

Travaux pratiques : Remplissage d'un entrepôt de données

Temps estimé nécessaire : 15 minutes

Objectifs

Dans ce laboratoire, vous allez :

- Créer une instance d'IBM DB2 sur le cloud
- Créer des informations d'identification pour l'accessibilité externe
- Créer un DSN db2cli
- Vérifier un DSN db2cli
- Créer le schéma sur l'entrepôt de données de production
- Remplir l'entrepôt de données de production
- Travailler avec la ligne de commande interactive db2cli

À propos de Skills Network Cloud IDE

Skills Network Cloud IDE (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 terminer 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 – Créer une instance d'IBM DB2 sur le cloud

Nous utiliserons l'instance cloud d'IBM DB2 comme entrepôt de données de production dans ce laboratoire.

Si vous ne disposez pas d'une instance d'IBM DB2 sur le cloud, suivez les instructions de ce [laboratoire](#) pour en créer une.

Exercice 2 – Créer des informations d'identification de service

Pour accéder à votre instance cloud IBM DB2 à partir de programmes externes, vous avez besoin d'informations d'identification de service.

Si vous ne disposez pas d'informations d'identification de service, suivez les instructions de ce [laboratoire](#) pour créer vos informations d'identification de service.

Prenez note des détails suivants :

- utilisateur
- mot de passe
- hôte
- port
- nom de la base de données

Vous en aurez besoin plus tard dans ce laboratoire.

Exercice 3 – Créer un DSN db2cli

Vous pouvez accéder à l'instance cloud IBM DB2 à l'aide de l'interface utilisateur du navigateur Web.

En utilisant le, db2cli vous pouvez accéder à votre instance cloud IBM DB2 à partir de la ligne de commande.

db2cli peut être très utile pour automatiser les tâches de chargement de données.

Dans cet exercice, nous allons créer un DSN (nom de source de données). En bref, un DSN est un nom simple qui fait référence à une source de données.

La création d'un DSN est un processus en deux étapes.

Étape 1 : nous ajoutons la base de données, l'hôte, le port et les détails du mode de sécurité. Un exemple de commande est fourni pour votre référence ci-dessous :

```
db2cli writecfg add -database dbname -host hostname -port 50001 -parameter "SecurityTransportMode=SSL"
```

Étape 2 : nous donnons un nom à la source de données que nous venons de créer. Ce nom DSN nous aide à pointer facilement vers l'instance IBM DB2. Un exemple de commande est fourni pour votre référence ci-dessous.

```
db2cli writecfg add -dsn dsn_name -database dbname -host hostname -port 50001
```

Exécutez les commandes ci-dessous sur le terminal pour créer un DSN nommé production. Assurez-vous d'utiliser le nom de la base de données, l'hôte et les détails du port que vous avez notés dans l'exercice 2.

```
db2cli writecfg add -database BLUDB -host 0c77d6f2-5da9-48a9-81f8-86b520b87518.bs2io90l08kqb1od8lcg.databases.appdomain.cloud -port 50001
db2cli writecfg add -dsn production -database BLUDB -host 0c77d6f2-5da9-48a9-81f8-86b520b87518.bs2io90l08kqb1od8lcg.databases.appdomain.cloud
```

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

Exercice 4 – Vérifier un DSN db2cli

Maintenant que le DSN est créé, nous devons vérifier s'il fonctionne avant de commencer à l'utiliser.

La syntaxe générique de la commande permettant de vérifier le DSN est donnée ci-dessous :

```
db2cli validate -dsn alias -connect -user userid -passwd password
```

Exécutez la commande ci-dessous pour vérifier le productionDSN. Assurez-vous d'utiliser le nom d'utilisateur et le mot de passe que vous avez notés dans l'exercice 2.

```
db2cli validate -dsn production -connect -user jrg38634 -passwd SuWySBe5Y4MsYnh9
```

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

Votre DSN est validé. Vous pouvez désormais l'utiliser pour accéder à l'instance cloud IBM DB2.

Exercice 5 – Créer le schéma sur l'entrepôt de données de production

Étape 1 : Téléchargez le fichier de schéma.

Exécutez la commande ci-dessous pour télécharger le fichier de schéma.

```
wget https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBM-DB0260EN-SkillsNetwork/labs/Populating%20a%20Data%20Ware
```

Étape 2 : créer le schéma.

Exécutez la commande ci-dessous pour créer le schéma sur l'entrepôt de données de production. Assurez-vous d'utiliser votre nom d'utilisateur et votre mot de passe que vous avez notés dans l'exercice 2.

```
db2cli execsql -dsn production -user jrg38634 -passwd SuWySBe5Y4MsYnh9 -inputsq1 star-schema.sql
```

La commande ci-dessus indique à db2cli d'exécuter les commandes du fichier star-schema.sql sur l'entrepôt de données de production.

Exercice 6 – Remplir l'entrepôt de données de production

Étape 1 : Téléchargez les fichiers de données.

Exécutez les commandes ci-dessous pour télécharger les fichiers de données.

```
wget https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBM-DB0260EN-SkillsNetwork/labs/Populating%20a%20Data%20Ware
wget https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBM-DB0260EN-SkillsNetwork/labs/Populating%20a%20Data%20Ware
wget https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBM-DB0260EN-SkillsNetwork/labs/Populating%20a%20Data%20Ware
ls *.sql
```

Étape 2 : Chargez les données dans l'entrepôt de données.

Exécutez les commandes ci-dessous pour charger les données dans l'entrepôt de données de production. Assurez-vous d'utiliser le nom d'utilisateur et le mot de passe que vous avez notés dans l'exercice 2.

```
db2cli execsql -dsn production -user jrg38634 -passwd SuWySBe5Y4MsYnh9 -inputsq1 DimCustomer.sql
db2cli execsql -dsn production -user jrg38634 -passwd SuWySBe5Y4MsYnh9 -inputsq1 DimMonth.sql
db2cli execsql -dsn production -user jrg38634 -passwd SuWySBe5Y4MsYnh9 -inputsq1 FactBilling.sql
```

Exercice 7 – Vérifier les données de l'entrepôt de données de production

Étape 1 : Téléchargez le fichier SQL de vérification.

Exécutez la commande ci-dessous pour télécharger le fichier SQL afin de vérifier les données.

```
wget https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBM-DB0260EN-SkillsNetwork/labs/Populating%20a%20Data%20Ware
```

Étape 2 : Vérifiez les données dans l'entrepôt de données.

Exécutez la commande ci-dessous pour vérifier les données de l'entrepôt de données de production. Assurez-vous d'utiliser le nom d'utilisateur et le mot de passe que vous avez notés dans l'exercice 2.

```
db2cli execsql -dsn production -user jrg38634 -passwd SuWySBe5Y4MsYnh9 -inputsq1 verify.sql
```

Vous avez chargé les données avec succès, si vous voyez une sortie similaire à celle ci-dessous.

Exercice 8 – Travailler avec la ligne de commande interactive db2cli

db2cli peut également être utilisé de manière interactive.

Exécutez la commande ci-dessous pour ouvrir un shell de commande SQL interactif sur votre entrepôt de données de production. Assurez-vous d'utiliser votre nom d'utilisateur et votre mot de passe que vous avez notés dans l'exercice 2.

```
db2cli execsql -dsn production -user jrg38634 -passwd SuWySBe5Y4MsYnh9
```

Exécutez la commande ci-dessous sur db2cli.

```
select count(*) from DimMonth;
```

Vous devriez voir un résultat comme celui montré dans l'image ci-dessous.

Nous vous encourageons à exécuter davantage de requêtes SQL. Une fois terminé, tapez quit pour quitter db2cli.

Exercices pratiques

1. Problème:

À l'aide du shell interactif db2cli, recherchez le nombre de lignes dans la table FactBilling

- Cliquez ici pour un indice
- Cliquez ici pour la solution

2. Problème:

À l'aide de l'interface utilisateur Cloud (et non de db2cli), créez un MQT simple nommé avg_customer_bill avec les champs customerid et averagebillamount.

- ▼ Cliquez ici pour un indice

utiliser la create table commande

- ▼ Cliquez ici pour la solution

Accédez à l'interface utilisateur de DB2, accédez à l'écran Exécuter SQL, dans l'éditeur, copiez la commande suivante :

```
CREATE TABLE avg_customer_bill (customerid, averagebillamount) AS
(select customerid, avg(billedamount)
from factbilling
group by customerid
)
DATA INITIALLY DEFERRED
REFRESH DEFERRED
MAINTAINED BY SYSTEM;
```

Cliquez sur le bouton « Exécuter tout » pour exécuter l'instruction. Vous devriez voir le statut « Succès » dans la section Résultat.

3. Problème:

Actualiser le MQT nouvellement créé

- ▼ Cliquez ici pour un indice

utiliser la refresh table commande

- ▼ Cliquez ici pour la solution

À l'invite db2cli, exécutez la commande suivante :

```
refresh table avg_customer_bill;
```

4. Problème:

À l'aide du MQT nouvellement créé, recherchez les clients dont la facturation moyenne est supérieure à 11 000.

- Cliquez ici pour un indice
- ▼ Cliquez ici pour la solution

À l'invite db2cli, exécutez la commande suivante :

```
select * from avg_customer_bill where averagebillamount > 11000;
```

Félicitations ! Vous avez terminé avec succès le laboratoire Remplissage d'un entrepôt de données.

Auteurs

Ramesh Sannareddy

Autres contributeurs

Rav Ahuja

© IBM Corporation 2023. Tous droits réservés.