

Bigdata : architecture et technologies

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 : Fondamentaux

Code de formation : BD001

PRÉSENTATION

Objectifs & compétences

Comprendre les concepts du BigData et savoir quelles sont les technologies implémentées.

Public visé

Chefs de projets, architectes, et toute personne souhaitant connaître les outils et solutions pour mettre en place une architecture BigData.

Pré-requis

Il est demandé aux participants d'avoir une bonne culture générale sur les systèmes d'information.

€ Tarifs

Prix public : 1270 €

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'information

PROGRAMME

Module 1 : Introduction

- Le besoin : volumes importants de données, traitements optimisés de flux de données au fil de l'eau, liés aux nouvelles technologies et aux nouveaux usages.
- Domaines concernés : recherche scientifique, médical, e-commerce, sécurité...
- Développement des techniques sur différents aspects : stockage, indexation/recherche, calcul.
- Définition ETL : Extract Transform Load.
- Les acteurs.

Module 2 : Stockage

- Caractéristiques NoSQL :
- Structure de données proches des utilisateurs, développeurs
- Données structurées et non structurées, documents, images, fichiers XML, JSON, CSV, ...
- Les différents modes et formats de stockage.
- Stockage réparti : réplication, sharding, gossip protocol, hachage,
- Systèmes de fichiers distribués : GFS, HDFS, BigTable, ...
- Les bases de données.
- Quelques exemples de produits et leurs caractéristiques : cassandra, MongoDB, CouchDB, DynamoDB.

Module 3 : Indexation et recherche

- Moteurs de recherche.
- Principe de fonctionnement.
- Méthodes d'indexation.
- Exemple de Lucene, et mise en oeuvre avec solr.
- Recherche dans les bases de volumes importants :
- Exemples de produits et comparaison : dremel, drill, elasticsearch, MapReduce,

Module 4 : Calcul et restitution, intégration

- Différentes solutions : calculs en mode batch, ou en temps réel, sur des flux de données ou des données statiques.
- Les produits : langage de calculs statistiques, R Statistics Language
- Outils de calcul sur des volumes importants : storm en temps réel, hadoop en mode batch.
- Zoom sur Hadoop : complémentarité de HDFS et MapReduce.

📍 Lieux & Horaires

Campus : Ensemble des sites

Durée : 14 heures

Délai d'accès :

Jusqu'à 8 jours avant le début de la formation

Distanciel possible : Oui

📅 Prochaines sessions

Cliquez sur la date choisie pour vous inscrire :

■ 17 / 10 / 2024

📍 : Ensemble des sites

✓ : Distanciel possible

🕒 : 14 heures

📅 : 2 jours

- Evolutions
 - Les offres Saas BigData comme Google BigQuery.
 - Les limites.
- Les nouveautés annoncées

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, autodiagnosics, 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)