Taller de Big Data.

Dashboard con TWEETS e integración de WS de la EMT en tiempo real con ELK: Elasticsearch, Logstash y Kibana

Alberto Jose Gallardo

MSMK

09 de Julio de 2018, Madrid

Índice general

[Capítulo 1 2](#_Toc519083218)

[1. Creación Twitter API 2](#_Toc519083219)

[1.1. Twitter Streaming API 2](#_Toc519083220)

[Capítulo 2 2](#_Toc519083221)

[2. Configuración de ELK 2](#_Toc519083222)

[2.1. Actualización de Java 2](#_Toc519083223)

[2.2. Eliminar Java 2](#_Toc519083224)

[2.3. ElasticSearch 3](#_Toc519083225)

[2.4. Configuración de Kibana 3](#_Toc519083226)

[2.5. Logstash 4](#_Toc519083227)

[2.6. Instalación X-pack 5](#_Toc519083228)

[2.7. Configuración de SUPERUSUARIO en Kibana 5](#_Toc519083229)

[2.8. Configuración del Template ElasticSearch 5](#_Toc519083230)

[2.9. Ejecución 6](#_Toc519083231)

[Capítulo 3 7](#_Toc519083232)

[3. Cuadro de mandos en tiempo real 7](#_Toc519083233)

# Capítulo 1

## Creación Twitter API

## Twitter Streaming API

Para ello se creó una cuenta:

<https://apps.twitter.com/>

Master\_Big\_Data\_ELK\_Twitter

<https://apps.twitter.com/app/15468627/show>

Necesitaremos las claves secretas más adelante para la configuración.

Consumer Key (API Key)

zxXijnGtO8RboC2ECOujpHR3X

Consumer Secret (API Secret)

2IJGLULqCofeDqKvGpeFdynB9kJ4qVRYv0Awc7uYCOsEOIhE2I

# Capítulo 2

## Configuración de ELK

## Actualización de Java

En primer lugar, tenemos que actualizar la versión de Java de la máquina virtual, ya que ElasticSearch funciona con Java8.

## Eliminar Java

Primero eliminaremos la versión instalada, para después instalar la que necesitamos.

sudo yum remove java

sudo yum install java-1.8.0-openjdk

Para que las aplicaciones encuentren la versi n de Java correctamente debemos ejecutar lo siguiente:

export JAVACMD=‘which java‘

Por último, vamos a comprobar si se ha instalado correctamente

java -version

Tenemos que observar: [cloudera@quickstart ~]$ java -version openjdk version "1.8.0\_171"

## ElasticSearch

En primer lugar, importamos una clave que nos permitir instalarnos ElasticSearch

sudo rpm --import <http://packages.elastic.co/GPG-KEY-elasticsearch>

Vamos a crear un nuevo repositorio yum para ElasticSearch. (yum es una herramienta libre para sistemas Linux que sirve para eliminar, actualizar e instalar paquetes)

sudo vi /etc/yum.repos.d/elasticsearch.repo

Ahora hemos abierto un editor vi, como el chero no exista automáticamente se crear. Tenemos que copiar y pegar el texto de abajo (antes de copiar pulsar la tecla ’i’ para que se habilite la escritura). Para salir del vi y guardar los cambios, ejecutar ’**esc**’ + ’**:wq**’.

[elasticsearch-5.x] name=Elasticsearch repository for 5.x packages baseurl=https://artifacts.elastic.co/packages/5.x/yum gpgcheck=1 gpgkey=https://artifacts.elastic.co/GPG-KEY-elasticsearch enabled=1 autorefresh=1 type=rpm-md

Ya estamos listos para instalar ElasticSearch.

sudo yum -y install elasticsearch

Vamos a editar la configuración de Elastic:

sudo vi /etc/elasticsearch/elasticsearch.yml

Buscamos la l nea que empieza por "network.host", la descomentamos y escribimos localhost.

Tiene que quedar así:

network.host: localhost

Para inicializar el servicio ejecutamos:

sudo service elasticsearch start

Por último, para comprobar que lo hemos hecho bien, ejecutamos:

curl <http://localhost:9200>

O directamente en copiando en el navegador: <http://localhost:9200>Tenemos que ver algo de este estilo:



## Configuración de Kibana

Vamos a crear un repositorio yum para Kibana también:

sudo vi /etc/yum.repos.d/kibana.repo

Ahora hemos abierto un editor vi, como el chero no exista automáticamente se creará. Tenemos que copiar y pegar el texto de abajo (antes de copiar pulsar la tecla ’i’ para que se habilite la escritura). Para salir del vi y guardar los cambios, ejecutar ’esc’ + ’:wq’.

[kibana-5.x] name=Kibana repository for 5.x packages baseurl=https://artifacts.elastic.co/packages/5.x/yum gpgcheck=1 gpgkey=https://artifacts.elastic.co/GPG-KEY-elasticsearch enabled=1 autorefresh=1

type=rpm-md

Ya podemos instalar Kibana

sudo yum -y install kibana

Inicializamos el servicio

sudo service kibana start

Para probar si lo hemos hecho bien, buscamos la siguiente web en el navegador y ya debemos estar dentro de la página de inicio de Kibana: <http://localhost:5601>

Usuario: elastic

Password: changeme

## Logstash

Por último, vamos a crear un repositorio yum para Logstash:

sudo vi /etc/yum.repos.d/logstash.repo

Ahora hemos abierto un editor vi, como el chero no exista automáticamente se creará. Tenemos que copiar y pegar el texto de abajo (antes de copiar pulsar la tecla ’i’ para que se habilite la escritura). Para salir del vi y guardar los cambios, ejecutar **’esc’** + **’:wq’**.

[logstash-5.x] name=Elastic repository for 5.x packages baseurl=https://artifacts.elastic.co/packages/5.x/yum gpgcheck=1 gpgkey=https://artifacts.elastic.co/GPG-KEY-elasticsearch enabled=1 autorefresh=1 type=rpm-md

Instalamos Logstash

sudo yum -y install logstash

Ahora vamos a crear el chero de configuración que utilizaremos para que Logstash se conecte a twitter:

sudo vi /etc/logstash/conf.d/EMT\_logstash.conf

Tenemos abierto el editor con un fichero vacío, hay que copiar el fichero de abajo, sustituyendo las claves de twitter por las nuestras.



## Instalación X-pack

Vamos ahora a instalar esta nueva herramienta de Elastic que nos va a permitir visualizar algoritmos de Machine Learning sobre nuestros datos.

sudo /usr/share/elasticsearch/bin/elasticsearch-plugin install x-pack

Y ahora en Kibana

sudo /usr/share/kibana/bin/kibana-plugin install x-pack

Reiniciamos los servicios y ya estar a listo:

sudo service elasticsearch restart

sudo service kibana restart

Volvemos a cargar la web de Kibana <http://localhost:5601> y nos logueamos con:

User: Elastic

Password: changeme

## Configuración de SUPERUSUARIO en Kibana

En Management🡪Users

Nombre: logstashinternal

Password: changeme;

roles: superuser;

fullname: Internal Logstash User (No importa este nombre)

## Configuración del Template ElasticSearch

Hasta ahora ya tenemos las tres herramientas instaladas. Solo nos queda la configuración del template de ElasticSearch que interpretará los tweets. Ejecutamos el siguiente comando y copiamos todo el chero de abajo.

sudo vi /etc/elasticsearch/EMT\_template.json



## Ejecución

Ya tenemos toda la configuración hecha, para que Logstash empiece a leer la información de Twitter y a enviársela a Elasticsearch, ejecutamos el siguiente comando.

sudo /usr/share/logstash/bin/logstash -f /etc/logstash/conf.d/EMT\_logstash.conf

Cuando empiecen a aparecer puntos en la consola significa que el proceso ya se ha iniciado y podemos ir a Kibana a ver los datos.

En caso de tener un fichero .csv que se almacena en HDFS y se desea consumir la información desde Kibana, y se presentan problemas cuando se configura en la MV, consultar el archivo siguiente:

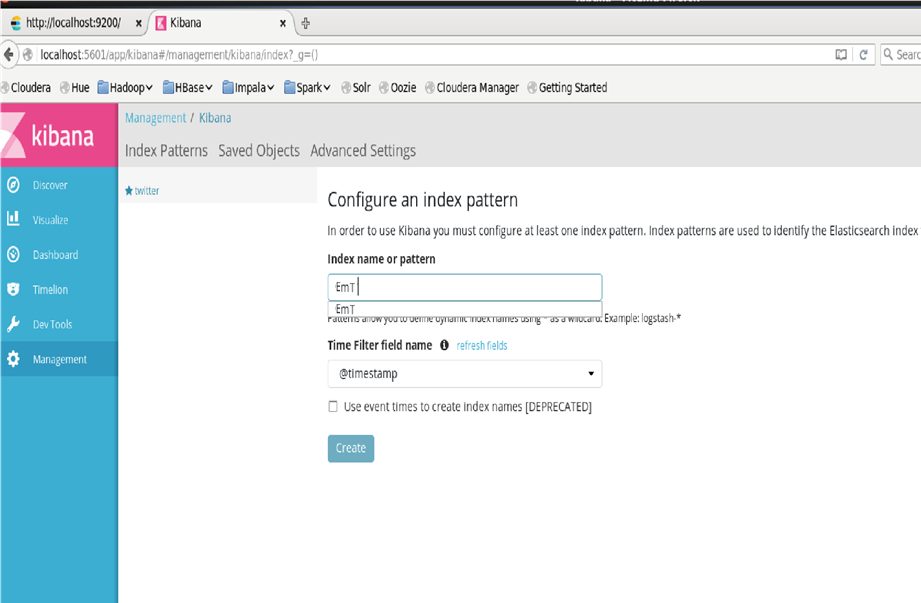


# Capítulo 3

## Cuadro de mandos en tiempo real

En primer lugar, tenemos que cargar el índice (**EMT**) que ha creado Elasticsearch en Kibana, como en la siguiente imagen:

Figura 3.1: Creación del Index “EmT” configurado anteriormente



A partir de aquí ya tenemos un cuadro de mando para visualizar información en tiempo real.