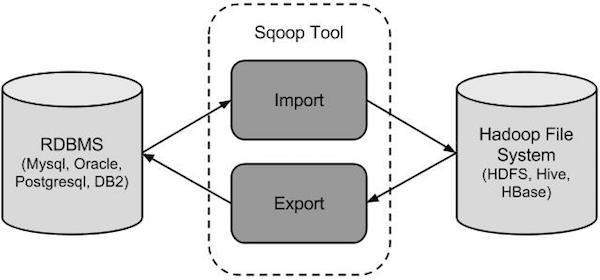
**Round2Day4 – Sqoop**

* Sqoop is a tool designed to transfer data between Hadoop and relational database servers. It is used to import data from relational databases such as MySQL, Oracle to Hadoop HDFS, and export from Hadoop file system to relational databases.
* Initial release: 1 June 2009
* Written in: Java
* OS:l Cross-platform
* Official website: https://sqoop.apache.org/
* Sqoop import – the import tool imports individual tables from RDBMS to HDFS. Each row in a table is treated as a record in HDFS. All records are stored as text data in text files or as binary data in Avro and Sequence files
* Sqoop Export – the export tool exports a set of files from HDFS back to an RDBMS. The files given an input to Sqoop contains records, which are called as rows in table. Those are read and parsed into a set of records and delimited with user-specified delimiter
* Sqoop provides many salient features like:
  + Full load
  + Incremental load
  + Parallel import/export
  + Import results of SQL query
  + Compression
  + Connectors for all major RDBMS databases
  + Kerberos Security Integration
  + Load data directly into Hive/Hbase
  + Support for Accumulo
* Sqoop providess command line interface to the end users
* Can also be accessed using Java APIs
* Sqoop vs Flume vs HDFS in Hadoop
  + Sqoop data load is not event-driven
  + Flume data load can be driven by event
  + HDFS just stores data provided to it by whatsoever means
* Sqoop – in order to impact data from structured data sources, one has to use Sqoop only, because its connectors know how to interact with structured data sources and fetch data from them
* Flume – in order to load streaming data such as tweets generated on Twitter or log files of a web server, Flume should be used. Flume agents are build for fetching streaming data
* HDFS – has its own built-in shell commands to store data into it. HDFS cannot import streaming data

Resources referred:

<https://www.tutorialspoint.com/sqoop/index.htm>

<https://www.dezyre.com/hadoop-tutorial/hadoop-sqoop-tutorial>

https://www.dezyre.com/hadoop-tutorial/hadoop-sqoop-tutorial