



## Elasticsearch : indexation de contenu

CB021

Durée: 2 jours

1 520 €

23 au 24 janvier  
23 au 24 avril

9 au 10 juillet  
22 au 23 octobre  
17 au 18 décembre

### Public :

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

### Objectifs :

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

### Connaissances préalables nécessaires :

disposer de solides bases en administration système Unix/Linux

### Objectifs pédagogiques :

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

### 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



# Phirio

---

## 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 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