

T

Programmes de formation

DataStage Enterprise Edition Essentials

DataStage (ETL proposé par IBM) est une solution d'importation, de manipulation et de mappages de données. Sa mise en œuvre permet d'effectuer des synchronisations massives d'informations d'une base de données vers une autre et d'alimenter ainsi les DataWarehouses.

Objectifs

*Devenir autonome sur DataStage Enterprise Edition Essentials
Concevoir et développer des traitements d'extraction, de nettoyage, de transformation et d'agrégation des données destinées à l'alimentation du DataWarehouse*

Public concerné

Développeurs débutants

Pré-requis

Connaissance des bases de données relationnelles
Maîtrise du langage de requêtes SQL

Programme

Présentation de DataStage

Architecture de DataStage
Composants de DataStage

Prise en main de module DataStage Administrator

Création du projet, maintenance et sauvegarde
Sécurité d'accès et partage des Projets

Prise en main du module DataStage Designer

Création jobs servers et jobs parallèles
Stage et liens
Les parallélismes
Principes de traitement parallèle
Les différentes formes de parallélisme
Utilisation de SGBD
Fonctions fournies avec DataStage
Stratégies de mise à jour
Chargement des données
Ordonnancement des jobs DataStage

Prise en main du module DataStage Manager

Création des catégories
Import et export de données du référentiel
Configuration du serveur
Création et compilé des Routines
Définition des sources et cibles physiques

Prise en main du module DataStage Director

Visualisation de l'exécution des traitements et des logs
Monitoring des jobs, tuning, reprise sur incident
Planifications des jobs

Best Practices

Rappel des justes pratiques d'utilisation de DataStage PX dans le cadre d'un projet impliquant un ETL