

## 1. SQOOP

Араше Sqoop — это инструмент, предназначенный для передачи данных между Hadoop и реляционными базами данных или мэйнфреймами.

Запрашиваю список БД в PosrgeSQL, чтобы в дальнейшем выполнить импорт БД PSQL в hive

```
[student3_16@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ sqoop list-databases --connect jdbc:postgresql://10.0.0.7/lesson5 --username exporter --password exporter_pass
Warning: /usr/hdp/3.1.4.0-315/accumulo does not exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/hdp/3.1.4.0-315/hadoop/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/hdp/3.1.4.0-315/phoenix/phoenix-5.0.0.3.1.4.0-315-server.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/hdp/3.1.4.0-315/hive/lib/log4j-slf4j-impl-2.10.0.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
21/04/25 16:37:59 INFO sqoop.Sqoop: Running Sqoop version: 1.4.7.3.1.4.0-315
21/04/25 16:37:59 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.
21/04/25 16:37:59 INFO manager.SqlManager: Using default fetchSize of 1000
postgres
ambari
template1
template0
hive
oozie
lesson5
student21
[student3_16@bigdataanalytics2-head-shdpt-v31-1-0 ~]$
```

Запрашиваю список таблиц

```
[student3_16@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ sqoop list-tables --connect jdbc:postgresql://10.0.0.7/lesson5 --username exporter --password exporter_pass
Warning: /usr/hdp/3.1.4.0-315/accumulo does not exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/hdp/3.1.4.0-315/hadoop/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/hdp/3.1.4.0-315/phoenix/phoenix-5.0.0.3.1.4.0-315-server.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/hdp/3.1.4.0-315/hive/lib/log4j-slf4j-impl-2.10.0.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
21/04/25 16:44:14 INFO sqoop.Sqoop: Running Sqoop version: 1.4.7.3.1.4.0-315
21/04/25 16:44:14 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.
21/04/25 16:44:14 INFO manager.SqlManager: Using default fetchSize of 1000
test
[student3_16@bigdataanalytics2-head-shdpt-v31-1-0 ~]$
```

Выполняю команду импорта с помощью sqoop из БД PSQL в Hive используя параметр -m 1

```
[student3_16@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ sqoop import --connect jdbc:postgresql://10.0.0.7/lesson5 --username exporter --password exporter_pass --table test --hive-import --hive-database studen
t3_16 --hive-table test
Warning: /usr/hdp/3.1.4.0-315/accumulo does not exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/hdp/3.1.4.0-315/hadoop/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/hdp/3.1.4.0-315/phoenix/phoenix-5.0.0.3.1.4.0-315-server.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/hdp/3.1.4.0-315/hive/lib/log4j-slf4j-impl-2.10.0.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
21/04/25 16:46:34 INFO sqoop.Sqoop: Running Sqoop version: 1.4.7.3.1.4.0-315
21/04/25 16:46:34 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.
21/04/25 16:46:34 INFO tool.BaseSqoopTool: Using Hive-specific delimiters for output. You can override
21/04/25 16:46:34 INFO tool.BaseSqoopTool: delimiters with --fields-terminated-by, etc.
21/04/25 16:46:35 INFO manager.SqlManager: Using default fetchSize of 1000
21/04/25 16:46:35 INFO tool.CodeGenTool: Beginning code generation
21/04/25 16:46:35 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM "test" AS t LIMIT 1
21/04/25 16:46:35 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /usr/hdp/3.1.4.0-315/hadoop-mapreduce
21/04/25 16:46:37 INFO org.apache.hadoop.mapreduce.Compiler: Writing jar file: /tmp/sqoop-student3_16/compile/992170a1c5a1cf6d4b758091ef4150c4/test.jar
21/04/25 16:46:37 WARN manager.PostgresqlManager: It looks like you are importing from postgresql.
21/04/25 16:46:37 WARN manager.PostgresqlManager: This transfer can be faster! Use the --direct
21/04/25 16:46:37 WARN manager.PostgresqlManager: option to exercise a postgresql-specific fast path.
21/04/25 16:46:37 INFO mapreduce.ImportJobBase: Beginning import of test
21/04/25 16:46:39 INFO client.NMProxy: Connecting to ResourceManager at bigdataanalytics2-head-shdpt-v31-1-0.novalocal/10.0.0.7:8050
21/04/25 16:46:39 INFO client.AHSProxy: Connecting to Application History server at bigdataanalytics2-head-shdpt-v31-1-0.novalocal/10.0.0.7:10200
21/04/25 16:46:40 INFO mapreduce.JobResourceUploader: Disabling Erasure coding for path: /user/student3_16/staging/job_1618415779628_0964
21/04/25 16:46:42 INFO db.DbInputFormat: Using read committed transaction isolation
21/04/25 16:46:42 INFO mapreduce.JobSubmitter: number of splits:1
21/04/25 16:46:43 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1618415779628_0964
21/04/25 16:46:43 INFO mapreduce.JobSubmitter: Executing with tokens: {}
21/04/25 16:46:43 INFO conf.Configuration: Found resource resource-types.xml at file:/etc/hadoop/3.1.4.0-315/0/resource-types.xml
21/04/25 16:46:43 INFO impl.YarnClientImpl: Submitted application application_1618415779628_0964
21/04/25 16:46:43 INFO mapreduce.Job: The url to track the job: http://bigdataanalytics2-head-shdpt-v31-1-0.novalocal:8088/proxy/application_1618415779628_0964/
21/04/25 16:46:43 INFO mapreduce.Job: Running job: job_1618415779628_0964
21/04/25 16:46:50 INFO mapreduce.Job: Job job_1618415779628_0964 running in uber mode : false
21/04/25 16:46:56 INFO mapreduce.Job: map 0% reduce 0%
21/04/25 16:51:52 INFO mapreduce.Job: Job job_1618415779628_0964 completed successfully
21/04/25 16:51:52 INFO mapreduce.Job: Counters: 32
File System Counters
FILE: Number of bytes read=0
FILE: Number of bytes written=243957
FILE: Number of read operations=0
FILE: Number of large read operations=0
FILE: Number of write operations=0
HDFS: Number of bytes read=5
HDFS: Number of bytes written=19
HDFS: Number of read operations=0
HDFS: Number of large read operations=0
HDFS: Number of write operations=2
Job Counters
Launched map tasks=1
Other local map tasks=1
Total time spent by all maps in occupied slots (ms)=14852
```

```

Peak Map Physical memory (bytes)=217870316
Peak Map Virtual memory (bytes)=3687260160
File Input Format Counters
  Bytes Read:
File Output Format Counters
  Bytes Written:
21/04/25 16:51:50 INFO mapreduce.ImportJobBase: Transferred 19 bytes in 313.7232 seconds (0.0006 bytes/sec)
21/04/25 16:51:52 INFO mapreduce.ImportJobBase: Retrieved 2 records.
21/04/25 16:51:52 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM "test" AS t LIMIT 1
21/04/25 16:51:52 INFO hive.HiveImport: Loading uploaded data into hive
21/04/25 16:51:54 INFO hive.HiveImport: SLF4J: Class path contains multiple SLF4J bindings.
21/04/25 16:51:54 INFO hive.HiveImport: SLF4J: Found binding in [jar:file:/usr/hdp/3.1.4.0-315/hive/lib/log4j-slf4j-impl-2.10.0.jar!/org/slf4j/impl/StaticLoggerBinder.class]
21/04/25 16:51:54 INFO hive.HiveImport: SLF4J: Found binding in [jar:file:/usr/hdp/3.1.4.0-315/hadoop/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
21/04/25 16:51:54 INFO hive.HiveImport: SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
21/04/25 16:51:54 INFO hive.HiveImport: SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]
21/04/25 16:51:57 INFO hive.HiveImport: Connecting to jdbc:hive2://bigdataanalytics2-worker-shdpt-v31-1-5.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-0.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-4.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-3.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-2.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-1.novalocal:2181/default;password=hive;serviceDiscoveryMode=zooKeeper;user=hive;zooKeeperNamespace=hiveserver2
21/04/25 16:51:58 INFO hive.HiveImport: 21/04/25 16:51:58 [main]: INFO jdbc.HiveConnection: Connected to bigdataanalytics2-head-shdpt-v31-1-0.novalocal:10000
21/04/25 16:51:58 INFO hive.HiveImport: Connected to: Apache Hive (version 3.1.0.3.1.4.0-315)
21/04/25 16:51:58 INFO hive.HiveImport: Driver: Hive JDBC (version 3.1.0.3.1.4.0-315)
21/04/25 16:51:58 INFO hive.HiveImport: Transaction isolation: TRANSACTION_REPEATABLE_READ
21/04/25 16:51:58 INFO hive.HiveImport: 0: jdbc:hive2://bigdataanalytics2-worker-shdp: CREATE TABLE IF NOT EXISTS 'student3_16`.`test' ( `id` INT, `str` STRING) COMMENT 'Imported by sqoop on 2021/04/25 16:51:52' ROW FORMAT DELIMITED FIELDS TERMINATED BY '\012' STORED AS TEXTFILE;
21/04/25 16:51:58 INFO hive.HiveImport: ATED BY '\012' LINES TERMINATED BY '\012' STORED AS TEXTFILE;
21/04/25 16:51:58 INFO hive.HiveImport: INFO : Compiling command(queryId=hive_20210425165158_f6b347a5-ccd1-4ec0-93c6-555894fe5c44): CREATE TABLE IF NOT EXISTS 'student3_16`.`test' ( `id` INT, `str` STRING) COMMENT 'Imported by sqoop on 2021/04/25 16:51:52' ROW FORMAT DELIMITED FIELDS TERMINATED BY '\012' LINES TERMINATED BY '\012' STORED AS TEXTFILE
21/04/25 16:51:58 INFO hive.HiveImport: INFO : Semantic Analysis Completed (retrial = false)
21/04/25 16:51:58 INFO hive.HiveImport: INFO : Returning Hive schema: Schema(fieldsSchemas:null, properties:null)
21/04/25 16:51:58 INFO hive.HiveImport: INFO : Completed compiling command(queryId=hive_20210425165158_f6b347a5-ccd1-4ec0-93c6-555894fe5c44): CREATE TABLE IF NOT EXISTS 'student3_16`.`test' ( `id` INT, `str` STRING) COMMENT 'Imported by sqoop on 2021/04/25 16:51:52' ROW FORMAT DELIMITED FIELDS TERMINATED BY '\012' LINES TERMINATED BY '\012' STORED AS TEXTFILE
21/04/25 16:51:58 INFO hive.HiveImport: INFO : Starting task [Stage-0:DDL] in serial mode
21/04/25 16:51:58 INFO hive.HiveImport: INFO : Executing command(queryId=hive_20210425165158_f6b347a5-ccd1-4ec0-93c6-555894fe5c44): CREATE TABLE IF NOT EXISTS 'student3_16`.`test' ( `id` INT, `str` STRING) COMMENT 'Imported by sqoop on 2021/04/25 16:51:52' ROW FORMAT DELIMITED FIELDS TERMINATED BY '\012' LINES TERMINATED BY '\012' STORED AS TEXTFILE
21/04/25 16:51:58 INFO hive.HiveImport: INFO : OK
21/04/25 16:51:58 INFO hive.HiveImport: No rows affected (0.208 seconds)
21/04/25 16:51:58 INFO hive.HiveImport: 0: jdbc:hive2://bigdataanalytics2-worker-shdp: LOAD DATA INPATH 'hdfs://bigdataanalytics2-head-shdpt-v31-1-0.novalocal:8020/user/student3_16/test' INTO TABLE 'student3_16`.`test';
21/04/25 16:51:59 INFO hive.HiveImport: INFO : Compiling command(queryId=hive_20210425165158_114bdcca-5e38-4491-9192-1739b6b47f20): LOAD DATA INPATH 'hdfs://bigdataanalytics2-head-shdpt-v31-1-0.novalocal:8020/user/student3_16/test' INTO TABLE 'student3_16`.`test'
21/04/25 16:51:59 INFO hive.HiveImport: INFO : Semantic Analysis Completed (retrial = false)
21/04/25 16:51:59 INFO hive.HiveImport: INFO : Returning Hive schema: Schema(fieldsSchemas:null, properties:null)
21/04/25 16:51:59 INFO hive.HiveImport: INFO : Completed compiling command(queryId=hive_20210425165158_114bdcca-5e38-4491-9192-1739b6b47f20): LOAD DATA INPATH 'hdfs://bigdataanalytics2-head-shdpt-v31-1-0.novalocal:8020/user/student3_16/test' INTO TABLE 'student3_16`.`test'
21/04/25 16:51:59 INFO hive.HiveImport: INFO : Starting task [Stage-0:MOVE] in serial mode
21/04/25 16:51:59 INFO hive.HiveImport: INFO : Loading data to table student3_16.test from hdfs://bigdataanalytics2-head-shdpt-v31-1-0.novalocal:8020/user/student3_16/test
21/04/25 16:51:59 INFO hive.HiveImport: INFO : Starting task [Stage-1:STATS] in serial mode
21/04/25 16:51:59 INFO hive.HiveImport: INFO : Completed executing command(queryId=hive_20210425165158_114bdcca-5e38-4491-9192-1739b6b47f20): LOAD DATA INPATH 'hdfs://bigdataanalytics2-head-shdpt-v31-1-0.novalocal:8020/user/student3_16/test' INTO TABLE 'student3_16`.`test'
21/04/25 16:51:59 INFO hive.HiveImport: INFO : OK
21/04/25 16:51:59 INFO hive.HiveImport: No rows affected (0.204 seconds)
21/04/25 16:51:59 INFO hive.HiveImport: 0: jdbc:hive2://bigdataanalytics2-worker-shdp:
21/04/25 16:51:59 INFO hive.HiveImport: 0: jdbc:hive2://bigdataanalytics2-worker-shdp: Closing: 0: jdbc:hive2://bigdataanalytics2-worker-shdpt-v31-1-5.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-0.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-4.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-3.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-2.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-1.novalocal:2181/default;password=hive;serviceDiscoveryMode=zooKeeper;user=hive;zooKeeperNamespace=hiveserver2
21/04/25 16:51:59 INFO hive.HiveImport: Hive import complete.
21/04/25 16:51:59 INFO hive.HiveImport: Export directory is not empty, keeping it.
21/04/25 16:51:59 INFO tool.ImportTool: ps: Pushing hive/ncat import job data to listeners for table test
[student3_16@bigdataanalytics2-head-shdpt-v31-1-0 ~]$

```

## Захожу Hive в БД student3\_16 и проверяю содержимое таблицы Test

```

[student3_16@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hive
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/hdp/3.1.4.0-315/hive/lib/log4j-slf4j-impl-2.10.0.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/hdp/3.1.4.0-315/hadoop/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]
Connecting to jdbc:hive2://bigdataanalytics2-worker-shdpt-v31-1-5.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-0.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-4.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-3.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-2.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-1.novalocal:2181/default;password=hive;serviceDiscoveryMode=zooKeeper;user=hive;zooKeeperNamespace=hiveserver2
21/04/25 16:55:21 [main]: INFO jdbc.HiveConnection: Connected to bigdataanalytics2-head-shdpt-v31-1-0.novalocal:10000
Connected to: Apache Hive (version 3.1.0.3.1.4.0-315)
Driver: Hive JDBC (version 3.1.0.3.1.4.0-315)
Transaction isolation: TRANSACTION_REPEATABLE_READ
Beeline version 3.1.0.3.1.4.0-315 By Apache Hive
0: jdbc:hive2://bigdataanalytics2-worker-shdp: use student3_16;
INFO : Compiling command(queryId=hive_20210425165532_40aba3c8-1ca4-4092-bda4-87ab7e494b6f): use student3_16
INFO : Semantic Analysis Completed (retrial = false)
INFO : Returning Hive schema: Schema(fieldsSchemas:null, properties:null)
INFO : Completed compiling command(queryId=hive_20210425165532_40aba3c8-1ca4-4092-bda4-87ab7e494b6f): use student3_16
INFO : Executing command(queryId=hive_20210425165532_40aba3c8-1ca4-4092-bda4-87ab7e494b6f): use student3_16
INFO : Starting task [Stage-0:DDL] in serial mode
INFO : Completed executing command(queryId=hive_20210425165532_40aba3c8-1ca4-4092-bda4-87ab7e494b6f): use student3_16
INFO : OK
No rows affected (0.173 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp: show tables;
INFO : Compiling command(queryId=hive_20210425165540_179a705d-92fb-4f48-8031-e7ef24741b5b): show tables
INFO : Semantic Analysis Completed (retrial = false)
INFO : Returning Hive schema: Schema(fieldsSchemas:[FieldSchema(name:tab_name, type:string, comment:from deserializer)], properties:null)
INFO : Completed compiling command(queryId=hive_20210425165540_179a705d-92fb-4f48-8031-e7ef24741b5b): show tables
INFO : Executing command(queryId=hive_20210425165540_179a705d-92fb-4f48-8031-e7ef24741b5b): show tables
INFO : Starting task [Stage-0:DDL] in serial mode
INFO : Completed executing command(queryId=hive_20210425165540_179a705d-92fb-4f48-8031-e7ef24741b5b): show tables
INFO : OK
+-----+
| tab_name |
+-----+
| test     |
| vehicles |
| vehicles_2 |
+-----+
3 rows selected (0.059 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp:

```

```

0: jdbc:hive2://bigdataanalytics2-worker-shdp: select * from test limit 10;
INFO : Compiling command(queryId=hive_20210425165956_306a90a0-8dc9-459f-af98-ff1c3fd5ee5b): select * from test limit 10
INFO : Semantic Analysis Completed (retrial = false)
INFO : Returning Hive schema: Schema(fieldsSchemas:[FieldSchema(name:test.id, type:int, comment:null), FieldSchema(name:test.str, type:string, comment:null)], properties:null)
INFO : Completed compiling command(queryId=hive_20210425165956_306a90a0-8dc9-459f-af98-ff1c3fd5ee5b): select * from test limit 10
INFO : Executing command(queryId=hive_20210425165956_306a90a0-8dc9-459f-af98-ff1c3fd5ee5b): select * from test limit 10
INFO : Completed executing command(queryId=hive_20210425165956_306a90a0-8dc9-459f-af98-ff1c3fd5ee5b): select * from test limit 10
INFO : OK
+-----+
| test.id | test.str |
+-----+
| 1       | myText  |
| 2       | myText2 |
+-----+
2 rows selected (0.218 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp:

```



Чтобы удостовериться что в импортируемой базе 2 записи, захожу в PSQL и проверяю содержимое таблицы test

```
[student3_16@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ psql -d lesson5 -U exporter
psql (9.2.24, server 9.6.21)
WARNING: psql version 9.2, server version 9.6.
        Some psql features might not work.
Type "help" for help.

lesson5=> select * from test limit 10;
 id |  str
----+-----
   1 | myText
   2 | myText2
(2 rows)

lesson5=> █
```

Захожу в hdfs dfs и проверяю содержимое папки test. Наличие файла \_SUCCESS говорит об успешном импорте БД из PSQL в Hive

```
[student3_16@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -ls test
Found 2 items
-rw-r--r--  2 student3_16 student3_16      0 2021-04-25 16:46 test/_SUCCESS
-rw-r--r--  2 student3_16 student3_16    19 2021-04-25 16:46 test/part-m-00000
[student3_16@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ █
```

Удаляю папку test в hdfs

```
[student3_16@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -rm -r -skipTrash test
Deleted test
[student3_16@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ █
```

Удаляю таблицу test в Hive

```
Beeline version 3.1.0.3.1.4.0-315 By Apache Hive
0: jdbc:hive2://bigdataanalytics2-worker-shdp> use student3_16;
INFO : Compiling command(queryId=hive_20210425170846_89bc9d38-9748-4321-9723-2f811397ccdd): use student3_16
INFO : Semantic Analysis Completed (retrial = false)
INFO : Returning Hive schema: Schema(fieldSchemas:null, properties:null)
INFO : Completed compiling command(queryId=hive_20210425170846_89bc9d38-9748-4321-9723-2f811397ccdd); Time taken: 0.004 seconds
INFO : Executing command(queryId=hive_20210425170846_89bc9d38-9748-4321-9723-2f811397ccdd): use student3_16
INFO : Starting task [Stage-0:DDL] in serial mode
INFO : Completed executing command(queryId=hive_20210425170846_89bc9d38-9748-4321-9723-2f811397ccdd); Time taken: 0.012 seconds
INFO : OK
No rows affected (0.124 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp> drop table test;
INFO : Compiling command(queryId=hive_20210425170856_84fdbefa-4935-4820-b84a-618c62115551): drop table test
INFO : Semantic Analysis Completed (retrial = false)
INFO : Returning Hive schema: Schema(fieldSchemas:null, properties:null)
INFO : Completed compiling command(queryId=hive_20210425170856_84fdbefa-4935-4820-b84a-618c62115551); Time taken: 0.041 seconds
INFO : Executing command(queryId=hive_20210425170856_84fdbefa-4935-4820-b84a-618c62115551): drop table test
INFO : Starting task [Stage-0:DDL] in serial mode
INFO : Completed executing command(queryId=hive_20210425170856_84fdbefa-4935-4820-b84a-618c62115551); Time taken: 0.115 seconds
INFO : OK
No rows affected (0.191 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp> █
```

Выполняю команду импорта с помощью sqoop из БД PSQL в Hive используя параметр `-split-by id`

```

hive> cd /home/hadoopanalytics2-head-shdgt-v31-0-0 && sqoop import --connect jdbc:postgresql://10.0.0.7/lessons -username exporter -password exporter_pass --table test --hive-import --hive-database student3 --hive-table test --split-by id
Warning: user/hdp/3.1.4.0-315/accumulo does not exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jarfile:/usr/hdp/3.1.4.0-315/hadoop/lib/slf4j-log4j12-1.7.25.jar/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jarfile:/usr/hdp/3.1.4.0-315/phoenix/phoenix-5.0.0.3.1.4.0-315-server.jar/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jarfile:/usr/hdp/3.1.4.0-315/hive/lib/log4j-slf4j-impl-2.10.0.jar/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: The following classes are also found, but no binding was loaded for them:
org.apache.hadoop.hdfs.tools.DfsAdminTool$DfsAdminToolFactory
21/04/25 17:10:16 INFO sqoop.Sqoop: Running Sqoop version: 1.4.7.3.1.4.0-315
21/04/25 17:10:16 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.
21/04/25 17:10:16 INFO tool.BaseSqoopTool: Using native-specific delimiters for output. You can override
21/04/25 17:10:16 INFO tool.BaseSqoopTool: delimiters with --fields-terminated-by, etc..
21/04/25 17:10:16 INFO manager.SqlManager: Using default fetchSize of 1000
21/04/25 17:10:16 INFO tool.CodeGenTool: Beginning code generation
21/04/25 17:10:17 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `test` AS t LIMIT 1
21/04/25 17:10:17 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /usr/hdp/3.1.4.0-315/hadoop-mapreduce
21/04/25 17:10:19 INFO orm.CompilationManager: Could not rename /tmp/sqoop-student3_16/compile/e398f3f6e1263704ee420cfe7b10/test.java to /home/student3_16/.//test.java. Error: Destination '/home/student3_16/.//test.java' already exists.
21/04/25 17:10:19 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-student3_16/compile/e398f3f6e1263704ee420cfe7b10/test-jar
21/04/25 17:10:19 WARN manager.PostgresManager: It looks like you are importing from postgresql.
21/04/25 17:10:19 WARN manager.PostgresManager: This transfer can be faster! Use the --direct
21/04/25 17:10:19 WARN manager.PostgresManager: option to exercise a postgresql-specific fast path.
21/04/25 17:10:19 INFO mapreduce.JobImportBase: Beginning import of test
21/04/25 17:10:21 INFO client.HBProxy: Connecting to ResourceManager at bigdataanalytics2-head-shdgt-v31-0-nvalocal/10.0.0.7:8050
21/04/25 17:10:21 INFO client.AMSPROXY: Connecting to Application History server at bigdataanalytics2-head-shdgt-v31-0-nvalocal/10.0.0.7:10200
21/04/25 17:10:22 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /user/student3_16/staging/job_1618415779628_0965
21/04/25 17:10:24 INFO DBInputFormat: Using read committed transaction isolation
21/04/25 17:10:24 INFO db.DatabaseInputFormat: BoundingValuesQuery: SELECT MIN('id'), MAX('id') FROM `test`
21/04/25 17:10:24 INFO db.IntegersSplitter: Split size: 0; Num splits: 4 from: 1 to: 2
21/04/25 17:10:24 INFO mapreduce.JobSubmitter: number of splits=2
21/04/25 17:10:24 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1618415779628_0965
21/04/25 17:10:24 INFO mapreduce.JobSubmitter: Executing with tokens: []
21/04/25 17:10:24 INFO conf.Configuration: Found resource types.xml at file:/etc/hadoop/3.1.4.0-315/or/resource-types.xml
21/04/25 17:10:25 INFO url.URLClient: submitted application mapreduce.job_1618415779628_0965
21/04/25 17:10:25 INFO mapreduce.Job: The url to track the job: http://bigdataanalytics2-head-shdgt-v31-0-nvalocal:8088/proxy/application/1618415779628_0965/
21/04/25 17:10:25 INFO mapreduce.Job: Running job: job_1618415779628_0965
21/04/25 17:10:32 INFO mapreduce.Job: Job job_1618415779628_0965 running in uber mode : false
21/04/25 17:10:32 INFO mapreduce.Job: map 0% reduce 0%
21/04/25 17:10:37 INFO mapreduce.Job: map 50% reduce 0%
21/04/25 17:10:38 INFO mapreduce.Job: map 100% reduce 0%
21/04/25 17:15:40 INFO mapreduce.Job: Job job_1618415779628_0965 completed successfully
21/04/25 17:15:40 INFO mapreduce.Job: Counters: 32
File System Counters
FILE: Number of bytes read=0
FILE: Number of bytes written=488240
FILE: Number of larger read operations=0
FILE: Number of write operations=0
HDFS: Number of bytes read=193

```

```

21/04/25 17:15:40 INFO mapreduce.JobSubmitter: Retried 2 records.
21/04/25 17:15:40 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM test AS t LIMIT 1
21/04/25 17:15:40 INFO hive.metastore: Loading uploaded data into Hive
21/04/25 17:15:42 INFO hive.HiveImport: SLF4J: Class path contains multiple SLF4J bindings.
21/04/25 17:15:42 INFO hive.HiveImport: SLF4J: Found binding in [jarfile:/usr/hdp/3.14.0-315/hive/lib/log4j-slf4j-impl-2.10.0.jar/org/slf4j/impl/StaticLoggerBinder.class]
21/04/25 17:15:42 INFO hive.HiveImport: SLF4J: Found binding in [jarfile:/usr/hdp/3.14.0-315/hadoop/lib/slf4j-log4j12-1.7.25.jar/org/slf4j/impl/StaticLoggerBinder.class]
21/04/25 17:15:42 INFO hive.HiveImport: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
21/04/25 17:15:42 INFO hive.HiveImport: SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]
21/04/25 17:15:46 INFO hive.HiveImport: Connecting to jdbc:hive2://bgdataanalytics2-worker-shdpt-v31-1-0.novalocal:2181,bgdataanalytics2-worker-shdpt-v31-1-0.novalocal:2181,bgdataanalytics2-worker-shdpt-v31-1-0.novalocal:2181/default;password=hive;serviceDiscoveryMode=zooKeeper;user=hive;zookeeperNameSpace=hiveserver2
21/04/25 17:15:46 INFO hive.HiveImport: 21/04/25 17:15:46 [main-SendThread(bgdataanalytics2-worker-shdpt-v31-1-0.novalocal:2181)] WARN zookeeper.ClientCnxn: Session 0x0 for server null, unexpected error, closing connection and attempting reconnect
21/04/25 17:15:46 INFO hive.HiveImport: java.net.ConnectException: Connection refused
21/04/25 17:15:46 INFO hive.HiveImport: at sun.nio.ch.SocketChannelImpl.finishConnect(Native Method) ~[?:1.8.0_91]
21/04/25 17:15:46 INFO hive.HiveImport: at sun.nio.ch.SocketChannelImpl.finishConnect(SocketChannelImpl.java:719) ~[?:1.8.0_91]
21/04/25 17:15:46 INFO hive.HiveImport: at org.apache.zookeeper.client.doRawRequest(ClientCnxn$Krb5ReqHandler.java:261) ~[zookeeper-3.4.6.3-1.4.0-315.jar:3.4.6-315-1]
21/04/25 17:15:46 INFO hive.HiveImport: at org.apache.zookeeper.ClientCnxn$SendThread.run(ClientCnxn.java:1119) [zookeeper-3.4.6.3-1.4.0-315.jar:3.4.6-315-1]
21/04/25 17:15:47 INFO hive.HiveImport: 21/04/25 17:15:47 [main]: INFO jdbc.HiveConnection: Connected to bgdataanalytics2-head-shdpt-v31-1-0.novalocal:10000
21/04/25 17:15:47 INFO hive.HiveImport: Connected to Apache Hive (version 3.14.0-315)
21/04/25 17:15:47 INFO hive.HiveImport: Driver: Hive JDBC (version 3.14.0-315)
21/04/25 17:15:47 INFO hive.HiveImport: Transaction isolation: TRANSACTION_REPEATABLE_READ
21/04/25 17:15:47 INFO hive.HiveImport: 0 jdbc:hive2://bgdataanalytics2-worker-shdptv CREATE TABLE IF NOT EXISTS student3_16_test ('id' INT, str STRING) COMMENT 'Imported by sqoop on 2021/04/25 17:15:46'
21/04/25 17:15:47 INFO hive.HiveImport: ATED BY '0001' LINES TERMINATED BY '\0012' STORED AS TEXTFILE;
21/04/25 17:15:47 INFO hive.HiveImport: INFO : Compiling command(queryId=20210425171547_c667f56b-f1cb-4704-a3ef-c466a9852f64): CREATE TABLE IF NOT EXISTS student3_16_test ('id' INT, str STRING) COMMENT 'Imported by sqoop on 2021/04/25 17:15:46' ROW FORMAT DELIMITED FIELDS TERMINATED BY '\0012' LINES TERMINATED BY '\0012' STORED AS TEXTFILE
21/04/25 17:15:47 INFO hive.HiveImport: INFO : Semantic Analysis Completed (retrial = false)
21/04/25 17:15:47 INFO hive.HiveImport: INFO : Returning Hive schema: Schema(fields=Schemas:null, properties=null)
21/04/25 17:15:47 INFO hive.HiveImport: INFO : Completed compiling command(queryId=20210425171547_c667f56b-f1cb-4704-a3ef-c466a9852f64): Time taken: 0.022 seconds
21/04/25 17:15:47 INFO hive.HiveImport: INFO : Executing command(queryId=20210425171547_c667f56b-f1cb-4704-a3ef-c466a9852f64): CREATE TABLE IF NOT EXISTS student3_16_test ('id' INT, str STRING) COMMENT 'Imported by sqoop on 2021/04/25 17:15:40' ROW FORMAT DELIMITED FIELDS TERMINATED BY '\0012' LINES TERMINATED BY '\0012' STORED AS TEXTFILE
21/04/25 17:15:47 INFO hive.HiveImport: INFO : Starting task [Stage=0:DML] in serial mode
21/04/25 17:15:47 INFO hive.HiveImport: INFO : Completed executing command(queryId=20210425171547_f152ad90-24f1-4d90-9c73-b2ad02a470ab): Time taken: 0.065 seconds
21/04/25 17:15:47 INFO hive.HiveImport: INFO : OK
21/04/25 17:15:47 INFO hive.HiveImport: No rows affected (0.283 seconds)
21/04/25 17:15:47 INFO hive.HiveImport: 0 jdbc:hive2://bgdataanalytics2-worker-shdptv LOAD DATA INPATH 'hdfs://bgdataanalytics2-head-shdpt-v31-1-0.novalocal:8020/user/student3_16/test' INTO TABLE student3_16_test
21/04/25 17:15:48 INFO hive.HiveImport: INFO : Compiling command(queryId=20210425171547_f152ad90-24f1-4d90-9c73-b2ad02a470ab): LOAD DATA INPATH 'hdfs://bgdataanalytics2-head-shdpt-v31-1-0.novalocal:8020/user/student3_16/test' INTO TABLE student3_16_test
21/04/25 17:15:48 INFO hive.HiveImport: INFO : Semantic Analysis Completed (retrial = false)
21/04/25 17:15:48 INFO hive.HiveImport: INFO : Returning Hive schema: Schema(fields=Schemas:null, properties=null)
21/04/25 17:15:48 INFO hive.HiveImport: INFO : Completed compiling command(queryId=20210425171547_f152ad90-24f1-4d90-9c73-b2ad02a470ab): Time taken: 0.04 seconds
21/04/25 17:15:48 INFO hive.HiveImport: INFO : Executing command(queryId=20210425171547_f152ad90-24f1-4d90-9c73-b2ad02a470ab): LOAD DATA INPATH 'hdfs://bgdataanalytics2-head-shdpt-v31-1-0.novalocal:8020/user/student3_16/test' INTO TABLE student3_16_test
21/04/25 17:15:48 INFO hive.HiveImport: INFO : Starting task [Stage=0:MOVE] in serial mode
21/04/25 17:15:48 INFO hive.HiveImport: INFO : Loading data to table student3_16_test from hdfs://bgdataanalytics2-head-shdpt-v31-1-0.novalocal:8020/user/student3_16/test
21/04/25 17:15:48 INFO hive.HiveImport: INFO : Starting task [Stage=1:TS] in serial mode
21/04/25 17:15:48 INFO hive.HiveImport: INFO : Completed executing command(queryId=20210425171547_f152ad90-24f1-4d90-9c73-b2ad02a470ab): Time taken: 0.201 seconds
21/04/25 17:15:48 INFO hive.HiveImport: INFO : OK
21/04/25 17:15:48 INFO hive.HiveImport: No rows affected (0.277 seconds)
21/04/25 17:15:48 INFO hive.HiveImport: 0 jdbc:hive2://bgdataanalytics2-worker-shdptv
21/04/25 17:15:48 INFO hive.HiveImport: 0 jdbc:hive2://bgdataanalytics2-worker-shdptv Closing: 0 jdbc:hive2://bgdataanalytics2-worker-shdpt-v31-1-0.novalocal:2181,bgdataanalytics2-worker-shdpt-v31-1-0.novalocal:2181,bgdataanalytics2-worker-shdpt-v31-1-0.novalocal:2181/default;password=hive;serviceDiscoveryMode=zooKeeper;user=hive;zookeeperNameSpace=hiveserver2
21/04/25 17:15:48 INFO hive.HiveImport: Hive import complete.
21/04/25 17:15:48 INFO hive.HiveClientConnector: Export directory is not empty, keeping it.
21/04/25 17:15:48 INFO tool.ReportTool: Publishing HiveCat export job data to Listeners for table test
student3_16_bkgdataanalytics2-head-shdpt-v31-1-0-j5

```

Захожу в hdfs dfs и проверяю содержимое папки test. Наличие файла \_SUCCESS говорит об успешном импорте БД из PSQL в Hive. Т.к. в таблице БД PSQL было 2 записи (2 id), то соответственно и 2 файла part-m-

```
[student3_16@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -ls test
Found 3 items
-rw-r--r--  2 student3_16 student3_16      0 2021-04-25 17:10 test/_SUCCESS
-rw-r--r--  2 student3_16 student3_16      9 2021-04-25 17:10 test/part-m-00000
-rw-r--r--  2 student3_16 student3_16     10 2021-04-25 17:10 test/part-m-00001
[student3_16@bigdataanalytics2-head-shdpt-v31-1-0 ~]$
```



Захожу Hive в БД student3\_16 и проверяю содержимое таблицы Test

```
Beeline version 3.1.0.3.1.4.0-315 by Apache Hive
0: jdbc:hive2://bigdataanalytics2-worker-shdp> use student3_16;
INFO : Compiling command(queryId=hive_20210425171934_d116351a-0d7a-466a-a40a-55cd2a68b900): use student3_16
INFO : Semantic Analysis Completed (retrial = false)
INFO : Returning Hive schema: Schema(fieldsSchemas:null, properties:null)
INFO : Completed compiling command(queryId=hive_20210425171934_d116351a-0d7a-466a-a40a-55cd2a68b900); Time taken: 0.005 seconds
INFO : Executing command(queryId=hive_20210425171934_d116351a-0d7a-466a-a40a-55cd2a68b900): use student3_16
INFO : Starting task [Stage-0:DOL] in serial mode
INFO : Completed executing command(queryId=hive_20210425171934_d116351a-0d7a-466a-a40a-55cd2a68b900); Time taken: 0.009 seconds
INFO : OK
No rows affected (0.158 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp> select * from test limit 10;
INFO : Compiling command(queryId=hive_20210425171950_45242c97-1065-426f-9cfa-f634b4ddb84a): select * from test limit 10
INFO : Semantic Analysis Completed (retrial = false)
INFO : Returning Hive schema: Schema(fieldsSchemas:[FieldSchema(name:test.id, type:int, comment:null), FieldSchema(name:test.str, type:string, comment:null)], properties:null)
INFO : Completed compiling command(queryId=hive_20210425171950_45242c97-1065-426f-9cfa-f634b4ddb84a); Time taken: 0.105 seconds
INFO : Executing command(queryId=hive_20210425171950_45242c97-1065-426f-9cfa-f634b4ddb84a): select * from test limit 10
INFO : Completed executing command(queryId=hive_20210425171950_45242c97-1065-426f-9cfa-f634b4ddb84a); Time taken: 0.001 seconds
INFO : OK
+-----+-----+
| test.id | test.str |
+-----+-----+
| 1       | myText   |
| 2       | myText2  |
+-----+-----+
2 rows selected (0.207 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp> █
```

## 2. FLUME

Араше Flume – распределенная и высоконадежная система для эффективного сбора, агрегации и сохранения больших объемов логов из множества различных источников в централизованное хранилище данных

Подготавливаю скрипт heartbeat.sh в директории flm и файл конфигурации flm.conf

```
[student3_16@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ cd flm
[student3_16@bigdataanalytics2-head-shdpt-v31-1-0 flm]$ vim heartbeat.sh
[student3_16@bigdataanalytics2-head-shdpt-v31-1-0 flm]$ chmod a+x heartbeat.sh
[student3_16@bigdataanalytics2-head-shdpt-v31-1-0 flm]$ ./heartbeat.sh
./heartbeat.sh: line 12: syntax error: unexpected end of file
[student3_16@bigdataanalytics2-head-shdpt-v31-1-0 flm]$ vim heartbeat.sh
[student3_16@bigdataanalytics2-head-shdpt-v31-1-0 flm]$ ./heartbeat.sh
./heartbeat.sh: line 12: syntax error: unexpected end of file
[student3_16@bigdataanalytics2-head-shdpt-v31-1-0 flm]$ vim heartbeat.sh
[student3_16@bigdataanalytics2-head-shdpt-v31-1-0 flm]$ ./heartbeat.sh
./heartbeat.sh: line 11: syntax error: unexpected end of file
[student3_16@bigdataanalytics2-head-shdpt-v31-1-0 flm]$ vim heartbeat.sh
[student3_16@bigdataanalytics2-head-shdpt-v31-1-0 flm]$ ./heartbeat.sh
I live for 0 seconds;2021-04-25:17.47.30;2021-04-25:17.47.30;I did it 1
0m0.002s 0m0.002s
0m0.007s 0m0.007s
I live for 10 seconds;2021-04-25:17.47.30;2021-04-25:17.47.30;I did it 2
0m0.003s 0m0.003s
0m0.014s 0m0.014s
^C[student3_16@bigdataanalytics2-head-shdpt-v31-1-0 flm]$ █
```

Выполняю команду импорта с помощью flume с использованием скрипта и конфигурационного файла, созданных на предыдущем этапе

[illegible]

## Проверяю наличие созданных логов

```
[student3_16@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -ls flm
Found 5 items
-rw-r--r-- 2 student3_16 student3_16 784 2021-04-25 18:00 flm/FlumeData.1619373600662.log
-rw-r--r-- 2 student3_16 student3_16 770 2021-04-25 18:01 flm/FlumeData.1619373600663.log
-rw-r--r-- 2 student3_16 student3_16 771 2021-04-25 18:01 flm/FlumeData.1619373600664.log
-rw-r--r-- 2 student3_16 student3_16 833 2021-04-25 18:02 flm/FlumeData.1619373696895.log
-rw-r--r-- 2 student3_16 student3_16 397 2021-04-25 18:02 flm/FlumeData.1619373731754.log
[student3_16@bigdataanalytics2-head-shdpt-v31-1-0 ~]$
```

## Создаю таблицу в hive

```
create external table flm_logs (first string, second string, third string,
fourth string, fifth string)
ROW FORMAT DELIMITED
FIELDS TERMINATED BY ';'
STORED AS INPUTFORMAT 'org.apache.hadoop.mapred.SequenceFileInputFormat'
OUTPUTFORMAT 'org.apache.hadoop.hive ql.io.HiveSequenceFileOutputFormat'
location '/user/student3_16/flm'
;
```

```

beeline version 3.1.0.3.1.4.0-315 by Apache Hive
0: jdbc:hive2://bigdataanalytics2-worker-shdp: use student3_16;
INFO : Compiling command(queryId=hive_20210426151822_8a4485c2-050e-4af5-af2a-6767de6ffa6c): use student3_16
INFO : Semantic analysis Completed (retrial = false)
INFO : Returning Hive schema: Schema(fieldsSchemas:null, properties:null)
INFO : Completed compiling command(queryId=hive_20210426151822_8a4485c2-050e-4af5-af2a-6767de6ffa6c): Time taken: 0.007 seconds
INFO : Executing command(queryId=hive_20210426151822_8a4485c2-050e-4af5-af2a-6767de6ffa6c): use student3_16
INFO : Starting task [Stage-0:DDL] in serial mode
INFO : Completed executing command(queryId=hive_20210426151822_8a4485c2-050e-4af5-af2a-6767de6ffa6c): Time taken: 0.012 seconds
INFO : OK
No rows affected (0.217 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp: create external table flm_logs (first string, second string, third string, fourth string, fifth string)
. . . . .> ROW FORMAT DELIMITED
. . . . .> FIELDS TERMINATED BY ','
. . . . .> STORED AS INPUTFORMAT 'org.apache.hadoop.mapred.SequenceFileInputFormat'
. . . . .> OUTPUTFORMAT 'org.apache.hadoop.hive ql.io.HiveSequenceFileOutputFormat'
. . . . .> location '/user/student3_16/flm'
INFO : Compiling command(queryId=hive_20210426151834_462a7d05-04cd-413c-baa3-75dfeaf7b1247): create external table flm_logs (first string, second string, third string, fourth string, fifth string)
ROW FORMAT DELIMITED
FIELDS TERMINATED BY ','
STORED AS INPUTFORMAT 'org.apache.hadoop.mapred.SequenceFileInputFormat'
OUTPUTFORMAT 'org.apache.hadoop.hive ql.io.HiveSequenceFileOutputFormat'
location '/user/student3_16/flm'
INFO : Semantic analysis Completed (retrial = false)
INFO : Returning Hive schema: Schema(fieldsSchemas:null, properties:null)
INFO : Completed compiling command(queryId=hive_20210426151834_462a7d05-04cd-413c-baa3-75dfeaf7b1247): Time taken: 0.021 seconds
INFO : Executing command(queryId=hive_20210426151834_462a7d05-04cd-413c-baa3-75dfeaf7b1247): create external table flm_logs (first string, second string, third string, fourth string, fifth string)
ROW FORMAT DELIMITED
FIELDS TERMINATED BY ','
STORED AS INPUTFORMAT 'org.apache.hadoop.mapred.SequenceFileInputFormat'
OUTPUTFORMAT 'org.apache.hadoop.hive ql.io.HiveSequenceFileOutputFormat'
location '/user/student3_16/flm'
INFO : Starting task [Stage-0:DDL] in serial mode
INFO : Completed executing command(queryId=hive_20210426151834_462a7d05-04cd-413c-baa3-75dfeaf7b1247): Time taken: 0.065 seconds
INFO : OK
No rows affected (0.131 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp:

```

## Проверяю содержимое таблицы

```
select * from flm logs;
```

```

0: jdbc:hive2://bigdataanalytics2-worker-shdp> select * from flm_logs;
INFO : Compiling command(queryId=hive_20210426151921_4c7c0d67-9f8f-4559-8379-d794e408eb44): select * from flm_logs
INFO : Semantic Analysis Completed (retrial = false)
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:flm_logs.first, type:string, comment:null), FieldSchema(name:flm_logs.second, type:string, comment:null), FieldSchema(name:flm_logs.third, type:string, comment:null), FieldSchema(name:flm_logs.fourth, type:string, comment:null), FieldSchema(name:flm_logs.fifth, type:string, comment:null)], tableName:flm_logs)
INFO : Completed compiling command(queryId=hive_20210426151921_4c7c0d67-9f8f-4559-8379-d794e408eb44); Time taken: 0.152 seconds
INFO : Executing command(queryId=hive_20210426151921_4c7c0d67-9f8f-4559-8379-d794e408eb44): select * from flm_logs
INFO : Completed executing command(queryId=hive_20210426151921_4c7c0d67-9f8f-4559-8379-d794e408eb44); Time taken: 0.001 seconds
INFO : OK

```

flm_logs.first	flm_logs.second	flm_logs.third	flm_logs.fourth	flm_logs.fifth
I live for 0 seconds	2021-04-25:17.59.56	2021-04-25:17.59.56	I did it 1	NULL
0m0.002s 0m0.007s	NULL	NULL	NULL	NULL
0m0.011s 0m0.004s	NULL	NULL	NULL	NULL
I live for 10 seconds	2021-04-25:17.59.56	2021-04-25:17.59.56	I did it 2	NULL
0m0.005s 0m0.007s	NULL	NULL	NULL	NULL
0m0.015s 0m0.014s	NULL	NULL	NULL	NULL
I live for 20 seconds	2021-04-25:17.59.56	2021-04-25:17.59.56	I did it 3	NULL
0m0.006s 0m0.008s	NULL	NULL	NULL	NULL
0m0.024s 0m0.019s	NULL	NULL	NULL	NULL
I live for 30 seconds	2021-04-25:17.59.56	2021-04-25:17.59.56	I did it 4	NULL
0m0.006s 0m0.010s	NULL	NULL	NULL	NULL
0m0.028s 0m0.024s	NULL	NULL	NULL	NULL
I live for 40 seconds	2021-04-25:17.59.56	2021-04-25:17.59.56	I did it 5	NULL
0m0.007s 0m0.011s	NULL	NULL	NULL	NULL
0m0.033s 0m0.033s	NULL	NULL	NULL	NULL
I live for 50 seconds	2021-04-25:17.59.56	2021-04-25:17.59.56	I did it 6	NULL
0m0.007s 0m0.013s	NULL	NULL	NULL	NULL
0m0.037s 0m0.041s	NULL	NULL	NULL	NULL
I live for 60 seconds	2021-04-25:17.59.56	2021-04-25:17.59.56	I did it 7	NULL
0m0.009s 0m0.014s	NULL	NULL	NULL	NULL
0m0.043s 0m0.049s	NULL	NULL	NULL	NULL
I live for 70 seconds	2021-04-25:17.59.56	2021-04-25:17.59.56	I did it 8	NULL
0m0.009s 0m0.016s	NULL	NULL	NULL	NULL
0m0.050s 0m0.054s	NULL	NULL	NULL	NULL
I live for 80 seconds	2021-04-25:17.59.56	2021-04-25:17.59.56	I did it 9	NULL
0m0.010s 0m0.017s	NULL	NULL	NULL	NULL
0m0.055s 0m0.063s	NULL	NULL	NULL	NULL
I live for 90 seconds	2021-04-25:17.59.56	2021-04-25:17.59.56	I did it 10	NULL
0m0.010s 0m0.018s	NULL	NULL	NULL	NULL
0m0.060s 0m0.068s	NULL	NULL	NULL	NULL
I live for 100 seconds	2021-04-25:17.59.56	2021-04-25:17.59.56	I did it 11	NULL
0m0.012s 0m0.018s	NULL	NULL	NULL	NULL
0m0.065s 0m0.075s	NULL	NULL	NULL	NULL
I live for 110 seconds	2021-04-25:17.59.56	2021-04-25:17.59.56	I did it 12	NULL
0m0.013s 0m0.020s	NULL	NULL	NULL	NULL
0m0.070s 0m0.082s	NULL	NULL	NULL	NULL
I live for 120 seconds	2021-04-25:17.59.56	2021-04-25:17.59.56	I did it 13	NULL
0m0.013s 0m0.022s	NULL	NULL	NULL	NULL
0m0.079s 0m0.086s	NULL	NULL	NULL	NULL
I live for 130 seconds	2021-04-25:17.59.56	2021-04-25:17.59.56	I did it 14	NULL
0m0.013s 0m0.024s	NULL	NULL	NULL	NULL
0m0.089s 0m0.088s	NULL	NULL	NULL	NULL
I live for 140 seconds	2021-04-25:17.59.56	2021-04-25:17.59.56	I did it 15	NULL
0m0.013s 0m0.027s	NULL	NULL	NULL	NULL
0m0.093s 0m0.099s	NULL	NULL	NULL	NULL

45 rows selected (0.318 seconds)