



Formation Building Modern Data Analytics Solutions on AWS

N° ACTIVITÉ : 11 92 18558 92

TÉLÉPHONE : 01 85 77 07 07

E-MAIL : inscription@hubformation.com

Développez les compétences nécessaires pour moderniser l'architecture des données afin de prendre des décisions avec rapidité et agilité à grande échelle grâce à un instructeur AWS expert.

Référence	AWSBMDAS
Durée	4 jours (28h)
Tarif	2 800 €HT

Objectifs

- | Comparer les fonctionnalités et les avantages des entrepôts de données, des lacs de données et des architectures de données modernes
- | Concevoir et mettre en oeuvre une solution d'analyse de données par lots
- | Identifier et appliquer les techniques appropriées
- | Ingérer, transformer et stocker des données en temps réel et en temps quasi réel
- | Sécuriser les données au repos et en transit
- | Identifier et résoudre les problèmes
- | Comparer les fonctionnalités et les avantages des entrepôts de données, des lacs de données et des architectures de données modernes
- | Concevoir et mettre en oeuvre une solution d'analyse d'entrepôt de données
- | Sécuriser les données au repos et en transit
- | Appliquer les méthodologies de lac de données dans la planification et la conception d'un lac de données
- | Articuler les composants et les services requis pour créer un lac de données AWS
- | Sécuriser un lac de données avec les autorisations appropriées
- | Interroger, analyser et visualiser des données dans un lac de données

- | Identifier les fonctionnalités et les avantages d'une architecture de données moderne.
- | Concevoir et mettre en oeuvre une solution d'analyse de données en continu
- | Choisir les flux, les clusters, les sujets, l'approche de mise à l'échelle et la topologie de réseau appropriés pour un cas d'utilisation métier particulier

Prérequis

- | Avoir suivi les cours :
- | AWS Technical Essentials ou Architecting sur AWS
- | Building Data Lakes on AWS

Public

- | Architectes et opérateurs qui construisent et gèrent des pipelines d'analyse de données

- | Ingénieurs et architectes de données
- | Ingénieurs plateformes de données
- | Ingénieurs d'entrepôt de données
- | Développeurs qui souhaitent créer et gérer des applications en temps réel et des solutions d'analyse de données en continu
- | Architectes de solutions

Programme de la formation

SESSIONS PROGRAMMÉES

A DISTANCE (FRA)

- du 14 au 17 mai 2024
- du 9 au 12 juillet 2024
- du 24 au 27 septembre 2024
- du 20 au 23 novembre 2024

[VOIR TOUTES LES DATES](#)

Building Batch Data Analytics Solutions on AWS

Présentation de l'analyse des données et du pipeline de données

| Cas d'utilisation de l'analyse de données

Utilisation du pipeline de données pour l'analyse

Présentation d'Amazon EMR

| Utilisation d'Amazon EMR dans les solutions d'analyse

| Architecture de cluster Amazon EMR

| Stratégies de gestion des coûts

Pipeline d'analyse de données à l'aide d'Amazon EMR : ingestion et stockage

| Optimisation du stockage avec Amazon EMR

| Techniques d'ingestion de données

Analyse de données par lots hautes performances à l'aide d'Apache Spark sur Amazon EMR

| Cas d'utilisation d'Apache Spark sur Amazon EMR

| Pourquoi Apache Spark sur Amazon EMR

| Concepts de Spark

| Transformation, traitement et analytique

| Utilisation de blocs-notes avec Amazon EMR

| Mise en pratique 1 : Analyse de données à faible latence à l'aide d'Apache Spark sur Amazon EMR

Traitement et analyse des données de lot avec Amazon EMR et Apache Hive

| Utilisation d'Amazon EMR avec Hive pour traiter les données par lots

| Transformation, traitement et analytique

| Introduction à Apache HBase sur Amazon EMR

| Mise en pratique 2 : traitement de données par lots à l'aide d'Amazon EMR avec Hive

Building Data Analytics Solutions Using Amazon Redshift

Présentation de l'analyse des données et du pipeline de données

| Cas d'utilisation de l'analyse de données

| Utilisation du pipeline de données pour l'analyse

Utilisation d'Amazon Redshift dans le pipeline d'analyse de données

| Pourquoi Amazon Redshift pour l'entreposage de données ?

| Présentation d'Amazon Redshift

Présentation d'Amazon Redshift

| Architecture Amazon Redshift

| Démo interactive 1 : Visite de la console Amazon Redshift

| Fonctionnalités Amazon Redshift

| Mise en pratique : Charger et interroger des données dans un cluster Amazon Redshift

Ingestion et stockage

| Ingestion

| Distribution et stockage des données

| Interroger des données dans Amazon Redshift

| Mise en pratique : Analyse de données à l'aide d'Amazon Redshift Spectrum

Traitement et optimisation des données

| Transformation des données

| Interrogation avancée

| La gestion des ressources

| Automatisation et optimisation

| Mise en pratique : Transformation et interrogation des données dans Amazon Redshift

Sécurité et surveillance des clusters Amazon Redshift

| Sécurisation du cluster Amazon Redshift

| Surveillance et dépannage des clusters Amazon Redshift

Conception de solutions d'analyse d'entrepôt de données

| Revue de cas d'utilisation de l'entrepôt de données

| Activité : Conception d'un workflow d'analyse d'entrepôt de données

Développement d'architectures de données modernes sur AWS

| Architectures de données modernes

Building Data Lakes on AWS

Introduction aux lacs de données

- | Décrire la valeur des lacs de données
- | Comparer les lacs de données et les entrepôts de données
- | Décrire les composants d'un lac de données
- | Reconnaître les architectures communes construites sur des lacs de données

Ingestion, catalogage et préparation des données

- | Décrire la relation entre le stockage du lac de données et l'ingestion de données
- | Décrire les robots AWS Glue et comment ils sont utilisés pour créer un catalogue de données
- | Identifier le formatage, le partitionnement et la compression des données pour un stockage et une requête efficaces
- | Mise en pratique : Configurer un lac de données simple

Traitement et analyse des données

- | Reconnaître comment le traitement des données s'applique à un lac de données
- | Utiliser AWS Glue pour traiter les données dans un lac de données
- | Décrire comment utiliser Amazon Athena pour analyser les données dans un lac de données

Création d'un lac de données avec AWS Lake Formation

- | Décrire les fonctionnalités et les avantages d'AWS Lake Formation
- | Utiliser AWS Lake Formation pour créer un lac de données
- | Comprendre le modèle de sécurité AWS Lake Formation
- | Mise en pratique : Créer un lac de données à l'aide d'AWS Lake Formation

Configurations supplémentaires de Lake Formation

- | Automatiser AWS Lake Formation à l'aide de plans et de flux de travail
- | Appliquer des contrôles de sécurité et d'accès à AWS Lake Formation
- | Faites correspondre les enregistrements avec AWS Lake Formation FindMatches
- | Visualisez les données avec Amazon QuickSight
- | Mise en pratique : Automatiser la création d'un lac de données à l'aide des plans AWS Lake Formation
- | Mise en pratique : Visualisation des données à l'aide d'Amazon QuickSight

Architecture et révision des cours

- | Vérification des connaissances après le cours
- | Revue d'architecture

Building Streaming Data Analytics Solutions on AWS

Présentation de l'analyse des données et du pipeline de données

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

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

- | L'importance de l'analyse des données en continu
- | Le pipeline d'analyse de données en continu
- | Concepts de streaming

Introduction aux services de diffusion en continu AWS

- | Services de données en continu dans AWS
- | Amazon Kinesis dans les solutions d'analyse
- | Utilisation d'Amazon Kinesis Data Analytics
- | Présentation d'Amazon MSK
- | Présentation de Spark Streaming
- | Mise en pratique : Configuration d'un pipeline de diffusion en continu avec Amazon Kinesis

Utilisation d'Amazon Kinesis pour l'analyse de données en temps réel

- | Explorer Amazon Kinesis à l'aide d'une charge de travail clickstream
- | Création de données Kinesis et de flux de diffusion
- | Construire des producteurs de flux
- | Construire des consommateurs de flux
- | Création et déploiement d'applications Flink dans Kinesis Data Analytics
- | Mise en pratique : analyses en continu avec Amazon Kinesis Data Analytics et Apache Flink

Sécurisation, surveillance et optimisation d'Amazon Kinesis

- | Optimisez Amazon Kinesis pour obtenir des informations commerciales exploitables
- | Bonnes pratiques de sécurité et de surveillance

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

- | Cas d'utilisation pour Amazon MSK
- | Création de clusters MSK
- | Ingestion de données dans Amazon MSK
- | Transformation et traitement dans Amazon MSK
- | Mise en pratique : Introduction au contrôle d'accès avec Amazon MSK

Sécurisation, surveillance et optimisation d'Amazon MSK

- | Optimisation d'Amazon MSK
- | Démonstration : Mise à l'échelle du stockage Amazon MSK
- | Sécurité et surveillance
- | Mise en pratique : pipeline de diffusion en continu Amazon MSK et déploiement d'applications

Conception de solutions d'analyse de données en continu

- | Examen des cas d'utilisation
- | Exercice de classe : Conception d'un flux de travail d'analyse de données en continu

Développement d'architectures de données modernes sur AWS

- | Architectures de données modernes

Méthode pédagogique

Chaque participant travaille sur un poste informatique qui lui est dédié. Un support de cours lui est remis soit en début soit en fin de cours. La théorie est complétée par des cas pratiques ou exercices corrigés et discutés avec le formateur. Le formateur projette une présentation pour animer la formation et reste disponible pour répondre à toutes les questions.

Méthode d'évaluation

Tout au long de la formation, les exercices et mises en situation permettent de valider et contrôler les acquis du stagiaire. En fin de formation, le stagiaire complète un QCM d'auto-évaluation.

Suivre cette formation à distance

Voici les prérequis techniques pour pouvoir suivre le cours à distance :

- | Un ordinateur avec webcam, micro, haut-parleur et un navigateur (de préférence Chrome ou Firefox). Un casque n'est pas nécessaire suivant l'environnement.
- | Une connexion Internet de type ADSL ou supérieure. Attention, une connexion Internet ne permettant pas, par exemple, de recevoir la télévision par Internet, ne sera pas suffisante, cela engendrera des déconnexions intempestives du stagiaire et dérangera toute la classe.
- | Privilégier une connexion filaire plutôt que le Wifi.
- | Avoir accès au poste depuis lequel vous suivrez le cours à distance au moins 2 jours avant la formation pour effectuer les tests de connexion préalables.
- | Votre numéro de téléphone portable (pour l'envoi du mot de passe d'accès aux supports de cours et pour une messagerie instantanée autre que celle intégrée à la classe virtuelle).
- | Selon la formation, une configuration spécifique de votre machine peut être attendue, merci de nous contacter.
- | Pour les formations incluant le passage d'une certification la dernière journée, un voucher vous est fourni pour passer l'examen en ligne.
- | Pour les formations logiciel (Adobe, Microsoft Office...), il est nécessaire d'avoir le logiciel installé sur votre machine, nous ne fournissons pas de licence ou de version test.
- | Horaires identiques au présentiel.

Accessibilité

Les sessions de formation se déroulent sur des sites différents selon les villes ou les dates, merci de nous contacter pour vérifier l'accessibilité aux personnes à mobilité réduite.
Pour tout besoin spécifique (vue, audition...), veuillez nous contacter au 01 85 77 07 07.