

Hi,

I am trying to run the tutorial provided in TOS BD v5.4.1 titled Step_2_Hive_ELT. (I have completed the prior step to this as well). Each time I try and run the job provided in the tutorial I receive the following error:

```
java.io.IOException: Cannot run program "\usr\bin\hadoop.cmd" (in directory "C:\Users\dgoverman\Documents\TMO\HDL\Talend\TOS_BD-r111943-V5.4.1"): CreateProcess error=2, The system cannot find the file specified
```

I have simply downloaded and unzipped TOS, so I am unclear as to what file it is looking for and why it is missing. I am running this against HDP 1.3 (to avoid issues between Windows & Linux present with HDP 2.0). The use case I am trying to demonstrate is just a simple ELT job where two tables are joined to produce an output. Other jobs I have created to prove this simple case have encountered the exact same error. I have provided the full job log below for reference:

Starting job Step_2_Hive_ELT at 11:40 20/02/2014.

connecting to socket on port 3775

connected

: org.apache.hadoop.hive.metastore.HiveMetaStore - 0: Opening raw store with implementation class rg.apache.hadoop.hive.metastore.ObjectStore

: org.apache.hadoop.hive.metastore.ObjectStore - ObjectStore, initialize called

: DataNucleus.Persistence - Property datanucleus.cache.level2 unknown - will be ignored

: org.apache.hadoop.hive.metastore.ObjectStore - Setting MetaStore object pin classes with

hive.metastore.cache.pinobjtypes="Table,StorageDescriptor,SerdeInfo,Partition,Database,Type,FieldSchema,Order"

: org.apache.hadoop.hive.metastore.ObjectStore - Initialized ObjectStore

: org.apache.hadoop.security.UserGroupInformation - No groups available for user dgoverman

: org.apache.hadoop.security.UserGroupInformation - No groups available for user dgoverman

: org.apache.hadoop.hive.service.HiveServer - Putting temp output to file \tmp\dgoverman\dgoverman_12444@USSEADGOVERMAN1_2014022011403487408250001585367.pipeout

: org.apache.hadoop.hive.service.HiveServer - Running the query: set hive.fetch.output.serde = org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe

: org.apache.hadoop.hive.service.HiveServer - Putting temp output to file \tmp\dgoverman\dgoverman_12444@USSEADGOVERMAN1_2014022011403487408250001585367.pipeout

: org.apache.hadoop.hive.service.HiveServer - Running the query: SET mapred.job.map.memory.mb=1000

: org.apache.hadoop.hive.service.HiveServer - Putting temp output to file \tmp\dgoverman\dgoverman_12444@USSEADGOVERMAN1_2014022011403487408250001585367.pipeout

: org.apache.hadoop.hive.service.HiveServer - Running the query: SET mapred.job.reduce.memory.mb=1000

: org.apache.hadoop.hive.service.HiveServer - Putting temp output to file \tmp\dgoverman\dgoverman_12444@USSEADGOVERMAN1_2014022011403487408250001585367.pipeout

: org.apache.hadoop.hive.service.HiveServer - Running the query: INSERT OVERWRITE TABLE customerWithState SELECT customers.id, customers.First_Name, customers.Last_Name, customers.Address, customers.idState, customers.Revenue, us_states.State FROM customers LEFT OUTER JOIN us_states ON(us_states.stateid = customers.idState)

: org.apache.hadoop.hive ql.Driver - <PERFLOG method=Driver.run>

: org.apache.hadoop.hive ql.Driver - <PERFLOG method=TimeToSubmit>

: org.apache.hadoop.hive ql.Driver - <PERFLOG method=compile>

: hive ql.parse.ParseDriver - Parsing command: INSERT OVERWRITE TABLE customerWithState SELECT customers.id, customers.First_Name, customers.Last_Name, customers.Address, customers.idState, customers.Revenue, us_states.State FROM customers LEFT OUTER JOIN us_states ON(us_states.stateid = customers.idState)

: hive ql.parse.ParseDriver - Parse Completed

: org.apache.hadoop.hive ql.parse.SemanticAnalyzer - Starting Semantic Analysis

: org.apache.hadoop.hive ql.parse.SemanticAnalyzer - Completed phase 1 of Semantic Analysis

: org.apache.hadoop.hive ql.parse.SemanticAnalyzer - Get metadata for source tables

: hive.metastore - Trying to connect to metastore with URI thrift://sandbox:9083

: org.apache.hadoop.security.UserGroupInformation - No groups available for user dgoverman

: hive.metastore - Waiting 1 seconds before next connection attempt.

: hive.metastore - Connected to metastore.

: org.apache.hadoop.hive ql.parse.SemanticAnalyzer - Get metadata for subqueries

: org.apache.hadoop.hive ql.parse.SemanticAnalyzer - Get metadata for destination tables

: org.apache.hadoop.hive ql.parse.SemanticAnalyzer - Completed getting MetaData in Semantic Analysis

: org.apache.hadoop.hive ql.parse.TypeCheckProcFactory - Invalid type entry TOK_TABLE_OR_COL=null

: org.apache.hadoop.hive ql.parse.TypeCheckProcFactory - Invalid type entry TOK_TABLE_OR_COL=null

: org.apache.hadoop.hive ql.parse.TypeCheckProcFactory - Invalid type entry TOK_TABLE_OR_COL=null

: org.apache.hadoop.hive ql.parse.TypeCheckProcFactory - Invalid type entry TOK_TABLE_OR_COL=null

: org.apache.hadoop.hive ql.parse.TypeCheckProcFactory - Invalid type entry TOK_TABLE_OR_COL=null

: org.apache.hadoop.hive ql.parse.TypeCheckProcFactory - Invalid type entry TOK_TABLE_OR_COL=null

: org.apache.hadoop.hive ql.parse.TypeCheckProcFactory - Invalid type entry TOK_TABLE_OR_COL=null

: org.apache.hadoop.hive ql.parse.TypeCheckProcFactory - Invalid type entry TOK_TABLE_OR_COL=null

: org.apache.hadoop.hive ql.parse.TypeCheckProcFactory - Invalid type entry TOK_TABLE_OR_COL=null

: org.apache.hadoop.hive ql.parse.TypeCheckProcFactory - Invalid type entry TOK_TABLE_OR_COL=null

```
: org.apache.hadoop.hive.ql.ppd.OpProcFactory - Processing for FS(6)
: org.apache.hadoop.hive.ql.ppd.OpProcFactory - Processing for SEL(5)
: org.apache.hadoop.hive.ql.ppd.OpProcFactory - Processing for JOIN(4)
: org.apache.hadoop.hive.ql.ppd.OpProcFactory - Processing for RS(2)
: org.apache.hadoop.hive.ql.ppd.OpProcFactory - Processing for TS(0)
: org.apache.hadoop.hive.ql.ppd.OpProcFactory - Processing for RS(3)
: org.apache.hadoop.hive.ql.ppd.OpProcFactory - Processing for TS(1)
: org.apache.hadoop.util.NativeCodeLoader - Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
: org.apache.hadoop.io.compress.snappy.LoadSnappy - Snappy native library not loaded
: org.apache.hadoop.hive.ql.exec.Utilities - Cache Content Summary for hdfs://sandbox:8020/apps/hive/warehouse/customers length: 549588 file count: 1 directory count: 1
: org.apache.hadoop.hive.ql.exec.Utilities - Cache Content Summary for hdfs://sandbox:8020/apps/hive/warehouse/us_states length: 688 file count: 1 directory count: 1
: org.apache.hadoop.hive.ql.optimizer.physical.MetadataOnlyOptimizer - Looking for table scans where optimization is applicable
: org.apache.hadoop.hive.ql.optimizer.physical.MetadataOnlyOptimizer - Found 0 metadata only table scans
: org.apache.hadoop.hive.ql.optimizer.physical.MetadataOnlyOptimizer - Looking for table scans where optimization is applicable
: org.apache.hadoop.hive.ql.optimizer.physical.MetadataOnlyOptimizer - Found 0 metadata only table scans
: org.apache.hadoop.hive.ql.parse.SemanticAnalyzer - Completed plan generation
: org.apache.hadoop.hive.ql.Driver - Semantic Analysis Completed
Total MapReduce jobs = 1
: org.apache.hadoop.hive.ql.Driver - Returning Hive schema: Schema(fieldSchemas:, properties:null)
: org.apache.hadoop.hive.ql.Driver - </PERFLOG method=compile start=1392925242000 end=1392925244112 duration=2112>
: org.apache.hadoop.hive.ql.Driver - <PERFLOG method=Driver.execute>
: org.apache.hadoop.hive.ql.Driver - Starting command: INSERT OVERWRITE TABLE customerWithState SELECT customers.id, customers.First_Name, customers.Last_Name,
customers.Address, customers.idState, customers.Revenue, us_states.State FROM customers LEFT OUTER JOIN us_states ON( us_states.stateid = customers.idState )
: org.apache.hadoop.hive.ql.Driver - Total MapReduce jobs = 1
: org.apache.hadoop.hive.ql.Driver - </PERFLOG method=TimeToSubmit start=1392925242000 end=1392925244115 duration=2115>
: org.apache.hadoop.hive.ql.exec.MapredLocalTask - Generating plan file file:/tmp/dgoverman/hive_2014-02-20_11-40-42_039_5064415122519008878/-local-10004/plan.xml
setting HADOOP_USER_NAME dgoverman
: org.apache.hadoop.hive.ql.exec.Task - setting HADOOP_USER_NAME dgoverman
: org.apache.hadoop.hive.ql.exec.MapredLocalTask - Executing: \usr\bin\hadoop.cmd jar
/C:/Users/dgoverman/Documents/TMO/HDL/Talend/TOS_BD-r111943-V5.4.1/workspace/.Java/lib/hive-exec-0.11.0.1.3.0.0-107.jar org.apache.hadoop.hive.ql.exec.ExecDriver -localtask -plan
file:/tmp/dgoverman/hive_2014-02-20_11-40-42_039_5064415122519008878/-local-10004/plan.xml -jobconf file
file:/tmp/dgoverman/hive_2014-02-20_11-40-42_039_5064415122519008878/-local-10005/jobconf.xml
java.io.IOException: Cannot run program "\usr\bin\hadoop.cmd" (in directory "C:\Users\dgoverman\Documents\TMO\HDL\Talend\TOS_BD-r111943-V5.4.1"): CreateProcess error=2, The
system cannot find the file specified
at java.lang.ProcessBuilder.start(Unknown Source)
at java.lang.Runtime.exec(Unknown Source)
at java.lang.Runtime.exec(Unknown Source)
at org.apache.hadoop.hive.ql.exec.MapredLocalTask.execute(MapredLocalTask.java:240)
at org.apache.hadoop.hive.ql.exec.Task.executeTask(Task.java:144)
at org.apache.hadoop.hive.ql.exec.TaskRunner.runSequential(TaskRunner.java:57)
at org.apache.hadoop.hive.ql.Driver.launchTask(Driver.java:1355)
at org.apache.hadoop.hive.ql.Driver.execute(Driver.java:1139)
at org.apache.hadoop.hive.ql.Driver.run(Driver.java:945)
at org.apache.hadoop.hive.service.HiveServer$HiveServerHandler.execute(HiveServer.java:198)
at org.apache.hadoop.hive.jdbc.HiveStatement.executeQuery(HiveStatement.java:192)
at org.apache.hadoop.hive.jdbc.HiveStatement.execute(HiveStatement.java:132)
at bigdatademos.step_2_hive_elt_0_1.Step_2_Hive_ELt.tELTHiveMap_1Process(Step_2_Hive_ELt.java:602)
: org.apache.hadoop.hive.ql.exec.MapredLocalTask - Exception: Cannot run program "\usr\bin\hadoop.cmd" (in directory
"C:\Users\dgoverman\Documents\TMO\HDL\Talend\TOS_BD-r111943-V5.4.1"): CreateProcess error=2, The system cannot find the file specified
: org.apache.hadoop.hive.ql.Driver - FAILED: Execution Error, return code 1 from org.apache.hadoop.hive.ql.exec.MapredLocalTask
: org.apache.hadoop.hive.ql.Driver - </PERFLOG method=Driver.execute start=1392925244112 end=1392925244384 duration=272>
: org.apache.hadoop.hive.ql.Driver - <PERFLOG method=releaseLocks>
: org.apache.hadoop.hive.ql.Driver - </PERFLOG method=releaseLocks start=1392925244384 end=1392925244384 duration=0>
```

```
at bigdatademos.step_2_hive_elt_0_1.Step_2_Hive_ELt.tHiveConnection_2Process(Step_2_Hive_ELt.java:1723)
at bigdatademos.step_2_hive_elt_0_1.Step_2_Hive_ELt.runJobInTOS(Step_2_Hive_ELt.java:1978)
at bigdatademos.step_2_hive_elt_0_1.Step_2_Hive_ELt.main(Step_2_Hive_ELt.java:1801)
Caused by: java.io.IOException: CreateProcess error=2, The system cannot find the file specified
at java.lang.ProcessImpl.create(Native Method)
at java.lang.ProcessImpl.<init>(Unknown Source)
at java.lang.ProcessImpl.start(Unknown Source)
... 16 more
FAILED: Execution Error, return code 1 from org.apache.hadoop.hive.ql.exec.MapredLocalTask
Exception in component tELTHiveOutput_1
java.sql.SQLException: Query returned non-zero code: 1, cause: FAILED: Execution Error, return code 1 from org.apache.hadoop.hive.ql.exec.MapredLocalTask
at org.apache.hadoop.hive.jdbc.HiveStatement.executeQuery(HiveStatement.java:194)
at org.apache.hadoop.hive.jdbc.HiveStatement.execute(HiveStatement.java:132)
at bigdatademos.step_2_hive_elt_0_1.Step_2_Hive_ELt.tELTHiveMap_1Process(Step_2_Hive_ELt.java:602)
at bigdatademos.step_2_hive_elt_0_1.Step_2_Hive_ELt.tHiveConnection_2Process(Step_2_Hive_ELt.java:1723)
at bigdatademos.step_2_hive_elt_0_1.Step_2_Hive_ELt.runJobInTOS(Step_2_Hive_ELt.java:1978)
at bigdatademos.step_2_hive_elt_0_1.Step_2_Hive_ELt.main(Step_2_Hive_ELt.java:1801)
disconnected
Job Step_2_Hive_ELt ended at 11:40 20/02/2014.
```

Thanks in advance for any help you can offer in resolving this error so I can complete the tutorial and demonstrate my use case.
David