ELK Stack: Elastic Search, Log stash and Kibana

Total Hours: 25 and Days: 5

Hardware Requirement:

1. Windows Machine with minimum 32 GB RAM
2. All the systems(Laptop) should be in the intranet
3. JDK17 should be installed and JAVA\_HOME path should be set.
4. Eclipse JEE Latest or Intellij Ultimate

Prerequisite:

1. Knowledge of Python OR Java
2. Understanding of SQL queries

Understanding Elastic eco system for log collection, aggregation and visualization (ELK stack)

Making your content searchable is not about exposing content to Google for indexing; it's about providing your own tailor made search, within your system to your users. Large players like Twitter and LinkedIn have been using Lucene for years, and today is the day you will see how easy it is to join their club by learning Elasticsearch from the ground up.

Table of Contents

Day 1

* Introduction
  + Elastic Stack Overview
  + Logs and Problems associated with it
* Role of Logstash for performing ETL
  + Logstash Installation and Configuration
  + Understanding and using Inputs, Filters, and Outputs
  + How to apply Conditionals in filters

Elasticsearch

* + Understanding Elastricsearch NoSQL architecture
  + CAP theorm for elasticserch
  + Comparing and contrasting Solr, Lucene and Elasticseach

JSON Structure

* + Nested JSON
  + Different Data Types
  + Understanding Elastricsearch NoSQL architecture
  + CAP theorm for elasticserch
  + Comparing and contrasting Solr, Lucene and Elasticseach
  + Understanding the Elasticsearch Architecture and different types of Nodes
  + Different nodes and its uses
    - Master node
    - Data nodes
    - Ingest node
    - Tribe node
  + Understanding the Elasticsearch Architecture and different types of Nodes
  + Inside the shards , node and performing cluster installation
  + Index Management :
* Multi Tenancy – Multiple indices, index aliases, cross index operations and introduction to data flow

Day 2

* Elasticsearch Index – Analysis, mappings, index operations, versioning, APIs and settings
* Create  
  Read  
  Update  
  Delete  
  Data Modelling  
  Mapping  
  Aliases
  + Performing different kinds of search : term, terms, match,multi\_match, fuzzy,regexp and wildcard
  + Query DSL  
    Full Text Searches  
    Structured Searches  
    Proximity  
    Partial Matching(wild card)

**Analysis (Advance Topic)**

Analyzers

Number of queries comes to each node / API (query)  
Tokenizers  
Token Filters  
Char Filters

Day 3

Aggregations : Performing different kinds of aggregation such as

• percentile\_ranks

• cardinality

• significant\_terms

• top hits

• scripted metric

• filters

• range

• geohash

• terms

• histogram

• date\_histogram

• stats

• extended stats

• min / max

• sum

Suggestions (Advance Topic) :

• Term suggestion

• Phrase Suggestion

• Completion Suggestion

• Context Suggestion

Highlighting : Showing snippets of text matching the query

• highlighter

• FVH: fast-vector-highlighter

• Postings

* + Writing JAVA programme for performing search and aggregation
  + Understand relevancy (Advance Topic):

Boosting

Custom Scoring

• Boost

• Random

• Decay Function

• Script

• Field Value Factor

• Weight

* + Index Template
  + Index Alias
  + Re-indexing from one index into another
  + Plugin management
  + Elasticseach backup and restore mechanism

Day 4

# [How To](https://www.elastic.co/guide/en/elasticsearch/reference/current/how-to.html) Improve performances

* + [General recommendations](https://www.elastic.co/guide/en/elasticsearch/reference/current/general-recommendations.html)
  + [Recipes](https://www.elastic.co/guide/en/elasticsearch/reference/current/recipes.html)
  + [Mixing exact search with stemming](https://www.elastic.co/guide/en/elasticsearch/reference/current/mixing-exact-search-with-stemming.html)
  + [Getting consistent scoring](https://www.elastic.co/guide/en/elasticsearch/reference/current/consistent-scoring.html)
  + [Tune for indexing speed](https://www.elastic.co/guide/en/elasticsearch/reference/current/tune-for-indexing-speed.html)
  + [Tune for search speed](https://www.elastic.co/guide/en/elasticsearch/reference/current/tune-for-search-speed.html)
  + [Tune for disk usage](https://www.elastic.co/guide/en/elasticsearch/reference/current/tune-for-disk-usage.html)

# Monitoring clusters in Elasticsearch

[Cluster APIs](https://www.elastic.co/guide/en/elasticsearch/reference/current/cluster.html)

* + [Cluster Health](https://www.elastic.co/guide/en/elasticsearch/reference/current/cluster-health.html)
  + [Cluster State](https://www.elastic.co/guide/en/elasticsearch/reference/current/cluster-state.html)
  + [Cluster Stats](https://www.elastic.co/guide/en/elasticsearch/reference/current/cluster-stats.html)
  + [Pending cluster tasks](https://www.elastic.co/guide/en/elasticsearch/reference/current/cluster-pending.html)
  + [Cluster Reroute](https://www.elastic.co/guide/en/elasticsearch/reference/current/cluster-reroute.html)
  + [Cluster Update Settings](https://www.elastic.co/guide/en/elasticsearch/reference/current/cluster-update-settings.html)
  + [Cluster Get Settings](https://www.elastic.co/guide/en/elasticsearch/reference/current/cluster-get-settings.html)
  + [Nodes Stats](https://www.elastic.co/guide/en/elasticsearch/reference/current/cluster-nodes-stats.html)
  + [Nodes Info](https://www.elastic.co/guide/en/elasticsearch/reference/current/cluster-nodes-info.html)
  + [Nodes Feature Usage](https://www.elastic.co/guide/en/elasticsearch/reference/current/cluster-nodes-usage.html)
  + [Remote Cluster Info](https://www.elastic.co/guide/en/elasticsearch/reference/current/cluster-remote-info.html)
  + [Task Management API](https://www.elastic.co/guide/en/elasticsearch/reference/current/tasks.html)
  + [Nodes hot\_threads](https://www.elastic.co/guide/en/elasticsearch/reference/current/cluster-nodes-hot-threads.html)
  + [Cluster Allocation Explain API](https://www.elastic.co/guide/en/elasticsearch/reference/current/cluster-allocation-explain.html)

Day 5

Logstash :

* + [Installing Logstash](https://www.elastic.co/guide/en/logstash/5.4/installing-logstash.html)
  + [Stashing Your First Event](https://www.elastic.co/guide/en/logstash/5.4/first-event.html)
* [Structure of a Config File](https://www.elastic.co/guide/en/logstash/5.4/configuration-file-structure.html)
* [Accessing Event Data and Fields in the Configuration](https://www.elastic.co/guide/en/logstash/5.4/event-dependent-configuration.html)
* [Using Environment Variables in the Configuration](https://www.elastic.co/guide/en/logstash/5.4/environment-variables.html)
* [Logstash Configuration Examples](https://www.elastic.co/guide/en/logstash/5.4/config-examples.html)
* [Reloading the Config File](https://www.elastic.co/guide/en/logstash/5.4/reloading-config.html)
* [Managing Multiline Events](https://www.elastic.co/guide/en/logstash/5.4/multiline.html)
  + [Glob Pattern Support](https://www.elastic.co/guide/en/logstash/5.4/glob-support.html)
  + [Advanced Pipeline](https://www.elastic.co/guide/en/logstash/5.4/advanced-pipeline.html)
  + [Multiple Output Plugins](https://www.elastic.co/guide/en/logstash/5.4/multiple-input-output-plugins.html)

FileBeat (Advance Topic)

* [Installing Filebeat](https://www.elastic.co/guide/en/beats/filebeat/current/filebeat-installation.html)
* [Configuring Filebeat](https://www.elastic.co/guide/en/beats/filebeat/current/filebeat-configuration.html)
* [Configuring Filebeat to Use Logstash](https://www.elastic.co/guide/en/beats/filebeat/current/config-filebeat-logstash.html)
* [Loading the Index Template in Elasticsearch](https://www.elastic.co/guide/en/beats/filebeat/current/filebeat-template.html)
* [Starting Filebeat](https://www.elastic.co/guide/en/beats/filebeat/current/filebeat-starting.html)
* [Loading the Kibana Index Pattern](https://www.elastic.co/guide/en/beats/filebeat/current/filebeat-index-pattern.html)

Load Balancing File Beats

Kibana

* + Using Kibana to visualize the ES data
  + Settings
  + Time Picker, Search, and Filters
  + Kibana Discover, Visualization, and Dashboard Interfaces
  + Lab
  + Handling Back Pressure in ELK stack
  + Deployment Architectures
  + Hardware Best Practices
  + Operation Checklist
  + Debugging and Monitoring
  + APM
  + Stack Monitoring
  + Saved Objects - importing / exporting
  + Canvas
  + Security Setting - roles and users
  + REST APIs
  + Spaces
  + 3rd Party plugin

# [How To](https://www.elastic.co/guide/en/elasticsearch/reference/current/how-to.html) Improve performances

* + [General recommendations](https://www.elastic.co/guide/en/elasticsearch/reference/current/general-recommendations.html)
  + [Recipes](https://www.elastic.co/guide/en/elasticsearch/reference/current/recipes.html)
  + [Mixing exact search with stemming](https://www.elastic.co/guide/en/elasticsearch/reference/current/mixing-exact-search-with-stemming.html)
  + [Getting consistent scoring](https://www.elastic.co/guide/en/elasticsearch/reference/current/consistent-scoring.html)
  + [Tune for indexing speed](https://www.elastic.co/guide/en/elasticsearch/reference/current/tune-for-indexing-speed.html)
  + [Tune for search speed](https://www.elastic.co/guide/en/elasticsearch/reference/current/tune-for-search-speed.html)
  + [Tune for disk usage](https://www.elastic.co/guide/en/elasticsearch/reference/current/tune-for-disk-usage.html)

Additional Topics

Batch Query Configurations

### APM

* What is the Purpose of APM?
* Key Components in APM
* Using Docker Containers as APM Nodes
* Install APM
* Significance of APM Programs
* Configure APM
* Understanding APM System Paths
* Manage APM

### Metricbeat

* What is the Purpose of Metricbeat?
* Install Metricbeat
* Significance of Metricbeat Programs
* Configure Metricbeat
* Understanding Metricbeat System Paths
* Manage Metricbeat

### Heartbeat

* What is the Purpose of Heartbeat?
* Install Heartbeat
* Significance of Heartbeat Programs
* Configure Heartbeat
* Understanding Heartbeat System Paths
* Manage Heartbeat

Observability

* Collecting logs and metrics.
* Collecting APM data.
* Working with observability data.
* Structuring and processing data.
* Actionable observability data.
* Visualizing observability data.
* Managing observability data.