

Data Engineering

## Ingestion et indexation de données avec Logstash et Elasticsearch

2 jours (14h00) | ★★★★★ 4,6/5 | BIG-LOG | Évaluation qualitative de fin de stage | Formation délivrée en présentiel ou distanciel <sup>(1)</sup>

Formations Informatique > Big Data > Data Engineering



### À l'issue de ce stage vous serez capable de :

- Comprendre la stack ELK
- Identifier les enjeux et les cas d'utilisation d'un moteur de recherche
- Savoir comment indexer des données Log
- Pouvoir analyser les données Log.

### Niveau requis

Avoir des bases en développement.

### Public concerné

Développeurs de logiciels, ingénieurs en logiciels, architectes de données, sur Data Analyst et Data Engineer.

### Cette formation :

- Est animée par un consultant-formateur dont les compétences techniques, professionnelles et pédagogiques ont été validées par des diplômes et/ou testées et approuvées par l'éditeur et/ou par M2i Formation
- Bénéficie d'un suivi de son exécution par une feuille de présence élargée par demi-journée par les stagiaires et le formateur.

#### (1) Modalité et moyens pédagogique :

Formation délivrée en présentiel ou distanciel \* (e-learning, classe virtuelle, présentiel à distance). Le formateur alterne entre méthodes \*\* démonstrative, interrogative et active (via des travaux pratiques et/ou des mises en situation). La validation des acquis peut se faire via des études de cas, des quiz et/ou une certification.

Les moyens pédagogiques mis en oeuvre (variables suivant les formations) sont : ordinateurs Mac ou PC (sauf pour les cours de l'offre Management), connexion internet fibre, tableau blanc ou paperboard, vidéoprojecteur ou écran tactile interactif (pour le distanciel). Environnements de formation installés sur les postes de travail ou en ligne. Supports de cours et exercices.

\* Nous consulter pour la faisabilité en distanciel. \*\* Ratio variable selon le cours suivi.

# Programme

## Introduction NoSQL

- Nature et exigence de stockage
- Regard sur les BDD relationnelles
- Théorème de CAP
- Les familles des BDD NoSQL
- NoSQL, comparatif et use case

## Stack ELK, présentation

- L'histoire de Elastic Stack
- Pourquoi Elastic Stack
- Les composants de Elastic Stack
- La donnée sur ES
- Cas d'utilisation
- Les fichiers de configuration

## Exemple de travaux pratiques (à titre indicatif)

- Installation de Elasticsearch

## Logstash

- Fonctionnement et concepts
- Différence entre Logstash et Beats
- Logstash vs outil ingestion Big Data
- Installation et configuration de base
- Type de données à ingérer
- Création et configuration d'un pipeline

## Exemple de travaux pratiques (à titre indicatif)

- Installation de Elasticsearch

## Gestion et analyseur de documents

- Indexation, récupération et suppression de documents
- Composants d'un index
- Analyse de texte pour l'indexation et la recherche
- Mapping et la configuration des index
- Les analyzers
- Recherche multi-champs
- Tri
- Pagination

## Cas d'utilisation réel avec Elasticsearch et Logstash

- Mettre en place une architecture Logstash, Kibana
- Utilisation des Filebeat avec Logstash
- Utilisation de Logstash pour l'indexation des fichiers CSV
- Visualiser sur Kibana le résultat de l'indexation

## Modalités d'évaluation des acquis

L'évaluation des acquis se fait :

- En cours de formation, par des études de cas ou des travaux pratiques
- Et, en fin de formation, par un questionnaire d'auto-évaluation ou une certification (M2i ou éditeur)