Experiment No.2	
Use of Sqoop tool	
Date of Performance:	
Date of Submission:	



<u>AIM</u>: To install SQOOP and execute basic commands of Hadoop eco system component Sqoop.

THEORY:

Installation and configuration of SQOOP

- 1) Download SQOOP from https://sqoop.apache.org
- 2) Unzip and Install SQOOP
 - After Downloading the SQOOP, we need to Unzip the sqoop-1.4.7.bin hadoop-2.6.0.tar.gz file.
- 3) Create a folder and move the final extracted file in it.
 - Set up the environment variables
 - Set SQOOP HOME
- 4) Set up path variable
- 5) Configure SQOOP

Basic SQOOP commands:

1. List Table

This command lists the particular table of the database in MYSQL server.

sqoop list - tables --connect jdbc:mysql://localhost/payment --username gatner

2. Target directory

This command import table in a specific directory in HDFS. -m denotes mapper argument.

They have an integer value.



sqoop import --connect jdbc:mysql://localhost/inventory --username jony -table inventory --m 1 --target-dir/inv

3. sqoop-eval

This command runs SQL queries of the respective database.

\$ sqoop eval --connect --query "SQLQuery"

4. sqoop – version

This command displays a version of the sqoop.

\$ sqoop version sqoop {revnumber}

5. sqoop-job

This command allows us to create a job, the parameters that are created can be invoked at any time. They take options like (-create,-delete,-show,-exit).

sqoop job --create --import --connect --table

6. code gen

This Sqoop command creates java class files which encapsulate the imported records. All the java files are recreated, and new versions of a class are generated. They generate code to interact with database records. Retrieves a list of all the columns and their data types.

\$ sqoop codegen --connect -table

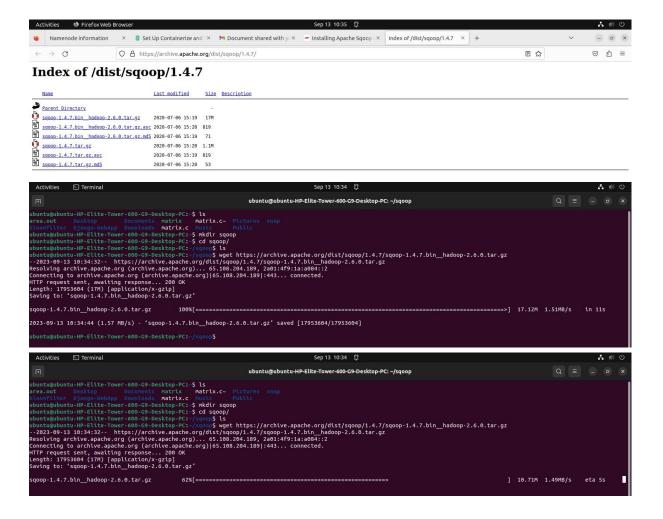
7. List Database

This Sqoop command lists all the available databases in the RDBMS server.

\$ sqoop list - database -- connect



Sqoop is a command-line interface application for transferring data between relational databases and Hadoop.





CONCLUSION:

During this experiment we installed and used the Apache Sqoop tool for efficiently transferring data between Hadoop and relational databases. Once installed, Sqoop offers a simple command-line interface and can be easily integrated into data workflows, making it a valuable asset for data engineers and analysts. Whether you need to ingest data from a database into Hadoop or export data from Hadoop to a relational database, Sqoop simplifies the process, ensuring the seamless movement of data, and contributing to more effective big data processing and analysis.