Importing a single table:

sqoop import \

--connect jdbc:oracle:thin:@//hostname:port/service\_name \

--username your\_username \

--password your\_password \

--table table\_name \

--target-dir /path/to/target\_directory \

--as-textfile

1. Importing a subset of data with a custom query:

bash  
sqoop import \

--connect jdbc:oracle:thin:@//hostname:port/service\_name \

--username your\_username \

--password your\_password \

--query "SELECT \* FROM table\_name WHERE column\_name > some\_value AND \$CONDITIONS" \

--target-dir /path/to/target\_directory \

--as-textfile

In these examples:

* Replace **hostname**, **port**, and **service\_name** with the appropriate values for your Oracle database connection.
* Replace **your\_username** and **your\_password** with your Oracle credentials.
* Replace **table\_name** with the name of the table you want to import.
* Replace **/path/to/target\_directory** with the directory where you want to store the imported data.
* You can also customize the SQL query in the second example to import only the subset of data you need.

Sure, let's walk through an example step by step:

Let's say we want to import data from an Oracle database hosted on a server with the following details:

* Hostname: **example.com**
* Port: **1521**
* Service Name: **ORCL**
* Username: **my\_username**
* Password: **my\_password**

We want to import data from a table named **employee** into HDFS directory **/user/hadoop/employee\_data**.

The command for importing data using Sqoop would be:

sqoop import \

--connect jdbc:oracle:thin:@//example.com:1521/ORCL \

--username my\_username \

--password my\_password \

--table employee \

--target-dir /user/hadoop/employee\_data \

--as-textfile

* **sqoop import**: This is the Sqoop command to import data.
* **--connect**: Specifies the JDBC connection string to connect to the Oracle database.
* **--username**: Specifies the username to use when connecting to the Oracle database.
* **--password**: Specifies the password to use when connecting to the Oracle database.
* **--table**: Specifies the name of the table (**employee**) from which to import data.
* **--target-dir**: Specifies the HDFS directory where the imported data will be stored (**/user/hadoop/employee\_data**).
* **--as-textfile**: Specifies that the data should be stored as text files in HDFS.

When you run this command in your terminal or command prompt, Sqoop will connect to the Oracle database, retrieve data from the **employee** table, and import it into the specified HDFS directory **/user/hadoop/employee\_data**.

sqoop import --connect jdbc:oracle:thin:@//example.com:1521/ORCL --username my\_username --password my\_password --table employee --target-dir /user/hadoop/employee\_data --as-textfile

Exporting data using Sqoop involves transferring data from Hadoop, typically stored in HDFS, to a relational database. Below is an example of how you can export data from HDFS to an Oracle database using Sqoop:

sqoop export \

--connect jdbc:oracle:thin:@//hostname:port/service\_name \

--username your\_username \

--password your\_password \

--table target\_table \

--export-dir /path/to/source\_directory \

--input-fields-terminated-by ',' \

--input-lines-terminated-by '\n'

Exporting data using Sqoop involves transferring data from Hadoop, typically stored in HDFS, to a relational database. Below is an example of how you can export data from HDFS to an Oracle database using Sqoop:

bashCopy code

sqoop export \ --connect jdbc:oracle:thin:@//hostname:port/service\_name \ --username your\_username \ --password your\_password \ --table target\_table \ --export-dir /path/to/source\_directory \ --input-fields-terminated-by ',' \ --input-lines-terminated-by '\n'

Explanation of the parameters used:

* **sqoop export**: This is the Sqoop command used to export data.
* **--connect**: Specifies the JDBC connection string to connect to the Oracle database.
* **--username**: Specifies the username to use when connecting to the Oracle database.
* **--password**: Specifies the password to use when connecting to the Oracle database.
* **--table**: Specifies the name of the target table in the Oracle database where the data will be exported.
* **--export-dir**: Specifies the HDFS directory containing the data to be exported.
* **--input-fields-terminated-by**: Specifies the field delimiter used in the input data (source data in HDFS). Replace ',' with the appropriate delimiter if it's different.
* **--input-lines-terminated-by**: Specifies the line terminator used in the input data (source data in HDFS). Replace '\n' with the appropriate line terminator if it's different.

Ensure to replace **hostname**, **port**, **service\_name**, **your\_username**, **your\_password**, **target\_table**, and **/path/to/source\_directory** with the appropriate values for your Oracle database and HDFS setup.

This command will export the data from the specified HDFS directory to the target Oracle database table. Make sure that the structure of the data matches the schema of the target table in the Oracle database. Additionally, ensure that you have the necessary permissions to write to the Oracle database from Sqoop.

sqoop export \

--connect jdbc:oracle:thin:@//example.com:1521/ORCL \

--username oracle\_username \

--password oracle\_password \

--table target\_table \

--export-dir /user/hadoop/source\_data \

--input-fields-terminated-by ',' \

--input-lines-terminated-by '\n'

Explanation:

sqoop export: This is the Sqoop command to export data.

--connect: Specifies the JDBC connection string to connect to the Oracle database.

--username: Specifies the username to use when connecting to the Oracle database.

--password: Specifies the password to use when connecting to the Oracle database.

--table: Specifies the name of the target table in the Oracle database where the data will be exported (target\_table).

--export-dir: Specifies the HDFS directory containing the data to be exported (/user/hadoop/source\_data).

--input-fields-terminated-by: Specifies the field delimiter used in the input data (source data in HDFS). Here, we assume the data is comma-separated (',').

--input-lines-terminated-by: Specifies the line terminator used in the input data (source data in HDFS). We assume it's newline ('\n').

When you run this command, Sqoop will read the data from the specified HDFS directory (/user/hadoop/source\_data), format it according to the specified delimiters, and export it to the target\_table in the Oracle database. Make sure the Oracle database is accessible from the machine where you're running Sqoop, and you have the necessary permissions to write to the target\_table.

Sqoop provides an option to import a single table from an Oracle database directly into Hive. You can use the **--table** parameter along with **--hive-table** to achieve this. Here's an example command:

sqoop import \

--connect jdbc:oracle:thin:@//example.com:1521/ORCL \

--username oracle\_username \

--password oracle\_password \

--table your\_table \

--hive-import \

--hive-table hive\_table\_name \

--hive-overwrite

To import all tables from an Oracle database into HDFS using Sqoop, you can use the import-all-tables command. This command imports all tables from the specified database into HDFS. Here's an example command:

sqoop import-all-tables \

--connect jdbc:oracle:thin:@//example.com:1521/ORCL \

--username oracle\_username \

--password oracle\_password \

--warehouse-dir /user/hadoop/oracle\_tables \

--as-textfile

Explanation:

* **sqoop import-all-tables**: This Sqoop command imports all tables from the specified database into HDFS.
* **--connect**: Specifies the JDBC connection string to connect to the Oracle database.
* **--username**: Specifies the username to use when connecting to the Oracle database.
* **--password**: Specifies the password to use when connecting to the Oracle database.
* **--warehouse-dir**: Specifies the base directory in HDFS where Sqoop will store the imported tables. Each table will have its directory under this base directory.
* **--as-textfile**: Specifies that the data should be stored as text files in HDFS.

When you run this command, Sqoop will import all tables from the Oracle database into HDFS, with each table's data stored in a separate directory under the specified base directory (**/user/hadoop/oracle\_tables** in this example).