



AWS – Building Batch Data Analytics Solutions on AWS

Description

Concevoir des solutions d'analyse de données par lots avec AWS

L'analyse de données est devenue un pilier essentiel pour les entreprises qui cherchent à tirer parti de leurs informations pour des décisions stratégiques. La formation AWS « Building Batch Data Analytics Solutions on AWS » vous permettra d'acquérir les compétences nécessaires pour concevoir des pipelines de traitement par lots sur AWS. Tout en utilisant des services comme Amazon EMR. Vous apprendrez à exploiter Apache Spark et Apache Hadoop pour optimiser la collecte, l'ingestion et le traitement des données. Ce cours s'adresse à ceux qui veulent maîtriser les outils d'AWS tout en intégrant des pratiques de sécurité et de gestion des coûts.

Contenu du cours

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

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

Module 1 : Introduction à Amazon EMR

- Utiliser Amazon EMR dans des solutions analytiques
- Architecture des clusters Amazon EMR
- Démo interactive : Lancement d'un cluster Amazon EMR
- Stratégies de gestion des coûts

Module 2 : Pipeline analytique avec Amazon EMR : ingestion et stockage

- Optimisation du stockage avec Amazon EMR
- Techniques d'ingestion de données

Module 3 : Analyse de données par lots à haute performance avec Apache Spark sur Amazon EMR

- Cas d'utilisation d'Apache Spark sur Amazon EMR
- Pourquoi choisir Apache Spark sur Amazon EMR
- Concepts de Spark
- Démo interactive : Connexion à un cluster EMR et exécution de commandes Scala avec Spark shell

- Transformation, traitement et analyses
- Utilisation des notebooks avec Amazon EMR
- Laboratoire pratique : Analyse de données à faible latence avec Apache Spark sur Amazon EMR

Module 4 : Traitement et analyse des données par lots avec Amazon EMR et Apache Hive

- Utilisation d'Amazon EMR avec Hive pour traiter les données par lots
- Transformation, traitement et analyses
- Laboratoire pratique : Traitement des données par lots avec Amazon EMR et Hive
- Introduction à Apache HBase sur Amazon EMR

Module 5 : Traitement des données sans serveur

- Traitement, transformation et analyses de données sans serveur
- Utilisation d'AWS Glue avec des charges de travail Amazon EMR
- Laboratoire pratique : Orchestration du traitement des données dans Spark avec AWS Step Functions

Module 6 : Sécurisation et surveillance des clusters Amazon EMR

- Sécurisation des clusters EMR
- Démo interactive : Chiffrement côté client avec EMRFS
- Surveillance et dépannage des clusters Amazon EMR
- Démo : Revue de l'historique des clusters Apache Spark

Module 7 : Conception de solutions d'analyse de données par lots

- Cas d'utilisation de l'analyse de données par lots
- Activité : Conception d'un flux de travail d'analyse de données par lots

Module B : Développement d'architectures de données modernes sur AWS

- Architectures de données modernes

Documentation

- Support de cours numérique inclus

Profils des participants

- Ingénieurs de plateformes de données
- Architectes de solutions de données
- Responsables de pipelines analytiques
- Opérateurs AWS spécialisés dans le Big Data

Connaissances Préalables

- Expérience avec Apache Spark et Apache Hadoop (minimum 1 an)
- Notions de base en gestion de clusters et réseaux sur AWS
- Connaissances en sécurité des données et chiffrement
- Compréhension des concepts d'architecture de données
- Familiarité avec les outils de surveillance et d'optimisation de performance

Objectifs

- Concevoir des pipelines de données analytiques

- Optimiser le stockage et l'ingestion des données
- Utiliser Amazon EMR avec Apache Spark et Hive
- Appliquer des pratiques de gestion des coûts
- Sécuriser les clusters et les données sur AWS
- Superviser les charges de travail analytiques
- Automatiser le traitement des données par lots

Description

Formation AWS - Building Batch Data Analytics Solutions on AWS

Niveau

Fondamental

Prix de l'inscription en Présentiel (CHF)

850

Prix de l'inscription en Virtuel (CHF)

850

Durée (Nombre de Jours)

1

Reference

AWS-301