ELK Logging Platform

info@crozaint.com
www.crozaint.com

Microsoft | AZURE PARTNER

amazon web services | Partner Network
CONSULTING PARTNER

crozaint

Summary

The Client had an infrastructure which had many micro-services running in docker containers. The log assessment was becoming a hectic task which used many admin hours, even days to get a basic assessment. So many resources hours were lost as a result of these. This was a gradual process thus it was contributing slow business cycles.

Objectives

To centralize log aggregation/parsing system which can provide valuable business insights through various metrics collected from the services/servers used in the client ' s infrastructure.

Assessments

Although the leading industrial standard tool to accomplish is Splunk , we have created an open source solution to implement the core features in the client requirement. The ELK stack could ship the logs from a variety of sources which includes Microservices logs, Server logs such as Auditd, SysLogs, Firewall/IDS logs, Network and Security Data, Infra/App Data etc.

Technology Stack

- ElasticSearch
- LogStash
- Kibana
- Beats
- Apache Kafka

Elasticsearch is the heart of the ELK Stack. The following features makes the Elasticsearch a worthy search engine. The features include Fuzzy Search Support, Autocomplete Capabilities, Speed, Scalability, Restful API, Faceted Search.

Logstash is a data-collection and log-parsing engine which ingests data from multiple sources and forward to a stash of choice. Most common use cases uses elasticsearch to index the data collected.

A Logstash config file three sections: input , filter and output sections.

As Sample Logstash Configuration file is as follows.

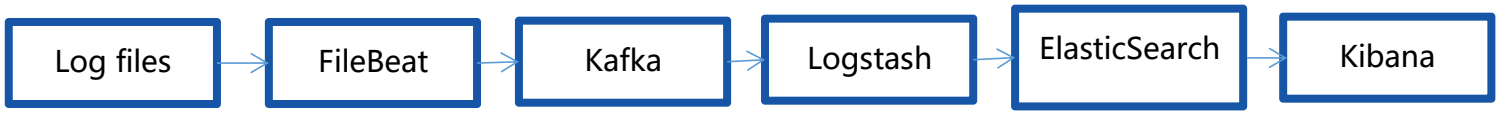
```
# This is a comment. You should use comments to describe
# parts of your configuration.
input {
  ...
}

filter {
  ...
}

output {
  ...
}
```

Kibana produces visualization from the content indexed in Elasticsearch . Kibana displays the data set indexed in elasticsearch using using time series bar, line and scatter plots, pie charts and can also used to visualize the Geospatial Data on Maps.

By adding **Beats** to the ELK Stack, all kinds of data can be shipped to the Logstash for Data Aggregation and Data Transformation Processes.

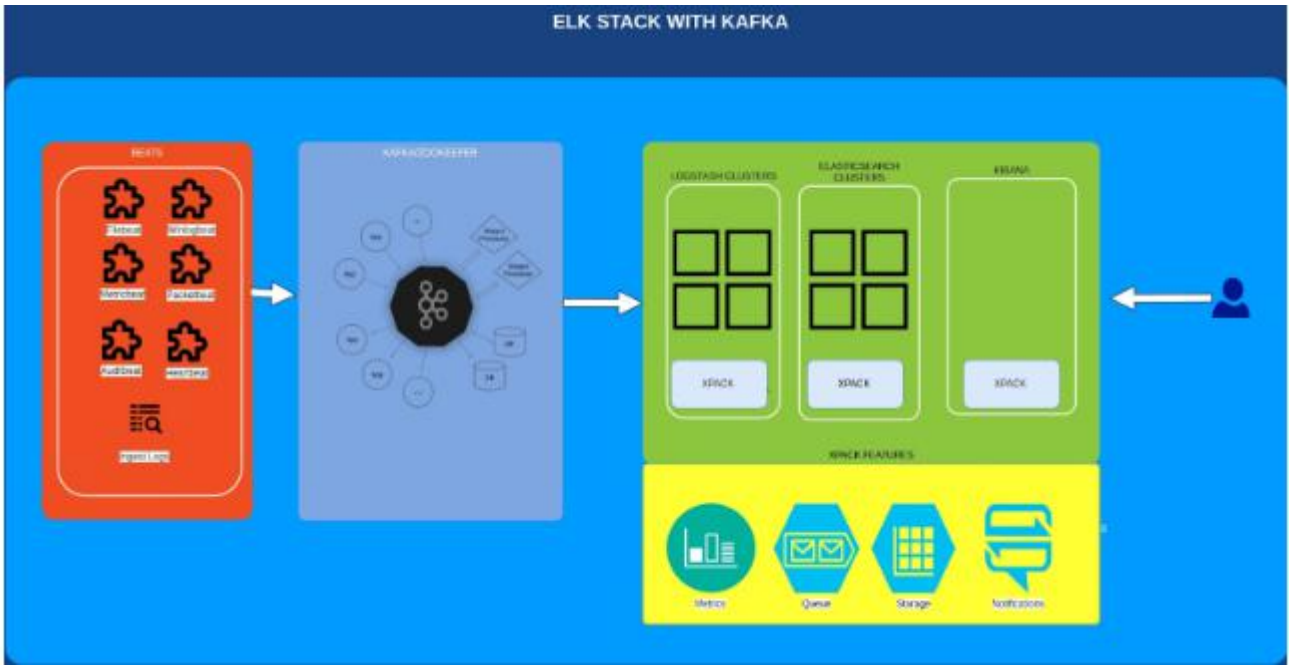
Kafka is an open-source distributed stream-processing software which can be used as a data buffer between the log shippers and the data parsing section

Event Flow



- The log files are shipped using Beatplugin which is forwarded to the Kafka. The filebeat.yml file specifies the port and the host for Kafka server. This will send events to the topic created in the Kafka. The Beat plugin will act as the Producer.
- The consumer console collects all the events from the producer The Kafka then forwards the data to the Logstash using the Logstash Input plugin.
- The filter section of the Logstash performs the parsing, filtering and transformation
- The output section of logstash.conf which contains host and the port of the elasticsearch

Architecture Diagram



Bussiness Impact

- Brought down the Error Resolution Time from a couple of hours to within an Hour
- Real Time Anomalies are spotted on and contributed to business insights
- The insights revealed many critical information which were then used to tweak mission critical operations