

Phoenix 安装与使用文档

作者：宋亚飞

2015 年 3 月 27 日

目录

Phoenix 安装与使用文档.....	1
目录.....	1
描述.....	2
Phoenix 安装.....	2
1、 下载 Phoenix.....	2
2、 上传压缩包.....	3
3、 解压缩文件.....	3
4、 配置 Phoenix.....	4
5、 修改权限.....	4
6、 验证是否成功.....	4
Phoenix 使用.....	5
1、 建表.....	5
2、 导入数据.....	5
3、 查询数据.....	5
查询 1、 查询全部记录.....	5
查询 2、 查询记录总条数.....	6
查询 3、 查询结果分组排序.....	6
查询 4、 求平均值.....	6
查询 5、 多字段分组，排序，别名。.....	7
查询 6、 查询日期类型字段.....	7
查询 7、 字符串，日期类型转换.....	7
4、 Phoenix 基本 shell 命令.....	8
5、 用 Phoenix Java api 操作 HBase.....	9
Phoenix 支持.....	12

描述

现有 hbase 的查询工具有很多如：Hive，Tez，Impala，Shark/Spark，Phoenix 等。今天主要记录 Phoenix。

phoenix，中文译为“凤凰”，很美的名字。Phoenix 是由 salesforce.com 开源的一个项目，后又捐给了 Apache 基金会。它相当于一个 Java 中间件，提供 jdbc 连接，操作 hbase 数据表。

但是在生产环境中，不可以用在 OLTP 中。在线事务处理的环境中，需要低延迟，而 Phoenix 在查询 HBase 时，虽然做了一些优化，但延迟还是不小。所以依然是用在 OLAT 中，再将结果返回存储下来。

[Phoenix](#) 官网上，对 Phoenix 讲解已经很详细了。如果英语好，可以看官网，更正式一些。

Phoenix 安装

1、下载 Phoenix

下载地址：<http://mirror.bit.edu.cn/apache/phoenix/>

phoenix 与 HBase 版本对应关系

Phoenix 2.x - HBase 0.94.x








Phoenix 3.x - HBase 0.94.x

Phoenix 4.x - HBase 0.98.1+

我目前测试使用版本概况：

Hadoop2.2.0--HBase0.98.2

所以我可以用 phoenix4.x。下载的压缩包为 phoenix-4.3.0-bin.tar.gz

Index of /apache/phoenix			
<u>Name</u>	<u>Last modified</u>	<u>Size</u>	<u>Description</u>
 Parent Directory		-	
 phoenix-3.1.0/	29-Aug-2014 08:51	-	
 phoenix-3.2.2/	11-Dec-2014 02:45	-	
 phoenix-3.3.0/	24-Feb-2015 13:56	-	
 phoenix-4.1.0/	29-Aug-2014 08:52	-	
 phoenix-4.2.2/	11-Dec-2014 02:45	-	
 phoenix-4.3.0/	24-Feb-2015 13:57	-	

2、上传压缩包

将 phoenix-4.3.0-bin.tar.gz 上传 hbase 集群的其中一个服务器的一个目录下
我上传的目录为/usr/lib

3、解压缩文件

```
tar -zxvf phoenix-4.3.0-bin.tar.gz
```

可看到有个 phoenix-4.3.0-bin/目录，里面包含了 Phoenix 的所有文件。

4、配置 Phoenix

- 4.1、将 phoenix-4.3.0-bin/目录下 phoenix-core-4.3.0.jar、phoenix-4.3.0-client.jar 拷贝到各个 hbase 的 lib 目录下。
- 4.2、重启 hbase 集群，使 Phoenix 的 jar 包生效。
- 4.3、将 hbase 的配置文件 hbase-site.xml 放到 phoenix-4.3.0-bin/bin/下，替换 Phoenix 原来的配置文件。

5、修改权限

切换到下，修改 psql.py 和 sqlline.py 的权限为 777
命令：chmod 777 文件名

6、验证是否成功

6.1、在 phoenix-4.3.0-bin/bin/下输入命令： `./sqlline.py localhost`

如果看到如下界面表示启动成功。

```
at SqlLine.SqlLineMain(SqlLine.java:252)
Connected to: Phoenix (version 4.3)
Driver: PhoenixEmbeddedDriver (version 4.3)
Autocommit status: true
Transaction isolation: TRANSACTION_READ_COMMITTED
Building list of tables and columns for tab-completion (set fastconnect to
79/79 (100%) Done
Done
sqlline version 1.1.8
0: jdbc:phoenix:localhost>
```

6.2、输入 `!tables`，查看都有哪些表。红框部分是用户建的表，其他为 Phoenix 系统表，系统表中维护了用户表的元数据信息。

```
0: jdbc:phoenix:localhost> !tables
```

TABLE_CAT	TABLE_SCHEM	TABLE_NAME
	SYSTEM	CATALOG
	SYSTEM	SEQUENCE
	SYSTEM	STATS
		STOCK_SYMBOL
		WEB_STAT

6.3、退出 Phoenix。输入 `!exit` 命令(PS: Phoenix 早期版本如(2.11 版本)需输入 `!quit` 才可退出，目前高版本已改为 `!exit` 命令)

```
0: jdbc:phoenix:localhost> !quit
Unknown command: quit
0: jdbc:phoenix:localhost> !exit
Closing: org.apache.phoenix.jdbc.PhoenixConnection
[root@centos041 bin]#
```

Phoenix 使用

1、建表

`./psql.py localhost:2181 ../examples/stock_symbol.sql`,其中 `../examples/stock_symbol.sql` 是建表的 sql 语句

```
CREATE TABLE IF NOT EXISTS WEB_STAT (
  HOST CHAR(2) NOT NULL,
  DOMAIN VARCHAR NOT NULL,
  FEATURE VARCHAR NOT NULL,
  DATE DATE NOT NULL,
  USAGE.CORE BIGINT,--usage 指定列族名
```

```
USAGE.DB BIGINT,--usage 指定列族名
STATS.ACTIVE_VISITOR INTEGER
CONSTRAINT PK PRIMARY KEY (HOST, DOMAIN, FEATURE, DATE)--指定主键
);
```

2、导入数据

命令: `./psql.py -t WEB_STAT localhost:2181 ../examples/web_stat.csv`

PS: 其中 `-t` 后面是表名, `../examples/web_stat.csv` 是 csv 数据 (注意数据的分隔符需要是逗号)。

3、查询数据

首先使用 `sqlline` 查看(截图为部分列的数据), 查询表名不区分大小写。

查询 1、查询全部记录

语句: `select * from web_stat;`

```
3 rows selected (0.112 seconds)
0: jdbc:phoenix:localhost> select * from WEB_STAT;
```

HOST	DOMAIN	FEATURE	DATE	
EU	Apple.com	Mac	2013-01-01 01:01:01.000	35
EU	Apple.com	Store	2013-01-03 01:01:01.000	345
EU	Google.com	Analytics	2013-01-13 08:06:01.000	25
EU	Google.com	Search	2013-01-09 01:01:01.000	395
EU	Salesforce.com	Dashboard	2013-01-06 05:04:05.000	12
EU	Salesforce.com	Login	2013-01-12 01:01:01.000	5
EU	Salesforce.com	Reports	2013-01-02 12:02:01.000	25
EU	Salesforce.com	Reports	2013-01-02 14:32:01.000	125
EU	Salesforce.com	Reports	2013-01-05 03:11:12.000	75
EU	Salesforce.com	Reports	2013-01-05 04:14:12.000	475
EU	Salesforce.com	Reports	2013-01-13 08:04:04.000	355
NA	Apple.com	Login	2013-01-01 01:01:01.000	35
NA	Apple.com	Login	2013-01-04 01:01:01.000	135
NA	Apple.com	Mac	2013-01-02 04:01:01.000	345
NA	Apple.com	Mac	2013-01-08 01:01:01.000	3
NA	Apple.com	iPad	2013-01-05 01:01:01.000	85
NA	Apple.com	iPad	2013-01-06 01:01:01.000	35
NA	Apple.com	iPad	2013-01-07 01:01:01.000	9
NA	Google.com	Analytics	2013-01-07 06:01:01.000	23
NA	Google.com	Analytics	2013-01-11 01:02:01.000	7
NA	Google.com	Analytics	2013-01-14 01:01:01.000	65
NA	Google.com	Search	2013-01-08 08:01:01.000	345

查询 2、查询记录总条数

语句: `select count(1) from web_stat;`

```
0: jdbc:phoenix:localhost> select count(1) from web_stat;
+-----+
|          COUNT(1)          |
+-----+
|          39                 |
+-----+
1 row selected (0.054 seconds)
0: jdbc:phoenix:localhost>
```

查询 3、查询结果分组排序

语句: select domain,count(1) as num from web_stat group by domain order by num desc;

```
0: jdbc:phoenix:localhost> select domain,count(1) as num from web_stat group by domain order by num desc;
+-----+-----+
|          DOMAIN          |          NUM          |
+-----+-----+
| Salesforce.com           | 22                    |
| Apple.com                | 9                     |
| Google.com               | 8                     |
+-----+-----+
3 rows selected (0.081 seconds)
0: jdbc:phoenix:localhost>
```

查询 4、求平均值

语句: select avg(core) from web_stat;

```
0: jdbc:phoenix:localhost> select avg(CORE) from web_stat;
+-----+
|          AVG(USAGE.CORE)          |
+-----+
|          217.0769                  |
+-----+
1 row selected (0.056 seconds)
0: jdbc:phoenix:localhost>
```

查询 5、多字段分组，排序，别名。

语句: select domain,count(1) as num,avg(core) as core,avg(db) as db from web_stat group by domain order by num desc;

```
0: jdbc:phoenix:localhost> select domain,count(1) as num,avg(core) as core,avg(db) as db from web_stat group by domain order by num desc;
+-----+-----+-----+-----+
|          DOMAIN          |          NUM          |          CORE          |          DB          |
+-----+-----+-----+-----+
| Salesforce.com           | 22                    | 260.7272               | 257.6363             |
| Apple.com                | 9                     | 114.1111               | 119.5555             |
| Google.com               | 8                     | 212.875                | 213.75               |
+-----+-----+-----+-----+
3 rows selected (0.133 seconds)
```

查询 6、查询日期类型字段

语句: select host,domain,date from web_stat where TO_CHAR(date)='2013-01-15 07:09:01.000';

```
0: jdbc:phoenix:localhost> select host,domain,date from web_stat where TO_CHAR(date)='2013-01-15 07:09:01.000';
```

HOST	DOMAIN	DATE
NA	Salesforce.com	2013-01-15 07:09:01.000

```
1 row selected (0.111 seconds)
```

查询 7、字符串，日期类型转换

语句: `select TO_DATE('20131125','yyyyMMdd') from web_stat;`

Ps: 输入的日期字符串会被转换为 hbase 表 date 的日期类型。

```
0: jdbc:phoenix:localhost> select TO_DATE('20131125','yyyyMMdd') from web_stat;
```

DATE '2013-11-25 00:00:00.000'
2013-11-25 00:00:00.000
2013-11-25 00:00:00.000
2013-11-25 00:00:00.000
2013-11-25 00:00:00.000
2013-11-25 00:00:00.000
2013-11-25 00:00:00.000
2013-11-25 00:00:00.000
2013-11-25 00:00:00.000
2013-11-25 00:00:00.000

总结: Phoenix 还支持了很多函数和 sql 语法, 在这里不再一一列举。

更多请参考 [Phoenix 支持](#) 部分

4、Phoenix 基本 shell 命令

PS: 以下, 可能有部分命令在 Phoenix 更高版本中已失效, 改为其他命令代替, 请注意。

```
0: jdbc:phoenix:localhost> help
```

!all	Execute the specified SQL against all the current connections
!autocommit	Set autocommit mode on or off
!batch	Start or execute a batch of statements
!brief	Set verbose mode off
!call	Execute a callable statement
!close	Close the current connection to the database
!closeall	Close all current open connections
!columns	List all the columns for the specified table
!commit	Commit the current transaction (if autocommit is off)
!connect	Open a new connection to the database.
!dbinfo	Give metadata information about the database
!describe	Describe a table
!dropall	Drop all tables in the current database
!exportedkeys	List all the exported keys for the specified table
!go	Select the current connection

!help	Print a summary of command usage
!history	Display the command history
!importedkeys	List all the imported keys for the specified table
!indexes	List all the indexes for the specified table
!isolation	Set the transaction isolation for this connection
!list	List the current connections
!manual	Display the SQLLine manual
!metadata	Obtain metadata information
!nativesql	Show the native SQL for the specified statement
!outputformat	Set the output format for displaying results (table,vertical,csv,tsv,xmlattrs,xmlelements)
!primarykeys	List all the primary keys for the specified table
!procedures	List all the procedures
!properties	Connect to the database specified in the properties file(s)
!quit	Exits the program 此命令在 Phoenix4.3 版本已改为!exit
!reconnect	Reconnect to the database
!record	Record all output to the specified file
!rehash	Fetch table and column names for command completion
!rollback	Roll back the current transaction (if autocommit is off)
!run	Run a script from the specified file
!save	Save the current variables and aliases
!scan	Scan for installed JDBC drivers
!script	Start saving a script to a file
!set	Set a sqlline variable

5、用 Phoenix Java api 操作 HBase

开发环境准备：eclipse3.5、jdk1.7、window8、hadoop2.2.0、hbase0.98.0.2、phoenix4.3.0

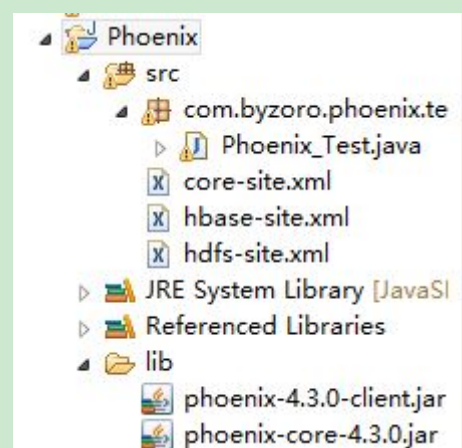
1、从集群拷贝以下文件：core-site.xml、hbase-site.xml、hdfs-site.xml 文件放到工程 src 下

2、把 phoenix 的 phoenix-4.3.0-client.jar 和 phoenix-core-4.3.0.jar 添加到工程 classpath

3、配置集群中各节点的 hosts 文件，把客户端的 hostname: IP 添加进去

4、在客户端 host 文件中加入集群的 hostname 和 IP

5、工程截图



6、工程代码

```
package com.byzoro.phoenix.test;

import java.sql.Connection;
import java.sql.DriverManager;
import java.sql.ResultSet;
import java.sql.Statement;

public class Phoenix_Test {
    /**
     * 使用 phoenix 提供的 api 操作 hbase 读取数据
     * @author songyafei
     * @param args
     * @throws Throwable
     * @date 2015-03-25
     */
    public static void main(String[] args) throws Throwable {
        try {
            //下面的驱动为 Phoenix 老版本使用 2.11 使用，对应 hbase0.94+
            //Class.forName("com.salesforce.phoenix.jdbc.PhoenixDriver");

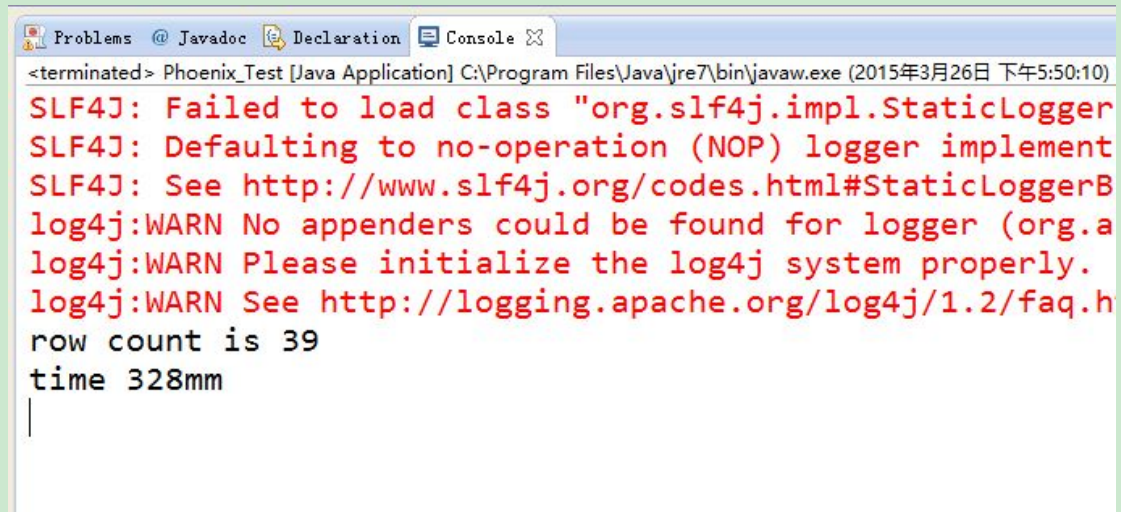
            //phoenix4.3 用下面的驱动对应 hbase0.98+
            Class.forName("org.apache.phoenix.jdbc.PhoenixDriver");
        } catch (Exception e) {
            e.printStackTrace();
        }
        //这里配置 zookeeper 的地址，可单个，也可多个。可以是域名或者 ip
        String url = "jdbc:phoenix:10.14.9.41:2181";
        //String url = "jdbc:phoenix:41.byzoro.com,42.byzoro.com,43.byzoro.com:2181";
        Connection conn = DriverManager.getConnection(url);
        Statement statement = conn.createStatement();
        String sql = "select count(1) as num from web_stat";
        long time = System.currentTimeMillis();
        ResultSet rs = statement.executeQuery(sql);
        while (rs.next()) {
            int count = rs.getInt("num");
            System.out.println("row count is " + count);
        }
        long timeUsed = System.currentTimeMillis() - time;
        System.out.println("time " + timeUsed + "mm");
        //关闭连接
        rs.close();
        statement.close();
    }
}
```

```

        conn.close();
    }
}

```

7、运行截图：



实例二：

```
package com.byzoro.phoenix.test;
```

```

import java.sql.Connection;
import java.sql.DriverManager;
import java.sql.ResultSet;
import java.sql.Statement;

```

```

public class Phoenix_Test {
    /**
     * 使用 phoenix 提供的 api 操作 hbase 中读取数据
     * @author songyafei
     * @param args
     * @throws Throwable
     * @date 2015-03-25
     */
    public static void main(String[] args) throws Throwable {
        try {
            //下面的驱动为 Phoenix 老版本使用 2.11 使用，对应 hbase0.94+
            //Class.forName("com.salesforce.phoenix.jdbc.PhoenixDriver");

            //phoenix4.3 用下面的驱动对应 hbase0.98+
            Class.forName("org.apache.phoenix.jdbc.PhoenixDriver");
        } catch (Exception e) {
            e.printStackTrace();
        }
        //这里配置 zk 的地址，可单个，也可多个。可以是域名或者 ip
    }
}

```

```

String url = "jdbc:phoenix:10.14.9.41:2181";
//String url = "jdbc:phoenix:41.byzoro.com,42.byzoro.com,43.byzoro.com:2181";
Connection conn = DriverManager.getConnection(url);
Statement statement = conn.createStatement();
//String sql = "select count(1) as num from web_stat";
String sql = "select * from web_stat where core = 35";
long time = System.currentTimeMillis();
ResultSet rs = statement.executeQuery(sql);
while (rs.next()) {
    //获取 core 字段值
    int core = rs.getInt("core");
    //获取 core 字段值
    String host = rs.getString("host");
    //获取 domain 字段值
    String domain = rs.getString("domain");
    //获取 feature 字段值
    String feature = rs.getString("feature");
    //获取 date 字段值,数据库中字段为 Date 类型, 这里代码会自动转化为 string
    类型
    String date = rs.getString("date");
    //获取 db 字段值
    String db = rs.getString("db");

    System.out.println("host:"+host+"\tdomain:"+domain+"\tfeature:"+feature+"\tdate:"+date+"\t"
    core:" + core+"\tdb:"+db);
}
long timeUsed = System.currentTimeMillis() - time;
System.out.println("time " + timeUsed + "mm");
//关闭连接
rs.close();
statement.close();
conn.close();
}
}

```

Phoenix 支持

Phoenix 中的[语法](#)

Phoenix 中的[数据类型](#)

Phoenix 中的[方法](#)

本文档为原创，转载请注明出处，尊重作者劳动成果。

更多大数据技术知识请访问：[宋亚飞.中国](#) www.songyafei.cn