[root@quickstart ~]# sqoop import --connect jdbc:mysql://quickstart.cloudera/tron --username root -P --table employee3

Warning: /usr/lib/sqoop/../accumulo does not exist! Accumulo imports will fail.

Please set $ACCUMULO\_HOME to the root of your Accumulo installation.

22/05/07 05:06:26 INFO sqoop.Sqoop: Running Sqoop version: 1.4.5-cdh5.4.0

Enter password:

22/05/07 05:06:31 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.

22/05/07 05:06:31 INFO tool.CodeGenTool: Beginning code generation

22/05/07 05:06:32 INFO manager.SqlManager: Executing SQL statement: SELECT t.\* FROM `employee3` AS t LIMIT 1

22/05/07 05:06:32 INFO manager.SqlManager: Executing SQL statement: SELECT t.\* FROM `employee3` AS t LIMIT 1

22/05/07 05:06:32 INFO orm.CompilationManager: HADOOP\_MAPRED\_HOME is /usr/lib/hadoop-mapreduce

Note: /tmp/sqoop-root/compile/325e1358b5566ae55fcb090de96a0669/employee3.java uses or overrides a deprecated API.

Note: Recompile with -Xlint:deprecation for details.

22/05/07 05:06:34 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-root/compile/325e1358b5566ae55fcb090de96a0669/employee3.jar

22/05/07 05:06:34 WARN manager.MySQLManager: It looks like you are importing from mysql.

22/05/07 05:06:34 WARN manager.MySQLManager: This transfer can be faster! Use the --direct

22/05/07 05:06:34 WARN manager.MySQLManager: option to exercise a MySQL-specific fast path.

22/05/07 05:06:34 INFO manager.MySQLManager: Setting zero DATETIME behavior to convertToNull (mysql)

22/05/07 05:06:34 INFO mapreduce.ImportJobBase: Beginning import of employee3

22/05/07 05:06:34 INFO Configuration.deprecation: mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address

22/05/07 05:06:34 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar

22/05/07 05:06:35 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps

22/05/07 05:06:35 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032

22/05/07 05:06:37 INFO db.DBInputFormat: Using read commited transaction isolation

22/05/07 05:06:37 INFO db.DataDrivenDBInputFormat: BoundingValsQuery: SELECT MIN(`id`), MAX(`id`) FROM `employee3`

22/05/07 05:06:37 INFO mapreduce.JobSubmitter: number of splits:4

22/05/07 05:06:38 INFO mapreduce.JobSubmitter: Submitting tokens for job: job\_1651922031586\_0002

22/05/07 05:06:38 INFO impl.YarnClientImpl: Submitted application application\_1651922031586\_0002

22/05/07 05:06:38 INFO mapreduce.Job: The url to track the job: http://quickstart.cloudera:8088/proxy/application\_1651922031586\_0002/

22/05/07 05:06:38 INFO mapreduce.Job: Running job: job\_1651922031586\_0002

22/05/07 05:06:47 INFO mapreduce.Job: Job job\_1651922031586\_0002 running in uber mode : false

22/05/07 05:06:47 INFO mapreduce.Job: map 0% reduce 0%

22/05/07 05:07:12 INFO mapreduce.Job: map 25% reduce 0%

22/05/07 05:07:13 INFO mapreduce.Job: map 75% reduce 0%

22/05/07 05:07:14 INFO mapreduce.Job: map 100% reduce 0%

22/05/07 05:07:14 INFO mapreduce.Job: Job job\_1651922031586\_0002 completed successfully

22/05/07 05:07:15 INFO mapreduce.Job: Counters: 30

File System Counters

FILE: Number of bytes read=0

FILE: Number of bytes written=539648

FILE: Number of read operations=0

FILE: Number of large read operations=0

FILE: Number of write operations=0

HDFS: Number of bytes read=393

HDFS: Number of bytes written=30

HDFS: Number of read operations=16

HDFS: Number of large read operations=0

HDFS: Number of write operations=8

Job Counters

Launched map tasks=4

Other local map tasks=4

Total time spent by all maps in occupied slots (ms)=87083

Total time spent by all reduces in occupied slots (ms)=0

Total time spent by all map tasks (ms)=87083

Total vcore-seconds taken by all map tasks=87083

Total megabyte-seconds taken by all map tasks=89172992

Map-Reduce Framework

Map input records=4

Map output records=4

Input split bytes=393

Spilled Records=0

Failed Shuffles=0

Merged Map outputs=0

GC time elapsed (ms)=1523

CPU time spent (ms)=8140

Physical memory (bytes) snapshot=660996096

Virtual memory (bytes) snapshot=6263934976

Total committed heap usage (bytes)=569901056

File Input Format Counters

Bytes Read=0

File Output Format Counters

Bytes Written=30

22/05/07 05:07:15 INFO mapreduce.ImportJobBase: Transferred 30 bytes in 39.9972 seconds (0.7501 bytes/sec)

22/05/07 05:07:15 INFO mapreduce.ImportJobBase: Retrieved 4 records.

[root@quickstart ~]# hdfs dfs -ls

Found 3 items

drwxr-xr-x - root supergroup 0 2022-05-03 04:37 employee

drwxr-xr-x - root supergroup 0 2022-05-07 05:02 employee2

drwxr-xr-x - root supergroup 0 2022-05-07 05:07 employee3

[root@quickstart ~]# hdfs dfs -ls employee2

Found 2 items

-rw-r--r-- 1 root supergroup 0 2022-05-07 05:02 employee2/\_SUCCESS

-rw-r--r-- 1 root supergroup 6 2022-05-07 05:02 employee2/part-m-00000

[root@quickstart ~]# hdfs dfs -ls employee2/part-m-00000

-rw-r--r-- 1 root supergroup 6 2022-05-07 05:02 employee2/part-m-00000

[root@quickstart ~]# hdfs dfs -cat employee2/part-m-00000

2,ram

[root@quickstart ~]# hdfs dfs -cat employee2/part-m-0000\*

2,ram

[root@quickstart ~]# hdfs dfs -cat employee2/part-m-00000\*

2,ram

[root@quickstart ~]# hdfs dfs -cat employee3/part-m-00000\*

1,raj

[root@quickstart ~]# hdfs dfs -cat employee3/part-m-0000\*

1,raj

2,ram

3,parth

4,shubham

[root@quickstart ~]# hdfs dfs -cat employee3

cat: `employee3': Is a directory

[root@quickstart ~]# hdfs dfs -ls employee3

Found 5 items

-rw-r--r-- 1 root supergroup 0 2022-05-07 05:07 employee3/\_SUCCESS

-rw-r--r-- 1 root supergroup 6 2022-05-07 05:07 employee3/part-m-00000

-rw-r--r-- 1 root supergroup 6 2022-05-07 05:07 employee3/part-m-00001

-rw-r--r-- 1 root supergroup 8 2022-05-07 05:07 employee3/part-m-00002

-rw-r--r-- 1 root supergroup 10 2022-05-07 05:07 employee3/part-m-00003

[root@quickstart ~]# hdfs dfs -cat employee3/part-m-00001

2,ram

[root@quickstart ~]# hdfs dfs -cat employee3/part-m-00002

3,parth

[root@quickstart ~]# hdfs dfs -cat employee3/part-m-00003

4,shubham

[root@quickstart ~]# hdfs dfs -cat employee3/part-m-0000\*

1,raj

2,ram

3,parth

4,shubham

mysql> describe employee;

+--------+-------------+------+-----+---------+-------+

| Field | Type | Null | Key | Default | Extra |

+--------+-------------+------+-----+---------+-------+

| id | int(11) | NO | PRI | NULL | |

| name | varchar(10) | YES | | NULL | |

| salary | int(11) | YES | | NULL | |

+--------+-------------+------+-----+---------+-------+

3 rows in set (0.00 sec)

mysql> use tron;

Database changed

mysql> select \* from employee;

+----+----------+--------+

| id | name | salary |

+----+----------+--------+

| 1 | raj | 99999 |

| 2 | ram | 98988 |

| 3 | parth | 98568 |

| 4 | abhijeet | 95459 |

| 5 | pradip | 93259 |

| 6 | akash | 93259 |

| 7 | Tim | 94559 |

+----+----------+--------+

7 rows in set (0.00 sec)

mysql> show tables;

+----------------+

| Tables\_in\_tron |

+----------------+

| employee |

| employee2 |

| employee3 |

+----------------+

3 rows in set (0.00 sec)

mysql> select \* from employee2;

+----+------+

| id | name |

+----+------+

| 2 | ram |

+----+------+

1 row in set (0.00 sec)

mysql> select \* from employee3;

+----+---------+

| id | name |

+----+---------+

| 1 | raj |

| 2 | ram |

| 3 | parth |

| 4 | shubham |

+----+---------+

4 rows in set (0.00 sec)

-------------------------------------------------------------------------------------------------------------------------------------------------------------------

-------------------------------------------------------------------------------------------------------------------------------------------------------------------

--target-dir using

[cloudera@quickstart sqoop]$ hdfs dfs -mkdir abhi

[cloudera@quickstart sqoop]$ hdfs dfs -ls

Found 13 items

drwxr-xr-x - cloudera cloudera 0 2019-12-29 00:57 \_sqoop

drwxr-xr-x - cloudera cloudera 0 2022-05-07 05:49 abhi

drwxr-xr-x - cloudera cloudera 0 2019-06-26 01:29 bank

drwxr-xr-x - cloudera cloudera 0 2019-06-29 06:38 btm

drwxr-xr-x - cloudera cloudera 0 2022-05-04 05:12 cloudera

-rw-r--r-- 3 cloudera cloudera 275 2018-03-11 01:47 emp

drwxr-xr-x - root cloudera 0 2022-05-03 04:55 employee

drwxr-xr-x - cloudera cloudera 0 2018-03-02 02:56 input

drwxr-xr-x - cloudera cloudera 0 2017-03-02 18:13 spark

drwxr-xr-x - cloudera cloudera 0 2018-08-24 09:57 spark\_data

-rw-r--r-- 1 cloudera cloudera 36 2018-03-02 01:58 spark\_inp

drwxr-xr-x - cloudera cloudera 0 2018-07-15 02:03 sparklab

drwxr-xr-x - cloudera cloudera 0 2018-07-14 21:54 test

[cloudera@quickstart sqoop]$ pwd

/usr/lib/sqoop

[cloudera@quickstart sqoop]$ cd ../..

[cloudera@quickstart usr]$ cd ../..

[cloudera@quickstart /]$ cd ..

[cloudera@quickstart /]$ cd

[cloudera@quickstart ~]$ pwd

/home/cloudera

[cloudera@quickstart ~]$ hdfs dfs -mkdir abhi

^[[Amkdir: `abhi': File exists

[cloudera@quickstart ~]$ pwd

/home/cloudera

[cloudera@quickstart ~]$ hdfs dfs -ls

Found 13 items

drwxr-xr-x - cloudera cloudera 0 2019-12-29 00:57 \_sqoop

drwxr-xr-x - cloudera cloudera 0 2022-05-07 05:49 abhi

drwxr-xr-x - cloudera cloudera 0 2019-06-26 01:29 bank

drwxr-xr-x - cloudera cloudera 0 2019-06-29 06:38 btm

drwxr-xr-x - cloudera cloudera 0 2022-05-04 05:12 cloudera

-rw-r--r-- 3 cloudera cloudera 275 2018-03-11 01:47 emp

drwxr-xr-x - root cloudera 0 2022-05-03 04:55 employee

drwxr-xr-x - cloudera cloudera 0 2018-03-02 02:56 input

drwxr-xr-x - cloudera cloudera 0 2017-03-02 18:13 spark

drwxr-xr-x - cloudera cloudera 0 2018-08-24 09:57 spark\_data

-rw-r--r-- 1 cloudera cloudera 36 2018-03-02 01:58 spark\_inp

drwxr-xr-x - cloudera cloudera 0 2018-07-15 02:03 sparklab

drwxr-xr-x - cloudera cloudera 0 2018-07-14 21:54 test

[cloudera@quickstart ~]$ sqoop import --connect jdbc:mysql://quickstart.cloudera/tron --username root -P --table employee4 --target-dir /home/cloudera/abhi

Warning: /usr/lib/sqoop/../accumulo does not exist! Accumulo imports will fail.

Please set $ACCUMULO\_HOME to the root of your Accumulo installation.

22/05/07 05:51:21 INFO sqoop.Sqoop: Running Sqoop version: 1.4.5-cdh5.4.0

Enter password:

22/05/07 05:51:24 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.

22/05/07 05:51:24 INFO tool.CodeGenTool: Beginning code generation

22/05/07 05:51:24 INFO manager.SqlManager: Executing SQL statement: SELECT t.\* FROM `employee4` AS t LIMIT 1

22/05/07 05:51:24 INFO manager.SqlManager: Executing SQL statement: SELECT t.\* FROM `employee4` AS t LIMIT 1

22/05/07 05:51:24 INFO orm.CompilationManager: HADOOP\_MAPRED\_HOME is /usr/lib/hadoop-mapreduce

Note: /tmp/sqoop-cloudera/compile/2eb36afbfd072ca06f7dda84763867f6/employee4.java uses or overrides a deprecated API.

Note: Recompile with -Xlint:deprecation for details.

22/05/07 05:51:26 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-cloudera/compile/2eb36afbfd072ca06f7dda84763867f6/employee4.jar

22/05/07 05:51:26 WARN manager.MySQLManager: It looks like you are importing from mysql.

22/05/07 05:51:26 WARN manager.MySQLManager: This transfer can be faster! Use the --direct

22/05/07 05:51:26 WARN manager.MySQLManager: option to exercise a MySQL-specific fast path.

22/05/07 05:51:26 INFO manager.MySQLManager: Setting zero DATETIME behavior to convertToNull (mysql)

22/05/07 05:51:26 INFO mapreduce.ImportJobBase: Beginning import of employee4

22/05/07 05:51:26 INFO Configuration.deprecation: mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address

22/05/07 05:51:27 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar

22/05/07 05:51:28 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps

22/05/07 05:51:28 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032

22/05/07 05:51:30 INFO db.DBInputFormat: Using read commited transaction isolation

22/05/07 05:51:30 INFO db.DataDrivenDBInputFormat: BoundingValsQuery: SELECT MIN(`id`), MAX(`id`) FROM `employee4`

22/05/07 05:51:30 INFO mapreduce.JobSubmitter: number of splits:4

22/05/07 05:51:31 INFO mapreduce.JobSubmitter: Submitting tokens for job: job\_1651922031586\_0003

22/05/07 05:51:31 INFO impl.YarnClientImpl: Submitted application application\_1651922031586\_0003

22/05/07 05:51:31 INFO mapreduce.Job: The url to track the job: http://quickstart.cloudera:8088/proxy/application\_1651922031586\_0003/

22/05/07 05:51:31 INFO mapreduce.Job: Running job: job\_1651922031586\_0003

22/05/07 05:51:44 INFO mapreduce.Job: Job job\_1651922031586\_0003 running in uber mode : false

22/05/07 05:51:44 INFO mapreduce.Job: map 0% reduce 0%

22/05/07 05:52:02 INFO mapreduce.Job: map 75% reduce 0%

22/05/07 05:52:03 INFO mapreduce.Job: map 100% reduce 0%

22/05/07 05:52:03 INFO mapreduce.Job: Job job\_1651922031586\_0003 completed successfully

22/05/07 05:52:03 INFO mapreduce.Job: Counters: 30

File System Counters

FILE: Number of bytes read=0

FILE: Number of bytes written=542052

FILE: Number of read operations=0

FILE: Number of large read operations=0

FILE: Number of write operations=0

HDFS: Number of bytes read=393

HDFS: Number of bytes written=30

HDFS: Number of read operations=16

HDFS: Number of large read operations=0

HDFS: Number of write operations=8

Job Counters

Launched map tasks=4

Other local map tasks=4

Total time spent by all maps in occupied slots (ms)=59042

Total time spent by all reduces in occupied slots (ms)=0

Total time spent by all map tasks (ms)=59042

Total vcore-seconds taken by all map tasks=59042

Total megabyte-seconds taken by all map tasks=60459008

Map-Reduce Framework

Map input records=4

Map output records=4

Input split bytes=393

Spilled Records=0

Failed Shuffles=0

Merged Map outputs=0

GC time elapsed (ms)=675

CPU time spent (ms)=7930

Physical memory (bytes) snapshot=672169984

Virtual memory (bytes) snapshot=6277914624

Total committed heap usage (bytes)=440926208

File Input Format Counters

Bytes Read=0

File Output Format Counters

Bytes Written=30

22/05/07 05:52:03 INFO mapreduce.ImportJobBase: Transferred 30 bytes in 35.6116 seconds (0.8424 bytes/sec)

22/05/07 05:52:03 INFO mapreduce.ImportJobBase: Retrieved 4 records.

[cloudera@quickstart ~]$ hdfs dfs -ls

Found 13 items

drwxr-xr-x - cloudera cloudera 0 2019-12-29 00:57 \_sqoop

drwxr-xr-x - cloudera cloudera 0 2022-05-07 05:49 abhi

drwxr-xr-x - cloudera cloudera 0 2019-06-26 01:29 bank

drwxr-xr-x - cloudera cloudera 0 2019-06-29 06:38 btm

drwxr-xr-x - cloudera cloudera 0 2022-05-04 05:12 cloudera

-rw-r--r-- 3 cloudera cloudera 275 2018-03-11 01:47 emp

drwxr-xr-x - root cloudera 0 2022-05-03 04:55 employee

drwxr-xr-x - cloudera cloudera 0 2018-03-02 02:56 input

drwxr-xr-x - cloudera cloudera 0 2017-03-02 18:13 spark

drwxr-xr-x - cloudera cloudera 0 2018-08-24 09:57 spark\_data

-rw-r--r-- 1 cloudera cloudera 36 2018-03-02 01:58 spark\_inp

drwxr-xr-x - cloudera cloudera 0 2018-07-15 02:03 sparklab

drwxr-xr-x - cloudera cloudera 0 2018-07-14 21:54 test

[cloudera@quickstart ~]$ hdfs dfs -ls abhi

[cloudera@quickstart ~]$ hdfs dfs -ls /abhi

ls: `/abhi': No such file or directory

[cloudera@quickstart ~]$ hdfs dfs -ls /home/cloudera/abhi

Found 5 items

-rw-r--r-- 1 cloudera supergroup 0 2022-05-07 05:52 /home/cloudera/abhi/\_SUCCESS

-rw-r--r-- 1 cloudera supergroup 6 2022-05-07 05:52 /home/cloudera/abhi/part-m-00000

-rw-r--r-- 1 cloudera supergroup 6 2022-05-07 05:52 /home/cloudera/abhi/part-m-00001

-rw-r--r-- 1 cloudera supergroup 8 2022-05-07 05:52 /home/cloudera/abhi/part-m-00002

-rw-r--r-- 1 cloudera supergroup 10 2022-05-07 05:52 /home/cloudera/abhi/part-m-00003

[cloudera@quickstart ~]$ hdfs dfs -cat /home/cloudera/abhi/part-m-0000\*

1,raj

2,ram

3,parth

4,shubham

[cloudera@quickstart ~]$ hdfs dfs -cat /home/cloudera/abhi/part-m-00001

2,ram

[cloudera@quickstart ~]$ hdfs dfs -cat /home/cloudera/abhi/part-m-00002

3,parth

[cloudera@quickstart ~]$ hdfs dfs -cat /home/cloudera/abhi/part-m-00003

4,shubham

mysql> create table employee4(id int primary key,name varchar(10));

Query OK, 0 rows affected (0.01 sec)

mysql> insert into employee4 values(1,'raj');

Query OK, 1 row affected (0.01 sec)

mysql> insert into employee4 values(2,'ram');

Query OK, 1 row affected (0.00 sec)

mysql> insert into employee4 values(3,'parth');

Query OK, 1 row affected (0.00 sec)

mysql> insert into employee4 values(4,'shubham');

Query OK, 1 row affected (0.00 sec)

mysql> select \* from employee4;

+----+---------+

| id | name |

+----+---------+

| 1 | raj |

| 2 | ram |

| 3 | parth |

| 4 | shubham |

+----+---------+

4 rows in set (0.00 sec)

-----------------------------------------------------------------------------------------------------------------------------------------------------------------

-----------------------------------------------------------------------------------------------------------------------------------------------------------------

--using only two map

[cloudera@quickstart ~]$ sqoop import --connect jdbc:mysql://quickstart.cloudera/tron --username root -P --table employee5 -m 2 --target-dir /home/cloudera/data

Warning: /usr/lib/sqoop/../accumulo does not exist! Accumulo imports will fail.

Please set $ACCUMULO\_HOME to the root of your Accumulo installation.

22/05/08 02:22:32 INFO sqoop.Sqoop: Running Sqoop version: 1.4.5-cdh5.4.0

Enter password:

22/05/08 02:22:34 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.

22/05/08 02:22:34 INFO tool.CodeGenTool: Beginning code generation

22/05/08 02:22:35 INFO manager.SqlManager: Executing SQL statement: SELECT t.\* FROM `employee5` AS t LIMIT 1

22/05/08 02:22:35 INFO manager.SqlManager: Executing SQL statement: SELECT t.\* FROM `employee5` AS t LIMIT 1

22/05/08 02:22:35 INFO orm.CompilationManager: HADOOP\_MAPRED\_HOME is /usr/lib/hadoop-mapreduce

Note: /tmp/sqoop-cloudera/compile/20785a14de5ab97650c64bd0c49120f4/employee5.java uses or overrides a deprecated API.

Note: Recompile with -Xlint:deprecation for details.

22/05/08 02:22:37 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-cloudera/compile/20785a14de5ab97650c64bd0c49120f4/employee5.jar

22/05/08 02:22:37 WARN manager.MySQLManager: It looks like you are importing from mysql.

22/05/08 02:22:37 WARN manager.MySQLManager: This transfer can be faster! Use the --direct

22/05/08 02:22:37 WARN manager.MySQLManager: option to exercise a MySQL-specific fast path.

22/05/08 02:22:37 INFO manager.MySQLManager: Setting zero DATETIME behavior to convertToNull (mysql)

22/05/08 02:22:37 INFO mapreduce.ImportJobBase: Beginning import of employee5

22/05/08 02:22:37 INFO Configuration.deprecation: mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address

22/05/08 02:22:38 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar

22/05/08 02:22:39 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps

22/05/08 02:22:39 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032

22/05/08 02:22:42 INFO db.DBInputFormat: Using read commited transaction isolation

22/05/08 02:22:42 INFO db.DataDrivenDBInputFormat: BoundingValsQuery: SELECT MIN(`id`), MAX(`id`) FROM `employee5`

22/05/08 02:22:42 INFO mapreduce.JobSubmitter: number of splits:2

22/05/08 02:22:42 INFO mapreduce.JobSubmitter: Submitting tokens for job: job\_1652001167505\_0001

22/05/08 02:22:43 INFO impl.YarnClientImpl: Submitted application application\_1652001167505\_0001

22/05/08 02:22:44 INFO mapreduce.Job: The url to track the job: http://quickstart.cloudera:8088/proxy/application\_1652001167505\_0001/

22/05/08 02:22:44 INFO mapreduce.Job: Running job: job\_1652001167505\_0001

22/05/08 02:22:57 INFO mapreduce.Job: Job job\_1652001167505\_0001 running in uber mode : false

22/05/08 02:22:57 INFO mapreduce.Job: map 0% reduce 0%

22/05/08 02:23:11 INFO mapreduce.Job: map 100% reduce 0%

22/05/08 02:23:12 INFO mapreduce.Job: Job job\_1652001167505\_0001 completed successfully

22/05/08 02:23:13 INFO mapreduce.Job: Counters: 30

File System Counters

FILE: Number of bytes read=0

FILE: Number of bytes written=271026

FILE: Number of read operations=0

FILE: Number of large read operations=0

FILE: Number of write operations=0

HDFS: Number of bytes read=197

HDFS: Number of bytes written=30

HDFS: Number of read operations=8

HDFS: Number of large read operations=0

HDFS: Number of write operations=4

Job Counters

Launched map tasks=2

Other local map tasks=2

Total time spent by all maps in occupied slots (ms)=21290

Total time spent by all reduces in occupied slots (ms)=0

Total time spent by all map tasks (ms)=21290

Total vcore-seconds taken by all map tasks=21290

Total megabyte-seconds taken by all map tasks=21800960

Map-Reduce Framework

Map input records=4

Map output records=4

Input split bytes=197

Spilled Records=0

Failed Shuffles=0

Merged Map outputs=0

GC time elapsed (ms)=195

CPU time spent (ms)=3470

Physical memory (bytes) snapshot=331341824

Virtual memory (bytes) snapshot=3143335936

Total committed heap usage (bytes)=220725248

File Input Format Counters

Bytes Read=0

File Output Format Counters

Bytes Written=30

22/05/08 02:23:13 INFO mapreduce.ImportJobBase: Transferred 30 bytes in 33.6228 seconds (0.8923 bytes/sec)

22/05/08 02:23:13 INFO mapreduce.ImportJobBase: Retrieved 4 records.

[cloudera@quickstart ~]$ hdfs dfs -ls data

[cloudera@quickstart ~]$ hdfs dfs -ls

Found 14 items

drwxr-xr-x - cloudera cloudera 0 2019-12-29 00:57 \_sqoop

drwxr-xr-x - cloudera cloudera 0 2022-05-07 05:49 abhi

drwxr-xr-x - cloudera cloudera 0 2019-06-26 01:29 bank

drwxr-xr-x - cloudera cloudera 0 2019-06-29 06:38 btm

drwxr-xr-x - cloudera cloudera 0 2022-05-04 05:12 cloudera

drwxr-xr-x - cloudera cloudera 0 2022-05-08 02:22 data

-rw-r--r-- 3 cloudera cloudera 275 2018-03-11 01:47 emp

drwxr-xr-x - root cloudera 0 2022-05-03 04:55 employee

drwxr-xr-x - cloudera cloudera 0 2018-03-02 02:56 input

drwxr-xr-x - cloudera cloudera 0 2017-03-02 18:13 spark

drwxr-xr-x - cloudera cloudera 0 2018-08-24 09:57 spark\_data

-rw-r--r-- 1 cloudera cloudera 36 2018-03-02 01:58 spark\_inp

drwxr-xr-x - cloudera cloudera 0 2018-07-15 02:03 sparklab

drwxr-xr-x - cloudera cloudera 0 2018-07-14 21:54 test

[cloudera@quickstart ~]$ hdfs dfs -ls /home/cloudera/data

Found 3 items

-rw-r--r-- 1 cloudera supergroup 0 2022-05-08 02:23 /home/cloudera/data/\_SUCCESS

-rw-r--r-- 1 cloudera supergroup 12 2022-05-08 02:23 /home/cloudera/data/part-m-00000

-rw-r--r-- 1 cloudera supergroup 18 2022-05-08 02:23 /home/cloudera/data/part-m-00001

[cloudera@quickstart ~]$ hdfs dfs -ls /home/cloudera/data/part-m-00000

-rw-r--r-- 1 cloudera supergroup 12 2022-05-08 02:23 /home/cloudera/data/part-m-00000

[cloudera@quickstart ~]$ hdfs dfs -cat /home/cloudera/data/part-m-00000

1,raj

2,ram

[cloudera@quickstart ~]$ hdfs dfs -cat /home/cloudera/data/part-m-00001

3,parth

4,shubham

mysql> use tron;

Reading table information for completion of table and column names

You can turn off this feature to get a quicker startup with -A

Database changed

mysql> show tables;

+----------------+

| Tables\_in\_tron |

+----------------+

| employee |

| employee2 |

| employee3 |

| employee4 |

+----------------+

4 rows in set (0.00 sec)

mysql> create table employee5(id int primary key,name varchar(10));

Query OK, 0 rows affected (0.02 sec)

mysql> insert into employee5 values(1,'raj');

Query OK, 1 row affected (0.02 sec)

mysql> insert into employee5 values(2,'ram');

Query OK, 1 row affected (0.02 sec)

mysql> insert into employee5 values(3,'parth');

Query OK, 1 row affected (0.01 sec)

mysql> insert into employee5 values(4,'shubham');

Query OK, 1 row affected (0.02 sec)

mysql> select \* from employee5;

+----+---------+

| id | name |

+----+---------+

| 1 | raj |

| 2 | ram |

| 3 | parth |

| 4 | shubham |

+----+---------+

4 rows in set (0.00 sec)

mysql>

================================================================================================================================================================

--fields-terminated-by '\t'

[cloudera@quickstart ~]$ sqoop import --connect jdbc:mysql://quickstart.cloudera/tron --username root -P --table employee5 -m 2 --target-dir import1 --fields-terminated-by '\t'

Warning: /usr/lib/sqoop/../accumulo does not exist! Accumulo imports will fail.

Please set $ACCUMULO\_HOME to the root of your Accumulo installation.

22/05/08 10:21:05 INFO sqoop.Sqoop: Running Sqoop version: 1.4.5-cdh5.4.0

Enter password:

22/05/08 10:21:07 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.

22/05/08 10:21:07 INFO tool.CodeGenTool: Beginning code generation

22/05/08 10:21:08 INFO manager.SqlManager: Executing SQL statement: SELECT t.\* FROM `employee5` AS t LIMIT 1

22/05/08 10:21:08 INFO manager.SqlManager: Executing SQL statement: SELECT t.\* FROM `employee5` AS t LIMIT 1

22/05/08 10:21:08 INFO orm.CompilationManager: HADOOP\_MAPRED\_HOME is /usr/lib/hadoop-mapreduce

Note: /tmp/sqoop-cloudera/compile/748b82b76ab94fe9c1bc8263ee73b584/employee5.java uses or overrides a deprecated API.

Note: Recompile with -Xlint:deprecation for details.

22/05/08 10:21:10 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-cloudera/compile/748b82b76ab94fe9c1bc8263ee73b584/employee5.jar

22/05/08 10:21:10 WARN manager.MySQLManager: It looks like you are importing from mysql.

22/05/08 10:21:10 WARN manager.MySQLManager: This transfer can be faster! Use the --direct

22/05/08 10:21:10 WARN manager.MySQLManager: option to exercise a MySQL-specific fast path.

22/05/08 10:21:10 INFO manager.MySQLManager: Setting zero DATETIME behavior to convertToNull (mysql)

22/05/08 10:21:10 INFO mapreduce.ImportJobBase: Beginning import of employee5

22/05/08 10:21:10 INFO Configuration.deprecation: mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address

22/05/08 10:21:11 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar

22/05/08 10:21:12 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps

22/05/08 10:21:12 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032

22/05/08 10:21:12 WARN security.UserGroupInformation: PriviledgedActionException as:cloudera (auth:SIMPLE) cause:org.apache.hadoop.mapred.FileAlreadyExistsException: Output directory hdfs://quickstart.cloudera:8020/user/cloudera/import1 already exists

22/05/08 10:21:12 ERROR tool.ImportTool: Encountered IOException running import job: org.apache.hadoop.mapred.FileAlreadyExistsException: Output directory hdfs://quickstart.cloudera:8020/user/cloudera/import1 already exists

at org.apache.hadoop.mapreduce.lib.output.FileOutputFormat.checkOutputSpecs(FileOutputFormat.java:146)

at org.apache.hadoop.mapreduce.JobSubmitter.checkSpecs(JobSubmitter.java:562)

at org.apache.hadoop.mapreduce.JobSubmitter.submitJobInternal(JobSubmitter.java:432)

at org.apache.hadoop.mapreduce.Job$10.run(Job.java:1306)

at org.apache.hadoop.mapreduce.Job$10.run(Job.java:1303)

at java.security.AccessController.doPrivileged(Native Method)

at javax.security.auth.Subject.doAs(Subject.java:415)

at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1671)

at org.apache.hadoop.mapreduce.Job.submit(Job.java:1303)

at org.apache.hadoop.mapreduce.Job.waitForCompletion(Job.java:1324)

at org.apache.sqoop.mapreduce.ImportJobBase.doSubmitJob(ImportJobBase.java:196)

at org.apache.sqoop.mapreduce.ImportJobBase.runJob(ImportJobBase.java:169)

at org.apache.sqoop.mapreduce.ImportJobBase.runImport(ImportJobBase.java:266)

at org.apache.sqoop.manager.SqlManager.importTable(SqlManager.java:668)

at org.apache.sqoop.manager.MySQLManager.importTable(MySQLManager.java:118)

at org.apache.sqoop.tool.ImportTool.importTable(ImportTool.java:497)

at org.apache.sqoop.tool.ImportTool.run(ImportTool.java:605)

at org.apache.sqoop.Sqoop.run(Sqoop.java:143)

at org.apache.hadoop.util.ToolRunner.run(ToolRunner.java:70)

at org.apache.sqoop.Sqoop.runSqoop(Sqoop.java:179)

at org.apache.sqoop.Sqoop.runTool(Sqoop.java:218)

at org.apache.sqoop.Sqoop.runTool(Sqoop.java:227)

at org.apache.sqoop.Sqoop.main(Sqoop.java:236)

[cloudera@quickstart ~]$ sqoop import --connect jdbc:mysql://quickstart.cloudera/tron --username root -P --table employee5 -m 2 --target-dir import2 --fields-terminated-by '\t'

Warning: /usr/lib/sqoop/../accumulo does not exist! Accumulo imports will fail.

Please set $ACCUMULO\_HOME to the root of your Accumulo installation.

22/05/08 10:21:25 INFO sqoop.Sqoop: Running Sqoop version: 1.4.5-cdh5.4.0

Enter password:

22/05/08 10:21:28 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.

22/05/08 10:21:28 INFO tool.CodeGenTool: Beginning code generation

22/05/08 10:21:29 INFO manager.SqlManager: Executing SQL statement: SELECT t.\* FROM `employee5` AS t LIMIT 1

22/05/08 10:21:29 INFO manager.SqlManager: Executing SQL statement: SELECT t.\* FROM `employee5` AS t LIMIT 1

22/05/08 10:21:29 INFO orm.CompilationManager: HADOOP\_MAPRED\_HOME is /usr/lib/hadoop-mapreduce

Note: /tmp/sqoop-cloudera/compile/9ff69d12c6b5dc902b1e29db98104e94/employee5.java uses or overrides a deprecated API.

Note: Recompile with -Xlint:deprecation for details.

22/05/08 10:21:31 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-cloudera/compile/9ff69d12c6b5dc902b1e29db98104e94/employee5.jar

22/05/08 10:21:31 WARN manager.MySQLManager: It looks like you are importing from mysql.

22/05/08 10:21:31 WARN manager.MySQLManager: This transfer can be faster! Use the --direct

22/05/08 10:21:31 WARN manager.MySQLManager: option to exercise a MySQL-specific fast path.

22/05/08 10:21:31 INFO manager.MySQLManager: Setting zero DATETIME behavior to convertToNull (mysql)

22/05/08 10:21:31 INFO mapreduce.ImportJobBase: Beginning import of employee5

22/05/08 10:21:31 INFO Configuration.deprecation: mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address

22/05/08 10:21:31 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar

22/05/08 10:21:32 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps

22/05/08 10:21:32 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032

22/05/08 10:21:34 INFO db.DBInputFormat: Using read commited transaction isolation

22/05/08 10:21:34 INFO db.DataDrivenDBInputFormat: BoundingValsQuery: SELECT MIN(`id`), MAX(`id`) FROM `employee5`

22/05/08 10:21:34 INFO mapreduce.JobSubmitter: number of splits:2

22/05/08 10:21:35 INFO mapreduce.JobSubmitter: Submitting tokens for job: job\_1652029794918\_0002

22/05/08 10:21:35 INFO impl.YarnClientImpl: Submitted application application\_1652029794918\_0002

22/05/08 10:21:35 INFO mapreduce.Job: The url to track the job: http://quickstart.cloudera:8088/proxy/application\_1652029794918\_0002/

22/05/08 10:21:35 INFO mapreduce.Job: Running job: job\_1652029794918\_0002

22/05/08 10:21:44 INFO mapreduce.Job: Job job\_1652029794918\_0002 running in uber mode : false

22/05/08 10:21:44 INFO mapreduce.Job: map 0% reduce 0%

22/05/08 10:21:54 INFO mapreduce.Job: map 100% reduce 0%

22/05/08 10:21:55 INFO mapreduce.Job: Job job\_1652029794918\_0002 completed successfully

22/05/08 10:21:55 INFO mapreduce.Job: Counters: 30

File System Counters

FILE: Number of bytes read=0

FILE: Number of bytes written=271006

FILE: Number of read operations=0

FILE: Number of large read operations=0

FILE: Number of write operations=0

HDFS: Number of bytes read=197

HDFS: Number of bytes written=30

HDFS: Number of read operations=8

HDFS: Number of large read operations=0

HDFS: Number of write operations=4

Job Counters

Launched map tasks=2

Other local map tasks=2

Total time spent by all maps in occupied slots (ms)=15605

Total time spent by all reduces in occupied slots (ms)=0

Total time spent by all map tasks (ms)=15605

Total vcore-seconds taken by all map tasks=15605

Total megabyte-seconds taken by all map tasks=15979520

Map-Reduce Framework

Map input records=4

Map output records=4

Input split bytes=197

Spilled Records=0

Failed Shuffles=0

Merged Map outputs=0

GC time elapsed (ms)=294

CPU time spent (ms)=3590

Physical memory (bytes) snapshot=336953344

Virtual memory (bytes) snapshot=3145191424

Total committed heap usage (bytes)=350224384

File Input Format Counters

Bytes Read=0

File Output Format Counters

Bytes Written=30

22/05/08 10:21:55 INFO mapreduce.ImportJobBase: Transferred 30 bytes in 23.1987 seconds (1.2932 bytes/sec)

22/05/08 10:21:55 INFO mapreduce.ImportJobBase: Retrieved 4 records.

[cloudera@quickstart ~]$ hdfs dfs -ls

Found 16 items

drwxr-xr-x - cloudera cloudera 0 2019-12-29 00:57 \_sqoop

drwxr-xr-x - cloudera cloudera 0 2022-05-07 05:49 abhi

drwxr-xr-x - cloudera cloudera 0 2019-06-26 01:29 bank

drwxr-xr-x - cloudera cloudera 0 2019-06-29 06:38 btm

drwxr-xr-x - cloudera cloudera 0 2022-05-04 05:12 cloudera

drwxr-xr-x - cloudera cloudera 0 2022-05-08 02:22 data

-rw-r--r-- 3 cloudera cloudera 275 2018-03-11 01:47 emp

drwxr-xr-x - root cloudera 0 2022-05-03 04:55 employee

drwxr-xr-x - cloudera cloudera 0 2022-05-08 10:19 import1

drwxr-xr-x - cloudera cloudera 0 2022-05-08 10:21 import2

drwxr-xr-x - cloudera cloudera 0 2018-03-02 02:56 input

drwxr-xr-x - cloudera cloudera 0 2017-03-02 18:13 spark

drwxr-xr-x - cloudera cloudera 0 2018-08-24 09:57 spark\_data

-rw-r--r-- 1 cloudera cloudera 36 2018-03-02 01:58 spark\_inp

drwxr-xr-x - cloudera cloudera 0 2018-07-15 02:03 sparklab

drwxr-xr-x - cloudera cloudera 0 2018-07-14 21:54 test

[cloudera@quickstart ~]$ hdfs dfs -ls import2

Found 3 items

-rw-r--r-- 1 cloudera cloudera 0 2022-05-08 10:21 import2/\_SUCCESS

-rw-r--r-- 1 cloudera cloudera 12 2022-05-08 10:21 import2/part-m-00000

-rw-r--r-- 1 cloudera cloudera 18 2022-05-08 10:21 import2/part-m-00001

[cloudera@quickstart ~]$ hdfs dfs -ls import2/part-m-0000\*

-rw-r--r-- 1 cloudera cloudera 12 2022-05-08 10:21 import2/part-m-00000

-rw-r--r-- 1 cloudera cloudera 18 2022-05-08 10:21 import2/part-m-00001

[cloudera@quickstart ~]$ hdfs dfs -cat import2/part-m-0000\*

1 raj

2 ram

3 parth

4 shubham

=====================================================================================================================================================================

--where

[cloudera@quickstart ~]$ sqoop import --connect jdbc:mysql://quickstart.cloudera/tron --username root -P --table employee5 -m 2 --where 'id=1' --target-dir import3

Warning: /usr/lib/sqoop/../accumulo does not exist! Accumulo imports will fail.

Please set $ACCUMULO\_HOME to the root of your Accumulo installation.

22/05/08 10:27:22 INFO sqoop.Sqoop: Running Sqoop version: 1.4.5-cdh5.4.0

Enter password:

22/05/08 10:27:28 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.

22/05/08 10:27:28 INFO tool.CodeGenTool: Beginning code generation

22/05/08 10:27:28 INFO manager.SqlManager: Executing SQL statement: SELECT t.\* FROM `employee5` AS t LIMIT 1

22/05/08 10:27:29 INFO manager.SqlManager: Executing SQL statement: SELECT t.\* FROM `employee5` AS t LIMIT 1

22/05/08 10:27:29 INFO orm.CompilationManager: HADOOP\_MAPRED\_HOME is /usr/lib/hadoop-mapreduce

Note: /tmp/sqoop-cloudera/compile/e17e8f185fab02b5b05bd9582abcb9f7/employee5.java uses or overrides a deprecated API.

Note: Recompile with -Xlint:deprecation for details.

22/05/08 10:27:31 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-cloudera/compile/e17e8f185fab02b5b05bd9582abcb9f7/employee5.jar

22/05/08 10:27:31 WARN manager.MySQLManager: It looks like you are importing from mysql.

22/05/08 10:27:31 WARN manager.MySQLManager: This transfer can be faster! Use the --direct

22/05/08 10:27:31 WARN manager.MySQLManager: option to exercise a MySQL-specific fast path.

22/05/08 10:27:31 INFO manager.MySQLManager: Setting zero DATETIME behavior to convertToNull (mysql)

22/05/08 10:27:31 INFO mapreduce.ImportJobBase: Beginning import of employee5

22/05/08 10:27:31 INFO Configuration.deprecation: mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address

22/05/08 10:27:31 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar

22/05/08 10:27:32 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps

22/05/08 10:27:32 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032

22/05/08 10:27:34 INFO db.DBInputFormat: Using read commited transaction isolation

22/05/08 10:27:34 INFO db.DataDrivenDBInputFormat: BoundingValsQuery: SELECT MIN(`id`), MAX(`id`) FROM `employee5` WHERE ( id=1 )

22/05/08 10:27:34 INFO mapreduce.JobSubmitter: number of splits:1

22/05/08 10:27:34 INFO mapreduce.JobSubmitter: Submitting tokens for job: job\_1652029794918\_0003

22/05/08 10:27:35 INFO impl.YarnClientImpl: Submitted application application\_1652029794918\_0003

22/05/08 10:27:35 INFO mapreduce.Job: The url to track the job: http://quickstart.cloudera:8088/proxy/application\_1652029794918\_0003/

22/05/08 10:27:35 INFO mapreduce.Job: Running job: job\_1652029794918\_0003

22/05/08 10:27:44 INFO mapreduce.Job: Job job\_1652029794918\_0003 running in uber mode : false

22/05/08 10:27:44 INFO mapreduce.Job: map 0% reduce 0%

22/05/08 10:27:51 INFO mapreduce.Job: map 100% reduce 0%

22/05/08 10:27:52 INFO mapreduce.Job: Job job\_1652029794918\_0003 completed successfully

22/05/08 10:27:53 INFO mapreduce.Job: Counters: 30

File System Counters

FILE: Number of bytes read=0

FILE: Number of bytes written=135784

FILE: Number of read operations=0

FILE: Number of large read operations=0

FILE: Number of write operations=0

HDFS: Number of bytes read=99

HDFS: Number of bytes written=6

HDFS: Number of read operations=4

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)=4764

Total time spent by all reduces in occupied slots (ms)=0

Total time spent by all map tasks (ms)=4764

Total vcore-seconds taken by all map tasks=4764

Total megabyte-seconds taken by all map tasks=4878336

Map-Reduce Framework

Map input records=1

Map output records=1

Input split bytes=99

Spilled Records=0

Failed Shuffles=0

Merged Map outputs=0

GC time elapsed (ms)=87

CPU time spent (ms)=1440

Physical memory (bytes) snapshot=175951872

Virtual memory (bytes) snapshot=1576300544

Total committed heap usage (bytes)=175112192

File Input Format Counters

Bytes Read=0

File Output Format Counters

Bytes Written=6

22/05/08 10:27:53 INFO mapreduce.ImportJobBase: Transferred 6 bytes in 20.5827 seconds (0.2915 bytes/sec)

22/05/08 10:27:53 INFO mapreduce.ImportJobBase: Retrieved 1 records.

[cloudera@quickstart ~]$ hdfs dfs -ls import3

Found 2 items

-rw-r--r-- 1 cloudera cloudera 0 2022-05-08 10:27 import3/\_SUCCESS

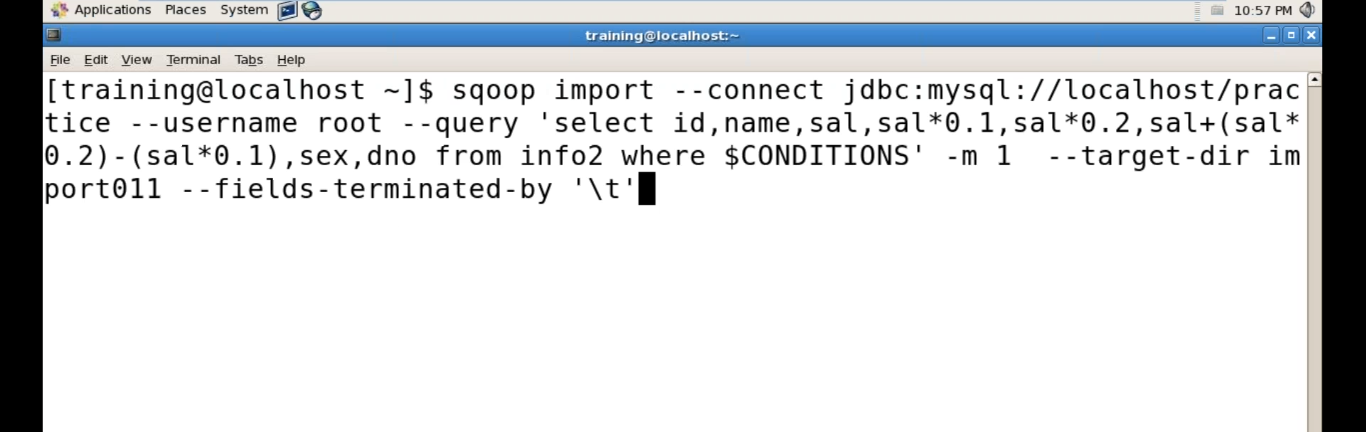
-rw-r--r-- 1 cloudera cloudera 6 2022-05-08 10:27 import3/part-m-00000

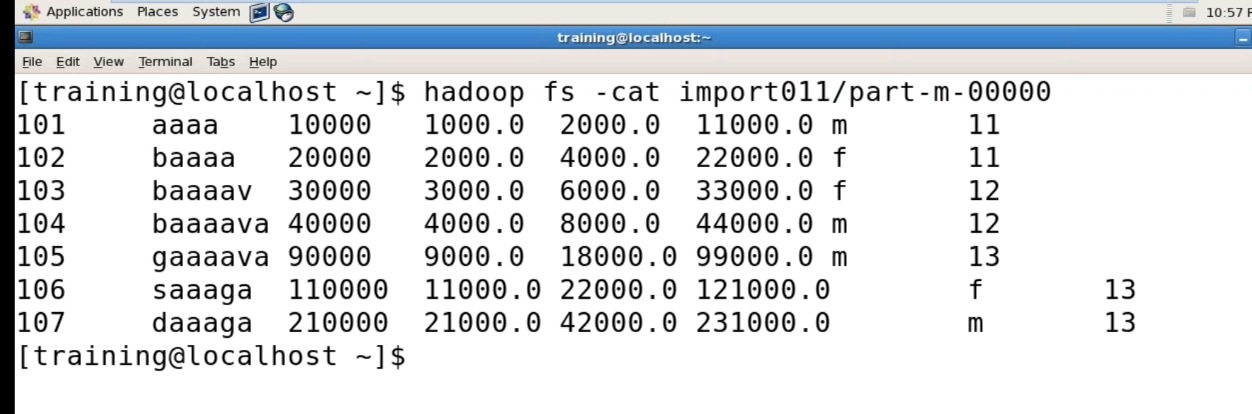
[cloudera@quickstart ~]$ hdfs dfs -ls import3/part-m-00000

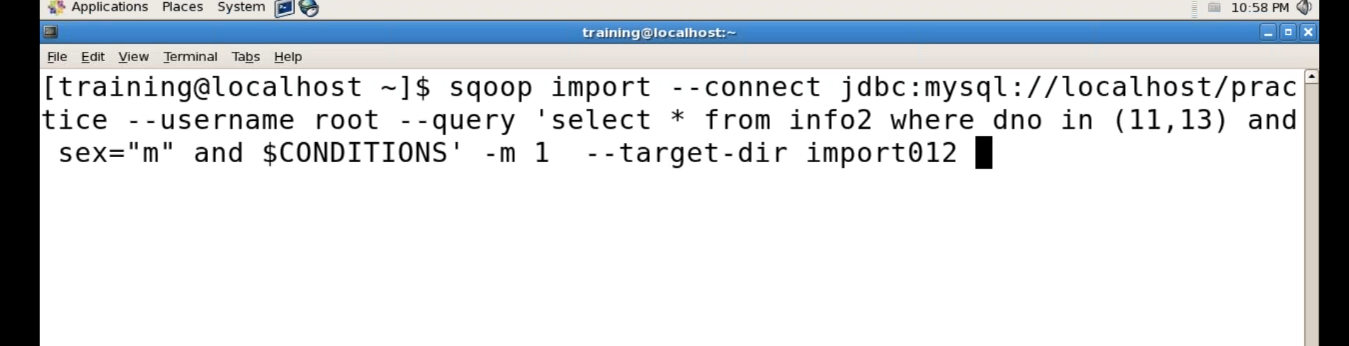
-rw-r--r-- 1 cloudera cloudera 6 2022-05-08 10:27 import3/part-m-00000

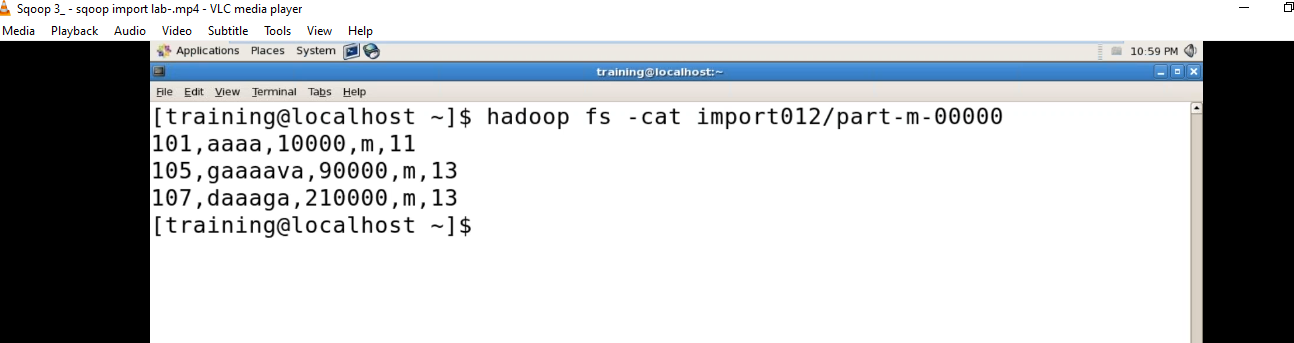
[cloudera@quickstart ~]$ hdfs dfs -cat import3/part-m-00000

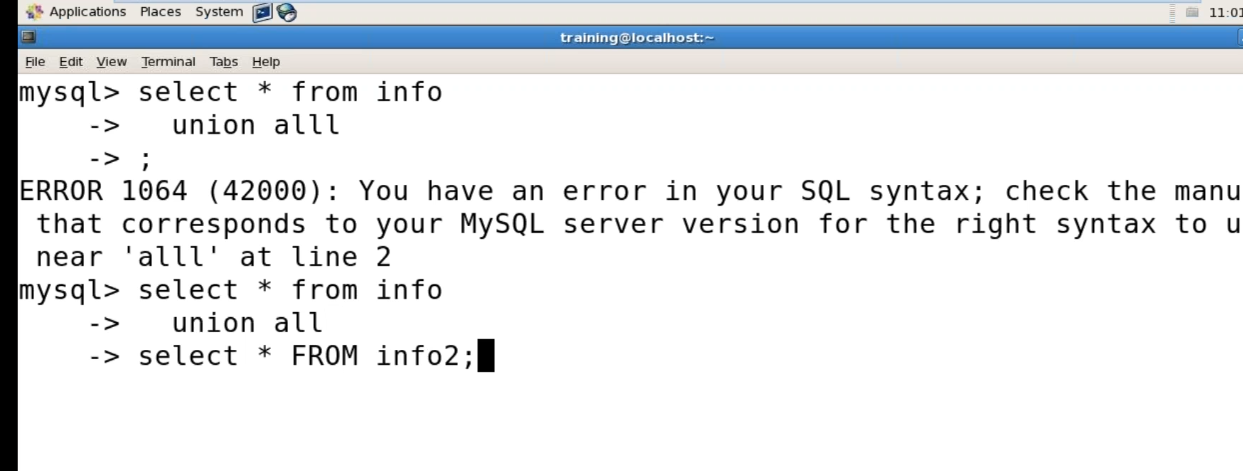
1,raj

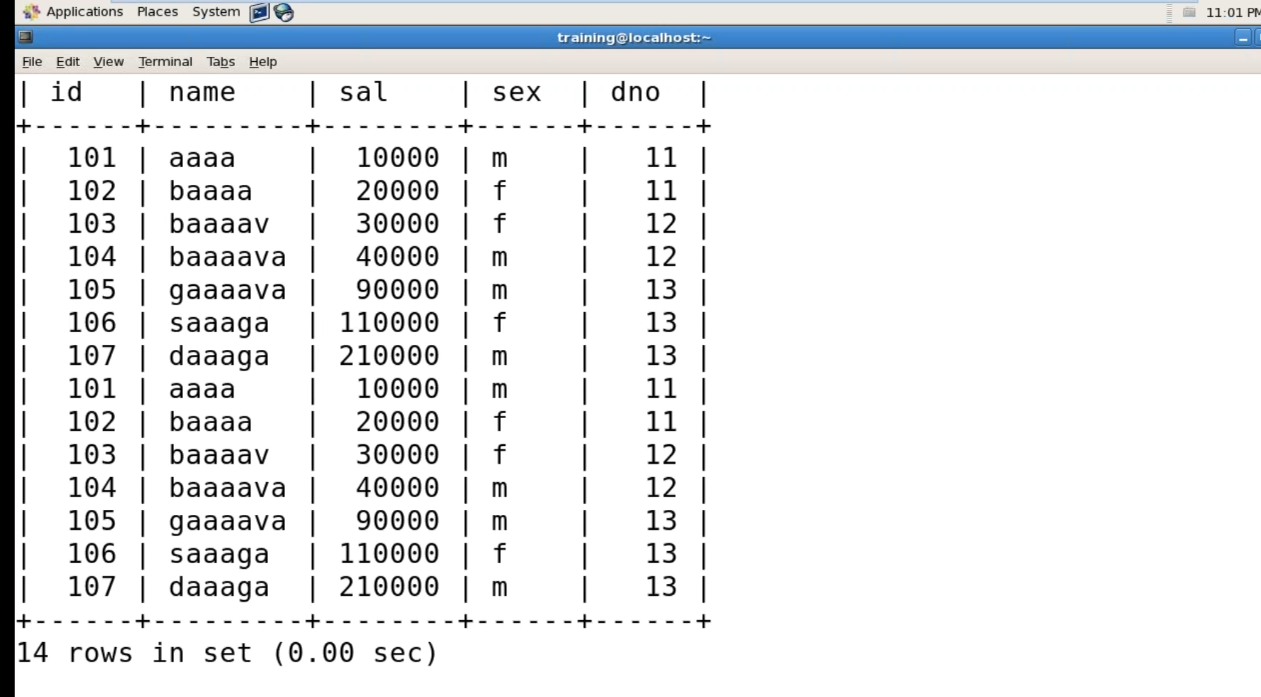


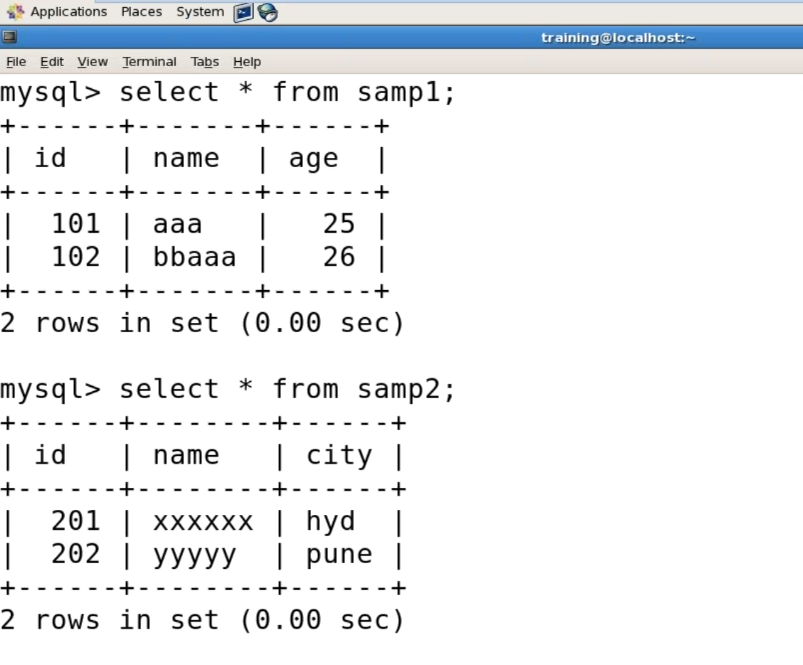


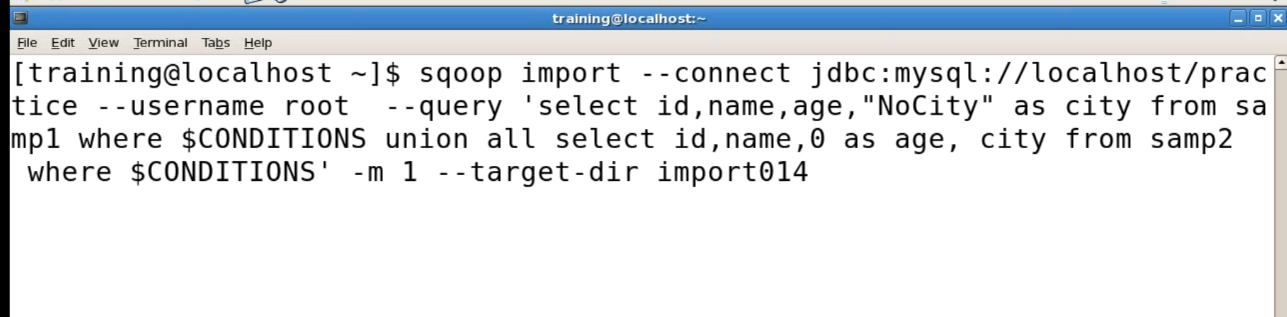


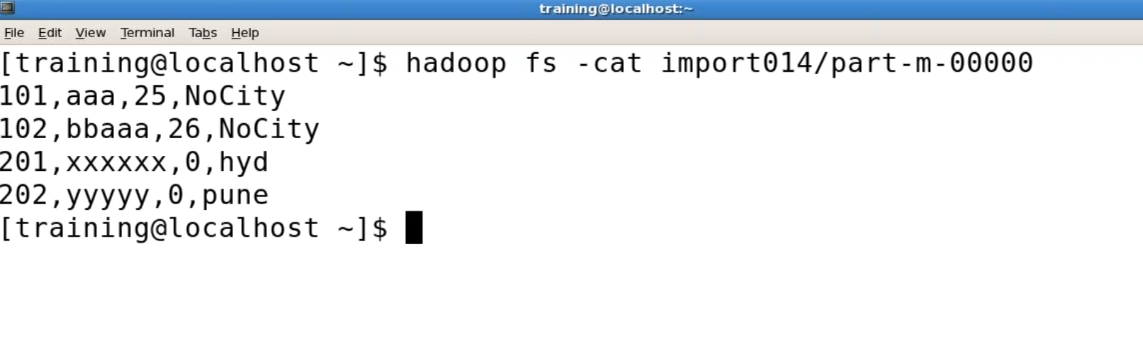


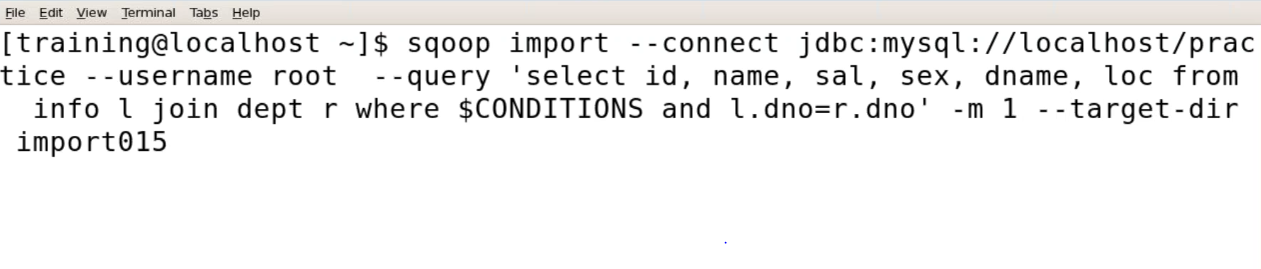


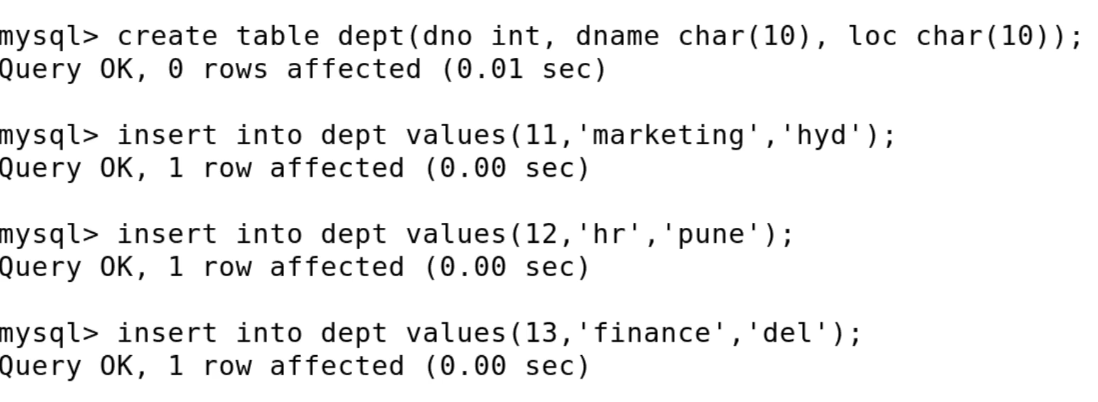


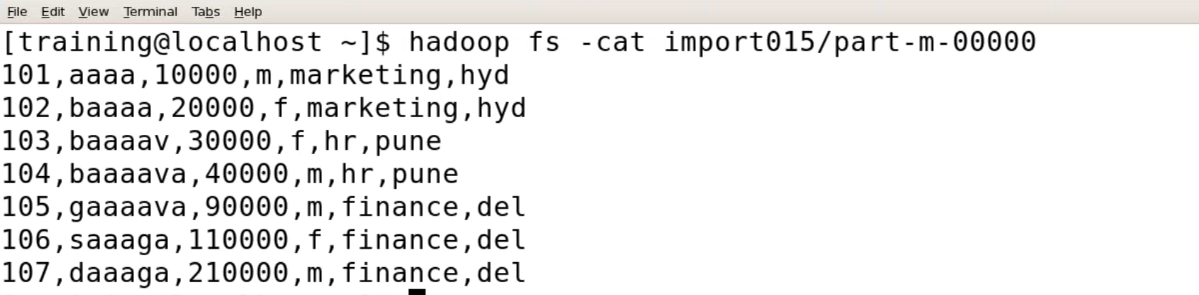


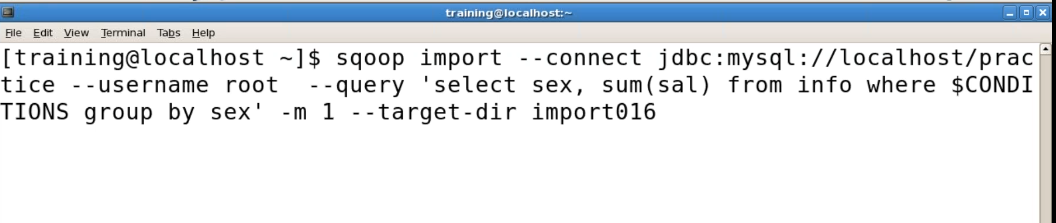


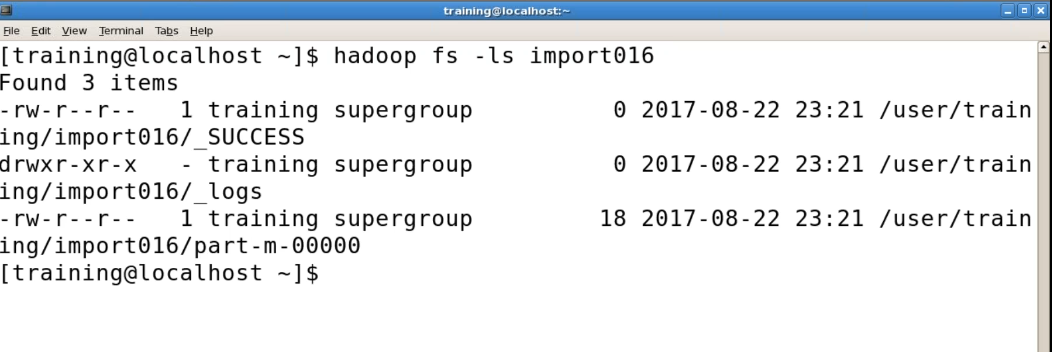


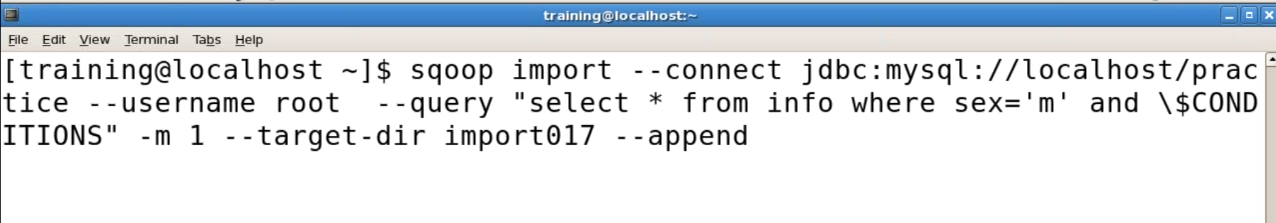






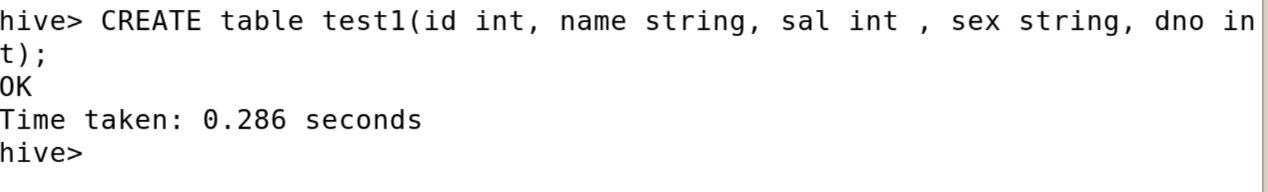


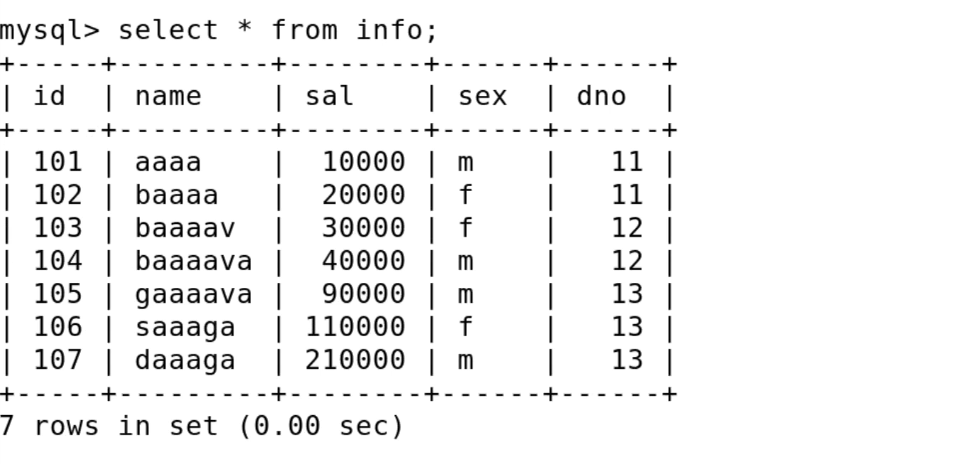




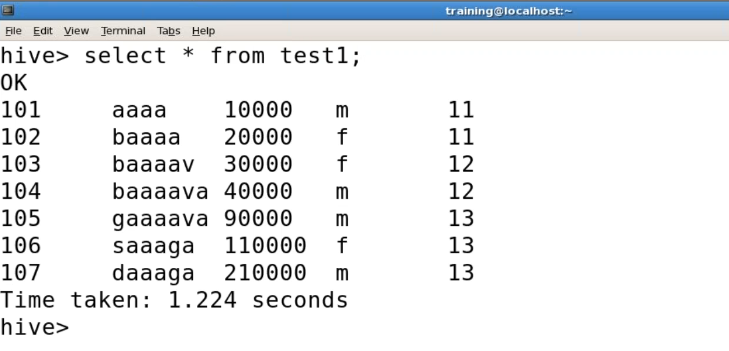


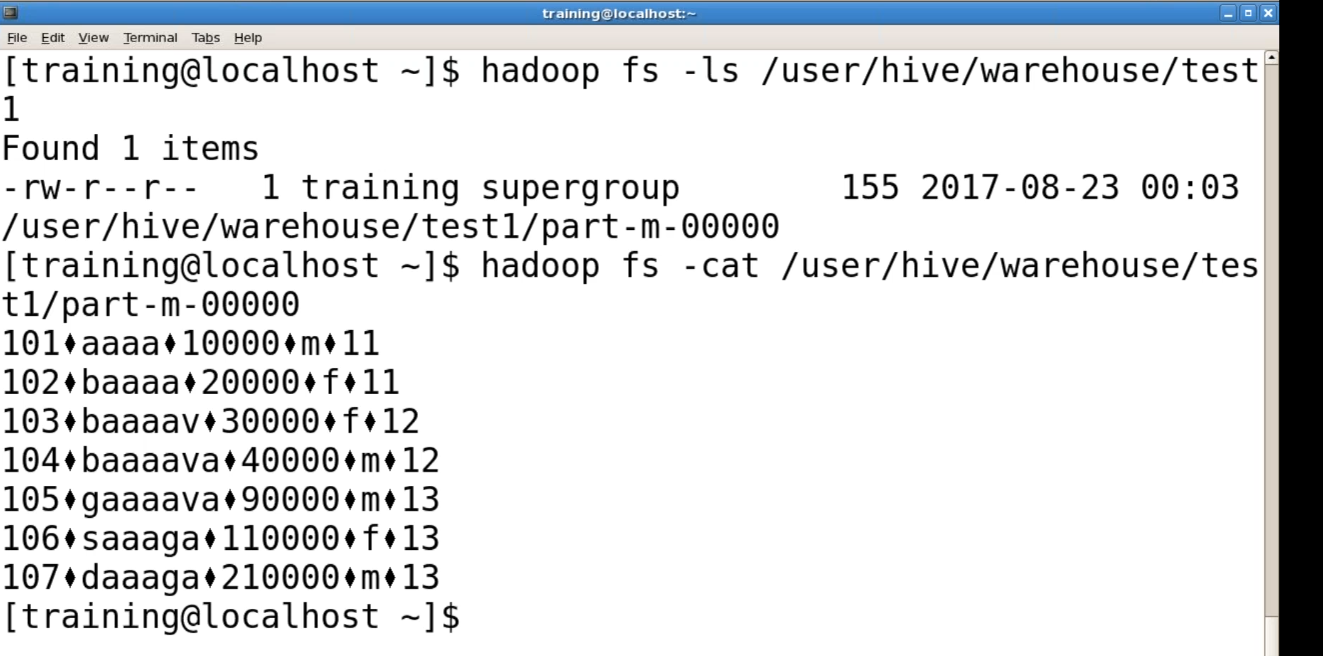
Import mysql to hive

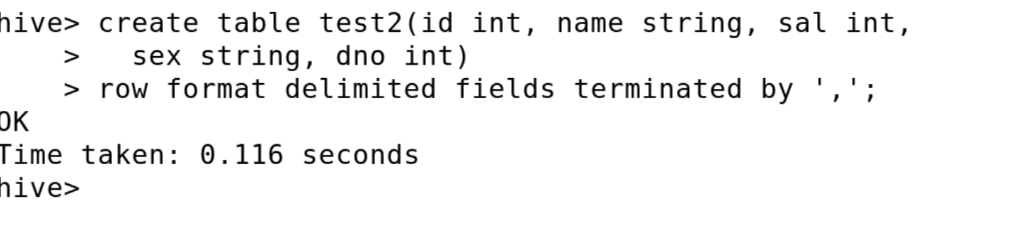


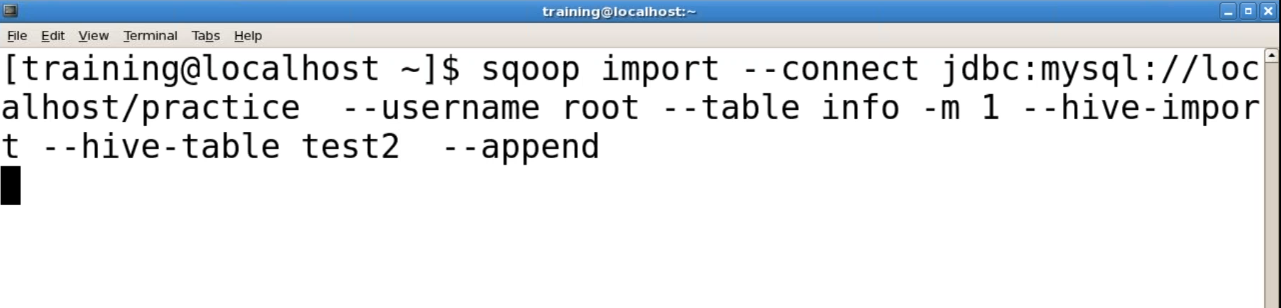


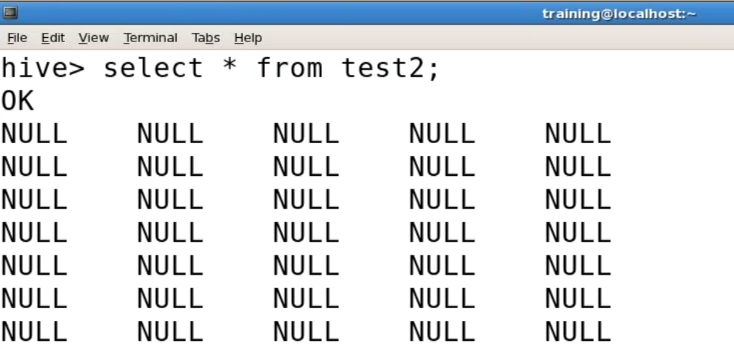


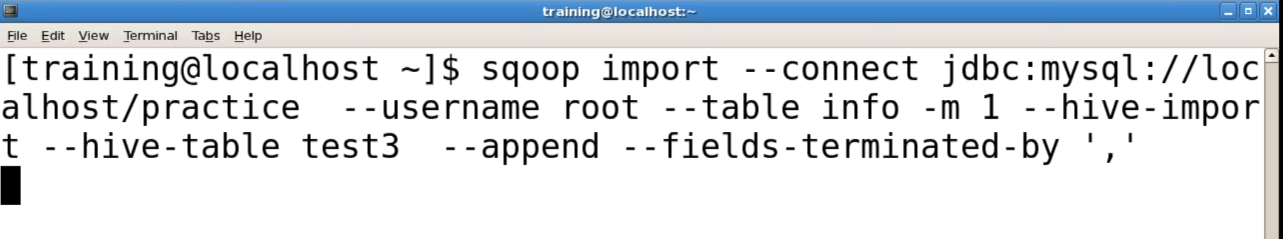


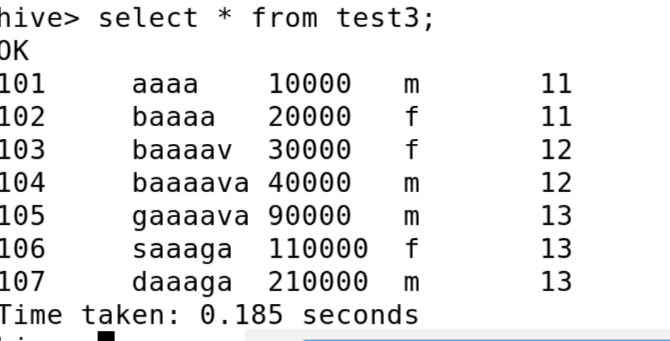












[cloudera@quickstart ~]$ hdfs dfs -ls /

Found 8 items

drwxr-xr-x - cloudera supergroup 0 2019-06-26 01:46 /bank

drwxr-xr-x - hbase supergroup 0 2022-05-09 03:30 /hbase

drwxr-xr-x - cloudera supergroup 0 2022-05-07 05:51 /home

drwxr-xr-x - solr solr 0 2015-04-23 07:49 /solr

-rw-r--r-- 1 cloudera supergroup 60 2022-05-09 03:44 /test

drwxrwxrwx - hdfs supergroup 0 2016-12-16 23:21 /tmp

drwxr-xr-x - hdfs supergroup 0 2022-05-08 02:48 /user

drwxr-xr-x - hdfs supergroup 0 2015-04-23 07:47 /var

[cloudera@quickstart ~]$ hdfs dfs -ls /test

-rw-r--r-- 1 cloudera supergroup 60 2022-05-09 03:44 /test

[cloudera@quickstart ~]$ hdfs dfs -cat /test

100,200,300

400,500,600

700,800,900

390,490,590

690,790,890

[cloudera@quickstart ~]$ sqoop export --connect jdbc:mysql://quickstart.cloudera/tron --username root -P --table lab1 --export-dir /test

Warning: /usr/lib/sqoop/../accumulo does not exist! Accumulo imports will fail.

Please set $ACCUMULO\_HOME to the root of your Accumulo installation.

22/05/09 03:53:21 INFO sqoop.Sqoop: Running Sqoop version: 1.4.5-cdh5.4.0

Enter password:

22/05/09 03:53:23 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.

22/05/09 03:53:23 INFO tool.CodeGenTool: Beginning code generation

22/05/09 03:53:24 INFO manager.SqlManager: Executing SQL statement: SELECT t.\* FROM `lab1` AS t LIMIT 1

22/05/09 03:53:24 INFO manager.SqlManager: Executing SQL statement: SELECT t.\* FROM `lab1` AS t LIMIT 1

22/05/09 03:53:24 INFO orm.CompilationManager: HADOOP\_MAPRED\_HOME is /usr/lib/hadoop-mapreduce

Note: /tmp/sqoop-cloudera/compile/47887e0b44ca724f4b05be8e3e61d86d/lab1.java uses or overrides a deprecated API.

Note: Recompile with -Xlint:deprecation for details.

22/05/09 03:53:26 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-cloudera/compile/47887e0b44ca724f4b05be8e3e61d86d/lab1.jar

22/05/09 03:53:26 INFO mapreduce.ExportJobBase: Beginning export of lab1

22/05/09 03:53:26 INFO Configuration.deprecation: mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address

22/05/09 03:53:27 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar

22/05/09 03:53:28 INFO Configuration.deprecation: mapred.reduce.tasks.speculative.execution is deprecated. Instead, use mapreduce.reduce.speculative

22/05/09 03:53:28 INFO Configuration.deprecation: mapred.map.tasks.speculative.execution is deprecated. Instead, use mapreduce.map.speculative

22/05/09 03:53:28 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps

22/05/09 03:53:28 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032

22/05/09 03:53:31 INFO input.FileInputFormat: Total input paths to process : 1

22/05/09 03:53:31 INFO input.FileInputFormat: Total input paths to process : 1

22/05/09 03:53:31 INFO mapreduce.JobSubmitter: number of splits:4

22/05/09 03:53:31 INFO Configuration.deprecation: mapred.map.tasks.speculative.execution is deprecated. Instead, use mapreduce.map.speculative

22/05/09 03:53:32 INFO mapreduce.JobSubmitter: Submitting tokens for job: job\_1652092179823\_0001

22/05/09 03:53:33 INFO impl.YarnClientImpl: Submitted application application\_1652092179823\_0001

22/05/09 03:53:33 INFO mapreduce.Job: The url to track the job: http://quickstart.cloudera:8088/proxy/application\_1652092179823\_0001/

22/05/09 03:53:33 INFO mapreduce.Job: Running job: job\_1652092179823\_0001

22/05/09 03:53:47 INFO mapreduce.Job: Job job\_1652092179823\_0001 running in uber mode : false

22/05/09 03:53:47 INFO mapreduce.Job: map 0% reduce 0%

22/05/09 03:54:06 INFO mapreduce.Job: map 50% reduce 0%

22/05/09 03:54:07 INFO mapreduce.Job: map 100% reduce 0%

22/05/09 03:54:08 INFO mapreduce.Job: Job job\_1652092179823\_0001 completed successfully

22/05/09 03:54:08 INFO mapreduce.Job: Counters: 30

File System Counters

FILE: Number of bytes read=0

FILE: Number of bytes written=540468

FILE: Number of read operations=0

FILE: Number of large read operations=0

FILE: Number of write operations=0

HDFS: Number of bytes read=630

HDFS: Number of bytes written=0

HDFS: Number of read operations=16

HDFS: Number of large read operations=0

HDFS: Number of write operations=0

Job Counters

Launched map tasks=4

Data-local map tasks=4

Total time spent by all maps in occupied slots (ms)=63314

Total time spent by all reduces in occupied slots (ms)=0

Total time spent by all map tasks (ms)=63314

Total vcore-seconds taken by all map tasks=63314

Total megabyte-seconds taken by all map tasks=64833536

Map-Reduce Framework

Map input records=5

Map output records=5

Input split bytes=468

Spilled Records=0

Failed Shuffles=0

Merged Map outputs=0

GC time elapsed (ms)=644

CPU time spent (ms)=6790

Physical memory (bytes) snapshot=646492160

Virtual memory (bytes) snapshot=6268309504

Total committed heap usage (bytes)=505937920

File Input Format Counters

Bytes Read=0

File Output Format Counters

Bytes Written=0

22/05/09 03:54:08 INFO mapreduce.ExportJobBase: Transferred 630 bytes in 39.4987 seconds (15.9499 bytes/sec)

22/05/09 03:54:08 INFO mapreduce.ExportJobBase: Exported 5 records.

mysql> use tron;

Reading table information for completion of table and column names

You can turn off this feature to get a quicker startup with -A

Database changed

mysql> show tables;

+----------------+

| Tables\_in\_tron |

+----------------+

| employee |

| employee2 |

| employee3 |

| employee4 |

| employee5 |

+----------------+

5 rows in set (0.00 sec)

mysql> create table lab1(A int,B int,C int);

Query OK, 0 rows affected (0.01 sec)

mysql> describe lab1;

+-------+---------+------+-----+---------+-------+

| Field | Type | Null | Key | Default | Extra |

+-------+---------+------+-----+---------+-------+

| A | int(11) | YES | | NULL | |

| B | int(11) | YES | | NULL | |

| C | int(11) | YES | | NULL | |

+-------+---------+------+-----+---------+-------+

3 rows in set (0.00 sec)

mysql> select \* from lab1;

+------+------+------+

| A | B | C |

+------+------+------+

| 100 | 200 | 300 |

| 400 | 500 | 600 |

| 700 | 800 | 900 |

| 390 | 490 | 590 |

| 690 | 790 | 890 |

+------+------+------+

5 rows in set (0.00 sec)

mysql> create table lab2(A int,B int,C int);

Query OK, 0 rows affected (0.00 sec)

[cloudera@quickstart rwq]$ hdfs dfs -copyFromLocal lab1 /abhi12

[cloudera@quickstart rwq]$ hdfs dfs -ls /

Found 10 items

-rw-r--r-- 1 cloudera supergroup 72 2022-05-09 04:11 /abhi

drwxr-xr-x - cloudera supergroup 0 2022-05-09 04:13 /abhi12

drwxr-xr-x - cloudera supergroup 0 2019-06-26 01:46 /bank

drwxr-xr-x - hbase supergroup 0 2022-05-09 03:30 /hbase

drwxr-xr-x - cloudera supergroup 0 2022-05-07 05:51 /home

drwxr-xr-x - solr solr 0 2015-04-23 07:49 /solr

-rw-r--r-- 1 cloudera supergroup 60 2022-05-09 03:44 /test

drwxrwxrwx - hdfs supergroup 0 2016-12-16 23:21 /tmp

drwxr-xr-x - hdfs supergroup 0 2022-05-08 02:48 /user

drwxr-xr-x - hdfs supergroup 0 2015-04-23 07:47 /var

[cloudera@quickstart rwq]$ hdfs dfs -ls /abhi12

Found 1 items

-rw-r--r-- 1 cloudera supergroup 72 2022-05-09 04:13 /abhi12/lab1

[cloudera@quickstart rwq]$ sqoop export --connect jdbc:mysql://quickstart.cloudera/tron --username root -P --table lab2 --export-dir /abhi12/lab1 --input-fields-terminated-by '\t'

Warning: /usr/lib/sqoop/../accumulo does not exist! Accumulo imports will fail.

Please set $ACCUMULO\_HOME to the root of your Accumulo installation.

22/05/09 04:14:29 INFO sqoop.Sqoop: Running Sqoop version: 1.4.5-cdh5.4.0

Enter password:

22/05/09 04:14:31 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.

22/05/09 04:14:31 INFO tool.CodeGenTool: Beginning code generation

22/05/09 04:14:32 INFO manager.SqlManager: Executing SQL statement: SELECT t.\* FROM `lab2` AS t LIMIT 1

22/05/09 04:14:32 INFO manager.SqlManager: Executing SQL statement: SELECT t.\* FROM `lab2` AS t LIMIT 1

22/05/09 04:14:32 INFO orm.CompilationManager: HADOOP\_MAPRED\_HOME is /usr/lib/hadoop-mapreduce

Note: /tmp/sqoop-cloudera/compile/41db7bfb7e097ee1fa40f9ed039f7adf/lab2.java uses or overrides a deprecated API.

Note: Recompile with -Xlint:deprecation for details.

22/05/09 04:14:34 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-cloudera/compile/41db7bfb7e097ee1fa40f9ed039f7adf/lab2.jar

22/05/09 04:14:34 INFO mapreduce.ExportJobBase: Beginning export of lab2

22/05/09 04:14:34 INFO Configuration.deprecation: mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address

22/05/09 04:14:34 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar

22/05/09 04:14:35 INFO Configuration.deprecation: mapred.reduce.tasks.speculative.execution is deprecated. Instead, use mapreduce.reduce.speculative

22/05/09 04:14:35 INFO Configuration.deprecation: mapred.map.tasks.speculative.execution is deprecated. Instead, use mapreduce.map.speculative

22/05/09 04:14:35 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps

22/05/09 04:14:35 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032

22/05/09 04:14:37 INFO input.FileInputFormat: Total input paths to process : 1

22/05/09 04:14:37 INFO input.FileInputFormat: Total input paths to process : 1

22/05/09 04:14:37 INFO mapreduce.JobSubmitter: number of splits:4

22/05/09 04:14:37 INFO Configuration.deprecation: mapred.map.tasks.speculative.execution is deprecated. Instead, use mapreduce.map.speculative

22/05/09 04:14:38 INFO mapreduce.JobSubmitter: Submitting tokens for job: job\_1652092179823\_0005

22/05/09 04:14:38 INFO impl.YarnClientImpl: Submitted application application\_1652092179823\_0005

22/05/09 04:14:38 INFO mapreduce.Job: The url to track the job: http://quickstart.cloudera:8088/proxy/application\_1652092179823\_0005/

22/05/09 04:14:38 INFO mapreduce.Job: Running job: job\_1652092179823\_0005

22/05/09 04:14:48 INFO mapreduce.Job: Job job\_1652092179823\_0005 running in uber mode : false

22/05/09 04:14:48 INFO mapreduce.Job: map 0% reduce 0%

22/05/09 04:15:07 INFO mapreduce.Job: map 75% reduce 0%

22/05/09 04:15:08 INFO mapreduce.Job: map 100% reduce 0%

22/05/09 04:15:09 INFO mapreduce.Job: Job job\_1652092179823\_0005 completed successfully

22/05/09 04:15:09 INFO mapreduce.Job: Counters: 30

File System Counters

FILE: Number of bytes read=0

FILE: Number of bytes written=540524

FILE: Number of read operations=0

FILE: Number of large read operations=0

FILE: Number of write operations=0

HDFS: Number of bytes read=688

HDFS: Number of bytes written=0

HDFS: Number of read operations=16

HDFS: Number of large read operations=0

HDFS: Number of write operations=0

Job Counters

Launched map tasks=4

Data-local map tasks=4

Total time spent by all maps in occupied slots (ms)=60648

Total time spent by all reduces in occupied slots (ms)=0

Total time spent by all map tasks (ms)=60648

Total vcore-seconds taken by all map tasks=60648

Total megabyte-seconds taken by all map tasks=62103552

Map-Reduce Framework

Map input records=6

Map output records=6

Input split bytes=496

Spilled Records=0

Failed Shuffles=0

Merged Map outputs=0

GC time elapsed (ms)=691

CPU time spent (ms)=7380

Physical memory (bytes) snapshot=643244032

Virtual memory (bytes) snapshot=6248783872

Total committed heap usage (bytes)=505937920

File Input Format Counters

Bytes Read=0

File Output Format Counters

Bytes Written=0

22/05/09 04:15:09 INFO mapreduce.ExportJobBase: Transferred 688 bytes in 33.7859 seconds (20.3635 bytes/sec)

22/05/09 04:15:09 INFO mapreduce.ExportJobBase: Exported 6 records.

mysql> select \* from lab2;

+------+------+------+

| A | B | C |

+------+------+------+

| 700 | 800 | 900 |

| 340 | 450 | 550 |

| 100 | 200 | 300 |

| 400 | 500 | 600 |

| 230 | 430 | 330 |

| 110 | 210 | 310 |

+------+------+------+

6 rows in set (0.00 sec)

-rw-r--r-- 1 root root 99 May 9 04:33 emp

[cloudera@quickstart ~]$ pwd

/home/cloudera

[cloudera@quickstart ~]$ cat emp

1,abhijeet,90000,m,11

2,prasad,90000,m,12

3,parth,84000,m,13

4,ram,89000,m,11

5,shubham,87000,m,12

[cloudera@quickstart ~]$

hive> load data local inpath '/home/cloudera/emp' into table emp2;

Loading data to table default.emp2

Table default.emp2 stats: [numFiles=1, totalSize=99]

OK

Time taken: 0.725 seconds

hive> select \* from emp2;

OK

1 abhijeet 90000 m 11

2 prasad 90000 m 12

3 parth 84000 m 13

4 ram 89000 m 11

5 shubham 87000 m 12

Time taken: 0.082 seconds, Fetched: 5 row(s)

[cloudera@quickstart ~]$ hdfs dfs -put emp /user

[cloudera@quickstart ~]$ sqoop export --connect jdbc:mysql://quickstart.cloudera/tron --username root -P --table lab3 --export-dir /user/emp

Warning: /usr/lib/sqoop/../accumulo does not exist! Accumulo imports will fail.

Please set $ACCUMULO\_HOME to the root of your Accumulo installation.

22/05/09 20:30:34 INFO sqoop.Sqoop: Running Sqoop version: 1.4.5-cdh5.4.0

Enter password:

22/05/09 20:30:37 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.

22/05/09 20:30:37 INFO tool.CodeGenTool: Beginning code generation

22/05/09 20:30:37 INFO manager.SqlManager: Executing SQL statement: SELECT t.\* FROM `lab3` AS t LIMIT 1

22/05/09 20:30:37 INFO manager.SqlManager: Executing SQL statement: SELECT t.\* FROM `lab3` AS t LIMIT 1

22/05/09 20:30:37 INFO orm.CompilationManager: HADOOP\_MAPRED\_HOME is /usr/lib/hadoop-mapreduce

Note: /tmp/sqoop-cloudera/compile/2abd92d44a200e774ba4f545dc79d0e7/lab3.java uses or overrides a deprecated API.

Note: Recompile with -Xlint:deprecation for details.

22/05/09 20:30:39 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-cloudera/compile/2abd92d44a200e774ba4f545dc79d0e7/lab3.jar

22/05/09 20:30:39 INFO mapreduce.ExportJobBase: Beginning export of lab3

22/05/09 20:30:39 INFO Configuration.deprecation: mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address

22/05/09 20:30:40 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar

22/05/09 20:30:41 INFO Configuration.deprecation: mapred.reduce.tasks.speculative.execution is deprecated. Instead, use mapreduce.reduce.speculative

22/05/09 20:30:41 INFO Configuration.deprecation: mapred.map.tasks.speculative.execution is deprecated. Instead, use mapreduce.map.speculative

22/05/09 20:30:41 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps

22/05/09 20:30:41 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032

22/05/09 20:30:44 INFO input.FileInputFormat: Total input paths to process : 1

22/05/09 20:30:44 INFO input.FileInputFormat: Total input paths to process : 1

22/05/09 20:30:45 INFO mapreduce.JobSubmitter: number of splits:4

22/05/09 20:30:45 INFO Configuration.deprecation: mapred.map.tasks.speculative.execution is deprecated. Instead, use mapreduce.map.speculative

22/05/09 20:30:45 INFO mapreduce.JobSubmitter: Submitting tokens for job: job\_1652152070587\_0005

22/05/09 20:30:45 INFO impl.YarnClientImpl: Submitted application application\_1652152070587\_0005

22/05/09 20:30:45 INFO mapreduce.Job: The url to track the job: http://quickstart.cloudera:8088/proxy/application\_1652152070587\_0005/

22/05/09 20:30:45 INFO mapreduce.Job: Running job: job\_1652152070587\_0005

22/05/09 20:30:54 INFO mapreduce.Job: Job job\_1652152070587\_0005 running in uber mode : false

22/05/09 20:30:54 INFO mapreduce.Job: map 0% reduce 0%

22/05/09 20:31:09 INFO mapreduce.Job: map 50% reduce 0%

22/05/09 20:31:10 INFO mapreduce.Job: map 100% reduce 0%

22/05/09 20:31:11 INFO mapreduce.Job: Job job\_1652152070587\_0005 completed successfully

22/05/09 20:31:12 INFO mapreduce.Job: Counters: 30

File System Counters

FILE: Number of bytes read=0

FILE: Number of bytes written=540508

FILE: Number of read operations=0

FILE: Number of large read operations=0

FILE: Number of write operations=0

HDFS: Number of bytes read=616

HDFS: Number of bytes written=0

HDFS: Number of read operations=16

HDFS: Number of large read operations=0

HDFS: Number of write operations=0

Job Counters

Launched map tasks=4

Data-local map tasks=4

Total time spent by all maps in occupied slots (ms)=52223

Total time spent by all reduces in occupied slots (ms)=0

Total time spent by all map tasks (ms)=52223

Total vcore-seconds taken by all map tasks=52223

Total megabyte-seconds taken by all map tasks=53476352

Map-Reduce Framework

Map input records=6

Map output records=6

Input split bytes=484

Spilled Records=0

Failed Shuffles=0

Merged Map outputs=0

GC time elapsed (ms)=500

CPU time spent (ms)=6290

Physical memory (bytes) snapshot=658407424

Virtual memory (bytes) snapshot=6269681664

Total committed heap usage (bytes)=569901056

File Input Format Counters

Bytes Read=0

File Output Format Counters

Bytes Written=0

22/05/09 20:31:12 INFO mapreduce.ExportJobBase: Transferred 616 bytes in 30.5082 seconds (20.1913 bytes/sec)

22/05/09 20:31:12 INFO mapreduce.ExportJobBase: Exported 6 records.

mysql> select \* from lab3;

+----+----------+

| id | name |

+----+----------+

| 1 | abhijeet |

| 2 | ram |

| 3 | parth |

| 4 | tron |

| 5 | som |

| 6 | shubham |

+----+----------+

6 rows in set (0.00 sec)