https://www.liangzl.com/get-article-detail-134667.html

Solr是一个用java开发的基于lucene引擎的 全文搜索 平台，原理是倒排索引，从关键字到文档的过程，Solr的资源都是以Doc为对象进行存储的，每个文档都由一系列的field构成，每个field表示资源的一个属性，文档的fied可以被索引，用于提高性能的搜索效率，一般情况下文档都包含一个能唯一表示该文档的id字段。

**Solr原理？不太明白**

倒排索引

通过value查找key，假设文章1 内容为123，文章2 内容为 456，文章3 内容为 789，

Solr有个功能是分词，先把文章123导入到服务器后，文章1切成1,2,3 ,原本123是文章内容，分词后，123变为三个 单独的索引,1,2,3相当于key，文章为value,文章2 和3与1类似，切成456789六个索引，他们的value就是文章序号2和3.

**索引为什么快？**  
索引以文件的形式存储在磁盘上，所以只使用更少的磁盘io 次数的数据结构更适合做索引。b 树和b+树是是多叉树，树的度大，所以高度低。内存和磁盘交互的单位是页，将b 树和b+树的一个节点的大小设置为一个页，能保证一次io 就能读到一个页，同时磁盘采用预读策略，一次性读取相邻的几个页，读入内存后在进行二分查找。

简单说就是因为索引使用B+树，空间换时间 。

**为什么要用Solr？**

随意的数据不可能是根据数据库的字段查询的，那是怎么查询出来的呢，为什么千奇百怪的关键字都可以查询出来呢？  
答案就是全文检索工具的实现，luncence采用了词元匹配和切分词。举个例子：北京天安门----luncence切分词：北京  京天  天安  安门  等等这些分词。所以我们搜索的时候都可以检索到。

有一种分词器就是IKanalyzer中文分词器，它有细粒度切分和智能切分，即根据某种智能算法。

这就使用solr的最大的好处：检索功能的实现。

**Solr为什么比较快？**

**倒排索引，使用solr的分词过后吗，通过二分法与排序好的索引进行匹配，进而快速查到定位文档**

**空间换时间**

**操作：**

主要是SolrTemplate完成文档操作和搜索

SolrTemplate、Query、Criteria

solr会考虑集群  
solr索引库在商品状态发生改变的时候（审核通过、删除），会通过activeMQ通知更新。  
solr搜索结果，可以通过redis缓存