Comparison of DevOps Log Monitoring Tools

Nisarga Jagadish S[1] Prof. Shantala C P[2]
[1]PG Scholar [2]Professor and HOD
[1,2]Department of Computer Science & Engineering
[1,2]CIT, Gubbi

Abstract— Devops is an approach in which operation team and development engineers team work together throughout entire life cycle. Devops can be divided into different domains namely Deployment and configuration, Build and test, Monitoring and Log monitoring. In this paper, author gives the description of log monitoring, different types of log monitoring tools and comparison between log monitoring tools.

Key words: Log Monitoring, DevOps

I. INTRODUCTION

DevOps is a combination of development and operation. It is a software development method that is comprised of communication, sharing of information and usage of web services, integration, automation and it also measures the co-operation between the developers and professionals. It is emerged from two major trends. [1] Firstly, agile system administration or agile operations and secondly, the value of co-operation between development and operation team at each stages of DevOps life cycle. DevOps has good understanding with agile approach and Lean approach. ‘Dev’ denotes makers of software and ‘Ops’ deals with the work that is after the creation of software. [1]

A. *Log Monitoring*

In debugging process log data is used by developers. Log data is treated as valuable tool in each stages of DevOps life cycle. Log data are used as debugging tool and also used for system load and performance testing in first stage of DevOps life cycle. [2] Log data is used for production monitoring and production trouble shooting in second stage of DevOps life cycle. Log data is used for web analytics and business metrics in third stage of DevOps life cycle. Log data is difficult in case of measuring the success of service and to troubleshoot the issue of the problem when aroused. As log data understands the issues faster, it analyzes the issue before it grows and maintains the healthy communication with all the groups in organization. So it is said log data improves the QA process. Real scalability and security misses the attention while managing the log data. [3]

II. LOG MONITORING TOOLS

These log monitoring tools are used to monitor the log data of any application, organization, system or web browser. The main job of these tools is to monitor the logging sytem of the system.There are many log monitoring tools available. Some of the DevOps logs monitoring tools are described. Tools described are Splunk, Sumo Logic, Logstash, Kibana, Paper Trail and Loggly.

A. *Splunk*

Splunk is an American based multinational cooperation company which is situated at San Francisco, California. Splunk tool is used for searching, analyzing and monitoring huge amount of data generated by machine. It generates the data, indexes the data and relates the data in searchable repository so that it can generates graphs, reports, alerts, dashboards and visualization. Main aim of Splunk is make the data accessible throughout the organization, also identifies the data patterns to be searched, provides metrics to it and troubleshoots the problem raised in the organization. [4]

Splunk server is scalable software which is written in C/C++ and python. Mainly it takes data generated by any server, applications or it can be any device. Splunkd and Splunkweb is called the Splunk developer API and this developer API is accessible by REST and SOAP or CLI (Command Line Interface). Splunk main function is to index the collected data and also to search for logs or log data. [5]

B. *Sumo Logic*

Sumo Logica was started in 2010 by ArcSight Veterans Kumar Saurabh and Christian Beedgen. Sumo logic was funded by Accel partners, Greylock partners, Sequoia capital, Sutter Hill Ventures and Angel investor Shlomo Kramer. Sumo Logic is situated at Redwood City of California. It is Cloud based log management system which converts the machine generated big data to real time IT insights. In June 2012 Sumo Logic free was introduced which has full freemium functionality and in August 2012 Sumo Logic for VMware was introduced which allowed the users to search, view and analyze all VMware log data. [6]

After collecting the log data, it creates single largest enterprise data set and Sumo Logic service helps the users to get never done operational efficiency, a security posture and Sumo logic security team to uncover security on large amount of machine generated log data. It has a feature called data retention which keeps log data available always so that cost and complexity for storing, archiving the data is saved. [6]

C. *Logstash*

Logstash is situated at both US and Europe. Logstash is originated from Jordan Sessils’s background in DevOps and system administration. [7]

Logstash is an open source tool which is used for managing event or log data. It is integrated framework which is used for collecting the log, indexing the log, parsing the log, searching the log data and also for storing the data. It collects the log data from various sources, parse them, store them and search them. Logstash contains the various types of filters which is used for modifying, manipulating, transforming the log data or events so that user gets the information which they needed. [7]

D. *Kibana*

Kibana is a web based dashboard which is used for visualizing the elastic search log data. Kibana is a data visualization engine of open source stack (ELK stack). [8] It interacts with all the log data in Elastic search via custom dashboards. It senses the log data to create real time dashboards and then share them with other users in organizations. Kibana converts the log data into visualization like graph, dashboards. It has strong and efficient user interface so that users can sense, search, store and analyze the log data of Elastic search. And the main function of this is to visualize the Logstash data. Kibana is an open source tool which has limit for storing the daily log data; it has the retention time of at least 30days; it has the paid storage system i.e.to store the log data the user pay; it alerts to the admin whenever the log data arrives and also alerts whenever the data retention time has reached; alerting is through email; then coming to searching, it takes the input from Elastic search and it also supports regular expression while searching; it also supports on demand analysis. [8]

E. Paper Trail

Paper Trail is headquartered at Seattle, WA. It was founded in 1999 by Troy Davis. [9]

Paper trail collects all the log files/data from different application, OS or it may be a simple text data file. All collected log data are kept in one place. In Paper trail a light weight daemon Ruby is used to combine one or more log files and also to transfer UDP syslog files to centralized log aggregation. As remote syslog generates UDP packets without depending on the daemon, so it's (remote syslog) configuration does not affect the systems log system. After collecting the log data and placed in a centralized log aggregation, user can the search for the log data. User can search the log data through browser or command line arguments or an API in real time. It also supports alerting. User gets alerted whenever log data/ any error in log system/ whenever the log system is full through nightly email (to registered email only). [9]

F. Loggly

Loggly is headquartered at San Francisco, California. It was founded in 2009 by Jon Gifford, Raffael Marty and Kord Campbell. [10]

Loggly is an open source technologies which has daily limits for storing log data; it is capable of storing log data of 30 days; it does not alert the user or customer whenever data arrives; when compared to other open source tools it has more number of applications and plugins to support this tool; it is used only to manage the log data of cloud based systems; in search option it automatically parse the data according to the input given by user but it does not support regular expression in searching the log data; It does not conduct more number of training and events for users. [10]

III. COMPARISON OF LOG MONITORING TOOLS

This section describes the comparison between log monitoring tools. Tools which are taken for comparison are Splunk, Sumo Logic, Logstash, Kibana, Paper Trail and Loggly. Table 1 shows the comparison between the log monitoring tools.

Combination of Logstash, Elastic search and Kibana is considered to the best log monitoring tool. [7][8] These three tools are open source, written in C/C++, Ruby and python. Firstly Logstash shipper is used for collecting the log data from various sources/ system. [7] After collecting the log data from various sources it stores the log data as a JSON document in Elastic search. This Elastic search is distributed all over the system to store the JSON document. Also it provides the functionality for querying and searching the log data within the JSON documents. Kibana is used for visualizing the Elastic search data also in real time. [8]. Below diagram shows the architecture of this.
Logstash is an open source tool which has daily limits for storing the data; it has the retention time nearly about 30 days; it alerts the user whenever the log data arrives or whenever the data retention time is reached through email; it reduces the thousands of log data into meaningful information of one page; it also manages the log of cloud environment; it automatically parse the data according to the input given by the user and also support regular expression while searching. [7]

Kibana is an open source tool which has limit for storing the daily log data; it has the retention time of at least 30days; it has the paid storage system i.e.to store the log data the user pay; it alerts to the admin whenever the log data arrives and also alerts whenever the data retention time has reached; alerting is through email; then coming to searching, it takes the input from Elastic search and it also supports regular expression while searching; it also supports on demand analysis. [8]

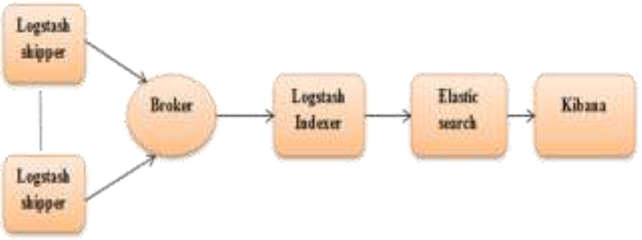

Fig. 1: Architecture of Logstash, Elastic search and Kibana

Tools → Features ↓	Splunk	Sumo Logic	Logstash	Kibana	Paper Trail	Loggly
Software type	Proprietary software	Proprietary software	Open source	Open source	Open source	Open source
Programming language	C/C++ and python	Scala	Ruby, python and C/C++	Ruby, python and C/C++	Python and JavaScript	C/C++ and python
Data retention	Unlimited storage	Limited storage	Limited storage	Limited storage	Limited storage	Limited storage
Daily limits for storing data	No	Yes	Yes	Yes	Yes	Yes
Data confidentiality	Yes	Yes	No	No	No	No
Data visualization	Yes	Yes	Yes	Yes	Yes	Yes

Alerting	Yes	Yes	Yes	Yes	Yes	No
Anomaly Detection	No	Yes	Yes	Yes	No	No
Apps & Plugins	More	Fewer	Fewer	Fewer	Fewer	Fewer
Auto source typing(Parse)	Yes	No	Yes	Yes	No	Yes
Log reduce feature	No	Yes	Yes	Yes	No	Now
On demand data analysis	No	Yes	No	Yes	Yes	Yes
On premise solution	Yes	No	No	Yes	Yes	Yes
Regular Expression support(in search)	Yes	Yes	Yes	Yes	No	No
Rolling Window Alert	Yes	Yes	No	Yes	Yes	No
Scheduled Alert	Yes	Yes	No	No	Yes	Yes

Table 1: Comparison between log monitoring tools

IV. Conclusion

The main job log monitoring tools is to collect, analyze and manage the log data. In these six tools taken, combination of Logstash, Elastic search and Kibana is considered to be best tool. Logstash is used to collect the data, Elastic search is used to analyze the data and Kibana is used to visualize the data.

References

[1] http://theagileadmin.com/what-is-devops/
[2] http://devops.com/features/log-data-valuable-tool-devops-lifecycle-beyond/
[3] https://blog.logentries.com/2014/12/connected-qa-selenium-log-analysis/
[4] http://www.splunk.com/
[5] http://www.splunk.com/view/SP-CAAABF9
[6] http://venturebeat.com/2014/12/04/log-management-company-splunk-gets-competitor-sumo-logic-to-change-its-messaging/
[7] http://logstash.net/
[8] https://www.elastic.co/products/kibana
[9] https://papertrailapp.com/
[10] https://www.loggly.com/blog/loggly-qa-talking-james-urquhart-continuous-integration-deployment-devops-role-log-monitoring/