

Travaux pratiques : mise en place d'une zone de préparation

Temps estimé nécessaire : 30 minutes

Objectif du laboratoire :

L'objectif de ce laboratoire est de vous fournir des compétences pratiques pour configurer et gérer un serveur de préparation pour un entrepôt de données, en particulier à l'aide de PostgreSQL. Le laboratoire se concentre sur l'enseignement de la conception et de la mise en œuvre d'un schéma de base de données, du chargement de données dans des tables et de l'exécution d'exemples de requêtes pour interagir avec les données. L'objectif est de vous fournir une compréhension pratique des subtilités impliquées dans la préparation et la gestion d'un environnement d'entrepôt de données.

Avantages de l'apprentissage en laboratoire :

Le principal avantage de ce laboratoire est l'expérience pratique qu'il offre en matière de concepts et de pratiques d'entreposage de données. En travaillant dans l'IDE Skills Network Cloud, un environnement interactif, vous pouvez mettre en pratique et perfectionner vos compétences en gestion de base de données. Cette exposition pratique est essentielle pour développer une base solide en gestion et analyse de données, des compétences très appréciées dans des scénarios commerciaux réels. Le laboratoire vous aide à comprendre les aspects opérationnels des entrepôts de données, vous permettant ainsi de mieux vous préparer aux défis de la gestion et de l'analyse des données dans les environnements professionnels.

Objectifs

Dans ce laboratoire, vous allez :

- Configurer un serveur de préparation pour un entrepôt de données
- Créer le schéma pour stocker les données
- Charger les données dans les tables
- Exécuter un exemple de requête

À propos de Skills Network Cloud IDE

L'IDE Cloud de Skills Network (basé sur Theia et Docker) fournit un environnement pour les travaux pratiques liés aux cours et aux projets. Theia est un IDE open source (environnement de développement intégré), qui peut être exécuté sur un ordinateur de bureau ou sur le cloud. Pour réaliser ce laboratoire, nous utiliserons l'IDE Cloud basé sur Theia exécuté dans un conteneur Docker.

Avis important concernant cet environnement de laboratoire

Veuillez noter que les sessions de cet environnement de laboratoire ne sont pas permanentes. Un nouvel environnement est créé pour vous à chaque fois que vous vous connectez à ce laboratoire. Toutes les données que vous avez pu enregistrer lors d'une session précédente seront perdues. Pour éviter de perdre vos données, prévoyez de terminer ces laboratoires en une seule session.

Exercice 1 – Démarrer le serveur PostgreSQL

Nous utiliserons le serveur PostgreSQL comme serveur de préparation.

Démarrez le serveur PostgreSQL à partir de la boîte à outils SN comme indiqué dans l'image ci-dessous.

Exercice 2 – Créer une base de données

Créer la base de données sur l'entrepôt de données.

Ouvrez maintenant un nouveau terminal en cliquant sur la barre de menu et en sélectionnant Terminal->Nouveau terminal. Cela ouvrira un nouveau terminal en bas de l'écran.

Exécutez les commandes ci-dessous sur le terminal nouvellement ouvert. (Vous pouvez copier le code en cliquant sur le petit bouton de copie en bas à droite du bloc de code ci-dessous, puis le coller où vous le souhaitez.)

En utilisant la commande `createdb` du serveur PostgreSQL, nous pouvons directement créer la base de données depuis le terminal. Avant cela, exportez votre mot de passe en remplaçant `<votre_mot_de_passe>` par le mot de passe de votre serveur PostgreSQL (que vous trouverez sous l'onglet des informations de connexion) dans la commande ci-dessous et exécutez-la.

```
export PGPASSWORD=<your_password>;
```

Exécutez ensuite la commande ci-dessous qui créera une base de données nommée `billingDW`.

```
createdb -h postgres -U postgres -p 5432 billingDW
```

Dans la commande ci-dessus

- `-h` mentionne que le serveur de base de données est accessible en utilisant le nom d'hôte « `postgres` »
- `-U` mentionne que nous utilisons le nom d'utilisateur `postgres` pour nous connecter à la base de données
- `-p` mentionne que le serveur de base de données fonctionne sur le port numéro 5432

Vous devriez voir un résultat comme celui-ci.

Exercice 3 – Créer un schéma d'entrepôt de données

Étape 1 :

Téléchargez les fichiers de schéma.

Les commandes pour créer le schéma sont disponibles dans le fichier ci-dessous.

<https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBM-DB0260EN-SkillsNetwork/labs/Setting%20up%20a%20staging%20area/billing-datawarehouse.tgz>

Exécutez les commandes ci-dessous pour télécharger et extraire les fichiers de schéma.

```
wget https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBM-DB0260EN-SkillsNetwork/labs/Setting%20up%20a%20staging%20area/billing-datawarehouse.tgz
tar -xvzf billing-datawarehouse.tgz
ls *.sql
```

Vous devriez voir 4 fichiers .sql répertoriés dans la sortie

Étape 2 : Créer le schéma

Exécutez la commande ci-dessous pour créer le schéma dans la billingDWbase de données.

```
psql -h postgres -U postgres -p 5432 billingDW < star-schema.sql
```

Vous devriez voir un résultat similaire à celui ci-dessous.

Exercice 4 – Charger des données dans des tables de dimension

Lorsque nous chargeons des données dans les tables, il est recommandé de charger d’abord les données dans les tables de dimension.

Étape 1 : charger les données dans la table DimCustomer

Exécutez la commande ci-dessous pour charger les données dans la table DimCustomer de billingDWla base de données.

```
psql -h postgres -U postgres -p 5432 billingDW < DimCustomer.sql
```

Étape 2 : Charger les données dans la table DimMonth

Exécutez la commande ci-dessous pour charger les données dans la table DimMonth de billingDWla base de données.

```
psql -h postgres -U postgres -p 5432 billingDW < DimMonth.sql
```

Exercice 5 – Charger des données dans la table de faits

Charger les données dans la table FactBilling

Exécutez la commande ci-dessous pour charger les données dans la table FactBilling de billingDWla base de données.

```
psql -h postgres -U postgres -p 5432 billingDW < FactBilling.sql
```

Exercice 6 – Exécuter un exemple de requête

Exécutez la commande ci-dessous pour vérifier le nombre de lignes dans toutes les tables de la billingDWbase de données.

```
psql -h postgres -U postgres -p 5432 billingDW < verify.sql
```

Vous devriez voir un résultat similaire à celui ci-dessous.

Votre zone de préparation d’entrepôt de données est maintenant prête.

Exercices pratiques

Dans cette séance pratique, vous allez créer une base de données nommée practiceet y charger les données.

Si vous utilisez un nouveau terminal pour cet exercice pratique, n'oubliez pas d'exporter à nouveau le mot de passe, comme vous l'avez fait dans les étapes précédentes.

1. Problème:

Créez une base de données nommée practice.

▼ Cliquez ici pour un indice

utiliser la commande createdb

▼ Cliquez ici pour la solution

```
createdb -h postgres -U postgres -p 5432 practice
```

2. Problème:

Dans la practicebase de données, créez un schéma en utilisant star-schema.sql.

▼ Cliquez ici pour un indice

Utilisez la commande psql, assurez-vous que vous définissez la base de données sur practice.

▼ Cliquez ici pour la solution

```
psql -h postgres -U postgres -p 5432 practice < star-schema.sql
```

3. Problème:

Dans la practicebase de données, chargez les données dans toutes les tables à l'aide de DimMonth.sql, DimCustomer.sql et FactBilling.sql.

▼ Cliquez ici pour un indice

Utilisez la commande psql, assurez-vous que vous définissez la base de données sur practice.

▼ Cliquez ici pour la solution

```
psql -h postgres -U postgres -p 5432 practice < DimMonth.sql
psql -h postgres -U postgres -p 5432 practice < DimCustomer.sql
psql -h postgres -U postgres -p 5432 practice < FactBilling.sql
```

4. Problème:

_Vérifiez que vous avez correctement chargé les données dans la practicebase de données.

▼ Cliquez ici pour un indice

Utilisez verify.sql et exécutez-le sur la practicebase de données.

▼ Cliquez ici pour la solution

```
psql -h postgres -U postgres -p 5432 practice < verify.sql
```

Félicitations !! Vous avez terminé avec succès le laboratoire de configuration d'un serveur de préparation.

Auteurs

Ramesh Sannareddy

Autres contributeurs

Rav Ahuja

© IBM Corporation 2023. Tous droits réservés.