ETL traitements des BIGDATA

Niveau :
Initiation

Durée :
5 journées

Prérequis
Aucun

Moyens techniques
Théorie et pratique

Objectifs :
Talend est la solution d'intégration de données Open Source permettant de répondre efficacité

Programme de formation :

Introduction

- Qu'est qu'un ETL ?

- Présentation de la société Talend

Talend Open Studio

- Installation

- Découverte de l'environnement

- Premier pas avec TOS

- Vue d'ensemble des composants

Mon premier flux

- Le Business Model

- Le job

- Les bonnes pratiques

- tRowGenerator

- tFileInputExcel & tFileOutputExcel

- tMsgBox

- tSendMail

- Type de liaisons entre composants

- Exécuter un job

Traiter les données

- tFilterRow

- tUniqRow

- tAgregateRow

- tSortRow

- tMap

Les métadonnées

- Vue d'ensemble

- Créer une métadonnée Excel

- Utiliser une métadonnée Excel

- Créer une métadonnée MySQL

- Récupérer les tables d'une BDD

Travailler avec les bases de données

- tMysqlConnection

- tMysqlInput

- tMysqlOutput

- tMysqlCommit

Autres composants et fonctionnalités

- variables globales

- tBuffer

- tSystem

- tServerAlive

- tUnite

- tReplicate

Les itérations

- tFileList

- tInfiniteLoop

- tLoop

- tPOP

- tFlowToIterate

Gestion des logs et des erreurs

- tLogRow

- tDie & tLogCatcher

- tWarn

- tStatCatcher

- tFlowMeter & tFlowMeterCatcher

Fonctionnalités avancées

- Variables de contexte

- Créer une variable de contexte

- Utiliser une variable de contexte

- Modifier une variable de contexte

- Cascade de jobs

- Les routines

Automatisation & Documentation

- Documenter ses jobs

- Générer des scripts