Sqoop

Hdfs文件与关系性数据库之间的数据迁移

# 一、作用

|  |
| --- |
| Hdfs文件与关系性数据库之间的数据迁移 |

# 二、安装

|  |
| --- |
| 1、解压  2、重命名并且配置  mv sqoop-env-template.sh sqoop-env.sh  vim sqoop-env.sh 新增如下配置（指定hadoop，hive的安装目录）  export HADOOP\_COMMON\_HOME=/opt/module/hadoop-2.7.2/  export HADOOP\_MAPRED\_HOME=/opt/module/hadoop-2.7.2/  export HIVE\_HOME=/opt/module/hive-1.2.1  3、拷贝数据库相应的驱动到sqoop/lib下  4、验证sqoop与数据库的连接  #查看数据库 同show databases  ./sqoop-list-databases --connect jdbc:mysql://hadoop101/hive --username root --password 123456  #查看表 同show tables  ./sqoop-list-tables --connect jdbc:mysql://hadoop101/hive --username root --password 123456 |

# 三、操作

## 1、将数据库个导入到HDFS

|  |
| --- |
| Sqoop import \  --connect jdbc:mysql://hadoop101/test \  --username root \  --password 123456 \  --target-dir /queryresult \  --fields-terminated-by ‘,’ \  #hdfs是分块的，指定某个字段作为分块的原则 –-m 为1时，可不指定  --split-by id \  --table test #指定导入的表  --m 2 #指定mapreduce程序的个数  --hive-import #导入hive映射。可选  --where ‘city = cq‘ #满足条件的数据导入 可选  --query ‘sql‘ #将查询出来的数据插入 可选 |
|  |

## 2、将HDFS导入到数据库表

|  |
| --- |
| Sqoop export\  --connect jdbc:mysql://hadoop101/test \  --username root \  --password 123456 \  --export-dir /queryresult \  --input-fields-terminated-by ‘,’ \  --table test #数据库中的表 |
|  |