### Elasticsearch

**注意存储肯定不能选es，es快是因为启动把索引加到了内存，es只能做搜。**

Elasticsearch 是一个实时的分布式搜索分析引擎，它能让你以前所未有的速度和规模，去探索你的数据。 它被用作全文检索、结构化搜索、分析以及这三个功能的组合：

* Wikipedia 使用 Elasticsearch 提供带有高亮片段的全文搜索，还有 search-as-you-type 和 did-you-mean 的建议。
* blog 使用 Elasticsearch 将网络社交数据结合到访客日志中，为它的编辑们提供公众对于新文章的实时反馈。
* Stack Overflow 将地理位置查询融入全文检索中去，并且使用 more-like-this 接口去查找相关的问题和回答。
* GitHub 使用 Elasticsearch 对1300亿行代码进行查询。

### 核心概念

有几个概念是ElasticSearch的核心。从一开始就理解这些概念将极大地帮助简化学习过程。

#### 2.1 Cluster

集群是一个或多个节点（服务器）的集合，它们共同保存整个数据，并在所有节点上提供联合索引和搜索功能。集群由一个唯一的名称标识，该名称默认为“elasticsearch”。该集群名非常重要，因为节点可以通过这个集群名加入群集，一个节点只能是群集的一部分。

#### 2.2 Node

节点是单个服务器实例，它是群集的一部分，可以存储数据，并参与群集的索引和搜索功能。

#### 2.3 Index

索引是具有某种相似特征的文档的集合，在索引文档、搜索、更新、删除时都需要指定索引的名称，并且index的名称必须为小写字母。为方便理解可以将其看作mysql中的一个库。

#### 2.4 type

类型是索引的逻辑类别/分区，允许您在同一索引中存储不同类型的文档，例如一种用于用户，另一种用于博客文章。为方便理解可以将其看作mysql中的一张表。

#### 2.5 document

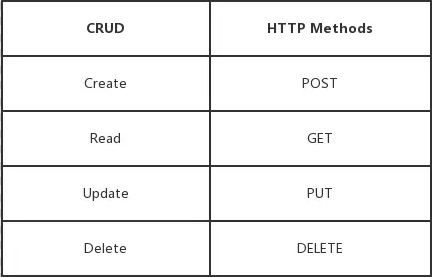
document是可以被索引的最小单元，可以将其看作为mysql中某张表的一条数据；其包含多个字段，比如用户表中可以包含用户名、密码、邮箱等；

在elasticsearch中，存储document数据需使用json格式数据与服务端交互；

#### 2.6 **RESTful API**

要弄清楚什么是RESTful API，首先要知道REST是什么。

REST，表示性状态转移（representation state transfer）。URL定位资源，HTTP请求方式(GET, POST, PUT, DELETE)描述对资源的操作。



### 使用

#### 启动Elasticsearch

Win:进入elasticsearch-xxx\bin,点击elasticsearch.bat启动

访问：<http://localhost:9200/>，出现下面信息则启动成功

|  |
| --- |
| {  "name" : "DESKTOP-EUOK1S3",  "cluster\_name" : "elasticsearch",  "cluster\_uuid" : "EDfybmb1QzCReF-H0dTxdQ",  "version" : {  "number" : "7.3.2",  "build\_flavor" : "default",  "build\_type" : "zip",  "build\_hash" : "1c1faf1",  "build\_date" : "2019-09-06T14:40:30.409026Z",  "build\_snapshot" : false,  "lucene\_version" : "8.1.0",  "minimum\_wire\_compatibility\_version" : "6.8.0",  "minimum\_index\_compatibility\_version" : "6.0.0-beta1"  },  "tagline" : "You Know, for Search"  } |

#### 与ES交互的URL要求

curl -X<VERB> '<PROTOCOL>://<HOST>:<PORT>/<PATH>?<QUERY\_STRING>' -d '<BODY>'

|  |  |
| --- | --- |
| VERB | 适当的 HTTP 方法 或 谓词 : GET、 POST、 PUT、 HEAD 或者 DELETE。 |
| PROTOCOL | http 或者 https（如果你在 Elasticsearch 前面有一个 https 代理） |
| HOST | Elasticsearch 集群中任意节点的主机名，或者用 localhost 代表本地机器上的节点 |
| PORT | 运行 Elasticsearch HTTP 服务的端口号，默认是 9200 |
| PATH | API 的终端路径（例如 \_count 将返回集群中文档数量）。Path 可能包含多个组件，例如：\_cluster/stats 和 \_nodes/stats/jvm 。 |
| QUERY\_STRING | 任意可选的查询字符串参数 (例如 ?pretty 将格式化地输出 JSON 返回值，使其更容易阅读) |
| BODY | 一个 JSON 格式的请求体 (如果请求需要的话) |

例如，计算集群中文档的数量

<http://localhost:9200/_count?pretty>

返回：

|  |
| --- |
| {  "count" : 0,  "\_shards" : {  "total" : 0,  "successful" : 0,  "skipped" : 0,  "failed" : 0  }  } |

#### CRUD