

Certifiante

Formation Building Batch Data Analytics Solutions on AWS (BBDAS)

Réf. SF34457

Construisez des pipelines d'analytique batch performants sur AWS avec Amazon EMR, Spark et Hadoop !

1 j
Durée

850 € HT
Inter / participant

0
Sessions à venir

Objectifs pédagogiques

- ✓ Distinguer les apports des data warehouses, data lakes et des architectures data modernes.
- ✓ Concevoir et déployer une solution d'analytique batch de bout en bout.
- ✓ Optimiser le stockage des données (dont l'usage de la compression) selon les besoins.
- ✓ Choisir les approches adaptées pour ingérer, transformer et stocker les données.
- ✓ Sélectionner les bons paramètres EMR (types d'instances/nœuds, clusters, auto scaling, topologie réseau) selon le cas d'usage.
- ✓ Expliquer l'impact des choix de stockage et de traitement sur l'analyse et la visualisation des résultats.
- ✓ Mettre en place la sécurisation des données au repos et en transit.
- ✓ Surveiller les charges analytiques, diagnostiquer les incidents et appliquer les bonnes pratiques de pilotage des coûts.

Public visé & prérequis

Public visé Architectes, Ingénieurs Big Data

Prérequis Disposer idéalement d'au moins 1 an d'expérience sur des frameworks data open source (ex. Apache Spark ou Apache Hadoop). Recommandé : avoir suivi Architecting on AWS (AWSA) ou AWS Technical Essentials (AWSE). Recommandé : avoir suivi Building Data Lakes on AWS (BDLA) ou Getting Started with AWS Glue.

Programme détaillé

Module A : Panorama de l'analytique et du pipeline de données

Cas d'usage de l'analytique de données
Exploiter un pipeline data pour l'analyse

Module 1 : Prise en main d'Amazon EMR

Positionnement d'Amazon EMR dans des solutions analytiques
Architecture d'un cluster Amazon EMR
Démon : démarrer un cluster Amazon EMR
Stratégies de gestion et d'optimisation des coûts

Module 2 : Pipeline analytique avec Amazon EMR – Ingestion et stockage

Optimiser le stockage avec Amazon EMR
Techniques d'ingestion de données

Module 3 : Analytique batch haute performance avec Apache Spark sur Amazon EMR

Cas d'usage de Spark sur Amazon EMR
Pourquoi utiliser Spark sur Amazon EMR
Concepts clés de Spark
Démon : connexion à un cluster EMR et exécution de commandes Scala via le Spark shell
Transformation, traitement et analyse

Utilisation des notebooks avec Amazon EMR

TP : analytique à faible latence avec Apache Spark sur Amazon EMR

Module 4 : Traitement et analyse batch avec Amazon EMR et Apache Hive

Traiter des données batch avec Hive sur Amazon EMR

Transformation, traitement et analyse

TP : traitement batch avec Amazon EMR et Hive

Introduction à Apache HBase sur Amazon EMR

Module 5 : Traitement serverless des données

Traitement, transformation et analyse en mode serverless

Exploiter AWS Glue avec des workloads Amazon EMR

TP : orchestrer un traitement Spark avec AWS Step Functions

Module 6 : Sécurité et supervision des clusters Amazon EMR

Sécurisation des clusters EMR

Démo : chiffrement côté client avec EMRFS

Supervision et dépannage des clusters Amazon EMR

Démo : analyse de l'historique d'un cluster Apache Spark

Module 7 : Concevoir des solutions d'analytique batch

Cas d'usage d'analytique batch

Atelier : concevoir un workflow d'analytique batch

Module B : Développer des architectures data modernes sur AWS

Architectures data modernes sur AWS

Pourquoi choisir Sparks Formation ?

20 ans d'expertise IT, des formateurs experts et des sessions garanties contre l'annulation.

20 ans
d'expertise

1 068+
formations

11 178
apprenants formés/an

+67
Net Promoter Score

Avis des participants

Note moyenne des participants de l'ensemble de nos formations : 4,9/5 — 735 avis

Berenice P. — formatrice agréable à l'écoute bonne pédagogie j'ai beaucoup apprécié nos échanges

Anne-Lise C. — Très bonne adaptation du formateur au public et au niveau d'attente.

Ils nous font confiance

Cette année notre filière Cloud Computing a accompagné plus de 34 entreprises dont :

HARDIS GROUP SILAE EPSILON FRANCE GERFLOR SAS HARDIS SUPPLY CHAIN LEMONWAY MALAKOFF HUMANIS AMAP ACENSI SANA ALIAPUR ALTECA +24 autres

Financement & certifications

Cette formation est éligible aux financements OPCO et FNE-Formation. Nos équipes vous accompagnent dans le montage de votre dossier.



Réservez votre place dès maintenant

20 ans d'expertise · 1 000+ formations · NPS +67 · réponse sous 24h ouvrées

0 805 950 800

demande@sparks-formation.com