# 1.2搭建hbase

## 一、大数据平台的技术选择

### 1. HBase技术

HBase是一个分布式，版本化，面向列的开源数据库，构建在 Apache Hadoop和 Apache ZooKeeper之上。HBase是一个高可靠性、高性能、面向列、可伸缩的分布式存储系统，利用HBase技术可在廉价PC Server上搭建起大规模结构化存储集群。HBase不同于一般的关系数据库，它是一个适合于非结构化数据存储的数据库，HBase基于列的而不是基于行的模式。

### 2. Phoenix技术

Phoenix是给HBase添加了一个语法表示层，允许我们用SQL语句读写HBase中的数据，可以做联机事务处理，拥有低延迟的特性，这就让我方便多了。Phoenix会把SQL编译成一系列的Hbase的scan操作，然后把scan结果生成标准的JDBC结果集，处理千万级行的数据也只用毫秒或秒级就搞定。而且Phoenix还支持MyBatis框架，正好可以和华夏代驾项目整合到一起。

## 二、部署HBase与Phoenix

### 1. 导入镜像文件

docker load < phoenix.tar.gz

### 2. 创建容器

因为镜像中已经包含和HBase和Phoenix，所以我们只需要创建出容器即可。由于HBase需要使用的内存较大，这里我没有规定具体的内存大小，容器会自动使用空闲的内存。容器中数据目录是/tmp/hbase-root/hbase/data，我把这个目录映射到宿主机的/root/hbase/data目录。

docker run -it -d -p 2181:2181 -p 8765:8765 -p 15165:15165 \  
-p 16000:16000 -p 16010:16010 -p 16020:16020 \  
-v /root/hbase/data:/tmp/hbase-root/hbase/data \  
--name phoenix \  
boostport/hbase-phoenix-all-in-one:2.0-5.0

### 3. 开放端口

我们要把Linux的2181、8765、15165、16000、16010、16020端口，映射到Windows的相应端口上面。

### 4. 初始化Phoenix

运行命令，进入到Phoenix容器中，然后执行命令设置HBASE\_CONF\_DIR环境变量。

代码块：

docker exec -it phoenix bash  
  
  
export HBASE\_CONF\_DIR=/opt/hbase/conf/

接下来我们要连接Phoenix的命令行客户端。虽然IDEA也内置了Phoenix客户端，但是Bug挺多的，用着并不方便，所以我建议大家使用Phoenix自带的命令行客户端更好一些。而且我们要执行的SQL语句也并不多，命令行客户端已经足够用了。

代码块：

/opt/phoenix-server/bin/sqlline.py localhost

## 三、创建逻辑库和数据表

### 1. 创建逻辑库

为了存储数据，我们需要像操作MySQL一样，先创建逻辑库，然后定义数据表。在Phoenix的命令行客户端我们先来执行创建逻辑库的命令。

CREATE SCHEMA hospital;  
USE hospital;

### 2. 创建数据表与导入数据

在本课程的git上面有数据库脚本.SQL文件，你把这个SQL文件中的语句复制粘贴到Phoenix的命令行中执行即可。

## 四、配置JDBC连接信息

把hospital-api和patient-wx-api两个项目的application.yml文件打开，确认其中JDBC数据库连接信息。因为连接的是本地Phoenix，所以写localhost是可以的。

重启  
docker ps  
docker restart xxxx

