

Vue d'ensemble de technologies Big-Data

Programme

Introduction

- La philosophie NoSQL
- Différences entre bases relationnelles et NoSQL
- Architecture distribuée
- Le format JSON comme format d'échange
- Exemples de bases NoSQL :
 - Base orientée clés/valeurs
 - Base orientée « colonnes »
 - Base volatile
 - Base orientée « documents »
 - Base de données orientées graphes
 - Base de données temps-réel
 - Base de données avec recherche « full

Présentation de NEO4J (base de données orientée graphes)

- Installation rapide
- Langage d'interrogation
- Prise en main avec TP

Présentation de INFLUXDB (base de données temps-réel)

- Installation rapide
- Langage d'interrogation
- Prise en main avec TP

Présentation de ELASTICSEARCH (base de données de recherche full-text)

- Installation rapide
- Prise en main avec TP
- Présentation de Kibana
- Langage de requêtage

Prise en main de HADOOP-v2

- Mise en oeuvre d'un Cluster
- Le gestionnaire YARN
- Utilisation de HDFS
- 1er exemple de MapReduce

Ecriture d'un ALGO.MAP/REDUCE

- Application Java
- Les différentes étapes
- Utilisation d'autres langages

Application connexes

- Présentation de HBase
- Présentation de Hive
- Présentation de Pig
- Collecte de données avec FLUME
- Présentation et mise en oeuvre de Spark

Référence

THBI1128

Durée

2 jours / 14 heures

Prix HT / stagiaire

1250€

Objectifs pédagogiques

- Analyser le principe du Big-Data avec les technologies phares
- Définir les technologies : le cluster Hadoop, MongoDB, Neo4J (ou OrientDB), InfluxDB, ElasticSearch
- Analyser les difficultés propres à un projet Big Data, les freins et les apports, tant sur les aspects techniques que sur ceux liés à la gestion du projet

Niveau requis

- Garantir posséder des connaissances de base Linux
- Garantir posséder une bonne culture générale des systèmes d'information

Public concerné

- Chefs de projets, architectes, développeurs, data scientists et toute personne souhaitant connaître les outils et solutions pour concevoir et mettre en oeuvre une architecture Big Data

Formateur

Les formateurs intervenants pour Themanis sont qualifiés par notre Responsable Technique Olivier Astre pour les formations informatiques et bureautiques et par Didier Payen pour les formations management.

Moyens pédagogiques et techniques

Salles de formation (accessibles et adaptables aux besoins des personnes en situation de handicap) équipée d'un ordinateur de dernière génération par stagiaire, réseau haut débit et vidéo-projection UHD

Documents supports de formation projetés
Apports théoriques, étude de cas concrets et exercices

Mise à disposition en ligne de documents supports à la suite de la formation

Dispositif de suivi de l'exécution de l'évaluation des résultats de la formation

Feuilles d'émargement (signature électronique privilégiée)

Evaluations formatives et des acquis sous forme de questions orales et/ou écrites (QCM) et/ou mises en situation

Prise en main de MONGODB

- Concepts de base
- Documents
- Types intégrés
- Collections
- Bases de données
- Manipulation de documents
- Indexation
- Requêtes d'agrégations
- Pipelines
- Architecture d'un Cluster H-A avec sharding