EduGroupe
Accompagner pour réussir

ElasticStack : indexation de contenu

Référence : **PYCB021** Durée : **2 jours (14 heures)** Certification : **Aucune**

Connaissances préalables

- 1- Connaissances générales des systèmes d'information, et des systèmes d'exploitation (Linux ou Windows)
- 2- Les travaux pratiques sont réalisés sur Linux
- 3- Connaissance d'un langage de programmation structuré

Profil des stagiaires

- Architectes techniques, développeurs, analystes

Objectifs

- Comprendre le fonctionnement et les apports d'Elasticsearch dans le traitement de données, et savoir le mettre en oeuvre pour l'analyse de données

Certification préparée

- Aucune

Méthodes pédagogiques

- La formation est constituée d’apports théoriques, d’exercices pratiques et de réflexions
- Remise d’une documentation pédagogique papier ou numérique pendant le stage
- Mise à disposition d’un poste de travail par participant

Formateur·rice

- Consultant-Formateur expert Bigdata

Méthodes d'évaluation des acquis

- Auto-évaluation des acquis par le stagiaire via un questionnaire
- Attestation des compétences acquises envoyée au stagiaire
- Attestation de fin de stage adressée avec la facture

Contenu du cours

1. Introduction

- Présentation de la pile elastic
- Positionnement d'Elasticsearch et des produits complémentaires : Watcher, Marvel, Kibana, Logstash, Beats, X-Pack
- Les apports de la version 7.x
- Principe : base technique Lucene et apports d'ElasticSearch. Fonctionnement distribué
- Cas d'usage classiques : analyse de logs et securité, analyse de métriques, recherches web, etc.

2. Installation et configuration

- Prérequis techniques
- Premiers pas dans la console DevTools de Kibana

3. Concepts clés

- Présentation des concepts clés d'ElasticSearch : index, types, documents, noeuds, clusters, shards et replica
- Notions de datatypes et mappings
- Opérations CRUD : exemples d'opérations basiques, création d'index et mappings

4. Format et stockage des données

- Format des données. Conversion au format JSON des données à traiter.
- Structure des données. Stockage, indexation. Terminologie Elasticsearch : notions de document, type, index
- Métadonnées : _index, _ID
- Choix de l'identifiant par l'application avec l'API index, ou genération automatique d'un identifiant
- Indexation inversée

5. Outils d'interrogation

- API RESTful en HTTP
- Exemples de requêtes simples et plus complexes : recherche de «phrases», extraction de plusieurs documents, etc.
- Notion de pertinence du résultat : «score»
- Requêtes avec Search Lite et avec Query DSL (domain-specific language)
- Utilisation de 'filtre' pour affiner des requêtes
- Aggrégation de résultats

6. Gestion des accès concurrents

- Utilisation du numéro de version
- Gestion par l'application : différentes méthodes selon les contraintes fonctionnelles
- Utilisation d'un numéro de version externe

7. Analyse et visualisation de données

- Principes de base de l'analyse de texte
- Recherche dans des données structurées, recherche full text
- Ecriture de requêtes complexes
- Notions d'aggrégations
- Mise en oeuvre : préparation des données, aggregation de mesures, bucket aggregation

8. Flux logstash et présentation Kibana

- Traitement de logs avec logstash
- Introduction à beats, installation et configuration
- Présentation Kibana et démonstrations
- Fonctionnalités : recherche, visualisation, création de tableaux de bord et graphiques à partir des données fournies par Elacticsearch

Notre référent handicap se tient à votre disposition au 01.71.19.70.30 ou par mail à referent.handicap@edugroupe.com pour recueillir vos éventuels besoins d'aménagements, afin de vous offrir la meilleure expérience possible.