**分布式ELK（分析日志）平台**

基础知识

ELK是什么？

sina、饿了么、携程、华为、美团、这些公司都在使用elk。ELK不是一个软件，而是一套解决方案，是三个软件产品的首字母缩写

Elasticsearch：负责日志检索和储存

logstash：负责日志的收集和分析、处理

kibana：负责日志的可视化

这三款软件都是开源软件，通常是配合使用，而且又先后归于Elastic.co公司名下，故被简称为ELK

ELK能做什么

ELK组件在海量日志之系统的运维中，可用于解决：

分布式日志数据集中式查询和管理

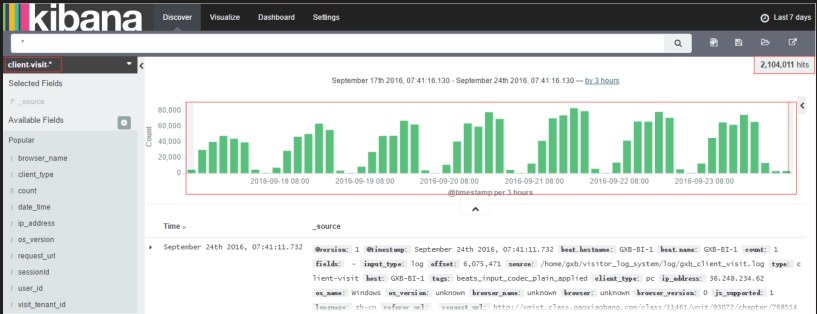
系统监控，包含系统硬件和应用各个组件的监控

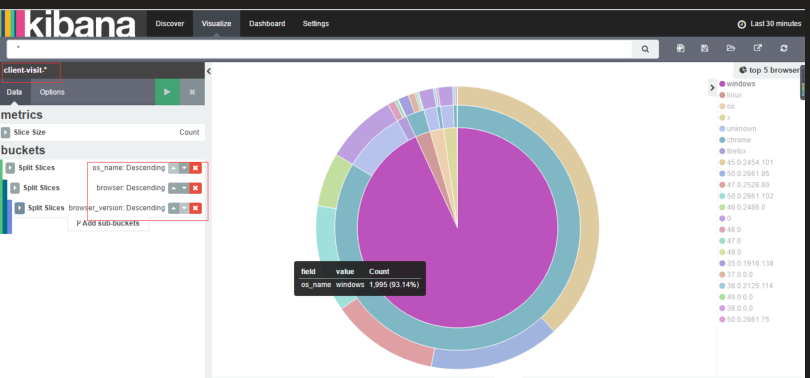
故障排查

安全信息和事件管理

报表功能

ELK是什么样子的







elasticsearch（数据库）

elasticsearch部分

elasticsearch是一个基于lucene的搜索服务器，它提供了一个分布式多用户能力的全文搜索引擎，基于RESTful API的web接口

elasticsearch是用java开发的，并作为apache许可条件下的开放源码发布，是当前流行的企业级搜索引擎。设计用于云计算中，能够达到实时搜索，稳定，可靠，快速，安装使用方便

主要特点：实时分析

分布式实时文件存储，并将每一个字段都编入索引

文档导向，所有的对象全部是文档

高可用性，易扩展，支持集群、分片和复制

接口友好，支持JSON

ES没有什么

Elasticsearch没有典型意义的事务

Elastivsearch是一种面向文档的数据库

Elastivsearsh没有提供授权和认证特性

相关概念

node：装有一个ES服务器的节点

cluster：有多个node组成的集群

document：一个可被搜索的基础信息单元，相当于数据库中的行

index：拥有相似特征的文档的集合，相当于数据库中的库

type：一个索引中可以定义一种或多种类型，相当于数据库中的表

filed：一个索引中可以定义一种或多种类型

shards：索引的分片，每个分片就是一个shard

replicas：索引的拷贝

sql与nosql

ES与关系型数据库的对比

在ES中，文档归属于一种类型（type），而这些类型存在于索引中，类比传统关系型数据库

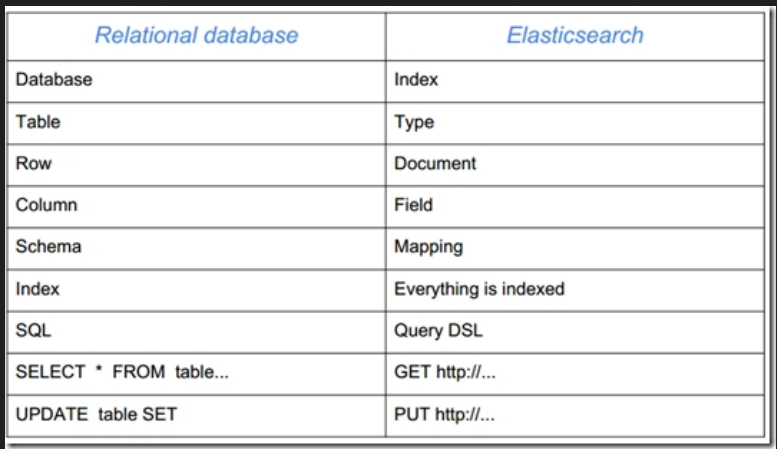
DB->Databases->tables->rows->columns

关系型 数据库 表 行 列

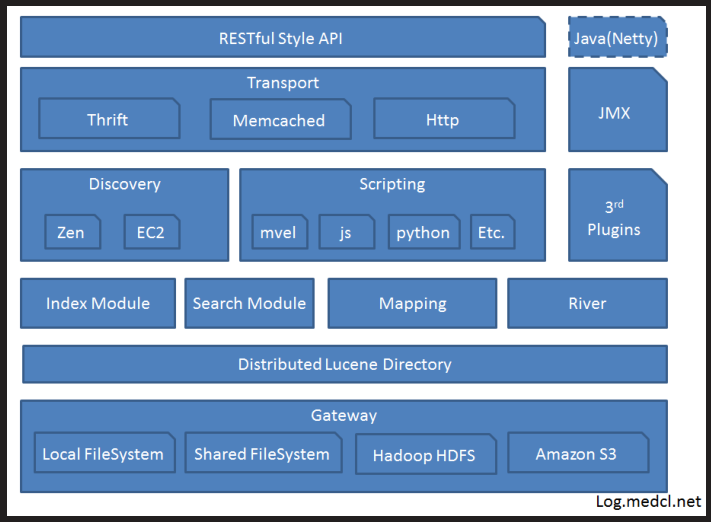
ES->indices->types->documents->fields

ES 索引 类型 文档 域（字段）

ES与关系型数据库的对比



Elasticsearch架构图



ES集群搭建

ES集群安装

安装第一台ES服务器

设置主机名称和ip对应关系

解决依赖关系

安装软件包

修改配置文件

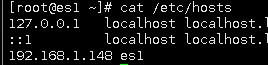
启动服务

检查服务

步骤1，设置ip与主机名称对应关系

配置/etc/hosts

192.168.4.11 node1



步骤2，安装JDK

Elasticsearch要求至少java7

一般推荐使用openjdk1.8

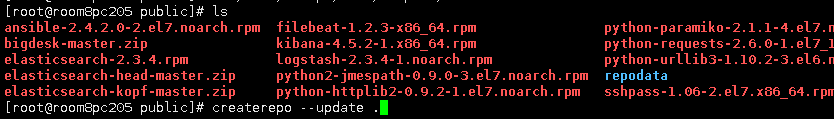
配置好安装源以后，先解决依赖关系

yum -y install java-1.8.0-openjdk



步骤3，安装ES

rpm -ivh elasticsearch-2.3.4.-1.noarch





步骤4，修改配置文件

vim elasticsearch.yml

network.host:0.0.0.0

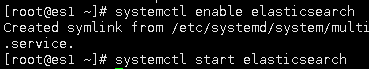




步骤5，启动服务，设置自启动

systemctl enable elasticsearch

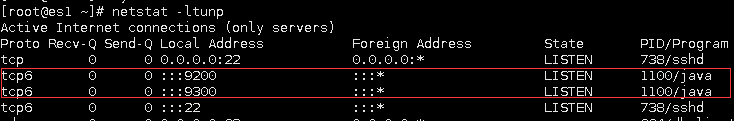
systemctl start elasticsearch



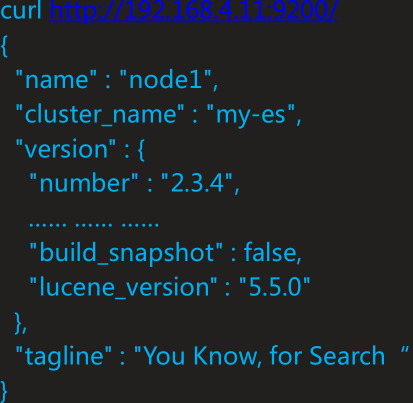
验证：

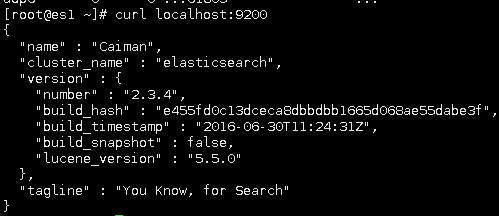
netstat -ltunp

能够看到9200 9300被监听



通过浏览器或curl访问9200端口



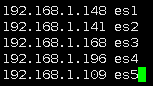


ES集群配置

ES集群配置也很简单，只需要对配置文件做少量的修改即可，其他步骤和单机完全一致

安装5台数据库节点

集群中的所有节点要相互能够ping通，要在所有集群机器上配置/etc/hosts中的主机名与ip对应关系

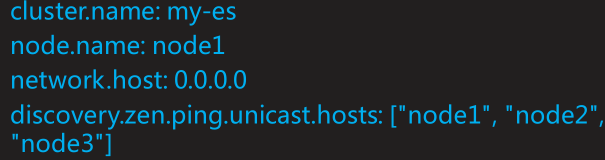


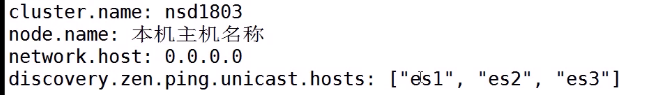


集群中所有机器都要安装java环境



ES集群配置文件





cluster.name集群名称配置要求完全一致

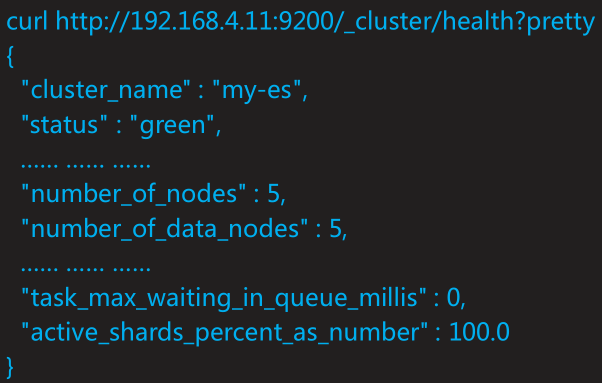
node.name为当前节点标识，应配置本机的主机名、

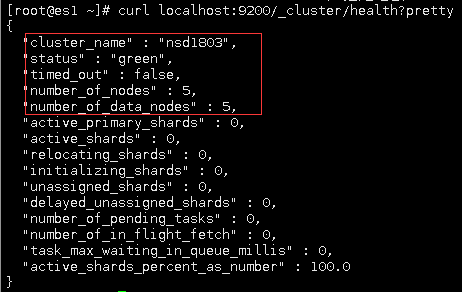
discovery为集群节点信息，不需要全部配置，启动时，会查找这些主机，定义的先启动，否则查找不到，会自定义为一个集群

配置完成以后启动所有节点服务（有可能会有一定的延迟，需要等待几十秒）



验证集群，使用ES内置字段\_cluster/health?pretty





ES集群验证

返回字段解析

status”:”green”集群状态，绿色为正常，黄色表示有问题但不是很严重，红色表示严重故障

“number\_of\_nodes”:5,表示集群中节点的数量

"number\_of\_data\_nodes" : 5,

…… …… ……

"task\_max\_waiting\_in\_queue\_millis" : 0,

"active\_shards\_percent\_as\_number" : 100.0

HTTP协议简介

分别是：请求行、消息包头、请求正文

请求行以一个方法符号开头，以空格分开，后面跟着请求的URI和协议版本，格式如下：

Method Request-URI HTTP-Version CRLF

http请求方法

常用方法GET,POST,HEAD

其他方法OPTIONS,PUT,DELETE,TRACE和CONNECT

ES常用

PUT ---増

DELETE ---删

POST ---改

GET ---查

系统命令curl

在Linux中curl是一个利用URL规则在命令行下工作的文件传输工具，可以说是一款很强大的http命令行工具。它支持多种请求模式，自定义请求头等强大功能，是一款综合工具

curl常用参数介绍

-A 修改请求agent

-X 设置请求方法

-i显示返回头信息

ES插件的安装与使用

ES插件的使用

ES常用插件

head插件：它展现ES集群的拓扑结构，并且可以通过它来进行索引和节点级别的操作。它提供一组针对集群的查询API，并将结果以json和表格形式返回。它提供一些快捷菜单，用以展现集群的各种状态

kopf插件：是一个elasticsearch的管理工具。它提供了对ES集群操作的API

bigdesk插件：是elasticsearch的一个集群监控工具，可以通过它来查看es集群的各种状态，如：CPU、内存、使用情况、索引数据、搜索情况，http连接数等

ES插件安装、查看

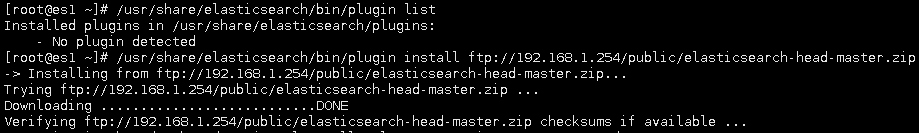
查看安装的插件

/usr/share/elasticsearch/bin/plugin list

安装插件

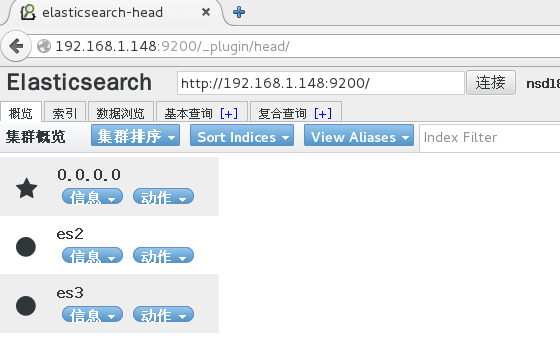
/usr/share/elasticsearch/bin/plugin install

<ftp://192.168.4.254/head.zip>



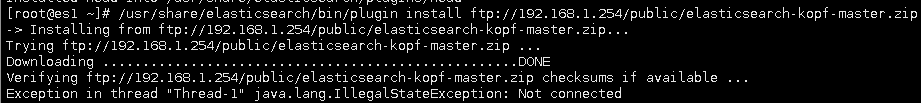
查看

192.168.1.148:9200/\_plugin/head



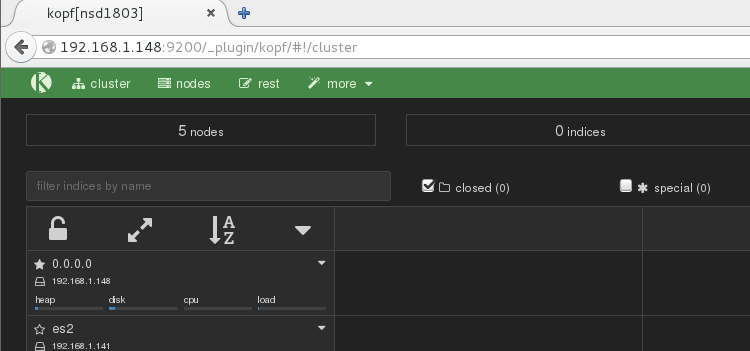
/usr/share/elasticsearch/bin/plugin install

[file:///tmp/kopf.zip](file://tmp\\kopf.zip)



查看

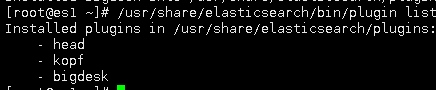
192.168.1.148:9200/\_plugin/kopf



/usr/share/elasticsearch/bin/plugin install

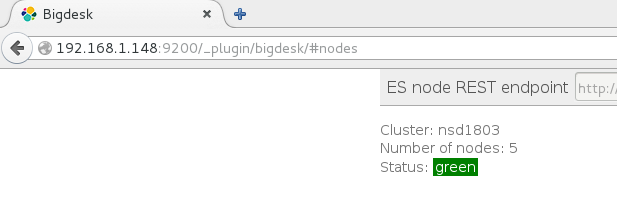
[file:///tmp/bigdesk.zip](file://tmp\\kopf.zip)





查看

192.168.1.148:9200/\_plugin/bigdesk



这里必须以url的方式进行安装，如果文件在本地，我们也需要使用file://的方式指定路径，例如文件在/tmp/xxx下面，要写成file:///tmp/xxx删除使用remove指令

HTTP与RESTful API

RESTful API调用

elsticsearch提供了一系列RESTful的API

检查集群、节点、索引的健康度、状态和统计

管理集群、节点、索引的数据及元数据

对索引进程CRUD操作及查询操作

执行其他高级操作如分页、排序、过滤等

POST或PUT数据使用json格式

json

json的全称是”javascript object notation”,意思是javascript对象表示法，它是一种基于文本，独立于语言的轻量级数据交换格式

json传输的就是一个字符串

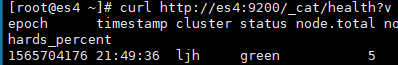
python中对应的字符串，列表，字典都可以转换成对应的json格式

Rest API的简单使用

\_cat API查询集群状态，节点信息

v参数显示详细信息

<http://192.168.4.15:9200/_cat/health?v>



help显示帮助信息

<http://192.168.4.15:9200/_cat/health?help>

nodes查询节点状态信息

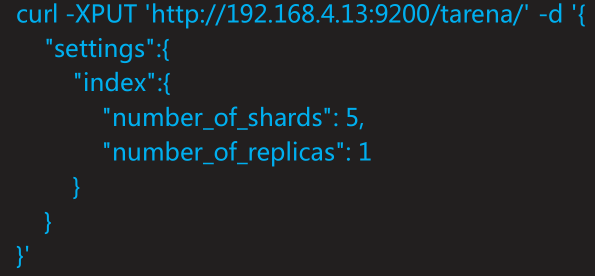
<http://192.168.4.15:9200/_cat/nodes?v>

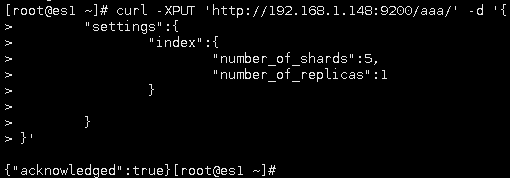
索引信息

<http://192.168.4.15:9200/_cat/indices?v>

Rest API增加

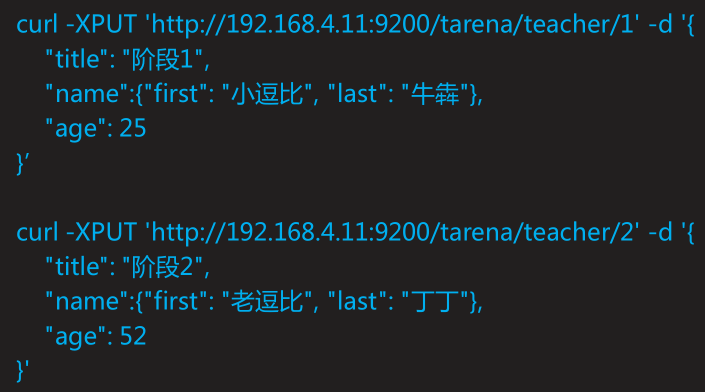
创建一个索引，并设置分片数量与副本数量

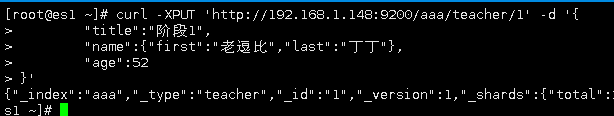






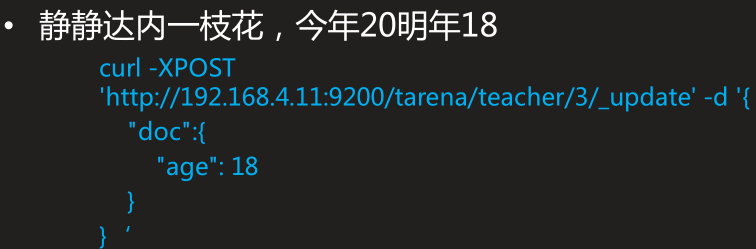
Rest API插入数据

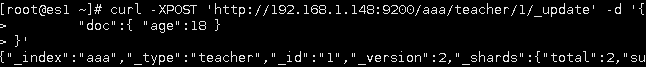






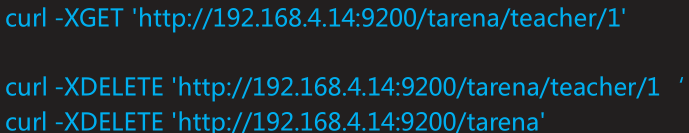
更改数据





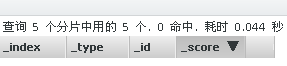


查询与删除









kibana安装

kibana是什么

数据可视化平台

特点：

灵活的分析和可视化平台

实时总结和流数据的图标

为不同的用户显示直观的界面

即时分享和嵌入的仪表盘

kibana安装（使用rpm方式安装）

rpm –ivh kibana-4.5.2-1.x86\_64.rpm

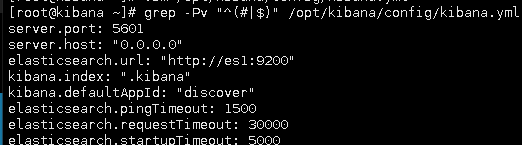




kibana默认安装在/opt/kibana下面，配置文件在/opt/kibana/config/kibana.yml，只需做少量的修改即可启动







kibana.yml的配置

除elasticsearch.yml需要配置为ES集群的地址之外，其他保持默认值就可以了

设置开机启动：systemctl enable kibana

启动服务：systemctl start kibana



web访问kibana：<http://192.168.4.20:5601/>

