**What is Splunk?**

Splunk is data collection, analysis & Visualization Tool. Companies use this tool to collect and monitor various types of data such as Application metrics, Application logs, Transactional data, Customer data, network data, configuration settings, csv, Security data and more.

Splunk is Monitoring Tool. It monitors the applications running on server. If application not working properly then Splunk provide us error log and gives indexing to error log and convert it into visualization like graph.

**Type of monitoring:**

* System performance
* Process monitoring
* Integration
* Application performance
* Business monitoring
* Splunk is widely used tool.

\*Documentation from the Splexicon: <https://docs.splunk.com/Splexicon>

**Splunk Benefits:**

It helps to monitor the multiple servers are working fine or not in organization. On servers there are lots of application are running, if application running slow or stuck or not working then it sends the message to users. And user must act on it. Or wrong password enters multiple time to login the application running on server then it also generates error log. This monitoring is real time.

1.Real-time performance Monitoring

2.Logging tool

Metrics provides the information that software or application is working or not. And log provides the information why software or application is not working or working slow.

Stack Security and Alerting - it focusses on security. It sends mail or messages to user where we have configured the alert like slack or Email.

* Dashboards and Visualizations
* Data Analysis
* Easy to use.
* Good customer support

**Splunk products:**

1.Enterprise - Enterprise help to generate logs of application running on server.

2.Cloud - It is hosted platform. Same as enterprise. It can be use through cloud like AWS and Splunk

3.Light - It is free version and has some limitations

* Splunk core - collecting tool like Spunk enterprise.
* Splunk IT Operations - IT environment
* Splunk Security - overall security
* Splunk DevOps- software development

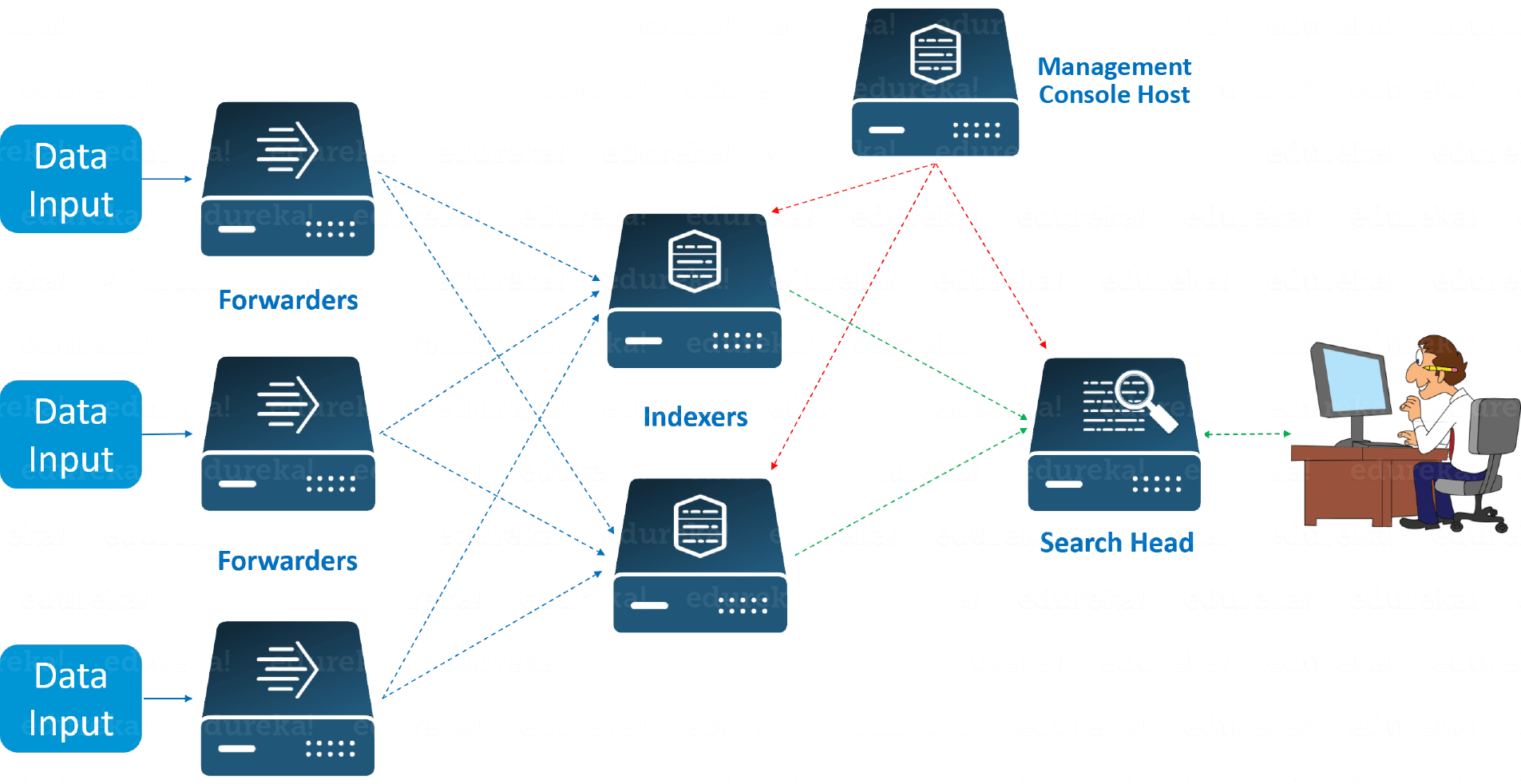
**Splunk Enterprise set of tools:**

* Forwarder - install forwarder on server, then it Collect data from server and forward data to indexer.
* Indexer - process data, Add metadata to Data and store it.
* Search Head - read, analyse, query, visualize data in Search Head.
* Management tools - 1. Deployment Server

2.index cluster master

3.Search Head cluster

4.License Master- for payment method for data storage



**Splunk Enterprise Role:**

* Administrator - admin management, user access, violation management
* Knowledge Manager- knowledge object management
* Search User - read data.
* Pivot User - visualization using pivot table without using query.
* Developer - app and website development, Splunk app development.

**Splunk Server Roles:**

A Splunk instance can play any set of these roles (and a few others) simultaneously, but they are distinctly different things and could be deployed on separate Splunk instances.

**Indexers (Where all the data resides) -**

An indexer is the Splunk instance that indexes data. The indexer transforms the raw data into events and stores the events into an index. The indexer also searches the indexed data in response to search requests. The search peers are indexers that fulfil search requests from the search head.

**Search Head (Used for searches) -**

In a distributed search environment, the search head is the Splunk instance that directs search requests to a set of search peers and merges the results back to the user. If the instance does only search and not indexing, it is usually referred to as a dedicated search head. At Ellie Mae we have a distributed environment with Multi-site clustering provides HA and Scalability to users.

**Forwarders** -

A Splunk instance that forwards data to another Splunk instance is referred to as a forwarder.  These can be further classified as Universal Forwarder and Heavy Forwarder.  Universal Forwarder is different entirely different installer package and only reads and forwards data, whereas a Heavy Forwarder is a full Splunk instance (regular Enterprise Installer)  which is used to read data and forward; it can also index data and then forward but we avoid doing it to lessen the load.  Heavy Forwarder is often used as single point to read data in scenarios where we require to pull data using an API and avoid data duplication across the cluster. Examples - Qualys, Evident.io , ServiceNow etc.

**Deployment Server -**

A deployment server is used to deploy apps to forwarders and technically could be used to deploy apps to other Splunk servers as well but with a number of caveats. All the configuration files are controlled and changed on the deployment Server , which then distributes to forwarders across the site and other Splunk instances.

**Deployer -**

A deployer is used to deploy apps to a search head cluster.

**Cluster Master -**

A cluster master is used to deploy apps and manage replication within an indexer cluster (single or multi-site)

**License Server -**

Stores the license utilization information with number of dashboards.

**Splunk Apps:**

Apps are downloaded from “Splunkbase.com”

Most apps are free, however there are a few premium apps for which need a license.

These include:

* Splunk App for Enterprise Security
* Splunk App for PCI Compliance
* Splunk App for VMware
* Splunk App for Microsoft Exchange

**Splunk consume data from:**

* Virtual machine
* Servers
* IoT
* Communications
* Logs
* Configurations
* Scripts
* Tickets
* Alerts

**Process for data in Splunk:**

* Upload files
* Monitor files and directories - Local and remote.
* Syslog - UDP or TCP, Local and remote
* SNMP (port udp:162)
* Scripted inputs from APIs

**Forwarding and Receiving:**

Types forwarder:

* Universal Forwarder

1. Most Popular
2. Install easily in Windows, Linux, Unix (AIX and HP-UX), Solaris and FreeBSD
3. Installed at local machine and can be configured using deployment server.
4. Default ports:8089 for management and 9997 for indexing

* Heavy Forwarder

1. A complete installation of Splunk software, but with a forwarder licence applied.
2. Does much of the ‘heavy lifting’ at the source.
3. Can be configured at the source, and through a deployment server.

* Configure Receiving

1. Forwarding data to Splunk indexer or search head won’t work unless you configure that indexer or search head to receive data.
2. Setting **>** forwarding and receiving **>** Add New

Virtual Machine > Splunk web > check ip of machine from cmd “ip config”> setting > forwarder and receiver > receiver > add new > 9997 port >save.

Another Virtual machine > Splunk universal forwarder for windows> 64 bit > download and save file > run installer > check all boxes of windows event log, Performance Monitor and Active directory monitoring > enter ip add Splunk for forwarder and receiver > install.

Search and Reporting:

Splunk > Search and Reporting > Data Summary > check Host for Machine Name

Check configuration:

C:/programfiles/SplunkUniversalForwarder/etc/system/local/output.conf

**Why Heavy Forwarders?**

* Index and parse data locally
* Load balance
* Specific data routing rules

**Installation**

* The same installation use for an indexer or search head
* Apply a forwarding license.

Setting > licensing > Change license group

* Configuration

Setting > Forwarding and Receiving > Add new (Forward data) > Enter new hostname:port number, or IP:port number And Separate multiple with commas.

For save report locally:

Setting > Forwarding and Receiving > Forwarding defaults > yes > Save

**Difference between Universal and Heavy forwarder:**

Universal:

* Light agent
* Event parsing available in some cases.
* No event routing.

Heavy:

* Full Splunk Enterprise instance
* Event parsing available.
* Event routing available

**Splunk enterprise**

After downloading Splunk Enterprise:

Setting > license > change license group > forward license change >save

Configure Forwarder for Splunk enterprise:

Setting forwarder and receiver > configure forwarder > add new > add host machine or ip > save.

Monitoring:

Add data > monitor > browse > file and directory > select source > directory select to monitor > host > host field name > review > submit > start forwarder.

Search Pipeline:

* Broad search - Host: myhost, Source type: csv
* Keywords/Booleans/ fields - Fail OR failure, Locked, User: b123
* Commands - Count, Sum, Eval
* Table/vitz - Table, Timechart

E.g.,



Basic Searching:

* Basic search terms (building blocks)
* Keywords - failed, error
* Wildcards - \*ailed, fail\*, user=\*
* Phrases - “failed login”
* Fields - key value pairs, user=user1.domain.com
* Booleans - Case sensitive (upper), AND, OR, NOT
* Basic search commands
* Chart / Timechart - return result in tabular output for charting
* Stats - provides statistics
* Rename - renames a specific field
* Eval - calculate an expression
* Sort - sorts results by specified fields
* Dedup - remove duplicates
* Table - builds a table with the specified fields

Diagram

Description automatically generated

Example for search and reporting on Splunk application:

Graphical user interface, application, website

Description automatically generated

Dealing with time:

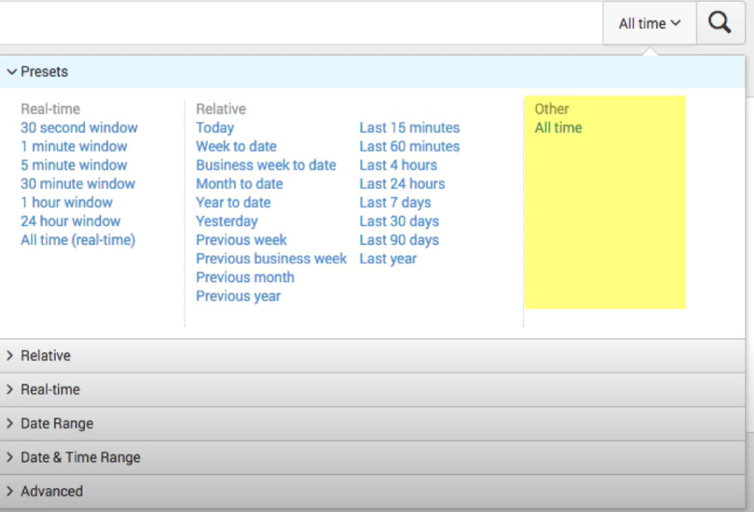
Timestamp -

(n) a default field that represents time information in an event. Most events contain timestamps. In cases where an event does not contain timestamps information, Splunk Enterprise attempts to assign a timestamp value to the event at index time.

The\_time field

* A spunk-generated default field that represents time.
* Timestamps are usually added automatically based on the event raw data.
* If time and date information are not included in the event raw data, Splunk attempts to “guess” at a timestamp.
* As a last resort, Splunk will set the timestamp to the system time.

Splunk uses the timestamp information for the time selector in the search bar.



**Time Conversion:**

* Time can be converted from Splunk’s default to a format of your choice using the strtime() eval function.

| eval time=strftime(\_time, “%H:%M”)

16:34

* Time variable:
* %H = Hour (24-hour clock)
* %I = Hour (12-hour clock)
* %M = Minute
* %S = Second
* %p = AM or PM
* Date Conversion:
* Time variable
* %A = Full day name
* %d = Day of the month (01-31)
* %e = Day of the month without leading zero(1-31)
* %B = Full month name (January)
* %b = Abbreviated month name (Jan)
* %m = Month as a number (01 - 12)
* %y = Four-digit year (2017)
* %Y = Two-digit year (17)

Table

Description automatically generated

Table

Description automatically generated

**Search Modes, Fields, Discovery:**

1.Modes:

* Fast-
* No field discovery, except the default metadata fields.
* Use if you know exactly which fields you need and can specify them in the search string.
* Smart -
* Returns the best results for whatever search you are running.
* Verbose -
* Discovers all fields it can.
* Use if you are not sure what fields you will want to report on

2.Field Discovery:

* During field discovery, Splunk detects key=value pairs
* error=failed
* level=critical
* The first 50 key=value pairs are displayed on the field browser on the left

**Field Extraction tool for Splunk:**

* Field extraction works by using regular expression.
* Splunk comes with built-in tools to help you.

Intermediate Searching Commands:

* Top
* top <field>

E.G., “top <user>”

* returns the most common values of a given field.
* defaults to 10 fields
* Can be combined with limit=<number>
* Automatically builds a table with count and percent columns.
* Can be used with multiple fields.

“return the top value for a field organized by another field”.

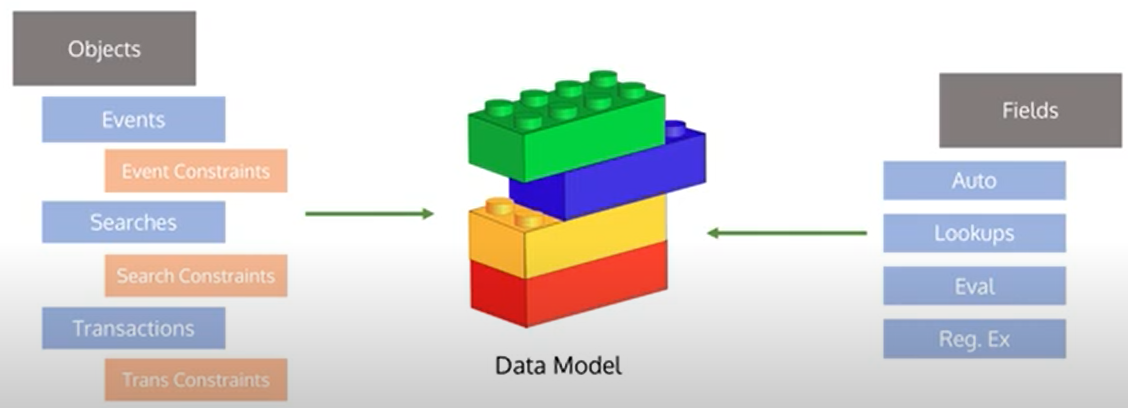
* Rare
* rare <field>
* Opposite of top
* Return the least common values of a field.
* Options are identical to top.
* Stats
* Stats <function(field)> BY <field(s)>
* Some common functions -
* count, avg, max, mean, median, sum, stdev, values, list.

**Data Models:**

* Make machine data easier to use.
* Simplify complex data through abstraction.
* Group specific types of data.

Hierarchically structured data set that includes:

* Objects
* Object Constraints
* Object Attributes



Objects

**Events**: most commonly used

* Event constraints - constrained to the “first part” of a search, i.e. “host=”

**Searches**: Splunk searches that includes transforming commands, etc.

* Search constraints - constrained to the full search string.

**Transactions**: Combine multiple events from one or many sources into single event

* Transaction constraints - must be legally formed transaction search.

Data Models

* Object Attributes (Field)
* Not tied to a specific object
* Auto extracted: Splunk automatically discovered fields.
* Eval: A field generated because of an eval expression.
* Lookup: fields that are the result of a lookup
* Regular expression: Fields extracted by regex.
* Power the Splunk Pivot tool
* Building data models abstracts the complexities of Splunk from the end user so they can “drag and drop” fields into a report.
* To get to the Pivot tool
* Through the datasets page
* Through the data model page in settings
* Through the search results page, visualization tab

**Add Data to Splunk:**

Splunk Web > Setting > Add Data > choose Upload Option > browse > file

**Pivot on Uploaded Data:**

Splunk Web > Settings > Data Models > pivot > Select Dataset > Select Chart type from Left Menu > review the Data for Pivot > Build Pivot > Save Pivot > Title line chart

**The Chart and Timechart Commands:**

* host = Hostname
* usr = User from data
* \*= all
* state = State from data
* chart = for chart
* timechart = chart based on time
* count(usr) BY state = count user from data based on state
* rename = renaming count column
* AS = aliases

1. host=hostname usr=\* state=\* | chart count(usr) BY state | rename count(usr) AS “Aliases name”
2. host=hostname usr=\* state=\* | timechart count(usr) BY state

**Reporting and Alerting:**

* Reports and alerts are knowledge objects in Splunk.
* To create reports and alerts, you need a Splunk Enterprise license
* The free license disables these features
* Saved searches that can run on a schedule and perform an action.
* Send an e-mail to report consumers
* Embed on a web page
* Update a dashboard panel
* Run a script

**Report:**

* Scheduled reports can run
* Every hour
* Every day
* Every week
* Every month
* On a chron schedule that you define
* You can stagger the report running window
* Useful if you have a lot of reports running at the same time

**Alert:**

* Can be scheduled or in real-time
* Triggered when the results of a search meet specific condition that you define
* For example, if the search host=firewall1 user=\* authentication=failed returns anything, trigger an alert
* Alert action can include
* Send an email
* Trigger a script
* Use a webhook
* List in triggered alerts
* Use an app (like PagerDuty or Slack)