**ELK Stack**

**What is the ELK Stack?**

The ELK Stack is a collection of three open-source products i,e Elasticsearch, Logstash, and Kibana.

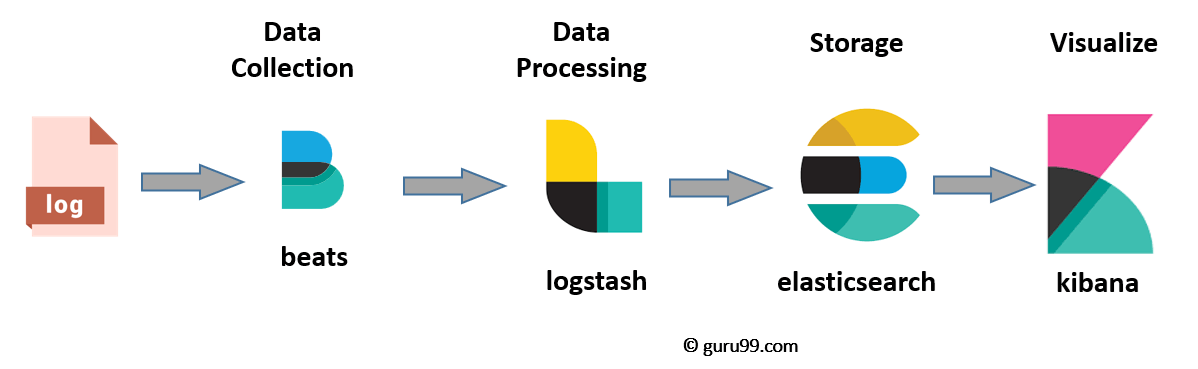
* E - ElasticSearch: used for storing logs.
* L - LogStash : used for both shipping as well as processing and storing logs.
* K - Kibana: is a visutalization tool (a web interface) which is hosted through Nginx or Apache.

ELK Stack is designed to allow users to take to data from any source, in any format, and to search, analyze, and visualize that data in real time.

ELK provides centralized logging that be useful when attempting to identify problems with servers or applications. It allows you to search all your logs in a single place. It also helps to find issues that occur in multiple servers by connecting their logs during a specific time frame.

## ELK Stack Architecture

Here is the simple architecture of ELK stack



* Logs: Server logs that need to be analyzed are identified
* Logstash: Collect logs and events data. It even parses and transforms data
* ElasticSearch: The transformed data from Logstash is Store, Search, and indexed.
* Kibana: Kibana uses Elasticsearch DB to Explore, Visualize, and Share
* Beats:Beats is the platform for single-purpose data shippers. They send data from hundreds or thousands of machines and systems to Logstash or Elasticsearch

## What is Logstash?

Logstash is the data collection pipeline tool. It collects data inputs and feeds into the Elasticsearch. It gathers all types of data from the different source and makes it available for further use.

Logstash can unify data from disparate sources and normalize the data into your desired destinations. It allows you to cleanse and democratize all your data for analytics and visualization of use cases.

It consists of three components:

* **Input**: passing logs to process them into machine understandable format.
* **Filters**: It is a set of conditions to perform a particular action or event
* **Output**: Decision maker for processed event or log

**Features of Logstash**

* Events are passed through each phase using internal queues
* Allows different inputs for your logs
* Filtering/parsing for your logs

**Advantage of Logstash**

* Offers centralize the data processing
* It analyzes a large variety of structured/unstructured data and events
* Offers plugins to connect with various types of input sources and platforms

## What is Elasticsearch?

## Elasticsearch is a NoSQL database. It is based on Lucene search engine, and it is built with RESTful APIS. It offers simple deployment, maximum reliability, and easy management. It is helpful for executing a quick search of the documents. Elasticsearch also allows you to store, search and analyze big volume of data. It is mostly used as the underlying engine to powers applications that completed search requirements.

**Features of Elastic search:**

* Open source search server is written using Java
* Used to index any kind of heterogeneous data
* Has REST API web-interface with JSON output
* Full-Text Search
* Near Real Time (NRT) search
* Sharded, replicated searchable, JSON document store
* Schema-free, REST & JSON based distributed document store
* Multi-language & Geolocation support.

**Advantages of Elasticsearch**

* Store schema-less data and also creates a schema for your data
* Manipulate your data record by record with the help of Multi-document APIs
* Perform filtering and querying your data for insights
* Based on Apache Lucene and provides RESTful API
* Provides horizontal scalability, reliability, and multitenant capability for real time use of indexing to make it faster search
* Helps you to scale vertically and horizontally.

**Important Terms used in Elastic Search**

* **Cluster:** A cluster is a collection of nodes which together holds data and provides joined indexing and search capabilities.
* **Node:** A node is an elasticsearch Instance. It is created when an elasticsearch instance begins.
* **Index:** An index is a collection of documents which has similar characteristics. e.g., customer data, product catalog. It is very useful while performing indexing, search, update, and delete operations. It allows you to define as many indexes in one single cluster.
* **Document:** It is the basic unit of information which can be indexed. It is expressed in JSON (key: value) pair. '{"user": "nullcon"}'. Every single Document is associated with a type and a unique id.
* **Shard:** Every index can be split into several shards to be able to distribute data. The shard is the atomic part of an index, which can be distributed over the cluster if you want to add more nodes.

## What is Kibana?

## Kibana is a data visualization which completes the ELK stack. This tool is used for visualizing the Elasticsearch documents and helps developers to have a quick insight into it. It can be used for search, view, and interact with data stored in Elasticsearch directories. Kibana helps you to perform advanced data analysis and visualize your data in a variety of tables, charts, and maps.

## the most common search types:

## Free text searches: It is used for searching a specific string.

* Field-levelsearches: It is used for searching for a string within a specific field.

## Logical statements: It is used to combine searches into a logical statement.

## Proximity searches: It is used for searching terms within specific character proximity.

**Features of Kibana:**

* Powerful front-end dashboard which is capable of visualizing indexed information from the elastic cluster
* Enables real-time search of indexed information
* You can search, View, and interact with data stored in Elasticsearch
* Execute queries on data & visualize results in charts, tables, and maps
* Configurable dashboard to slice and dice logstash logs in elasticsearch
* Capable of providing historical data in the form of graphs, charts, etc.
* Real-time dashboards which is easily configurable
* Enables real-time search of indexed information

**Advantages and Disadvantages of Kibana:**

* Easy visualizing
* Fully integrated with Elasticsearch
* Offers real-time analysis, charting, summarization, and debugging capabilities
* Provides instinctive and user-friendly interface
* Allows sharing of snapshots of the logs searched
* Permits saving the dashboard and managing multiple dashboards.

## Advantages and Disadvantages of ELK stack

**Advantages**

* ELK works best when logs from various Apps of an enterprise converge into a single ELK instance
* It provides amazing insights for this single instance and also eliminates the need to log into hundred different log data sources
* Rapid on-premise installation
* Easy to deploy Scales vertically and horizontally
* Elastic offers a host of language clients which includes Ruby. Python. PHP, Perl, .NET, Java, and JavaScript, and more
* Availability of libraries for different programming and scripting languages

**Disadvantages**

* Different components In the stack can become difficult to handle when you move on to complex setup
* There's nothing like trial and error. Thus, the more you do, the more you learn along the way