

**TASK-2 : copy only person-id = 8 from db to hdfs and export the data from hds to a new table**

**STEP-1 :** Check the person table data exist or not.

```
mysql> select * from person;
```

person_id	person_name	person_address	person_mobile
1	satya	Tavarekera	8989890000
2	priya	BTM	9898002323
3	muna	Agra	7878999900
4	kuna	Silk Board	7676889900
5	swapna	HSR	8787990012
6	prakash	HSR	9889007878
7	jhuna	HSR	9876500123

```
7 rows in set (0.00 sec)
```

**STEP-2 :** create another table from which we can copy data from hdfs

```
mysql> create table person2 like person;
Query OK, 0 rows affected (0.03 sec)

mysql> select * from person2;
Empty set (0.00 sec)
```

**STEP-3 :** check the hdfs, where the person data copied from the db and make sure there is no 8<sup>th</sup> record present in it.

```
[acadgild@localhost ~]$ hadoop fs -cat /satya/sqoop/SQ00P0UT/p*
18/11/08 19:02:03 WARN util.NativeCodeLoader: Unable to load nat
Tavarekera,1,8989890000,satya
BTM,2,9898002323,priya
Agra,3,7878999900,muna
Silk Board,4,7676889900,kuna
HSR,5,8787990012,swapna
HSR,6,9889007878,prakash
HSR,7,9876500123,jhuna
```

**STEP-4 :** Now insert the 8<sup>th</sup> record to the existing table **Person**

```
mysql> insert into person values(8,'Rahul','BTM',8585767700);
Query OK, 1 row affected (0.00 sec)
```

```
mysql> select * from person;
```

person_id	person_name	person_address	person_mobile
1	satya	Tavarekera	8989890000
2	priya	BTM	9898002323
3	muna	Agra	7878999900
4	kuna	Silk Board	7676889900
5	swapna	HSR	8787990012
6	prakash	HSR	9889007878
7	jhuna	HSR	9876500123
8	Rahul	BTM	8585767700

```
8 rows in set (0.00 sec)
```

STEP-5 : Now we have additional record with person\_id as 8 which need to be imported in the Hadoop file. For that we will use the **sqoop incremental import**

```
[acadvild@localhost ~]$ sqoop import --connect jdbc:mysql://localhost/emp --username root --password Root@123 --table person --target-dir /satya/sqoop/SQOOP/OUT/ -m 1 --incremental append --check-column person_id --last-value 7
```

```
[acadvild@localhost ~]$ sqoop import --connect jdbc:mysql://localhost/emp --username root --password Root@123 --table person --target-dir /satya/sqoop/SQOOP/OUT/ -m 1 --incremental append --check-column person_id --last-value 7
Warning: /home/acadvild/install/sqoop/sqoop-1.4.6/bin_hadoop-2.0.4-alpha../hcatalog does not exist! HCatalog jobs will fail.
Please set $HCAT_HOME to the root of your HCatalog installation.
Warning: /home/acadvild/install/sqoop/sqoop-1.4.6/bin_hadoop-2.0.4-alpha../accumulo does not exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
18/11/08 18:20:29 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6
18/11/08 18:20:29 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.
18/11/08 18:20:29 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
18/11/08 18:20:29 INFO tool.CodeGenTool: Beginning code generation
Thu Nov 08 18:20:29 IST 2018 WARN: Establishing SSL connection without server's identity verification is not recommended. According to MySQL 5.5.45+, 5.6
.26+ and 5.7.6+ requirements SSL connection must be established by default if explicit option isn't set. For compliance with existing applications not us
ing SSL the verifyServerCertificate property is set to 'false'. You need either to explicitly disable SSL by setting useSSL=false, or set useSSL=true and
provide truststore for server certificate verification.
18/11/08 18:20:31 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `person` AS t LIMIT 1
18/11/08 18:20:31 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `person` AS t LIMIT 1
18/11/08 18:20:31 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /home/acadvild/install/hadoop/hadoop-2.6.5
Note: /tmp/sqoop-acadvild/compile/cb5b3580a618d1d0b253e835446645ce/person.java uses or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
18/11/08 18:20:36 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-acadvild/compile/cb5b3580a618d1d0b253e835446645ce/person.jar
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/acadvild/install/hadoop/hadoop-2.6.5/share/hadoop/common/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLogger
Binder.class]
SLF4J: Found binding in [jar:file:/home/acadvild/install/hbase/hbase-1.2.6/lib/slf4j-log4j12-1.7.5.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]
18/11/08 18:20:36 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
18/11/08 18:20:38 INFO tool.ImportTool: Maximal id query for free form incremental import: SELECT MAX(`person_id`) FROM `person`
18/11/08 18:20:38 INFO tool.ImportTool: Incremental import based on column `person_id`
18/11/08 18:20:38 INFO tool.ImportTool: Lower bound value: 7
18/11/08 18:20:38 INFO tool.ImportTool: Upper bound value: 8
18/11/08 18:20:38 WARN manager.MySQLManager: It looks like you are importing from mysql.
18/11/08 18:20:38 WARN manager.MySQLManager: This transfer can be faster! Use the --direct
18/11/08 18:20:38 WARN manager.MySQLManager: option to exercise a MySQL-specific fast path.
18/11/08 18:20:38 INFO manager.MySQLManager: Setting zero DATETIME behavior to convertToNull (mysql)
18/11/08 18:20:38 INFO mapreduce.ImportJobBase: Beginning import of person
18/11/08 18:20:38 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar
18/11/08 18:20:38 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps
18/11/08 18:20:38 INFO client.RMPProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
```

```
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)=10513
  Total time spent by all reduces in occupied slots (ms)=0
  Total time spent by all map tasks (ms)=10513
  Total vcore-milliseconds taken by all map tasks=10513
  Total megabyte-milliseconds taken by all map tasks=10765312
Map-Reduce Framework
  Map input records=1
  Map output records=1
  Input split bytes=87
  Spilled Records=0
  Failed Shuffles=0
  Merged Map outputs=0
  GC time elapsed (ms)=177
  CPU time spent (ms)=2160
  Physical memory (bytes) snapshot=108212224
  Virtual memory (bytes) snapshot=2063437824
  Total committed heap usage (bytes)=32571392
File Input Format Counters
  Bytes Read=0
File Output Format Counters
  Bytes Written=23
18/11/08 18:21:20 INFO mapreduce.ImportJobBase: Transferred 23 bytes in 42.5579 seconds (0.5404 bytes/sec)
18/11/08 18:21:20 INFO mapreduce.ImportJobBase: Retrieved 1 records.
18/11/08 18:21:20 INFO util.AppendUtils: Appending to directory SQ00POUT
18/11/08 18:21:20 INFO util.AppendUtils: Using found partition 1
18/11/08 18:21:20 INFO tool.ImportTool: Incremental import complete! To run another incremental import of all data following this import, supply the following arguments:
18/11/08 18:21:20 INFO tool.ImportTool: --incremental append
18/11/08 18:21:20 INFO tool.ImportTool: --check-column person_id
18/11/08 18:21:20 INFO tool.ImportTool: --last-value 8
18/11/08 18:21:20 INFO tool.ImportTool: (Consider saving this with 'sqoop job --create')
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$
```

STEP-6 : Now we will check once again the data using cat command in the same file

```
[acadgild@localhost ~]$ hadoop fs -cat /satya/sqoop/SQ00POUT/p*
18/11/08 19:07:24 WARN util.NativeCodeLoader: Unable to load native
Tavarekera,1,8989890000,satya
BTM,2,9898002323,priya
Agra,3,7878999900,muna
Silk Board,4,7676889900,kuna
HSR,5,8787990012,swapna
HSR,6,9889007878,prakash
HSR,7,9876500123,jhuna
BTM,8,8585767700,Rahul
[acadgild@localhost ~]$
```

STEP-7:

Now we will export data from hadoop filesystem to DB

```
[acadgild@localhost ~]$ sqoop export --connect jdbc:mysql://localhost:3306/emp --username root --password Root@123 --table person2 --export-dir /satya/sqoop/SQ00POUT -m 1
Warning: /home/acadgild/install/sqoop/sqoop-1.4.6-bin-hadoop2.0.4-alpha1/lib/catalog does not exist! HCatalog jobs will fail.
```

```

[acacgild@localhost ~]$ sqoop export --connect jdbc:mysql://localhost:3306/emp --username root --password Root@123 --table person2 --export-dir /satya/sq
oop/SQOOPOUT -m 1
Warning: /home/acacgild/install/sqoop/sqoop-1.4.6.bin__hadoop-2.0.4-alpha/./hcatalog does not exist! Hcatalog jobs will fail.
Please set $HCAT_HOME to the root of your HCatalog installation.
Warning: /home/acacgild/install/sqoop/sqoop-1.4.6.bin__hadoop-2.0.4-alpha/./accumulo does not exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
18/11/08 19:10:34 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6
18/11/08 19:10:34 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.
18/11/08 19:10:35 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
18/11/08 19:10:35 INFO tool.CodeGenTool: Beginning code generation
Thu Nov 08 19:10:35 IST 2018 WARN: Establishing SSL connection without server's identity verification is not recommended. According to MySQL 5.5.45+, 5.6
.26+ and 5.7.6+ requirements SSL connection must be established by default if explicit option isn't set. For compliance with existing applications not us
ing SSL the verifyServerCertificate property is set to 'false'. You need either to explicitly disable SSL by setting useSSL=false, or set useSSL=true and
provide truststore for server certificate verification.
18/11/08 19:10:36 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM 'person2' AS t LIMIT 1
18/11/08 19:10:36 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM 'person2' AS t LIMIT 1
18/11/08 19:10:36 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /home/acacgild/install/hadoop/hadoop-2.6.5
Note: /tmp/sqoop-acacgild/compile/f49838d76263eb9ba494f2b2cc44f56e/person2.java uses or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
18/11/08 19:10:41 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-acacgild/compile/f49838d76263eb9ba494f2b2cc44f56e/person2.jar
18/11/08 19:10:41 INFO mapreduce.ExportJobBase: Beginning export of person2
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/acacgild/install/hadoop/hadoop-2.6.5/share/hadoop/common/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLogger
Binder.class]
SLF4J: Found binding in [jar:file:/home/acacgild/install/hbase/hbase-1.2.6/lib/slf4j-log4j12-1.7.5.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]
18/11/08 19:10:41 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
18/11/08 19:10:41 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar
18/11/08 19:10:43 INFO Configuration.deprecation: mapred.reduce.tasks.speculative.execution is deprecated. Instead, use mapreduce.reduce.speculative
18/11/08 19:10:43 INFO Configuration.deprecation: mapred.map.tasks.speculative.execution is deprecated. Instead, use mapreduce.map.speculative
18/11/08 19:10:43 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps
18/11/08 19:10:43 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps

```

```

18/11/08 19:11:00 INFO mapreduce.Job: map 0% reduce 0%
18/11/08 19:11:11 INFO mapreduce.Job: map 100% reduce 0%
18/11/08 19:11:12 INFO mapreduce.Job: Job job_1541677863265_0006 completed successfully
18/11/08 19:11:13 INFO mapreduce.Job: Counters: 30
  File System Counters
    FILE: Number of bytes read=0
    FILE: Number of bytes written=127575
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=414
    HDFS: Number of bytes written=0
    HDFS: Number of read operations=7
    HDFS: Number of large read operations=0
    HDFS: Number of write operations=0
  Job Counters
    Launched map tasks=1
    Data-local map tasks=1
    Total time spent by all maps in occupied slots (ms)=7888
    Total time spent by all reduces in occupied slots (ms)=0
    Total time spent by all map tasks (ms)=7888
    Total vcore-milliseconds taken by all map tasks=7888
    Total megabyte-milliseconds taken by all map tasks=8077312
  Map-Reduce Framework
    Map input records=8
    Map output records=8
    Input split bytes=208
    Spilled Records=0
    Failed Shuffles=0
    Merged Map outputs=0
    GC time elapsed (ms)=82
    CPU time spent (ms)=2260
    Physical memory (bytes) snapshot=107286528
    Virtual memory (bytes) snapshot=2061332480
    Total committed heap usage (bytes)=32571392
  File Input Format Counters
    Bytes Read=0
  File Output Format Counters
    Bytes Written=0
18/11/08 19:11:13 INFO mapreduce.ExportJobBase: Transferred 414 bytes in 29.7962 seconds (13.8944 bytes/sec)
18/11/08 19:11:13 INFO mapreduce.ExportJobBase: Exported 8 records.
You have new mail in /var/spool/mail/acacgild

```

STEP-8: Check the table person2 whether all the data copied or not

```
mysql> select * from person2;
```

person_id	person_name	person_address	person_mobile
1	satya	Tavarekera	8989890000
2	priya	BTM	9898002323
3	muna	Agra	7878999900
4	kuna	Silk Board	7676889900
5	swapna	HSR	8787990012
6	prakash	HSR	9889007878
7	jhuna	HSR	9876500123
8	Rahul	BTM	8585767700

```
8 rows in set (0.00 sec)
```