

Formation Apache Airflow

Présentation

Cette formation de 3 jours sur Apache Airflow vous permettra de maîtriser la supervision, la planification et la gestion de workflows de données complexes. À travers des sessions pratiques et théoriques, vous apprendrez à créer des workflows dynamiques et évolutifs, ainsi qu'à utiliser les fonctionnalités avancées d'Airflow. Destinée aux ingénieurs Big Data, développeurs et architectes techniques, cette formation vous donnera les compétences nécessaires pour optimiser vos pipelines de données.

Durée : 21,00 heures (3 jours)

Tarif INTRA : Nous consulter

Objectifs de la formation

- Découvrir les fonctionnalités d'Apache Airflow
- Concevoir, planifier et surveiller des workflows
- Comprendre le fonctionnement interne d'Airflow
- Apprendre à mettre à l'échelle Airflow pour gérer de nombreux pipelines
- Connaître les meilleures pratiques et les limitations d'Airflow

Prérequis

Connaissance de base en Python

Public

- Ingénieurs Big Data
- Développeurs
- Tech Leads
- Architectes Techniques

Programme de la formation



1.Introduction

- Présentation des objectifs de la formation
- Configuration de l'environnement de développement (installation des outils)

2.Découverte d'Apache Airflow

- Utilité et cas d'utilisation d'Airflow
- Concurrents d'Airflow
- Architecture d'Airflow
- Fonctionnement des workflows
- [Pratique] Exploration de l'interface utilisateur
- [Pratique] Utilisation de la CLI et de l'API REST

3.Créer son premier Workflow

- Définir un DAG (Directed Acyclic Graph)
- [Pratique] Création d'un premier DAG
- Types d'opérateurs
- [Pratique] Ajout des principaux opérateurs dans un DAG
- Dépendances entre opérateurs et composition
- [Pratique] Ajouter des dépendances à un DAG
- Fonctionnement du planificateur d'Airflow
- Définition d'un Workflow

4.Exécuteurs et leurs backends

- Types d'exécuteurs et backends
- Avantages et inconvénients de chaque type

5.Concepts dynamiques d'Airflow

- Minimiser les répétitions avec les Groupes de tâches
- [Pratique] Simplification d'un DAG avec les Groupes de tâches
- Utilisation des XCOMs pour échanger des données
- [Pratique] Création de tâches conditionnelles
- Gestion des SLAs pour les tâches
- [Pratique] Définition des SLAs
- Utilisation des variables d'Airflow
- [Pratique] Utilisation des variables dans un DAG

- Utilisation des connexions d'Airflow
- [Pratique] Utilisation des connexions dans un DAG
- Jinja Templating, Macros et Variables
- [Pratique] Ajout de données au run time avec les variables
- [Pratique] Utiliser les templates pour personnaliser les notifications
- Utilisation des Pools

6. Interactions avec le reste du monde via les Airflow providers

- Installation et utilisation des providers
- Choix des opérateurs en fonction des exécuteurs et de l'échelle
- Personnalisation d'un opérateur existant
- Quiz

7. Administration d'Airflow : Sécurité, RBAC, Métriques et Logging

- Sécurisation d'Airflow
- Création d'utilisateurs et authentification
- Contrôle d'accès basé sur les rôles (RBAC)

Organisation

Formateur

Les formateurs de Docaposte Institute sont des experts de leur domaine, disposant d'une expérience terrain qu'ils enrichissent continuellement. Leurs connaissances techniques et pédagogiques sont rigoureusement validées en amont par nos référents internes. Riches de leur expérience sur le sujet, ils sauront accompagner vos collaborateurs dans leur montée en compétence.

Moyens pédagogiques et techniques

- Apports des connaissances communes.
- Mises en situation sur le thème de la formation et des cas concrets.
- Méthodologie d'apprentissage attractive, interactive et participative.
- Equilibre théorie / pratique : 60 % / 40 %.
- Supports de cours fournis au format papier et/ou numérique.
- Ressources documentaires en ligne et références mises à disposition par le formateur.

- Pour les formations en présentiel dans les locaux mis à disposition, les apprenants sont accueillis dans une salle de cours équipée d'un réseau Wi-Fi, d'un tableau blanc ou paperboard. Un ordinateur avec les logiciels appropriés est mis à disposition (le cas échéant).

Dispositif de suivi de l'exécution et de l'évaluation des résultats de la formation

En amont de la formation

- Recueil des besoins des apprenants afin de disposer des informations essentielles au bon déroulé de la formation (profil, niveau, attentes particulières...).
- Auto-positionnement des apprenants afin de mesurer le niveau de départ.

Tout au long de la formation

- Évaluation continue des acquis avec des questions orales, des exercices, des QCM, des cas pratiques ou mises en situation...

A la fin de la formation

- Auto-positionnement des apprenants afin de mesurer l'acquisition des compétences.
- Evaluation par le formateur des compétences acquises par les apprenants.
- Questionnaire de satisfaction à chaud afin de recueillir la satisfaction des apprenants à l'issue de la formation.
- Questionnaire de satisfaction à froid afin d'évaluer les apports ancrés de la formation et leurs mises en application au quotidien.

Accessibilité

Nos formations peuvent être adaptées à certaines conditions de handicap. Nous contacter pour toute information et demande spécifique.