

Sessions garanties disponibles

Formation Spark

Réf. SF33388

Traitez les données massives en un temps record grâce au moteur d'analyse Apache Spark et à notre formation Spark !

3 j Durée	2 400 € HT Inter / participant	6 150 € HT Intra (≤12 pers.)	9 Sessions à venir
--------------	-----------------------------------	---------------------------------	-----------------------

Objectifs pédagogiques

- ✓ Expliquer le fonctionnement d'Apache Spark et ses relations avec le Big Data et Hadoop
- ✓ Utiliser les RDD pour gérer les opérations et MapReduce
- ✓ Intégrer et utiliser HDFS avec Spark
- ✓ Mettre en place et déployer une structure de cluster Spark
- ✓ Appliquer le partitionnement et la programmation parallèle sur les données HDFS
- ✓ Concevoir et générer une application Spark
- ✓ Implémenter Spark Streaming pour le traitement des données en temps réel
- ✓ Appliquer des algorithmes de Machine Learning avec MLlib et optimiser Spark

Public visé & prérequis

Public viséAdministrateurs de base de données, Administrateurs systèmes et réseaux, Architectes techniques, Data analysts, Data scientists, Développeurs, Ingénieurs Big Data

PrérequisConnaissances basiques en Java, Scala ou Python, ainsi que sur Apache Hadoop

Programme détaillé

Présentation de Apache Spark

Spark et le Big Data
Fonctionnement de Apache Spark
Spark et Hadoop: deux solutions différentes et complémentaires
Spark Shell
Environnement de travail
Outils

RDD (Resilient Distributed Datasets)

Fonctionnement
Cas d'usages
Gestion des opérations de RDD
MapReduce pour le RDD
Spark SQL

Combiner Spark à HDFS (Hadoop Distributed File System)

Intérêts
Cas d'usages
Intégration à Spark de HDFS
Utilisation de HDFS

Spark en cluster

Création d'une structure en clusters
Héberger la structure

Déployer la structure
Interface Web de Spark

Programmation parallèle, Partitionnement

Localiser les données du HDFS
Partitionnement de RDD
Programmation d'opérations parallèles
Exécution
Mise en cache du partitionnement
Gestion de la persistance des données

Conception d'une application à l'aide de Spark

Présentation de Spark
Configuration
Prototyper des opérations avec Spark Shell
Compiler
Générer l'application

Spark Streaming pour le traitement des données en temps réel

Fonctionnement
Principes de base
DStream
Intégrer, transformer des DStreams
Opérations de sortie
Gérer les performances

Implémenter des algorithmes, Machine Learning

Résoudre les problèmes à l'aide d'algorithmes itératifs
Machine Learning Library MLlib
Opérations sur les données graphiques

Optimiser Spark

Gérer les variables partagées
Données broadcastées
Accumulateurs
Optimiser les performances: méthodes et outils

Prochaines sessions

DATE	LIEU	STATUT
25/11/2026	Lyon	Sur confirmation
25/11/2026	à distance	Sur confirmation
25/11/2026	Grenoble	Sur confirmation
25/11/2026	Toulouse	Sur confirmation
25/11/2026	Aix-en-Provence	Sur confirmation
30/12/2026	Paris	Sur confirmation

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 la filière Data Engineering & Big Data : 4,9/5 — 13 avis

Arnaud M. — That was really useful for me, even though I am a PDA. The trainer was very patient with me (I needed more explanation as I was the only...

Thierry R. — Formation idéale pour appréhender la technologie nosql Bon support de formation. Bonne pédagogie

Votre formateur

Pierre P. — Formateur Machine Learning & Data Science

Pierre intervient sur des formations portant sur la Data Science et le Machine Learning. Parmi les formations qui lui sont confiées figurent notamment « Spark » et « Machine Learning...

Ils nous font confiance

Cette année notre filière Data Engineering & Big Data a accompagné plus de 73 entreprises dont :

SOPRA STERIA CAPGEMINI TECHNOLOGY SERVICES AMADEUS SAS GROUPAMA SUPPORTS & SERVICES AUBAY FRAIKIN FRANCE LBC FRANCE
ORANGE BUSINESS SERVICES SA ACATUS INFORMATIQUE ACENSI SANA +63 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