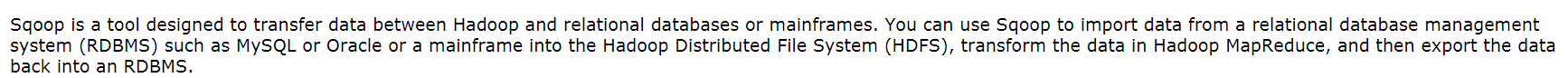
# Sqoop

* 概述
* 预备环境

Hadoop

Zookeeper

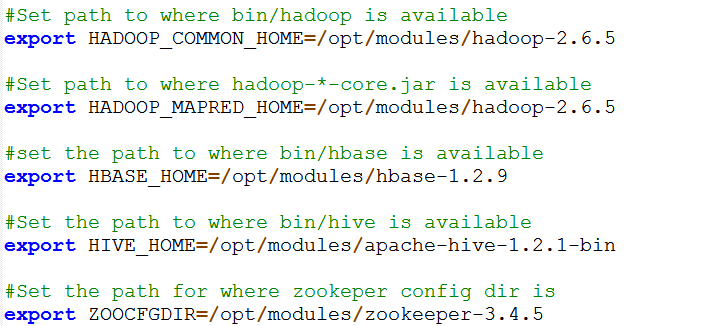
MySQL

Hive

HBase

* 安装

编辑sqoop-env.sh



在.bash\_profile中配置HADOOP\_CLASSPATH



拷贝hive/conf/hive-site.xml到sqoop/conf下

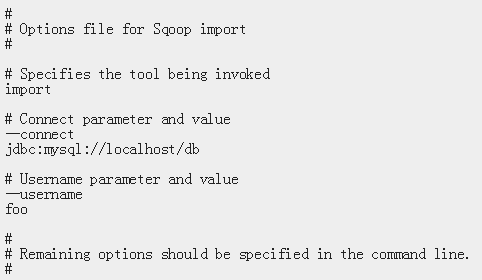
拷贝相应数据库驱动包到lib目录下

* 基本操作

**sqoop-import**

表查询 

脚本查询





指定存储目录



设置map的数量



导入条件

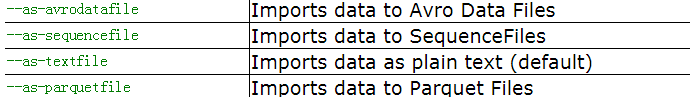




自定义查询语句



导入文件格式



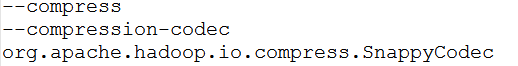
启用压缩



指定压缩方式



例：



设置导入内容分割方式



删除已存在目标目录

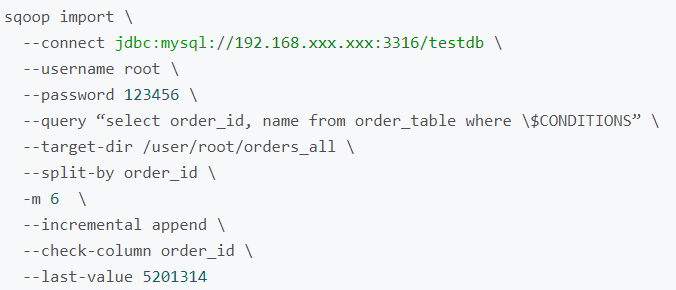


导入数据到Hive

**在hive中创建相应表**

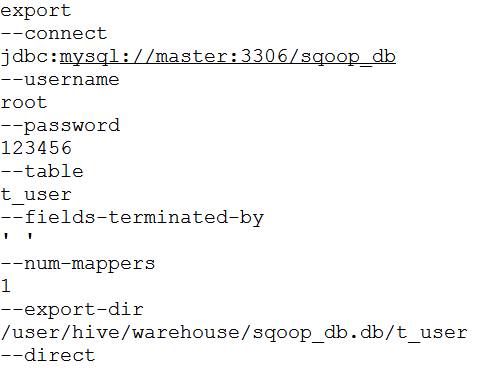


增量数据导入



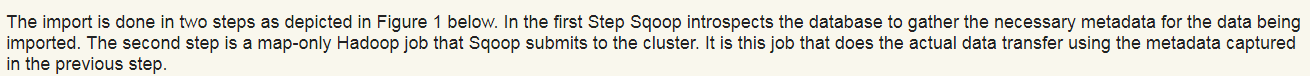
**sqoop-export**

Hive导出

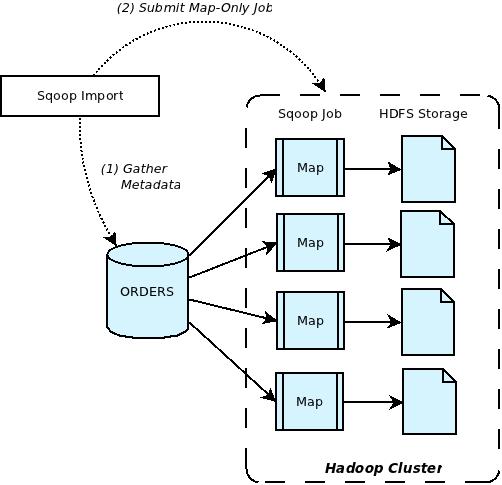


* 架构

**<https://blogs.apache.org/sqoop/entry/apache_sqoop_overview>**



**Import Data**



**Export Data**

