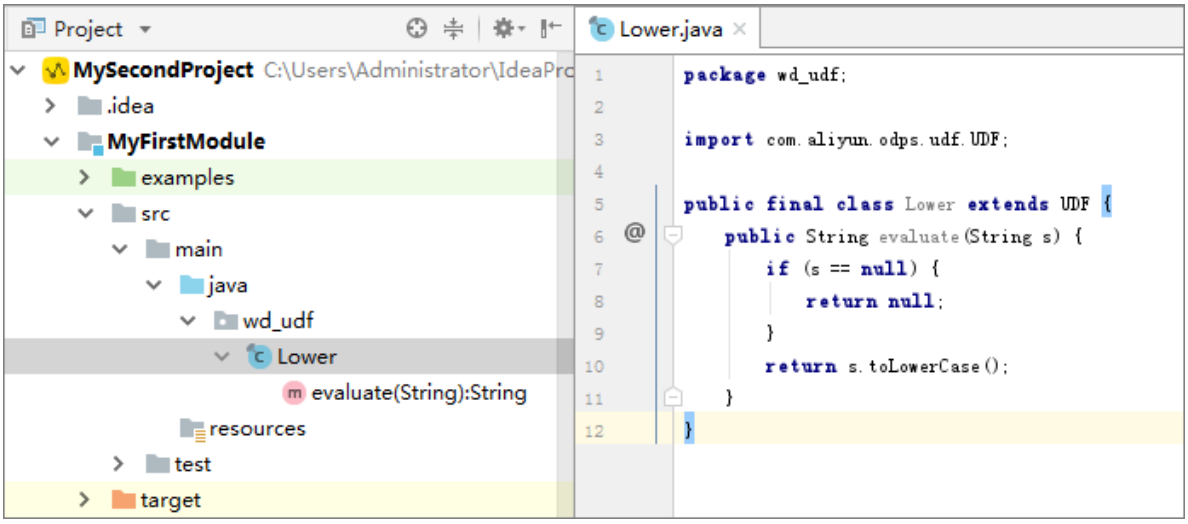


> 说明：

- Name：填写创建的MaxCompute Java Class名称，如果还没创建package，可以在此处填写packagename.classname，会自动生成package。
- Kind：选择类型。目前支持的类型有：自定义函数（UDF/UDAF/UDTF）、MapReduce（Driver/Mapper/Reducer）、非结构化开发（StorageHandler/Extractor）等。

2. 编辑Java UDF代码

在您新建的Java UDF项目（本例中的porjectLower）中编辑代码，如下图。



示例代码如下。

```
package <package名称>;
import com.aliyun.odps.udf.UDF;
public class Lower extends UDF {
   public String evaluate(String s) {
      if (s == null) {
          return null;
      }
      return s.toLowerCase();
   }
}
```

说明：

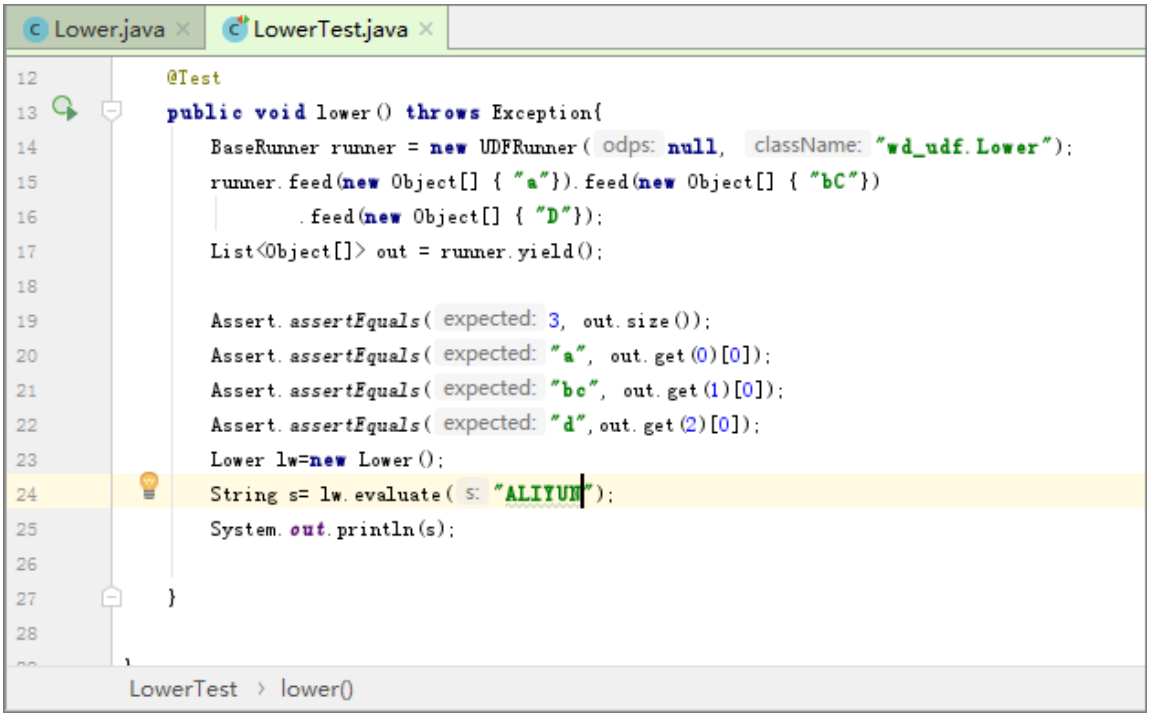
这里的代码模板可在您的Intellij IDEA中自定义，具体操作路径：Settings > Editor > File Code Templates，然后在Code标签页中寻找MaxCompute对应的模板修改。

3. 测试UDF

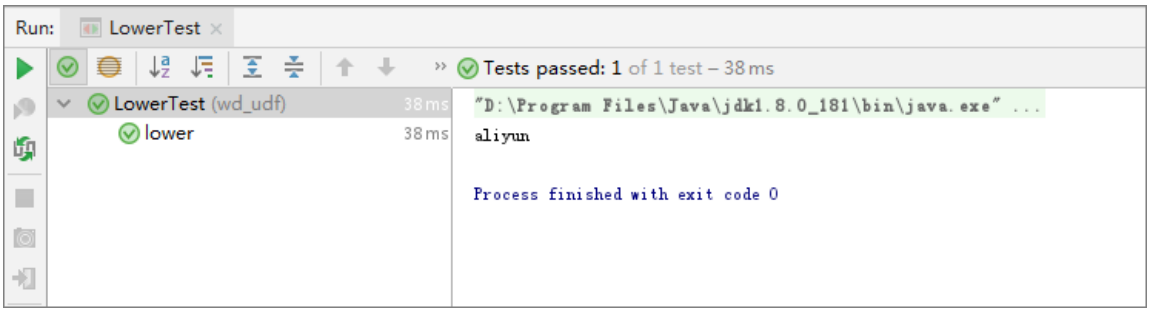
开发UDF完成后，可通过单元测试和本地运行两种方式进行测试，看是否符合预期结果，操作如下。

a) 单元测试

在您的Modul项目中examples目录下有各种类型的单元测试示例，您可参考示例编写自己的Unit Test。



测试结果如下。



我们可以看到大写字母"ALIYUN"已经成功转换成小写字母"aliyun"输出。

b) 本地运行

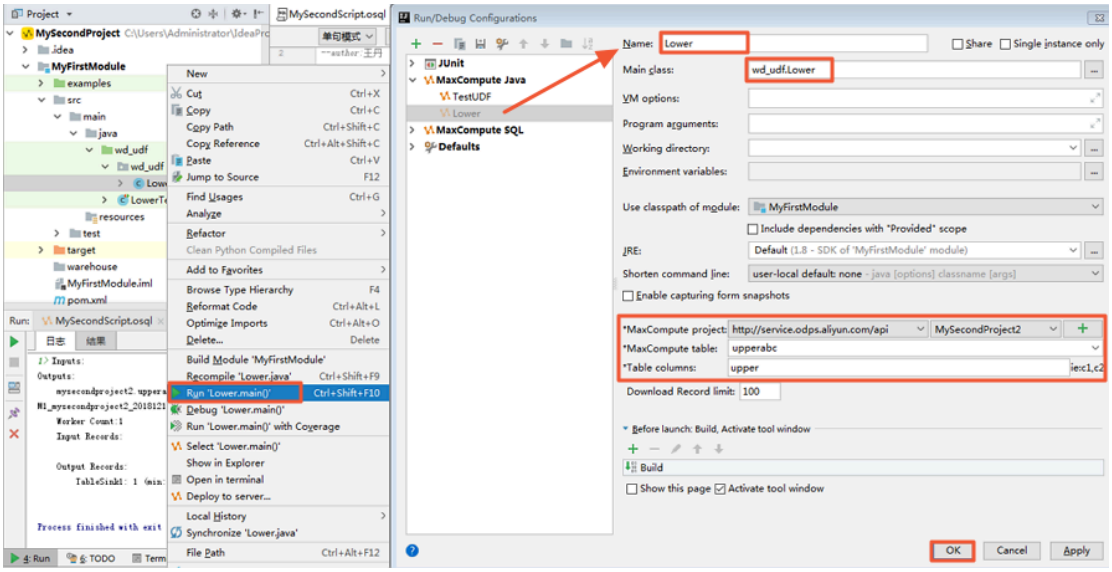
在您的IntelliJIDEA中本地运行UDF时，需要指定运行数据源，有以下两种方式设定测试数据源：

- MaxCompute Studio通过Tunnel服务自动下载指定项目下的表数据到warehouse目录下。
- 提供Mock项目及表数据，即您可参考warehouse下的example_project自行设置。

操作步骤

A. 为了测试Java UDF代码，我们可以首先在MaxCompute上存放一些大写字母作为输入数据。您可以利用script脚本文件或者odpscmd客户端使用SQL语句`create table upperABC(upper string);`新建一个名为upperABC的测试表格，如图。

B. 右击UDF类，单击Run '类名.main()'，弹出run configurations对话框，如下图。

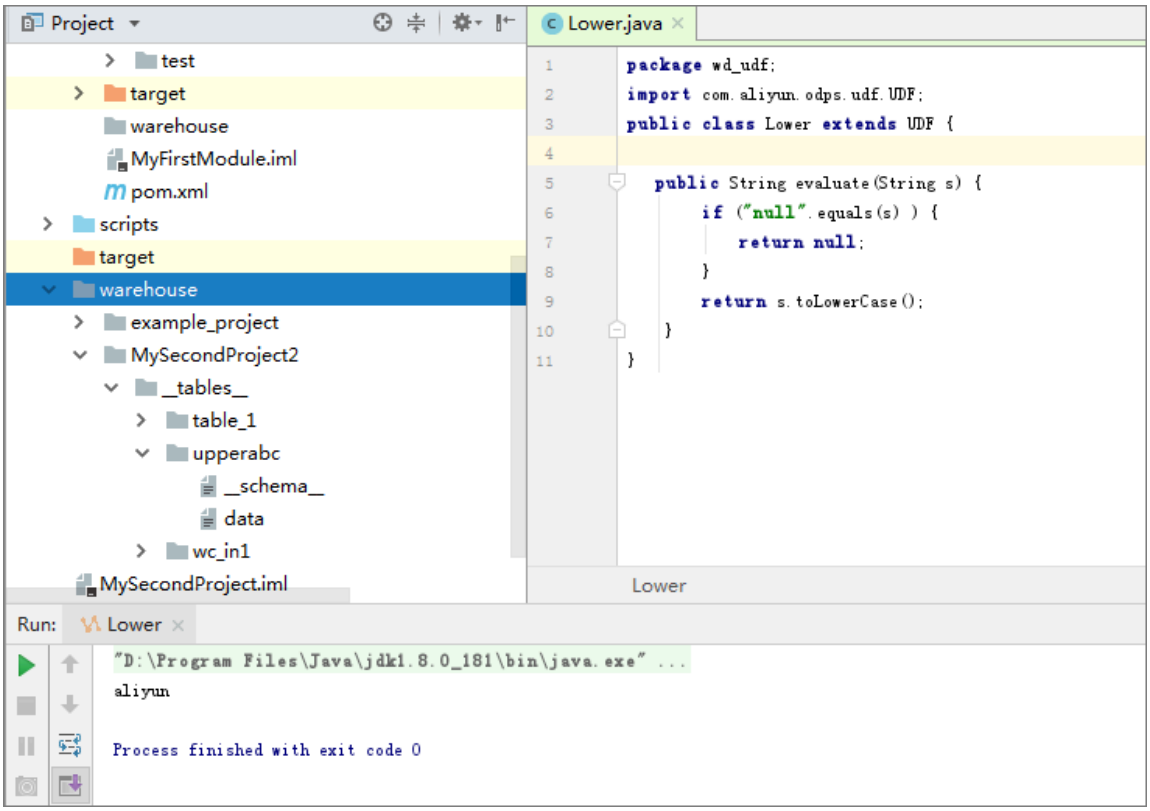


说明：

- UDF/UDAF/UDTF一般作用于select子句中表的某些列，需要配置MaxCompute project，table和column（元数据来源于project explorer和warehouse下的Mock项目）。复杂类型的调试也是支持的。
- 如果指定项目下的表数据未被下载到warehouse中，需要先下载数据，默认下载100条。默认下载100条，如需更多数据，可配置Download record limit项。
- UDF的local run框架会将warehouse中指定列的数据作为UDF的输入，开始本地运行UDF，您可以在控制台看到日志输出和结果打印。

- 如果采用Mock项目或已下载数据，则直接运行。

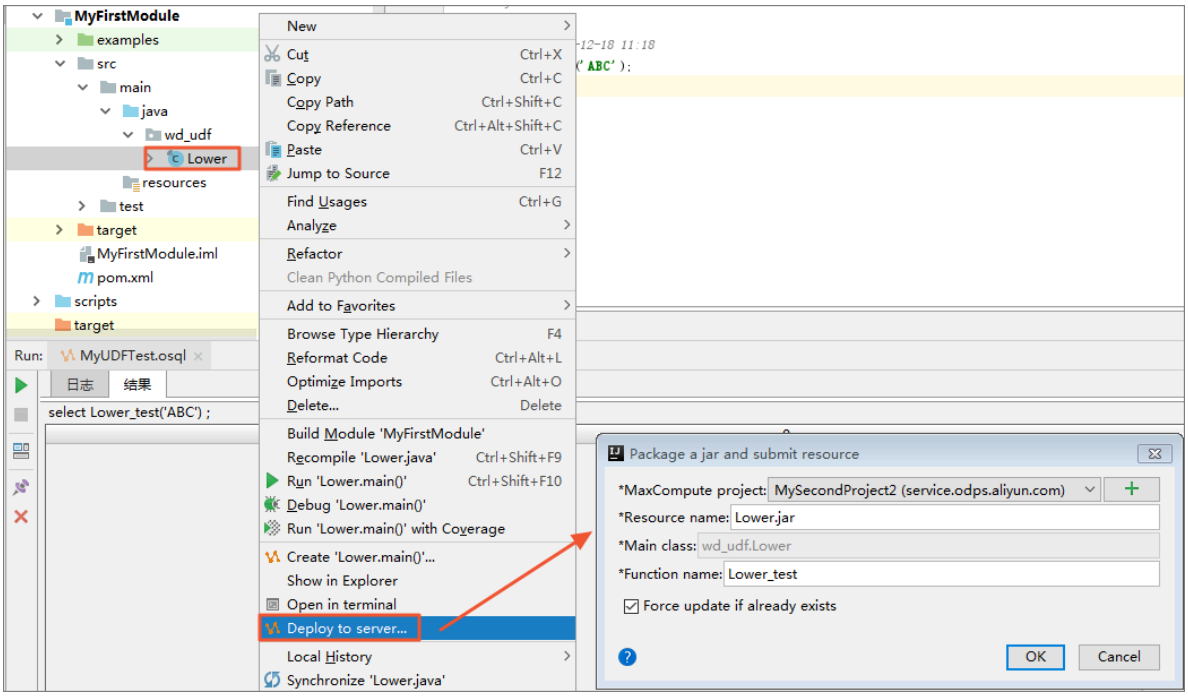
单击OK，运行结果如下图。



4. 发布UDF

此时我们的*Lower.java*测试通过，接下来我们要将其打包成jar资源上传到MaxCompute服务端上。一个UDF要想发布到服务端供生产使用，要经过打包 > 上传 > 注册三个步骤。针对此，IntelliJ IDEA MaxCompute Studio提供了一键发布功能（Studio会依次执行maven clean package，上传jar和注册UDF三个步骤，一次完成）。具体的操作如下。右键单击UDF的Java文件，选择Deploy to server，弹框里选择注册到哪一个MaxCompute

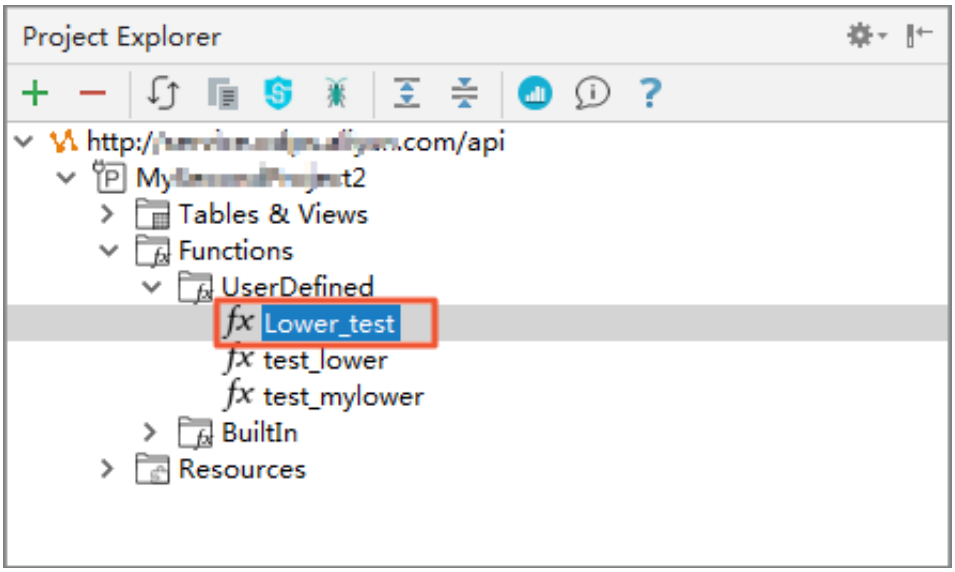
project，依次输入Function name和Resource name，Resource name可以修改，如下图。



说明：

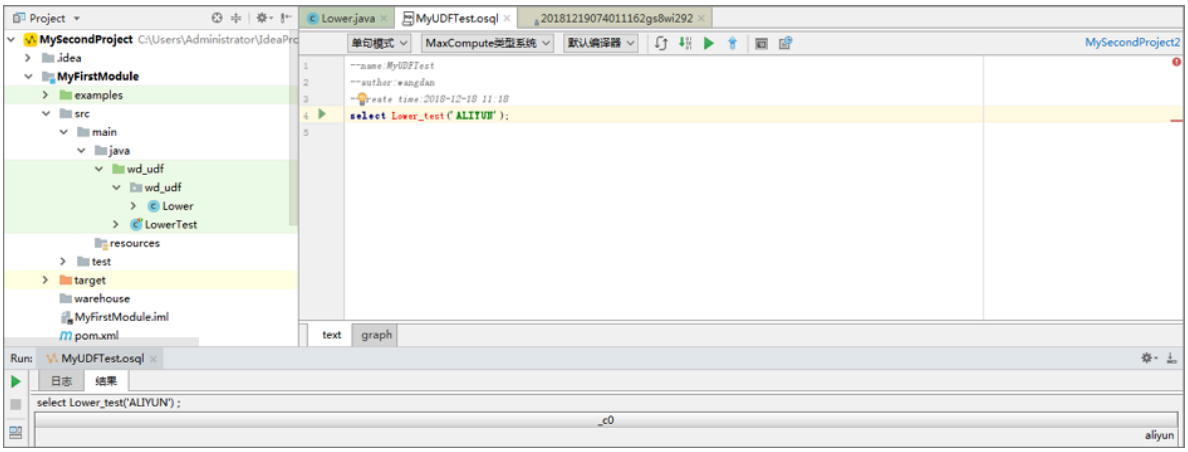
如果您想了解打包、上传和注册的详细操作步骤，请参见#unique_80。

填写完成后，单击OK即可完成注册，成功后会有提示。您可在连接的MaxCompute项目下找到已经注册好的Function函数，如图所示。



5. 试用UDF

成功注册UDF后，即可试用UDF。在您的Module项目中打开SQL脚本，执行命令`select Lower_test('ALIYUN');`，显示结果如下图所示。



您也可以在odpscmd客户端使用`select Lower_test('ALIYUN') from uppperABC;`命令测试您的Java UDF函数。到此，您使用IntelliJIDEA上开发的Java UDF函数Lower_test已经可用了。

后续步骤

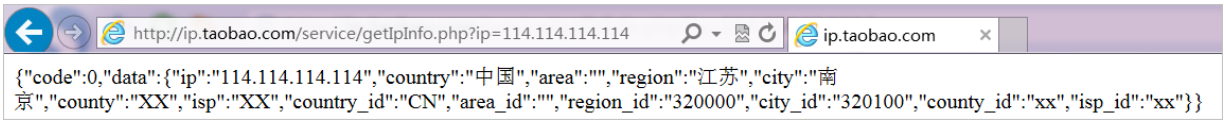
如果您要使用Eclipse开发工具完成完整的Java UDF开发流程，请参见#unique_81。

3.3 使用MaxCompute分析IP来源最佳实践

本文将为您介绍如何在MaxCompute上分析IP来源，包括下载、上传IP地址库数据、编写UDF函数和编写SQL四个步骤。

背景介绍

淘宝IP地址库的查询接口为IP地址字串，使用示例如下。



由于在MaxCompute中禁止使用HTTP请求，目前可以通过以下三种方式，实现在MaxCompute中查询IP。

- 用SQL将数据查询到本地，再发起HTTP请求查询。

  说明:
  效率低下，且淘宝IP库查询频率需小于10QPS，否则拒绝请求。

- 下载IP地址库到本地，进行查询。

  说明:
  同样效率低，且不利于数仓等分析使用。

- 将IP地址库定期维护上传至MaxCompute，进行连接查询。本文重点为您介绍该方式。

  说明:
  比较高效，但是IP地址库需自己定期维护。

下载IP地址库

1. 首先您需要获取地址库数据。地址库您可以自行获取，本文仅提供一个UTF8格式的不完整的地址库demo。
2. 下载UTF-8地址库数据到本地后，检查数据格式，示例如下。

```
0,16777215,"0.0.0.0","0.255.255.255","","","内网IP","内网IP","内网IP"
16777216,16777471,"1.0.0.0","1.0.0.255","澳大利亚","","","",""
16777472,16778239,"1.0.1.0","1.0.3.255","中国","福建省","福州市","","电信"
```

前四个数据是IP地址的起始地址与结束地址：前两个是十进制整数形式，后两个是点分形式。这里我们使用整数形式，以便计算IP是否属于这个网段。

说明:
如果您需要使用真实IP地址，请自行下载IP地址库，具体的下载地址和使用方式请参见云栖社区。

上传IP地址库数据

1. 创建表DDL，您可以使用MaxCompute客户端进行操作，也可以使用DataWorks进行图形化建表。

```
DROP TABLE IF EXISTS ipresource ;

CREATE TABLE IF NOT EXISTS ipresource
(
    start_ip BIGINT
    ,end_ip BIGINT
    ,start_ip_arg string
    ,end_ip_arg string
    ,country STRING
    ,area STRING
```

```
    ,city STRING
    ,county STRING
    ,isp STRING
);
```

2. 使用#unique_83上传您的文件，本例中ipdata.txt.utf8文件存放在D盘。

```
odps@ workshop_demo>tunnel upload D:/ipdata.txt.utf8 ipresource;
```

可以通过SQL语句`select count(*) from ipresource;`查看表中上传的数据条数（由于地址库有人更新维护，所以条目数会不断增长）。

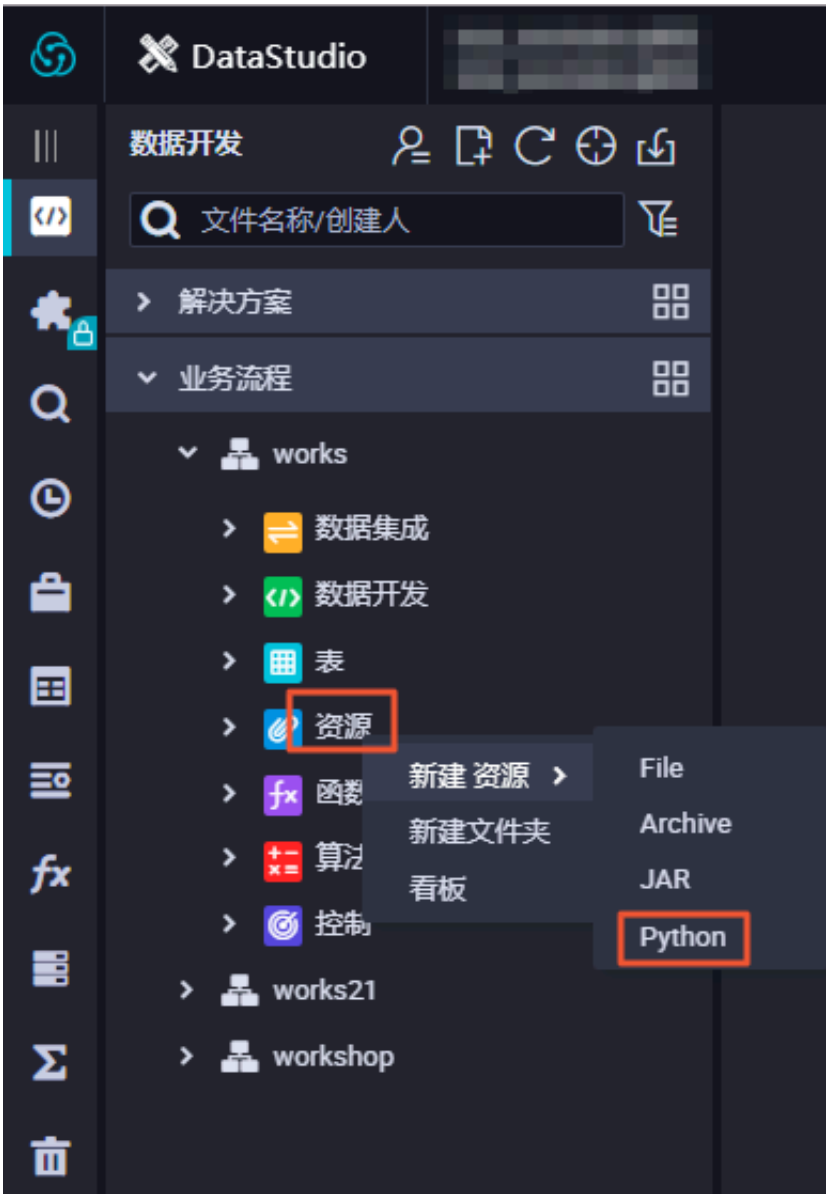
3. 使用SQL语句`select * from ipresource limit 10;`查看ipresource表前10条的样本数据，示例如下。

```
Job Queueing...
+------------+------------+--------------+------------+---------+------+------+--------+-----+
| start_ip   | end_ip     | start_ip_arg | end_ip_arg | country | area | city | county | isp |
+------------+------------+--------------+------------+---------+------+------+--------+-----+
| 3395369026 | 3395369026 | "202.97.56.66" | "202.97.56.66" | "中国" | "湖南省" | "长沙市" | ""     | "电信" |
| 3395369027 | 3395369028 | "202.97.56.67" | "202.97.56.68" | "中国" | "黑龙江省" | ""   | ""     | "电信" |
| 3395369029 | 3395369029 | "202.97.56.69" | "202.97.56.69" | "中国" | "安徽省" | "合肥市" | ""     | "电信" |
| 3395369030 | 3395369030 | "202.97.56.70" | "202.97.56.70" | "中国" | "湖南省" | "长沙市" | ""     | "电信" |
| 3395369031 | 3395369033 | "202.97.56.71" | "202.97.56.73" | "中国" | "黑龙江省" | ""   | ""     | "电信" |
| 3395369034 | 3395369034 | "202.97.56.74" | "202.97.56.74" | "中国" | "湖南省" | "长沙市" | ""     | "电信" |
| 3395369035 | 3395369036 | "202.97.56.75" | "202.97.56.76" | "中国" | "黑龙江省" | ""   | ""     | "电信" |
| 3395369037 | 3395369037 | "202.97.56.77" | "202.97.56.77" | "中国" | "江苏省" | "南京市" | ""     | "电信" |
| 3395369038 | 3395369038 | "202.97.56.78" | "202.97.56.78" | "中国" | "湖南省" | "长沙市" | ""     | "电信" |
| 3395369039 | 3395369040 | "202.97.56.79" | "202.97.56.80" | "中国" | "黑龙江省" | ""   | ""     | "电信" |
+------------+------------+--------------+------------+---------+------+------+--------+-----+
```

编写UDF函数

通过编写Python UDF，将点号分割的IP地址转化为整数类型的IP地址，本示例使用DataWorks的PyODPS节点完成。

1. 右键单击相应业务流程下的资源，选择新建资源 > Python。

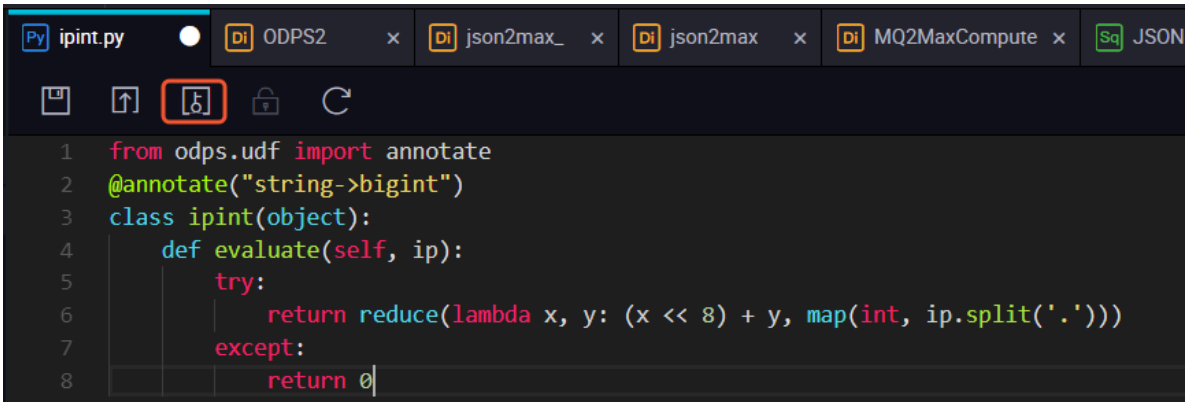


2. 在新建资源对话框中，填写资源名称，并勾选上传为ODPS资源，单击确定。
3. 在新建的Python资源内，编写Python资源代码，示例如下。

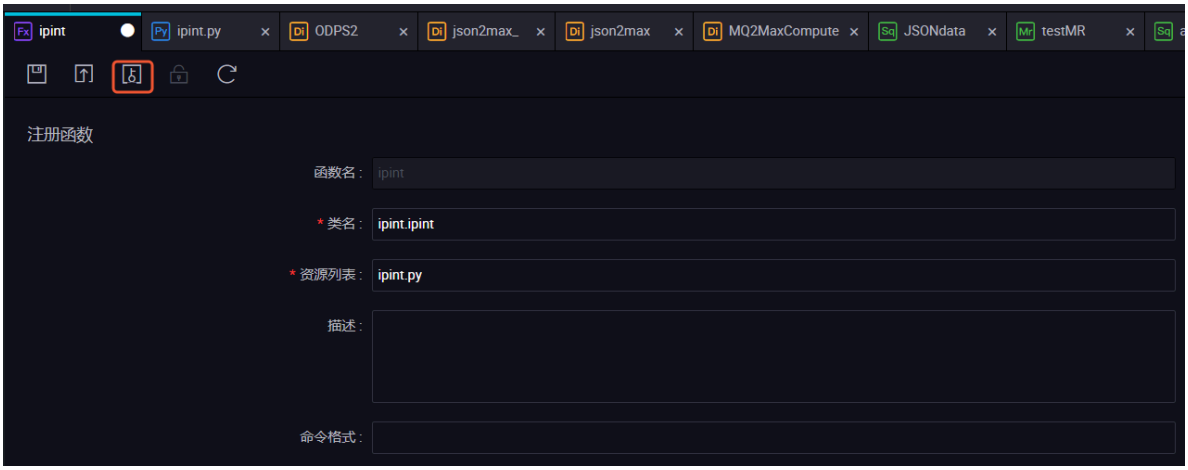
```
from odps.udf import annotate
@annotate("string->bigint")
class ipint(object):
    def evaluate(self, ip):
        try:
            return reduce(lambda x, y: (x << 8) + y, map(int, ip.
split('.')))
        except:
```

```
            return 0
```

单击提交并解锁。

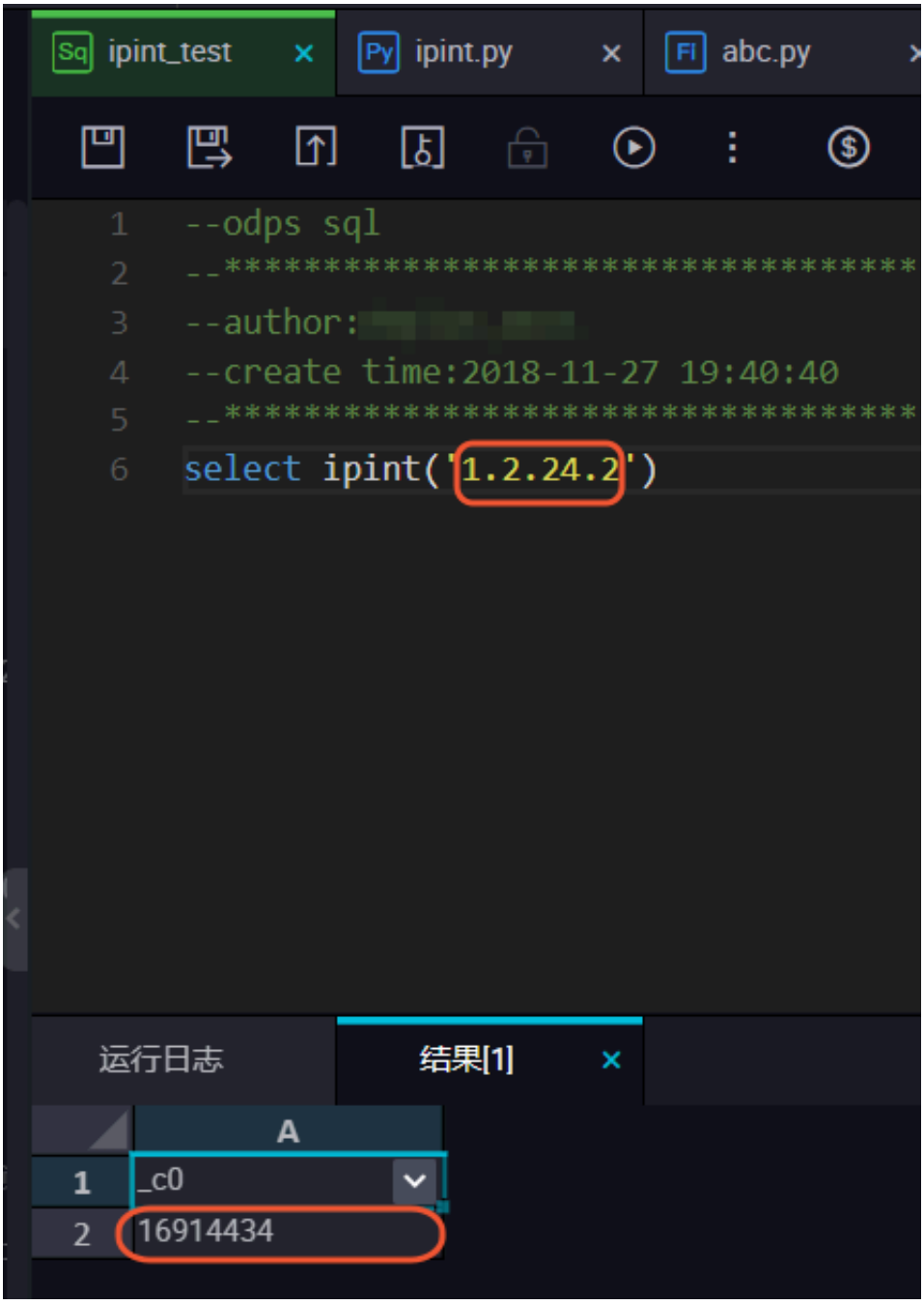


4. 右键单击相应业务流程下的函数，选择新建函数。
5. 在新建函数对话框中，填写函数名称，单击提交。
6. 编辑函数配置，单击提交并解锁。



本示例中，函数的类名为`ipint.ipint`，资源列表填写上文提交的资源的名称。

7. 验证ipint函数是否生效并满足预期，您可以在DataWorks上新建一个ODPS SQL类型节点，执行SQL语句进行查询，示例如下。



您也可以在本地创建*ipint.py*文件，使用MaxCompute客户端上传资源。

```
odps@ MaxCompute_DOC>add py D:/ipint.py;
OK: Resource 'ipint.py' have been created.
```

完成上传后，使用客户端注册函数。

```
odps@ MaxCompute_DOC>create function ipint as ipint.ipint using ipint.
py;
```

```
Success: Function 'ipint' have been created.
```

完成注册后，即可使用该函数。您可以在客户端运行`select ipint('1.2.24.2');`进行测试。

说明:

如果同一主账号下其他项目需要使用这个UDF，您可以进行跨项目授权。

1. 创建名为ipint的package。

```
odps@ MaxCompute_DOC>create package ipint;
OK
```

2. 将已经创建好的UDF函数加入package。

```
odps@ MaxCompute_DOC>add function ipint to package ipint;
OK
```

3. 允许另外一个项目bigdata_DOC安装这个package。

```
odps@ MaxCompute_DOC> allow project bigdata_DOC to install package
ipint;
OK
```

4. 切换到另一个需要使用UDF的项目bigdata_DOC，安装package。

```
odps@ MaxCompute_DOC>use bigdata_DOC;
odps@ bigdata_DOC>install package MaxCompute_DOC.ipint;
OK
```

5. 现在您就可以使用这个UDF函数了， 如果项目空间bigdata_DOC的用户Bob需要访问这些资源，那么管理员可以通过ACL给Bob自主授权。

```
odps@ bigdata_DOC>grant Read on package MaxCompute_DOC.ipint to
user aliyun$bob@aliyun.com; --通过ACL授权Bob使用package
```

在SQL中使用

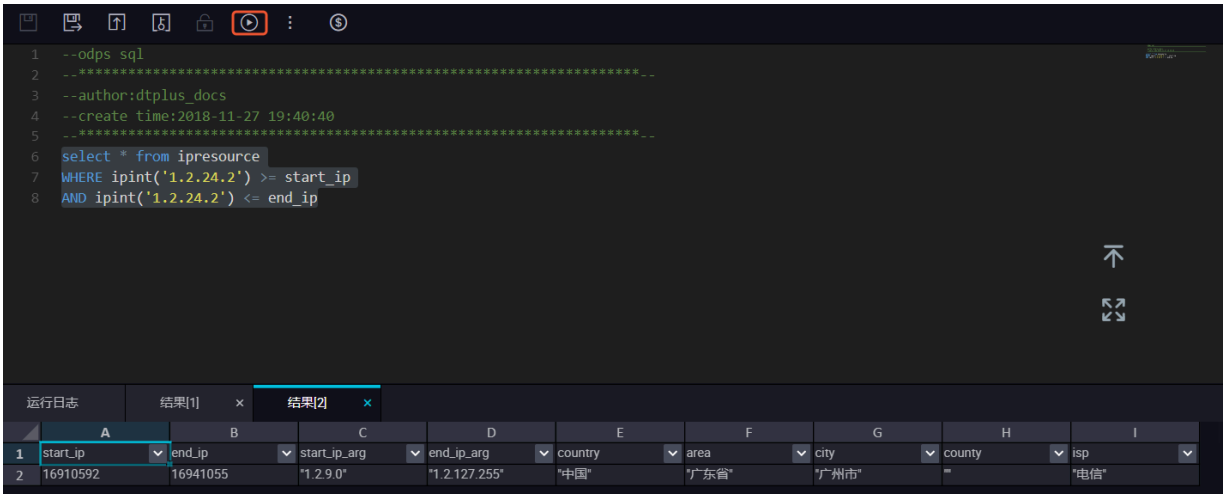
说明:

本文以一个随机的IP 1.2.24.2地址为例，您在使用时可以用具体表的字段来读入。

测试使用的SQL代码如下，单击运行即可查看结果。

```
select * from ipresource
WHERE ipint('1.2.24.2') >= start_ip
```

```
AND ipint('1.2.24.2') <= end_ip
```



通过为保证数据准确性，您可以定期从淘宝IP库获取数据来维护ipresource表。

3.4 解决DataWorks 10M文件限制问题最佳实践

本文向您介绍如何解决用户在DataWorks上执行MapReduce作业的时候，文件大于10M的JAR和资源文件不能上传到DataWorks的问题，方便您使用调度功能定期执行MapReduce作业。

解决方案：

1. 将大于10M的resources通过MaxCompute客户端上传。
   - 客户端下载地址：#unique_62。
   - 客户端配置AK、EndPoint请参考：#unique_88和#unique_89。

   向客户端添加资源命令：

   ```
   //添加资源
   add jar C:\test_mr\test_mr.jar -f;
   ```

2. 目前通过MaxCompute客户端上传的资源，在DataWorks左侧资源列表是找不到的，只能通过list resources查看确认资源：

   ```
   //查看资源
   list resources;
   ```

3. 减小Jar文件(瘦身Jar)，因为DataWorks执行MR作业的时候，一定要在本地执行，所以保留一个main函数就可以。

   ```
   jar
   -resources test_mr.jar,test_ab.jar --resources在客户端注册后直接引用
   -classpath test_mr.jar --瘦身策略：在gateway上提交要有main和相关的mapper和reducer，额外的三方依赖可以不需要，其他都可以放到resources
   ```

```
com.aliyun.odps.examples.mr.test_mr wc_in wc_out;
```

通过上述方法，我们可以在DataWorks上运行大于10M的MR作业。

3.5 实现指定用户访问指定UDF最佳实践

本文将为您介绍如何实现将具体的某个资源（表、UDF等）设置为仅能被指定的用户访问。此UDF涉及到数据的加密解密算法，属于数据安全管控范畴。

前提条件

您需要提前安装MaxCompute客户端，以实现指定UDF被指定用户访问的操作。详情请参见#unique_88。

常见方案

- Package方案，通过打包授权进行权限精细化管控。

  Package通常是为了解决跨项目空间的共享数据及资源的用户授权问题。当通过Package授予用户开发者角色后，用户则拥有所有权限，风险不可控。

  - 首先，用户熟知的DataWorks开发者角色的权限如下所示。

```
odps@ sz_mc>desc role role_project_dev;

Authorization Type: Policy
A       projects/sz_mc: *
A       projects/sz_mc/instances/*: *
A       projects/sz_mc/jobs/*: *
A       projects/sz_mc/offlinemodels/*: *
A       projects/sz_mc/packages/*: *
A       projects/sz_mc/registration/functions/*: *
A       projects/sz_mc/resources/*: *
A       projects/sz_mc/tables/*: *
A       projects/sz_mc/volumes/*:
```

    由上图可见，开发者角色对工作空间中的Package、Functions、Resources和Table默认有全部权限，明显不符合权限配置的要求。

  - 其次，通过DataWorks添加子账号并赋予开发者角色，如下所示。

```
odps@ sz_mc>show grants for RAM$y[redacted].pt@aliyun-test.com:ramtest;

[roles]
role_project_dev

Authorization Type: Policy
[role/role_project_dev]
A       projects/sz_mc: *
A       projects/sz_mc/instances/*: *
A       projects/sz_mc/jobs/*: *
A       projects/sz_mc/offlinemodels/*: *
A       projects/sz_mc/packages/*: *
A       projects/sz_mc/registration/functions/*: *
A       projects/sz_mc/resources/*: *
A       projects/sz_mc/tables/*: *
A       projects/sz_mc/volumes/*: *
```

    由此可见，通过打包授权和DataWorks默认的角色都不能满足我们的需求。例如将子账号`RAM$xxxxx.pt@aliyun-test.com:ramtest`授予开发者角色，则默认拥有当前工作空间中所有Object的所有操作权限，详情请参见用户授权。

- 在DataWorks中新建角色来进行高级管控。

  您可以进入DataWorks控制台中的工作空间配置 > MaxCompute高级配置 > 自定义用户角色页面，进行高级管控。但是在MaxCompute高级配置中只能针对某个表/某个项目进行授权，不能对资源和UDF进行授权。

- Role Policy方案，通过Role Policy自定义Role的权限集合。

  通过Policy可以精细化地管理到具体用户针对具体资源的具体权限粒度，下文将为您详细描述如何通过Role Policy方案实现指定UDF被指定用户访问。

说明:

为了安全起见，建议初学者使用测试项目来验证Policy。

通过Policy自定义Role的权限集合

1. 创建默认拒绝访问UDF的角色。

   a. 在客户端输入`create role denyudfrole;`，创建一个role denyudfrole。

   b. 创建Policy授权文件，如下所示。

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

[{
"Effect":"Deny",
"Action":["odps:Read","odps:List"],
"Resource":"acs:odps:*:projects/sz_mc/resources/getaddr.jar"
},
{
"Effect":"Deny",
"Action":["odps:Read","odps:List"],
"Resource":"acs:odps:*:projects/sz_mc/registration/functions/
getregion"
}
```

```
  ] }
```

c. **设置和查看Role Policy。**

   在客户端输入`put policy /Users/yangyi/Desktop/role_policy.json on role denyudfrole;`**命令，设置Role Policy文件的存放路径等配置。**

   通过`get policy on role denyudfrole;`**命令，查看Role Policy。**

```
odps@sz_mc>get policy on role denyudfrole;
{
    "Statement": [{
            "Action": ["odps:Read",
                "odps:List"],
            "Effect": "Deny",
            "Resource": ["acs:odps:*:projects/sz_mc/resources/getaddr.jar"]},
        {
            "Action": ["odps:Read",
                "odps:List"],
            "Effect": "Deny",
            "Resource": ["acs:odps:*:projects/sz_mc/registration/functions/getre
gion"]}],
    "Version": "1"}
```

d. **在客户端输入**`grant denyudfrole to RAM$xxxx.pt@aliyun-test.com:ramtest;`**，添加子账号至role denyudfrole。**

2. 验证拒绝访问UDF的角色是否创建成功。

   以子账号RAM$xxxx.pt@aliyun-test.com:ramtest登录MaxCompute客户端。

   a. 登录客户端输入whoami;确认角色。

```
odps@ sz_mc>whoami;
Name: RAM$y    pt@aliyun-test.com:ramtest
End_Point: http://service.odps.aliyun.com/api
Tunnel_End_Point: http://dt.cn-shanghai.maxcompute.aliyun.com
Project: sz_mc
```

   b. 通过show grants;查看当前登录用户权限。

```
odps@ sz_mc>show grants;

[roles]
role_project_dev, denyudfrole

Authorization Type: Policy
[role/denyudfrole]
D       projects/sz_mc/registration/functions/getregion: List | Read
D       projects/sz_mc/resources/getaddr.jar: List | Read
[role/role_project_dev]
A       projects/sz_mc: *
A       projects/sz_mc/instances/*: *
A       projects/sz_mc/jobs/*: *
A       projects/sz_mc/offlinemodels/*: *
A       projects/sz_mc/packages/*: *
A       projects/sz_mc/registration/functions/*: *
A       projects/sz_mc/resources/*: *
A       projects/sz_mc/tables/*: *
A       projects/sz_mc/volumes/*: *
```

   通过查询发现该RAM子账号有两个角色，一个是role_project_dev（即DataWorks默认的开发者角色），另一个是刚自定义创建的denyudfrole。

   c. 验证自建UDF以及依赖的包的权限。

```
odps@ sz_mc>desc function getregion;
FAILED: ODPS-0420095: Access Denied - Authorization Failed [4011], You have NO privilege 'odps:Read' on {acs:odps:*
 ID:38747231-386f-4018-b792-1da11700dc7e.       --->Tips: Pricipal:RAM$y    .pt@aliyun-test.com:ramtest by policy
odps@ sz_mc>desc resource getaddr.jar;
FAILED: ODPS-0420095: Access Denied - Authorization Failed [4011], You have NO privilege 'odps:Read' on {acs:odps:*:pro
1-5788-49b1-82b8-1f4a95c2f9c5.  --->Tips: Pricipal:RAM$y    .pt@aliyun-test.com
[DEBUG]: com.aliyun.odps.OdpsException: ODPS-0420095: Access Denied - Authorization Failed [4011], You have NO privileg
```

   通过上述验证发现，该子账号在拥有DataWorks开发者角色的前提下并没有自建UDF：getregion的读权限。但还需要结合Project Policy来实现该UDF只能被指定的用户访问。

3. 配置Project Policy。

   a. 编写Policy。

```
{
"Version": "1", "Statement":
[{
"Effect":"Allow",
"Principal":"RAM$yangyi.pt@aliyun-test.com:yangyitest",
"Action":["odps:Read","odps:List","odps:Select"],
"Resource":"acs:odps:*:projects/sz_mc/resources/getaddr.jar"
},
{
"Effect":"Allow",
 "Principal":"RAM$xxxx.pt@aliyun-test.com:yangyitest",
"Action":["odps:Read","odps:List","odps:Select"],
"Resource":"acs:odps:*:projects/sz_mc/registration/functions/
getregion"
```

```
}] }
```

b. 设置和查询Policy。

通过put policy /Users/yangyi/Desktop/project_policy.json;命令设置Policy文件的存放路径。

通过get policy;命令查看Policy。

```
odps@ sz_mc>get policy;
{
    "Statement": [{
            "Action": ["odps:Read",
                "odps:List",
                "odps:Select"],
            "Effect": "Allow",
            "Principal": ["RAM$[redacted].pt@aliyun-test.com:yangyitest"],
            "Resource": ["acs:odps:*:projects/sz_mc/resources/getaddr.jar"]},
        {
            "Action": ["odps:Read",
                "odps:List",
                "odps:Select"],
            "Effect": "Allow",
            "Principal": ["RAM$yangyi.pt@aliyun-test.com:yangyitest"],
            "Resource": ["acs:odps:*:projects/sz_mc/registr[redacted]
    "Version": "1"}
```

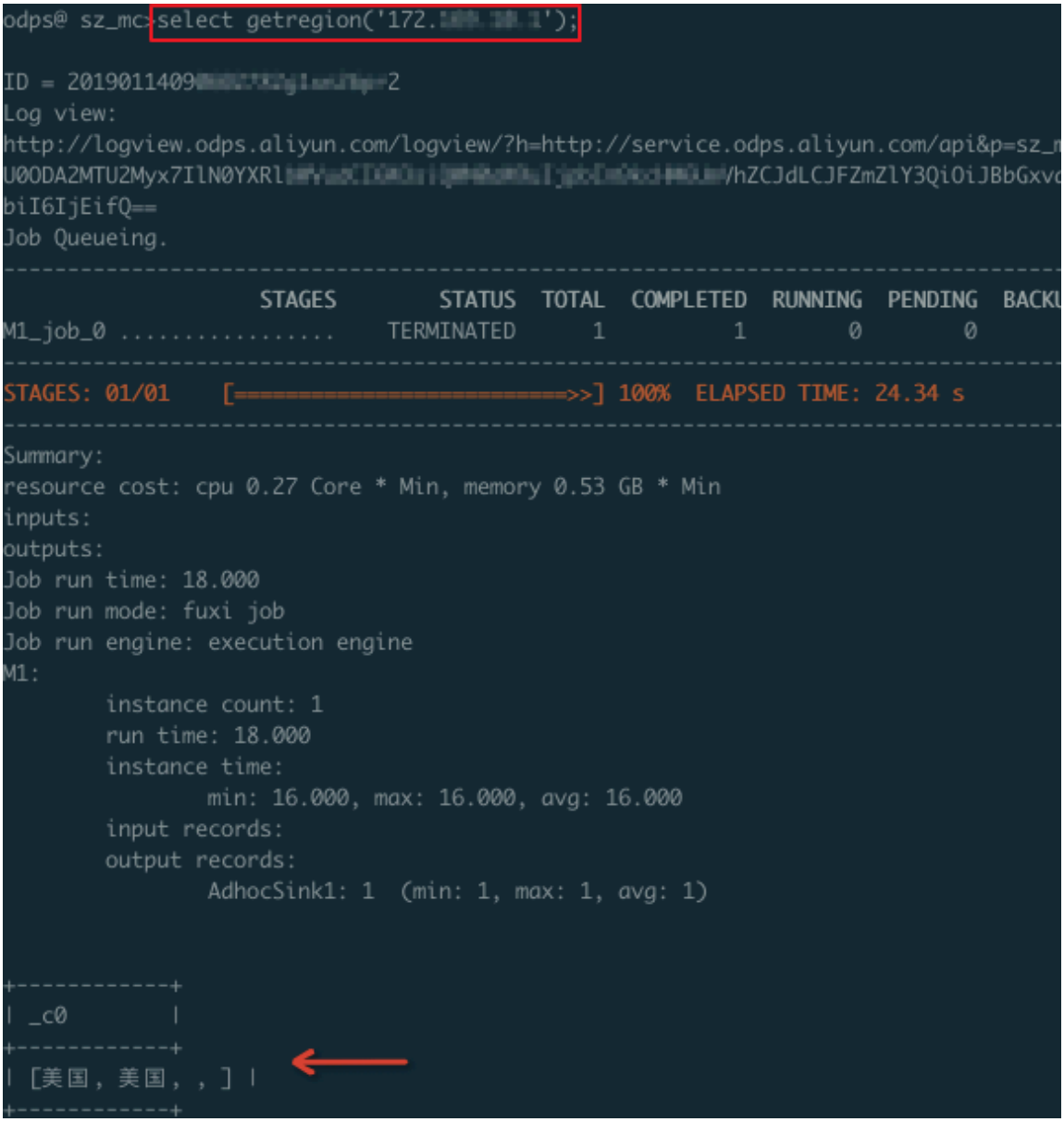
c. 通过whoami;和show grants;进行验证。

```
odps@ sz_mc>whoami;
Name: RAM$[redacted].pt@aliyun-test.com:yangyitest
End_Point: http://service.odps.aliyun.com/api
Tunnel_End_Point: http://dt.cn-shanghai.maxcompute.aliyun.com
Project: sz_mc
odps@ sz_mc>show grants;

[roles]
role_project_dev

Authorization Type: Policy
[role/role_project_dev]
A       projects/sz_mc: *
A       projects/sz_mc/instances/*: *
A       projects/sz_mc/jobs/*: *
A       projects/sz_mc/offlinemodels/*: *
A       projects/sz_mc/packages/*: *
A       projects/sz_mc/registration/functions/*: *
A       projects/sz_mc/resources/*: *
A       projects/sz_mc/tables/*: *
A       projects/sz_mc/volumes/*: *
[user/RAM$yangyi.pt@aliyun-test.com:yangyitest]
A       projects/sz_mc/registration/functions/getregion: List | Read | Select
A       projects/sz_mc/resources/getaddr.jar: List | Read | Select
```

d. 运行SQL任务，查看是否仅指定的RAM子账号能够查看指定的UDF和依赖的包。



总结

关于DataWorks和MaxCompute如何实现指定用户访问指定UDF，总结如下：

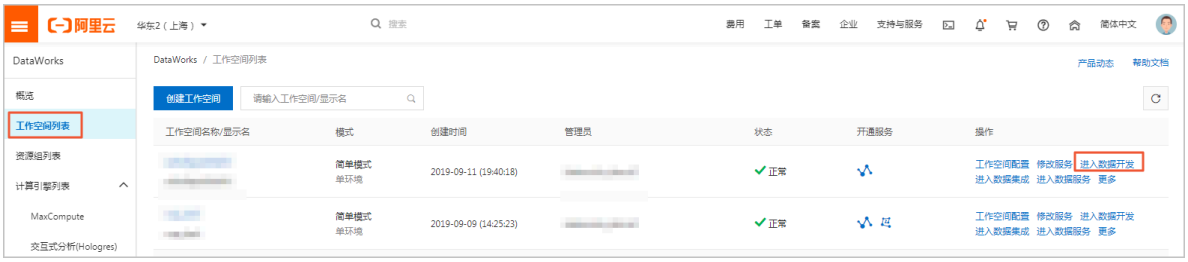
- 如果您不想让其他用户访问工作空间内具体的资源，在DataWorks中添加数据开发者权限后，再根据Role Policy的操作，在MaxCompute客户端将其配置为拒绝访问权限。
- 如果您要指定用户访问指定资源，通过Role Policy在DataWorks中配置数据开发者权限后，再根据Project Policy的操作，在MaxCompute客户端将其配置为允许访问权限。

3.6 PyODPS节点实现结巴中文分词

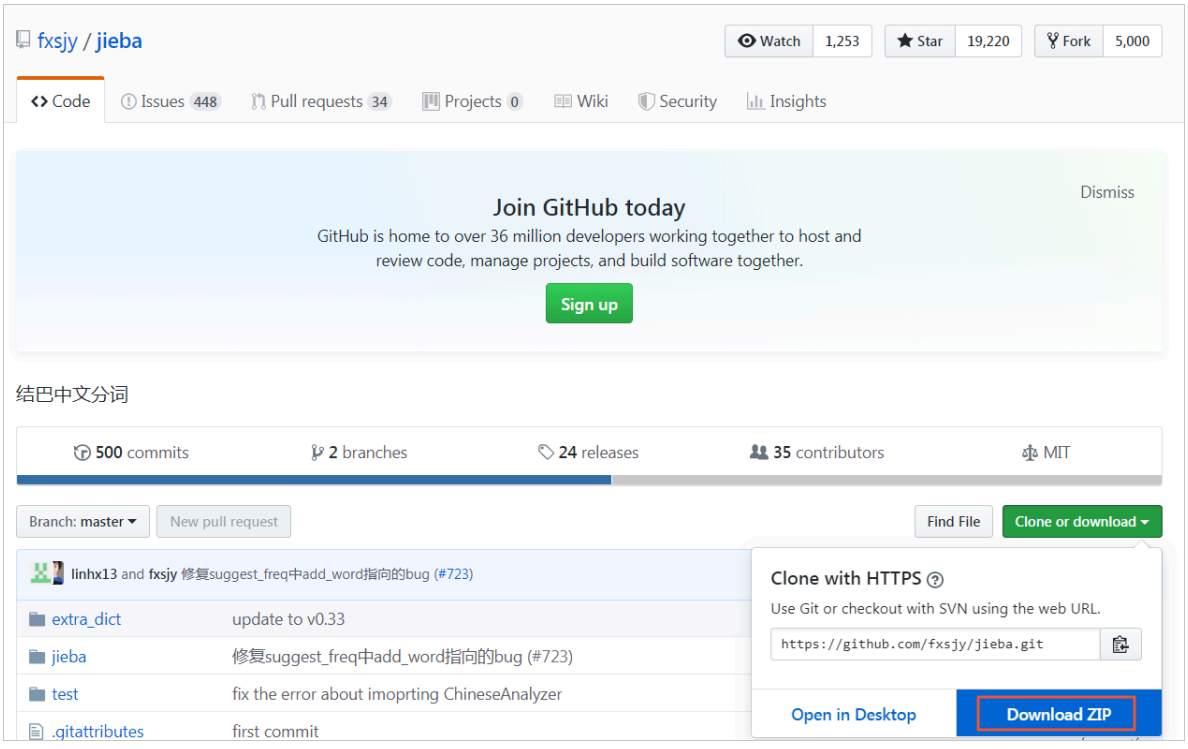
本文为您介绍如何使用DataWorks的PyODPS类型节点，借助开源结巴中文分词包实现对中文字段的分词并写入新的表。

前提条件

- 请首先确保您已经完成DataWorks工作空间的创建，本例使用简单模式的工作空间，详情请参见#unique_94。完成工作空间创建后，单击进入数据开发。



- 请在GitHub下载开源结巴分词中文包。



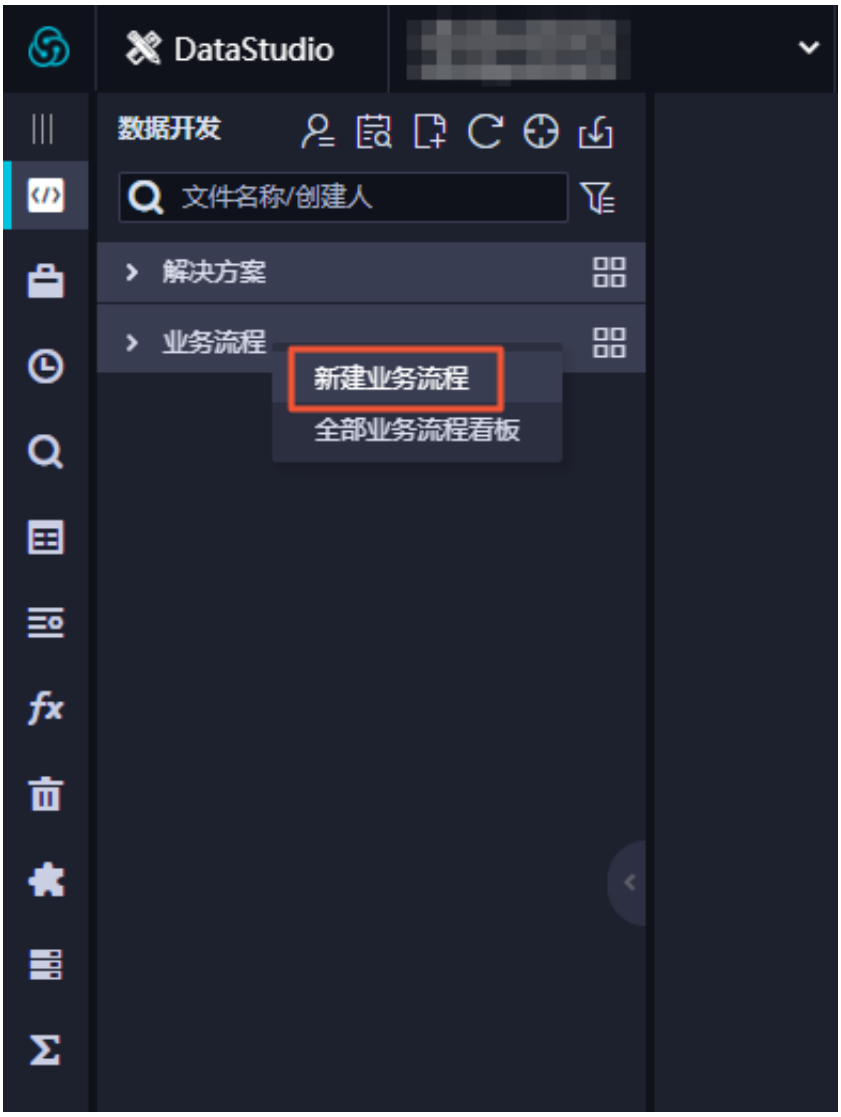
背景信息

PyODPS集成了Maxcompute的Python SDK。您可以在DataWorks的PyODPS节点上直接编辑Python代码并使用Maxcompute的Python SDK。关于PyODPS节点的详情请参见#unique_95。

操作步骤

1. 创建业务流程。

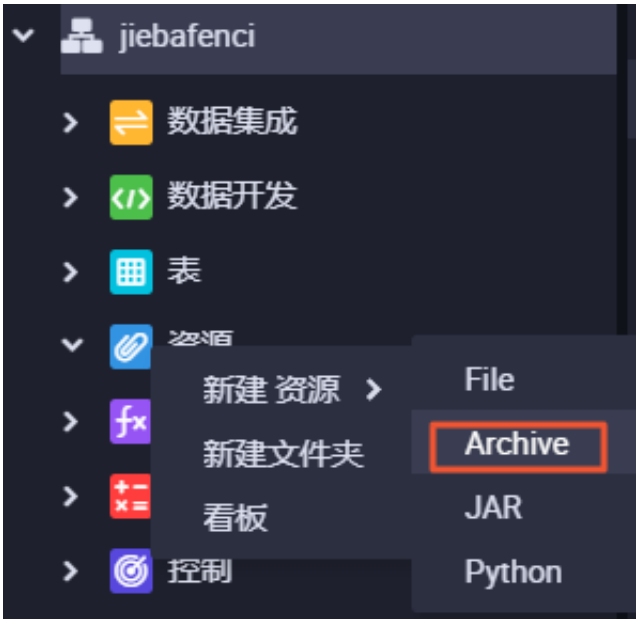
   a) 右键单击业务流程，选择新建业务流程。



   b) 输入您的业务流程名称后，单击新建。

2. 上传jieba-master.zip包。

a) 右键单击资源，选择Archive。

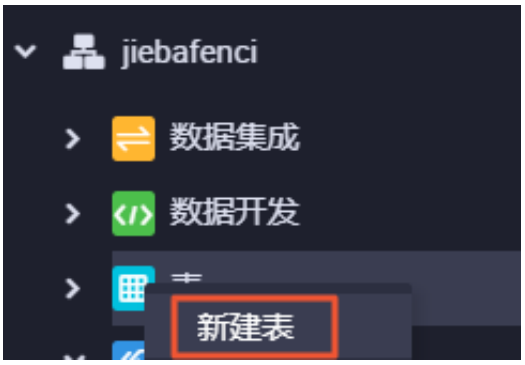


b) 上传您已下载到本地的jieba-master.zip，勾选上传为ODPS资源，单击确定。

c) 提交资源。

3. 创建测试数据表。

a) 右键单击表，选择新建表。输入表名jieba_test。



b) 单击DDL模式，输入建表DDL语句如下。

本教程准备了两列测试数据，您在后续开发过程中可以选择一列进行分词。

```
CREATE TABLE `jieba_test` (
    `chinese` string,
    `content` string
```

```
);
```

c) 单击提交到生产环境。

4. 创建测试结果存放表。

本例仅对测试数据的chinese列做分词处理，因此结果表仅有一列，创建方法同上。DDL语句如下所示。

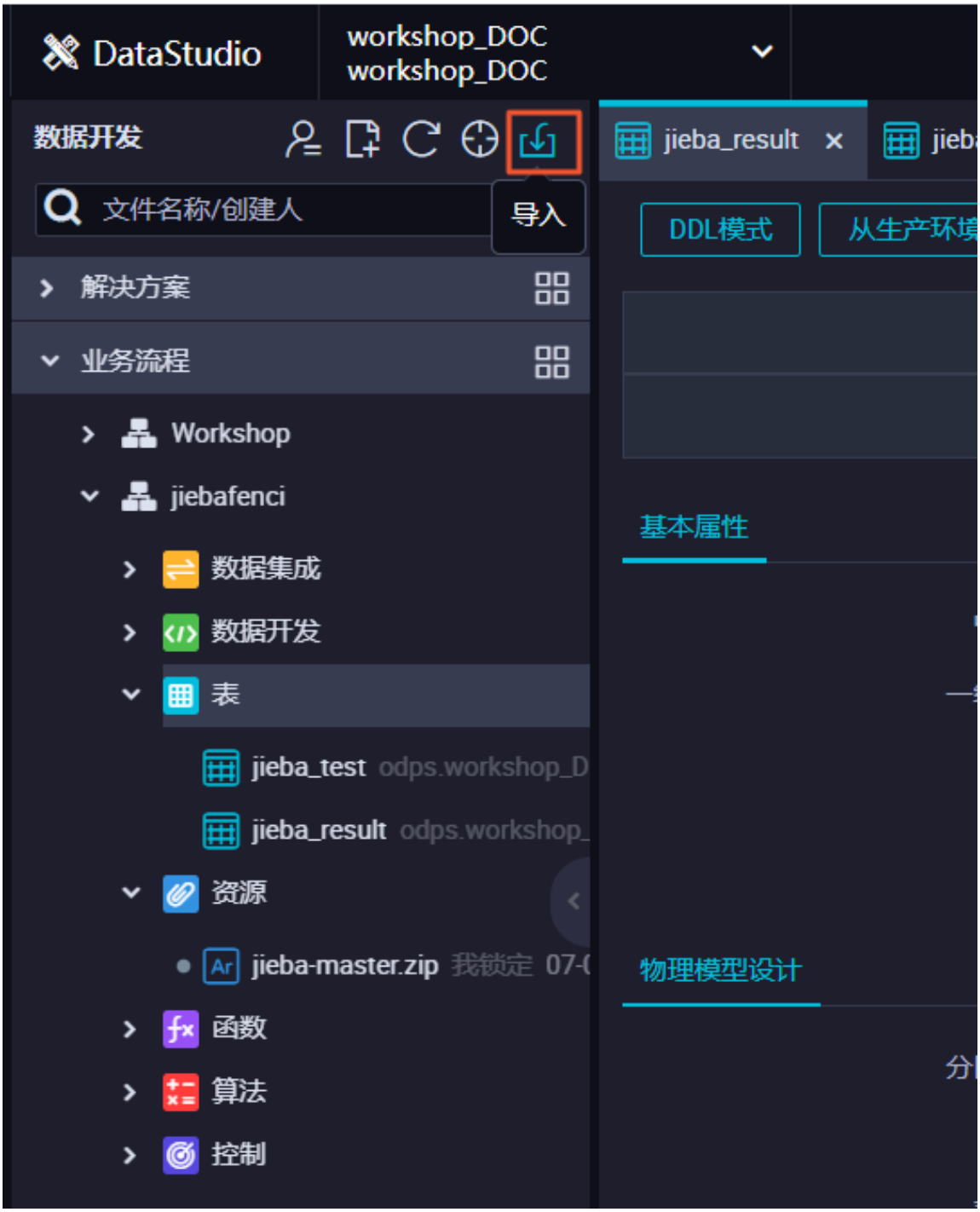
```
CREATE TABLE `jieba_result` (
    `chinese` string
```

```
);
```

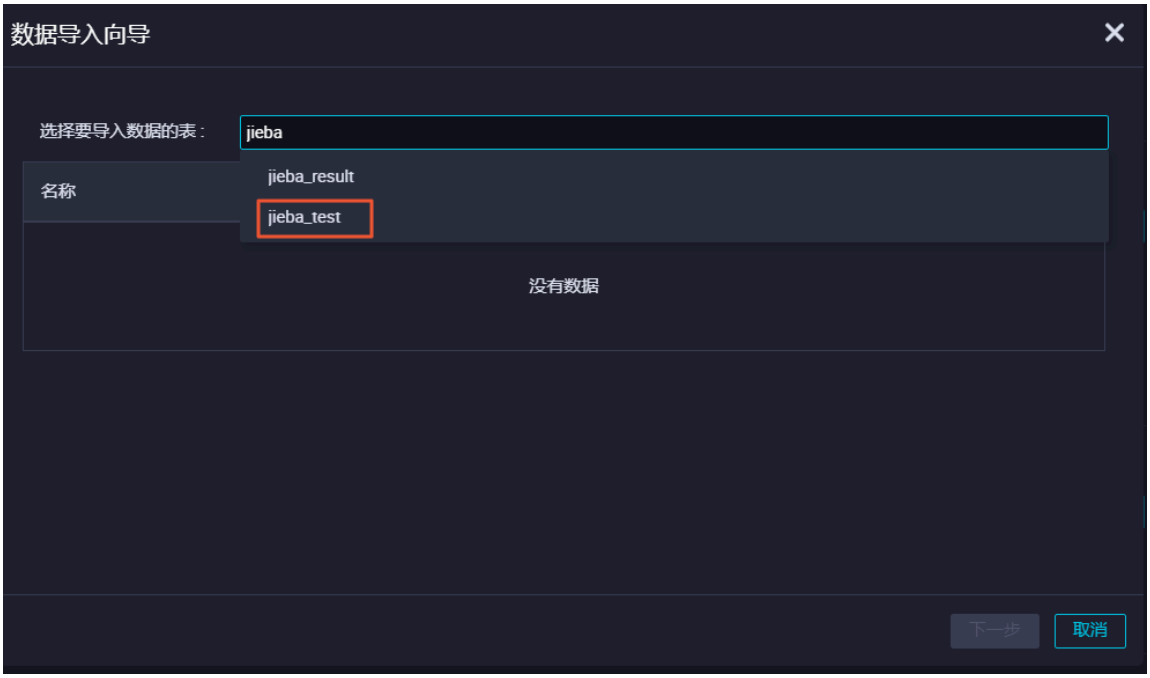
5. 上传测试数据。

   本例已为您准备好分词测试数据，请点击此处下载。

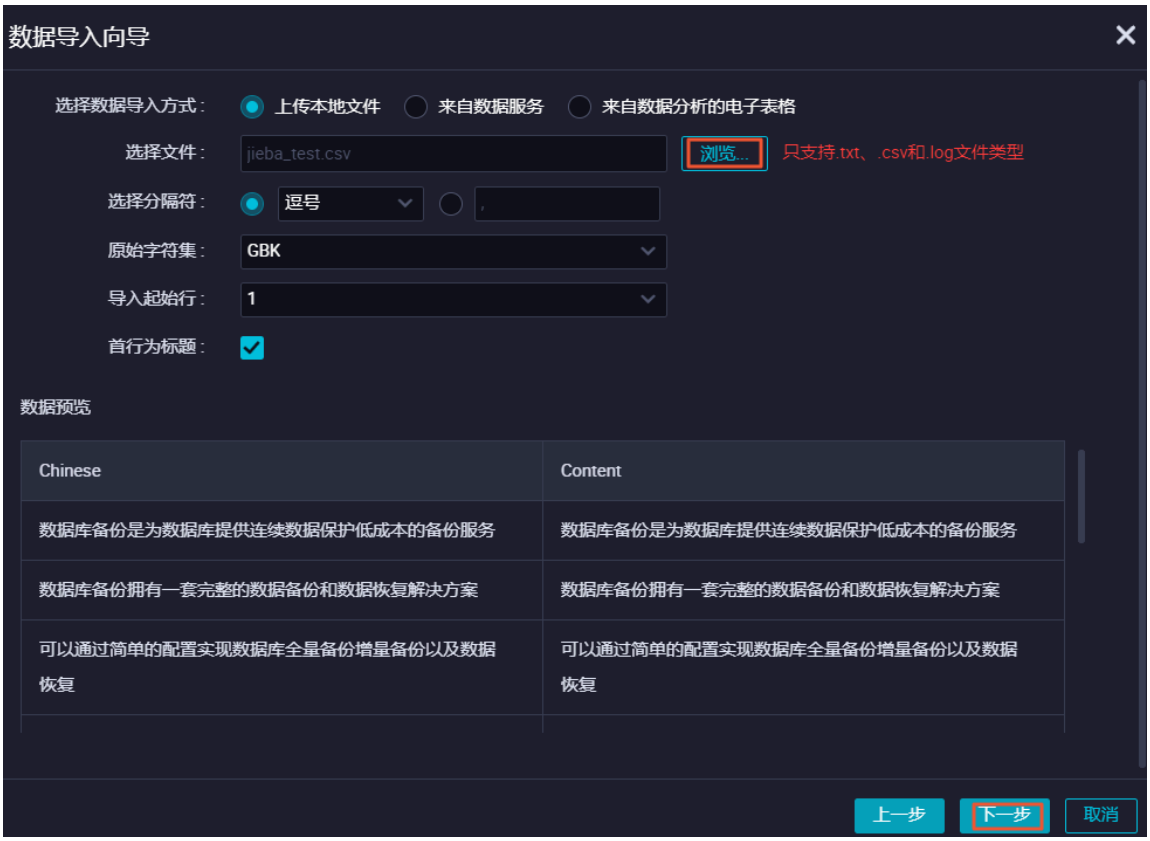
   a) 单击导入。



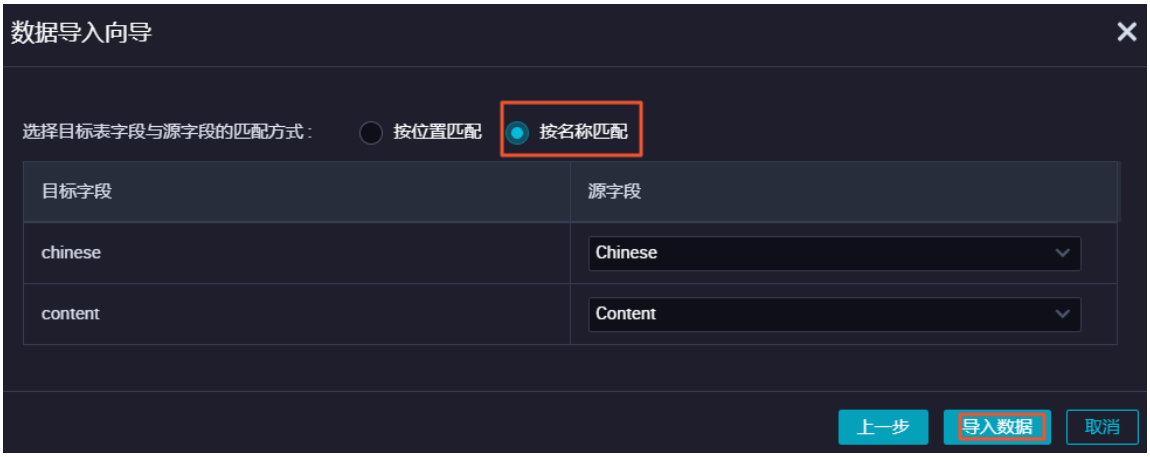
   b) 输入测试数据表名`jieba_test`，单击下一步。



c) 单击浏览，上传您下载到本地的jieba_test.csv文件，单击下一步。

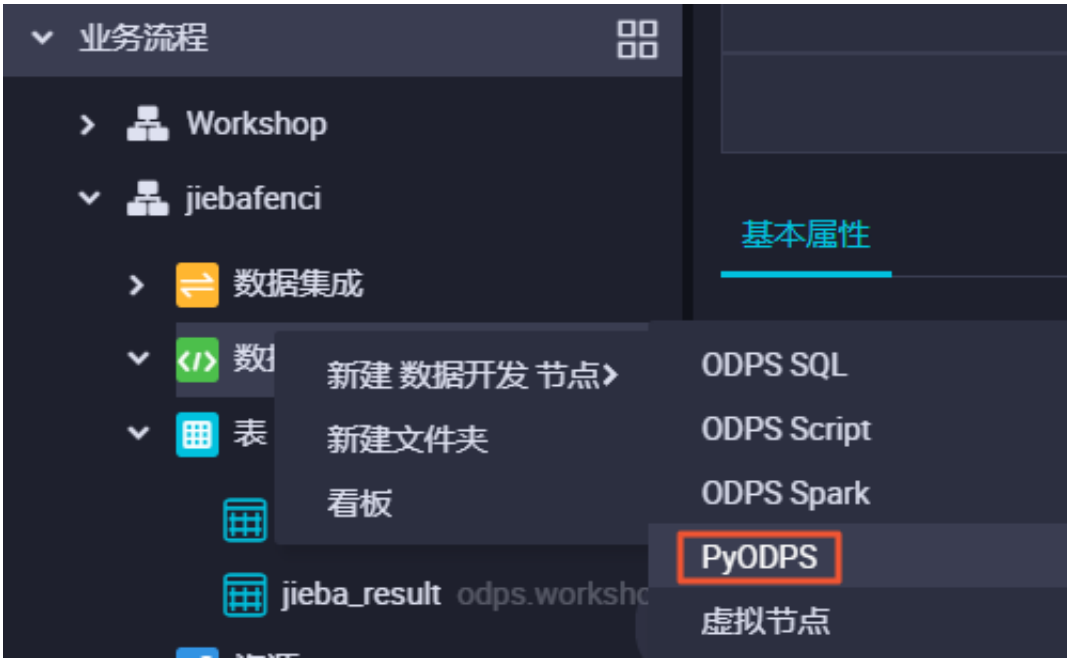


d) 勾选按名称匹配，单击导入数据。



6. 创建PyODPS节点。

a) 在业务流程中右键单击数据开发，选择新建数据开发节点 > PyODPS 。



b) 输入节点名称word_split。

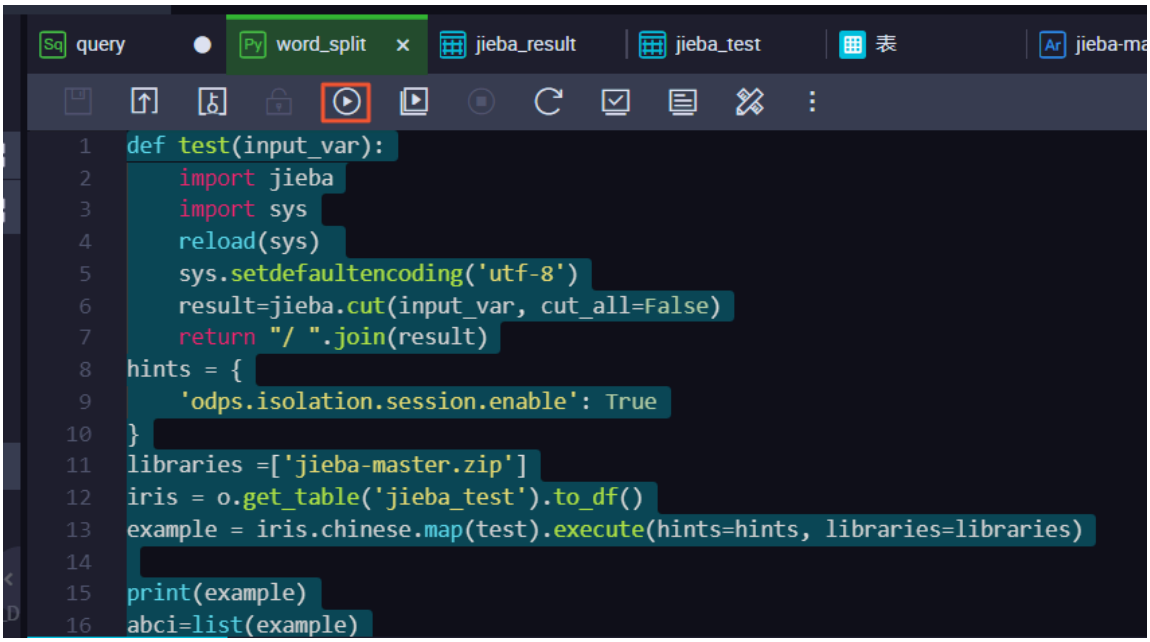
c) 输入您的PyODPS代码。

输入代码如下，释义请见代码注释。

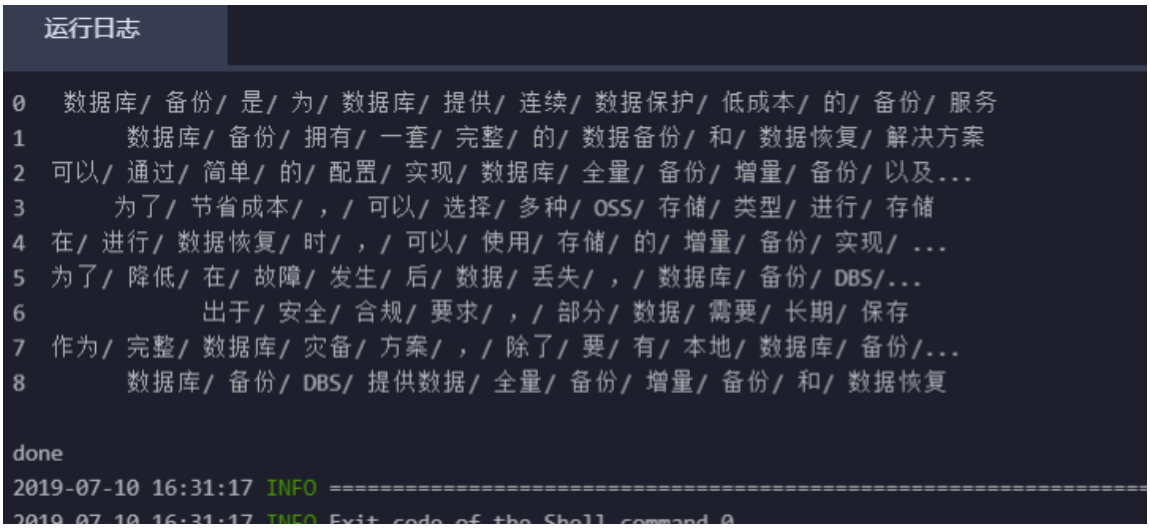
```
def test(input_var):
    import jieba
    import sys
    reload(sys)
    sys.setdefaultencoding('utf-8')
    result=jieba.cut(input_var, cut_all=False)
    return "/ ".join(result)
hints = {
    'odps.isolation.session.enable': True
}
libraries =['jieba-master.zip']  #引用您的jieba-master.zip压缩包。
iris = o.get_table('jieba_test').to_df()  #引用您的jieba_test表中的数据。
```

```
example = iris.chinese.map(test).execute(hints=hints, libraries=
libraries)
print(example)  #查看分词结果，分词结构为MAP类型数据。
abci=list(example)   #将分词结果转为list类型数据。
i = 0
for i in range(i,len(abci)):
    pq=str(abci[i])
    o.write_table('jieba_result',[pq])  #通过循环，将数据逐条写入您的结
果表jieba_result中。
    i+=1
else:
    print("done")
```

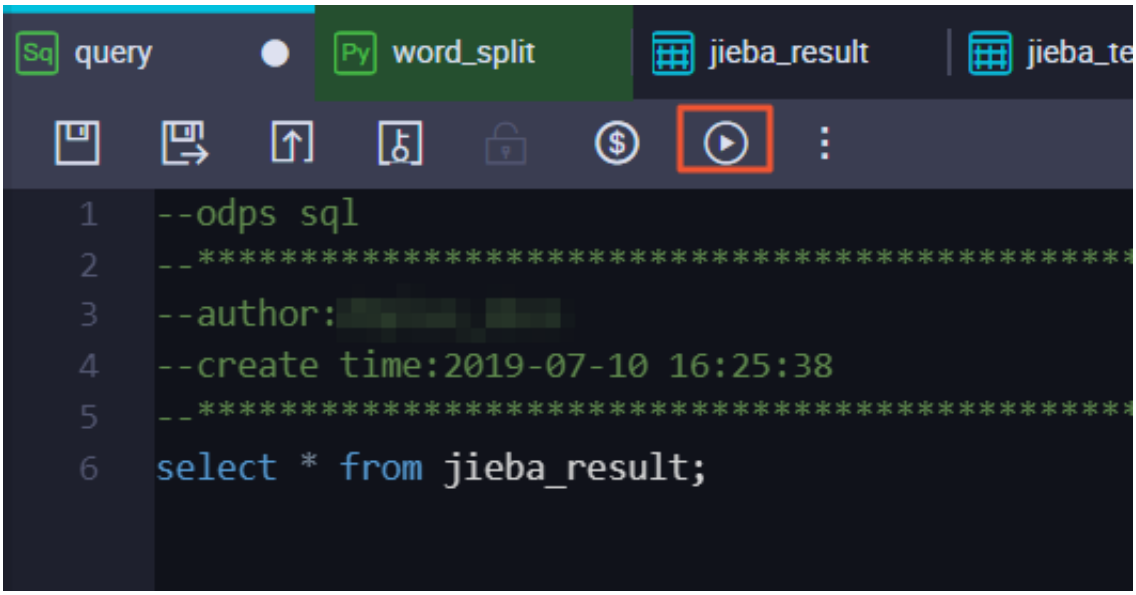
d) 运行代码进行测试。



e) 您可以在运行日志查看结巴分词的程序运行结果。



7. 在数据开发创建一个ODPS SQL类型节点，输入`select * from jieba_result`，单击运行。



8. 查看运行结果，验证数据是否已写入结果表中。

4 计算优化

4.1 SQL优化示例

- **Join语句中Where条件的位置**

  当两个表进行Join操作时，主表的Where限制可以写在最后，但从表分区限制条件不要写在Where条件中，建议写在ON条件或者子查询中。主表的分区限制条件可以写在Where条件中（最好先用子查询过滤）。示例如下：

```
select * from A join (select * from B where dt=20150301)B on B.id=A.
id where A.dt=20150301;
select * from A join B on B.id=A.id where B.dt=20150301;  --不允许
select * from (select * from A where dt=20150301)A join (select *
from B where dt=20150301)B on B.id=A.id;
```

  第二个语句会先Join，后进行分区裁剪，数据量变大，性能下降。在实际使用过程中，应该尽量避免第二种用法。

- 数据倾斜

  产生数据倾斜的根本原因是有少数Worker处理的数据量远远超过其他Worker处理的数据量，从而导致少数Worker的运行时长远远超过其他的平均运行时长，从而导致整个任务运行时间超长，造成任务延迟。

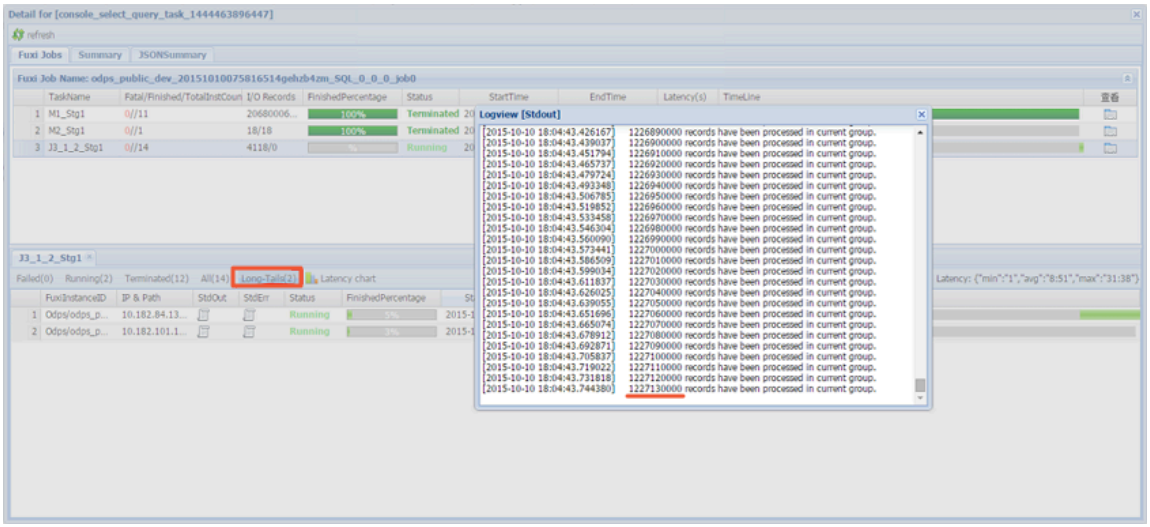
  更多数据倾斜优化的详情请参见#unique_98。

  - Join造成的数据倾斜

    造成Join数据倾斜的原因是Join on的key分布不均匀。假设还是上述示例语句，现在将大表A和小表B进行Join操作，运行如下语句。

    ```
    select * from A join B on A.value= B.value;
    ```

    此时，复制logview的链接并打开webcosole页面，双击执行Join操作的fuxi job，可以看到此时在[Long-tails]区域有长尾，表示数据已经倾斜了。

    此时您可通过如下方法进行优化。

    ■ 由于表B是个小表并且没有超过512MB，您可将上述语句优化为mapjoin语句再执行，语句如下。

    ```
    select /*+ MAPJOIN(B) */ * from A join B on A.value= B.value;
    ```

    ■ 您也可将倾斜的key用单独的逻辑来处理，例如经常发生两边的key中有大量null数据导致了倾斜。则需要在Join前先过滤掉null的数据或者补上随机数，然后再进行Join，示例如下。

    ```
    select * from A join B
    ```

```
on case when A.value is null then concat('value',rand() ) else
A.value end = B.value;
```

在实际场景中，如果您知道数据倾斜了，但无法获取导致数据倾斜的key信息，那么可以使用一个通用的方案，查看数据倾斜，如下所示。

```
例如: select * from a join b on a.key=b.key; 产生数据倾斜。
您可以执行:
   ```sql
select left.key, left.cnt * right.cnt from
(select key, count(*) as cnt from a group by key) left
join
(select key, count(*) as cnt from b group by key) right
on left.key=right.key;
```

查看key的分布，可以判断a join b时是否会有数据倾斜。

- group by倾斜

  造成group by倾斜的原因是group by的key分布不均匀。

  假设表A内有两个字段（key，value），表内的数据量足够大，并且key的值分布不均，运行语句如下所示：

  ```
  select key,count(value) from A group by key;
  ```

  当表中的数据足够大时，您会在webcosole页面看见长尾。若想解决这个问题，您需要在执行SQL前设置防倾斜的参数，设置语句为`set odps.sql.groupby.skewindata=true`。

- 错误使用动态分区造成的数据倾斜

  动态分区的SQL，在MaxCompute中会默认增加一个Reduce，用来将相同分区的数据合并在一起。这样做的好处，如下所示。

  - 可减少MaxCompute系统产生的小文件，使后续处理更快速。
  - 可避免一个Worker输出文件很多时占用内存过大。

  但也正是因为这个Reduce的引入，导致分区数据如果有倾斜的话，会发生长尾。因为相同的数据最多只会有10个Worker处理，所以数据量大，则会发生长尾，示例如下。insert overwrite table A2 partition(dt) select split_part(value,'\t',1) as field1, split_part(value,'\t',2) as field2, dt from A where dt='20151010';

  这种情况下，没有必要使用动态分区，所以可以改为如下语句：

  ```
  insert overwrite table A2 partition(dt='20151010')
  select
  split_part(value,'\t',1) as field1,
  split_part(value,'\t',2) as field2
  from A
  ```
  ```

```
where dt='20151010';
```

- 窗口函数的优化

  如果您的SQL语句中用到了窗口函数，一般情况下每个窗口函数会形成一个Reduce作业。如果窗口函数略多，那么就会消耗资源。在某些特定场景下，窗口函数是可以进行优化的。

  - 窗口函数over后面要完全相同，相同的分组和排序条件。
  - 多个窗口函数在同一层SQL执行。

  符合上述两个条件的窗口函数会合并为一个Reduce执行。SQL示例如下所示。

```
select
rank()over(partition by A order by B desc) as rank,
row_number()over(partition by A order by B desc) as row_num
from MyTable;
```

- 子查询改Join

  例如有一个子查询，如下所示。

```
SELECT * FROM table_a a WHERE a.col1 IN (SELECT col1 FROM table_b b
WHERE xxx);
```

  当此语句中的table_b子查询返回的col1的个数超过1000个时，系统会报错为`records returned from subquery exceeded limit of 1000`。此时您可以使用Join语句来代替，如下所示。

```
SELECT a.* FROM table_a a JOIN (SELECT DISTINCT col1 FROM table_b b
WHERE xxx) c ON (a.col1 = c.col1)
```

   说明:

  - 如果没用Distinct，而子查询c返回的结果中有相同的col1的值，可能会导致a表的结果数变多。
  - 因为Distinct子句会导致查询全落到一个Worker里，如果子查询数据量比较大的话，可能会导致查询比较慢。
  - 如果已经从业务上控制了子查询里的col1不可能会重复，比如查的是主键字段，为了提高性能，可以把Distinct去掉。

4.2 计算长尾调优

长尾问题是分布式计算中最常见的问题之一，也是典型的疑难杂症。究其原因，是因为数据分布不均，导致各个节点的工作量不同，整个任务需要等最慢的节点完成才能结束。

处理这类问题的思路就是把工作分给多个 Worker 去执行，而不是一个 Worker 单独运行最重的那份工作。本文将为您介绍平时工作中遇到的一些典型的长尾问题的场景及其解决方案。

Join 长尾

问题原因：

Join 出现长尾，是因为 Join 时出现某个 Key 里的数据特别多的情况。

解决办法：

您可以从以下四方面进行考虑：

- 排除两张表都是小表的情况，若两张表里有一张大表和一张小表，可以考虑使用 mapjoin，对小表进行缓存，具体的语法和说明请参见#unique_56。如果是 MapReduce 作业，可以使用资源表的功能，对小表进行缓存。
- 但是如果两张表都比较大，就需要先尽量去重。
- 若还是不能解决，就需要从业务上考虑，为什么会有这样的两个大数据量的 Key 要做笛卡尔积，直接考虑从业务上进行优化。
- 小表leftjoin大表，odps直接leftjoin较慢。此时可以先小表和大表mapjoin，这样能拿到小表和大表的交集中间表，且这个中间表一定是不大于大表的（只要不是有很大的key倾斜就不会膨胀的很大）。然后小表再和这个中间表进行leftjoin，这样效果等于小表leftjoin大表。

Group By 长尾

问题原因：

Group By Key 出现长尾，是因为某个 Key 内的计算量特别大。

解决办法：

您可以通过以下两种方法解决：

- 可对 SQL 进行改写，添加随机数，把长 Key 进行拆分。如下所示：

```
SELECT Key,COUNT(*) AS Cnt FROM TableName GROUP BY Key;
```

不考虑 Combiner，M 节点会 Shuffle 到 R 上，然后 R 再做 Count 操作。对应的执行计划是 M > R。但是如果对长尾的 Key 再做一次工作再分配，就变成如下语句：

```
-- 假设长尾的Key已经找到是KEY001
SELECT a.Key
  , SUM(a.Cnt) AS Cnt
FROM (
  SELECT Key
    , COUNT(*) AS Cnt
FROM TableName
GROUP BY Key,
  CASE
```

```
    WHEN Key = 'KEY001' THEN Hash(Random()) % 50
    ELSE 0
   END
) a
GROUP BY a.Key;
```

由上可见，这次的执行计划变成了 M > R > R。虽然执行的步骤变长了，但是长尾的 Key 经过 2 个步骤的处理，整体的时间消耗可能反而有所减少。

说明:

若数据的长尾并不严重，用这种方法人为地增加一次 R 的过程，最终的时间消耗可能反而更大。

- 使用通用的优化策略 — 系统参数，设置如下：

```
set odps.sql.groupby.skewindata=true。
```

但是通用性的优化策略无法针对具体的业务进行分析，得出的结果不总是最优的。您可以根据实际的数据情况，用更加高效的方法来改写 SQL。

Distinct 长尾

对于 Distinct，上述 Group By 长尾时，把长 Key 进行拆分的策略已经不生效了。对这种场景，您可以考虑通过其他方式解决。

解决办法：

```
--原始SQL,不考虑Uid为空
SELECT COUNT(uid) AS Pv
    , COUNT(DISTINCT uid) AS Uv
FROM UserLog;
```

可以改写成如下语句：

```
SELECT SUM(PV) AS Pv
    , COUNT(*) AS UV
FROM (
    SELECT COUNT(*) AS Pv
      , uid
    FROM UserLog
    GROUP BY uid
) a;
```

该解法是把 Distinct 改成了普通的 Count，这样的计算压力不会落到同一个 Reducer 上。而且这样改写后，既能支持前面提到的 Group By 优化，系统又能做 Combiner，性能会有较大的提升。

动态分区长尾

问题原因：

- 动态分区功能为了整理小文件，会在最后启一个 Reduce，对数据进行整理，所以如果使用动态分区写入数据时若有倾斜，就会发生长尾。
- 一般情况下，滥用动态分区的功能也是产生这类长尾的一个常见原因。

解决办法：

若写入的数据已经确定需要把数据写入某个具体分区，那可以在 Insert 的时候指定需要写入的分区，而不是使用动态分区。

通过 Combiner 解决长尾

对于 MapRedcuce 作业，使用 Combiner 是一种常见的长尾优化策略。在 WordCount 的示例中，已提到这种做法。通过 Combiner，减少 Mapper Shuffle 往 Reducer 的数据，可以大大减少网络传输的开销。对于 MaxCompute SQL，这种优化会由系统自动完成。

 说明：

Combiner 只是 Map 端的优化，需要保证是否执行 Combiner 的结果是一样的。以 WordCount 为例，传 2 个 (KEY,1) 和传 1 个 (KEY,2) 的结果是一样的。但是比如在做平均值时，便不能直接在 Combiner 里把 (KEY,1) 和 (KEY,2) 合并成 (KEY,1.5)。

通过系统优化解决长尾

针对长尾这种场景，除了前面提到的 Local Combiner，MaxCompute 系统本身还做了一些优化。比如在跑任务的时候，日志里突然打出如下的内容（+N backups 部分）：

```
M1_Stg1_job0:0/521/521[100%] M2_Stg1_job0:0/1/1[100%] J9_1_2_Stg5_job0
:0/523/523[100%] J3_1_2_Stg1_job0:0/523/523[100%] R6_3_9_Stg2_job0:1/
1046/1047[100%]
M1_Stg1_job0:0/521/521[100%] M2_Stg1_job0:0/1/1[100%] J9_1_2_Stg5_job0
:0/523/523[100%] J3_1_2_Stg1_job0:0/523/523[100%] R6_3_9_Stg2_job0:1/
1046/1047[100%]
M1_Stg1_job0:0/521/521[100%] M2_Stg1_job0:0/1/1[100%] J9_1_2_Stg5_job0
:0/523/523[100%] J3_1_2_Stg1_job0:0/523/523[100%] R6_3_9_Stg2_job0:1/
1046/1047(+1 backups)[100%]
M1_Stg1_job0:0/521/521[100%] M2_Stg1_job0:0/1/1[100%] J9_1_2_Stg5_job0
:0/523/523[100%] J3_1_2_Stg1_job0:0/523/523[100%] R6_3_9_Stg2_job0:1/
1046/1047(+1 backups)[100%]
```

可以看到 1047 个 Reducer，有 1046 个已经完成了，但是最后一个一直没完成。系统识别出这种情况后，自动启动了一个新的 Reducer，跑一样的数据，然后看两个哪个快，取快的数据归并到最后的结果集里。

通过业务优化解决长尾

虽然前面的优化策略有很多，但仍然不能解决所有问题。有时碰到的长尾问题，还需要从业务角度上去考虑是否有更好的解决方法，示例如下：

- 实际数据可能包含非常多的噪音。比如：需要根据访问者的 ID 进行计算，看每个用户的访问记录的行为。需要先去掉爬虫的数据（现在的爬虫已越来越难识别），否则爬虫数据很容易长尾计算的长尾。类似的情况还有根据 xxid 进行关联的时候，需要考虑这个关联字段是否存在为空的情况。
- 一些业务特殊情况。比如：ISV 的操作记录，在数据量、行为方式上都会和普通的个人会有很大的区别。那么可以考虑针对大客户，使用特殊的分析方式进行单独处理。
- 数据分布不均匀的情况下，不要使用常量字段做 Distribute by 字段来实现全排序。

4.3 长周期指标的计算优化方案

实验背景

电子商务公司（如淘宝）对用户数据分析的角度和思路可谓是应有尽有、层出不穷，所以在电商数据仓库和商业分析场景中，经常需要计算最近 N 天的访客数、购买用户数、老客数等类似的指标。

这些指标有一个共同点：都需要根据用户在电商平台上（或网上店铺）一段时间积累的数据进行计算（这里讨论的前提是数据都存储在 MaxCompute 上）。

一般情况下，这些指标的计算方式就是从日志明细表中计算就行了，如下代码计算商品最近 30 天的访客数：

```
select item_id --商品id
  ,count(distinct visitor_id) as ipv_uv_1d_001
from 用户访问商品日志明细表
where ds <= ${bdp.system.bizdate}
and ds >=to_char(dateadd(to_date(${bdp.system.bizdate},'yyyymmdd'),-29
,'dd'),'yyyymmdd')
 group by item_id;
```

说明：

代码中的变量都是DataWorks的调度变量，仅适用于DataWorks的调度任务。为了方便后文不再提醒。

当每天的日志量很大时，上面代码存在一个严重的问题，需要的 Map Instance 个数太多，甚至会超过 99999 个 Instance 个数的限制，Map Task 就没有办法顺利执行，更别说后续的操作了。

为什么 Instance 个数需要那么多呢？是因为每天的日志数据很大，30 天的数据量更是惊人。此时 Select 操作需要大量的 Map Instance，结果超过了 Instance 的上限，导致代码无法运行。

实验目的

如何计算长周期的指标，又不影响性能呢？通常有以下两种思路：

- 多天汇总的问题根源是数据量的问题，如果把数据量给降低了，便可解决此问题。

- 减少数据量最直接的办法是把每天的数据量都给减少，因此需要构建临时表，对 1d 的数据进行轻度汇总，这样便可去掉很多重复数据，减少数据量。

实验方案

操作步骤

1. 构建中间表，每天汇总一次。

   比如对于上面的例子，可以构建一个 item_id+visitor_id 粒度的中间表。即构建item_id+visitior_id 粒度的日汇总表，记作 A。如下所示：

```
insert overwrite table mds_itm_vsr_xx(ds='${bdp.system.bizdate} ')
select item_id,visitor_id,count(1) as pv
  from
   (
  select item_id,visitor_id
  from 用户访问商品日志明细表
  where ds =${bdp.system.bizdate}
  group by item_id,visitor_id
  ) a;
```

2. 计算多天的数据，依赖中间表进行汇总。

   对 A 进行 30 天的汇总，如下所示：

```
select item_id
        ,count(distinct visitor_id) as uv
        ,sum(pv) as pv
  from mds_itm_vsr_xx
  where ds <= '${bdp.system.bizdate} '
  and ds >= to_char(dateadd(to_date('${bdp.system.bizdate} ','
yyyymmdd'),-29,'dd'),'yyyymmdd')
  group by item_id;
```

影响及思考

上面讲述的方法，对每天的访问日志明细数据进行单天去重，从而减少了数据量，提高了性能。缺点是每次计算多天的数据的时候，都需要读取 N 个分区的数据。

那么是否有一种方式，不需要读取 N 个分区的数据，而是把 N 个分区的数据压缩合并成一个分区的数据，让一个分区的数据包含历史数据的信息呢?

业务上是有类似场景的，可以通过 增量累计方式计算长周期指标。

场景示例

求最近 1 天店铺商品的老买家数。老买家数的算法定义为：过去一段时间有购买的买家（比如过去 30 天）。

一般情况下，老买家数计算方式如下所示：

```
select item_id --商品id
        ,buyer_id as old_buyer_id
from 用户购买商品明细表
where ds < ${bdp.system.bizdate}
and ds >=to_char(dateadd(to_date(${bdp.system.bizdate},'yyyymmdd'),-29
,'dd'),'yyyymmdd')
group by item_id
        ,buyer_id;
```

改进思路：

- 维护一张店铺商品和买家购买关系的维表记作表 A，记录买家和店铺的购买关系，以及第一次购买时间，最近一次购买时间，累计购买件数，累计购买金额等信息。
- 每天使用最近 1 天的支付明细日志更新表 A 的相关数据。
- 计算老买家时，最需要判断最近一次购买时间是否是 30天 之内就行了，从而做到最大程度上的数据关系对去重，减少了计算输入数据量。

4.4 MaxCompute账单分析最佳实践

本文问您介绍如何分析MaxCompute账单。

背景信息

阿里云大数据计算服务MaxCompute是一款商业化的大数据分析平台，其计算资源的计费方式分为预付费和后付费两种，产品每天会以Project为维度进行计量计费（账单会在第二天上午6点前生成）。

关于MaxCompute计量计费说明，详情请参见计量计费说明文档。

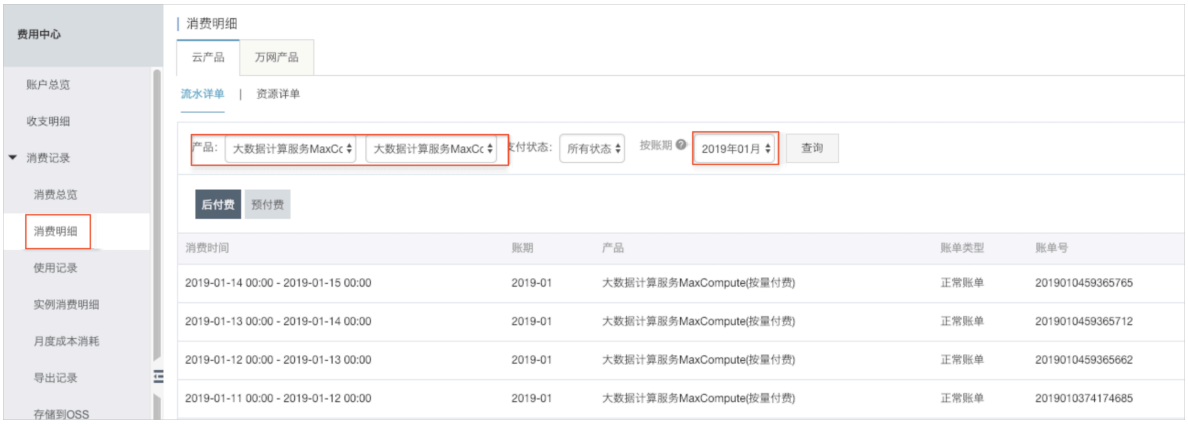
我们会在数据开发阶段或者在产品上线前夕发布账单波动（通常情况下为消费增长）信息。用户可以通过自助的方式来分析账单波动情况，再对自己的作业进行优化。您可以在阿里云费用中心下载阿里云所有商业化收费的产品的费用明细。

获取账单信息

通常您需要使用主账号查看账单详情。如果您需要使用子账号查看账单信息，请首先参考费用中心RAM配置策略进行子账号授权。

1. **使用主账号或者被授权的RAM子账号登录阿里云管控台。**
2. **单击进入费用中心，如下图所示。**

3. 在费用中心>消费记录>消费明细中，选择产品和账单日期，如下图所示。

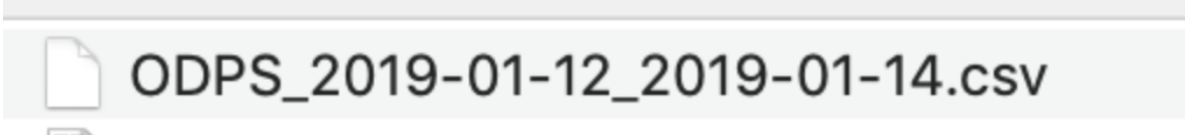


说明:

包年包月中的后付费是指项目开通包年包月计算计费模式后，产生的存储、下载对应的费用（存储、下载费用只有后付费）。

4. 在使用记录中下载csv文件，如下图所示。

5. 在本地打开下载的csv文件，进行批量数据分析。




```
--csv的表头
"项目编号","计量信息编号","数据分类","存储（Byte）","结束时间","SQL/交互式分析 读取量（Byte）","SQL复杂度","公网上行流量（Byte）","公网下行流量（Byte）","MR
```

```
/Spark作业计算（Core*Second）","SQL读取量_访问OTS（Byte）","SQL读取量_访问OSS
（Byte）","开始时间","计算资源规格","DataWorks调度任务ID"
```

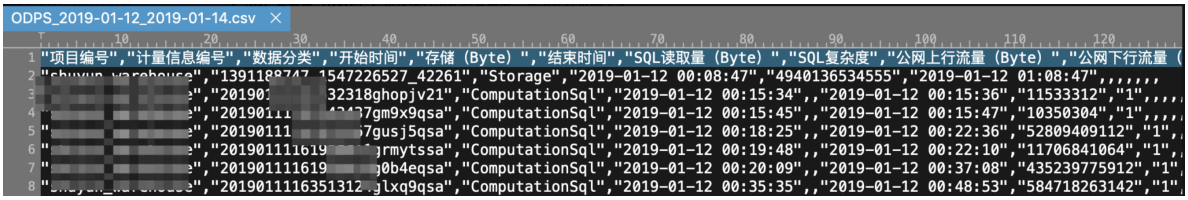
上传账单明细至MaxCompute

使用记录明细字段解释：

- 项目编号：当前账号或子账号对应的主账号的MaxCompute Project列表。
- 计量信息编号：其中会以存储、计算、上传和下载的任务ID为计费信息编号，SQL为instanceid，上传和下载为Tunnel sessionid。
- 数据分类：Storage（存储）、ComputationSql（计算）、UploadIn（内网上传）、UploadEx（外网上传）、DownloadIn（内网下载）、DownloadEx（外网下载）。按照计费规则其中只有红色为实际计费项目。
- 存储（Byte）：每小时读取的存储量单位为Byte。
- 开始时间/结束时间：按照实际作业执行时间进行计量，只有Storage是按照每个小时取一次数据。
- SQL读取量（Byte）：SQL计算项，每一次SQL执行时SQL的input数据量，单位为Byte。
- SQL复杂度（Byte）：每次执行SQL的复杂度，为SQL计费因子之一。
- 公网上行流量（Byte）、公网下行流量（Byte）：分别为公网上传和下载的数据量，单位Byte。
- MR/Spark作业计算（CoreSecond）：MR/Spark作业的计算时单位为coresecond，需要转换为计算时Hour。
- SQL读取量_访问OTS（Byte）、SQL读取量_访问OSS（Byte）：外部表实施收费后的读取数据量，单位Byte。
- 计算资源规格（后付费）：对应计量信息所在项目所属的计算资源规格，若值为null，则表示后付费标准版；若值为OdpsDev，则表示后付费开发者版。
- 计算资源规格（预付费）：对应计量信息所在项目所属的计算资源规格。若值为null，则表示预付费标准版；若值为‘OdpsPlus160CU150TB'、'OdpsPlus320CU300TB'、'OdpsPlus600CU500TB'，则分别表示存储密集型160套餐、存储密集型320套餐，存储密集型600套餐。
- DataWorks调度任务ID：计量中对应作业在DataWorks上的调度节点ID。若值为null，则表示非DataWorks调度节点提交的job；若值为一串数字ID，则表示job对应DataWorks调度节点ID。您可以在DataWorks对应的项目中使用该ID搜索到具体任务。

详细的实施步骤为：

1. 确认CSV文件数据，尤其是列分隔符等。



> **说明：**
> 数据以逗号分隔，且单元格值都带有双引号。

2. 数据预处理：替换掉文档所有双引号，以方便使用Tunnel等工具，如下图所示。

> **说明：**
> 替换为中的内容可以为空。直接点击全部替换。

3. 创建MaxCompute表，存储下载的消费明细。

```
DROP TABLE IF EXISTS maxcomputefee ;
```

```
CREATE TABLE IF NOT EXISTS maxcomputefee
(
    projectid STRING COMMENT '项目编号'
    ,feeid STRING COMMENT '计费信息编号'
    ,type STRING COMMENT '数据分类，包括Storage、ComputationSQL、
DownloadEx等'
    ,storage BIGINT COMMENT '存储（Byte）'
    ,endtime DATETIME COMMENT '结束时间'
    ,computationsqlinput BIGINT COMMENT 'SQL/交互式分析 读取量（Byte）'
    ,computationsqlcomplexity DOUBLE COMMENT 'sql复杂度'
    ,uploadex BIGINT COMMENT '公网上行流量Byte'
    ,download BIGINT COMMENT '公网下行流量Byte'
    ,cu_usage DOUBLE COMMENT 'MR计算时*second'
    ,input_ots BIGINT COMMENT '访问OTS的数据输入量'
    ,input_oss BIGINT COMMENT '访问OSS的数据输入量'
    ,starttime DATETIME COMMENT '开始时间'
    ,source_type String COMMENT '计算资源'
    ,source_id String COMMENT 'DataWorks调度任务ID'
);
```

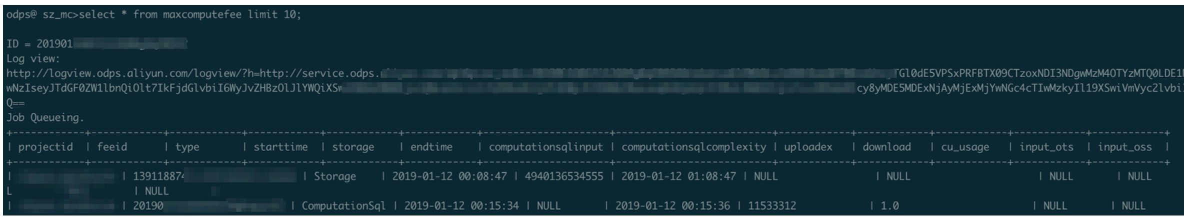
4. Tunnel上传数据，Tunnel的配置详情请参见#unique_83。

```
odps@ sz_mc>tunnel upload /Users/yangyi/Desktop/ODPS_2019-01-12_2019
-01-14.csv maxcomputefee -c "UTF-8" -h "true" -dfp "yyyy-MM-dd HH:mm
:ss";
```

说明:

用户也可以通过DataWorks数据导入的功能来进行，具体请参见操作步骤。

5. 验证数据。



通过SQL分析账单数据

1. 分析SQL费用

说明:

云上用户使用MaxCompute，95%的用户通过SQL即可满足需求，SQL也在消费增长中占了很大比例。 SQL费用=一次SQL计算费用 = 计算输入数据量*SQL复杂度*单价（0.3元/GB）

```
--分析SQL消费，按照sqlmoney行排行。
SELECT  to_char(endtime,'yyyymmdd') as ds,feeid as instanceid
        ,projectid
        ,computationsqlcomplexity  --复杂度
        ,SUM((computationsqlinput / 1024 / 1024 / 1024)) as
computationsqlinput  --数据输入量（GB）
        ,SUM((computationsqlinput / 1024 / 1024 / 1024)) *
computationsqlcomplexity * 0.3 AS sqlmoney
FROM    maxcomputefee
WHERE   TYPE = 'ComputationSql'
AND to_char(endtime,'yyyymmdd') >= '20190112'
GROUP BY to_char(endtime,'yyyymmdd'),feeid
         ,projectid
         ,computationsqlcomplexity
ORDER BY sqlmoney DESC
LIMIT   10000
```

```
;
```

查询结果

	A	B	C	D	E	F
1	ds	instanceid	projectid	computationsqlcor	computationsqlinp	sqlmoney
2	20190114	20190113220731203g		1.5	939.2598592275754	422.6669366524089
3	20190113	20190112211606543g		1.5	939.0618489095941	422.5778320093173
4	20190114	20190113214053411g		1.5	797.8894629115239	359.0502583101857
5	20190113	20190112204652265g		1.5	797.6614840412512	358.94766781856305
6	20190114	20190113212053799g		1.5	700.250122227706	315.1125550024677
7	20190113	20190112202519334g		1.5	700.0137415388599	315.00618369248696
8	20190114	20190113222308193g		1.0	750.0447742938995	225.01343228816987
9	20190113	20190112212657773g		1.0	749.9437991976738	224.98313975930213
10	20190114	20190113173351510g		1.0	546.8404815010726	164.05214445032178

根据此段SQL执行结果可得到以下结论：

- 大作业可以优化的点：是否可以减小数据读取量、降低复杂度来优化费用成本。
- 也可以按照ds字段（按照天）进行汇总，分析某个时间段内的SQL消费金额走势。例如利用本地excel或云上QuickBI等工具绘制折线图等方式，更直观的反应作业的趋势。
- 根据执行结果定位到需要优化的点：

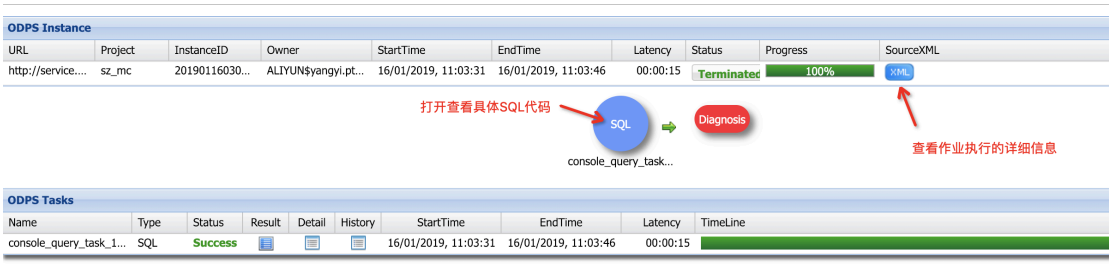
  a. 获取具体的instanceid

  在console或者DataWorks脚本执行`wait instanceid;`命令查看具体作业和SQL。

  b. 在浏览器中打开logview的url地址

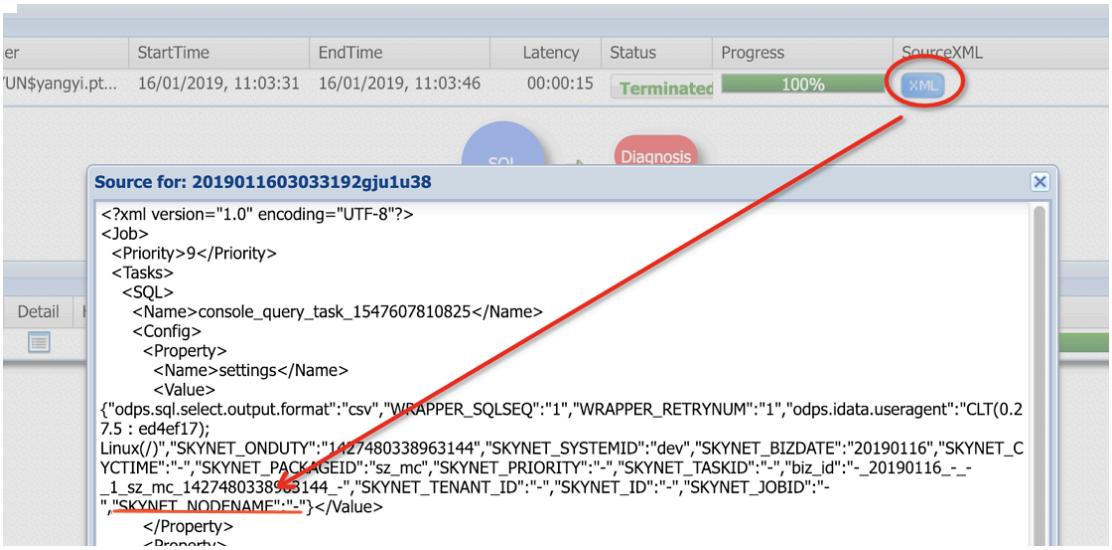
  关于logview的介绍请参见#unique_103。

  c. 从logview中获取DataWorks节点名称

  在logview中打开SourceXML可以查看到具体执行信息，例如SKYNET_NODENAME表示DataWorks的节点名称（只有被调度系统执行的作业才有

值，临时查询为空，如下图所示）。获得节点名称后，可以快速地在DataWorks上找到该节点并进行优化或查看责任人，如下图所示。



2. 分析作业增长趋势

> **说明:**
> 通常，费用的增长背后其实是作业量的暴涨，可能是重复执行，也可能是调度属性配置的不合理。

```
--分析作业增长趋势。
SELECT   TO_CHAR(endtime,'yyyymmdd') AS ds
         ,projectid
         ,COUNT(*) AS tasknum
FROM     maxcomputefee
WHERE    TYPE = 'ComputationSql'
AND      TO_CHAR(endtime,'yyyymmdd') >= '20190112'
GROUP BY TO_CHAR(endtime,'yyyymmdd')
          ,projectid
ORDER BY tasknum DESC
LIMIT    10000
;
```

执行结果

	A	B	C
1	ds	projectid	tasknum
2	20190112		311
3	20190113		304
4	20190114		282

从执行结果可以看出12~14日提交到MaxCompute且执行成功的作业数的波动趋势。

3. 分析存储费用

> **说明:**
> 存储费用的计费规则相对复杂。下载到的明细是按每个小时取一次得出的数据。按照MaxCompute存储计费规则，会先整体24小时求和，再将平均之后的值进行阶梯收费。详情请参见#unique_104。

```
--分析存储费用。
SELECT  t.ds
        ,t.projectid
        ,t.storage
        ,CASE    WHEN t.storage < 0.5 THEN 0.01
                 WHEN t.storage >= 0.5 AND t.storage <= 100 THEN t.
storage*0.0192
                 WHEN t.storage > 100 AND t.storage <= 1024 THEN (
100*0.0192+(t.storage-100)*0.0096)
                 WHEN t.storage > 1024 AND t.storage <= 10240 THEN (
100*0.0192+(1024-100)*0.0096+(t.storage-1024)*0.0084)
                 WHEN t.storage > 10240 AND t.storage <= 102400 THEN
 (100*0.0192+(1024-100)*0.0096+(10240-1024)*0.0084+(t.storage-10240
)*0.0072)
                 WHEN t.storage > 102400 AND t.storage <= 1048576
THEN (100*0.0192+(1024-100)*0.0096+(10240-1024)*0.0084+(102400-10240
)*0.0072+(t.storage-102400)*0.006)
         END storage_fee
FROM    (
            SELECT  to_char(starttime,'yyyymmdd') as ds
                    ,projectid
                    ,SUM(storage/1024/1024/1024)/24 AS storage
            FROM    maxcomputefee
            WHERE   TYPE = 'Storage'
            and to_char(starttime,'yyyymmdd') >= '20190112'
            GROUP BY to_char(starttime,'yyyymmdd')
                     ,projectid
        ) t
ORDER BY storage_fee DESC
;
```

执行结果

	A	B	C
1	ds	projectid	tasknum
2	20190112		311
3	20190113		304
4	20190114		282

根据计算结果可以分析得出结论：

- 存储在12日为最高有一个增长的过程，但在14日出现降低。
- 存储优化，建议为表设置生命周期，删除长期不使用的临时表等。

4. 分析下载费用

对于公网或者跨Region的数据下载，MaxCompute将按照下载的数据大小进行计费。计费公式为：

```
一次下载费用=下载数据量*单价（0.8元/GB）
```

```
--分析下载消费明细。
SELECT  TO_CHAR(starttime,'yyyymmdd') AS ds
        ,projectid
        ,SUM((download/1024/1024/1024)*0.8) AS download_fee
FROM    maxcomputefee
WHERE   type = 'DownloadEx'
AND     TO_CHAR(starttime,'yyyymmdd') >= '20190112'
GROUP BY TO_CHAR(starttime,'yyyymmdd')
         ,projectid
ORDER BY download_fee DESC
;
```

按照执行结果也可以分析出某个时间段内的下载费用走势。另外可以通过tunnel show history查看具体历史信息，具体命令详见#unique_83。

```
odps@ sz_mc>tunnel show history;
20190116101028c237dc0b0003e50b  success 'download inttable inttable.txt'
20190115193039d37dc0b00015e67  success 'upload /Users/yangyi/Desktop/ODPS_2019-01-12_2019-01-14.csv maxcomputefee -c "UTF-8" -h "true" -dfp "yyyy-MM-dd HH:mm:ss"'
```

以下几种计算作业与SQL类似，按照#unique_105编写SQL即可。

5. 分析MapReduce作业消费

说明：

MR任务当日计算费用=当日总计算时*单价（0.46元）

```
--分析MR作业消费。
SELECT  TO_CHAR(starttime,'yyyymmdd') AS ds
        ,projectid
        ,(cu_usage/3600)*0.46 AS mr_fee
FROM    maxcomputefee
WHERE   type = 'MapReduce'
AND     TO_CHAR(starttime,'yyyymmdd') >= '20190112'
GROUP BY TO_CHAR(starttime,'yyyymmdd')
         ,projectid
         ,cu_usage
ORDER BY mr_fee DESC
;
```

6. 分析外部表作业（OTS和OSS）

说明：

SQL外部表功能计费规则： 一次SQL计算费用=计算输入数据量*SQL复杂度（1）*单价（0.03元/GB）

```
--分析OTS外部表SQL作业消费。
SELECT  TO_CHAR(starttime,'yyyymmdd') AS ds
```

```
        ,projectid
        ,(input_ots/1024/1024/1024)*1*0.03 AS ots_fee
FROM    maxcomputefee
WHERE   type = 'ComputationSql'
AND     TO_CHAR(starttime,'yyyymmdd') >= '20190112'
GROUP BY TO_CHAR(starttime,'yyyymmdd')
         ,projectid
         ,input_ots
ORDER BY ots_fee DESC
;

--分析OSS外部表SQL作业消费。
SELECT  TO_CHAR(starttime,'yyyymmdd') AS ds
        ,projectid
        ,(input_oss/1024/1024/1024)*1*0.03 AS ots_fee
FROM    maxcomputefee
WHERE   type = 'ComputationSql'
AND     TO_CHAR(starttime,'yyyymmdd') >= '20190112'
GROUP BY TO_CHAR(starttime,'yyyymmdd')
         ,projectid
         ,input_oss
ORDER BY ots_fee DESC
;
```

7. 分析Lightning查询费用

说明:

一次Lightning查询费用 = 查询输入数据量*单价（0.03元/GB）

```
SELECT  to_char(endtime,'yyyymmdd') as ds,feeid as instanceid
        ,projectid
        ,computationsqlcomplexity
        ,SUM((computationsqlinput / 1024 / 1024 / 1024)) as
computationsqlinput
        ,SUM((computationsqlinput / 1024 / 1024 / 1024)) *
computationsqlcomplexity * 0.03 AS sqlmoney
FROM    maxcomputefee
WHERE   TYPE = 'LightningQuery'
--AND to_char(endtime,'yyyymmdd') >= '20190112'
GROUP BY to_char(endtime,'yyyymmdd'),feeid
         ,projectid
         ,computationsqlcomplexity
ORDER BY sqlmoney DESC
LIMIT   10000
;
```

8. 分析Spark计算费用

说明:

Spark任务当日计算费用 = 当日总计算时*单价（0.66元/计算时）

```
--分析Spark作业消费。
SELECT  TO_CHAR(starttime,'yyyymmdd') AS ds
        ,projectid
        ,(cu_usage/3600)*0.66 AS mr_fee
FROM    maxcomputefee
WHERE   type = 'spark'
AND     TO_CHAR(starttime,'yyyymmdd') >= '20190112'
```

```
GROUP BY TO_CHAR(starttime,'yyyymmdd')
         ,projectid
         ,cu_usage
ORDER BY mr_fee DESC
;
```

MaxCompute产品消费的增长（暴涨）大多是因为作业量有大幅度提升。如果您想要优化费用成本，需要先检查SQL等作业中存在的问题。至于要优化具体哪一个SQL，本文能够给予您一些帮助。

更多信息

如果您想了解更多关于费用成本优化的文章，请参见云栖社区帮助企业做好MaxCompute大数据平台成本优化的最佳实践。