Elasticsearch是一款分布式搜索引擎，像github，阿里，新浪等，可用于大数据开发（可处理PB级别（1024T）数据）

什么是分布式搜索引擎？

搜索引擎：

有个数据源（数据采集，数据库，可使用python等）进行处理，存放在搜索节点中，为用户提供检索服务，存在的目的是提供有效的搜索服务，对海量数据进行分析等

分布式存储与搜索：

由多个节点共同组成的搜索服务，提高可扩展性，横向扩展集群，提高存储容量

Lucene和solr和elasticsearch的对比

都是到倒排索引（源于实际应用中需要根据属性的值来查找记录。这种索引表中的每一项都包括一个属性值和包含该属性值的各个记录地址。由于不是根据记录来确定属性，而是根据属性来确定记录的位置，所以称之为倒排索引）

Lucene是类库，进行java编写的API库，jar包，只能使用java开发

Solr基于Lecene

ES基于Lecene

ES核心术语：

ES是面向文档进行检索的

索引index 相当于 表

类型type（在7.X版本后不使用） 相当于 表逻辑类型

文档document 相当于 行

字段fields 相当于 列

（索引）User[

(文档){

(字段)Id: 1,

Username: ‘111’

}，

{

Id: 2,

Username: ‘222’

}]

映射mapping 相当于表结构定义

近实时NRT

节点node

Shard（数据分片） replica 备份

Primary Shard 主备份

Replica shard（从备份）

ES的环境搭建

将文件上传至/usr/local

解压 elasticsearch-7.13.2-linux-x86\_64.tar.gz

kibana-7.13.2-linux-x86\_64

logstash-7.13.2-linux-aarch64.tar.gz

在elasticsearch目录下创建data目录

mkdir data

因为安全问题，elastic不允许root用户直接运行，所以要创建新用户

1.在root中创建新用户

useradd esuser

passwd esuser

chown -R esuser /usr/local/elasticsearch-7.4.2/

修改配置文件

/elasticsearch-7.13.2/config/elasticsearch.yml

/elasticsearch-7.13.2/config/jvm.options

修改用户可操控的线程数

vim /etc/security/limits.conf

最下面追加

esuser soft nofile 65536

esuser hard nofile 131072

esuser soft nproc 4096

esuser hard nproc 4096

修改/etc/sysctl.conf

追加 vm.max\_map\_count=262145

刷新 sysctl -p

后台启动命令：

./elasticsearch -d

快速关闭：

jps

杀掉对应进程