

BigData, architecture et technologies



Informatique

Référence formation : 5-PP-BIG - Durée : 2 jours

Objectifs

Comprendre les concepts du BigData et connaître les technologies associées.

Pré-requis

PUBLIC : Chefs de projets, architectes, et toute personne souhaitant connaître les outils et solutions pour mettre en place une architecture BigData.
PREREQUIS : bonne culture générale sur les systèmes d'information.

Contenu pédagogique

1. Introduction

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

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...



XXL Formation

34 rue Raymond Aron
76130 Mont Saint Aignan

Tél : 02 35 12 25 55 – Fax : 02 35 12 25 56

N° siret : 485 050 611 00014 – N° d'agrément : 23.76.03752.76



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...

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 HDMS et MapReduce.

5. Evolutions

- Les offres Saas BigData, type Google BigQuery.
- Les limites actuelles. Les avancées annoncées.

Méthodes pédagogiques	Présentation des concepts, démonstration, exécution, synthèse et exercices pratiques d'assimilation
Modalités pédagogiques	Présentiel - Distanciel - AFEST
Moyens pédagogiques	Formateur expert du domaine - 1 ordinateur, 1 support de cours version papier ou numérique, un bloc-note et un stylo par personne - vidéo projecteur - tableau blanc
Modalités d'évaluation	Positionnement préalable oral ou écrit - Evaluation formative tout au long de la formation - Evaluation sommative faite par le formateur ou à l'aide de la certification NULL : https://www.francecompetences.fr/recherche/NULL
Durée journée de formation	7h00
Public concerné	Salariés - Demandeur d'emploi - Reconversion professionnelle - Public en situation de handicap



XXL Formation

34 rue Raymond Aron
76130 Mont Saint Aignan

Tél : 02 35 12 25 55 – Fax : 02 35 12 25 56

N° siret : 485 050 611 00014 – N° d'agrément : 23.76.03752.76

