

Mis à jour le 20/11/2024

S'inscrire

Formation Samza

3 jours (21 heures)

Présentation

Samza est un moteur de traitement de données évolutif qui vous permet de traiter et d'analyser vos données en temps réel. Notre formation Samza va vous permettre de comprendre en profondeur et de simplifier le flux de vos données.

Dans cette formation pratique, vous apprendrez à gérer, mettre en place, administrer et configurer Samza.

Vous découvrirez comment utiliser Samza pour manipuler des pipelines de données, à maîtriser ses concepts, à utiliser les composants principaux, son architecture de base et à appeler ses API.

Au terme de cette formation, vous serez capable de configurer, d'optimiser et d'automatiser vos flux de données. Vous maîtriserez la configuration et le déploiement des API, l'optimisation des performances ainsi que la gestion de la latence et de la scalabilité de votre plateforme.

Comme pour toutes, la formation Samza vous sera présentée avec sa dernière version : [Samza 1.8.0](#).

Objectifs

- Comprendre l'architecture de Samza
- Comprendre le fonctionnement des flux de données
- Utiliser et comprendre les outils et les API
- **Sécuriser** les différents flux
- Comprendre le fonctionnement de Samza réussir à l'intégrer.

Public visé

- Data Scientists
- Développeurs
- Architectes
- Data analysts
- Data engineers

Pré-requis

- Bonne compréhension des bases de données et de la manipulation de fichiers structurés
- Expérience avec un langage de programmation, tel que Python ou Java
- Connaissances de base en traitement de données massives (Hadoop, Spark, etc.)
- Familiarité avec les concepts de stockage en colonnes et de compression

Programme de la formation Samza

INTRODUCTION A SAMZA

- Présentation de **Samza**?
- Avantages de l'utilisation de Samza pour traiter de grandes quantités de données
- Cas d'utilisation typiques
- Introduction au fonctionnement de Samza

PRÉSENTATION, CONCEPTS DE BASE

- Introduction à Samza
- Le flux et la partition
- **Les demandes de flux**
- Le traitement D'état
- Le temps de traitement et le temps d'un événement

ARCHITECTURE

- Modèle de threading et des commandes
- Gestion des états
- Tolérance aux pannes de l'état
- Distributed execution
 - Tâches
 - container
 - Coordinateur
- Différentes affinités de l'hôte

CONFIGURATION DE SAMZA

- Configuration d'application Samza

- Configuration d'application avancée
- Les différents points de contrôle, système et flux
- Configurations pour la consommation et la production sur Amazon Kinesis

PRÉSENTATION DE L'API

- Introduction des API fournies par Samza
- Concepts clés et application Samza
- Flux et table Descriptors
- Définir la logique du traitement de flux

API DE HAUT ET BAS NIVEAU

- Différence, définition des API et des exemples de code
- Concepts clés et opérateurs
- Sérialisation des données et application
- Samza TABLE API

DÉPLOIEMENT DE SAMZA

- Les options de déploiement
- Le fonctionnement sur YARN
- Intégration de Samza dans la bibliothèque et son exécution

LES DIFFÉRENTS CONNECTEURS DE SAMZA

- Présentation des connecteurs
- Apache Kafka
- Azure Event Hubs
- AWS Kinesis

Sociétés concernées

Cette formation s'adresse à la fois aux particuliers ainsi qu'aux entreprises, petites ou grandes, souhaitant former ses équipes à une nouvelle technologie informatique avancée ou bien à acquérir des connaissances métiers spécifiques ou des méthodes modernes.

Positionnement à l'entrée en formation

Le positionnement à l'entrée en formation respecte les critères qualité Qualiopi. Dès son inscription définitive, l'apprenant reçoit un questionnaire d'auto-évaluation nous permettant d'apprécier son niveau estimé sur différents types de technologies, ses attentes et objectifs personnels quant à la formation à venir, dans les limites imposées par le format sélectionné. Ce questionnaire nous permet également d'anticiper certaines difficultés de connexion ou de sécurité interne en entreprise (intraentreprise ou classe virtuelle) qui pourraient être

problématiques pour le suivi et le bon déroulement de la session de formation.

Méthodes pédagogiques

Stage Pratique : 60% Pratique, 40% Théorie. Support de la formation distribué au format numérique à tous les participants.

Organisation

Le cours alterne les apports théoriques du formateur soutenus par des exemples et des séances de réflexions, et de travail en groupe.

Validation

À la fin de la session, un questionnaire à choix multiples permet de vérifier l'acquisition correcte des compétences.

Sanction

Une attestation sera remise à chaque stagiaire qui aura suivi la totalité de la formation.