

大数据技术-第九章：Sqoop数据迁移 配置Sqoop环境



CONTENTS

01. 下载和解压Sqoop

02. 配置Sqoop环境

03. 启动Sqoop



01

下载和解压Sqoop



» 下载和解压Sqoop

Sqoop相关发行版本可以通过官网<https://mirror-hk.koddos.net/apache/sqoop/>来获取。

安装Sqoop组件需要与Hadoop环境适配。使用root用户在Master节点上进行部署，将/opt/software/sqoop-1.4.7.bin__hadoop-2.6.0.tar.gz压缩包解压到/usr/local/src目录下。

```
[root@master ~]#tar -zxvf sqoop-1.4.7.bin__hadoop-2.6.0.tar.gz -C /usr/local/src
```

将解压后生成的sqoop-1.4.7.bin__hadoop-2.6.0文件夹更名为sqoop。

```
[root@master ~]#cd /usr/local/src/  
[root@master src]#mv ./sqoop-1.4.7.bin__hadoop-2.6.0 sqoop
```

02

配置Sqoop环境



配置Sqoop环境

(1) 创建Sqoop的配置文件sqoop-env.sh。

复制sqoop-env-template.sh模板，并将模板重命名为sqoop-env.sh。

```
[root@master ~]cd /usr/local/src/sqoop/conf/  
[root@master conf]# cp sqoop-env-template.sh sqoop-env.sh
```

(2) 修改sqoop-env.sh文件，添加Hadoop、Hbase、Hive等组件的安装路径。

注意，下面各组件的安装路径需要与实际环境中的安装路径保持一致。

```
[root@master conf]# vi sqoop-env.sh  
  
export HADOOP_COMMON_HOME=/usr/local/src/hadoop  
export HADOOP_MAPRED_HOME=/usr/local/src/hadoop  
export HBASE_HOME=/usr/local/src/hbase  
export HIVE_HOME=/usr/local/src/hive
```

配置Sqoop环境

(3) 配置Linux系统环境变量，添加Sqoop组件的路径。

```
vi /etc/profile

#在文件末尾添加
# set sqoop environment
export SQOOP_HOME=/usr/local/src/sqoop
export PATH=$PATH:$SQOOP_HOME/bin
export CLASSPATH=$CLASSPATH:$SQOOP_HOME/lib

# 使系统环境变量生效
[root@master conf]# source /etc/profile
```

(4) 为了使Sqoop能够连接MySQL数据库，需要将/opt/software/mysql-connector-java-5.1.48.jar文件放入sqoop的lib目录中。该jar文件的版本需要与MySQL数据库的版本相对应，否则Sqoop导入数据时会报错。（mysql-connector-java-5.1.48.jar对应的是MySQL 5.7版本）

```
[root@master ~] cp /opt/software/mysql-connector-java-5.1.48.jar /usr/local/src/sqoop/lib/
```

03

启动Sqoop



» 启动Sqoop

(1) 执行Sqoop前需要先启动Hadoop集群。

在master节点切换到hadoop用户执行start-all.sh命令启动Hadoop集群。

```
[root@master ~] su - hadoop  
[root@master ~] start-all.sh
```

(2) 检查Hadoop集群的运行状态。

```
[hadoop@master ~]$ jps  
1457 NameNode  
1795 ResourceManager  
2060 Jps  
1646 SecondaryNameNode
```

(3) 测试Sqoop是否能够正常连接MySQL数据库。

```
[hadoop@master ~]
```

```
sqoop list-databases --connect jdbc:mysql://127.0.0.1:3306/ --username root -P    #
```

Sqoop连接MySQL数据库

Warning: /home/hadoop/sqoop/./hcatalog does not exist! HCatalog jobs will fail.

Please set \$HCAT_HOME to the root of your HCatalog installation.

Warning: /home/hadoop/sqoop/./accumulo does not exist! Accumulo imports will fail.

Please set \$ACCUMULO_HOME to the root of your Accumulo installation.

Warning: /home/hadoop/sqoop/./zookeeper does not exist! Accumulo imports will fail.

Please set \$ZOOKEEPER_HOME to the root of your Zookeeper installation.

Turing AI 万维图灵 | 大数据系列课程

大数据

BIG
DATA

智 / 能 / 科 / 技 放 / 眼 / 未 / 来

