## 环境搭建(linux)

1. 下载并安装**docker**服务
2. 使用docker安装es 7x，执行命令： docker pull elasticsearch:7.9.3
3. 启动es，执行以下命令

docker run --name elasticsearch -p 9200:9200 -p 9300:9300 -e "discovery.type=single-node" elasticsearch:7.9.3

1. 进入ES容器，docker exec -it "这里为es的docker id" bash
2. 为ES安装lk插件，执行以下命令

cd bin && ./elasticsearch-plugin install <https://github.com/medcl/elasticsearch-analysis-ik/releases/download/v7.9.3/elasticsearch-analysis-ik-7.9.3.zip>

1. 重启 elasticsearch:7.9.3 docker restart "你的elasticsearch 容器ID"
2. 查看es服务是否正常 curl -s "http://localhost:9200/" 看是否返回，es信息

{

"name" : "c7df55e22430",

"cluster\_name" : "docker-cluster",

"cluster\_uuid" : "do\_FH5vTTOm\_uiEcfFMgZw",

"version" : {

"number" : "7.9.3",

"build\_flavor" : "default",

"build\_type" : "docker",

"build\_hash" : "c4138e51121ef06a6404866cddc601906fe5c868",

"build\_date" : "2020-10-16T10:36:16.141335Z",

"build\_snapshot" : false,

"lucene\_version" : "8.6.2",

"minimum\_wire\_compatibility\_version" : "6.8.0",

"minimum\_index\_compatibility\_version" : "6.0.0-beta1"

},

"tagline" : "You Know, for Search"

}

8. 如果ES正常，就可以接着往下部署

## 部署爬虫，搜索服务器

1. 部署爬虫，crawler，默认配置文件为config.yaml 和应用一层目录
2. 启动 ./crawler , 如果看到 Save Result Count::xxx 说明抓取到的内容已经保存了elasticsearch
3. config.yaml可以配置网站和抓取的一些参数，里面有详细注解，也可以默认参数，如需更改，最好不要超过或者低于建议的数值
4. 部署search server ./search\_server 默认监听在 :9090端口，启动后，

curl -s "http://localhost:9090/" 看看是否有相应

1. 数据量足够后，浏览器打开 <http://ip:9090/> 就可以搜索试试想要的关键字

## 查询接口接入

接入: <http://ip:9090/search/关键字?pn=1> pn为页码，页面大小默认为10，暂时不支持页大小更改。返回数据为json数组，数据元数据结构如下：

{

"id": 0, // 网页ID

"domain": "", // 域名

"title": "", // 网页标题

"description": "", // 网页描述

"key\_words": "", // 网页关键字

"url": "", // 网页链接

"clicks": 0, // 点击数

"create\_at": 0, // 创建时间

"weight": 0, // 权重

"expiration": 0, // 过期时间

"ico": "" // 网页小图标

}

## 常见问题

1. centos错误 Cannot connect to the Docker daemon at unix:///var/run/docker.sock. Is the docker daemon

解决： systemctl daemon-reload && service docker restart

1. 如果是centos端口外网服务无法访问，可以暂时关闭centos的防火墙

解决： systemctl stop firewalld

1. driver failed programming external connectivity on endpoint elasticsearch

解决：

service docker stop

rm /var/lib/docker/network/files/local-kv.db

service docker start