# 1、elasticsearch安装搭建以及基本原理

搜索引擎包括全文索引（百度，谷歌）、目录索引、元搜索引擎、垂直搜索引擎（电商类的，专业领域类等）、集合式搜索引擎、门户搜索引擎与免费链接列表等，公司内部使用垂直搜索

## 1.1 es安装与基本原理

### 1.1 分布式索引概念介绍

1.number\_of\_shards：分片数量，类似于数据库里面分库分表，一经定义不可更改。主要响应写操作

2.number\_of\_replicas：副本数，用于备份分片的，和分片里面的数据保持一致，主要响应读操作，副本越多读取就越快。

3.分布式索引一定要注意分片数量不能更改，所以在创建的时候一定要预先估算好数据大小，一般在8CPU16G的机器上一个分片不要超过500g（不怕的话就500g，只要你索引数据结构优化的好，一般的是没问题）。索引会根据分片的配置来均匀的响应用户请求

4.如果调整了分片数那就要重建索引。

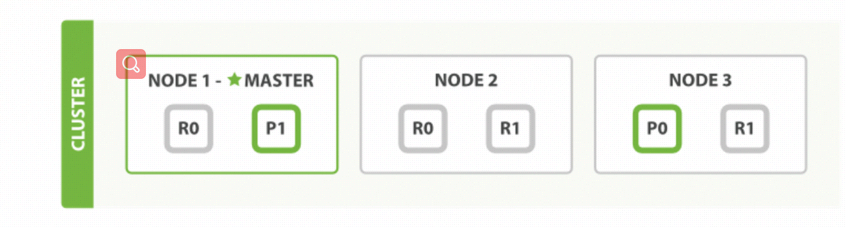
5.为什么shards不能修改：

因为我们采取的是分布式索引，假如我们有4篇文档id分别为：1，2，3，4。相信大家再数据库的分表中都知道有一个取模算法。用它来分配记录到对应的表中，其实es也采取的是这种思路。

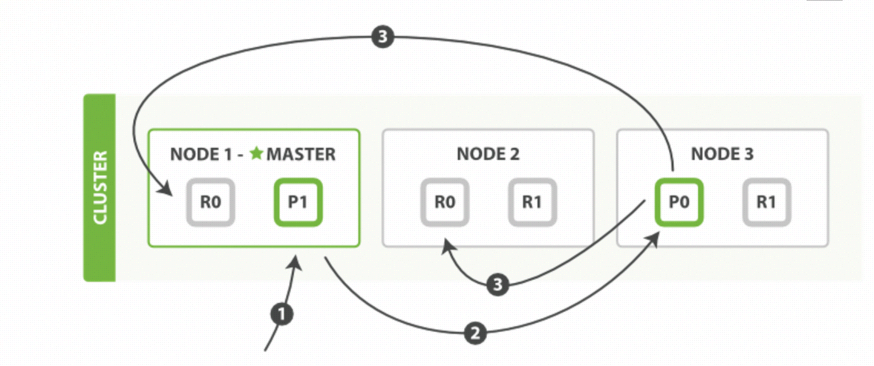
所以1 3 % 2 ==> 1那么就会在第二个分片上。2 4 % 2 ==> 0那么就会在第一个分片上。

如果这时候你把分片数变了，很显然数据就不对了。所以分片数一旦变化需要重新全部建索引

6.索引写操作在集群上的操作：



假设我们集群如上面所示：这上面每个节点都可以接收请求，那么新建的一个流程如下所示：



以下是在主副分片和任何副本分片上面 成功新建，索引和删除文档所需要的步骤顺序：

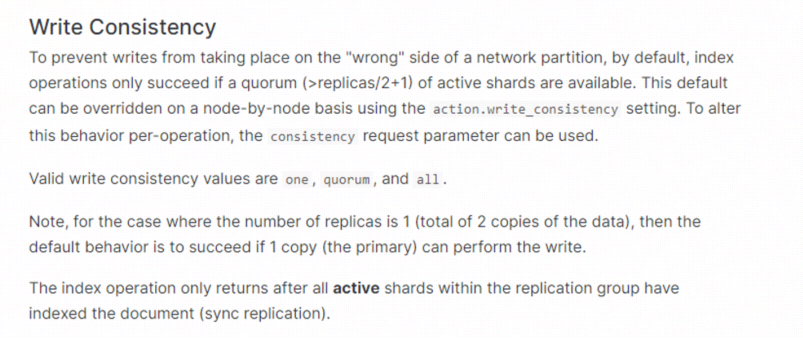
(1)客户端向Node1发送新建、索引或者删除请求。

(2)节点使用文档的\_id确定文档属于分 0。请求会被转发到Node3因为分片 0 的主分片目前被分配在 Node3上。

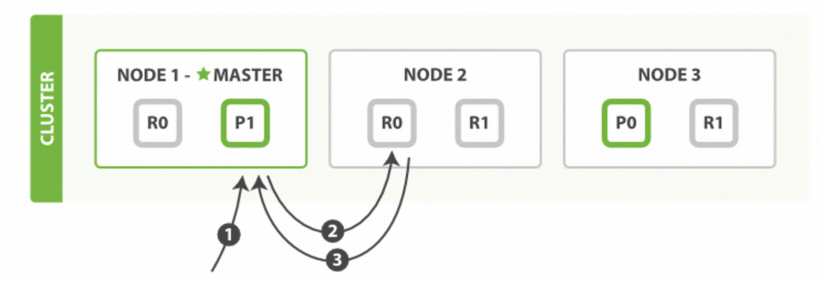
(3)Node 3在主分片上面执行请求。如果成功了，它将请求并行转发到 Node 1和 Node 2 的副本分片上。一旦所有的副本分片都报告成功, Node 3 将向协调节点报告成功，协调节点向客户端报告成功。

在客户端收到成功响应时，文档变更已经在主分片和所有副本分片执行完成，变更是安全的。

（这里课上讲的应该有点问题，现在的es版本是可用shard过半就可以写入，成功后会发送消息给master，我查了es最新的官方文档确实是这样的。所以如果书分片设置为1副本的时候，那默认的就是主分片写如成功了就会返回成功。）以下为es官方文档：



7、读操作



以下是从主分片或者副本分片检索文档的步骤顺序：

(1)客户端向Node1发送获取请求。

(2)节点使用文档的\_id来确定文档属于分片0,分片0的数据在三个节点上都有。 在这种情况下，它会根据负载均衡策略将请求转发到其中一个，比如Node2。

(3)Node2将文档返回给Node1然后将文档返回给客户端。

在处理读取请求时，协调结点在每次请求的时候都会通过轮询所有的副本分片来达到负载均衡，主分片也会接受请求，es有自适应副本功能，使用最合适的副本。

在文档被检索时，已经被索引的文档可能已经存在于主分片上但是还没有复制到副本分片。 在这种情况下，副本分片可能会报告文档不存在，但是主分片可能成功返回文档。

### 1.2 es的安装

### 1.3 es如何保证数据不丢失

在elasticsearch和磁盘之间还有一层cache也就是filesystem cache，大部分新增或者修改，删除的数据都在这层cache中，如果没有flush操作，那么就不能100%保证系统的数据不会丢失，比如突然断电或者机器宕机了，但实际情况是es中默认是30分钟才flush一次磁盘，这么长的时间内，如果发生不可控的故障，那么如何保证数据不丢失呢？

在es里面引入了transaction log（简称translog），这个log的作用就是每条数据的任何操作都会被记录到该log中，非常像Hadoop里面的edits log和hbase里面的WAL log

transaction log的工作流程如下：

（1）当一个文档被索引时，它会被添加到内存buffer里面同时也会在translog里面追加

（2）当每个shard每秒执行一次refresh操作完毕后，内存buffer会被清空但translog不会。

过程如下：

2.1 当refresh动作执行完毕后，内存buffer里面的数据会被写入到一个segment里面，这个还在cache中，并没有执行flush命令

2.2 新生成的segment在cache中，会被打开，这个时候就可以搜索新加的数据

2.3 最后内存buffer里面的数据会被清空

（3）随着更多的document添加，内存buffer区会不断的refresh，然后clear，但translog数量却越增越多

（4）当达到默认的30分钟时候，translog也会变得非常大，这个时候index要执行一次flush操作，同时会生成一个新的translog文件，并且要执行full commit操作，流程如下：

4.1 内存buffer里的所有document会被生成一个新的segment

4.2 然后segment被refresh到系统cache后，内存buffer会被清空

4.3 接着commit point会被写入到磁盘上

4.4 filesystem cache会被flush到磁盘上通过fsync操作

4.5 最后旧的translog会被删除，并会生成一个新的translog

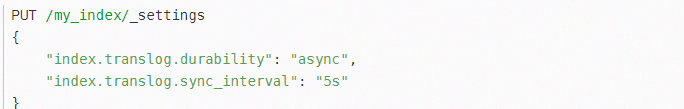
tanslog的作用就是给所有还没有flush到硬盘上的数据提供持久化记录，当es重启时，它首先会根据上一次停止时的commit point文件把所有已知的segments文件给恢复出来，然后再通过translog文件把上一次commit point之后的所有索引变化包括添加，删除，更新等操作给重放出来。

除此之外tanslog文件还用于提供一个近实时的CURD操作，当我们通过id读取，更新 除document时，es在从相关的segments里面查询d的优先访问最新版本的数据。

我们知道执行flush命令之后，所有系统cache中的数据会被同步到磁盘上并且会删除旧的translog然后生成新的translog，默认情况下es的shard会每隔30分钟自动执行一次flush命令，或者当translog变大超过一定的阈值后。

flush命令基本不需要我们手动操作，但当我们要重启节点或者关闭索引时，最好提前执行以下flush命令作为优化，因为es恢复索引或者重新打开索引时，它必须要先把translog里面的所有操作给恢复，所以也就是说translog越小，recovery恢复操作就越快。

如果在一个大数据量的集群中数据并不是很重要，那么就可以设置成每隔5秒进行异步fsync操作translog，配置如下：



## 1.2 es的集群搭建

### 2.1 集群节点的关闭与重启



节点关闭或者索引关闭前执行以下flush操作，避免translog恢复占用太多时间，重启前要关闭集群的reblance（生产中建议关闭reblance功能，避免因为网络抖动等原因造成数据的无用复制），否则重启节点后failed掉的节点会立即开始在其他节点之间开始复制，对恢复速度有影响，用 ***GET  /\_cat/recovery?active\_only=true***确认下高耗时的recovery是哪种类型，是否的确在拷贝文件，还是做translog recovery。

用GET /\_cat/pending\_tasks 确认下master是否有大量的pending tasks在排队等待处理，如果pending tasks很多，更改集群设置的操作有可能会超时没生效，所以GET一下集群的settings看更改是否生效。

集群节点不要太多，避免小的分片的存在，否则新结点加入时master需要重新对unassinged的shard做allocation路由计算，这个过程需要遍历所有unassigned的shard，并发请求到所有结点，询问是否有关于此shard的信息，还需要更新集群的状态数据，并且下发给所有结点。 当shard太多的时候，这个过程会很慢，并且可能会阻塞读取结点状态信息的请求，以至于xpack都无法获取到监控数据。master在做集群状态数据更新的时候，是单线程的，如果有很多状态更新，会非常非常慢。

Recovery是指将一个索引的未分配shard分配到一个结点的过程。 在快照恢复，更改索引复制片数量，结点故障或者结点启动时发生。由于master持有整个集群的状态信息，因此可以判断出哪些shard需要做再分配，以及分配到哪个结点。例如:

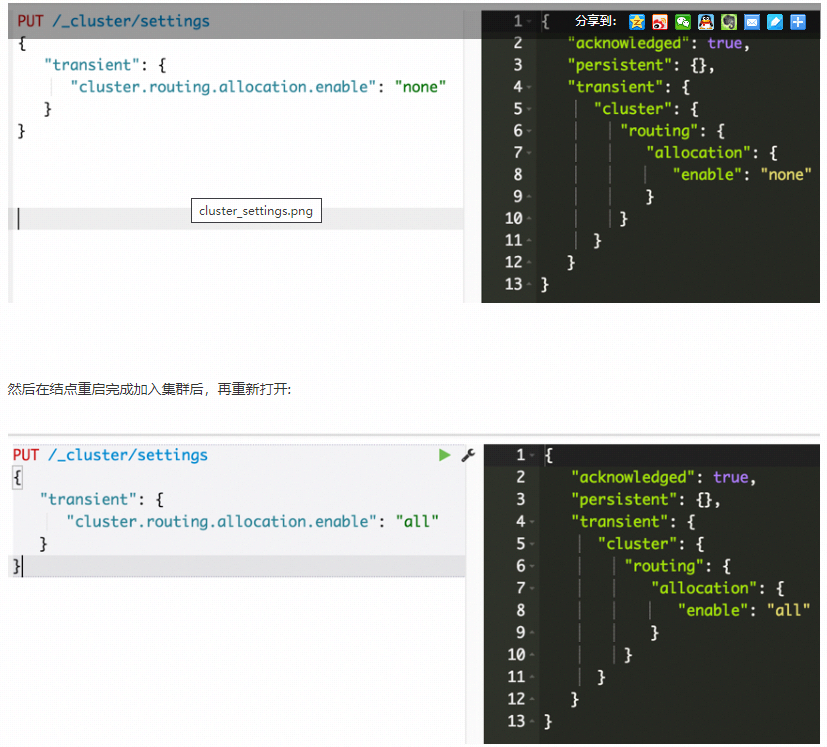
* 如果某个shard主片在，副片所在结点挂了，那么选择另外一个可用结点，将副片分配(allocate)上去，然后进行主从片的复制。
* 如果某个shard的主片所在结点挂了，副片还在，那么将副片升级为主片，然后做主副复制。
* 如果某个shard的主副片所在结点都挂了，则暂时无法恢复，等待持有相关数据的结点重新加入集群后，从结点上恢复主分片，再选择某个结点分配复制片，并从主分片同步数据。
* 

通过此命令查看集群状态

Recovery过程要消耗额外的资源，CPU、内存、结点之间的网络带宽等等。 这些额外的资源消耗，有可能会导致集群的服务能力降级，或者一部分功能暂时不可用。了解一些Recovery的过程和相关的配置参数，对于减小recovery带来的资源消耗，加快集群恢复过程都是很有帮助的。  
  
**减少集群Full Restart造成的数据来回拷贝**  
集群可能会有整体重启的需要，比如需要升级硬件、升级操作系统或者升级ES大版本。重启所有结点可能带来的一个问题: 某些结点可能先于其他结点加入集群。先加入集群的结点可能已经可以选举好master，并立即启动了recovery的过程，由于这个时候整个集群数据还不完整，master会指示一些结点之间相互开始复制数据。 那些晚到的结点，一旦发现本地的数据已经被复制到其他结点，则直接删除掉本地“失效”的数据。当整个集群恢复完毕后，数据分布不均衡显然是不均衡的，master会触发rebalance过程，将数据在结点之间挪动。整个过程无谓消耗了大量的网络流量。合理设置recovery相关参数则可以防范这种问题的发生。  
  
gateway.expected\_nodes  
gateway.expected\_master\_nodes  
gateway.expected\_data\_nodes

说集群里一旦有多少个结点就立即开始recovery过程。不同之处在于，第一个参数指的是master或者data都算在内，而后面两个参数则分指master和data node。  
  
在期待的节点数条件满足之前, recovery过程会等待gateway.recover\_after\_time (默认5分钟) 这么长时间，一旦等待超时，则会根据以下条件判断是否启动:  
gateway.recover\_after\_nodes  
gateway.recover\_after\_master\_nodes  
gateway.recover\_after\_data\_nodes  
  
举例来说，对于一个有10个data node的集群，如果有以下的设置:  
gateway.expected\_data\_nodes: 10  
gateway.recover\_after\_time: 5m  
gateway.recover\_after\_data\_nodes: 8

那么集群5分钟以内10个data node都加入了，或者5分钟以后8个以上的data node加入了，都会立即启动recovery过程。  
  
**减少主副本之间的数据复制**  
如果不是full restart，而是重启单个data node，仍然会造成数据在不同结点之间来回复制。为避免这个问题，可以在重启之前，先关闭集群的shard allocation与reblance（新增节点不要关闭）:



这样在结点重启完成后，尽量多的从本地直接恢复数据。  
但是在ES1.6版本之前，即使做了以上措施，仍然会发现有大量主副本之间的数据拷贝。从表面去看，这点很让人不能理解。 主副本数据完全一致，ES应该直接从副本本地恢复数据就好了，为什么要重新从主片再复制一遍呢？ 原因在于Recovery是简单对比主副本的segment file来判断哪些数据一致可以本地恢复，哪些不一致需要远端拷贝的。而不同结点的segment merge是完全独立运行的，可能导致主副本merge的深度不完全一样，从而造成即使文档集完全一样，产生的segment file却不完全一样。  
为了解决这个问题，ES1.6版本以后加入了synced flush的新特性。 对于5分钟没有更新过的shard，会自动synced flush一下，实质是为对应的shard加了一个synced flush ID。这样当重启结点的时候，先对比一下shard的synced flush ID，就可以知道两个shard是否完全相同，避免了不必要的segment file拷贝，极大加快了冷索引的恢复速度。  
需要注意的是synced flush只对冷索引有效，对于热索引（5分钟内有更新的索引）没有作用。 如果重启的结点包含有热索引，那么还是免不了大量的文件拷贝。因此在重启一个结点之前，最好按照以下步骤执行，recovery几乎可以瞬间完成:

1. 暂停数据写入程序
2. 关闭集群shard allocation
3. 手动执行POST /\_flush/synced
4. 重启结点
5. 重新开启集群shard allocation
6. 等待recovery完成，集群health status变成green
7. 重新开启数据写入程序

**(特别大的)热索引为何恢复慢**  
对于冷索引，由于数据不再更新，利用synced flush特性，可以快速直接从本地恢复数据。 而对于热索引，特别是shard很大的热索引，除了synced flush派不上用场需要大量跨结点拷贝segment file以外，translog recovery是导致慢的更重要的原因。  
  
从主片恢复数据到副片需要经历3个阶段:

1. 对主片上的segment file做一个快照，然后拷贝到复制片分配到的结点。数据拷贝期间，不会阻塞索引请求，新增索引操作记录到translog里。
2. 对translog做一个快照，此快照包含第一阶段新增的索引请求，然后重放快照里的索引操作。此阶段仍然不阻塞索引请求，新增索引操作记录到translog里。
3. 为了能达到主副片完全同步，阻塞掉新索引请求，然后重放阶段二新增的translog操作。  
   可见，在recovery完成之前，translog是不能够被清除掉的（禁用掉正常运作期间后台的flush操作）。如果shard比较大，第一阶段耗时很长，会导致此阶段产生的translog很大。重放translog比起简单的文件拷贝耗时要长得多，因此第二阶段的translog耗时也会显著增加。等到第三阶段，需要重放的translog可能会比第二阶段还要多。 而第三阶段是会阻塞新索引写入的，在对写入实时性要求很高的场合，就会非常影响用户体验。 因此，要加快大的热索引恢复速度，最好的方式是遵从上一节提到的方法: 暂停新数据写入，手动sync flush，等待数据恢复完成后，重新开启数据写入，这样可以将数据延迟影响可以降到最低。  
     
   万一遇到Recovery慢，想知道进度怎么办呢？ CAT Recovery API可以显示详细的recovery各个阶段的状态。 这个API怎么用就不在这里赘述了，参考: <https://www.elastic.co/guide/en/elasticsearch/reference/current/cat-recovery.html>  
   **其他Recovery相关的专家级设置**  
   还有其他一些专家级的设置（参见： <https://www.elastic.co/guide/en/elasticsearch/reference/current/recovery.html>）可以影响recovery的速度，但提升速度的代价是更多的资源消耗，因此在生产集群上调整这些参数需要结合实际情况谨慎调整，一旦影响应用要立即调整回来。 对于搜索并发量要求高，延迟要求低的场合，默认设置一般就不要去动了。 对于日志实时分析类对于搜索延迟要求不高，但对于数据写入延迟期望比较低的场合，可以适当调大indices.recovery.max\_bytes\_per\_sec，提升recovery速度，减少数据写入被阻塞的时长。  
      
   最后要说的一点是ES的版本迭代很快，对于Recovery的机制也在不断的优化中。 其中有一些版本甚至引入了一些bug，比如在ES1.4.x有严重的translog recovery bug，导致大的索引trans log recovery几乎无法完成 （[issue #9226](https://github.com/elastic/elasticsearch/issues/9226)）  。因此实际使用中如果遇到问题，最好在Github的issue list里搜索一下，看是否使用的版本有其他人反映同样的问题。

## 1.3索引的创建于数据添加

**参考资料：**

<https://www.elastic.co/guide/cn/elasticsearch/guide/cn/index.html>

这个是中文版的，虽然这个是es2.x版本的，但是大家可以发现其核心的技术点并没有太多的变化，仍然具备参考价值。

<https://www.elastic.co/guide/en/elasticsearch/reference/6.0/search-aggregations-metrics-avg-aggregation.html>

6.x英文版api资料

Es名词定义：

对外提供检索服务，http或者transport协议对外提供搜索。Restful的json。Es6.6.0 es2.x。

索引=数据库：很多个

类型（type）=表 =》 es6.x只有一个type，之前可以建很多，es7.x就没有这个type了（\_doc）。

文档=行数据 docment。

Filed

缺点：nosql 非关系型的，没有办法链接查询的，也就是跨索引查询。



Es基础类型：

Text：字符串类型，可以被分析；

Keyword：不能被分析，只可以精确匹配的字符串类型

Date：日期类型，通常配合format使用 比如{“type”:”date”,”format”:”yyyy-MM-dd”}

Long,integer,short…

Boolean

Array：数组类型

Object：一般是json

Ip：ip地址

geo\_point：地理位置 {

“lat”:

“lon”:

}

Put重建索引只会更新数据，不会改变分片副本等配置值

指定分片与副本

PUT /test2

{

"settings": {

"number\_of\_shards": 1,

"number\_of\_replicas": 1

}

}

更新副本数量

PUT /test2/\_settings

{

"number\_of\_replicas": 0

}

添加数据

PUT /test/\_doc/1

{

"name":"kevin",

"age1":"30"

}

结构化索引创建

PUT /test

{

"settings": {

"number\_of\_shards": 1,

"number\_of\_replicas": 1

},

"mappings": {

"\_doc":{

"properties":{

"name":{

"type":"text","analyzer":"ik\_max\_word","search\_analyzer":"ik\_smart"

},

"sname":{

"type":"text","analyzer":"ik\_smart"

},

"enname":{

"type":"text","analyzer":"english"

},

"age":{

"type":"intrger"

}

}

}

}

}

同义词部分的索引

PUT /test11

{

"settings": {

"number\_of\_replicas": 1,

"number\_of\_shards": 1,

"analysis": {

"filter": {

"my\_synonym\_filter": {

"type": "synonym",

"synonyms\_path": "analysis-ik/synonyms.txt"

}

},

"analyzer": {

"ik\_syno": {

"type": "custom",

"tokenizer": "ik\_smart",

"filter": [

"my\_synonym\_filter"

]

},

"ik\_syno\_max": {

"type": "custom",

"tokenizer": "ik\_max\_word",

"filter": [

"my\_synonym\_filter"

]

}

}

}

},

"mappings": {

"\_doc": {

"properties": {

"name": {

"type": "text",

"analyzer": "ik\_syno\_max",

"search\_analyzer": "ik\_syno"

}

}

}

}

}

使用put会重建索引，如果只修改数据可以使用post

POST /test/\_doc/1/\_update

{

"doc":{

"name":"kevin1"

}

}

如果post不加\_update会更新所有数据等同于put

POST /test/\_doc/2

{

"age":3

}

加上\_create后只能执行一次，否则会报错

POST /test/\_doc/3/\_create

{

"age":3

}

当post创建数据不指定id时，es会指定一个默认ID，并且可以重复创建

POST /test/\_doc{

"age":3

}

{

"\_index" : "test",

"\_type" : "\_doc",

"\_id" : "PmzHeW8BTfPcbtGGONN4",

"\_score" : 1.0,

"\_source" : {

"age" : 4

}

# 2、数据的查询

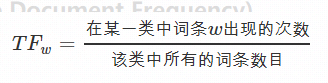
## 2.1 kibana查询操作

GET：term为精确匹配，match为模糊查询，query\_string

创建索引的时候，因为分析器的原因，大写的英文数据如DOG会自动变为小写的的dog，match与query\_string查询时内容也会被分析会自动判断查询为dog，而term会直接查询DOG，导致搜索不到

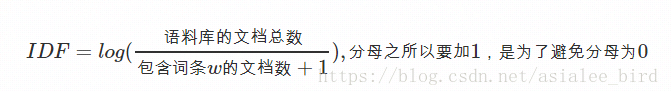
TF-IDF：

tf: ：词频这个document文档包含了多少个这个词n，包含越多表明越相关



df：有m个文档包含了这个词；这个词所有的文档都有，那这个就没有意义。

idf: ：逆文档频率1/m， 包含该词的文档总数目，idf值越大即m越小，则该词对文档的区分度越好

 TF-IDF：tf\*idf= n/m

tfNom: : 根据field长度做归一化，文档内出现频率越高，field越短越相关操作不管是字符与还是或，按照逻辑关系命中后相加得分

### 1 term精确匹配

GET /book/\_search

{

"query": {

"term": {

"bookName": {

"value": "童话故事大全"

}

}

}

Term查询值使用value，match查询值使用query

Terms查询可以使用include包含，exclude排除

POST /nba/\_search

{

"aggs": {

"aggsTeamName": {

"terms": {

"field": "teamNameEn",

"include": [

"Lakers",

"Rockets",

"Warriors"

],

"exclude": [

"Warriors"

],

"size": 30,

"order": {

"avgAge": "desc"

}

},

"aggs": {

"avgAge": {

"avg": {

"field": "age"

}

}

}

}

},

"size": 0

}

在my\_index 索引中查找与文档id为2的color字段有共同值得选项

GET my\_index/\_search?pretty

{

"query": {

"terms": {

"color" : {

"index" : "my\_index",

"id" : "2",

"path" : "color"

}

}

}

}

### 2 短语查询

Match\_phase:查询确切的phase即搜索的值，例如hello world，返回满足如下两个条件的数据，1、match\_phase中所有的term都出现在待查字段中 2、待查字段中所有term都必须和match\_phase具有相同的顺序

match\_phrase\_prefix:可实现即时的短语匹配查询，与match\_phrase查询相同，不同的是它将查询字符串的最后一个词作为前缀使用，即查询时先匹配最后一个词

"match\_phrase\_prefix" : {

"brand" : {

"query": "walker johnnie bl", (1)

"slop": 10，

"max\_expansions": 50

}

}

Slop设置为使搜索的字符串各个词的顺序匹配不那么严格

max\_expansions：prefix 查询存在严重的资源消耗问题，短语查询的这种方式也同样如此。前缀 a 可能会匹配成千上万的词，这不仅会消耗很多系统资源，而且结果的用处也不大。

可以通过设置 max\_expansions 参数来限制前缀扩展的影响，一个合理的值是可能是 50 ，参数 max\_expansions 控制着可以与前缀匹配的词的数量，它会先查找第一个与前缀 bl 匹配的词，然后依次查找搜集与之匹配的词（按字母顺序），直到没有更多可匹配的词或当数量超过 max\_expansions 时结束。

### 3.多字段查询

Multi\_match：多字段查询

GET /book/\_search

{

"query": {

"multi\_match": {

"query": "大自然的旅行故事",

"fields": ["bookName","discription"]

}

}

}

此时为默认best\_fields类型，即取两边分词分数高的那个字段，多字段查询中如果字段中匹配到多个分词，则此字段分数为多个分词分数之和

字段加权重，即在bookName字段的分数乘10

"fields": ["bookName^10","discription"]

添加tie\_breaker：

"tie\_breaker": 0.3 放在fields下面

表示单一分词在两个字段中的得分取最大值加上其余值的0.3倍

（单一字段下所有分词查询取最大值加上其余分词的总分数乘以tie\_breaker）

多字段分值相加：

"type": "most\_fields" 即该条数据的总分数为多个字段分数之和

以分词为单位分别在每个字段里面得分取最大相加，非常适用于以词作为权重的系统：

"type": "cross\_fields"，以分词为单位，单个分词的得分取该分词在所有字段中的最大值，然后多个分词的得分之和为该条数据的总得分

### 4.排序、分页

查询所有

GET /test/\_search

或者

GET /test/\_search

{

"query":{

"match\_all": {}

}

}

分页查询

GET /test/\_search

{

"query":{

"match\_all": {}

},

"from":0,

"size":2

}

带条件查询

GET /test/\_search

{

"query": {

"match": {

"name": "kevin"

}

}

}

### 5.最小匹配查询

GET /book/\_analyze

{

"field": "bookName",

"text": "安徒生的大自然童话故事"

}

GET /book/\_search

{

"query": {

"match": {

"bookName": {

"query": "安徒生的大自然童话故事",

"operator": "or",

"minimum\_should\_match": 2

}

}

}

}

"operator": "or"，or代表搜索的字符串所分的词用于检索式有一个即可，and代表需要包含所有分词

"minimum\_should\_match": 2 必须包含的分词个数，用于控制返回结果数，此时operator需要为or

### 6、analyze、分词器与validate-query

**analyze api** 根据所选择的分词器将text数据进行分词操作

GET /test/\_analyze

{

"field": "name",

"text": "my name is zhaoyun and i like eating apples and running"

}

从上面的结果中我们可以看出es对其进行了分词。Es中主要有三种分词：

(1)stander中文的默认分词，这是按照一个个字分开。此种分词非常占空间会导致倒排索引很大，而且在搜索的时候也会搜出很多不相干的东西。优点就是搜的多。。

(2)English分词器：会提取词干和去掉停用词，并将大写转为小写。假设name不采用英文分词：

PUT /test/\_doc/1

{

"name": "eating apples,

"age": 30

}

然后搜索

GET /test/\_search

{

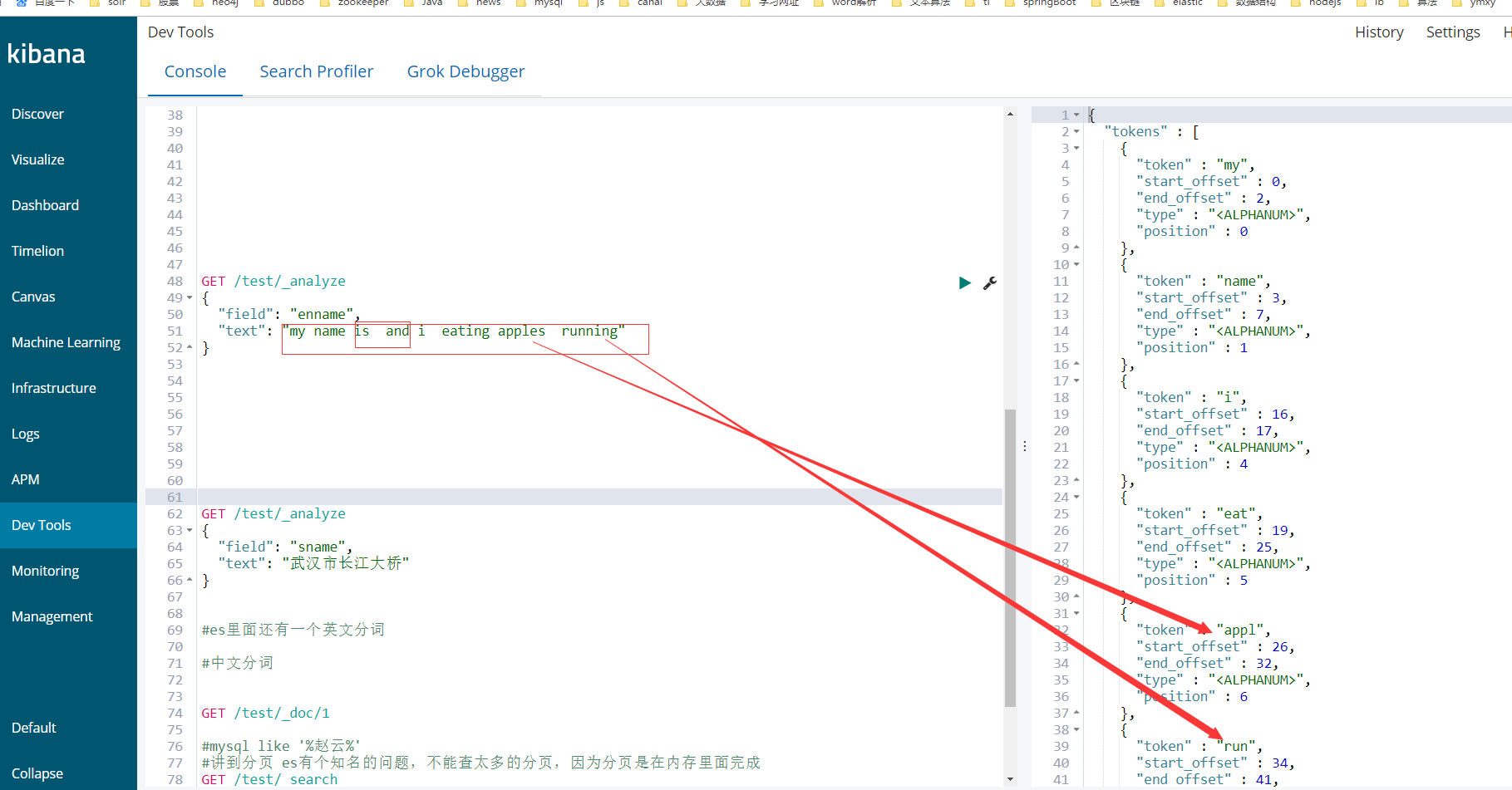
"query":{

"match": {"name":"eat"}

}

}

没搜到结果，我们使用后使用analyze api查看分析处理结果，可以看到没有分出eat，所以搜不到，改成用english分词器做。就可以看到以下的情况



可以看到running和apples被提取成了词干，提取算法其实就是一个映射感兴趣的同学可以自己取看看。而and 和 is被当作停用词去掉了。

（3）IK分词器：他会有两种模式一种smart，一种max\_word. 建索引的时候 用ik\_max\_word，搜索的时候用smart（查询时使用尽量长的分词）。如果需要详细理解里面的原理的，请看算法课的第23节课。讲的相当详细。

Ik分词器某个分词在存储包含该分词的文档id时使用的链表来保存，热点分词的链表可能很长，所以会进行压缩，查找会使用跳表进行查询

**分词的妙用：**

是不是就意味着stander分词没用了？其实并不是这样的哦

#托底，搜江大桥没有，在建了ik的字段，在建一个一样的stander的字段。如果ik搜不到 就可以搜这一个stander分词的，这样保证会有结果。但是慎用，因为占空间，有些特殊的系统可以使用。

#其实ik也还有一个解决的办法 叫砍词：江大桥 我可以砍掉一个词，我砍掉江 就出来了。砍词的策略可以自定义

#江大桥：电商中。我们系统假设有大桥这个品牌。 我会一个个的是去试一下，比如可以用字符串匹配找出大桥。也有很多系统很粗暴，直接从第一个字开始砍，一直砍到有为止。

#既有英文又有中文的 直接选ik

#如果不用砍词那就要去词库加词，比如加入江大桥就可以了。具体路径是在es的：

/apps/svr/elasticsearch-node1/config/analysis-ik/main.dic

注意集群的话那就要所有的es都需要加的哦

验证查询validate-query：

GET /book/\_validate/query?explain

{

"query": {

"multi\_match": {

"query": "童话故事的大选",

"fields": ["bookName","discription"]

}

}

}

当不加？explain时，返回查询是否合法的信息，当想了解具体原因时需要加上

也可以最直接在query上面加上"explain":true,

GET /book/\_search

{

"explain": true,

"query": {

"multi\_match": {

"query": "大自然的旅行故事",

"fields": ["bookName","discription"],

"type": "cross\_fields"

}

}

}

### 7、queryString

GET /book/\_search

{

"query": {

"query\_string": {

"fields": ["bookName"],

"query": "大自然 AND 旅行"

}

}

}

default\_field： 查询的字段，支持多字段查询"fields": ["bookName"，”description”]

default\_operator:AND代表包含两个分词，ＯＲ为或，必须大写,or为默认值，可以复合使用”(this AND that) OR thu\*”

### ８、bool过滤器

Bool过滤器由三部分组成：

Must: 所有语句都必须匹配，等价于and

Must\_not: 所有语句都不能匹配，与not等价

Should：至少有一个语句匹配，与or等价

Should与 must\_not 混合使用，should的两个条件满足一个即可，并且必须满足must\_not

GET /test/\_search

{

"query": {

"bool": {

"should": [

{

"term": {

"name": "java"

}

},

{

"term": {

"sname": "develop"

}

}

],

"must\_not": {

"term": {

"name": "java develop"

}

}

}

}

}

在should中嵌套一个bool查询，即bool的内部第一个条件或者嵌套bool为true都能查出来

GET /test/\_search

{

"query" : {

"bool" : {

"should" : [

{ "term" : {"name" : "java"}},

{ "bool" : {

"must" : [

{ "term" : {"sname": "develop"}},

{ "term" : {"age" : 23}}

]

}}

]

}

}

}

当bool查询为should条件时，下方的filter过滤should查出的数据，此部分数据有分数，filter仍然会过滤其他不符合should的数据，此部分数据分值为0，如果should没有结果返回，filter仍会执行所得结果分数为0

GET /test/\_search

{

"query" : {

"bool" : {

"should" : [

{ "term" : {"name" : "java"}},

{ "bool" : {

"must" : [

{ "term" : {"sname": "develop"}},

{ "term" : {"age" : 23}}

]

}}

]

,

"filter": {

"range":{

"age":{

"lte":34,

"gte":20

}

}

}

}

}

}

Filter查询是没有分数的：使用order过滤，避免没有分数

GET /book/\_search

{

"query": {

"bool": {

"filter": [

{

"range": {

"commentNum": {

"lte": 2000,

"gte": 1

}

}

},

{

"term": {

"author":"朱奎"

}

}

]

}

} ,

"sort": [

{

"commentNum": {

"order": "desc"

}

}

]

}

### 9、Function Score 自定义评分函数

score\_mode:指定如何计算组合得分是（即多个函数之间的分数计算）

multiply： 分数相乘（默认）即每个函数得分相乘

sum：分数相加

avg：分数是平均值

first：具有匹配过滤器的第一个函数被应用

max：使用最高分

min：使用最低分

max\_boost: 将新分数限定为不高于某值，如果大于则赋值为限定值

min\_score：将打分低于此值的数据过滤掉

boost\_mode:代表查询的原有得分（即主查询得分）

multiply：查询分数与功能分数（score\_mode）相乘（默认）

replace：仅使用功能分数，查询分数将被忽略

sum：查询分数与功能分数相加

avg：平均分

max：查询分数与功能分数的最大值

min：查询分数与功能分数的最小值

GET /\_search

{

"query": {

"function\_score": {

"query": { "match\_all": {} },

"boost": "5",

"functions": [

{

"filter": { "match": { "test": "bar" } },

"random\_score": {},

"weight": 23

},

{

"filter": { "match": { "test": "cat" } },

"weight": 42

}

],

"max\_boost": 42,

"score\_mode": "max",

"boost\_mode": "multiply",

"min\_score" : 42

}

}

}

**Es中预定义了一些函数，weight，field\_value\_factor,random\_score,衰减函数 linear、exp、gauss，script\_score**

Weight：权重值，单个函数得出的结果乘以权重值为当前函数的得分

**random\_score：**在相同的分片你片中，如果字段值相同，则得到的分值相同，所以字段最好选择唯一性字段，可以使用\_seq\_no字段，但是该字段在你数据更新时会变动，加入seed种子值可以使每次访问得到相同的分值

"random\_score": {

"seed": 10,

"field":"\_seq\_no"

}

**field\_value\_factor：主查询分值乘以该函数返回值**

factor: 分数乘以相应的倍数

modifier：

none：不使用任何算法公式（默认）

log：常用对数公式

log1p：给字段的值加1，然后使用普通的对数公式，加1使其更加服从高斯分布，避免重复值问题

log2p：给字段的值加2，然后使用普通的对数公式

ln：natural logarithm 自然对数，以e为底的对数

ln1p：给字段的值加1，然后使用自然对数公式

ln2p：给字段的值加2，然后使用自然对数公式

square：平方值

sqrt：开平方

reciprocal：取倒数

missing：如果文档中不包含此字段，则此字段的初始值为missing值

GET /book/\_search

{

"query": {

"function\_score": {

"query": {

"multi\_match": {

"query": "大自然的旅行故事",

"fields": ["bookName","discription"],

"operator": "or",

"type": "most\_fields"

}

},

"functions": [

{

"field\_value\_factor": {

"field": "commentNum",

"modifier": "log2p",

"factor": 8

“missing”:1

}

}

**script\_score:使用自定义的脚本实现控制评分计算**

script中使用的字段需要为keyword类型，text类型需要将其fielddata=true

如果要忽略主查询得分，可使用boost\_mode：replace

GET /\_search

{

"query": {

"function\_score": {

"query": {

"match": { "message": "elasticsearch" }

},

"script\_score" : {

"script" : {

"params": {

"a": 5,

"b": 1.2

},

"source": "params.a / Math.pow(params.b, doc['name'].value)"

}

}

}

}

}

],

"score\_mode": "sum",

"boost\_mode": "sum"

}

}

}

Decay functions：衰减函数

有三种衰减函数—— linear 、 exp 和 gauss （线性、指数和高斯函数），它们可以操作数值、时间以及经纬度地理坐标点这样的字段。所有三个函数都能接受以下参数

Origin：*中心点* 或字段可能的最佳值，落在原点 origin 上的文档评分 \_score 为满分 1.0

Scale：衰减率，即一个文档从原点 origin 下落时，评分 \_score 改变的速度。（例如，每 £10 欧元或每 100 米）。

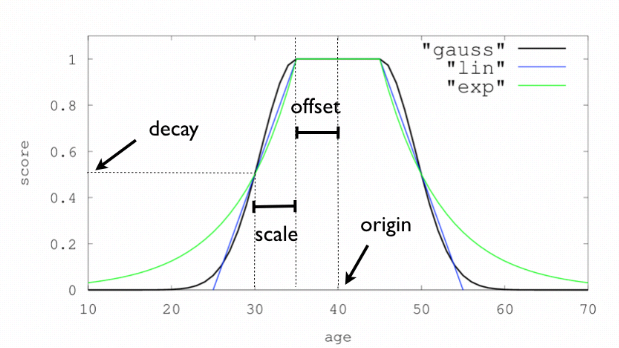
Decay：从原点 origin 衰减到 scale 所得的评分 \_score ，默认值为 0.5

Offset：以原点 origin 为中心点，为其设置一个非零的偏移量 offset 覆盖一个范围，而不只是单个原点。在范围 origin +/- offset 内的所有评分 \_score 都是 1.0

Liner:  线性函数是条直线，一旦直线与横轴 0 相交，所有其他值的评分都是 0.0

Exp:指数函数先剧烈衰减然后变缓

Gauss： 高斯函数是钟形的——它的衰减速率是先缓慢，然后变快，最后又放缓



### 10、通配符、正则、模糊查询

1、通配符

Query的内容支持？与\*，？可以代表任意一个字符，\*可以代表任意个字符（包括0个），例如\*elc，查询elc结尾的数据

2、正则

如果要在query的内容中使用正则表达式，在两端加上正斜杠/即可。比如name:/ob[am]{2}a/

也可以使用通配符进行匹配

GET /\_search

{

"query": {

"regexp": {

"user": {

"value": "k.\*y",

"flags" : "ALL",

"max\_determinized\_states": 10000,

"rewrite": "constant\_score"

}

}

}

}

Field：user

Flags：ALL、COMPLEMENT、INTERVAL、INTERSECTION、ANYSTRING

ALL:默认值，允许所有操作符

COMPLEMENT：使用~进行排除操作a~bc # matches 'adc' and 'aec' but not 'abc'

INTERVAL：<>进行数字排行匹配

foo<1-100> # matches 'foo1', 'foo2' ... 'foo99', 'foo100'

foo<01-100> # matches 'foo01', 'foo02' ... 'foo99', 'foo100'

INTERSECTION：&功能相当于and

aaa.+&.+bbb # matches 'aaabbb'

ANYSTRING：@，类似于\*可以匹配所有字符串，不同的操作符可以组合使用

@&~(abc.+) # matches everything except terms beginning with 'abc'

max\_determinized\_states：Lucene将每个正则表达式转换为一个包含若干确定状态的具体查询条件，此参数限制分解出的条件个数，默认值10000

rewrite：此参数值得改变会影响搜索性能与相关性，精通es之前不要改动此参数，es推荐使用constant\_score, constant\_score\_boolean, top\_terms\_boost\_N

constant\_score：默认值

constant\_score\_boolean：将原始查询改为bool查询

scoring\_boolean：

top\_terms\_blended\_freqs\_N：

top\_terms\_boost\_N：

top\_terms\_N：

3、模糊查询

在query的搜索值可以使用~来进行模糊查询如“query”:”name:obalv~”

在~后面加数字来指定允许有几个字母错误，如“query”:”name:obalv~1”只能允许一个字母错误

### 11、range范围查询

GET /\_search

{

"query": {

"range" : {

"age" : {

"gte" : 10,

"lte" : 20,

"boost" : 2.0

}

}

}

}

Gt：大于

Gte：大于等于

Lt：小于

Lte：小于等于

Format：用于解析时间值

Relation：

Time\_zone:

Boost:

时间参数释义：

+1h: Add one hour

-1d: Subtract one day

/d: Round down to the nearest day

Assuming now is 2001-01-01 12:00:00, some examples are:

now+1h: now in milliseconds plus one hour. Resolves to: 2001-01-01 13:00:00

now-1h: now in milliseconds minus one hour. Resolves to: 2001-01-01 11:00:00

now-1h/d: now in milliseconds minus one hour, rounded down to UTC 00:00. Resolves to: 2001-01-01 00:00:00

2001.02.01\|\|+1M/d: 2001-02-01 in milliseconds plus one month. Resolves to: 2001-03-01 00:00:00

gt

Rounds up to the first millisecond not covered by the rounded date.

For example, 2014-11-18||/M rounds up to 2014-12-01T00:00:00.000, excluding the entire month of November.

gte

Rounds down to the first millisecond.

For example, 2014-11-18||/M rounds down to 2014-11-01T00:00:00.000, including the entire month.

lt

Rounds down to the last millisecond before the rounded value.

For example, 2014-11-18||/M rounds down to 2014-10-31T23:59:59.999, excluding the entire month of November.

lte

Rounds up to the latest millisecond in the rounding interval.

For example, 2014-11-18||/M rounds up to 2014-11-30T23:59:59.999, including the entire month.

### 12、同义词

更新同义词词库、ik分词器main.dic词库都需要重建索引，ik分词器的更改需要先重启

在"analysis-ik/synonyms.txt"路径下新建同义词文档

PUT /test11

{

"settings": {

"number\_of\_replicas": 1,

"number\_of\_shards": 1,

"analysis": {//指定分析器

"filter": {

"my\_synonym\_filter": {//自定义同义词分析器

"type": "synonym",

"synonyms\_path": "analysis-ik/synonyms.txt"

}

},

"analyzer": {//分析器

"ik\_syno": {//用于搜索

"type": "custom",

"tokenizer": "ik\_smart",

"filter": [

"my\_synonym\_filter"

]

},

"ik\_syno\_max": {//用于分词

"type": "custom",

"tokenizer": "ik\_max\_word",

"filter": [

"my\_synonym\_filter"

]

}

}

}

},

"mappings": {

"\_doc": {

"properties": {

"name": {

"type": "text",

"analyzer": "ik\_syno\_max",

"search\_analyzer": "ik\_syno"

}

}

}

}

}

PUT /test11/\_doc/1

{

"name":"苹果"

}

GET /test11/\_analyze

{

"field": "name",

"text": "气质好"

}

GET /test11/\_search

{

"query": {

"match": {

"name": "apple"

}

}

}

### 13、聚合查询

聚合查询：

**指标聚合**：对⼀个数据集求最⼤、最⼩、和、平均值等指标的聚合，在ES中称为指标聚合

**桶聚合**：⽽关系型数据库中除了有聚合函数外，还可以对查询出的数据进⾏分组group by，再在组上进⾏指标聚合。在ES中称为桶聚合

Bucket，分桶类型，类似于group by，按照一定的规则将文档分配到不同的桶中，以达到分类分析的目的

Terms根据term分桶，如果是text类型，则按照分词后的结果分桶

range（指定数值范围来设置分桶）

data range（指定日期的范围）

histogram（以固定间隔的策略分割）

date histogram（针对日期)

Metric，指标分析类型，如最大值、最小值、平均值等等

单值分析 min、max、avg、sum、cardinality(distinct count)

多值分析 stats,extended stats percentile（百分位数） percentile rank top hits

Pipeline,管道分析类型，基于上一级的聚合分析结果

Parent结果内嵌到现有的聚合分析结果

Sibling结果与现有聚合分析结果同级

Matrix,矩阵分析类型

语法

aggs->聚合名字->聚合类型->聚合内容

#### 指标聚合：

min、max、avg、sum、cardinality(distinct count 去重)value\_count 统计非空字段的文档数

stats统计count max min avg sum 5个值

Extended stats ⽐stats多4个统计结果： 平⽅和、⽅差、标准差、平均值加/减两个标准差的区间

Percentiles 占比百分位对应的值统计，默认返回（1,5,25,50,75,95,99）分位上的最大值，可以了解百分位数的概念及运算，大意是将数据从小到大排列，1%即前1%的数

size 查询结果，查询的记录条数为0，不返回记录详情，也就是，只返回聚合结果

POST /nba/\_search

{

"query": {

"term": {

"teamNameEn": {

"value": "Rockets"

}

}

},

"aggs": {

"avgAge": {

"avg": {

"field": "age"

}

}

},

"size": 0

}

百分位数

POST /nba/\_search

{

"query": {

"term": {

"teamNameEn": {

"value": "Rockets"

}

}

},

"aggs": {

"pecentAge": {

"percentiles": {

"field": "age"

}

}

},

"size": 0

}

指定具体的百分位数

POST /nba/\_search

{

"query": {

"term": {

"teamNameEn": {

"value": "Rockets"

}

}

},

"aggs": {

"percentAge": {

"percentiles": {

"field": "age",

"percents": [

20,

50,

75

]

}

}

},

"size": 0

}

#### 桶聚合：类似于MySQL的group by

1、terms Aggregation 根据字段项分组聚合

Group by age

POST /nba/\_search

{

"query": {

"term": {

"teamNameEn": {

"value": "Rockets"

}

}

},

"aggs": {

"aggsAge": {

"terms": {

"field": "age",

"size": 10

}

}

},

"size": 0

}

⽕箭队根据年龄进⾏分组，分组信息通过年龄从⼤到⼩排序 ，组内升序，组之间根据组内最大年龄降序

POST /nba/\_search

{

"query": {

"term": {

"teamNameEn": {

"value": "Rockets"

}

}

},

"aggs": {

"aggsAge": {

"terms": {

"field": "age",

"size": 10,

"order": {

"\_key": "desc"

}

}

}

},

"size": 0

}

⽕箭队根据年龄进⾏分组，分组信息通过⽂档数从⼤到⼩排序 (通过⽂档数)

POST /nba/\_search

{

"query": {

"term": {

"teamNameEn": {

"value": "Rockets"

}

}

},

"aggs": {

"aggsAge": {

"terms": {

"field": "age",

"size": 10,

"order": {

"\_count": "desc"

}

}

}

},

"size": 0

}

嵌套聚合每⽀球队按该队所有球员的平均年龄进⾏分组排序 (通过分组指标值)

POST /nba/\_search

{

"aggs": {

"aggsTeamName": {

"terms": {

"field": "teamNameEn",

"size": 30,

"order": {

"avgAge": "desc"

}

},

"aggs": {

"avgAge": {

"avg": {

"field": "age"

}

}

}

}

},

"size": 0

}

湖⼈和⽕箭队按球队平均年龄进⾏分组排序 (指定值列表)

POST /nba/\_search

{

"aggs": {

"aggsTeamName": {

"terms": {

"field": "teamNameEn",

"include": [

"Lakers",

"Rockets",

"Warriors"

],

"exclude": [

"Warriors"

],

"size": 30,

"order": {

"avgAge": "desc"

}

},

"aggs": {

"avgAge": {

"avg": {

"field": "age"

}

}

}

}

},

"size": 0

}

湖⼈和⽕箭队按球队平均年龄进⾏分组排序 (正则表达式匹配值)

POST /nba/\_search

{

"aggs": {

"aggsTeamName": {

"terms": {

"field": "teamNameEn",

"include": "Lakers|Ro.\*|Warriors.\*",

"exclude": "Warriors",

"size": 30,

"order": {

"avgAge": "desc"

}

},

"aggs": {

"avgAge": {

"avg": {

"field": "age"

}

}

}

}

},

"size": 0

}

2、 Range Aggregation范围分组聚合

NBA球员年龄按20,20-35,35这样分组

POST /nba/\_search

{

"aggs": {

"ageRange": {

"range": {

"field": "age",

"ranges": [

{

"to": 20

},

{

"from": 20,

"to": 35

},

{

"from": 35

}

]

}

}

},

"size": 0

}

NBA球员年龄按20,20-35,35这样分组 (起别名)

POST /nba/\_search

{

"aggs": {

"ageRange": {

"range": {

"field": "age",

"ranges": [

{

"to": 20,

"key": "A"

},

{

"from": 20,

"to": 35,

"key": "B"

},

{

"from": 35,

"key": "C"

}

]

}

}

},

"size": 0

}

3、Date range Aggregation 时间范围分组聚合

NBA球员按出⽣年⽉分组

POST /nba/\_search

{

"aggs": {

"birthDayRange": {

"date\_range": {

"field": "birthDay",

"format": "MM-yyy",

"ranges": [

{

"to": "01-1989"

},

{

"from": "01-1989",

"to": "01-1999"

},

{

"from": "01-1999",

"to": "01-2009"

},

{

"from": "01-2009"

}

]

}

}

},

"size": 0

}

4、Date Histogram Aggregation 时间柱状图聚合

按天、月、年进行聚合统计。可按year，quarter季度，month，day，hour，minute，second间隔聚合

POST /nba/\_search

{

"aggs": {

"birthday\_aggs": {

"date\_histogram": {

"field": "birthDay",

"format": "yyyy",

"interval": "year"

}

}

},

"size": 0

}

排序加聚合查询

GET /test/\_search

{

"query": {

"match": {

"name": "kevin"

}

},

"sort": [

{

"age": {

"order": "desc"

}

}

],

"aggs":{

"group\_by\_age":{//自定义的名称

"terms": {//统计

"field": "age",

"size": 4 分组后显示的分组个数

}

}

}

}

# 3、lucene

Lucene与搜索引擎的关系：

Lucene是检索系统的核心框架，它类似spring Jar包。只要把Jar引入自己的工程就可以用，不依赖es 和 solr，单机版 -> 集群版。

特别适用于小规模的检索，比如10万左右的数据，用mysql达不到效果，用es和solr等成本太高，就可以直接用Lucene。直接嵌入到你的工程，不需要部署任何服务。

Mysql like ‘%四川省%’

Solr: 他们是基于Lucene封装的一个成熟的产品，提供很多的api尤其是提供了分布式索引，大大的提高了存储能力和搜索效率。适用于千万级以上分单机版和集群版：最大的区别

zookeeper:协调 管理服务 分布式锁 节点 统一动态配置的功能。分布式

Es:跟solr一样他也是基于Lucene开发的一款分布式搜索引擎。分单机版和集群版 这个集群版不需要zk;

## 3.1 trie树

Es中用于获取分词存不存在

1.什么是Trie树：trie树就是我们平常说的字典树又称前缀树，它是一种专门用来处理字符串匹配的数据结构。特别适合用来在很多字符串中快速查找某一个特定的字符串。前缀树，赫夫曼树，前缀编码

2.Trie的数据结构：假设我们有以下几个英文单词：my name apple age sex，假如我们要查找里面某一个字符串是否存在，你怎么去找呢？

散列表：hashMap

如果利用Trie树我们该怎么来解决上面的查找问题呢？

我们可以将上面的字符串变成以下结构，利用字符串的公共前缀，将重复的合在一起组成一颗树，即为我们所要讲的trie树。

3.Trie树的构建：我们要先将词分成一个个的字母，然后再依次插入到树中。如右图所示，根节点root，如果我们要插入app

则首先将app分成：a,p,p，然后从root点开始，一层层的插入，注意的是

P会挂在a下面，后面的一个p会挂在前面的p上。单词的末尾我们就用紫色表示。

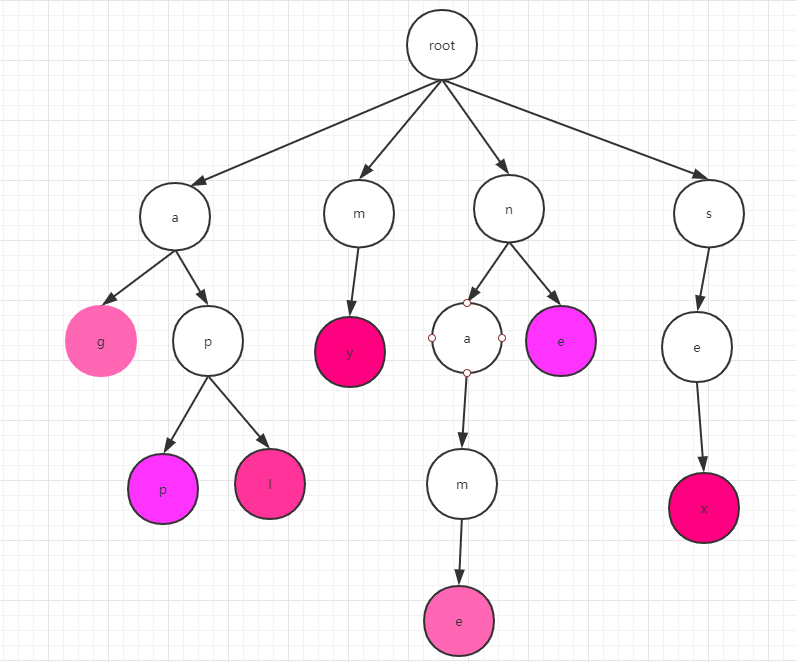
这里需要注意我们插入的时候每一层的字母都是有序的。

4.Trie的查找：

查找我们就从root点开始，再第一层找第一个字母，依次往下找到我们所要的单词。

注意要找到末尾的标记才算完成一个单词的查找。比如app，我们要找ap

虽然字典树里面有ap，但是这个p不是紫色那么ap还是不存在字典树中的。



5.Trie树的实现:其实Trie树就是一颗多叉树。这里我们应该要想到B+Tree&B-Tree，是有些类似的。

Trie树又是巧妙的利用了数组的下标，因为英文字母刚好是26个，所以我们可以开一个26长度的数组

A[] = new int[26];

A[0] = ‘a’ => 下标就是’a’-97 这里刚好就是0，利用的是ascii计算。

所以它的数据结构应该是

Class TrieNode{

Char c; //存储当前这个字符

TrieNode child[26]; //存储这个字符的子节点

}

6.Trie树的分析：

时间复杂度：非常高效 O(单词的长度)

空间复杂度：

以空间来换效率的数据结构。因为每个单词理论上都有26个子节点，所有它的空间复杂度就是26^n，n表示的是树的高度。

优化：

1.重复的字母不要重复建

2.因为我们每个node都开了26个空间来存储节点。但实际情况可能不需要这么多，所以这里其实我们可以考虑用散列表来实现，

这里大家可以去看IK的源码，当子节点少的时候是用的数组，但是节点大于3个是它是用的hashMap，这个再一定的程度上是可以节省很多的空间的。

