1:设计索引库的settings信息的mappings信息，并把这些配置信息保存到一个配置文件中。

vi bjsxt.json

{

"settings":{

"number\_of\_shards":5,

"number\_of\_replicas":0

},

"mappings":{

"doc":{

"dynamic":"strict",

"properties":{

"id":{"type":"integer","store":"yes"},

"title":{"type":"string","store":"yes","index":"analyzed","analyzer": "ik\_max\_word","search\_analyzer": "ik\_max\_word"},

"describe":{"type":"string","store":"yes","index":"analyzed","analyzer": "ik\_max\_word","search\_analyzer": "ik\_max\_word"},

"author":{"type":"string","store":"yes","index":"no"}

}

}

}

}

2：安装配置es服务。

1)：整合IK中文分词工具

3：启动es。

4：创建索引库

使用开始定义的文件进行创建，使用下面命令进行创建

curl -XPOST 'master:9200/bjsxt' -d @bjsxt.json

5：初始化数据

把原始数据在es中建立索引，并且还要在hbase中保存一份。

需要使用提供的代码里面的一个类(DataImportAndIndex.java)

1：使用hbaseutils工具类的时候，要注意修改zk的IP地址和hbase.rootdir的地址

conf.set("hbase.zookeeper.quorum", "192.168.57.4:2181");

/\*\*

\* 注意：如果使用这个地址无法访问的话请尝试使用这个

\* hdfs://chenkl/hbase

\* 因为hadoop这边做了HA，现在主为197，备为196

\*/

conf.set("hbase.rootdir", "hdfs://chenkl/hbase");

还需要修改本机的hosts文件,不然hbase工具类无法正常调用

还需要在集群中创建一个表：create 'doc','cf1'

2：使用es工具类的时候要注意：

修改es的ip地址

addIndex方法需要实现

6：确认Es工具类中的search方法是否已经完善。

7：启动tomcat

8：访问：http://localhost:8080/article

9：如果使用tomcat启动这个项目

注意：jdk使用1.7

tomcat使用7.0