

Ingestion de données avec NiFi

L'objectif de cette formation est d'apprendre à utiliser Apache NiFi pour créer des flux de données personnalisés, des concepts à la mise en oeuvre en entreprise : introduction, différence entre données froides et données chaudes, outils et technologies Big Data, installation et architecture, approches de développement, outils de développement d'applications et état d'esprit, ETL. **Pré-requis** : Python, ETL.



Modalité : à distance

Formule au choix : Live Training+ *ou bien*
Classe virtuelle Teams ou Zoom

Durée totale : 14 H (2 jours)

PLAN DETAILLE

Fondamentaux de NiFi : #1

Flux, pipeline de data, ETL
Pourquoi utiliser NiFi?
Architecture, cas d'usages
Développement et flow management
Processors, FlowFile, Connectors
Composants, fonctionnalités
Installer, configurer NiFi
Prendre en main l'interface
Créer des flux de données

NiFi : concepts : #2

User Interface
Processors et Connections
NiFi Processor Groups
Controller services
Flowfile

Démarrer avec NiFi, Dataflow : #3

Configurer un processor : GetFile
Créer un flux de données, PutFile
Configurations et connexions
Générer des Flow Files
Attributs et variables
Prioritizers
Labels

Sécurité NiFi : #5

Security configuration
User authentication
NiFi et LDAP
HDF and HDP

Concepts avancés: #4

NiFi Registry
Dataflow templates
Expression Language, routage
Format des FlowFiles
Optimiser les flux de données
Management du cluster Edge
Debugging, gestion des erreurs
Provenance des données
Monitoring, rapports, logs