

## Building Streaming Data Analytics Solutions on AWS

**Durée: 1 Jour**    **Réf de cours: GK7400**    **Méthodes d'apprentissage: Classe à distance**

### Résumé:

Cette formation permet aux participants d'acquérir les compétences et connaissances nécessaires pour construire des solutions d'analyse de données en continu à l'aide des services AWS, notamment Amazon Kinesis et Amazon Managed Streaming for Apache Kafka (Amazon MSK). Amazon Kinesis est un service de streaming de données en temps réel massivement évolutif et durable. Amazon MSK offre un service Apache Kafka sécurisé, entièrement géré et hautement disponible. Les participants apprendront également comment Amazon Kinesis et Amazon MSK s'intègrent aux services AWS tels que AWS Glue et AWS Lambda. La formation aborde les composants d'ingestion, de stockage et de traitement des données en continu du pipeline d'analyse des données. Les participants apprendront à appliquer les meilleures pratiques en matière de sécurité, de performance et de gestion des coûts au fonctionnement de Kinesis et d'Amazon MSK.

Cette formation est délivrée **en synchrone à distance tout en garantissant l'accès à un environnement d'apprentissage complet !**

Mise à jour : 13.03.2023

### Public visé:

Cette formation s'adresse aux Ingénieurs et architectes de données et aux développeurs qui veulent construire et gérer des applications en temps réel et des solutions d'analyse de données en continu.

### Objectifs pédagogiques:

- A l'issue de la formation, les participants seront capables de :
- Choisir les flux, les clusters, les thèmes, l'approche de mise à l'échelle et la topologie de réseau appropriés pour un cas d'utilisation particulier
- Comprendre les caractéristiques et les avantages d'une architecture de données moderne
- Comprendre comment le stockage et le traitement des données affectent les mécanismes d'analyse et de visualisation nécessaires pour obtenir des informations commerciales exploitables
- Apprendre comment les services de streaming AWS s'intègrent dans une architecture de données moderne
- Sécuriser les données en continu au repos et en transit
- Concevoir et mettre en œuvre une solution d'analyse de données en continu
- Surveiller les charges de travail analytiques pour identifier les problèmes et y remédier
- Identifier et appliquer les techniques appropriées, telles que la compression, le partage et le partitionnement, afin d'optimiser le stockage des données
- Appliquer les meilleures pratiques en matière de gestion des coûts
- Sélectionner et déployer les options appropriées pour ingérer, transformer et stocker les données en temps réel et en temps quasi réel

### Pré-requis:

Au moins un an d'expérience en analyse de données ou une expérience directe dans la création d'applications en temps réel ou de solutions d'analyse en continu.

Nous suggérons le livre blanc Streaming Data Solutions on AWS pour ceux qui ont besoin d'une remise à niveau sur les concepts de streaming.

- GK4502 - Architecture sur AWS
- GK7377 - Building Data Lakes on AWS
- GK7378 - Building Batch Data Analytics Solutions on AWS

### Test et certification

Afin de se préparer efficacement à l'obtention de la certification **AWS Certified Data Analytics Speciality**, nous vous recommandons de suivre également les formations suivantes :

[GK7377](#) - Building Data Lakes on AWS

[GK7378](#) - Building Batch Data Analytics Solutions on AWS

[GK7379](#) - Building Data Analytics Solutions using Amazon Redshift

\*\*

Cette formation mène également à la Certification professionnelle

## Après cette formation, nous vous conseillons le(s) module(s) suivant(s):

- GK7379 - Building Data Analytics Solutions Using Amazon Redshift

### Contenu:

Vue d'ensemble de l'analyse des données et du pipeline de données

- Cas d'utilisation de l'analyse des données
- Utilisation du pipeline de données pour l'analyse

Utilisation des services de streaming dans le pipeline d'analyse de données

- Importance de l'analyse de données en continu
- Le pipeline d'analyse de données en continu
- Concepts de streaming

Introduction aux services de streaming AWS

- Les services de streaming de données dans AWS
- Amazon Kinesis dans les solutions analytiques - Démonstration : Explorer Amazon Kinesis Data Streams

Atelier pratique : Mise en place d'un pipeline de diffusion en continu avec Amazon Kinesis

- Utilisation d'Amazon Kinesis Data Analytics
- Introduction à Amazon MSK - Aperçu de Spark Streaming

Utiliser Amazon Kinesis pour l'analyse de données en temps réel

- Exploration d'Amazon Kinesis à l'aide d'une charge de travail clickstream
- Création de flux de données et de livraison Kinesis
- Démonstration : Comprendre les producteurs et les consommateurs
- Création de producteurs de flux
- Création de consommateurs de flux
- Création et déploiement d'applications Flink dans Kinesis Data Analytics
- Démonstration : Explorer les ordinateurs portables Zeppelin pour Kinesis Data Analytics

Atelier pratique : Analyse de flux avec Amazon Kinesis Data Analytics et Apache Flink

Sécurisation, surveillance et optimisation d'Amazon Kinesis

- Optimiser Amazon Kinesis pour obtenir des informations commerciales exploitables
- Meilleures pratiques en matière de sécurité et de surveillance

Utilisation d'Amazon MSK dans les solutions d'analyse de données en continu

- Cas d'utilisation d'Amazon MSK
- Création de clusters MSK - Démonstration : Provisionnement d'un cluster MSK
- Ingérer des données dans Amazon MSK

Atelier pratique : Introduction au contrôle d'accès avec Amazon MSK - Transformation et traitement dans Amazon MSK

Sécuriser, surveiller et optimiser Amazon MSK

- Optimiser Amazon MSK
- Démonstration : Augmentation de la capacité de stockage d'Amazon MSK

Atelier pratique : Pipeline de streaming Amazon MSK et déploiement d'applications

- Sécurité et surveillance
- Démonstration : Surveillance d'un cluster MSK

Conception de solutions d'analyse de données en continu

- Examen des cas d'utilisation

Atelier pratique : Conception d'un flux de travail d'analyse de données en continu

Développer des architectures de données modernes sur AWS

- Architectures de données modernes

## Méthodes pédagogiques :

Pour le suivi de cette formation à distance depuis un site client équipé, il suffit d'avoir une bonne connexion internet, un casque avec micro et d'être dans un endroit au calme pour en profiter pleinement. Les participants inscrits à distance pourront vérifier en amont leur installation technique.

L'accès à l'environnement d'apprentissage (Support de cours officiel AWS en anglais au format électronique, labs), ainsi qu'aux feuilles d'émargement et d'évaluation est assuré.

---

## 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.
- Suivi d'exécution : Une feuille d'émargement par demi-journée de présence est signée par tous les participants et le formateur.
- Modalités d'évaluation : le participant est invité à s'auto-évaluer par rapport aux objectifs énoncés.
- Chaque participant, à l'issue de la formation, répond à un questionnaire de satisfaction qui est ensuite étudié par nos équipes pédagogiques en vue de maintenir et d'améliorer la qualité de nos prestations.

### Délais d'inscription :

- Vous pouvez vous inscrire sur l'une de nos sessions planifiées en inter-entreprises jusqu'à 5 jours ouvrés avant le début de la formation sous réserve de disponibilité de places et de labs le cas échéant.
- Votre place sera confirmée à la réception d'un devis ou ""booking form"" signé. Vous recevrez ensuite la convocation et les modalités d'accès en présentiel ou distanciel.
- Attention, si vous utilisez votre Compte Personnel de Formation pour financer votre inscription, vous devrez respecter un délai minimum et non négociable fixé à 11 jours ouvrés.