

Greenplum 转换 DATE 数据类型问题

| | |
|-------------------------------|---|
| Greenplum 转换 DATE 数据类型问题..... | 1 |
| 1 场景概述..... | 1 |
| 2 查看外表时的错误..... | 1 |
| 3 问题解答思路..... | 2 |
| 3.1 修改外表字段类型..... | 2 |
| 3.2 把外表映射成内表..... | 2 |
| 3.3 修改内表类型..... | 2 |
| 3.4 把内表修改成 DATE 类型..... | 3 |
| 3.5 更新字段效率查看..... | 3 |

1 场景概述

在工作中使用 Greenplum 外表时发现 date 类型中有 null 或空值,外表不识别类型,问题解答思路,先使用 varchar 类型把外表的数据加载到 Greenplum,在使用数据类型转化转化为 date 即可。

2 查看外表时的错误

```
[SQL]select * from test_external limit 100;
```

```
NOTICE: Found 8833 data formatting errors (8833 or more input rows). Rejected related input data.
```

```
[Err] ERROR: All 1000 first rows in this segment were rejected. Aborting operation regardless of REJECT LIMIT value. Last error was: invalid input syntax for type date: "null", column candate (seg17 slice1 192.168.209.14:40001 pid=417285)
```

```
DETAIL: External table xiaoxu_temp, line 1000 of gpfdfs://nameservice1/tmp/*****/***_all/*, column candate
```

以上问题是在查询外表时遇到了数据类型 date 为 null 的数值

3 问题解答思路

3.1 修改外表字段类型

在创建外表语句时把 `candate` 字段的类型修改成 `varchar`

3.2 把外表映射成内表

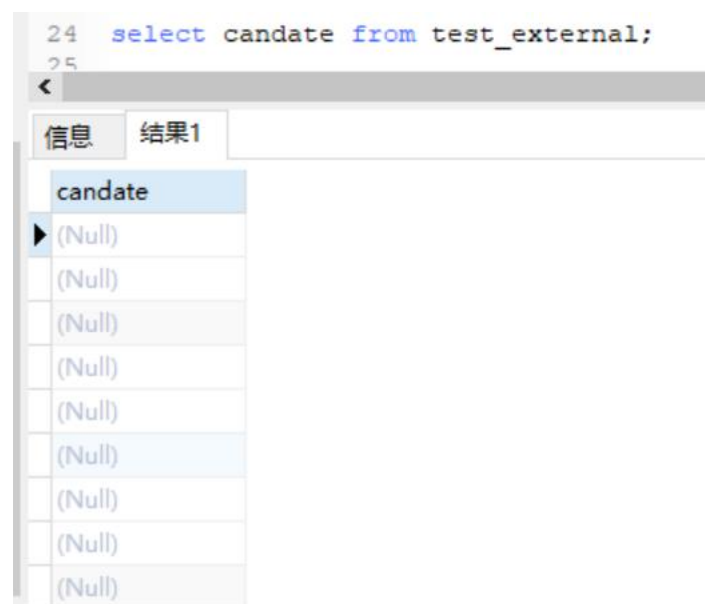
把外表的数据复制到内表中，映射语句如下：

```
create table temp_internal with (appendonly = true, compressstype = zlib, compresslevel = 5
,orientation=column) as
select * from test_external
Distributed by (id)
```

3.3 修改内表类型

```
update test_external set candate=null where candate ='null' or length(candate)=0;
```

以上语句是把内表中 candate 是 null 或空修改成 null，效果如下



3.4 把内表修改成 DATE 类型

以下语句是先把 candate 转换为 carchar 再转换为 date

```
alter table test_external alter column "candate" type date using ("candate"::VARCHAR::date);
```

| | | | |
|---|---------|------|--|
| 3 | | | |
| 4 | candate | date | |
| 5 | revdate | date | |

3.5 更新字段效率查看

```
[SQL]update table test_external set candate=null where candate='null';
```

时间: 145.664s

受影响的行: 45804388

```
[SQL]alter table table test_external alter column "candate" type date using ("candate"::VARCHAR::date);
```

时间: 214.938s

受影响的行: 0