800+ Q&As | Logout | Contact

Java-Success.com

Prepare to fast-track, choose & go places with 800+ Java & Big Data Q&As with lots of code & diagrams.

search here … Go

Home Why? ▾ 300+ Java FAQs ▾ 300+ Big Data FAQs ▾ Courses ▾

Membership ▾ Your Career ▾

Home › bigdata-success.com › Tutorials - Big Data › TUT - Cloudera on Docker › 17: Docker Tutorial: sqoop import – on Cloudera quickstart

17: Docker Tutorial: sqoop import – on Cloudera quickstart

Posted on June 5, 2019

This extends Docker Tutorial: BigData on Cloudera quickstart via Docker.

Step 1: Run the container on a command line.

```
~/projects/docker-hadoop]$ docker run --hostname=qu
--privileged=true -t -i -v /Users/arulkumarankumar
--publish-all=true -p 8888:8888 -p 80:80 -p 7180:71
```

Connect to MySQL DB

300+ Java Interview FAQs

300+ Java FAQs 🔥

16+ Java Key Areas Q&As

150+ Java Architect FAQs

80+ Java Code Quality Q&As

150+ Java Coding Q&As

300+ Big Data Interview FAQs

300+ Big Data FAQs 🔥

Tutorials - Big Data

TUT - ⏯ Starting Big Data

TUT - Starting Spark & Scala

Step 2: Hive & Impala store the metadata in a database like MySQL. The Hive **metastore service** connects to the metastore Database to store metadata. You can connect to the mysql RDBMS as shown below:

```
[root@quickstart /]# mysql -uroot -pcloudera
mysql> show databases;
+--------------------+
| Database           |
+--------------------+
| information_schema |
| cm                 |
| firehose           |
| hue                |
| metastore          |
| mysql              |
| nav                |
| navms              |
| oozie              |
| retail_db          |
| rman               |
| sentry             |
+--------------------+
12 rows in set (0.01 sec)


```

Create a DB, table & insert rows

Step 3: Create a new database and switch to that database.

```
mysql> create database test;
mysql> use test;
Database changed
mysql>

```

Step 4: Create a new table.

```
mysql> CREATE TABLE employee (firstname varchar(10
surname varchar(100) not null,
```

TUT - Starting with Python
TUT - Kafka
TUT - Pig
TUT - Apache Storm
TUT - Spark Scala on Zeppelin
TUT - Cloudera
TUT - Cloudera on Docker
TUT - File Formats
TUT - Spark on Docker
TUT - Flume
TUT - Hadoop (HDFS)
TUT - HBase (NoSQL)
TUT - Hive (SQL)
TUT - Hadoop & Spark
TUT - MapReduce
TUT - Spark and Scala
TUT - Spark & Java
TUT - PySpark on Databricks
TUT - Zookeeper

800+ Java Interview Q&As

300+ Core Java Q&As
300+ Enterprise Java Q&As
150+ Java Frameworks Q&As
120+ Companion Tech Q&As
Tutorials - Enterprise Java

```
department varchar(100) not null
);

```

Step 5: Insert records.

```
mysql> mysql> INSERT INTO employee (firstname, sur
mysql> INSERT INTO employee (firstname, surname, d
mysql> INSERT INTO employee (firstname, surname, d

```

```
 mysql> SELECT * FROM employee;
+-----------+---------+------------+
| firstname | surname | department |
+-----------+---------+------------+
| John      | Samuel  | IT         |
| Peter     | Smith   | Finance    |
| Sean      | Mendis  | Marketing  |
+-----------+---------+------------+
3 rows in set (0.00 sec)


```

import into HDFS from mysql

Step 6: Import data into HDFS from MySQL RDBMs.

```
root@quickstart /]# sqoop import --connect jdbc:my
> --username root --password cloudera \
> --table employee \
> --where "firstname is NOT NULL" -m 1 \
> --target-dir /user/root/employee

```

Alternatively:

Create a file **sqoop-import.txt**

```
import
--connect
jdbc:mysql://quickstart.cloudera/test
--username
root
```

```
--password
cloudera

```

```
[root@quickstart /]# sqoop --options-file sqoop-im
> --where "firstname is NOT NULL" -m 1 \
> --target-dir /user/root/employee

```

Step 7: Check the HDFS.

```
[root@quickstart /]# hdfs dfs -ls /user/root/emplo
Found 2 items
-rw-r--r--   1 root supergroup          0 2019-06-
-rw-r--r--   1 root supergroup         57 2019-06-
[root@quickstart /]#

```

```
[root@quickstart /]# hdfs dfs -cat /user/root/empl
John,Samuel,IT
Peter,Smith,Finance
Sean,Mendis,Marketing

```

What is next?

Sqoop export to export data from HDFS to MySql.

‹ 16: Docker Tutorial: Apache Spark (spark-shell) & parquet-tools – csv to parquet on Cloudera quickstart

18: Docker Tutorial: sqoop export – on Cloudera quickstart ›

Disclaimer

The contents in this Java-Success are copyrighted and from EmpoweringTech pty ltd. The EmpoweringTech pty ltd has the right to correct or enhance the current content without any prior notice. These are general advice only, and one needs to take his/her own circumstances into consideration. The EmpoweringTech pty ltd will not be held liable for any damages caused or alleged to be caused either directly or indirectly by these materials and resources. Any trademarked names or labels used in this blog remain the property of their respective trademark owners. Links to external sites do not imply endorsement of the linked-to sites. Privacy Policy