

Task 1

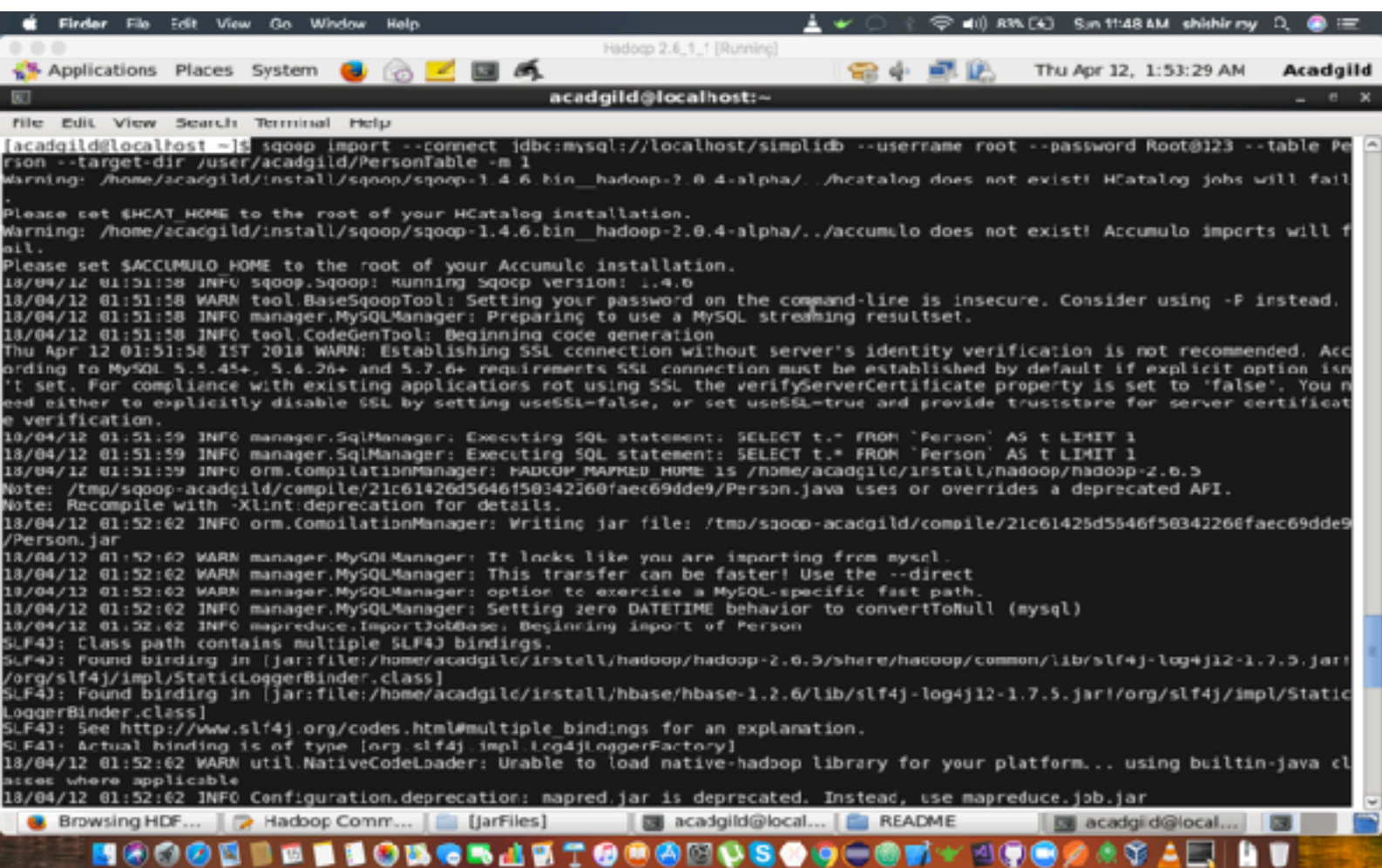
Use Sqoop tool to export data present in SQOOPOUT folder made while demo of Import table

Below commands to import mysql database to hdfs:

```
sqoop import --connect jdbc:mysql://localhost/simplidb --username root --password Root@123 --table Person --target-dir /user/acadgild/PersonTable -m 1
```

Display the Content

```
hadoop fs -cat /user/acadgild/PersonTable/part-m-00000
```



```
acadmild@localhost:~$ sqoop import --connect jdbc:mysql://localhost/simplidb --username root --password Root@123 --table Person --target-dir /user/acadgild/PersonTable -m 1
Warning: /home/acadgild/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/acadgild/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.
18/04/12 01:51:58 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6
18/04/12 01:51:58 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -F instead.
18/04/12 01:51:58 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
18/04/12 01:51:58 INFO tool.CodeGenTool: Beginning code generation
Thu Apr 12 01:51:58 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 using 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.
18/04/12 01:51:59 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM 'Person' AS t LIMIT 1
18/04/12 01:51:59 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM 'Person' AS t LIMIT 1
18/04/12 01:51:59 INFO orm.compilationManager: HADOOP_MAPRED_HOME is /home/acadgild/install/hadoop/hadoop-2.0.5
Note: /tmp/sqoop-acadmild/compile/21c61426d5646f50342260faec69dde9/Person.java uses or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
18/04/12 01:52:02 INFO orm.compilationManager: Writing jar file: /tmp/sqoop-acadmild/compile/21c61426d5646f50342260faec69dde9/Person.jar
18/04/12 01:52:02 WARN manager.MySQLManager: It looks like you are importing from mysql.
18/04/12 01:52:02 WARN manager.MySQLManager: This transfer can be faster! Use the --direct
18/04/12 01:52:02 WARN manager.MySQLManager: option to exercise a MySQL-specific fast path.
18/04/12 01:52:02 INFO manager.MySQLManager: Setting zero DATETIME behavior to convertToNull (mysql)
18/04/12 01:52:02 INFO mapreduce.ImportJobBase: Beginning import of Person
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/acadgild/install/hadoop/hadoop-2.0.5/share/hadoop/common/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/home/acadgild/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]
18/04/12 01:52:02 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
18/04/12 01:52:02 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar
```

```
File Edit View Go Window Help
Hadoop 2.6.1_1 [Running]
Applications Places System
acadmild@localhost:~
File Edit View Search Terminal Help
18/04/12 01:52:32 INFO mapreduce.Job: Counters: 30
File System Counters
  FILE: Number of bytes read=0
  FILE: Number of bytes written=127880
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
  HDFS: Number of bytes read=87
  HDFS: Number of bytes written=150
  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)=5932
  Total time spent by all reduces in occupied slots (ms)=0
  Total time spent by all map tasks (ms)=5932
  Total vcore-milliseconds taken by all map tasks=5932
  Total megabyte-milliseconds taken by all map tasks=6974908
Map-Reduce Framework
  Map input records=5
  Map output records=5
  Input split bytes=87
  Spilled Records=0
  Failed Shuffles=0
  Merged Map outputs=0
  GC time elapsed (ms)=77
  CPU time spent (ms)=1650
  Physical memory (bytes) snapshot=169178880
  Virtual memory (bytes) snapshot=2063437024
  Total committed heap usage (bytes)=32571392
File Input Format Counters
  Bytes Read=0
File Output Format Counters
  Bytes Written=150
18/04/12 01:52:32 INFO mapreduce.ImportJobBase: Transferred 150 bytes in 28.2634 seconds (5.3072 bytes/sec)
18/04/12 01:52:32 INFO mapreduce.ImportJobBase: Retrieved 5 records.
[acadmild@localhost ~]$
```

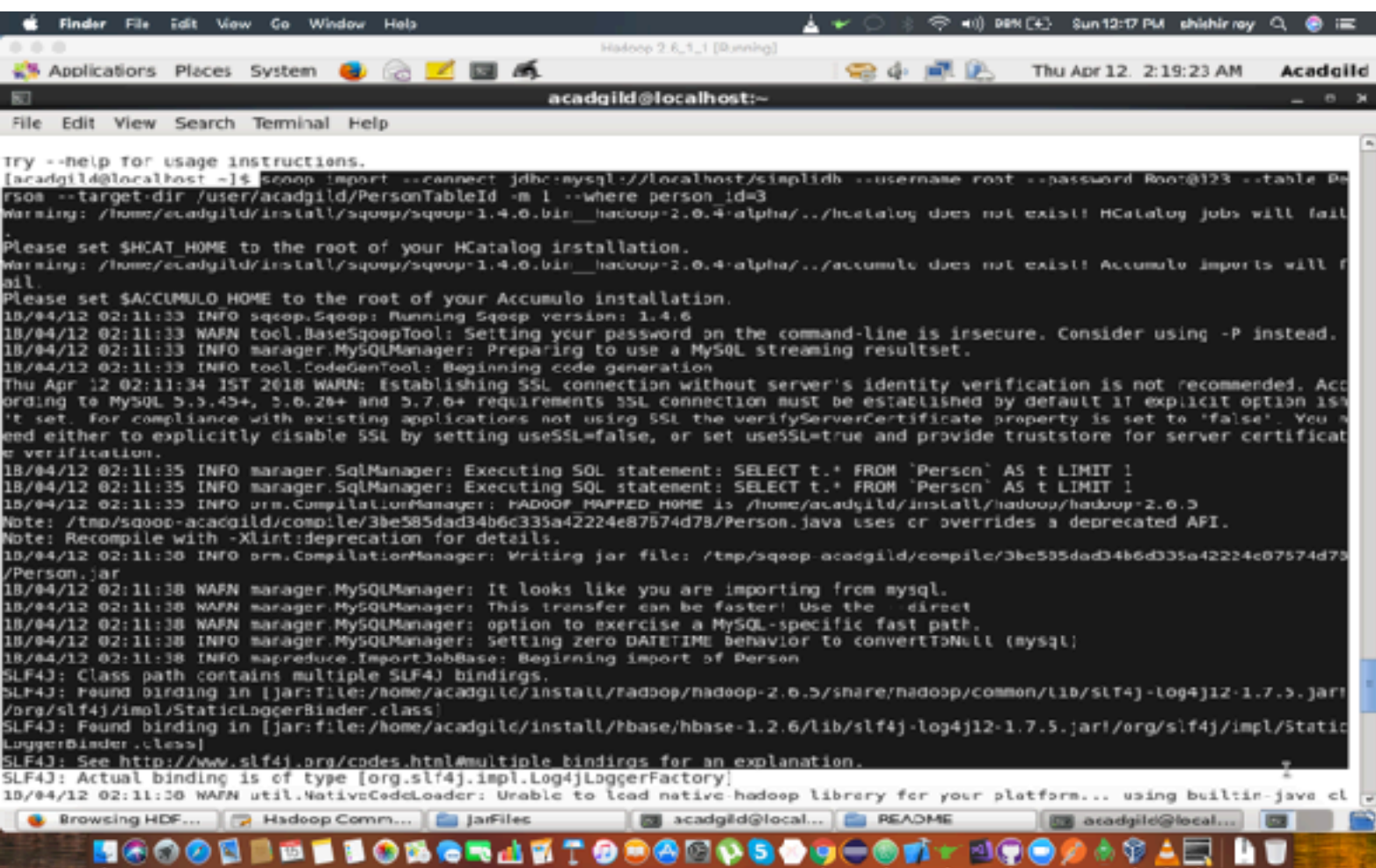
```
File Edit View Go Window Help
Hadoop 2.6.1_1 [Running]
Applications Places System
acadmild@localhost:~
File Edit View Search Terminal Help
  HDFS: Number of large read operations=0
  HDFS: Number of write operations=2
Job Counters
  Launched map tasks=1
  Other local map tasks=1
  Total time spent by all maps in occupied slots (ms)=5485
  Total time spent by all reduces in occupied slots (ms)=0
  Total time spent by all map tasks (ms)=6486
  Total vcore-milliseconds taken by all map tasks=6486
  Total megabyte-milliseconds taken by all map tasks=6541564
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)=00
  CPU time spent (ms)=1710
  Physical memory (bytes) snapshot=167577344
  Virtual memory (bytes) snapshot=2063437824
  Total committed heap usage (bytes)=32571392
File Input Format Counters
  Bytes Read=0
File Output Format Counters
  Bytes Written=37
18/04/12 02:12:05 INFO mapreduce.ImportJobBase: Transferred 37 bytes in 25.4837 seconds (1.4519 bytes/sec)
18/04/12 02:12:05 INFO mapreduce.ImportJobBase: Retrieved 1 records.
You have new mail in /var/spool/mail/acadmild
[acadmild@localhost ~]$ hadoop fs -cat /user/acadmild/PersonTable/part-8-00000
18/04/12 02:33:55 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Patna,Bihar,Ram,Shyam,1
Whitefield,Bangalore,Priya,Tanya,2
New York,United States,Brown,James,3
Los Angeles,United States,Miller,John,4
c,d,b,a,789
You have new mail in /var/spool/mail/acadmild
[acadmild@localhost ~]$
```

Below commands to import mysql database to hdfs:

```
sqoop import --connect jdbc:mysql://localhost/simplidb --username root --password Root@123 --table Person --target-dir /user/acadgild/PersonTableId -m 1 --where person_id=3
```

Display the content:

```
hadoop fs -cat /user/acadgild/PersonTableId/part-m-00000
```



The screenshot shows a Mac OS desktop with a terminal window open. The terminal window title is "acadgild@localhost:~". The command executed is `sqoop import --connect jdbc:mysql://localhost/simplidb --username root --password Root@123 --table Person --target-dir /user/acadgild/PersonTableId -m 1 --where person_id=3`. The output shows various warnings and informational messages, including a warning about HCatalog jobs failing due to missing HCatalog, a warning about Accumulo imports failing due to missing Accumulo, and a warning about SSL connection not being recommended. The output also shows the execution of SQL statements and the compilation of the PersonTableId.jar file.

```
Try --help for usage instructions.
[acadgild@localhost ~]$ sqoop import --connect jdbc:mysql://localhost/simplidb --username root --password Root@123 --table Person --target-dir /user/acadgild/PersonTableId -m 1 --where person_id=3
Warning: /home/acadgild/install/sqoop/sqoop-1.4.0/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/acadgild/install/sqoop/sqoop-1.4.0/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.
18/04/12 02:11:33 INFO Sqoop.Sqoop: Running Sqoop version: 1.4.0
18/04/12 02:11:33 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.
18/04/12 02:11:33 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
18/04/12 02:11:33 INFO tool.CodeGenTool: Beginning code generation
Thu Apr 12 02:11:34 IST 2018 WARN: Establishing SSL connection without server's identity verification is not recommended. According to MySQL 5.5.45+, 5.6.20+ and 5.7.0+ requirements SSL connection must be established by default if explicit option isn't set. For compliance with existing applications not using 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.
18/04/12 02:11:35 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM 'Person' AS t LIMIT 1
18/04/12 02:11:35 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM 'Person' AS t LIMIT 1
18/04/12 02:11:35 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /home/acadgild/install/hadoop/hadoop-2.0.3
Note: /tmp/sqoop-acadgild/compile/3be585dad34b6c335a42224e87574d78/Person.java uses or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
18/04/12 02:11:38 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-acadgild/compile/3be585dad34b6c335a42224e87574d78/Person.jar
18/04/12 02:11:38 WARN manager.MySQLManager: It looks like you are importing from mysql.
18/04/12 02:11:38 WARN manager.MySQLManager: This transfer can be faster! Use the --direct
18/04/12 02:11:38 WARN manager.MySQLManager: option to exercise a MySQL-specific fast path.
18/04/12 02:11:38 INFO manager.MySQLManager: Setting zero DATETIME behavior to convertToNull (mysql)
18/04/12 02:11:38 INFO mapreduce.ImportJobBase: Beginning import of Person
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/acadgild/install/hbase/hbase-1.2.6/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/home/acadgild/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]
18/04/12 02:11:38 WARN util.NativeCodeLoader: Unable to load native hadoop library for your platform... using builtin-java cl
```



```
File Edit View Go Window Help
Hadoop 2.6.1_1 [Running]
Applications Places System
acacgild@localhost:~
File Edit View Search Terminal Help
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
18/04/12 02:11:38 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
18/04/12 02:11:38 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar
18/04/12 02:11:40 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps
18/04/12 02:11:40 INFO client.NMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
Thu Apr 12 02:11:43 IST 2018 WARN: Establishing SSL connection without server's identity verification is not recommended. According to MySQL 5.5.45+, 5.6.20+ and 5.7.0+ requirements SSL connection must be established by default if explicit option isn't set. For compliance with existing applications not using 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.
18/04/12 02:11:43 INFO db.DBCInputFormat: Using read committed transaction isolation
18/04/12 02:11:43 INFO mapreduce.JobSubmitter: number of splits=1
18/04/12 02:11:44 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1521868459622_0028
18/04/12 02:11:45 INFO impl.YarnClientImpl: Submitted application application_1521868459622_0028
18/04/12 02:11:45 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1521868459622_0028/
18/04/12 02:11:45 INFO mapreduce.Job: Running job: job_1521868459622_0028
18/04/12 02:11:56 INFO mapreduce.Job: Job job_1521868459622_0028 running in uber mode : false
18/04/12 02:11:56 INFO mapreduce.Job: map 0% reduce 0%
18/04/12 02:12:05 INFO mapreduce.Job: map 100% reduce 0%
18/04/12 02:12:05 INFO mapreduce.Job: Job job_1521868459622_0028 completed successfully
18/04/12 02:12:05 INFO mapreduce.Job: Counters: 30
File System Counters
  FILE: Number of bytes read=0
  FILE: Number of bytes written=128034
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
  HDFS: Number of bytes read=87
  HDFS: Number of bytes written=37
  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)=5485
  Total time spent by all reduces in occupied slots (ms)=0
Browsing HDF... Hadoop Comm... JarFiles acacgild@local... README acacgild@local...
```

```
File Edit View Go Window Help
Hadoop 2.6.1_1 [Running]
Applications Places System
acacgild@localhost:~
File Edit View Search Terminal Help
File System Counters
  FILE: Number of bytes read=0
  FILE: Number of bytes written=128034
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
  HDFS: Number of bytes read=87
  HDFS: Number of bytes written=37
  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)=5485
  Total time spent by all reduces in occupied slots (ms)=0
  Total time spent by all map tasks (ms)=6486
  Total vcore-milliseconds taken by all map tasks=6486
  Total megabyte-milliseconds taken by all map tasks=6541564
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)=88
  CPU time spent (ms)=1710
  Physical memory (bytes) snapshot=167577344
  Virtual memory (bytes) snapshot=2063417824
  Total committed heap usage (bytes)=32571392
File Input Format Counters
  Bytes Read=0
File Output Format Counters
  Bytes Written=37
18/04/12 02:12:05 INFO mapreduce.ImportJobBase: Transferred 37 bytes in 25.4837 seconds (1.4519 bytes/sec)
18/04/12 02:12:05 INFO mapreduce.ImportJobBase: Retrieved 1 records.
You have new mail in /var/spool/mail/acacgild
[acacgild@localhost ~]$
Browsing HDF... Hadoop Comm... JarFiles acacgild@local... README acacgild@local...
```

```
Finder File Edit View Go Window Help
Hadoop 2.6.1_1 [Running]
Applications Places System
Thu Apr 12, 2:35:38 AM Acadgild
acadgild@localhost:~
File Edit View Search Terminal Help
Total time spent by all maps in occupied slots (ms)=6486
Total time spent by all reduces in occupied slots (ms)=0
Total time spent by all map tasks (ms)=6486
Total vcore-milliseconds taken by all map tasks=6486
Total megabyte-milliseconds taken by all map tasks=6641664
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)=88
CPU time spent (ms)=1710
Physical memory (bytes) snapshot=187577344
Virtual memory (bytes) snapshot=2063437824
Total committed heap usage (bytes)=32571392
File Input Format Counters
  Bytes Read=0
File Output Format Counters
  Bytes Written=37
18/04/12 02:12:05 INFO mapreduce.ImportJobBase: Transferred 37 bytes in 25.4837 seconds (1.4519 bytes/sec)
18/04/12 02:12:05 INFO mapreduce.ImportJobBase: Retrieved 1 records.
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ hadoop fs -cat /user/acadgild/PersonTable/part-m-00000
18/04/12 02:33:55 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
Patna,Bihar,Ram,Shyam,1
Whitefield,Bangalore,Priya,Tanya,2
New York,United States,Brown,James,3
Los Angeles ,United States,Miller,Jhon,4
c,d,b,a,789
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ hadoop fs -cat /user/acadgild/PersonTableId/part-m-00000
18/04/12 02:34:55 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
New York,United States,Brown,James,3
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$
```