

SQOOP-1533

shigeki.hirose

2014-09-19

# Contents

# Chapter 1

## Root issue SQOOP-1533

### 1.1 Summary

Java 8 : sqoop import --create-hcatalog-table for mysql causes NullPointerException when column name contains upper case char

### 1.2 Description

Sqoop command example:

```
sqoop import --connect "jdbc:mysql://127.0.0.1:3306/xxx" --username xxx
--password xxx --table SampleTable --hcatalog-database xxx --create-hcatalog-table
--hcatalog-storage-stanza 'stored as orc' --hcatalog-table sampletable
```

Call stack:

```
sqoop.Sqoop: Got exception running Sqoop: java.lang.NullPointerException
java.lang.NullPointerException
at org.apache.sqoop.mapreduce.hcat.SqoopHCatUtilities.createHCatTable(SqoopHCatUtilities.
at org.apache.sqoop.mapreduce.hcat.SqoopHCatUtilities.configureHCat(SqoopHCatUtilities.
at org.apache.sqoop.mapreduce.hcat.SqoopHCatUtilities.configureImportOutputFormat(Sqoop
at org.apache.sqoop.mapreduce.ImportJobBase.configureOutputFormat(ImportJobBase.java:98)
at org.apache.sqoop.mapreduce.ImportJobBase.runImport(ImportJobBase.java:240)
at org.apache.sqoop.manager.SqlManager.importTable(SqlManager.java:614)
at org.apache.sqoop.manager.MySQLManager.importTable(MySQLManager.java:118)
at org.apache.sqoop.tool.ImportTool.importTable(ImportTool.java:413)
at org.apache.sqoop.tool.ImportTool.run(ImportTool.java:506)
at org.apache.sqoop.Sqoop.run(Sqoop.java:147)
at org.apache.hadoop.util.ToolRunner.run(ToolRunner.java:70)
at org.apache.sqoop.Sqoop.runSqoop(Sqoop.java:183)
at org.apache.sqoop.Sqoop.runTool(Sqoop.java:222)
at org.apache.sqoop.Sqoop.runTool(Sqoop.java:231)
at org.apache.sqoop.Sqoop.main(Sqoop.java:240)
```

This seems to be because source MySQL table contains a column name that has an upper case letter.

Sample table below. The column name, 'A', causes the exception.

```
1 DROP TABLE `SampleTable`;
2 CREATE TABLE `SampleTable` (
3   `id` bigint(20) unsigned NOT NULL AUTO_INCREMENT,
4   `A` bigint(20) unsigned NOT NULL DEFAULT '0',
5   PRIMARY KEY (`id`)
6 ) ENGINE=InnoDB DEFAULT CHARSET=utf8;
```

Once column name is changed to 'a', it works properly:

```
1 DROP TABLE `SampleTable`;
2 CREATE TABLE `SampleTable` (
3   `id` bigint(20) unsigned NOT NULL AUTO_INCREMENT,
4   `a` bigint(20) unsigned NOT NULL DEFAULT '0',
5   PRIMARY KEY (`id`)
6 ) ENGINE=InnoDB DEFAULT CHARSET=utf8;
```

This is probably because LCKeYMap (and HCKeYMap) class in SqoopHCatUtilities.java line 160 - 177 are expecting HashMap to call \*put\* method for each element when \*putAll\* is invoked, which seem to have been the case in Java 6/7, but it doesn't behave that way in Java 8.

```
1  /**
2   * A Map using String as key type that ignores case of its key and stores the
3   * key in lower case.
4   */
5  private static class LCKeYMap<V> extends HashMap<String, V> {
6      private static final long serialVersionUID = -6751510232323094216L;
7      @Override
8      public V put(String key, V value) {
9          return super.put(key.toLowerCase(), value);
10     }
11
12     @Override
13     public V get(Object key) {
14         return super.get(((String) key).toLowerCase());
15     }
16 }
```

## 1.3 Attachments

No attachments

## 1.4 Comments

1. **harisekhon:** I also have this problem or similar issue when importing from SQL Server (we do a lot of SQL Server imports, the only change was using HCatalog to try to write directly

as ORC). We're on Oracle Java 7 and the stack trace is different implying a different bug so I've raised a separate Jira here: [SQOOP-2271](#), although I've noticed we also have uppercase chars in column names.