iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 7 **Issue:** VI **Month of publication:** June 2019

DOI: http://doi.org/10.22214/ijraset.2019.6115

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue VI, June 2019- Available at www.ijraset.com

Intrusion Detection System using ELK Stack

Haritha S Kumar, Nitesh Kumar, Manjula Devi T H
[1]*Student of M. Tech, Dept. Of TCE, DSCE, Bengaluru*
[2]*Scientist/Engineer –'E', ISTRAC, ISRO, Bengaluru*
[3]*Sr. Assistant Professor, Dept. Of TCE, DSCE, Bengaluru*

Abstract: ***Network security is an unavoidable scenario in present network driven world. with many new attacks emerging, securing these are tedious job, yet it developing drastically. with many tools and security system available, the need for more robust features and techniques is mandatory. For high end solution developing a customized tool is necessary, which will not only improve overall security system but also gives an easy monitoring ways. hence this paper aims to build a security system with intrusion detection using elk stack. the configuration, its advantages and disadvantages are mentioned.***

Keywords: ***Network Security, Intrusion Detection System, Open Source Tools, Log Monitoring, ELK Stack.***

I. INTRODUCTION

Network security is vital for an economical system administration. with the type of cyber treats two-faced by business networks within the present, network security is that the vital side of any security strategy. notwithstanding however huge or tiny your business is, they have to make sure a sound network security to stop data thievery and privacy spoofing

Attacks are ever increasing, active attacks are something where the network might get directly attacked for resources and services the user gets aware of it immediately, and passive attacks are silent, the data from the network is being monitored without the user's knowledge, hence detecting these are most needed task.

Many paid services provide excellent network security, but the constant updating and its cost might be a good option for usage, nowadays many paid antiviruses are been released for individual system, but due to its cost many do not opt to choose these.

Here free services are preferred, open source software is mostly used, even by companies for cost effectiveness

Open source tool is one which has a source code published openly, for any specific task. it be modified according to the user's needs. this makes the tool cost free, but configuration is time consuming and at times confusing. this difficulty needs to minimized as much as possible for better time management and higher performance.

There are plenty of software available, choosing among these are first step for creating productive tool which would need less maintenance, elk stack is an open source tool that is available with many features.

II. LITERATURE SURVEY

In computer network security, identifying the intrusion attacks is the most challenging issues. The methods of anomaly detection include predictive pattern generation, neural network, sequence matching, statistics and supervising [1]. IDS are becoming the main part for many organizations after deploying firewall technology at the network perimeter. IDS can offer protection from external users and internal attackers, where traffic doesn't go past the firewall at all. However, the following points are must to always keep in mind. If all of these points are not attached to, an IDS implementation along with a firewall alone cannot make a highly secured infrastructure [2]. Several tools are still having trouble in detection of accurate intruders with minimum hardware and sensor supports. So, there is a need to provide a comprehensive analysis to make a new and effective tool with high accuracy in detection and less in computational cost[3].The need to install logstash and elasticsearch tools in every virtual machine to collect and store log data, and then focus on data processing, instead of the deployment of intrusion detection system on each host. It reduces the use of system resources in each virtual machine, improves the utilization of system resources [4].SSH has turned out to be one of the famous focuses from the entire vulnerabilities which is existed. Assaults on SSH have different qualities [5]. Security Information and Event Management (SIEM) is the state-of-the-practice to address the complexity underlying the collection and normalization of diverse data sources for security analysis. SIEM is the core component of any typical Security Operations Center (SOC), [6], Depending on the goals of the attacker, he would use a specific technique rather than another. Most of port scan techniques give information about state of the targeted ports whereas other techniques give information about the service or the operating system.[7] Data collection is the basic part of the intrusion detection system, and the accuracy, reliability and effectiveness of the dataset directly affect the efficiency of detection. In this paper, we use logstash to collect log information from each virtual machine, and then store them into the elasticsearch cluster. The logstash is a real-time data collection engine that supports all log types, including

system logs, webserver logs, and application logs. The elasticsearch cluster is a distributed search and analysis engine, while supporting data storage [8]. Using logstash tool to collect the system logs from each virtual machine, and storing them into elasticsearch cluster centrally. After that, we analyze all these logs in the detection center and send the results to each virtual machine.[9]

III.LOGS

Logs plays a very vital role in monitoring the system, Log files are the records that Linux stores for administrators to keep track and monitor important events about the server, kernel, services, and applications running on it.
Linux provides a centralized repository of log files that can be located under the /var/log directory.
The log files generated in a Linux environment can typically be classified into four different categories:

A. Application Logs
B. Event Logs
C. Service Logs
D. System Logs

IV.METHODOLOGY

The system is build-up of different modules including one server (admin) system which is gets all the activities of its client systems i.e. client 1 and client 2. The logs are hardened to get all possible events in the system including, system logs, authentication logs, network logs, etc...These logs are difficult to be monitored and filtered from the normal ssh window, hence these are transferred to ELK Stack where kibana is the dashboard where the user interacts with the data, elastic search is the indexing unit which acts a pipeline between logstash and kibana. Logstash takes the logs as input and depending on the configuration sends it to kibana.
The data can be visualized using various graphs and charts according to the need. Alerts are being created or unusual events and abnormal events. Various alerts can be categorized using filters, instant logs of a time can be obtained by using the machine learning feature of the ELK Stack. Visualization of specific logged data has the following benefits as follows

A. Monitor the operations of the system remotely,
B. Communicate information clearly and efficiently via statistical graphics,
C. Plots and information graphics, extract knowledge from the data visualized in the form of different graphs
D. Take necessary actions to better the system.

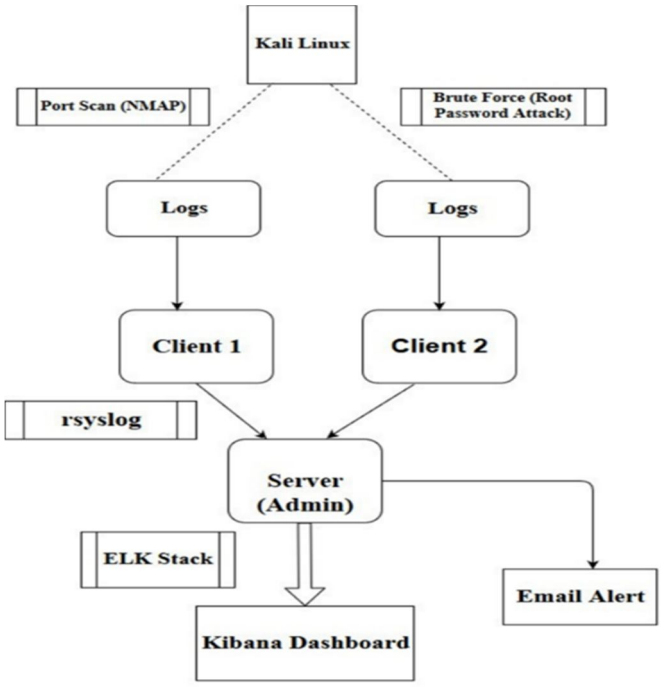


Fig. 1 Flowchart of the methodology

V. ELK STACK

ELASTIC SEARCH is open source analytics and full-text search engine. It's often used for enabling search functionality for different applications. Logstash is a log aggregator that collects data from various input sources, executes different transformations and enhancements and then ships the data to various supported output destinations. Logstash uses logstash conf language to code its config file according to the logs generated, each logs such as secure , messages ,httpd, and so on creates separate index as shown in Fig.2 for the analysis .Kibana is a visualization layer that works on top of Elasticsearch, providing users with the ability to analyze and visualize the data. Kibana lets the user decide the variables of the graphs displayed. Each visualization can be combined and made in a dashboard as shown in Fig.3 Various modules are inbuild in kibana for network security which can be configured only if needed and last but not least beats are lightweight agents that are installed on edge hosts to collect different types of data for forwarding into the stack. Together, these different components are most commonly used for monitoring, troubleshooting and securing IT environments.

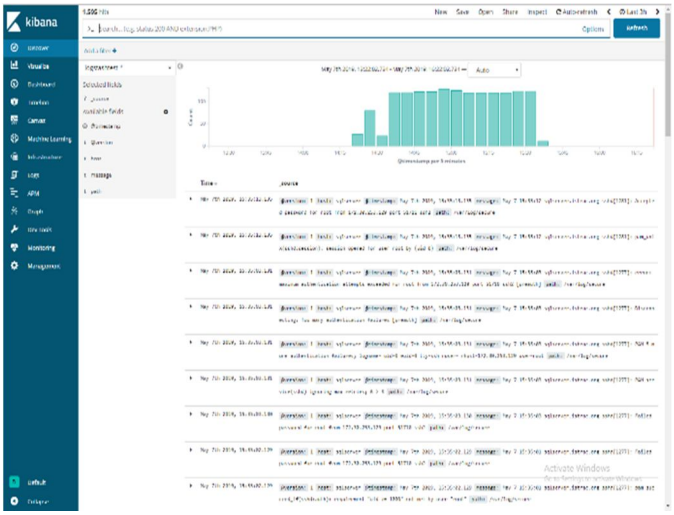


Fig. 2 Index pattern display on kibana

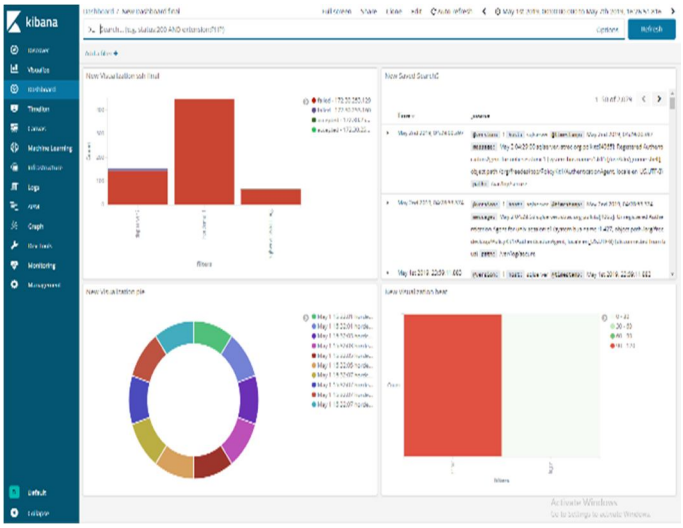


Fig. 3 Kibana Dashboard

VI. MERITS OF ELK STACK

A. Both analysis and monitoring using visualization can be done with the same tool
B. Self-customization and alerts depending on severity
C. Machine Learning feature inbuild
D. Can review on daily basis and store for future use.
E. Doesn't use much space while installation even though it is three different tools combined together.

VII. DEMERITS OF ELK STACK

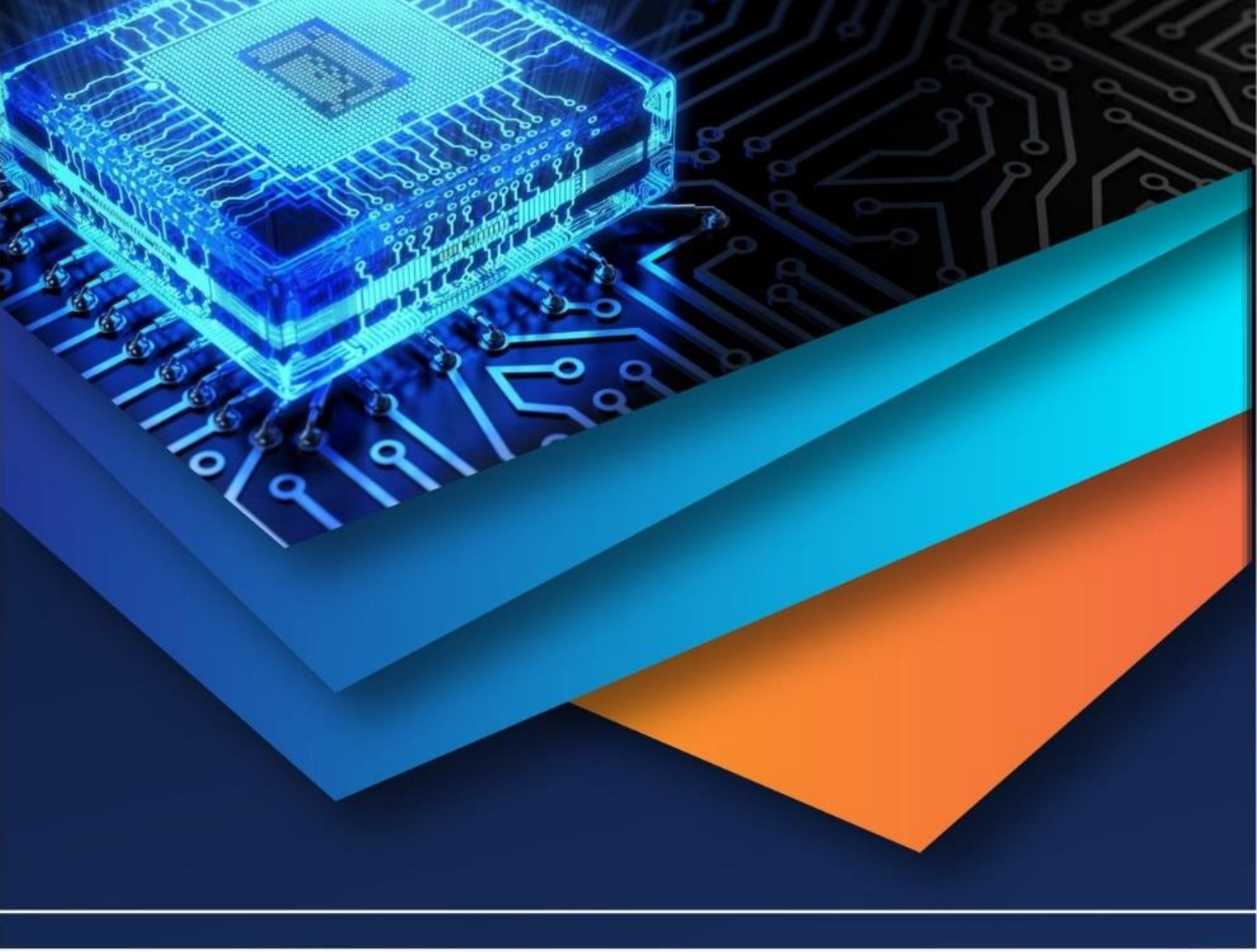
A. Installation may be timing consuming, with lots of extra packages to be installed
B. Indentation in config files are to be carefully done, any monitor mistake doesn't create a index
C. Firewall and network permission must be set well defined or else the kibana dashboard doesn't work.
D. Update if any should be done manually
E. Maintenance of three different modules is actually difficult, if any error occurs to find out which module has caused it might be quite tedious.

VIII. CONCLUSION

Logging can be an aid in fighting errors and debugging programs instead of using a print statement. Logs are beings interacted with each other using rsyslog. The visualisation of data is a necessary step in situations where a huge amount of data is generated every single moment. Data-Visualization tools ELK Stack which is used here and other techniques offer executives and other new approaches to dramatically improve their ability to grasp information hiding in their data from logs got as input. The alerts created are customized for the errors. Rapid identification of error logs, easy comprehension of data and highly customisable data visuals are some of the advantages which helps in preventing possible attacks.

REFERNCES

[1] Rafath Samrin, D Vasumathi , "Review on Anomaly based Network Intrusion Detection System" , 2017 International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEECCOT) pp 141-147
[2] Mohit Tiwari, Raj Kumar, Akash Bharti, Jai Kishan, "Intrusion Detection System", International Journal of Technical Research and Applications e-ISSN: 2320-8163, Volume 5, Issue 2 (March - April 2017), PP. 38-44
[3] Resmi AM, Dr. R Manicka chezian , "Intrusion Detection System Techniques and Tools: A Survey", Sch. J. Eng. Tech., Mar 2017; 5(3): pp 122-130
[4] Zhijian Wang, Yanqin Zhu, "A Centralized HIDS Framework for Private Cloud", 978-1-5090-5504-3/17/$31.00 ©2017 IEEE, pp 115- 120
[5] Arsalan Ali Shaikh, Heng Qi, Wei Jiang, Muhammad Tahir, "A Novel HIDS and Log Collection Based System for Digital Forensics in Cloud Environment" 2017 3rd IEEE International Conference on Computer and Communications, pp 1434- 1438
[6] Marcello Cinque, Domenico Cotroneo, Antonio Pecchia , "Challenges and Directions in Security Information and Event Management (SIEM)", 2018 IEEE International Symposium on Software Reliability Engineering Workshops
[7] Hassani Mohamed, Lebbat Adil , Tallal Saida "A Collaborative Intrusion Detection and Prevention System in Cloud Computing", 2018
[8] https://kibana.logstash.es/content
[9] https://www.elastic.co/products/logstash



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY