Major Project Report on

"Centralized logging using ELK stack"

In partial fulfillment of requirements for the degree of

Bachelor of Technology (B. Tech.)

in

Computer Science and Engineering

Submitted by

Aprajita Halder

Under the Guidance of

Mr. Anand Pillai

Department of Computer Science and Engineering

SCHOOL OF ENGINEERING AND TECHNOLOGY

Mody University and Science and Technology
Lakshmangarh, Distt. Sikar-332311

June 2021

A C K N O W L E D G E M E N T

I am thankful to societe generale, for providing me the opportunity to work on “Centralized logging using ELK stack”. Also, I would like to express my gratitude towards CDC department and Mody University for providing me the chance to work for such an organization. Their immense help and guidance helped me to gain experience and knowledge in real world scenario.

Aprajita Halder

CERTIFICATE

This is to certify that the major project report entitled "Centralized logging using ELK stack", submitted by Ms. Aprajita Halder, as a partial fulfillment for the requirement of B.Tech VIII Semester examination of the School of Engineering and Technology, Mody University of Science and Technology, Lakshmangarh for the academic session 2020-2021 is an original project work carried out under the supervision and guidance of Mr. Anand Pillai has undergone the requisite duration as prescribed by the institution for the project work.

PROJECT GUIDE:

Signature:

Name: Mr. Anand Pillai

Date:

HEAD OF DEPARTMENT

Signature:

Name: Dr. A Senthil

Date:

EXAMINER-I:

Signature:

Name:

Date:

EXAMINER-II

Signature:

Name:

Date:

ABSTRACT

This is a spring boot project, which generates a log file of the application. We gave the logfile to logstash. The data is documented to elasticseach. Elasticsearch is a NoSQL database, used for storinf unstructures data. It is based on apache lucene. The data is stored in JSON format. Then data from elasticsearch will be send to kibana to visualize index patterns. The project consists of single microserver. ELK stack can be storing and processing huge logs, generated by multiple microservers. We will need a spring boot starter web dependencies. In application we defined API's to generate logs. Spring framework is a powerful opensource, lightweight tool. It is popular because of it's huge resources and in-built frameworks, which allows developers to write code with minimal configuration. It makes development process faster and easier.

Table of Contents

Chapter 1: Introduction

1. INTRODUCTION

ELK Stack

ELK stack refers to Elastic search, logstash and kibana. These are open-source tools used for storing, processing and visualizing the data. They act as a log-management platform.

Because of it's open-source, it has outreached the popularity as compared to splunk, another popular log-management platform.

Earlier it only compromised of Elastic search, logstash and kibana. All these were product of company "ELASTIC". Later Beats was added to the trio.

Log Analysis

In competitive world, corporates and companies, handling records and data is important to provide meaningful insights, analysis, risk monitoring, security flaw, troubleshooting etc. They cannot afford to loose customers over downtime. Hence, log-analysis is very important for companies to provide smoother experience.

Elasticsearch

It is a search and analytics engine based on Apache Lucene. It is a NoSQL database, though recently, it can be used with sql database also. It helps in solving complex problems of logs. Generally, we use REST API to provide data to elastic seach.

Logstash

It is a processing pipeline that gets data from various sources, processes it, and then sends it to kibana for visualization. We can also apply filter in between input and output as per our requirements.

Kibana

It helps in visualizing the data. It helps in searching and analyzing in browser, the elastic indices. We can easily analyze using graphs in large volume data.

Rest API

REST API refers to representational state transfer, which provides an interface to interact with. It supports GET, POST, PUT, DELETE methods.

Spring Boot

Spring Boot enables rapid application development, with help of spring framework and embedded servers. It is commonly referred as “frameworks of frameworks”. Now, instead of writing lengthy and repetitive code, developers can directly use in-built frameworks. It increases the development process.

1.1 PRESENT SYSTEM

A spring boot application is created. It contains a class called 'User.java' in spring boot with two fields, 'id' and 'name'. There is a controller class named 'ElkStackExampleApplication.java', with a method, which will return list of objects for user. From controller, based on the id, it fetches the user from the list. If there is no user object with that id, return null and throw exception.

Elastic search batch file is run and kibana is configured. It takes a couple of second to up the nosql elasticsearch database. Now go to kibana folder. We open the 'kibana.yml' file to configure. We want to inform kibana where our elasticsearch is running. Now go to bin folder. There will be a 'kibana.bat' batch file. Then kibana and elasticsearch is verified in the browser using ‘localhost’ with their respective ports.

Then we generate the log file of this application. In 'logstash.conf', we gave input, filter and output. In input, the file path is given and the start_position as 'begginning'. In output section, we are giving the log from specified path of input to elasticsearch. In logstash console, you will be able to see logs with timestamp.

Now in kibana console we create index pattern. We go to management. Click on 'create an index pattern'. Give the index name which is current time stamp index. We do nnot enable the time filter. Now we will be able to see all logs in kibana console. Using ‘discover’ in kibana we will see the index pattern.

1.2 PROPOSED SYSTEM

Develop a spring boot application. Create a class called 'User.java' in spring boot with two fields, 'id' and 'name'. Then create a controller class named 'ElkStackExampleApplication.java'. There write a method, which will return list of objects for user. From controller, based on the id, fetch the user from the list. If there is no user object with that id, return null and throw exception.

Before developing this application, run the elasticsearch batch file. Also configure the kibana.yml file. Type 'cmd' to run that. In prompt type 'elasticsearch.bat'. It will take a couple of second to up the nosql elasticsearch database. Now go to kibana folder. Go to the config folder. Open 'kibana.yml' file to configure. Scroll down a bit. There you will find a localhost line. Uncomment that line. We want to inform kibana where our elasticsearch is running. Now go to bin folder. There will be a 'kibana.bat' batch file. Open command prompt and run kibana.bat in it. To verify elasticsearch, go to browser and type 9200 localhost. There you will see a response. To verify kibana console, check in prompt it's port no. and go to browser. There you will see an interactive kibana dashboard.

After that generate the log file of this application. To generate log file in logstash, go to application.yml and type their 'logging' and specify the file and give file path of logs and give name of log-file, 'elk-stack.log'. Go to logstash folder, go to bin, and run the logstash batch file. In 'logstash.conf', give input, filter and output. In input, specify the file path and start_position as 'begginning'. In output section, we are giving the log from specified path of input to elasticsearch. Mention the elasticsearch host port there

Download and unzip the logstash. Then create a 'logstash.conf' file. Inside that config file, we need to tell out logstash, where our log-file is located. Copy this conf file into bin folder of logstash and paste file. Now type command and run the 'logstash.bat' file. Then give file name, which is 'logstash.conf'. The command is 'logstash -f logstash.conf'.

There you can get the logstash running port. Now try to hit the API and get user info, with valid and invalid id's.

In logstash console, you will be able to see logs with timestamp.

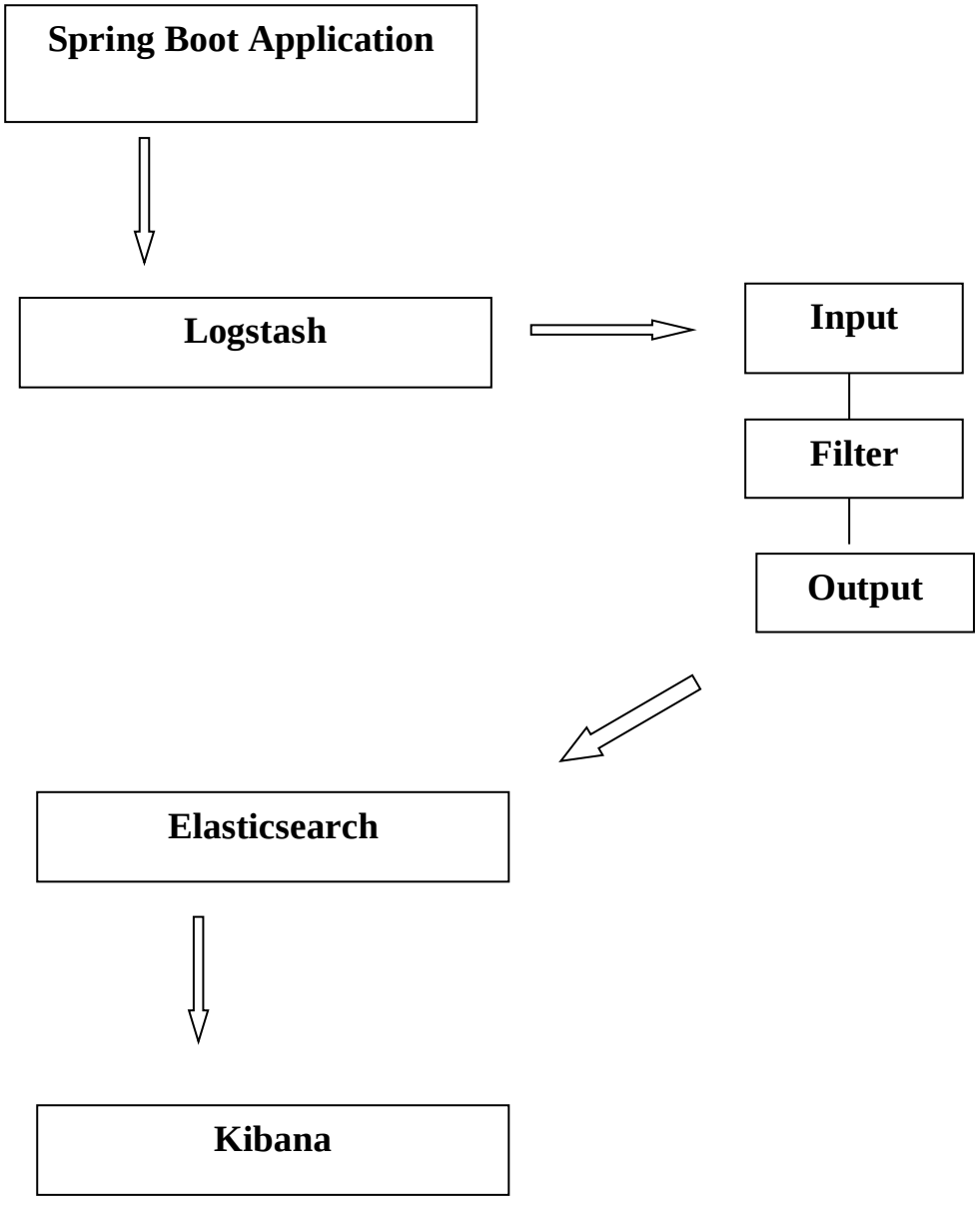
Now go to kibana and create index pattern with the same index. To create your own index, you need to make changes in 'logstash.conf' file, by giving index in output folder.

To create index pattern go to management. Click on 'create an index pattern'. Give your index name which is current time stamp index. You may or may not enable the time filter. Now we will be able to see all logs in kibana console. Click on 'Discover'. There you will see your index pattern.

Chapter 2: System Design

SYSTEM FLOWCHART

The spring boot application generates logs, which is given to logstash. The logstash, processes the data. It consists of input, output and filter. Then, the processed data is stored in the elasticseach, which is a NoSQL database. Then, the data is viewed and analyzed through user interactive tool kibana.



Chapter 3 : Hardware and Software details

Hardware Configuration

Processor (CPU):	Intel Core i3 (sixth generation or newer) or equivalent
Operating System:	Microsoft **Windows 10** x64
Memory:	8 GB RAM
Storage:	500 GB internal storage drive
Monitor/Display:	15" LCD monitor
Other:	802.11ac 2.4/5 GHz wireless adapter

Softwares Required

Eclipse/STS/Intellij (IDE)

Elasticsearch

Logstash,

Kibana

JDK 1.8+ or later

Maven 3.2+ / Gradle 4+

Chapter 4: Implementation Work details

4.1 Real Life Applications

- ELK stack is widely used in development and troubleshooting. It is used to deliver high-end applications. Also the generated logs help us to troubleshoot by showing exceptions and errors.

- It is also helpful in cloud operations. It act as one-time solution for complex devops and IT engineer tasks. Deployment is a complex task with so many steps involved. This saves the production time.

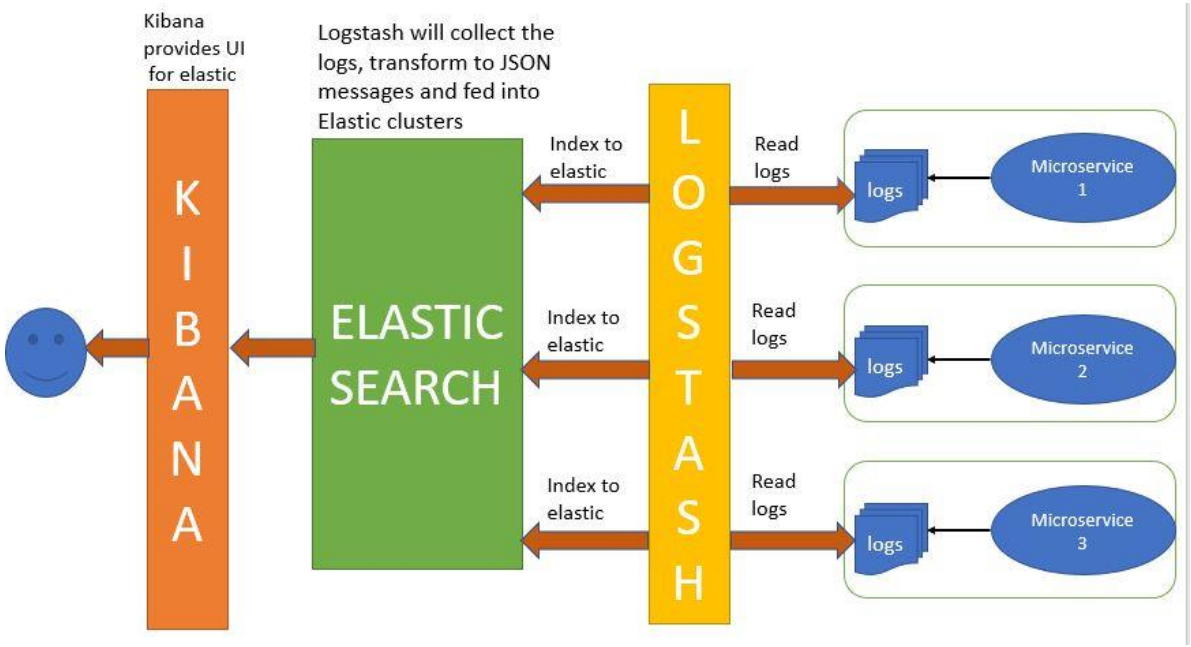
- Due to continuous integration of data, we need an application which can monitor the logs continuously. This saves the companies from upcoming problems and logging problems.

- It is also use for risk implementation and RIFT management. It has application in security analysis as well.

- This also has application in AI/ML field. In data science, we have to deal with huge amount of data. Elasticsearch can be used to store that data and kibana to visualize and draw meaningful conclusions from it.

- Many industries have adopted elk stack instead of it's competitor splunk, because it is open-source, free to use and easy configuration.

4.2 Data Implementation and Program Execution



- First download the three components which are elasticseach, logstash and kibana. Go the downloaded folder. Go to bin. There will be a batch file, named elasticseach.bat. Type 'cmd' to run that. In promt type 'elasticsearch.bat'. It will take a couple of second to up the nosql elasticsearch database.

- Now go to kibana folder. Go to the config folder. Open 'kibana.yml' file to configure. Scroll down a bit. There you will find a localhost line. Uncomment that line. We want to inform kibana where our elasticsearch is running. Now go to bin folder. There will be a 'kibana.bat' batch file. Open command prompt and run kibana.bat in it.

- To verify elasticsearch, go to browser and type 9200 localhost. There you will see a response. To verify kibana console, check in prompt it's port no. and go to browser. There you will see an interactive kibana dashboard.

- Create the spring boot application and controller class. Now generate log file of this application. Add log statement of the success as well as for failure.
- Logger.info() and logger.error() was used for that.

- To generate log file in logstash, go to application.yml and type their 'logging' and specify the file and give file path of logs and give name of log-file, 'elk-stack.log'.

- Now run the application. There you will see in logs folder, our log file is generated.

- Now configure the logstash, to push the entire generated log to logstash. Go to logstash folder, go to bin, and run the logstash batch file. Download and unzip the logstash. Then create a 'logstash.conf' file. Inside that config file, we need to tell out logstash, where our log-file is located.

- Copy this conf file into bin folder of logstash and paste file. Now type command and run the 'logstash.bat' file. Then give file name, which is 'logstash.conf'. The command is 'logstash -f logstash.conf'. There you can get the logstash running port, '9600'. Now try to hit the API and get user info, with valid and invalid id's. In logstash console, you will be able to see logs with timestamp.

- Now go to kibana and create index pattern with the same index. To create your own index, you need to make changes in 'logstash.conf' file, by giving index in output folder.

Chapter 5: System Testing

ELK

Check the ports of elasticsearch, kibana and logstash in browser, to see if they are up and running. Go to bin folders and run all the batch files. Also configure kibana.yml to let kibana know where elasticseach is running.

Generating Logs

In logs folder, ensure whether the logs are generated or not after starting the server. Logger.info() and logger.error() commands will be used.

Verifying Ids

Enter both valid and invalid id to check if all logs are generated properly. It should produce result in case of valid ids and throw exception and print error in console in case of invalid id.

Copy the URL 'get_user'. Check the port form application.yml file, which is 9898 in this case. Type 'localhost:9898/get_user/3'. If we have user id, it will display in browser. If we type a user id which is not present, the console will show exception with the error message.

There you can get the logstash running port, '9600'. Now try to hit the API and get user info, with valid and invalid id's.

Logstash Configuration

In 'logstash.conf', give inout, filter and output. In input, specify the file path and start_position as 'begginning'. In output section, we are giving the log from specified path of input to elasticsearch. Mention the elasticsearch host port there, which is 9200.

The command is 'logstash -f logstash.conf'.

In logstash console, you will be able to see logs with timestamp.

Kibana Verification

Go to browser. Type 'localhost:9200/cat', to verify indexes. There you will find 'indices'. Type it as 'localhost:9200/cat/indices', there you will find index internally created by elk.

For logging purpose, the logstash with current time-stamp was created. You can view that content in kibana console. Copy that, and give that index 'localhost:9200/logstash-202004-05-000001' and seacrh it. You will see logs.

Chapter 6: Individual Contribution

At first I was asked to study and grasp the concept of REST Api in spring boot framework. I was provided with eclipse IDE in my sumo desktop. Then, ELK logging mechanism was explained by someone in the company. We were asked to revert back to the team and gave presentation in our understanding of the topic.

I created the spring boot project in eclipse IDE. It compromised of two REST APIs. One which gave success result on valid user id and other threw exception and generate error in case of invalid user id. Then, I generated the logs from the application by providing valid and invalid ids in the server. Those logs were pushed to logstash. In logstash, I provided the output to elasticsearch, my mentioning the port no. in .conf file. The elasticseach acted as an storing agent. From there, I viewed the index patterns from kibana console.

Chapter 7: Conclusion

7.1 LIMITATION

- To generate this application, we have used spring boot, which increases the deployment file size due to it's added dependencies. The configuration and deployment is a complex task.

- Elk stack also requires complex management and maintenance cost is also high.

- Right now the application only has user information and generate logs on two API's.

7.2 SCOPE FOR FUTURE WORK

- Current system is based on one microserver. In future, we may add multiple microserver, which will generate enormous amount of logs.

- We will add more advanced features and complex task in our user class.

- In logstash, we will process the data in-depth for patterns, by adding filters.

- Deployment will be done of a production-ready application with enhanced testing and verification process.

Bibliography

https://spring.io/guides/gs/rest-service/

https://logz.io/learn/complete-guide-elk-stack/