

Elasticsearch搜索的过程?

Elasticsearch的搜索过程:

1. **提交查询请求**: 客户端通过 HTTP RESTful API 向 Elasticsearch 集群发送一个搜索请求。请求通常采用 GET 或 POST 方法, URL 形如 GET /<index>/_search 或 POST /<index>/_doc/_search。
2. **接收查询请求**: 该请求首先被发送到集群中的任意一个节点。这个节点被称为**协调节点** (Coordinating Node), 它负责管理整个搜索过程。Elasticsearch**首先需要解析查询语句**。查询语句可以是结构化查询 (如SQL语句) 或全文查询 (如关键词查询), 通常包括查询的索引、查询条件 (如关键词、过滤器等)、排序方式、分页信息以及是否需要聚合等。
3. **路由查询**: 协调节点根据查询请求中的索引名称确定要查询的索引。如果是别名或通配符查询, 协调节点会解析出具体的索引。索引的数据通常分布在多个**分片** (Shard) 中, 每个分片可能在不同的节点上。Elasticsearch 使用哈希算法和文档 ID 确定每个文档的分片位置。
4. **分发查询**: 协调节点将查询请求分发到所有相关的主分片 (Primary Shard) 或副本分片 (Replica Shard)。为了分摊负载, 协调节点可以选择主分片或副本分片来执行查询。每个分片独立执行查询, 生成**局部结果**。这包括匹配的文档 ID、相关度分数 (用于排序) 以及任何指定的聚合结果。每个分片将其局部结果返回给协调节点。
5. **合并结果**: 协调节点汇总所有分片的得分信息, 进行全局排序, 并合并得到最终的搜索结果集以及聚合数据。
6. **相应结果**: 协调节点将整理好的搜索结果 (包含文档的详细内容、匹配数量、分页信息以及聚合数据等) 以JSON格式返回给客户端。用户可以根据这些结果来进一步操作, 例如查看文档内容、进行二次查询等。

Elasticsearch的搜索过程是基于其强大的全文搜索引擎功能实现的。



知识星球

长按扫码领取优惠 ▶



公众号 · 旧时光大数据