

# Elasticsearch Service

词汇表

产品文档



腾讯云

---

**【版权声明】**

©2013-2019 腾讯云版权所有

本文档著作权归腾讯云单独所有，未经腾讯云事先书面许可，任何主体不得以任何形式复制、修改、抄袭、传播全部或部分本文档内容。

**【商标声明】**

及其它腾讯云服务相关的商标均为腾讯云计算（北京）有限责任公司及其关联公司所有。本文档涉及的第三方主体的商标，依法由权利人所有。

**【服务声明】**

本文档意在向客户介绍腾讯云全部或部分产品、服务的当时的整体概况，部分产品、服务的内容可能有所调整。您所购买的腾讯云产品、服务的种类、服务标准等应由您与腾讯云之间的商业合同约定，除非双方另有约定，否则，腾讯云对本文档内容不做任何明示或模式的承诺或保证。

# 词汇表

最近更新时间：2019-11-08 17:42:50

## 分片

分片 (Shards) 是数据的容器，文档保存在分片内，一个分片是一个底层的工作单元，它仅保存了全部数据中的一部分，分片又被分配到集群内的各个节点里，当您的集群规模扩大或者缩小时，Elasticsearch 会自动的在各节点中迁移分片，使得数据仍然均匀分布在集群里。

一个分片可以是主分片或者副本分片。索引内任意一个文档都归属于一个主分片，所以主分片的数目决定着索引能够保存的最大数据量（技术上来说，一个主分片最大能够存储 `Integer.MAX_VALUE - 128` 个文档）。

一个副本分片 (Replicas) 只是一个主分片的拷贝。副本分片作为硬件故障时保护数据不丢失的冗余备份，并为搜索和返回文档等读操作提供服务。在索引建立的时候就已经确定了主分片数，但是副本分片数可以随时修改。

## 集群和节点

一个运行中的 Elasticsearch 实例称为一个节点 (node)，同一网络内一个或者多个拥有相同集群名 (cluster.name) 配置并网络互通的节点组成 Elasticsearch 集群 (cluster)。集群内的节点共同承担数据的存储和查询请求，当有节点加入集群中或者从集群中移除时，集群将会重新平均分布所有的数据。每个节点都知道任意文档所处的位置，无论用户将请求发送到哪个节点，都能够将请求直接转发到存储所需文档的节点，并从各个包含所需文档的节点收集回数据，将最终结果返回给客户端。

## 索引

### 常用概念

一个索引 (名词) 类似于传统关系数据库中的一个数据库，是存储关系型文档的地方。索引的复数词为 `indices` 或 `indexes`。

索引 (动词) 一个文档就是存储一个文档到一个索引 (名词) 中以便它可以被检索和查询到。除了文档已存在时新文档会替换旧文档情况之外，这非常类似于 SQL 语句中的 `INSERT` 关键词。

### 倒排索引

关系型数据库通过增加一个索引，例如一个 B 树 (B-tree) 索引到指定的列上，来提升数据检索速度。Elasticsearch 和 Lucene 使用了一个倒排索引的结构来达到相同的目的。默认的，一个文档中的每一个属性都是被索引 (有一个倒排索引) 和可搜索的。一个没有倒排索引的属性是不能被搜索到的。

## 文档

Elasticsearch 是面向文档 (document) 的，它存储的是整个对象或文档，并且索引每个文档的内容使之可以被检索。Elasticsearch 使用 JSON 作为文档的序列化格式，简单、简洁、易于阅读。JSON 序列化被大多数编程语言所支持，并且已经成为 NoSQL 领域的标准格式。在 Elasticsearch 中，用户不是对行列数据进行检索，而是对文档进

---

行索引、检索、排序和过滤，这是一种完全不同的思考数据的方式，也是 Elasticsearch 能支持复杂全文检索的原因。