1. ElasticSearch
2. 概述

ElasticSearch，简称ES，是一个开源的分布式全文检索引擎，可以近乎实时地存储，检索数据。同时它可以扩展到上百台服务器之中，处理PB级别的数据。ES使用Java开发并且使用Lucene作为其核心来实现所有的索引和搜索功能，但是它通过简单的RESTful API来让搜索的调用更加简单。

1. ElasticSearch与Solr的对比
2. Solr利用Zookeeper进行分布式管理，而ElasticSerch自身带有分布式协调管理功能。
3. Solr支持更多格式的数据，而ElasticSearch仅支持json文件。
4. Solr在传统搜索中好于ElasticSearch，但是在实时搜索效率较低。
5. ElasticSearch的安装与启动

Windows中，直接解压下载的ElasticSearch压缩包即可，在bin目录下直接点击elasticseach.bat批处理文件就可以启动。9200是http协议的RESTful接口，9300是tcp通讯端口，集群间和TCPClient都执行该端口。通过访问127.0.0.1:9300就能访问到ElasticSearch服务器。

1. 安装ES的图形化界面插件

通过安装插件来使用图形化界面，查看索引的数据。插件的名称叫做head，它是基于node.js写的，因此还需要安装node.js。安装好node.js后，将head插件的压缩包解压后，在head插件的解压目录下打开cmd窗口，安装grunt，这是基于Node.js的项目构建工具，只有安装了grunt之后才能使用head插件。

还是在head插件的解压目录下的cmd窗口中，通过npm命令来安装grunt：

npm install -g grunt-cli

执行完毕后，再执行以下的命令就可以启用head插件了：

npm install

grunt server

访问127.0.0.1:9100地址，就能看到head插件的图形化界面了。之前我们已经启动了一个ElasticSearch服务在9200端口，因此在插件中如果我们想要连接到该服务中，就需要修改ElasticSeach的config目录下的配置文件：config/elasticsearch.yml，增加以下两行命令，用于跨域访问：

http.cors.enabled: true

http.cors.allow-origin: “\*”

重新启动ElasticSearch服务，通过head插件连接启动的ElasticSeach服务，可以看到显示出了一个节点，这就是之前启动的elasticsearch节点的信息。这样，你的ES图形化界面就可以使用了。

1. ElasticSearch核心概念
2. 索引index

一个索引是一个包含多份相似的文档的集合，类似于一个数据库。即每一个索引中包含的都是同一类型的文档，如一个客户数据的索引，一个订单数据的索引。索引由名字来标识，名字都是小写字母的，这个名字在对索引中的文档进行索引，搜索，更新和删除时使用。一个集群可以定义多个索引。

1. 类型type

类型是一个索引的逻辑分区，即将一个索引再分为多个类型，类似于一个数据表，每个类型的语义是自定义的。通常，会为具有一组共同字段的文档定义一个类型、

1. 字段field

对文档的数据根据不同属性进行的分类标识，每个文档为一条记录，每个文档中有多个字段，即每个文档有不同的属性分类。

1. 映射mapping

映射是对处理数据的方式和规则方面做出的限制，如某个字段的数据类型，默认值等等都是用映射设置的。

1. 文档document

一个文档是一个可被索引的基础信息单元，一般是JSON格式的。

1. 集群cluster

一个集群有多个节点组织在一起，共同持有整个的数据，并一起提供索引和搜索功能。每个集群由名字来标识，默认是“elasticsearch”。

1. 节点node

一个节点是集群中的一个服务器，每个节点由名字来标识。节点通过配置集群名称的方式来加入集群，默认情况下，所有节点都会加入到默认的“elasticsearch”集群中。

1. 分片和复制shards&replicas

当索引存储的数据量过大时，可以将索引划分为多个分片，在创建一个索引时，可以指定想要的分片的数量，每个分片可以放到集群中的任意节点中。分片非常重要，因为分片能够允许扩展内容容量，允许在分片之上进行分布式，并行的操作，提高性能。

同时，每个分片都可以创建拷贝，这些拷贝称为复制。在分片失败的情况下，复制能够及时地填补空缺，复制不能与原分片处于同一个节点上。