youtube video Elasticsearch Tutorial ELK stack Tutorial Intellipaat

ubuntu

Installation pre reqisites

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

2Gb Ram

20 Gb Memory

Installation

\*\*\*\*\*\*\*\*\*\*\*\*

launch ec2 instance

login

[ubuntu@ip172-31-25-152]sudo apt-get update

sudo apt-get install -y openjdk-8-jdk

sudo apt-get update

sudo apt-get install -y nginx

#nginx is used as webserver to access our kibana

since kibana cant be accessed from outside

so we will be using nginx to acces kibana from out cuurent host instead of our instance

sudo systemctl enable nginx

install elastic search

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

sudo wget https://artifacts.elastic.co/downloads/elastic search-7.2.0-amd64.deb

sudo dpkg -i elasticsearch-7.2.0-amd64.deb

#install kibnaa

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

sudo wget https://artifacts.elastic.co/downloads/kibana/kibana-7.2.0-amd64.deb

sudo dpkg -i kibana-7.2.0.deb

#before installing logstash we have to instll a pkg

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

sudo apt-get install apt-transport-https

#install logstash

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

sudo wget https://artifacts.elastic.co/downloads/logstash/logstash-7.2.0-amd64.deb

sudo dpkg -i logstash-7.2.0.deb

#install filebeats

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

sudo wget https://artifacts.elastic.co/downloads/filebeats/filebeats-7.2.0-amd64.deb

sudo dpkg -i filebeats-7.2.0.deb

Configure elastic search and kibana

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

elasticsearch

\*\*\*\*\*\*\*\*\*\*\*\*\*

vi /etc/elasticsearch/elasticsearch.yml

cluster name: my-application

node.name: node-1

network.host: localhost # 192.168.0.1

http.port: 9200

ctrl+O

ctrl+s

systemctl start elasticsearch

kibana

\*\*\*\*\*\*

vi /etc/kibana/kibana.yml

server.port: 5601

server.host: "localhost"

save and exit it

sudo systemctl start kibana

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Now we have to configure our elasticsearch and kibana

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

configure kibana such that when you access it you have a username and password using nginx

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

you need to use htpasswd for that

sudo apt-get install -y apache2-utils

sudo htpasswd -c /etc/nginx/htpasswd.users kibadmin

Newpasswd: password

it shows adding password for user kibadmin

sudo nano /etc/nginx/htpasswdd.users

you will see the hash password

Configure nginx default file

sudo nano /etc/nginx/sites-available/default

#paste the following data in this file

server {

listen 80;

server\_name Instance Private IP>; #privayte ip of the ELK server

auth\_basic "Restricted Access";

auth\_basic\_user\_file /etc/nginx/htpasswd.users;

location / {

proxy\_pass http://localhost:5601;

proxy\_http\_version 1.1;

proxy\_set\_header Upgrade $http\_upgrade;

proxy\_set\_header Connection 'upgrade';

proxy\_set\_header Host $host;

proxy\_cache\_bypass $http\_upgrade;

}

}

what is this file doing

when we type ip:80 in browser it directs to nginx, and the nginx directs you to localhost:5601

so we are acesing kibana indirectly

sudo systemctl start nginx

go to browser and enter ip, it asks for username and password

username: kibadmin

password: password

it will open kibana page

now clcik explore on your own

ELK hands on

\*\*\*\*\*\*\*\*\*\*\*\*\*

1. collect static apache logs using logstash and analyse them using kibana

2. collect static .csv using logstash and analyze them using kibana

3. collect real time web logs and configure beats to inject them into elasticsearch and analyse them using kibana

1. collect static apache logs using logstash and analyse them using kibana

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

first we have to get some static data

sudo wget https://logz.io/sample-data

ls #it shows sampe-data

sudo cp sample-data apache.log #we changed the name to apache.log

cat apache.log

#it shows data in apache.log

#now lets go and inject this inorder to vizualise this in kibana

#for that we have to first go to logstash

cd /etc/logstash/conf.d/

vi apachelog.conf #this is a new file and we will create a pipeline

#this content is available invodeo at 39.57 min

input {

file {

path => "/home/ubuntu/apache.log" // this is where the log is located

start\_position => "beginning"

sincedb\_path => "/dev/null" //since there is no DB we wont link one here

}

}

filter {

grok {

match => { "message" => "%{COMBINEDAPACHELOG}" }

}

date {

match => [ "timestamp" , "dd/MM/yyyy:HH:mm:ss Z" ]

}

geoip {

source => "clientip"

}

output {

elasticsearch {

hosts => ["localhost:9200"]

index => "petclinic-prd-1"

}

}

sudo systemctl start logstash

#go to browser: logstash must have taken the data from our instance and sent it to the elastic search and elastic search have indexed it under the title we have given it

got to browser

click discover

step1 click index pattern

type index pattern: same name as given earlier i.e petclinic-prd-1

step2 configure settngs: @timestamp

2. collect static .csv using logstash and analyze them using kibana

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

sudo curl -O https://raw.githubusercontent.com/PackPublishing/Kibana-7-Quick-Start-Guide/master/Chapter02/crimes\_2001.csv

# it downloads a csv file for us

cat crimes\_2001.csv

it shows the content of the file

cd /etc/logstash/conf.g/

ls

you will see already created apachelog.conf file

sudo nano crimes.conf

input {

files {

path => "/home/ubuntu/crimes\_2001.csv"

start\_position => beginning

}

}

filter {

csv {

column => [

"ID",

"Case Number",

"Date",

"Block",

"IUCR",

"Primary Type",

"Description",

"location Description",

"Arrest",

"Domestic",

"Beat",

"District",

"Beat",

"Ward",

"Community Area",

"Fbi",

"X Coordinate",

"y coordinate",

"year",

"updated on",

"lat",

"long",

"location"

]

seperator => ","

}

}

output {

elasticsearch {

action => "index"

hosts => "["localhost"]

index => "crimes"

}

}

sudo systemctl start logstash

go back to browser --> management ---> index pattern under kibana --> createnew index pattern

index pattern : crimes

configure settings: @timestamp

create index pattern

once its done then you can play with thee data as per the requirements

3. collect real time web logs and configure beats to inject them into elasticsearch and analyse them using kibana

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

sudo filebeat modules list

it shows enabled and disabled list of modules

sudo filebear modules enable nginx

sudo filebear modules enable system

sudo filebeat modules list

now it shows that ngninx and system are enabled

it means file beat is collecting data from nginx and system

we have to configure the beats so that it knows where to collect the data from and what all data is to be collected

cd /etc/filebeat/modules.d/

you can see yml files and most of them disabled

sudo nano nginx.yml

#here we have to mention var.path i.e the location of log files

#Access logs

access: #this is already present

enabled: true #this is already present

var.paths: [/var/log/nginx/access.log\*"]

#Error logs

error: #this is already present

enabled: true #this is already present

var.paths: [/var/log/nginx/error.log\*"]

sudo nanno system.yml

#sys logs

syslog: #this is already present

enabled: true #this is already present

var.paths: [/var/log/syslog/\*"]

#auth logs

auth: #this is already present

enabled: true #this is already present

var.paths: [/var/log/auth.log\*"]

sudo systemctl start filebeat

now check the kibana index list u will find

filebeat 56......aw376.

you can acces the data by entering index pattern and timestamp and then you can vizualise the data indiff formats

sudo filebear setup -e

the above command pulls different dashboard formats and makes them available for you in kibana

before using this command we have to make sure that kibana is running

tools

configuration

infra related

jeninspipepline

dependency

git project access

i need to access project reposiories

access to entire project

jira, bitbucket

urls of instances, name of

architecture diagram

project overview

files docs,

prepare kt plan