CPSC 8985 FA2015

BIG DATA INSIGHTS USING ANALYTICS

Naga Krishna Reddy, Muppidi [redacted]

Sai Kiran, Merugu [redacted]

Pramod chowdary, Khambampati [redacted]

Contents

Introduction

The main objective of this project is to find the data insights from the huge amount of data that is evolving around us day by day. In order to analyze the data we need an architecture that is suitable for all kinds of data that we see in 21st century. We are using SPLUNK architecture for analyzing the data and getting the insights that we need for taking better decisions. SPLUNK is google for datacenters. By using SPLUNK we can generate all kinds of DASHBOARDS,ALERTS,SCHEDULING,PIVOTS and a lot more important things that is very usable for managers to take a better decisions. We use SPL language for manipulating the data.

.

Purpose

1.It provides "better and efficient" service to members.

2. Reduce the workload of Data Analyst.

3. Faster retrieval of machine data and insights.

4.Provide facility for analyzing the data very fast and can manage load balancing.

5.Can create Dashboards, Pivots, alerts, and scheduling.

6.SPL(Search Processing Language) language is used for getting insights.

Product Functions

The functions of the system include the system providing different type of services based on the type of users.

- User accounts with login facilities.
- Can create dashboards for daily reporting easily .
- Easy to deploy and scalable to any size enterprise.

User Classes

The users of the system are

- Administrator
- Members

Assumptions

- The users should have sufficient knowledge of computers & of English language, as the user interface will be provided in English.
- The members should be either the student or faculty of the institute, where this s/w will be implemented.

Hardware Interface

- Processor: Pentium 3.0 GHz or higher
- RAM: 512 Mb or more
- Hard Drive: 10 GB or more

Software Interface

- Operating System : Windows 8
- Language : SPL
- Framework : Splunk Enterprise

Splunk Components

Data Presentation Layer– Search Head(s)

Universal Forwarder

Data Indexing Layer – Indexer(s).

Data Collection Layer - Universal Forwarders, syslog, API, TCP, Scripts, Wire, etc.

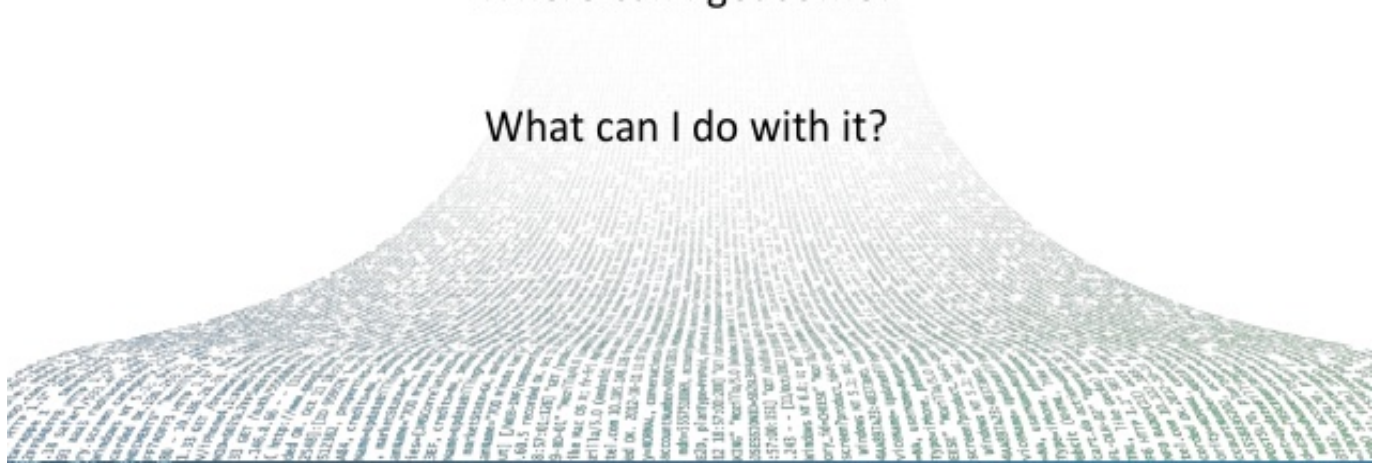
splunk> listen to your data

We'll tackle these questions:

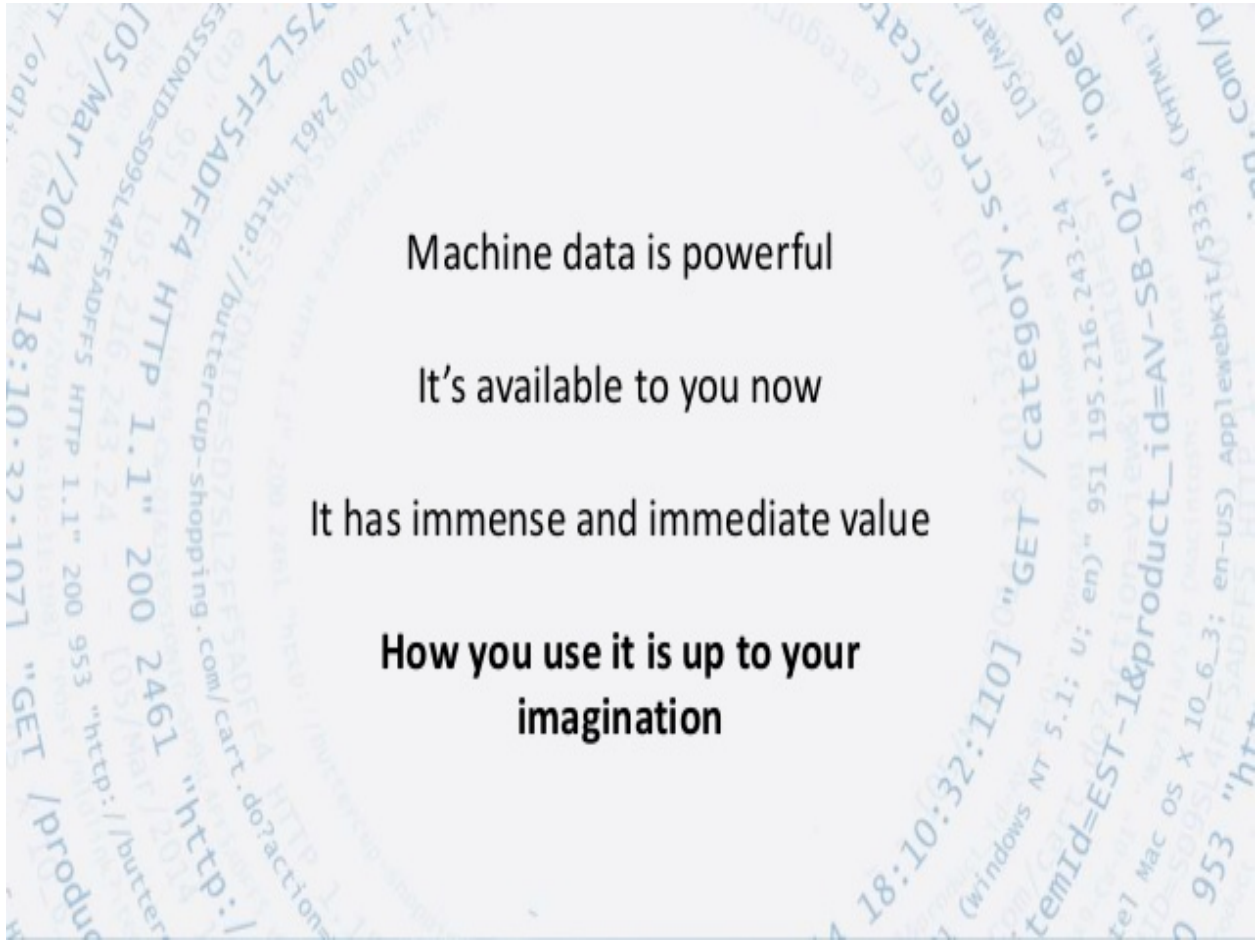
What is machine data?

Where can I get some?

What can I do with it?




What do most agencies use it for?
Improve Constituent Engagement
Mitigate Security Threats
Deliver Better Eligibility Experience
Reduce IT Costs
State & Local Government Use Cases
Ensure HIPPA/PCI/FISMA Compliance
Prevent Fraud
Accelerate Transaction Processing
Improve Operational Efficiency
6


Machine data is powerful
It's available to you now
It has immense and immediate value
How you use it is up to your imagination

What is machine data?

Logs files from computers and network devices

Desktop Laptop Datacenters

Usage and Access Data from Personal electronics

iPad iPhone Android Watches

Data from sensors and remote metering devices

RFID Energy Meters Weather Sensors Radar & Traffic Cameras

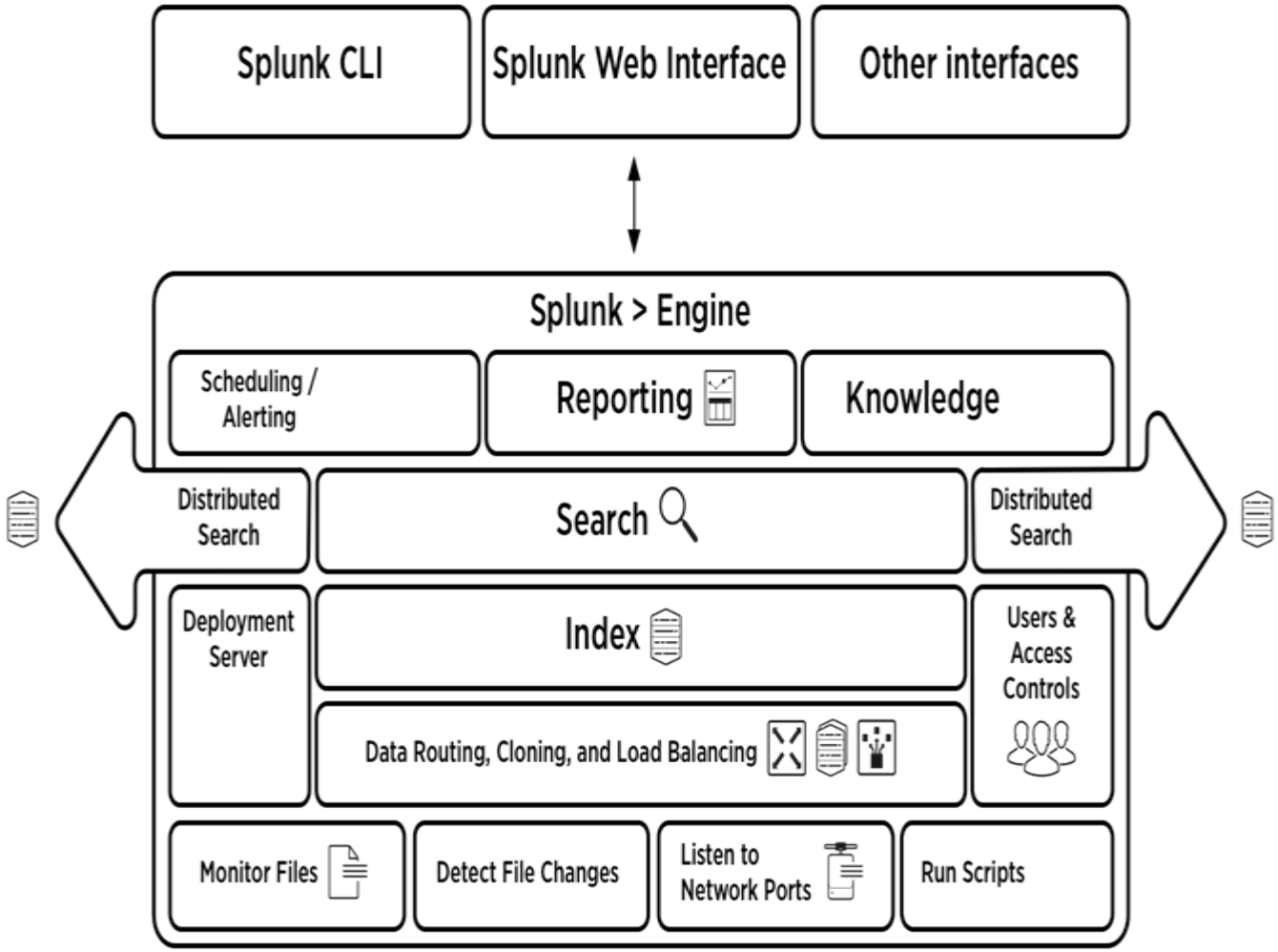
All of these produce data on their own

Lots of data

Valuable data

splunk> listen to your data

Analysis Model



Conclusion

Splunk offers a great platform for Web analytics with large logs andData models with acceleration worked best for us. The business users interest in web analytics increased – More and more colleagues want access to Splunk to solve a problem. Learning to do it on your own is recommendable – Not yet easy to find good and available consultants in USA.

THANK YOU

WELCOME QUERIES