

Talend OpenStudio intégration de données

Référence : PYCB060

Durée : 3 jours

Certification : **Aucune**

CONNAISSANCES PREALABLES

- Maîtrise des SGBDR et de SQL. La connaissance de la programmation en Java serait utile..

PROFIL DES STAGIAIRES

- Développeurs, chefs de projet et toute personne souhaitant utiliser Talend OpenStudio pour le traitement de données .

OBJECTIFS

- Savoir créer des jobs dans l'application ETL Talend, et les optimiser par l'utilisation des contextes.. • Savoir exécuter les jobs et en suivre l'exécution par l'utilisation de statistiques..

CERTIFICATION PREPAREE

Aucune

METHODES PEDAGOGIQUES

- Mise à disposition d'un poste de travail par stagiaire
- Remise d'une documentation pédagogique papier ou numérique pendant le stage
- La formation est constituée d'apports théoriques, d'exercices pratiques, de réflexions et de retours d'expérience
- Le suivi de cette formation donne lieu à la signature d'une feuille d'émargement

FORMATEUR

Consultant-Formateur expert Bigdata

METHODE D'EVALUATION DES ACQUIS

- Auto-évaluation des acquis par le stagiaire via un questionnaire
- Attestation de fin de stage adressée avec la facture

CONTENU DU COURS

Introduction

- Rappels sur les solutions ETL.
- Présentation Talend OpenStudio : installation, configuration des préférences utilisateurs. Documentation.
- Concevoir des jobs simples avec Talend OpenStudio.

Modélisation

- Présentation des outils : Business Modeler, JobDesigner Mise en oeuvre des principales connexions.
- Intégration de fichiers XML et CSV
- Etude des composants de transformation.
- Analyse du code et exécution des jobs

Optimisation des jobs

- Utilisation des métadonnées, import/export, propagation sur les jobs, configuration de connexions réutilisables
- Stockage des variables de contexte dans les fichiers .properties et .ini

Liens avec les bases de données

- Présentation des bases de données supportées
- Opérations sur les tables,
- Connexion à un schéma de bases de données
- Gestion des transactions
- Utilisation de SQLBuilder pour créer des requêtes

Traitements de données multi-sources

- Le composant tMap
- Mise en oeuvre : Création de jointures, transformations à l'aide des variables, expressions et jointures, qualification des données à l'aide de filtres.
- Génération de sorties multiples

- Extensions : décomposition de jobs, mise en oeuvre du tRunJob, debugging, analyse des statistiques d'exécution, reporting avec JasperOutput