

[今日课程大纲]

Solr 简介

Solr 单机版搭建

Analyzer 配置业务字段

SolrJ 操作 Solr

SolrCloud 搭建

[知识点详解]

一.Solr 简介

1. Solr 是什么？

1.1 就是一个 war 项目

2. 自己的项目如何和 Solr 进行交互？

2.1 特定的 API 叫做 SolrJ

3. 具备数据持久化功能.

3.1 Solr 中会存储需要进行搜索的数据

3.2 把所有数据都初始化到 Solr 中.

4. Solr 作用(什么时候使用 Solr)

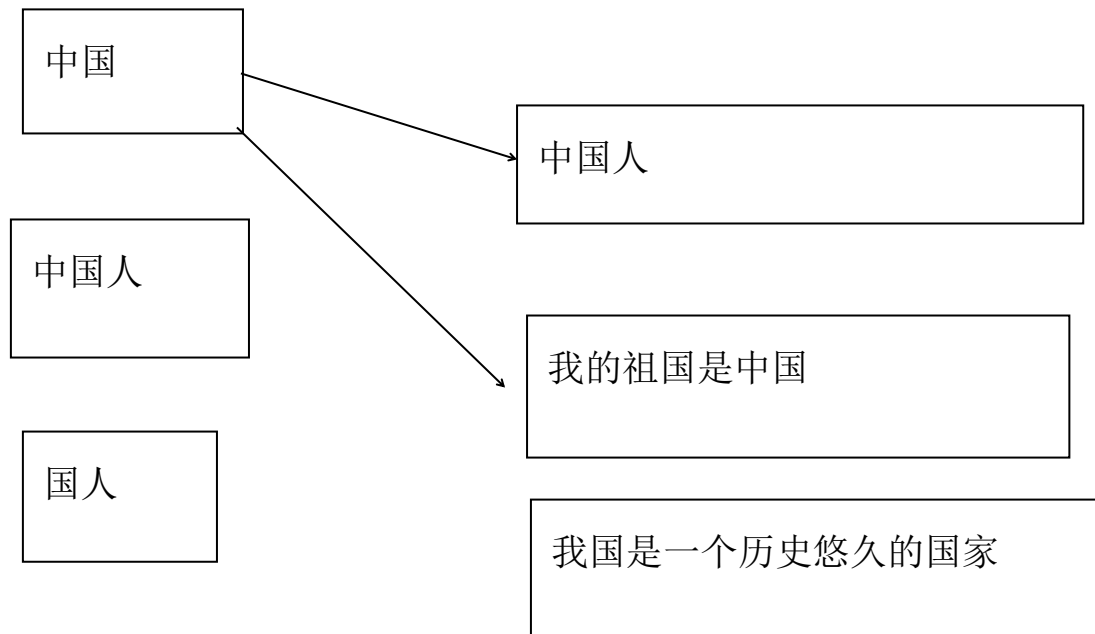
4.1 大量数据检索时使用 Solr,能提升检索效率.

5. Solr 是基于索引进行查询的.

5.1 顺序查询:从内容的最开始找到内容为止

5.2 反向键索引:

5.2.1 把内容中进行拆分.



6. 国内实现检索的常用方案.

6.1 apache lucene : 实现检索的解决方案(Solr 基于 lucene)

6.2 Baidu API:

6.3 Google API

二. IK Analyzer 中文拆词器

1. Solr 默认对中文拆词功能支持不好.

1.1 解决方案:使用 IK Analyzer 拆词器

2. 配置 IK Analyzer 时本质实际上是给 solr 新建了一个 fieldType,只要某个属性(field)是这个类型,这个属性就会使用 IK Analyzer 进行拆词

三. Solr 管理界面的 documents 菜单功能

1. 包含了 solr 数据的新增,删除,修改三个功能.
2. 数据支持很多种格式:json 或 xml 等
3. 新增时,必须包含对 id 的新增.
4. 每次新增,Solr 会新建一个 SolrDocument 对象.这个对象存储新增内容.

4.1 把 SolrDocument 理解成实体类.

4.2 包含了 Solr 的 schema.xml 中配置的所有 field

4.3 新增时有什么属性,最终显示时就只有什么属性.

5. 把 DocumentType 以 xml 方式举例

5.1 新增

```
<doc>

<field name="id">change.me</field>

<field name="title">change.me</field>

</doc>
```

5.2 修改: 只要 id 已经存在执行修改

5.3 删除:

5.3.1 根据 id 删除

```
<delete>

<id>id 内容</id>

</delete>
```

```
<commit/>
```

5.3.2 删除全部

```
<delete>  
<query>*:*</query>  
</delete>  
<commit/>
```

四. 管理界面 quey 菜单功能

1. q 表示查询条件

1.1 *:* 查询全部

1.2 bjsxt:java 培训

1.2.1 把 java 培训拆分成 java 和培训,按照 java 和培训在 bjsxt 属性中进行搜索.

1.3 bjsxt:"java 培训": java 培训不回被拆词

1.4 在 schema.xml 中配置复合属性,这个复合属性可以包含多个其他属性的值.

1.4.1 在 schema.xml 添加下面内容

```
<field name="bjsxtall" type="text_ik" indexed="true" stored="true"  
multiValued="true"/>
```

```
<copyField source="bjsxt" dest="bjsxtall"/>
```

```
<copyField source="bjsxt1" dest="bjsxtall"/>
```

五. SolrCloud

1. solr 是 web 项目,需要放入到 tomcat 中.依赖多个 tomcat,让多个 tomcat 之间能够通信.
2. 需要借助 zookeeper 实现 tomcat 直接通信
3. 结构图

