

Formation InfoSphere Advanced

DataStage - Parallel Framework 11.3

(Cours officiel IBM)

Objectifs de la formation InfoSphere DataStage Avancé v11.3

Cette formation InfoSphere DataStage Parallel Framework 11.3 présente les techniques de développement de jobs parallèles avancés dans DataStage 11.3. Cette technique permet d'intégrer des données sur de multiples systèmes à haute performance et de prendre en charge la gestion étendue des métadonnées et la connectivité d'entreprise.

Les participants vont renforcer leur compréhension de l'architecture de DataStage, y compris du développement de DataStage et des environnements de runtime. Ils seront ainsi en mesure de concevoir des jobs parallèles robustes, moins sujets à des erreurs, réutilisables, et optimisés pour de meilleures performances.

A la suite de cette formation InfoSphere DataStage, vous aurez acquis les capacités nécessaires pour:

- Comprendre l'architecture parallèle
- Comprendre le parallélisme du pipeline et de la partition
- Comprendre les rôles et les principaux éléments des fichiers de configuration de DataStage
- Comprendre le processus de compilation et comment il est représenté dans l'OHS
- Comprendre le processus d'exécution du runtime job et comment il est représenté dans le Score
- Comprendre comment fonctionne le partitionnement et la collecte de données dans la structure parallèle

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

- Lister et sélectionner les algorithmes de partitionnement et de collecte
- Comprendre le tri dans une structure parallèle
- Comprendre les techniques d'optimisation de tri et de partition
- Comprendre la fonction de buffer dans la structure parallèle
- Comprendre les techniques d'optimisation des buffers
- Comprendre et travailler avec les différents types de données et éléments de la structure parallèle, tels que les éléments de virtualisation des données et les schémas
- Comprendre le fonctionnement et l'utilisation de la Runtime Column Propagation (RCP)
- Créer des composants réutilisables à l'aide de conteneurs partagés
- Comprendre le fonctionnement et l'utilisation de l'optimisation répartie
- Optimiser les travaux parallèles DataStage en utilisant la répartition.

À qui s'adresse cette formation ?

Public :

Ce cours s'adresse aux développeurs expérimentés sur DataStage souhaitant comprendre l'architecture parallèle et progresser dans la technique de jobs.

Prérequis :

Il est recommandé d'avoir suivi la formation InfoSphere DataStage Essentials v11.3 (KM203G) ou de posséder des connaissances équivalentes et de posséder au moins un an d'expérience dans le développement de jobs parallèles avec DataStage.

Contenu du cours InfoSphere DataStage Avancé v11.3

Introduction à l'architecture Parallel Framework

Compilation et exécution

Partitionnement et collecte des données

Tri des données

Bufferisation dans les jobs parallèles

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

Parallel Framework Data Types

Composants réutilisables

Optimiser avec la répartition

Travaux Pratiques

De nombreux exercices pratiques permettront de mettre en oeuvre et d'assimiler les notions abordées.