Elasticsearch, Kibana and Logstash

40 academic hours

Course Description

This ELK training intends to provide a solid foundation in search and information retrieval. It starts with fundamental concepts and follows with internals, best-practices and key features. Each topic is followed by a hands-on lab. At the end of the training, attendee will understand how Elasticsearch works, how to load data into Elasticsearch using Logstash, will be able to reliably analyze data and will be ready to build search applications and present visualization using Kibana

Target Population

- Architects
- DBAs
- BI developers and analysts

Pre-requisites

Basic knowledge of database concepts and data presentation tools

Course Objectives

Upon completion of the course, participants will be able to:

- Describe and design an ELK environment
- Create Elasticsearch cluster

- Control and monitor the Elasticsearch cluster
- Use Kibana for visualizations

Course Topics

Module 1 - Introduction to Elasticsearch

- The Story of Elasticsearch
- Documents
- Indexes
- Indexing Data
- Searching Data
- The Bulk API

Module 2 - Logstash

- Introduction to Logstash
- Writing Logstash configuration
- Loading data from relational database to Elasticsearch

Module 3 - The Search API

- Introduction to the Search API
- URI Searches
- Request Body Searches
- The match Query
- The match phrase Query
- The range Query
- The bool Query
- Source Filtering

Module 4 - Text Analysis

- What is Analysis?
- Building an Inverted Index
- Analyzers
- Custom Analyzers
- Character Filters
- Tokenizers
- Token Filters
- Defining Analyzers

- Synonyms
- How to Choose an Analyzer
- Segments

Module 5 - Mappings

- What is a Mapping?
- Dynamic Mappings
- Defining Explicit Mappings
- Adding Fields
- Dive Deeper into Mappings
- Specifying Analyzers
- Dynamic Templates
- Index Templates

Module 6 - More Search Features

- Filters
- Term Filters
- The match_phrase_prefix Query
- The multi_match_Query
- Fuzziness
- Highlighting
- More Like This

Module 7 - The Distributed Model

- Starting a Node
- Creating an Index
- Starting a Second Node
- Shards: Distribution of an Index
- Distributing Documents
- Replication
- Split Brain
- Other Node Types
- Development vs. Production Mode

Module 8 - Working with Search Results

- The Anatomy of a Search
- Relevance
- Boosting Relevance
- DFS Query-then-fetch

- Sorting Results
- Doc Values and Fielddata
- Pagination
- Scroll Searches
- Choosing a Search Type

Module 9 - Aggregations

- What are Aggregations
- Types of Aggregations
- Buckets and Metrics
- Common Metrics Aggregations
- The range Aggregation
- The date_range Aggregation
- The terms Aggregation
- Nesting Buckets

Module 10 - More Aggregations

- Global Aggregation
- The missing Aggregation
- Histograms
- Date Histograms
- Percentiles
- Top Hits
- Significant Terms
- Sorting Buckets

Module 11 – Handling Relationships

- The Need for Data Modeling
- Denormalization
- The Need for Nested Types
- Nested Types
- Querying a Nested Type
- Sorting on a Nested Type
- The Nested Aggregation
- Parent/Child Types
- The has_child Query
- The has_parent Query

Module 12 - Kibana – Discover Data

- Introduction to Kibana and environment setup
- Searching Your Data
- Setting time filter
- Filter by field
- Viewing document data
- Kibana – Visualize Data
- Creating a Visualization
- Available Chart Types

Building a Dashboard