**Sqoop1 설치하기**

**Sqoop 이란?**

Apache Hadoop 과 관계형 데이터베이스와 같은 구조화 된 데이터 저장소 간에 대량 데이터를 효율적으로 전송하기 위해 설계된 도구입니다.

Sqoop 사이트 : <https://sqoop.apache.org/>

다운로드 경로 : <http://apache.tt.co.kr/sqoop/1.4.7/sqoop-1.4.7.bin__hadoop-2.6.0.tar.gz>

**Sqoop 다운로드 (다운로드 경로는 home)**

wget http://apache.tt.co.kr/sqoop/1.4.7/sqoop-1.4.7.bin\_\_hadoop-2.6.0.tar.gz

**압축해제**

tar -zxf sqoop-1.4.7.bin\_\_hadoop-2.6.0.tar.gz

**바로가기 생성**

ln-s sqoop-1.4.7.bin\_\_hadoop-2.6.0 sqoop

**환경변수 설정 (내용 추가.)**

sudo gedit ~/.bashrc

|  |
| --- |
| export SQOOP\_HOME=/home/yunhyuck(계정이름)/sqoop  export PATH=$PATH:$SQOOP\_HOME/bin |

**환경변수 적용**

source ~/.bashrc

**Sqoop 환경 설정 (/sqoop/sqoop-1.4.7)**

**폴더 이름 변경**

cp conf/sqoop-env-template.sh conf/sqoop-env.sh

|  |
| --- |
| #Set path to where bin/hadoop is available  export HADOOP\_COMMON\_HOME=/home/yunhyuck(계정이름)/sqoop  #Set path to where hadoop-\*-core.jar is available  export HADOOP\_MAPRED\_HOME=/home/yunhyuck(계정이름)/sqoop |

**Sqoop 실행에 필요한 파일.**

**\* mysql-connector-java-5.1.38-bin.jar 파일을 /home/yunhyuck/sqoop/sqoop-1.4.7/lib 폴더 안에 존재해야 함.**

wget <https://dev.mysql.com/get/Downloads/Connector-J/mysql-connector-java-5.1.38.tar.gz>

tar -zxvf mysql-connector-java-5.1.38.tar.gz

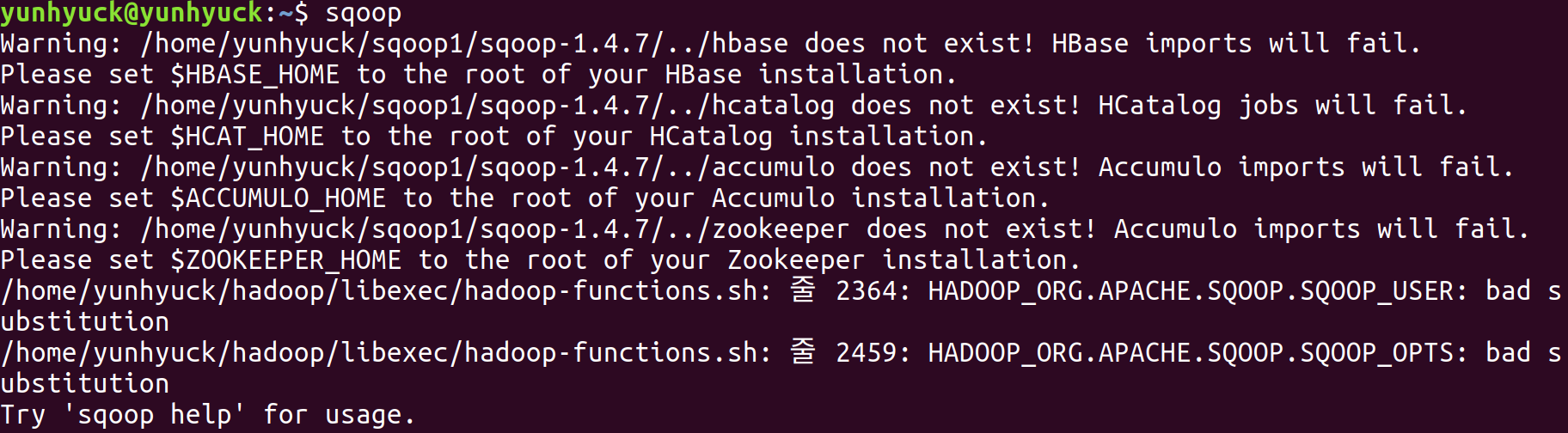
(압축 파일내에 있는 mysql-connector-java-5.1.38-bin 파일 이동)

**\* Lang4j 오류 발생을 방지하기 위해, commons-lang-2.6.jar 파일 역시 /home/yunhyuck/sqoop/sqoop-1.4.7/lib 폴더 안에 존재해야 함.**

**wget** [**http://www.java2s.com/Code/JarDownload/org.apache.commons/org.apache.commons.lang-2.6.jar.zip**](http://www.java2s.com/Code/JarDownload/org.apache.commons/org.apache.commons.lang-2.6.jar.zip)

**스쿱 실행**

|  |
| --- |
| yunhyuck@yunhyuck:~$ **sqoop**  Warning: /home/yunhyuck/sqoop1/sqoop-1.4.7/../hbase does not exist! HBase imports will fail.  Please set $HBASE\_HOME to the root of your HBase installation.  Warning: /home/yunhyuck/sqoop1/sqoop-1.4.7/../hcatalog does not exist! HCatalog jobs will fail.  Please set $HCAT\_HOME to the root of your HCatalog installation.  Warning: /home/yunhyuck/sqoop1/sqoop-1.4.7/../accumulo does not exist! Accumulo imports will fail.  Please set $ACCUMULO\_HOME to the root of your Accumulo installation.  Warning: /home/yunhyuck/sqoop1/sqoop-1.4.7/../zookeeper does not exist! Accumulo imports will fail.  Please set $ZOOKEEPER\_HOME to the root of your Zookeeper installation.  /home/yunhyuck/hadoop/libexec/hadoop-functions.sh: 줄 2364: HADOOP\_ORG.APACHE.SQOOP.SQOOP\_USER: bad substitution  /home/yunhyuck/hadoop/libexec/hadoop-functions.sh: 줄 2459: HADOOP\_ORG.APACHE.SQOOP.SQOOP\_OPTS: bad substitution  Try 'sqoop help' for usage. |

경고에 대해서는 각각, hcatalog, Accumulo, Zookeeeper 설정을 안 해주었기 때문에 발생한다.

**Sqoop 사용하기 (MariaDB에서 HDFS로 파일 전송.)**

**sqoop import --connect jdbc:mysql://localhost/mysql --username root --password xxxx --table test --columns a --target-dir** **hdfs://localhost:9000/user/yunhyuck/test4 -m 1**

**--connect 명령어 줄에는 JDBC 주소를 입력한다. MariaDB에서는 mysql 기반이기 때문에 jdbc:mariadb 입력하는 것을 주의 한다.**

**jdbc:mysql://localhost/mysql(테이블 )명**

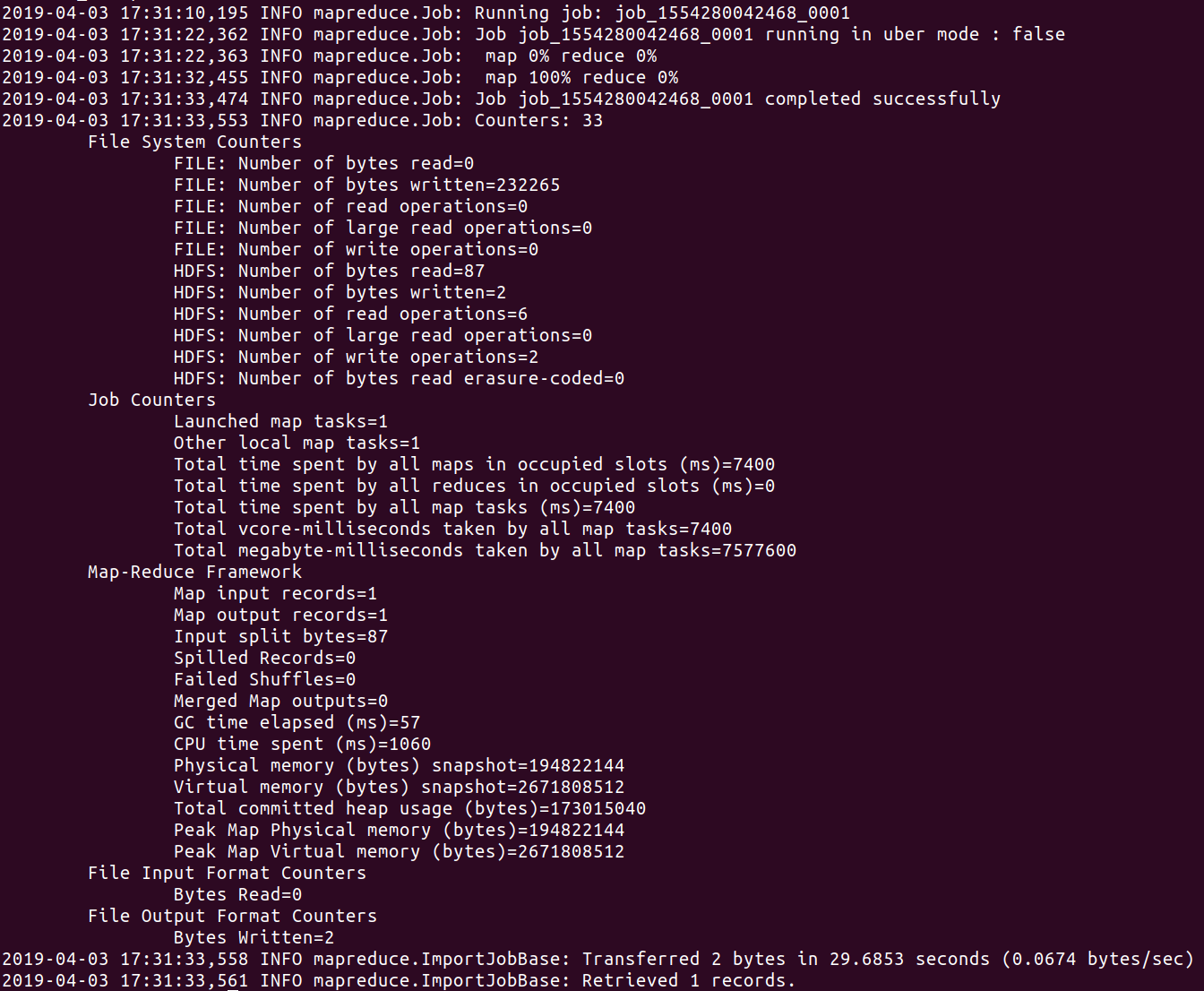
**--username root(DB 계정명)**

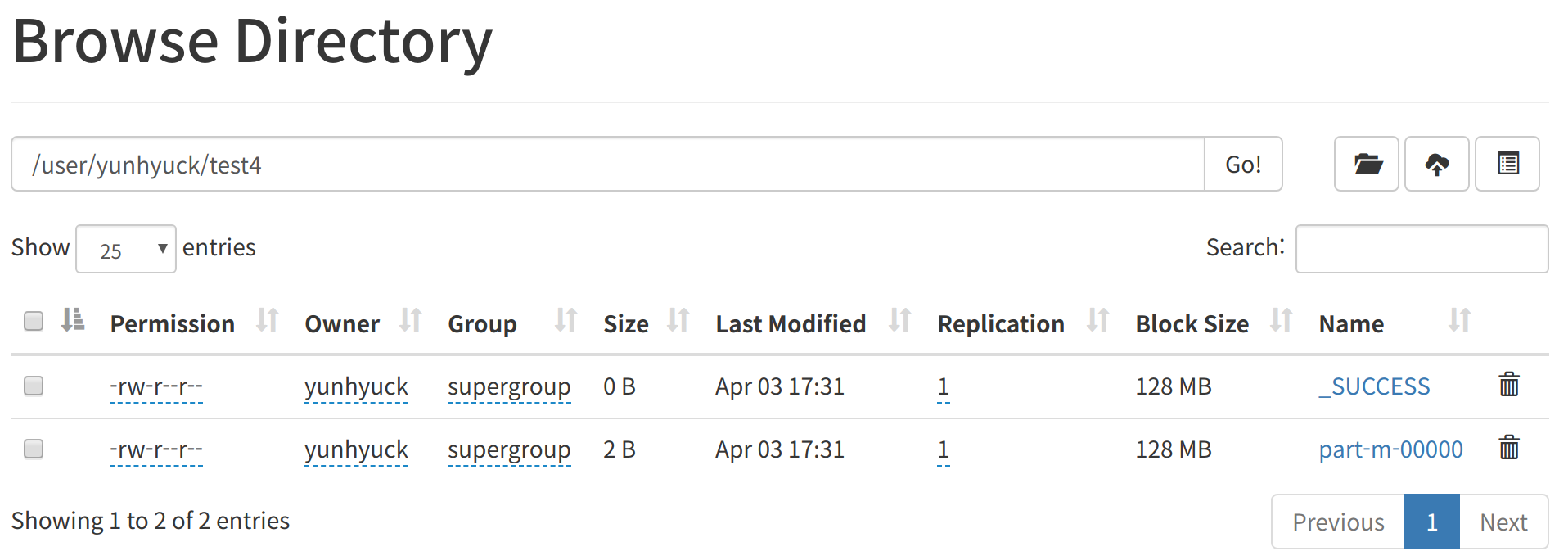
**--password xxxx(DB 패스워드)**

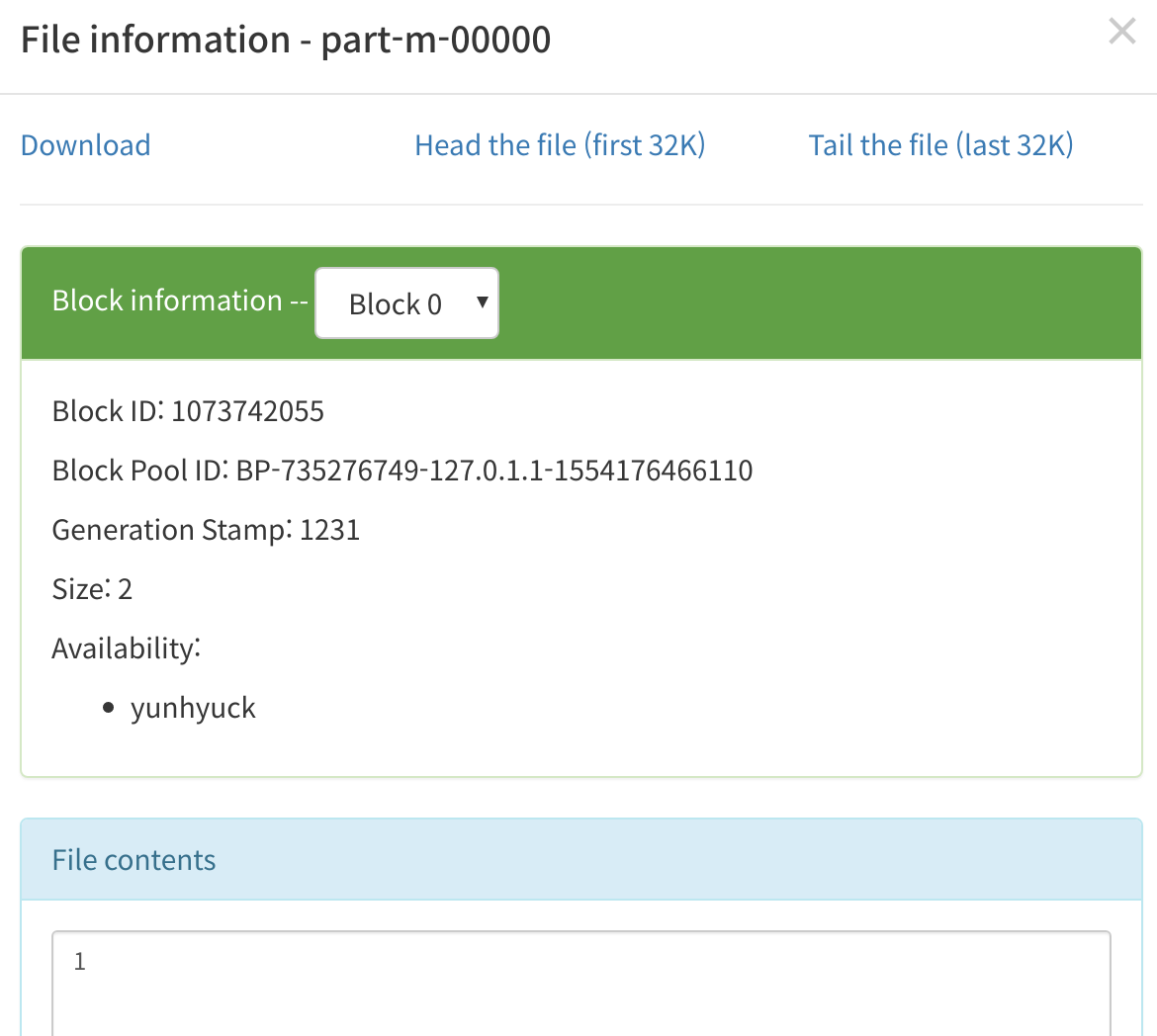
**--table test (DB 테이블 명)**

**--columns a(DB테이블에서 가져올 칼럼 명)**

**--target-dir hdfs://localhost:9000/user/yunhyuck/test4 -m 1 (HDFS에 저장하기 위해 HDFS의 주소 와 저장할 위치 지정. 단, 디렉토리는 존재하지 않는 폴더 여야 함.)**

****

**http://localhost:9870/**

****