目录

[2018.7.12 1](#_Toc28971)

[lucene 1](#_Toc20270)

[2018.8.25 1](#_Toc10991)

[2018.8.26 2](#_Toc6779)

[2018.8.27 2](#_Toc22237)

[2018.8.30 3](#_Toc1621)

[2018.8.31 3](#_Toc30976)

# 2018.7.12

现在手上确实没什么任务，研究下全文搜索。

先说好，是全文搜索，不是百度谷歌那些互联网搜索。我对全文搜索的理解就是通过某个关键词关联到数据库内的某条记录。不考虑那些框架，唯一的实现方式是用SQL去like。like关键字是没法走索引的，因此要扫全表，那么，一旦表的数据量变大，一次查询消耗的时间估计能有几十秒，这个是等不起的。

所以要用框架。

是什么，为什么，怎么办。

可以啊，三大问题已经解决两个了，先看看阮一峰的入门教程吧。

“Elastic 本质上是一个分布式数据库”

觉得这句话说得特别好，它本质上是个数据库，查询说到底，利用的就是索引。

# lucene

## 2018.8.25

很烦，看视频看的难受。

我不知道我为什要学这个，我也不知道我做这个有什么用。

行了，不去管它，我自己试试看能不能写demo，不过在这之前要先写周报。

// 周报一写完，劲头又没了、~~再休息一下~~至少看5分钟吧。

写一个什么demo呢，这个demo就用来查找我的blogbackup

## 2018.8.26

唉，写个demo写老半天写不出来。原本的计划是搜索blogbackup，但里面全中文，而且还是word。所以只能改成搜索棚车少年。

可是这个demo写不出来啊。

用lucene的第一步是根据文本构建索引，创建索引的类是IndexWriter。好了，这里我暂停一下，我写代码都是从网上找例子。但是他妈的，lucene版本迭代的太快了，而我又一根筋，非得下最新版，所以我找对的上版本的例子就很困难。后来索性自己看源码，这一看就30min过去了。

其实也挺好，熟悉熟悉怎么搞新框架，顺带着打发时间。

IndexWriter indexWriter = new IndexWriter(directory, indexWriterConfig);

创建writer需要设置config，这个好说，虽然背后的意义未必懂，但是会调方法，不报错。Directory就是索引存储目录，这是个接口，因为索引可以存储在硬盘上，也可以存储在内存里。我肯定是存在硬盘上的…… ……（省略一堆心路历程）

java.nio.file.Path

这个吊东西呢，也是接口，但我new不出实例来。

好吧，jdk自带静态方法。

## 2018.8.27

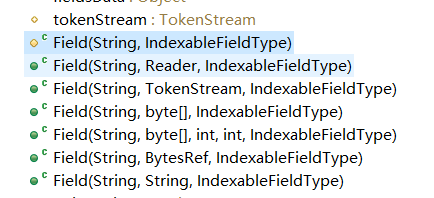
IndexWriter 已经new出来了，下一步，开始创建索引。

indexWriter.addDocument(document);

只要能调用这个方法就行。

Document这个类的结构很简单，就是用来保存Field类的一个容器，所以Document可以直接new，不需要配置什么东西。想来，一个文件一个Document对象。

Field类我理解是用来描述一个文本，名字是什么，路径是什么，内容是什么。



不过new Field类必须提供FieldType，那么这个对象又要怎么来呢？

## 2018.8.30

讲道理，我研究老半天的东西，看视频真的自然而然就过去了。

所以，还是老老实实先看视频好了。

添加文档，删除文档，更新文档。

## 2018.8.31

时间有，心态也还可以，但我不知道我学这个东西能干嘛，犹犹豫豫的，耗着耗着就一下午过去了。

lucene暂停一下，直接看elasticSearch的视频，因为我手里面有公司的项目，多少能有点实践空间，而不是对着demo琢磨。