开源实时日志分析ELK平台部署第一节课

主讲：tree老师

QQ：3059942361

ELK是指logstash,elasticsearch,kibana三件套，这三件套可以组成日志分析和监控工具

为什么学ELK?

本节课所包含单词：

logstash美 **[stæʃ]**

elastic美 [ɪ'læstɪk]

search 美 **[sɜːrtʃ]**

# Kibana

日志主要包括系统日志、应用程序日志和安全日志。系统运维和开发人员可以通过日志了解服务器软硬件信息、检查配置过程中的错误及错误发生的原因。经常分析日志可以了解服务器的负荷，性能安全性，从而及时采取措施纠正错误。

通常，日志被分散的储存不同的设备上。如果你管理数十上百台服务器，你还在使用依次登录每台机器的传统方法查阅日志。这样是不是感觉很繁琐和效率低下。当务之急我们使用集中化的日志管理，例如：开源的syslog，将所有服务器上的日志收集汇总。

集中化管理日志后，日志的统计和检索又成为一件比较麻烦的事情，一般我们使用grep、awk和wc等Linux命令能实现检索和统计，但是对于要求更高的查询、排序和统计等要求和庞大的机器数量再去使用这些方法就比较的麻烦了。

开源实时日志分析ELK平台能够完美的解决我们上述的问题，ELK由ElasticSearch、Logstash和Kiabana三个开源工具组成。

官方网站：<https://www.elastic.co/products>

  Elasticsearch是个开源分布式搜索引擎，它的特点有：分布式，零配置，自动发现，索引自动分片，索引副本机制，restful风格接口，多数据源，自动搜索负载等。

Elasticsearch概念及特点：

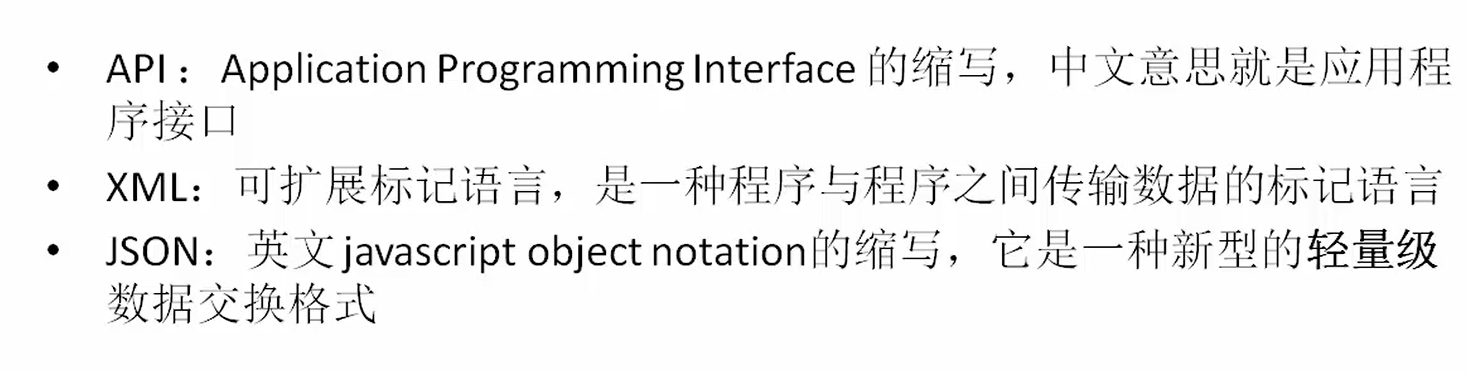
看下百度百科 ，开发人员看到LUCENE 就估计就明白了一半，运维人员就不怎么了解。这个是一个框架，这个框架的作者就是hadoop的作者。

elastic 可伸缩的。

一个采用RESTFUL api 标准的高扩展性和高可用性的实时数据分析的全文搜索工具。

RUStful的解释

做运维的可能不太了解rustful，但是做开发可能会了解的多一些，从这里就可以看出来，学一门语言是不是重要的。



有了API就 可以更好的调用和使用。XML为什么会被丢弃，因为XML文件格式比较庞大，而且不容易解析，，不同浏览器之间解析XML又不一样，而且现在微博，等web端的应用，XML越来越跟不上时代的发展。所以现在开始用轻量级的JSON

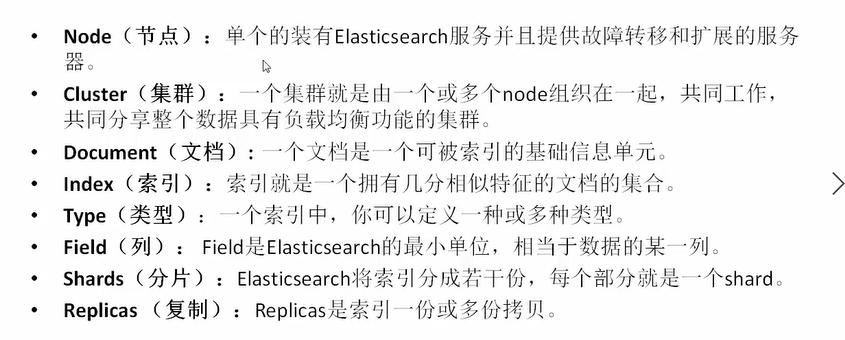


1.表现层：资源，比如说一段文本，一部电影，每个资源都有一个ID来表示，这个ID我们成为URL。每个资源仅代表一个信息，一个信息又可以用很多种形式来表示，比如文字，有text ，word格式

2.状态转化：客户端想要访问服务器，必须要通过某些手段与服务器进行交互，交互的过程又是基于表现层的。

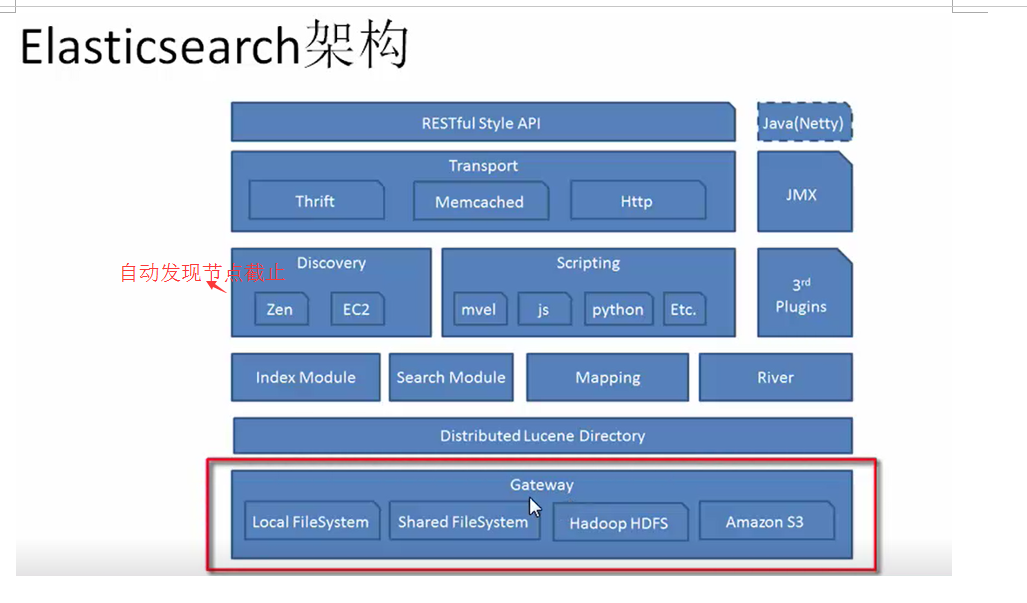
这里的某些手段指的是，GET用来获取资源， post 用来新建资源 PUT 用来更新资源。 delete删除资源。

高扩展性体现在拓展非常简单，新的节点，基本无需做复杂配置自动发现节点。高可用性呢，因为这个社区是分布式的，而且是一个实时搜索平台，支持PB级的大数据搜索能力。从一个文档到这个文档被搜索到呢，只有略微的延迟，所以说他的实时性非常高。



你可以把elasticearch当做一个数据库，来对比下与其他数据库的不同，elasticseach是非关系型数据库。





gateway：elasticsearch支持的索引数据的存储格式，当这个elasticsearch关闭再启动的时候，他就会从gateway里面读取索引数据。本地的，分布式，亚马逊的S3

第二行是lucene的框架

第三行：对数据的加工处理方式，创建index模块，搜索模块，mapping,river

第四行：

左边第一块是elasticsearch自动发现节点的机制。

zen 用来实现节点自动发现的，加入master发生故障，其他的节点就会自动选举，来确定一个master。

中间，方便对查询出来的数据进行数据处理。

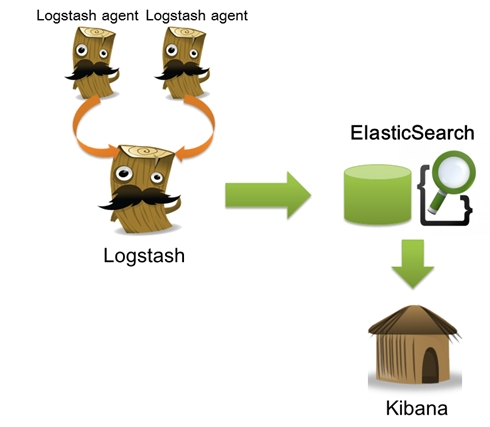
最右边：插件识别，中文分词插件，断点监控插件

倒数第二层，是交互方式，默认是用httpd协议来传输的。

最顶层右边是说可以使用java语言，本身就是用java开发的。

  Logstash是一个完全开源的工具，他可以对你的日志进行收集、分析，并将其存储供以后使用（如，搜索）。

  kibana 也是一个开源和免费的工具，他Kibana可以为 Logstash 和 ElasticSearch 提供的日志分析友好的 Web 界面，可以帮助您汇总、分析和搜索重要数据日志。



 在需要收集日志的所有服务上部署logstash，作为logstash agent（logstash shipper）用于监控并过滤收集日志，将过滤后的内容发送到logstash indexer，logstash indexer将日志收集在一起交给全文搜索服务ElasticSearch，可以用ElasticSearch进行自定义搜索通过Kibana 来结合自定义搜索进行页面展示。

ELK有两种安装方式

(1)集成环境：Logstash有一个集成包，里面包括了其全套的三个组件；也就是安装一个集成包。

(2)独立环境：三个组件分别单独安装、运行、各司其职。(比较常用)

本实验也以第二种方式独立环境来进行演示；单机版主机地址为：192.168.1.63

安装JDK

检查本机是否有JDK

[root@xuegod63 ~]# java -version

java version "1.7.0\_45"

OpenJDK Runtime Environment (rhel-2.4.3.3.el6-x86\_64 u45-b15)

OpenJDK 64-Bit Server VM (build 24.45-b08, mixed mode)

卸载JDK

[root@xuegod63 ~]# yum remove java

Loaded plugins: product-id, refresh-packagekit, security,

: subscription-manager

This system is not registered to Red Hat Subscription Management. You can use subscription-manager to register.

Setting up Remove Process

Resolving Dependencies

--> Running transaction check

---> Package java-1.6.0-openjdk.x86\_64 1:1.6.0.0-1.66.1.13.0.el6 will be erased

---> Package java-1.7.0-openjdk.x86\_64 1:1.7.0.45-2.4.3.3.el6 will be erased

--> Finished Dependency Resolution

Removed:

java-1.6.0-openjdk.x86\_64 1:1.6.0.0-1.66.1.13.0.el6

java-1.7.0-openjdk.x86\_64 1:1.7.0.45-2.4.3.3.el6

Complete!

[root@xuegod63 ~]#

[root@xuegod63 ~]# java -version

bash: /usr/bin/java: 没有那个文件或目录

注意：如果还有的话可以使用yum grouprmove java

JDK下载地址<http://www.oracle.com/technetwork/java/javase/downloads/jdk8-downloads-2133151.html>

Elasticsearch下载地址

<https://www.elastic.co/downloads/elasticsearch>

rz命令上传

安装JDK

[root@xuegod63 ~]# tar zxf jdk-8u111-linux-x64.tar.gz

[root@xuegod63 ~]# ls

anaconda-ks.cfg

elasticsearch-5.0.0.tar.gz vmware-tools-distrib

jdk1.8.0\_111

jdk-8u111-linux-x64.tar.gz

[root@xuegod63 ~]# cd jdk1.8.0\_111/

[root@xuegod63 jdk1.8.0\_111]# ll

总用量 10096

drwxr-xr-x. 2 uucp 143 4096 9月 23 07:24 bin

-r--r--r--. 1 uucp 143 3244 9月 23 07:23 COPYRIGHT

drwxr-xr-x. 4 uucp 143 4096 9月 23 07:23 db

-rwxr-xr-x. 1 uucp 143 5091318 9月 23 03:20 javafx-src.zip

drwxr-xr-x. 5 uucp 143 4096 9月 23 07:24 jre

-r--r--r--. 1 uucp 143 159 9月 23 07:23 README.html

-rw-r--r--. 1 uucp 143 526 9月 23 07:23 release

[root@xuegod63 jdk1.8.0\_111]#

[root@xuegod63 jdk1.8.0\_111]# vim /etc/profile

export JAVA\_HOME=/root/jdk1.8.0\_111

export JAVA\_BIN=/root/jdk1.8.0\_111/bin

export PATH=$PATH:$JAVA\_HOME/bin

export CLASSPATH=.:$JAVA\_HOME/lib/dt.jar:$JAVA\_HOME/lib/tools.jar

export JAVA\_HOME JAVA\_BIN PATH CLASSPATH

重新加载JDK

reboot或者source

[root@xuegod63 jdk1.8.0\_111]# source /etc/profile

[root@xuegod63 jdk1.8.0\_111]# java -version

java version "1.8.0\_111"

Java(TM) SE Runtime Environment (build 1.8.0\_111-b14)

Java HotSpot(TM) 64-Bit Server VM (build 25.111-b14, mixed mode)

[root@xuegod63 jdk1.8.0\_111]#

安装elasticserache

解压即可运行

[root@xuegod63 ~]# tar zxf elasticsearch-5.0.0.tar.gz

[root@xuegod63 ~]# cd elasticsearch-5.0.0

[root@xuegod63 elasticsearch-5.0.0]# ls

bin config lib LICENSE.txt modules NOTICE.txt README.textile

[root@xuegod63 elasticsearch-5.0.0]#

[root@xuegod63 elasticsearch-5.0.0]# ./bin/elasticsearch -d

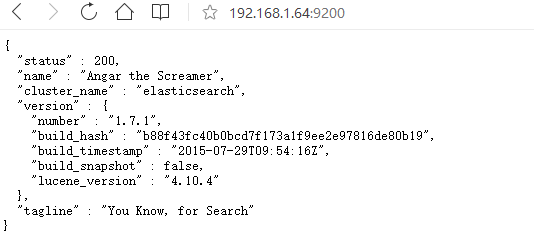
-d 表示让服务在后台运行

怎么确认elasticsearch 是否启动成功呢？

我们来访问下web服务地址就可以确认了

web服务地址

<http://localhost:9200>



这是一个json的格式，把数据返回回来了

200说明成功了，

name :节点的名称，这个名称是美国漫威漫画角色的名字，他会随机分配一个

只要java配对了，基本上都可以启动成功，如果启动不成功可以去log日志里面查看

接下来我们看下目录结构



[root@xuegod64 bin]# ls

elasticsearch elasticsearch.in.bat elasticsearch-service-mgr.exe elasticsearch-service-x86.exe plugin.bat

elasticsearch.bat elasticsearch.in.sh（做优化的时候可以用到） elasticsearch-service-x64.exe plugin service.bat

cd config

yml YAML的形式

最重要的是data文件

以上是安装

接下来我们看看elasticsearch相关的插件（插件很多，可以在官网搜）

head插件

bigdesk插件

Mavel插件 （收费的）但是用的还比较多。

head插件是一个集群管理工具，是由html5写的

GIThub 网址获取相关信息：

<https://github.com/mobz/elasticsearch-head>

安装步骤

1.进入elasticsearch bin目录下运行./plugin -install mobz/elasticsearch-head 命令

[root@xuegod64 bin]# pwd

/root/elasticsearch-1.7.1/bin

这里面的plugin就是用来装插件用的

运行命令

[root@xuegod64 bin]# ./plugin -install mobz/elasticsearch-head

-> Installing mobz/elasticsearch-head...

Trying https://github.com/mobz/elasticsearch-head/archive/master.zip...

Downloading .........................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................DONE

Installed mobz/elasticsearch-head into /root/elasticsearch-1.7.1/plugins/head

2.使用方法

登陆网址测试：http://192.168.1.64:9200/\_plugin/head/

* 

以后还会用到这个插件，今天我们简单了解下

2.Bigdesk插件：



[root@xuegod64 bin]# ./plugin -install lukas-vlcek/bigdesk/2.5.0

-> Installing lukas-vlcek/bigdesk/2.5.0...

Trying http://download.elasticsearch.org/lukas-vlcek/bigdesk/bigdesk-2.5.0.zip...

Trying http://search.maven.org/remotecontent?filepath=lukas-vlcek/bigdesk/2.5.0/bigdesk-2.5.0.zip...

Trying https://oss.sonatype.org/service/local/repositories/releases/content/lukas-vlcek/bigdesk/2.5.0/bigdesk-2.5.0.zip...

Trying https://github.com/lukas-vlcek/bigdesk/archive/2.5.0.zip...

Trying https://github.com/lukas-vlcek/bigdesk/archive/master.zip...

Downloading ........................................................................................................................................................................................................................................................DONE

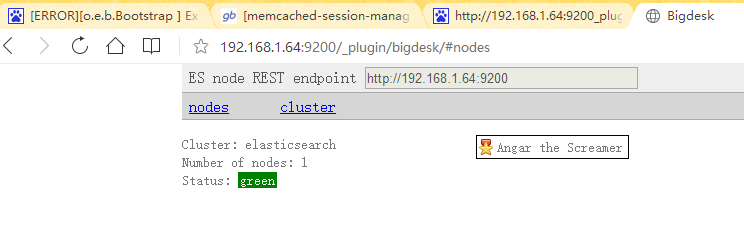
Installed lukas-vlcek/bigdesk/2.5.0 into /root/elasticsearch-1.7.1/plugins/bigdesk

Identified as a \_site plugin, moving to \_site structure ...

[root@xuegod64 bin]#

通过网页访问验证

<http://192.168.1.63:9200/_plugin/bigdesk/#nodes>



点击nodes 会有很多图标，都是html5的

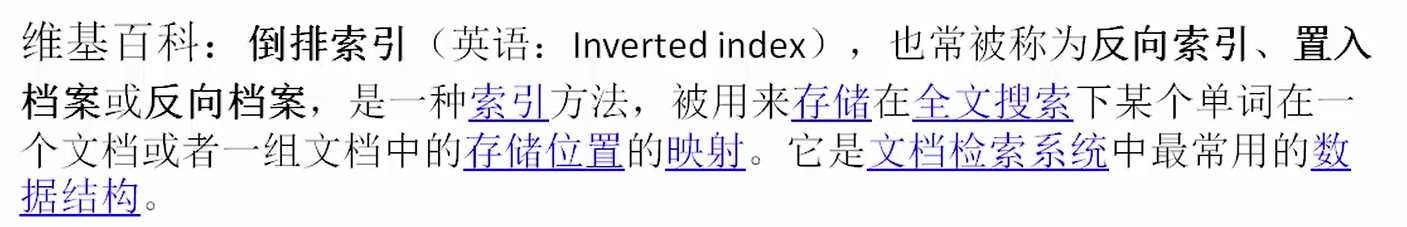
系统监控，cpu使用率，内存使用率，http的交互情况 ，io吞吐量

cluster 目前还是空的，等以后我们会用到的时候，来看看界面是啥样子。

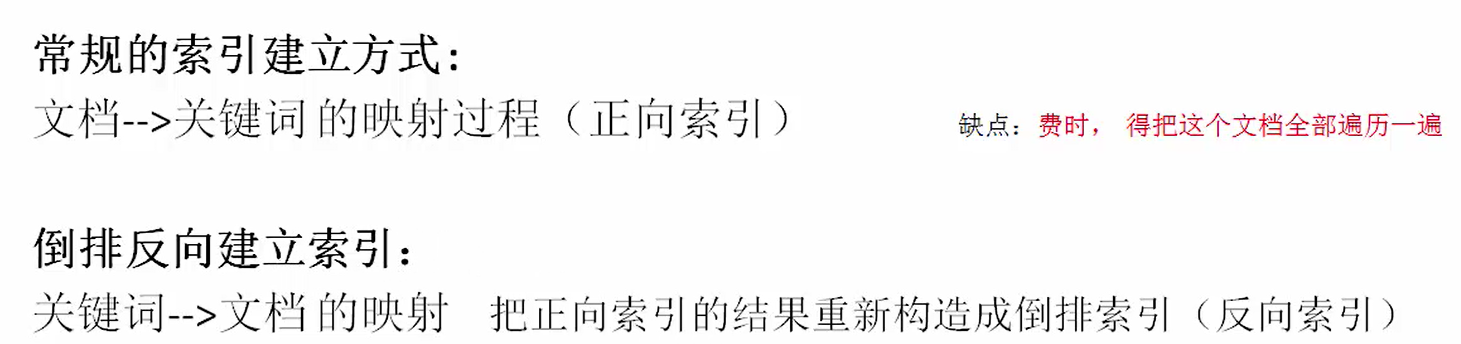
4.倒排索引

什么是倒排索引

维基百科：



正向索引是每个单词都要读取一遍，全部文档要遍历一遍





比如搜索乔布斯，那就只到文档一文档2，文档3里面查找

TF表示出现的频率，文档1，在3和11的位置出现过两次

MAVEL插件（收费）



[root@xuegod64 bin]# ./plugin -install elasticsearch/marvel/latest

-> Installing elasticsearch/marvel/latest...

Trying http://download.elasticsearch.org/elasticsearch/marvel/marvel-latest.zip...

Downloading ........................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................DONE

Installed elasticsearch/marvel/latest into /root/elasticsearch-1.7.1/plugins/marvel

[root@xuegod64 bin]#

测试：

<http://192.168.1.64:9200/_plugin/marvel/>