**SQOOP**

Su **objetivo fundamental es transferir datos entre RDBMS y Hadoop** (HDFS)

Los **datos son importados** a HDFS **como text files** delimitados o SequenceFiles

Los ficheros son importados mediante MapReduce, como hemos comentado, por lo tanto se almacenan en formato part-\*.0\*

**Mysql -u root -p**

**Password: cloudera**

Ejemplo: importar una tabla llamada *empleados* de una bbdd llamada *personal*

**sqoop import --username user --password pass \ --connect**

**jdbc:mysql://database.example.com/personal --table empleados --target-dir XXX**

* La sintaxis básica corresponde a lo siguiente
  + sqoop tool-name [tool-options]
* Ejemplos de tool-name
  + import
  + import-all-tables
  + list-tables
* Ejemplos de tool-options
  + --connect
  + --username
  + --password

Query:

sqoop import [% argumentos %] **--query 'SELECT a.\*, b.\* FROM a JOIN b on (a.id == b.id) WHERE $CONDITIONS'** --target-dir /user/foo/joinresults

create database pruebadb;

create table tabla\_prueba (nombre varchar(30), edad int);

Insert into tabla\_prueba (“juan ”, 25);

Insert into tabla\_prueba (“pedro”, 11);

Select \* from tabla\_prueba;

hadoop fs -ls /user/hive/warehouse

Create table tabla\_prueba\_hive (nombre string, edad int);

Listar tablas

sqoop list-tables --connect jdbc:mysql://localhost/pruebadb --username root --password cloudera