# Hive、MySQL、HBase数据互导

**【实验名称】**

Hive、MySQL、HBase数据互导

**【实验目的】**

掌握数据仓库Hive概念与基本原理、关系数据库概念与基本原理、SQL语句、列族数据库HBase概念与基本原理

**【实验要求】**

 数据仓库Hive的基本操作、关系数据库MySQL的基本操作、Sqoop工具的使用方法

**【实验设备】**

基于综合实验2在hadoop1主节点上进行操作

## Hive预操作

请先查看hadoop集群、HBase、Zookeeper、MySQL是否已经启动，如果未启动请先启动

启动hadoop集群：

[root@hadoop1 ~]# cd /opt/hadoop-3.1.0/sbin/

[root@hadoop1 sbin]# ./start-all.sh

启动HBase:

[root@hadoop1 ~]# cd /opt/hbase/bin/

[root@hadoop1 bin]# ./start-hbase.sh

启动Zookeeper:(集群的每台服务器都需要执行)

[root@hadoop1 ~]# cd /opt/zookeeper-3.4.10/conf/

[root@hadoop1 conf]# zkServer.sh start

启动MySQL：

[root@hadoop1 ~]# service mysqld start

### 1.1启动Hive

[root@hadoop1 ~]#hive

### 1.2创建临时表user\_action

命令执行完以后，Hive会自动在HDFS文件系统中创建对应的数据文件/user/hive/warehouse/dblab.db/user\_action

hive> create table dblab.user\_action(id STRING,uid STRING, item\_id STRING, behavior\_type STRING, item\_category STRING, visit\_date DATE, province STRING) COMMENT 'Welcome to XMU dblab! ' ROW FORMAT DELIMITED FIELDS TERMINATED BY '\t' STORED AS TEXTFILE;



### 1.3将bigdata\_user表中的数据插入到user\_action

hive> INSERT OVERWRITE TABLE dblab.user\_action select \* from dblab.bigdata\_user;



查看上面命令是否执行成功：

hive> use dblab;

hive> select \* from user\_action limit 10;



## 使用Sqoop将数据从Hive导入MySQL

### 2.1登录MySQL，这里在第10课中已经设置了MySQL的密码为root

[root@hadoop1 ~]# mysql -u root -p



### 2.2创建数据库

#### 2.2.1显示所有数据库

mysql> show databases;



#### 2.2.2创建dblab数据库

mysql> create database dblab;



#### 2.3.3使用dblab数据库

mysql> use dblab;



### 2.3修改MySQL数据编码为utf8

#### 2.3.1退出MySQL数据库

mysql> quit;



#### 2.3.2编辑my.cnf文件，加入以下参数

[root@hadoop1 ~]# vi /etc/my.cnf



#### 2.3.3重启MySQL数据库

[root@hadoop1 ~]# service mysqld restart



#### 2.3.4查看是否修改成功

进入MySQL数据库，密码为root

[root@hadoop1 ~]# mysql -u root -p

mysql> show variables like "char%";



### 2.4创建表

#### 2.4.1在MySQL的数据库dblab中创建一个新表user\_action，并设置其编码为utf-8

mysql> CREATE TABLE `dblab`.`user\_action` (`id` varchar(50),`uid` varchar(50),`item\_id` varchar(50),`behavior\_type` varchar(10),`item\_category` varchar(50), `visit\_date` DATE,`province` varchar(20)) ENGINE=InnoDB DEFAULT CHARSET=utf8;



#### 2.4.2退出MySQL数据库

mysql> quit;



### 2.5导入数据

#### 2.5.1将hive中的数据导入导MySQL中

[root@hadoop1 ~]# sqoop export --connect jdbc:mysql://hadoop1:3306/dblab --username root --password root --table user\_action --export-dir /user/hive/warehouse/dblab.db/user\_action/ --fields-terminated-by '\t';



#### 2.5.1查看MySQL数据库中user\_action表的数据

进入mysql

[root@hadoop1 ~]# mysql -u root -p



查询user\_action表中的数据

mysql> use dblab;

mysql> select \* from user\_action limit 10;



## 将MySQL的数据导入导HBase中

请先查看hadoop集群、HBase、Zookeeper、MySQL是否已经启动，如果未启动请先启动

启动hadoop集群：

[root@hadoop1 ~]# cd /opt/hadoop-3.1.0/sbin/

[root@hadoop1 sbin]# ./start-all.sh

启动HBase:

[root@hadoop1 ~]# cd /opt/hbase/bin/

[root@hadoop1 bin]# ./start-hbase.sh

启动Zookeeper:(集群的每台服务器都需要执行)

[root@hadoop1 ~]# cd /opt/zookeeper-3.4.10/conf/

[root@hadoop1 conf]# zkServer.sh start

启动MySQL：

[root@hadoop1 ~]# service mysqld start

### 3.1启动HBase并创建user\_action表

[root@hadoop1 ~]# hbase shell



创建user\_action表

hbase> create 'user\_action', { NAME => 'f1', VERSIONS => 5}



### 3.2将MySQL的数据导入导HBase中

[root@hadoop1 ~]# sqoop import --connect jdbc:mysql://hadoop1:3306/dblab --username root --password root --table user\_action --hbase-table user\_action --column-family f1 --hbase-row-key id --hbase-create-table -m 1



### 3.3查看HBase中user\_action表的前10条数据

hbase > scan 'user\_action',{LIMIT=>10}

