



Formation AWS Big Data

Programme de formation



Description

Grâce à la formation Amazon Web Services - Big Data on AWS, vous découvrirez des solutions de Big Data basées sur le cloud comme Amazon EMR, Amazon Redshift, Amazon Kinesis et le reste de la plateforme AWS Big Data. Vous découvrirez à travers du cours et des exercices comment utiliser Amazon EMR afin de traiter des données grâce au vaste écosystème d'outils Hadoop tels que Hive et Hue. Cette formation vous permettra également d'apprendre à créer des environnements de Big Data, à utiliser Amazon DynamoDB, Amazon Redshift, Amazon QuickSight, Amazon Athena et Amazon Kinesis. Enfin, vous découvrirez comment concevoir au mieux des environnements de Big Data sécurisés et économiques.

Coût et durée de la formation :

Durée

3 jours

Tarif

inter : 1990€ HT /pers

intra : selon la demande du client

Déroulement



Public

- Architectes de solutions
- Administrateurs SysOps
- Experts en science des données
- Analystes de donnée



Pré-requis

- Démarrage technique AWS ou AWS Cloud Practitioner
- Data Analytics Fundamentals
- Concepts de bases de données
- Concepts du stockage, du traitement et de l'analyse des données
- Infrastructure technique



Méthode pédagogique

- Formation théorique
- Ateliers pratiques

Programme détaillé de la formation

Journée 1 •

- Présentation du Big Data
- Ingestion et transferts Big Data
- Streaming Big Data et Amazon Kinesis
- Solutions de stockage
- Traitements et Analyses Big Data

Journée 2

- Apache Hadoop et Amazon EMR
- Utilisations d'Amazon EMR
- Cadres de programmation Hadoop
- Interfaces Web sur Amazon EMR
- Apache Spark sur Amazon EMR
-

Journée 3

- Utilisation d'Amazon Glue pour automatiser les traitements ETL
- Amazon Redshift et Big Data
- Sécuriser ses déploiements sur Amazon
- Gérer et surveiller les coûts de son infrastructure Big Data
- Visualisation et orchestration de ses traitements Big Data
- Design patterns Big Data

Objectif

À l'issue de la formation, les stagiaires seront à même de :

- Intégrer une solution AWS dans un écosystème de Big Data
- Utiliser Apache Hadoop dans le cadre d'Amazon EMR
- Identifier les composants d'un cluster Amazon EMR puis lancer et configurer un cluster Amazon EMR
- Utiliser les frameworks de programmation courants compatibles avec Amazon EMR, notamment Hive, Pig et Streaming
- Optimiser la facilité d'utilisation d'Amazon EMR avec Hadoop User Experience (Hue)
- Utiliser les analyses en mémoire avec Apache Spark sur Amazon EMR
- Choisir les options de stockage de données AWS appropriées
- Identifier les avantages de l'utilisation d'Amazon Kinesis pour le traitement de Big Data en temps quasi réel
- Utiliser Amazon Redshift pour stocker et analyser des données de manière efficace
- Comprendre et gérer les coûts et la sécurité pour une solution de Big Data
- Identifier les options pour l'importation, le transfert et la compression de données
- Exploiter Amazon Athena pour les analyses par requêtes ad hoc
- Utiliser AWS Glue pour automatiser l'extraction, la transformation et le chargement (ETL) de charges de travail
- Utiliser un logiciel de visualisation pour représenter les données et les requêtes à l'aide d'Amazon QuickSight



Lea Cesaratto

Responsable Training



06 85 10 85 36



lea.cesaratto@wescale.fr

Service Training



01 83 75 05 26



training@wescale.fr

wescale

23 rue Taitbout 75009 Paris

www.wescale.fr | blog.wescale.fr