

Elasticsearch : indexation de contenu

INFORMATIONS GÉNÉRALES

Type de formation : Formation continue

Éligible au CPF : Non

Domaine : IA, Big Data et Bases de données

Action collective : Non

Filière : Big Data

Rubrique : ElasticSearch,...

Code de formation : BD030

€ Tarifs

Prix public : 1520 €

Tarif & financement :

Nous vous accompagnons pour trouver la meilleure solution de financement parmi les suivantes :

- **Le plan de développement des compétences de votre entreprise** (rapprochez-vous de votre service RH).
- **Le dispositif FNE-Formation.**
- **L'OPCO** (opérateurs de compétences) de votre entreprise.
- **Pôle Emploi** sous réserve de l'acceptation de votre dossier par votre conseiller Pôle Emploi.
- **CPF -MonCompteFormation.**

Contactez nous pour plus d'informations.

PRÉSENTATION

Objectifs & compétences

A l'issue de la formation, le stagiaire sera capable de mettre en oeuvre une solution de recherche performante de données volumineuses avec ElasticSearch.

Situer Elasticsearch dans un écosystème Big Data
Identifier les enjeux et les cas d'utilisation d'un moteur de recherche
Appréhender le fonctionnement d'ElasticSearch
Savoir installer et configurer ElasticSearch
Indexer des volumes importants de données
Comprendre comment administrer le système et le surveiller afin de garantir sa disponibilité

Public visé

Administrateur système, Architectes, Développeurs, Data Analysts et/ou Data Engineers, ...

Pré-requis

Disposer de solides bases en administration système Unix/Linux

📍 Lieux & Horaires

Durée : 14 heures

Délai d'accès :

Jusqu'à 8 jours avant le début de la formation.

PROGRAMME

Situer Elasticsearch dans un écosystème Big Data

Positionnement d'Elasticsearch et des produits complémentaires :Kibana, Logstash, Beats, X-Pack
Présentation des concepts clés d'ElasticSearch

Atelier : comparaison d'ElasticSearch avec d'autres moteurs d'indexation en fonction du cas d'utilisation

Identifier les enjeux et les cas d'utilisation d'un moteur de recherche

Présentation de la pile elastic.
Cas d'usage classiques : analyse de logs et sécurité, analyse de métriques, recherches web, ...
Analyse et visualisation de données
Principes de base de l'analyse de texte, recherche dans des données structurées, recherche full text,

Atelier : choix d'un outil d'indexation en fonction des contraintes d'utilisation

Appréhender le fonctionnement d'ElasticSearch

Principe : base technique Lucene et apports d'ElasticSearch. Fonctionnement distribué
Concepts clés : index, types, documents, noeuds, clusters, shards et replica
Notions de datatypes et mappings
Opérations CRUD : exemples d'opérations basiques, création d'index et mappings

Atelier : Installation d'une pile Elastic et création d'un index

Savoir installer et configurer ElasticSearch

Prérequis techniques. Mode d'installation en grappe distribuée.
API RESTful en HTTP . Requêtes avec Search Lite et avec Query DSL (domain-specific

📅 Prochaines sessions

Consultez-nous pour les prochaines sessions.

language)

Atelier : premiers pas dans la console DevTools de Kibana.

Indexer des volumes importants de données

Format et stockage des données.

Conversion au format JSON des données à traiter.

Structure des données. Stockage, indexation. Terminologie Elasticsearch : notions de document, index.

Métadonnées : `_index`, `_id`

Choix de l'identifiant par l'application avec l'API index, ou génération automatique d'un identifiant, ...

Indexation inversée. Outils d'interrogation.

Atelier : exemples de requêtes simples et plus complexes : recherche de «phrases», extraction de plusieurs documents, ...

Notion de pertinence du résultat : «score»

Utilisation de 'filtre' pour affiner des requêtes.

Atelier : Mise en oeuvre d'une agrégation de résultats.

Ecriture de requêtes complexes.

Notions d'agrégations,

Atelier : préparation des données, agrégation de mesures, bucket aggregation.

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.

Comprendre comment administrer le système et le surveiller afin de garantir sa disponibilité

Flux logstash et présentation Kibana

Traitement de logs avec logstash

Introduction à beats, installation et configuration

Supervision du système avec MetricBeat et FileBeat

Atelier : recherche, visualisation, création de tableaux de bord et graphiques à partir des données fournies par Elasticsearch

MODALITÉS

Modalités

Modalités : en présentiel, distanciel ou mixte – Horaires de 9H à 12H30 et de 14H à 17H30 soit 7H – Intra et Inter entreprise

Pédagogie : essentiellement participative et ludique, centrée sur l'expérience, l'immersion et la mise en pratique. Alternance d'apports théoriques et d'outils pratiques.

Ressources techniques et pédagogiques : Support de formation au format PDF ou PPT Ordinateur, vidéoprojecteur, Tableau blanc, Visioconférence : Cisco Webex / Teams / Zoom

Pendant la formation : mises en situation, autodiagnostic, travail individuel ou en sous-groupe sur des cas réels

Méthode

Fin de formation : entretien individuel

Satisfaction des participants : questionnaire de satisfaction réalisé en fin de formation

Assiduité : certificat de réalisation (validation des acquis)