The Apache Hadoop Project Page **1** of **1**

**Exercise: Importing Data into HDFS Using Sqoop**

# Importing Data from the MySQL Database Using Sqoop

In the following exercise, you review two Sqoop scripts used for importing data into Hadoop. Then you import the MySQL table, **test\_table**, into HDFS.

* + Using xshell open **student** login to your virtual machine, if you did not already do so.
  + Review the Sqoop import table script, **c01ex04\_sqoop\_import.sh**. If necessary, modify your IP address
  + Review the Sqoop import query script, **c01ex04\_sqoop\_query.sh**, and determine how it differs from the previous Sqoop import table script.
  + Before running c01ex04\_sqoop\_import.sh, open HUE as student user and open the **Job Browser** to view the running job.
  + Select the **Xshell** student login that was started previously. Navigate to DIHPS/lab01.
    - Type **sh c01ex04\_sqoop\_import.sh** to execute the script from the linux command line.
  + In the Hue Job Browser, verify that the Sqoop script runs and completes successfully.
  + Using the File Browser, verify the contents /user/student/DIHPS/test\_table imported from MySQL.

**Question**: What type of connection is used in the Sqoop import scripts?

**Answer**:

**Question**: How many mappers were used during the Sqoop import?

**Answer**:

**Question**: What is the purpose of $CONDITIONS in a Sqoop query import?

# Answer:

Copyright © 2016, Mohamed Rahal