

1.1 概述

Phoenix是HBase的一个开源SQL框架。可以使用标准的JDBC api而不是常规的HBase客户端api来创建表、插入数据和查询HBase数据。

Phoenix会将用户编写的sql查询编译为一系列的scan操作，最终产生通用的JDBC结果集返回给客户端

下载地址：<http://phoenix.apache.org/download.html>

或者：<http://archive.apache.org/dist/phoenix/>

1.2 phoenix安装

1 将jar拷贝到hbase的lib目录下

phoenix-4.12.0-HBase-1.2-client.jar

phoenix-core-4.12.0-HBase-1.2.jar

完成以上步骤之后，重启Hbase

2 将hbase/conf下的hbase-site.xml文件拷贝到Phoenix/bin目录下，要么拷贝，要么直接修改(如果HA则需要把core-site.xml hdfs-site.xml 一起拷贝过来)

```
<configuration>
  <property>
    <name>hbase.regionserver.wal.codec</name>
    <value>org.apache.hadoop.hbase.regionserver.wal.IndexedWALEditCodec</value>
  </property>
  <!-->以上为phoenix 自带属性<-->
  <property>
    <name>hbase.tmp.dir</name>
    <value>/opt/app/hbase-1.3.1/hbase</value>
  </property>

  <property>
    <name>hbase.cluster.distributed</name>
    <value>true</value>
  </property>

  <property>
    <name>hbase.zookeeper.quorum</name>
    <value>hadoop01,hadoop02,hadoop03</value>
  </property>

  <property>
    <name>hbase.rootdir</name>
    <value>hdfs://mycluster/hbase</value>
  </property>
</configuration>
```

3 启动

启动Phoenix：指定ZK的地址以及端口号，作为hbase的访问入口

bin/sqlline.py hadoop01:2181

1.3 操作语法

语句大多数类似mysql语句，但是没有show database table

测试：show databases、list 都不行 用help帮助信息

!tables 当前phoenix的系统表，hbase中也会有这些表，表示关联成功一半

语法学习网站

<http://forcedotcom.github.io/phoenix/语法>

<http://phoenix.apache.org/language/index.html>语法

1.3.1 基本操作

在完成了以上的步骤之后，接下来对Phoenix进行测试

1 首先在phoenix中创建一张表

```
CREATE TABLE user ( id varchar PRIMARY KEY, name varchar , passwd varchar );
```

向表中插入数据

```
upsert into user(id, name, passwd) values('001', 'admin', 'admin');
```

hbase区分大小写

Phoenix不区分大小写，默认情况下都是大写，变小写加上双引号就可以

在hbase的shell中：describe 'USER' 查看

得知默认列族为0，id主键字段对应到hbase的ROW字段

2 删除表

```
drop table user;
```

3 更新数据

插入数据: updata+insert结合 -> upsert 更新数据与这个语句是一样的

```
upsert into user(id, name, passwd) values('001', 'admin', 'admin');
```

3 查询语句

```
SELECT *  
FROM user;
```

```
SELECT name  
FROM user  
WHERE id ='001';
```

4 删除语句

```
delete  
FROM user  
WHERE id='001';
```

在phoenix的CLI界面中进行简单的CRUD操作，基本上与RDBMS的操作没有太大区别

5 hbase与phoenix表与表进行关联

在phoenix中：列的顺序可以不一致，表名称必须一致，以及列名称一致

```
create table "stu_import"( "row" varchar primary key, "f1"."location" varchar,
"f2"."age" varchar, "f1"."name" varchar, "f3"."addr" varchar );
```

在hbase中出现这个不用管他，是Phoenix的处于性能问题加上的列，没有多大影响，不会影响查询
column=f1:_0, timestamp=1531625109300, value=

1.4 phoenix导入csv文件数据(批量)

注意：要加载的所有CSV文件都必须具有".csv"文件扩展名，同样phoenix也支持mr做一个批量数据导入

主键就是hbase对应的rowkey，为了防止hbase主键的唯一性，可以设置多个字段为主键

```
create table if NOT EXISTS bulkdata ( state char(2) NOT null, city varchar NOT
null, population bigint constraint my_pk primary key (state, city) );
```

测试数据

```
NY,New York,8143197
CA,Los Angeles,3844829
IL,Chicago,2842518
TX,Houston,2016582
PA,Philadelphia,1463281
AZ,Phoenix,1461575
TX,San Antonio,1256509
CA,San Diego,1255540
TX,Dallas,1213825
CA,San Jose,912332
```

没有指定表的名称，他根据文件名称来找对应的表

```
bin/psql.py qianfent.class /opt/modules/bulkdata.csv
```