



0.1. 安装 Sqoop

第一节 简介

Sqoop 是一种方便数据导入和导出的工具，方便将关系型数据库如 MySQL、Oracle 中的数据导入到 HDFS、Hive 或者是 HBase 中或从后者中导出。也是一个比较重要且实用的工具。

第二节 安装配置

sqoop-env.sh

```
#Set path to where bin/hadoop is available
export HADOOP_COMMON_HOME=/opt/hadoop
#Set path to where hadoop-*-core.jar is available
export HADOOP_MAPRED_HOME=/opt/hadoop
#set the path to where bin/hbase is available
export HBASE_HOME=/home/zhangyu/hbase
#Set the path to where bin/hive is available
export HIVE_HOME=/home/zhangyu/hive
export HCATA_HOME=/home/zhangyu/hive/hcatalog
#Set the path for where zookeeper config dir is
export ZOOCFGDIR=/home/zhangyu/zookeeper/conf
```

第三节 测试和运用

将配置好的sqoop/文件夹分发到各台机器上，即可开始使用。

可以输入sqoop help查看命令帮助。

Sqoop 的基本用法如下：

use-sqoop.sh

```
# 列出数据库
$ sqoop list-databases \
> connect jdbc:mysql://master:3306/ \
> -username root \
> -password password
# 列出数据库中的表
```

```
$ sqoop list-tables \  
> connect jdbc:mysql://master:3306/hive \  
> -username root \  
> -password password  
# 导入数据到HDFS  
$ sqoop import \  
> connect jdbc:mysql://master:3306/mysqoop \  
> -username root \  
> -password password \  
> -table student \  
> -m 1 \  
> -target-dir /mysqoop/test/mystu  
# 导入数据到HBase  
$ sqoop import \  
> connect jdbc:mysql://master:3306/mysqoop \  
> -username root \  
> -password password \  
> -table student \  
> -hbase-create-table \  
> -hbase-table mystu \  
> -column-family mycf \  
> -hbase-row-key id
```

使用 Sqoop 时可能会报一些错误,主要是 MySQL 数据库连接的错误,要提前将[mysql-connector-xxx.jar](#) 文件拷贝到[sqoop/lib](#)文件夹下,而且驱动的版本也要合适,否则也会报错。