

## BIGDATA, ARCHITECTURE ET TECHNOLOGIES

<b>Durée</b>	<b>2 jours</b>	<b>Référence Formation</b>	<b>4-PP-BIG</b>
--------------	----------------	----------------------------	-----------------

### Objectifs

Comprendre les concepts du BigData et connaître les technologies associées.

### Participants

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

### Pré-requis

Bonne culture générale sur les systèmes d'information.

### Moyens pédagogiques

Réflexion de groupe et apports théoriques du formateur

Travail d'échange avec les participants sous forme de réunion-discussion

Utilisation de cas concrets issus de l'expérience professionnelle

Validation des acquis par des questionnaires, des tests d'évaluation, des mises en situation et des jeux pédagogiques.

Remise d'un support de cours.

### PROGRAMME

#### 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

#### Stockage

- Caractéristiques NoSQL :
- <ul>- Structure de données proches des utilisateurs, développeurs
- Données structurées et non structurées, documents, images

#### 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...

#### 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

#### Evolutions

#### CAP ÉLAN FORMATION

[www.capelanformation.fr](http://www.capelanformation.fr) - Tél : 04.86.01.20.50

Mail : [contact@capelanformation.fr](mailto:contact@capelanformation.fr)

Organisme enregistré sous le N° 76 34 0908834

version 2024



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