

# 使用JDBC向Kudu表插入中文字符-双引号的秘密

原创

2017-09-09

Fayson

Hadoop实操

温馨提示：要看高清无码套图，请使用手机打开并单击图片放大查看。

## 1.问题描述

使用Impala JDBC向Kudu表中插入中文字符，插入的中文字符串乱码，中文字符串被截断。

## 2.问题复现

测试环境：

- CDH5.12.0
- Kudu1.4.0
- ImpalaJDBC41\_2.5.35

1.使用ImpalaJDBC代码进行测试，测试代码

```
static String JDBC_DRIVER = "com.cloudera.impala.jdbc41.Driver";
static String CONNECTION_URL = "jdbc:impala://ip-172-31-10-118:21050/default";
```

```
public static void main(String[] args) {
    Connection con = null;
    ResultSet rs = null;
    PreparedStatement ps = null;

    try {
        Class.forName(JDBC_DRIVER);
        con = DriverManager.getConnection(CONNECTION_URL);
        String insertsql = "insertinto my_first_table values(46, '测试中文字符')";

        ps = con.prepareStatement(insertsql);
        ps.execute();
        ps.close();

        ps = con.prepareStatement("select* from my_first_table order by id asc");
        rs = ps.executeQuery();
        while (rs.next()) {
            System.out.println(rs.getLong(1) + "\t" +rs.getString(2));
        }

    } catch (Exception e) {
        e.printStackTrace();
    } finally {
        try { // 关闭rs、ps和con
            rs.close();
            ps.close();
            con.close();
        } catch (SQLException e) {
            // TODO Auto-generated catch block
            e.printStackTrace();
        }
    }
}
```

2.向Kudu表中分别插入测试数据，如“测试”，“测试中文”，“测试中文字符”

```
String insertsql = "insert into my_first_table values(44, '测试')";
String insertsql = "insert into my_first_table values(45, '测试中文')";
String insertsql = "insert into my_first_table values(46, '测试中文字符')";
```

如下是按测试顺序插入的数据

5	44	测试
2	45	测试中文
6	46	测试中文字符

通过以上操作重现问题。

## 3.解决方法

修改程序中插入语句，将插入字符串的单引号修改为双引号

String insertsql = "insert into my\_first\_table values(51, \"测试中文字符\")";

String insertsql = "insert into my\_first\_table values(52, \"测试中文\")";

String insertsql = "insert into my\_first\_table values(53, \"测试\")";

修改后重新向Kudu中插入测试数据：“测试中文字符”，“测试中文”，“测试”  
使用Hue查询显示如下：

9	51	测试中文字符
24	52	测试中文
15	53	测试

中文字符串插入Kudu显示正常。

## 4.备注

1.使用Cloudera官网最新的JDBC驱动，插入中文字符时也有上述问题

下载地址：[https://downloads.cloudera.com/connectors/impala\\_jdbc\\_2.5.38.1058.zip](https://downloads.cloudera.com/connectors/impala_jdbc_2.5.38.1058.zip)

2.通过Impala-shell插入中文字符串正常

```
[172.31.10.118:21000] > insert into my_first_table values(66,'插入中文字符');
```

Modified 1 row(s), 0 row error(s) in 0.11s

```
[172.31.10.118:21000] > select * from my_first_table where id=66;
```

```
+----+-----+
| id | name      |
+----+-----+
| 66 | 插入中文字符 |
+----+-----+
```

Fetchd 1 row(s) in 0.21s

```
[172.31.10.118:21000] >
```

```
[172.31.10.118:21000] > insert into my_first_table values(77, "测试中文字符");
```

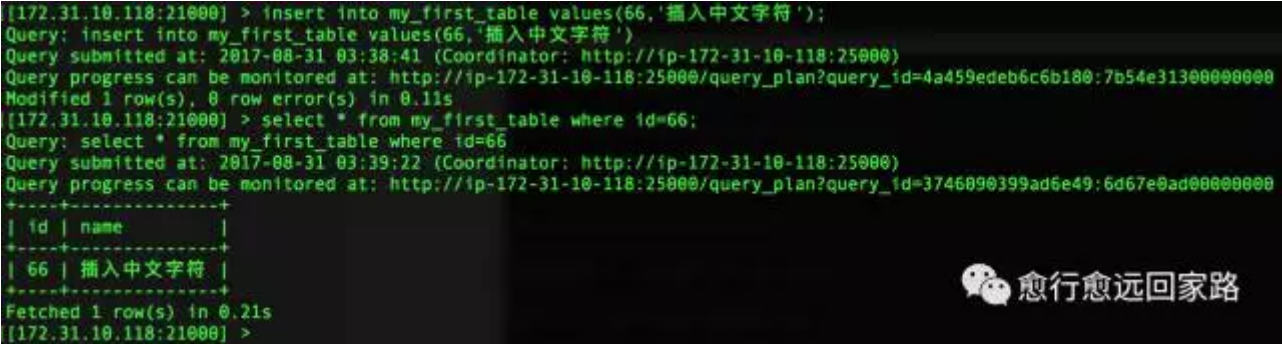
Modified 1 row(s), 0 row error(s) in 0.11s

```
[172.31.10.118:21000] > select * from my_first_table where id=77;
```

```
+----+-----+
| id | name      |
+----+-----+
| 77 | 测试中文字符 |
+----+-----+
```

Fetchd 1 row(s) in 0.18s

```
[172.31.10.118:21000] >
```



醉酒鞭名马，少年多浮夸！ 岭南浣溪沙，呕吐酒肆下！挚友不肯放，数据玩的花！  
温馨提示：要看高清无码套图，请使用手机打开并单击图片放大查看。

欢迎关注Hadoop实操，第一时间，分享更多Hadoop干货，喜欢请关注分享。