Splunk Dashboard: An Application Activities Presenter and Statistical Analyzer

Sukanta Sinha, PhD
Wipro Ltd
Brisbane, Queensland
Australia 4000

Debajyoti Mukhopadhyay, PhD
Bennett University
Greater Noida, Uttar Pradesh
India 201310

ABSTRACT

In the present era, application security has been given immense priority by the software industries; particularly those who are dealing with Credit or Debit card numbers. Application security has been distributed across every level,like firewall rules has configured to communicate between two servers, Card Data Environment has been created to maintain the application servers, Lightweight Directory Access ProtocolActive Directory used to access application by the authorized users.In this paper, we have presented Splunk, which give us time driven activities and application information without login to the application. Splunk is a very good web analytics and helped business to identify their statistics in any sector.

General Terms

Information Retrieval, Security

Keywords

Application Configuration, Splunk, Security.

1. INTRODUCTION

Splunk [1] is a tool which gives an overview of an application health, availability status, performance status, and generate alerts for the application support people based on some predefined conditions. It is a kind of log aggregator that allow users to analyze their machine data. Splunk allow to configure an aggregated dashboard based on the multiple machine logs. The dashboard helps non-technical users or business users to understand the application real time performance.

2. WHY SPLUNK?

In the year of 2012, Splunk [2] founders were rooting around in the logs of computers trying to understand why a website had crashed and getting data from different sources. They likened that to ferreting around in a cave so the name came from speleology in America it is called spelunking and then shortened that to Splunk. Splunk [3] is a horizontal technology used for application management, security and compliance, as well as business and web analytics. Splunk captures, indexes, and correlates real-time data in a searchable repository from which it can generate graphs, reports, alerts, dashboards, and visualizations. Splunk platform harnesses your machine data, which contains a definitive record of all user transactions, customer behavior, machine behavior, security threats, system health, fraudulent activity and more to provide operational intelligence. Splunk software helps you make business sense of this machine data [4],[5].

3. HOW TO CONFIGURE SPLUNK

To configure Splunk[6] dashboard for an application you need to make sure all the required parameters are in your hand. Those are briefly described below.

3.1 Splunk Forwarder

The universal forwarder collects data from a data source or another forwarder and sends it to a forwarder or a Splunk deployment. With a universal forwarder, you can send data to Splunk Enterprise, Splunk Light, or Splunk Cloud. It also replaces the Splunk Enterprise light forwarder. The universal forwarder is available as a separate installation package. The universal forwarder offers advantages over using a heavy or light forwarder. The most notable benefit is that it uses significantly fewer hardware resources than other Splunk software products. It can, for example, coexist on a host that runs a Splunk Enterprise instance. It also is more scalable than the other Splunk products, as you can install thousands of universal forwarders with little impact on network and host performance.

3.2 Application Host Name

Developer has to define the application host names with their fully qualified domain name for deploying the Splunk agents to redirect the application logs to Splunk server.

3.3 Application Description

It will be nice to have a good application description for the future references.

3.4 Application Host OS

Splunk doesn't care about the Operating System (OS) of application host machine. It supports any kind of operating system; like Windows, Linux, and Mainframe etc.

3.5 Application Host Zone

Application host zone depends on the organization level. For banking domain, if you are handling customer personal data, like card number, address, phone number, email, date of birth, etc. then you must keep them in securely and you should not allow to access those data to the outside world. To achieve that, people came up with zone concept like card data zone, developer zone, tester zone, etc. While configuring the Splunk, you must be define your application zone and based on that you need to burn the firewall for Splunk agents to send the logs from application server to Splunk server.

3.6 Application Host Region

Every application in the IT organizations, must have different host regions to build, test, and run. One of the most conspicuous standard of the regions are Development (Dev), Integration (INT), System Test (ST), User Acceptance Test (UAT), Production (Prod) and Digester Recovery (DR). The better approach is to segregate those regions into two parts; Production and Non-Production. Normally, people consider Prod and DR regions under Prod. Dev, INT, ST and UAT regions under Non-Prod region. And it will be better to apply two Splunk indexes one for Prod and other one for Non-Prod.

It will help to reduce the Splunk cost, as they normally charge based on number of indexes. Also we can easily segregate the host servers based on the host name.

3.7 Splunk Index Name

All incoming data must be assigned to an index for storage. Indexes [7] can be used a mechanism to group related data, similar as a file on a file-system or a table in a database. The number of indexes used is often dictated by the environment region and/or underlying required permission scheme.

3.8 Application Log Source Path

Application log source path [8] is one of the most important field to configure Splunk. It actually provides all raw data sources to the Splunk.

3.9 Application Source Type

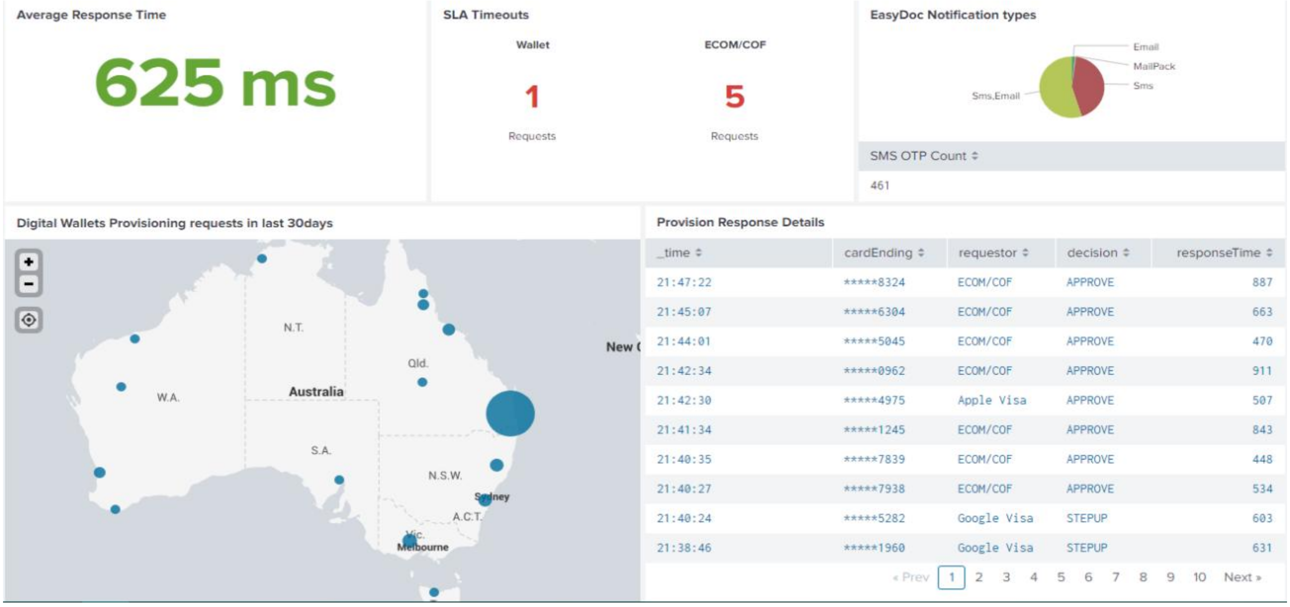
The source type [9] is one of the default fields that Splunk Enterprise assigns to all incoming data. It tells Splunk Enterprise what kind of data you've got, so that it can format the data intelligently during indexing. And it's a way to categorize your data, so that you can search it easily.

3.10 Application Source File

Source file should be the application log file. While configuring the Splunk you need to set all the log files.

3.11 Volume of Data (Optional)

It would be better, if we estimate the data volume beforehand, then we can manage the huge volume of raw log data in an efficient way.



Average Response Time

625 ms

SLA Timeouts

Wallet

1

Requests

ECOM/COF

5

Requests

EasyDoc Notification types

SMS OTP Count ⇕
461

Digital Wallets Provisioning requests in last 30days

Provision Response Details

_time ⇕	cardEnding ⇕	requestor ⇕	decision ⇕	responseTime ⇕
21:47:22	*****8324	ECOM/COF	APPROVE	887
21:45:07	*****6304	ECOM/COF	APPROVE	663
21:44:01	*****5045	ECOM/COF	APPROVE	470
21:42:34	*****0962	ECOM/COF	APPROVE	911
21:42:30	*****4975	Apple Visa	APPROVE	507
21:41:34	*****1245	ECOM/COF	APPROVE	843
21:40:35	*****7839	ECOM/COF	APPROVE	448
21:40:27	*****7938	ECOM/COF	APPROVE	534
21:40:24	*****5282	Google Visa	STEPUP	603
21:38:46	*****1960	Google Visa	STEPUP	631

« Prev 1 2 3 4 5 6 7 8 9 10 Next »

Fig 1: Sample Splunk Dashboard Snippet

4. SPLUNK DASHBOARD

Dashboards [10] are views that are made up of panels. The panels can contain modules such as search boxes, fields, charts, tables, and lists. Dashboard panels are usually connected to reports. After you create a search visualization or save a report, you can add it to a new or existing dashboard. In the dashboard you can mask or encrypt the sensitive data for the viewers. For example, in banking domain, user personal information, bank details, card data must be encrypted or masked before it displayed in the dashboard. The better approach is to mask those data while writing them in the application log, but incase if it not possible then mask those data while transferring to Splunk server. Using a simple regular expression you can mask the sensitive data. A sample Splunk dashboard snippet displayed in Figure.1.

5. SPLUNK ALERT

Splunk alerts are another important feature where predefined users were notified based on some predefined conditions. Alerts can be setup in two types of mode, i.e. email or direct call. Using these alerts application support people take necessary actions at their earliest. Some alerts, like high CPU usages, failure beyond a threshold value, application not running, etc. are applicable for any application and used everywhere.

6. SPLUNK ADVANTAGES

There are several advantages to use Splunk. Some of them are given below:

- ✓ Splunk supports heterogeneous operating systems, i.e., Linux, Windows, and Mainframe etc.
- ✓ Bring everything into one place
- ✓ Monitor application health without login into the application
- ✓ Protect application from the security violation by delivering the application log to Splunk server in an automated way
- ✓ Understand the problem across the whole application, not on a specific machine or a specific technology
- ✓ More cost effective problem resolutions-smaller teams and less reliance on resources
- ✓ Splunk provides a pictorial view, i.e., dashboard, to the business without logging into the application
- ✓ Apply Big Data concepts to all that supplemental information that aids in determining the health of your application or enterprise
- ✓ Capable to handle few years data
- ✓ Allow application server to maintain only a short span of log history and long history will be maintained by Splunk server

7. CONCLUSION

In conclusion, the article gives an idea about how to configure and use Splunk in our daily life. In the view of advantages outlined above, we can conclude that despite some drawbacks, like maintaining another server, the benefits of the Splunk in the Information Technology industry are huge and adopting this tool very fast. In future, we have a plan to work with Splunk for statistical analysis of COVID waves and prepare a dashboard for common people, which helps to identify different zones like red, green, orange, etc. and identify the hotspots area and their gathering threshold defined by the government. Overall, Splunk dashboard gives a very useful pictorial view without login to the application and maintain security standard.

8. ACKNOWLEDGMENTS

Our sincere thanks to Wipro Ltd. for providing opportunities to conduct this research in the laboratory on various applications.

9. REFERENCES

[1] Smith, K., 2015. Building Splunk Solutions: Splunk Developer Guide: 1

[2] Smith, K., 2016. Splunk Developer's Guide - Second Edition

[3] Zadrozny,P. and Kodali,R., 2013.Big Data Analytics using Splunk.

[4] Ohlhorst,F., 2013.Big Data Analytics: Turning Big Data into Big Money.

[5] Splunk, 2018. Using Healthcare Machine Data for Operational Intelligence

[6] James, H. B., 2018. Splunk 7.x Quick Start Guide

[7] Splunk Doc, 2015. Aboutindexesandindexers

[8] Splunk Doc, 2014. InputsConf

[9] Splunk Doc, 2013. WhySource Types Matter

[10] Splunk Doc, 2019.Create New Dashboard