

Formation InfoSphere DataStage

Essentials v11.3

Objectifs de la formation InfoSphere DataStage v11.3

Cette formation apporte toutes les connaissances nécessaires aux développeurs ETL sur IBM DataStage V11.3.

Plus concrètement après avoir suivi ce stage vous saurez:

- Décrire l'utilisation de DataStage et du Workflow DataStage
- Décrire l'architecture Information Server et comment DataStage s'y intègre
- Décrire les options de déploiement de Information Server et DataStage
- Utiliser la console Web de Information Server et le client d'administration DataStage pour créer les utilisateurs et configurer l'environnement DataStage
- Importer et exporter les objets DataStage dans un fichier
- Importer des définitions de tables pour des fichiers séquentiels et des tables relationnelles
- Concevoir, compiler, activer et surveiller des tâches en parallèle dans DataStage
- Concevoir des tâches qui lisent et écrivent des fichiers séquentiels
- Décrire l'architecture de traitement parallèle de DataStage
- Concevoir les tâches qui combinent des données avec des jointures et des recherches
- Concevoir des tâches qui trient et rassemblent les données
- Mettre en oeuvre un business logique complexe avec DataStage Transformer
- Debugger les tâches DataStage avec DataStage PX Debugger
- Lire et écrire dans les tables des Bases de données avec DataStage ODBC et les connecteurs DB2
- Travailler avec les fonctions "Repository" telles que la recherche et l'analyse d'impact
- Créer des séquences de tâches qui contrôlent les ensembles de tâches

Capital Social: 50000 DT **MF:** 1425253/M/A/M/000 **RC:** B91211472015

Tél. / Fax.: +216 73362 100 **Email:** contact@itgate-training.com **Web:** www.itgate-training.com

Adresse : 12 Rue Abdelkadeur Daghrrir - Hammam Sousse 4011 – Tunisie

À qui s'adresse cette formation ?

Public :

Cette formation de base s'adresse aux administrateurs de projets et aux développeurs ETL, responsables des extractions et de la transformation des données avec DataStage.

Prérequis :

Avant de suivre ce cours, vous devez avoir des connaissances sur Windows et sur les techniques d'accès aux bases de données.

Contenu du cours InfoSphere DataStage v11.3

Introduction à DataStage

Lister et décrire les utilisations de DataStage

Lister et décrire les clients DataStage

Décrire le Workflow DataStage

Décrire les deux types de parallélisme possibles avec les tâches en parallèle dans DataStage

Déploiement

Identifier les composants de Information Server qui doivent être installés

Décrire en quoi consiste un déploiement

Décrire les différentes options de déploiement du domaine

Décrire le processus d'installation

Démarrer Information Server

Administration DataStage

Ouvrir la console Web

Créer des nouveaux utilisateurs et des groupes

Assigner des rôles Suite et Composant aux utilisateurs et aux groupes

Donner les accès DataStage aux utilisateurs

Capital Social: 50000 DT **MF:** 1425253/M/A/M/000 **RC:** B91211472015

Tél. / Fax.: +216 73362 100 **Email:** contact@itgate-training.com **Web:** www.itgate-training.com

Adresse : 12 Rue Abdelkadeur Daghri - Hammam Sousse 4011 – Tunisie

Connexion à DataStage en tant qu'administrateur

Ajouter un utilisateur DataStage dans l'onglet Permission et spécifier le rôle de l'utilisateur

Spécifier les projets par défaut

Lister et décrire les variables d'environnement importantes

Travailler avec les Metadata

Connexion à DataStage

Se déplacer dans DataStage Designer

Importer et exporter les objets DataStage dans un fichier

Importer une définition de table pour un fichier séquentiel

Créer des tâches parallèles

Concevoir des tâches parallèles dans DataStage

Définir le paramètre de la tâche

Utiliser le Row Generator, Peek et Annotation dans une tâche

Compiler les tâches

Activer les tâches

Surveiller les journaux des tâches

Créer un jeu de paramètres et l'utiliser dans une tâche

Accéder aux données séquentielles

Comprendre les étapes pour accéder aux différents types de fichiers de données

Lire et écrire des séquences de fichiers via Sequential File Stage

Lire et écrire des fichiers d'ensemble de données via les étapes DataStage

Créer des liens de rejet

Travailler avec les valeurs NULL dans les fichiers séquentiels

Lire des fichiers séquentiels multiples via des modèles de fichiers

Utiliser des lecteurs multiples

Capital Social: 50000 DT **MF:** 1425253/M/A/M/000 **RC:** B91211472015

Tél. / Fax.: +216 73362 100 **Email:** contact@itgate-training.com **Web:** www.itgate-training.com

Adresse : 12 Rue Abdelkadeur Daghri - Hammam Sousse 4011 – Tunisie

Partitionner et collecter

Décrire l'architecture du processus parallèle

Décrire le pipeline et la partition du parallélisme

Lister et décrire le partitionnement et les algorithmes de rassemblement

Décrire les fichiers de configuration

Décrire le processus de compilation des tâches parallèles

Expliquer OSH

Expliquer l'étendue

Combiner les données

Combiner les données via les étapes Lookup

Définir les « range lookups »

Combiner les données via la fusion

Combiner les données via la jointure

Combiner les données via Funnel

Etapes de traitement de groupes

Trier les données

Combiner l'étape d'agrégation

Combiner les données et supprimer les doublons

Etape de transformation

Utiliser l'étape de transformation dans les tâches parallèles

Définir les contraintes

Définir les dérivations

Utiliser les variables

Créer un jeu de paramètres et utiliser ses paramètres dans les contraintes et les dérivations

Capital Social: 50000 DT **MF:** 1425253/M/A/M/000 **RC:** B91211472015

Tél. / Fax.: +216 73362 100 **Email:** contact@itgate-training.com **Web:** www.itgate-training.com

Adresse : 12 Rue Abdelkadeur Daghri - Hammam Sousse 4011 – Tunisie

Fonctions Repository

Améliorer la recherche simple

Améliorer la recherche avancée

Améliorer l'analyse d'impact

Comparer les différences entre deux définitions de table

Comparer les différences entre deux tâches

Travailler avec des données relationnelles

Importer les définitions de tables pour les tables relationnelles

Créer les connexions de données

Utiliser ODBC et le connecteur DB2 dans une tâche

Utiliser SQL Builder pour définir SQL SELECT et INSERT

Utiliser les liens d'entrée multiples dans le connecteur pour mettre à jour les tables multiples dans une transaction simple

Créer des liens de rejets à partir du Connecteur pour capturer les lignes avec les erreurs SQL

Contrôle des tâches

Utiliser le séquenceur de tâches DataStage pour créer une tâche qui contrôle une séquence de tâches

Utiliser les liens du séquenceur et les étapes pour contrôler la séquence d'un jeu de tâches

Utiliser les déclencheurs et les étapes pour contrôler les conditions selon lesquelles les tâches sont activées

Transmettre des informations dans les paramètres des tâches à partir du master de contrôle des tâches

Prise en main des erreurs et des exceptions