

Big Data on AWS (Amazon Web Services)

Durée: 3 Jours Réf de cours: GK4509 Version: 3.1

Résumé:

La formation Big Data sur AWS présente des solutions de Big Data basées sur le cloud ainsi qu'Amazon Elastic MapReduce (EMR), la plate-forme de Big Data d'AWS, Amazon Redshift et Amazon Kinesis. Dans le cadre de ce cours, vous découvrirez comment utiliser Amazon EMR afin de traiter des données grâce au vaste écosystème d'outils Hadoop tels que Hive et Hue. Vous apprendrez également à créer des environnements de Big Data, à utiliser Amazon DynamoDB, Amazon Redshift et Amazon Kinesis, Amazon QuickSight, Amazon Athena et à appliquer au mieux les bonnes pratiques afin de concevoir des environnements de Big Data sécurisés et économiques.

Public visé:

Personnes en charge de la conception et de l'implémentation de solutions de Big Data, c'est-à-dire les architectes de solutions ainsi que les analystes de données souhaitant découvrir les solutions de Big Data sur AWS.

Objectifs pédagogiques:

- Solutions Fit AWS au sein d'un écosystème Big Data
- Comprendre les applications Apache Hadoop dans le contexte d'Amazon EMR
- Identifier les composants d'un cluster Amazon EMR
- Lancer et configurer un cluster Amazon EMR
- Utiliser les frameworks de programmation courants disponibles pour Amazon EMR notamment Hive, Pig and Streaming.
- Utiliser Hue pour améliorer la facilité d'utilisation de Amazon EMR
- Utiliser les analyses en mémoire avec Spark sur Amazon EMR
- Choisir les options de stockage de données appropriées
- Identifier les avantages de l'utilisation de Amazon Kinesis pour le process Big Data proches du temps réel
- Utiliser Amazon Redshift pour stocker et analyser efficacement les données
- Comprendre et gérer les coûts et la sécurité d'une solution Big Data
- Sécuriser une solution Big Data
- Identifier les options pour récupérer, transférer et compresser des données
- Comprendre Amazon Athena pour les analyses de requêtes ad-hoc
- Utiliser le logiciel de visualisation pour représenter les données et les requêtes via Amazon QuickSight
- Orchestrer le flux de la Big Data via AWS Data Pipeline

Pré-requis:

- Avoir des connaissances de base sur les technologies Big Data, notamment Apache Hadoop et HDFS Pig, Hive et MapReduce.
- Savoir utiliser les principaux services AWS et l'implémentation de cloud public.
- Les participants doivent avoir suivi le cours Notions de base AWS ou posséder un niveau d'expérience équivalent.
- La compréhension des concepts d'entrepôt de données, de système de base de données relationnelle et de conception de base données est recommandée.

Test et certification

Les professionnels peuvent valider leurs compétences techniques dans ce domaine clé, grâce au passage de l'examen de niveau **Specialty** pour obtenir la certification AWS Certified Big Data. Notez que celle-ci requiert une certification de niveau Associate pour être validée.

Contenu:

Jour 1

- Vue d'ensemble de la Big data
- Récupérer et transférer la Big Data
- Streaming Big Data et Amazon Kinesis
- Solutions de stockage Big Data
- Process Big Data et Analyse

Jour 2

- Apache Hadoop et Amazon EMR
- Utiliser Amazon EMR
- Frameworks de programmation Hadoop
- Interfaces Web sur Amazon EMR
- Apache Spark sur Amazon EMR

Jour 3

- Amazon Redshift et Big Data
- Visualiser et Orchestrer la Big Data
- Gérer les coûts de la Big Data
- Sécuriser les déploiements de votre Amazon
- Modèles de conception Big Data

Méthodes pédagogiques :

Support de cours officiel AWS remis en anglais au format électronique. Pour profiter pleinement du support électronique dès le 1er jour, nous invitons les participants à se munir d'un PC ou d'une tablette, qu'ils pourront connecter en WiFi dans nos locaux de Rueil, Lyon ou Lille.

Autres moyens pédagogiques et de suivi:

- Compétence du formateur : Les experts qui animent la formation sont des spécialistes des matières abordées et ont au minimum cinq ans d'expérience d'animation. Nos équipes ont validé à la fois leurs connaissances techniques (certifications le cas échéant) ainsi que leur compétence pédagogique.
- Evaluation : Chaque participant, à l'issue de la formation, répond à un questionnaire d'évaluation qui est ensuite étudié par nos équipes pédagogiques en vue de maintenir et d'améliorer la qualité de nos prestations.
- Suivi d'exécution : Une feuille d'émargement par demi-journée de présence est signée par tous les participants et le formateur.