

Formation Industrie 4.0 : Real Time Financial & IoT Data Processing

3 jours (21 heures)

Présentation

Nous venons de rentrer dans l'ère de l'industrie 4.0. Grâce aux apports des technologies Big Data de nouvelles générations telles qu'Apache Kudu, Impala, Spark, et Flink de nouvelles approches ont émergé comprenant notamment le Real Time Financial & l'IoT Data Processing.

Cette formation adopte une approche globale et ce veut un guide pratique couvrant les aspects les plus importants du Big Data destinés aux entreprises modernes.

Vous allez pouvoir appréhender et résoudre une majorité de vos problèmes d'Architecture, de performance et de stockage issue de vos problématiques IT du quotidien de votre entreprise.

Nous mettrons en œuvre les outils et applications de nouvelle génération utilisés :

- Pour le stockage de données volumineuses
- L'optimisation de l'entrepôt de données
- L'ingestion et le traitement en temps réel et par lots de données
- La visualisation de données en temps réel
- La gouvernance de données volumineuses
- Calcul distribué Big Data en mémoire

Objectifs

- Installer Apache Kudu, Impala, Spark, Flink pour moderniser les environnements d'entreposage de données d'entreprise et de veille stratégique, avec des exemples concrets, faciles à suivre et des conseils pratiques.
- Intégration de HBase, Solr, Oracle, SQL Server, MySQL, Flume, Kafka, HDFS et Amazon S3 avec Apache Kudu, Impala et Spark
- Comprendre les sujets relatifs au Big Data d'entreprise tels que la gouvernance du Big Data, la gestion des métadonnées, le lignage des données, l'analyse d'impact et l'application des règles, et comment utiliser Cloudera Navigator pour effectuer des tâches courantes de gouvernance des données.
- Implémenter des cas d'utilisation de données volumineuses tels que l'entreposage de données volumineuses, l'optimisation de l'entrepôt de données, l'Internet des objets, l'ingestion et l'analyse de données en temps réel, le traitement des événements complexes et la modélisation prédictive évolutive.
- Être en mesure de déployer des solutions à travers différentes études de cas : données massives réalisées dans le monde réel auprès d'entreprises innovantes, notamment Navistar, Cerner, British Telecom, Shopzilla, Thomson Reuters et Mastercard.

Public visé

- Développeur, Lead Developer, Data Engineer

Pré-requis

- Bonne notion en programmation sur un langage récent
- Savoir les rudiments sur un système Linux
- Ingénieur Big Data confirmé

Pour aller plus loin

- Autour du sujet, nous vous recommandons également les formations :
 - [IBM Bluemix & Watson](#)
 - [Le Machine Learning avec Spark](#)
 - [Le Deep Learning avec TensorFlow](#)

Programme de notre Formation Apache Flink, Kudu, Impala & Spark

- Big Data de nouvelles générations
- Introduction au Kudu
- Introduction à Impala
- Analyse de données de haute performance avec Impala
- Introduction à Spark et à Flink
- Traitement de données haute performance avec Spark, Flink et Kudu
- Ingestion et traitement de données par lots et en temps réel
- Entreposage de données volumineuses
- Visualisation Big Data
- Big Data Computing
- Gouvernance et gestion
- Big Data dans le cloud
- Études de cas Big Data

Sociétés concernées

Cette formation s'adresse aux entreprises, petites ou grandes, souhaitant former ses équipes à une nouvelle technologie informatique avancée.

Méthodes pédagogiques

Stage Pratique : 60% Pratique, 40% Théorie. Support de la formation distribué au format numérique à tous les participants.

Organisation

Le cours alterne les apports théoriques du formateur soutenus par des exemples et des séances de réflexions, et de travail en groupe.

Validation

À la fin de la session, un questionnaire à choix multiple permet de vérifier l'acquisition correcte des compétences.

Sanction

Une attestation sera remise à chaque stagiaire qui aura suivi la totalité de la formation.