sudo docker run --hostname=sparkslave1 --net infs3208 --ip 172.45.0.3 -ti infs3208/slaveenv /bin/bash

bash /root/init.sh

source /etc/profile

sudo docker run --hostname=sparkslave2 --net infs3208 --ip 172.45.0.4 -ti infs3208/slaveenv /bin/bash

bash /root/init.sh

source /etc/profile

sudo docker run --hostname=sparkmaster --net infs3208 --ip 172.45.0.2 -p 8888:8888 -p 50070:50070 -p 8088:8088 -p 8042:8042 -p 4040:4040 -p 3307:3306 -ti -v ~/ccproject:/root/infs3208 infs3208/masterenv /bin/bashinfs3208/masterenv /bin/bash

source /etc/profile

bash /root/init.sh

然后ctrl c来退出

service mysql start

cd /root/spark/jars

wget <https://storage.googleapis.com/infs3208/dockerenv/mysql-connector-java-5.1.48-bin.jar>

chmod 777 mysql-connector-java-5.1.48-bin.jar

cd /root/spark/sbin

stop-all.sh

start-all.sh

• Now we will copy the file “demo.txt” from master docker to HDFS with a new name: “demo.hdfs.txt”.

cd infs3208/data

hdfs dfs –copyFromLocal ~/infs3208/data/librarydata.json /test/librarydata.hdfs.json

查看copyfrom的文件是否在：

hdfs dfs -ls /test

重新连接notebook：

bash /root/init.sh

import org.apache.spark.sql.SQLContext

import java.util.Properties

import java.sql.DriverManager

import java.sql.Connection

import org.apache.spark.sql.{Row, SparkSession}

import org.apache.spark.sql.types.\_

import scala.collection.mutable

import java.text.SimpleDateFormat

val spark = new SparkSession.Builder().appName("studentmangement").getOrCreate()

val sc = spark.sparkContext

var url = "jdbc:mysql://localhost:3306/hive"

val sqlContext = new SQLContext(sc)

val properties = new Properties()

properties.setProperty("user","hive")

properties.setProperty("password","hive")

properties.setProperty("useSSL","false")

import spark.sqlContext.implicits.\_

import org.apache.spark.sql.functions.\_

import org.apache.spark.mllib.linalg.Vectors

import org.apache.spark.mllib.stat.Statistics

var tableName = "ROLES"

val dataFrame = sqlContext.read.jdbc(url, tableName,properties)

dataFrame.select("ROLE\_ID","ROLE\_NAME").show()

再次退出notebook：

mysql -u root

create user 'project'@'%' identified by 'project';

create schema project default character set utf8 collate utf8\_general\_ci;

GRANT ALL PRIVILEGES ON project.\* TO 'project'@'%';

flush privileges ;

Create user table and comment table:

sql>

CREATE TABLE `project`.`comment` ( `id` INT(10) NOT NULL PRIMARY KEY, `user` VARCHAR(30) NOT NULL , `content` VARCHAR(100) NOT NULL , `time` DATETIME(0) NOT NULL ) ENGINE = MyISAM;

CREATE TABLE `project`.`user` ( `username` VARCHAR(30) NOT NULL , `password` VARCHAR(30) NOT NULL ) ENGINE = MyISAM;

> set up PHP connect to Spark SQL in GCP:

> set SQL remote access authority

1. Start SQL by entering 'service mysql start' in the sparkmaster,

2. Then enter into a document using 'vim /etc/mysql/mysql.conf.d/mysqld.cnf',

3. Find out the bind-address(127.0.0.1) and change it to '0.0.0.0'.

4. Press 'i' to edit the document, then press 'esc' and enter ':wq' and ENTER, it will save the document and exit.

重新连接notebook：

bash /root/init.sh