Core Elasticsearch Outline

Overview

This training will guide you into an in-depth, instructor-led training course with case study discussions held by Elasticsearch developers.

Objectives

The course aims to provide a solid foundation in search and information retrieval. Starting with fundamental concepts and covers best practices, key features, and distributed search application development with Elasticsearch. During the course there will be time for discussion as well as attendee case studies. At the end of the training you will have an in-depth understanding of how Elasticsearch works, you will be able to reliably analyze, understand, and solve common problems, and be ready to build state-of-the-art search applications.

Audience

Developers who would like to build real-time search solutions and analytics solutions.

Duration

2 Days Class is scheduled from 9 a.m. to 5 p.m.

Pre-requisites

Exposure to or interest in Elasticsearch, relational databases, distributed systems, or information retrieval.

Requirements

None.

Core Elasticsearch Outline

Introduction

Terminology, basic concepts, implementation, setup, and basic operations.
What is Elasticsearch?
Overview of best practices
What's in a distribution?
Understanding Elasticsearch cluster, shards, and replicas
Discussion of configuration, APIs, and local gateway

Multi-Tenancy

Value of multiple indices, index aliases, and cross-index operations
Introduction to data flow

Elasticsearch Index

In-depth analysis of mappings, indexing, and operations
Discussion of transaction logs and Lucene indexing
Understanding configuration options, mappings, APIs, and available settings

Search

Understanding search Query DSL
In-depth understanding of search components: aggregations, search types, highlighting and other options.
Overview of bitSets, filters and Lucene

Advanced Search and Mapping

Introduction to aggregations and nested document relations
Understanding nested objects and parent-child relationships
The importance of geolocation, mapping, indexing query percolation, relevancy, searching, and more

Advanced Distributed Model

Cluster state recovery, low level replication, low-level recovery, and shard allocation
How to approach data architecture
Index templates, features, and functionality

Big Data Design Pattern

In-depth content on multiple indices, overallocation, shard overallocation, node types, routing, replication, and aliases

Preparing for Production

Discussion on capacity planning and data flow
Performance tuning, more on data flow, and memory allocation.

Running in Production

Installation, configuration, memory file descriptions, and hardware
Monitoring, alerts, thread pools, information and stats API.