安装：

环境：centos7 、sqoop-1.4.6.bin__hadoop-2.0.4、mariadb
上传压缩包,解压文件:tar zxvf /var/www/html/sqoop-1.4.6.bin__hadoop-2.0.4-alpha.tar.gz -C /usr/local/
修改文件夹名称：mv /usr/local/sqoop-1.4.6.bin__hadoop-2.0.4-alpha/ /usr/local/sqoop
配置环境变量：vim /etc/profie

```
export JAVA_HOME=/usr/local/jdk
export JRE_HOME=/usr/local/jdk/jre
export MAVEN_HOME=/usr/local/maven
export ZOOKEEPER_HOME=/usr/local/zookeeper
export HADOOP_HOME=/usr/local/hadoop
export HIVE_HOME=/usr/local/hive
export HBASE_HOME=/usr/local/hbase
export SQOOP_HOME=/usr/local/sqoop
export ROCKETMQ_HOME=/usr/local/incubator-rocketmq/target/apache-rocketmq-all
export
PATH=$PATH:$JAVA_HOME/bin:$JRE_HOME/bin:$MAVEN_HOME/bin:$ZOOKEEPER_HOME/bin:$HADOOP_HOME/bin:$HADOOP_HOME/sbin:$$HBASE_HOME/bin:$HIVE_HOME/bin:$SQOOP_HOME/bin
```

是配置立即生效：source /etc/profile
将数据库的驱动程序复制到 sqoop 目录 lib 中： cp /var/www/html/mariadb-java-client-1.5.9.jar /usr/local/sqoop/lib/

mysql 数据导入 hdfs

建立测试数据表

```
DROP DATABASE bigdata;
CREATE DATABASE bigdata CHARACTER SET UTF8 ;
USE bigdata;
CREATE TABLE student(
sid INT AUTO_INCREMENT ,
name VARCHAR(50) ,
score INT,
CONSTRAINT pk_mid PRIMARY KEY(sid)
) ;
INSERT INTO student(name,score ) VALUES ('xiaoming',86) ;
INSERT INTO student(name,score ) VALUES ('xiaohong',83) ;
INSERT INTO student(name,score ) VALUES ('xiaolan',78);
```

复制数据库连接驱动到 sqoop lib 目录。

sqoop 连接数据库测试

```
sqoop list-tables --connect jdbc:mysql://192.168.28.130:3306/bigdata --username root --password 123456 --driver com.mysql.jdbc.Driver
```

数据导入到 hdfs

```
sqoop import --connect jdbc:mysql://192.168.28.130:3306/bigdata --username root --password 123456 --driver com.mysql.jdbc.Driver --table student --fields-terminated-by ',' --null-string '**' -m 1

17/04/15 15:05:07 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar
17/04/15 15:05:07 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM student AS t WHERE 1=0
17/04/15 15:05:08 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps
17/04/15 15:05:08 INFO client.RMProxy: Connecting to ResourceManager at centos01/192.168.28.130:8050
17/04/15 15:05:14 INFO db.DBInputFormat: Using read commited transaction isolation
17/04/15 15:05:14 INFO mapreduce.JobSubmitter: number of splits:1
17/04/15 15:05:14 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1492237742384_0001
17/04/15 15:05:15 INFO impl.YarnClientImpl: Submitted application application_1492237742384_0001
17/04/15 15:05:15 INFO mapreduce.Job: The url to track the job: http://centos01:8088/proxy/application_1492237742384_0001/
17/04/15 15:05:15 INFO mapreduce.Job: Running job: job_1492237742384_0001
17/04/15 15:05:24 INFO mapreduce.Job: Job job_1492237742384_0001 running in uber mode : false
17/04/15 15:05:24 INFO mapreduce.Job: map 0% reduce 0%
17/04/15 15:05:30 INFO mapreduce.Job: map 100% reduce 0%
17/04/15 15:05:32 INFO mapreduce.Job: Job job_1492237742384_0001 completed successfully
17/04/15 15:05:32 INFO mapreduce.Job: Counters: 30
File System Counters
FILE: Number of bytes read=0
FILE: Number of bytes written=136963
FILE: Number of read operations=0
FILE: Number of large read operations=0
FILE: Number of write operations=0
HDFS: Number of bytes read=87
HDFS: Number of bytes written=41
HDFS: Number of read operations=4
HDFS: Number of large read operations=0
```

```
HDFS: Number of write operations=2
Job Counters
Launched map tasks=1
Other local map tasks=1
Total time spent by all maps in occupied slots (ms)=2831
Total time spent by all reduces in occupied slots (ms)=0
Total time spent by all map tasks (ms)=2831
Total vcore-milliseconds taken by all map tasks=2831
Total megabyte-milliseconds taken by all map tasks=2898944
Map-Reduce Framework
Map input records=3
Map output records=3
Input split bytes=87
Spilled Records=0
Failed Shuffles=0
Merged Map outputs=0
GC time elapsed (ms)=62
CPU time spent (ms)=770
Physical memory (bytes) snapshot=170479616
Virtual memory (bytes) snapshot=2100588544
Total committed heap usage (bytes)=95944704
File Input Format Counters
Bytes Read=0
File Output Format Counters
Bytes Written=41
17/04/15 15:05:32 INFO mapreduce.ImportJobBase: Transferred 41 bytes in 24.0033
seconds (1.7081 bytes/sec)
17/04/15 15:05:32 INFO mapreduce.ImportJobBase: Retrieved 3 records.
```

从关系型数据到 hive

hive 是建立在 hdfs 上，所以也可将数据导入到 hive 中。

1、在 hive 创建对应数据表。

```
sqoop create-hive-table --connect jdbc:mysql://192.168.28.130:3306/bigdata --
username root --password 123456 --table student --hive-table hive_student --
fields-terminated-by ','
```

2、将数据导入 hive。

```
sqoop import --connect jdbc:mysql://192.168.28.130:3306/bigdata --username root
--password 123456 --table student --hive-table hive_student --hive-import --
fields-terminated-by ','
```

从 HDFS、Hive 数据导出关系型数据库。

新建一个测试文档 student.txt，并上传到 hdfs。

4,xya,22

5,slw,22

6,zcj,22

1、将数据导出到 mysql:

```
sqoop export --connect jdbc:mysql://192.168.28.130:3306/bigdata --username root --password 123456 --table student --export-dir '/input/student.txt' --input-fields-terminated-by ','
```

2、将数据导入到 hive:

```
sqoop export --connect jdbc:mysql://192.168.28.130:3306/bigdata --username root --password 123456 --table student --export-dir '/user/hive/warehouse/hive_student' --input-fields-terminated-by ','
```