### **QCM de Préparation - dbt Cloud Architect**

#### **Thème 1 : Configuration des connexions à l'entrepôt de données**

1. **Lors de la configuration d'une connexion à un entrepôt de données dans dbt Cloud, quelle est la première étape à réaliser après avoir entré les informations d'identification ?**

A. Configurer la liste blanche d'IP (IP Whitelist).

B. Lier le projet à un dépôt Git.

C. Tester la connexion pour s'assurer que dbt Cloud peut communiquer avec l'entrepôt.

D. Ajouter un Client ID pour OAuth.  
**Réponse correcte : C**

1. **À quoi servent le Client ID et le Client Secret lors de la configuration de l'authentification OAuth pour une connexion à un entrepôt ?**

A. Ils permettent à dbt Cloud de s'authentifier auprès de votre fournisseur Git.

B. Ils authentifient l'application dbt Cloud elle-même auprès du fournisseur d'identité de l'entrepôt, permettant aux utilisateurs de s'authentifier via OAuth.

C. Ils servent à chiffrer les données en transit entre dbt Cloud et l'entrepôt.

D. Ils sont utilisés comme informations d'identification pour un compte de service.  
**Réponse correcte : B**

1. **Pourquoi la configuration de la liste blanche d'adresses IP (IP Whitelisting) est-elle une étape de sécurité cruciale pour de nombreuses organisations ?**

A. Pour garantir que seuls les développeurs autorisés peuvent modifier le code dbt.

B. Pour s'assurer que seul le trafic provenant des serveurs de dbt Cloud peut atteindre l'entrepôt de données, bloquant ainsi tout autre accès externe non autorisé.

C. Pour limiter le nombre de threads qu'un job dbt peut utiliser.

D. Pour activer l'authentification unique (SSO).  
**Réponse correcte : B**

#### **Thème 2 : Configuration des connexions Git**

1. **Quelle est la méthode recommandée pour intégrer un dépôt GitHub avec dbt Cloud ?**

A. Utiliser une clé SSH personnelle.

B. Utiliser un jeton d'accès personnel (Personal Access Token).

C. Installer l'application GitHub dbt Cloud.

D. Cloner le dépôt manuellement dans l'IDE.  
**Réponse correcte : C**

1. **Que se passe-t-il si la connexion au dépôt Git est rompue ?**

A. Les jobs planifiés continueront de s'exécuter avec la dernière version du code extraite.

B. Les développeurs ne pourront plus ouvrir l'IDE, commiter du code, ou lancer des exécutions basées sur les dernières modifications.

C. La connexion à l'entrepôt de données sera automatiquement désactivée.

D. Les utilisateurs perdront leur accès à dbt Cloud.  
**Réponse correcte : B**

1. **Lors de la connexion d'un dépôt Git externe (par exemple, GitLab, Bitbucket), quelle information est généralement requise en plus de l'URL du dépôt ?**

A. Un webhook pour les notifications.

B. Le nom de la branche par défaut uniquement.

C. Une clé de déploiement (Deploy Key) ou un jeton d'accès pour autoriser l'accès en lecture au dépôt.

D. L'adresse e-mail de l'administrateur du dépôt.  
**Réponse correcte : C**

#### **Thème 3 : Création et maintenance des environnements**

1. **Dans quel scénario est-il préférable d'utiliser un compte de service (Service Account) pour un environnement de déploiement ?**

A. Lorsque les développeurs ont besoin de lancer des exécutions dans leur environnement de développement personnel.

B. Pour les jobs de production ou de CI/CD, afin que les exécutions ne soient pas liées aux informations d'identification d'un utilisateur individuel.

C. Pour accorder un accès temporaire à des consultants externes.

D. Pour l'authentification à l'interface de dbt Cloud.  
**Réponse correcte : B**

1. **Comment dbt gère-t-il les variables d'environnement si une même variable est définie à plusieurs endroits (ex: dans l'environnement et dans la définition du job) ?**

A. La variable de l'environnement a toujours la priorité.

B. La variable définie dans la configuration du job (Job override) a la priorité sur celle de l'environnement.

C. dbt demande à l'utilisateur de choisir laquelle utiliser à chaque exécution.

D. La première variable trouvée est utilisée.  
**Réponse correcte : B**

1. **À quoi sert la configuration d'une "custom branch" dans un environnement de déploiement ?**

A. À permettre aux développeurs de choisir n'importe quelle branche pour exécuter un job.

B. À lier de manière permanente un environnement (ex: Production) à une branche Git spécifique et stable (ex: main ou prod).

C. À créer automatiquement de nouvelles branches pour chaque nouveau job.

D. À tester les fonctionnalités de dbt Cloud.  
**Réponse correcte : B**

#### **Thème 4 : Création et maintenance des définitions de jobs**

1. **Qu'est-ce que le "job chaining" (enchaînement de jobs) dans dbt Cloud ?**

A. La possibilité d'exécuter plusieurs commandes dbt dans un seul job.

B. La configuration d'un job pour qu'il se déclenche automatiquement après la réussite d'un autre job.

C. L'utilisation de la commande dbt chain dans l'IDE.

D. Une méthode pour lier des modèles entre différents projets.  
**Réponse correcte : B**

1. **Dans quel ordre les commandes dbt seed, dbt run, et dbt test devraient-elles être implémentées dans un job de déploiement standard ?**

A. dbt run, dbt test, dbt seed

B. dbt test, dbt run, dbt seed

C. dbt seed, dbt run, dbt test

D. L'ordre n'a pas d'importance.  
**Réponse correcte : C**

1. **Quel est l'intérêt de spécifier une version de dbt (dbt version override) dans les paramètres d'un job ?**

A. Pour s'assurer que tous les développeurs utilisent la même version de dbt localement.

B. Pour garantir la stabilité et la reproductibilité d'un job en le liant à une version spécifique de dbt Core, même si l'environnement est mis à jour.

C. Pour accéder à des fonctionnalités bêta de dbt Core.

D. C'est une obligation pour tous les jobs de production.  
**Réponse correcte : B**

#### **Thème 5 : Configuration de la sécurité et des licences**

1. **Quelle fonctionnalité de dbt Cloud Enterprise permet de définir des permissions granulaires basées sur des groupes d'utilisateurs et des ensembles de permissions ?**

A. Les jetons de service (Service Tokens).

B. Le contrôle d'accès basé sur les rôles (RBAC - Role-Based Access Control).

C. Le mappage de licences (License Mappings).

D. L'authentification unique (SSO).  
**Réponse correcte : B**

1. **À quoi sert la configuration du mappage de licences (License Mappings) dans un compte dbt Cloud Enterprise ?**

A. À acheter de nouvelles licences dbt Cloud.

B. À attribuer automatiquement des licences (Developer ou Read-Only) aux utilisateurs qui se connectent via SSO en fonction de leurs groupes d'appartenance dans le fournisseur d'identité.

C. À définir les permissions pour les jetons de service.

D. À lier une licence à un projet spécifique.  
**Réponse correcte : B**

1. **Un jeton de service (Service Token) a été créé avec un ensemble de permissions "Job Admin". Que peut faire une application utilisant ce jeton ?**

A. Uniquement voir le statut des jobs.

B. Déclencher des jobs existants mais pas les modifier.

C. Créer, modifier et déclencher des jobs dans les projets autorisés.

D. Gérer les utilisateurs et les licences du compte.  
**Réponse correcte : C**

#### **Thème 6 : Mise en place du monitoring et des alertes**

1. **Pour une intégration avec un outil d'alerte comme PagerDuty ou pour envoyer des notifications personnalisées sur Slack, quelle est la méthode la plus flexible et recommandée ?**

A. Configurer les notifications par e-mail vers une adresse PagerDuty ou Slack.

B. Utiliser des Webhooks pour envoyer une charge utile (payload) JSON détaillée à un point de terminaison (endpoint) spécifique.

C. Consulter manuellement l'historique des exécutions.

D. Écrire un script personnalisé qui interroge l'API de dbt Cloud toutes les minutes.  
**Réponse correcte : B**

1. **Par défaut, quand dbt Cloud envoie-t-il une notification par e-mail pour un job ?**

A. Uniquement lorsque le job réussit.

B. À chaque fois que le job se termine, quel que soit son statut.

C. Uniquement lorsque le job échoue ou est annulé.

D. Uniquement pour les jobs qui contiennent des tests de données.  
**Réponse correcte : C**

1. **Quelle information N'EST PAS incluse dans la charge utile (payload) d'un webhook de job ?**

A. L'ID du job et l'ID de l'exécution (Run ID).

B. Le statut final de l'exécution (par exemple, "Success", "Error").

C. Les logs détaillés de chaque étape de la commande dbt.

D. Un lien vers la page de l'exécution dans dbt Cloud.  
**Réponse correcte : C** (Les logs complets ne sont généralement pas dans le payload initial pour des raisons de taille, mais peuvent être récupérés via l'API).

#### **Thème 7 : Mise en place d'un dbt Mesh et références inter-projets**

1. **Comment un projet dbt (Projet A) peut-il référencer un modèle public d'un autre projet (Projet B) au sein de la même organisation dbt Cloud ?**

A. En important le Projet B comme un package dans le fichier packages.yml du Projet A.

B. En utilisant la fonction source() pour pointer vers la table du modèle du Projet B.

C. En configurant une dépendance de projet dans dbt\_project.yml du Projet A, puis en utilisant la fonction ref() normalement.

D. En copiant le code SQL du modèle du Projet B dans le Projet A.  
**Réponse correcte : C**

1. **Comment un architecte peut-il contrôler quels modèles d'un projet peuvent être référencés par d'autres projets dans un dbt Mesh ?**

A. En utilisant des variables d'environnement pour cacher certains modèles.

B. En définissant le modificateur d'accès (access) d'un modèle à public. Les modèles private (par défaut) ou protected ne peuvent pas être référencés par d'autres projets.

C. En créant des permissions RBAC spécifiques pour chaque modèle.

D. En plaçant les modèles publics dans un sous-dossier spécial nommé /public.  
**Réponse correcte : B**

1. **Quel est le prérequis pour qu'une référence inter-projet (cross-project ref) fonctionne dans un job ?**

A. Les deux projets doivent utiliser la même connexion à l'entrepôt de données.

B. Le job doit avoir le deferral (report) activé, pointant vers un environnement où le projet dépendant a déjà été exécuté avec succès.

C. Les deux projets doivent partager le même dépôt Git.

D. L'option "cross-project references" doit être activée manuellement par le support dbt.  
**Réponse correcte : B**

#### **Thème 8 : Configuration et utilisation de dbt Explorer**

1. **Lequel des cas d'utilisation suivants est un objectif principal de dbt Explorer ?**

A. Écrire et tester du code SQL pour de nouveaux modèles.

B. Comprendre le lignage complet des données (lineage), identifier les dépendances critiques et visualiser l'impact des changements.

C. Gérer les licences et les permissions des utilisateurs.

D. Déclencher et surveiller les exécutions de jobs en temps réel.  
**Réponse correcte : B**

1. **Comment dbt Explorer peut-il aider à optimiser les coûts et les performances ?**

A. En compressant automatiquement les tables dans l'entrepôt.

B. En identifiant les modèles complexes (avec de nombreuses dépendances en amont) ou les chemins critiques qui pourraient être des goulots d'étranglement lors des exécutions.

C. En fournissant un profileur de requêtes SQL intégré.

D. En supprimant automatiquement les modèles non utilisés.  
**Réponse correcte : B**

1. **Un développeur souhaite utiliser un modèle d'un autre projet. Comment peut-il utiliser dbt Explorer pour faciliter cela ?**

A. En exécutant le modèle de l'autre projet directement depuis l'interface d'Explorer.

B. En copiant le code compilé affiché dans Explorer.

C. En recherchant le modèle, en vérifiant son statut d'accès (s'il est public), et en visualisant ses colonnes et sa documentation pour s'assurer qu'il correspond à ses besoins.

D. En modifiant directement le niveau d'accès du modèle de private à public.  
**Réponse correcte : C**

### **QCM de Préparation - dbt Cloud Architect (Partie 2)**

#### **Thème 1 : Configuration des connexions à l'entrepôt de données**

1. **Quel est l'objectif principal du bouton "Test Connection" dans les paramètres d'un projet dbt Cloud ?**

A. Vérifier que la syntaxe du fichier dbt\_project.yml est correcte.

B. Confirmer que les informations d'identification fournies permettent à dbt Cloud de s'authentifier et d'exécuter une requête simple sur l'entrepôt.

C. S'assurer que le fournisseur Git est accessible.

D. Valider la licence de l'utilisateur actuel.  
**Réponse correcte : B**

1. **Lorsqu'un développeur utilise l'IDE, l'authentification à l'entrepôt est gérée via OAuth. Comment l'authentification est-elle gérée pour un job de production s'exécutant sur un calendrier ?**

A. Le job utilise les identifiants OAuth du développeur qui a créé le job.

B. Le job utilise un ensemble fixe d'identifiants (par exemple, login/mot de passe ou paire de clés) configuré dans les paramètres de connexion de l'environnement.

C. Le job demande une authentification manuelle à chaque exécution.

D. Le job n'a pas besoin de s'authentifier car il s'exécute dans l'infrastructure de dbt Cloud.  
**Réponse correcte : B**

1. **Si une entreprise n'utilise pas de pare-feu restreignant l'accès à son entrepôt de données cloud (ex: Snowflake, BigQuery), est-il quand même nécessaire de configurer la liste blanche d'IP ?**

A. Oui, c'est une étape obligatoire pour toutes les connexions dbt Cloud.

B. Non, si l'entrepôt est accessible publiquement sur Internet, la liste blanche d'IP n'est pas nécessaire pour établir une connexion.

C. Oui, car cela améliore les performances des requêtes.

D. Uniquement si vous utilisez des jetons de service.  
**Réponse correcte : B**

#### **Thème 2 : Configuration des connexions Git**

1. **Quelle est la principale raison pour laquelle dbt Cloud est si étroitement intégré à Git ?**

A. Pour stocker les logs d'exécution des jobs.

B. Pour permettre aux développeurs de partager des requêtes SQL.

C. Pour assurer le contrôle de version du code, permettre la collaboration et automatiser les déploiements (CI/CD).

D. Pour gérer les permissions des utilisateurs.  
**Réponse correcte : C**

1. **Un développeur essaie de "pull" les derniers changements dans l'IDE de dbt Cloud mais rencontre une erreur de fusion (merge conflict). Où doit-il résoudre ce conflit ?**

A. Directement dans l'interface utilisateur de dbt Cloud, qui dispose d'outils de résolution de conflits.

B. En utilisant des commandes Git sur son ordinateur local, puis en poussant la branche résolue vers le dépôt distant.

C. En supprimant le projet dbt Cloud et en le recréant.

D. En contactant le support de dbt.  
**Réponse correcte : B**

1. **Pour connecter un dépôt depuis un fournisseur Git auto-hébergé (comme GitLab Self-managed), que faut-il garantir du point de vue du réseau ?**

A. Que le fournisseur Git peut atteindre les adresses IP de dbt Cloud.

B. Que les serveurs de dbt Cloud peuvent atteindre l'URL du dépôt Git sur Internet.

C. Que tous les développeurs ont un accès VPN au Git auto-hébergé.

D. Que le dépôt est public.  
**Réponse correcte : B**

#### **Thème 3 : Création et maintenance des environnements**

1. **Comment un architecte peut-il forcer la rotation d'une paire de clés d'authentification pour un projet de manière programmatique ?**

A. En supprimant et recréant l'environnement dans l'interface utilisateur.

B. En utilisant le point de terminaison (endpoint) approprié de l'API d'administration de dbt Cloud.

C. En modifiant directement le fichier profiles.yml, ce qui déclenche une nouvelle génération de clés.

D. Ce n'est pas possible ; la rotation doit être manuelle.  
**Réponse correcte : B**

1. **À quel niveau le schéma/dataset par défaut d'un environnement est-il défini ?**

A. Dans la définition du job.

B. Dans les paramètres de la connexion à l'entrepôt.

C. Dans les paramètres de l'environnement lui-même.

D. Dans le profil personnel de chaque développeur.  
**Réponse correcte : C**

1. **Quel est l'avantage de configurer dbt pour qu'il puisse se référer ("defer") à un autre environnement ?**

A. Cela permet d'exécuter des jobs plus rapidement en utilisant moins de threads.

B. Dans un job de CI, cela permet de ne construire que les modèles modifiés et de lire les modèles non modifiés directement depuis l'environnement de production.

C. Cela copie automatiquement les données de l'environnement de production vers l'environnement de développement.

D. Cela active le lignage des données entre les environnements.  
**Réponse correcte : B**

#### **Thème 4 : Création et maintenance des définitions de jobs**

1. **Un job de CI est configuré pour s'exécuter sur chaque Pull Request. Quel est l'état de référence le plus courant pour la commande dbt build --select state:modified+ dans ce scénario ?**

A. Le dernier commit de la branche de la Pull Request.

B. Un catalogue dbt généré localement par un développeur.

C. Le manifeste (manifest.json) d'un job de production récent et fiable.

D. Le dernier job de CI qui a réussi.  
**Réponse correcte : C**

1. **À quoi sert le paramètre "Target Name" dans la configuration d'un job ?**

A. À nommer le schéma de destination des modèles.

B. À spécifier quel profil de profiles.yml utiliser pour cette exécution, permettant d'altérer par exemple le nom d'utilisateur ou le rôle.

C. À définir le nom de la branche Git à utiliser.

D. À spécifier le nom de la base de données cible.  
**Réponse correcte : B**

1. **Quelle est la fonction de l'étape "Generate Docs" dans un job ?**

A. Envoyer la documentation par e-mail aux parties prenantes.

B. Mettre à jour et publier le site web de la documentation du projet avec les derniers résultats d'exécution.

C. Créer des fichiers PDF de la documentation.

D. Valider la documentation YAML des modèles.  
**Réponse correcte : B**

#### **Thème 5 : Configuration de la sécurité et des licences**

1. **Lors de la mise en place du SSO (Single Sign-On), que représentent l'"Identifier or Entity ID" et l'"Assertion Consumer Service (ACS) URL" fournis par dbt Cloud ?**

A. Les informations d'identification de l'administrateur dbt.

B. Les informations que vous devez fournir à votre fournisseur d'identité (ex: Okta, Azure AD) pour qu'il sache qui est dbt Cloud et où renvoyer l'utilisateur après authentification.

C. Les URLs pour accéder à l'API de dbt Cloud.

D. Les clés de licence pour activer le SSO.  
**Réponse correcte : B**

1. **Un utilisateur se voit attribuer une licence "Read-Only". À quoi peut-il accéder ?**

A. Il peut uniquement voir la page d'accueil de dbt Cloud.

B. Il peut consulter la documentation du projet, voir les définitions des jobs et l'historique des exécutions, mais ne peut pas utiliser l'IDE ni modifier les configurations.

C. Il peut utiliser l'IDE mais ne peut pas commiter de code.

D. Il a les mêmes droits qu'un développeur mais ne peut pas déclencher de jobs.  
**Réponse correcte : B**

1. **Quelle est la meilleure pratique pour gérer les permissions d'un jeton de service destiné à un outil d'orchestration externe ?**

A. Lui donner les permissions "Account Admin" pour éviter tout problème d'accès.

B. Créer un groupe de permissions personnalisé avec le principe du moindre privilège, n'accordant que les accès strictement nécessaires (ex: "Job Runner").

C. Utiliser le même jeton pour toutes les intégrations externes.

D. Ne pas utiliser de jetons de service et fournir des identifiants utilisateur à l'outil externe.  
**Réponse correcte : B**

#### **Thème 6 : Mise en place du monitoring et des alertes**

1. **En plus des notifications de début et de fin de job, quel autre type d'événement peut déclencher un webhook dans dbt Cloud ?**

A. Un commit est poussé sur la branche main.

B. Un utilisateur se connecte via SSO.

C. Un modèle spécifique rencontre une erreur de test de données.

D. Une nouvelle Pull Request est ouverte (si l'intégration GitHub est configurée).  
**Réponse correcte : D**

1. **Comment configurer une alerte par e-mail pour qu'elle soit envoyée à toute une équipe d'astreinte ?**

A. En ajoutant les adresses e-mail de chaque membre de l'équipe, séparées par des virgules, dans le champ de notification du job.

B. En utilisant l'adresse e-mail d'une liste de distribution ou d'un groupe (ex: equipe-data@entreprise.com).

C. Il n'est possible d'alerter qu'un seul utilisateur par job.

D. En utilisant un webhook qui déclenche un service d'envoi d'e-mails.  
**Réponse correcte : B**

1. **Un job échoue. Le webhook configuré est-il envoyé immédiatement au moment de l'échec de l'étape, ou à la fin de l'exécution complète du job ?**

A. Immédiatement après l'échec de la commande.

B. Uniquement si l'utilisateur annule manuellement le job.

C. À la fin de l'exécution du job, lorsque son statut final est déterminé comme "Error".

D. Le webhook n'est pas envoyé en cas d'échec, seulement en cas de succès.  
**Réponse correcte : C**

#### **Thème 7 : Mise en place d'un dbt Mesh et références inter-projets**

1. **Quelle est la principale différence entre un modèle avec un accès public et un modèle avec un accès protected ?**

A. public peut être référencé par n'importe quel autre projet, protected n'existe pas.

B. public peut être référencé par tous les autres projets, protected peut être référencé par d'autres projets mais pas par des projets en aval de ceux-ci.

C. public est visible dans la documentation, protected ne l'est pas.

D. public peut être référencé par n'importe quel autre projet, tandis que protected peut être référencé par d'autres projets du même group.  
**Réponse correcte : D** (Note : protected est une évolution de la gouvernance dans des versions plus récentes).

1. **Dans le fichier dbt\_project.yml, comment déclarez-vous une dépendance sur un autre projet dbt Cloud ?**

A. dependencies: { project: 'nom\_autre\_projet' }

B. models: { depends\_on: ['nom\_autre\_projet'] }

C. projects: { nom\_autre\_projet: {} }

D. dependencies: { projects: [{ project: 'nom\_autre\_projet' }] }  
**Réponse correcte : D**

1. **Si vous référencez un modèle d'un projet en amont, mais que votre job de CI ne se réfère pas ("defer") à un état de production, que se passera-t-il ?**

A. dbt trouvera intelligemment le modèle dans la production et l'utilisera.

B. L'exécution réussira mais affichera un avertissement.

C. L'exécution échouera car dbt ne trouvera pas le modèle référencé dans les nœuds de l'exécution actuelle.

D. dbt essaiera de reconstruire le modèle du projet en amont.  
**Réponse correcte : C**

#### **Thème 8 : Configuration et utilisation de dbt Explorer**

1. **Comment dbt Explorer représente-t-il une référence inter-projet dans le graphe de lignage ?**

A. Il ne peut pas afficher les références inter-projets.

B. Il affiche une icône spéciale ou une bordure de couleur différente sur le nœud du modèle en amont pour indiquer qu'il provient d'un autre projet.

C. Il affiche une ligne en pointillé vers le projet en amont.

D. Il affiche uniquement le nom du modèle, sans indiquer son origine.  
**Réponse correcte : B**

1. **Un test de données échoue sur un modèle. Comment un développeur peut-il utiliser dbt Explorer pour investiguer ?**

A. En relançant le test directement depuis l'interface d'Explorer.

B. En sélectionnant le modèle, en visualisant ses parents (modèles en amont) et ses sources pour comprendre d'où viennent les données qui pourraient causer l'échec.

C. En modifiant le code du test dans l'éditeur de code d'Explorer.

D. Explorer n'est pas utilisé pour le débogage, seulement pour la visualisation.  
**Réponse correcte : B**

1. **Quelle fonctionnalité d'Explorer permet de comprendre l'impact potentiel de la modification d'un modèle ?**

A. Le profileur de requêtes.

B. La capacité à voir tous les enfants (modèles, expositions, etc.) qui dépendent directement ou indirectement du modèle sélectionné.

C. L'intégration avec le fournisseur Git pour voir les diffs.

D. Le bouton "Simuler l'impact".  
**Réponse correcte : B**

### **QCM de Préparation - dbt Cloud Architect (Partie 3)**

#### **Thème 1 : Configuration des connexions à l'entrepôt de données**

1. **Un test de connexion échoue avec une erreur de "timeout". Quelle est la cause la plus probable ?**

A. Le mot de passe de l'utilisateur de l'entrepôt est incorrect.

B. Un problème réseau, comme un pare-feu ou une liste blanche d'IP mal configurée, empêche dbt Cloud d'atteindre l'entrepôt.

C. La version de dbt spécifiée dans le projet est incompatible.

D. Le dépôt Git associé est inaccessible.  
**Réponse correcte : B**

1. **Quelle est la différence fondamentale entre l'authentification via OAuth et l'authentification par paire de clés dans le contexte de dbt Cloud ?**

A. OAuth est moins sécurisé que la paire de clés.

B. OAuth est utilisé pour l'authentification des utilisateurs individuels (IDE), tandis que la paire de clés est utilisée pour l'authentification des services/jobs.

C. La paire de clés ne fonctionne qu'avec Snowflake, tandis qu'OAuth est universel.

D. Il n'y a pas de différence fondamentale, ce sont deux noms pour la même chose.  
**Réponse correcte : B**

1. **Si vous mettez à jour les informations d'identification de votre entrepôt (par exemple, le mot de passe), où devez-vous principalement effectuer ce changement dans dbt Cloud ?**

A. Dans chaque job qui utilise la connexion.

B. Dans les paramètres de la connexion au niveau du projet.

C. Dans votre profil utilisateur personnel.

D. Dans le fichier dbt\_project.yml.  
**Réponse correcte : B**

#### **Thème 2 : Configuration des connexions Git**

1. **Quelle est la fonction du bouton "Import Project" sur la page d'accueil de dbt Cloud ?**

A. Importer des données depuis un fichier CSV.

B. Cloner un dépôt Git existant pour créer un nouveau projet dbt Cloud.

C. Importer une configuration d'environnement depuis un autre projet.

D. Installer un package dbt.  
**Réponse correcte : B**

1. **Un développeur ajoute \*.csv au fichier .gitignore dans son projet dbt. Quel sera l'effet dans dbt Cloud ?**

A. Les fichiers CSV ne pourront plus être importés via la commande dbt seed.

B. Les fichiers CSV présents dans le dépôt ne seront pas clonés dans l'environnement de développement de dbt Cloud.

C. Cela n'a aucun effet, car dbt Cloud ignore le fichier .gitignore.

D. Les fichiers CSV seront automatiquement supprimés après chaque exécution de dbt seed.  
**Réponse correcte : B**

1. **Lors de la configuration de l'intégration avec GitHub, pourquoi est-il recommandé d'utiliser l'application GitHub plutôt qu'un jeton d'accès personnel ?**

A. L'application GitHub est la seule méthode qui fonctionne avec les dépôts privés.

B. L'application GitHub fournit des permissions plus granulaires et n'est pas liée à un compte utilisateur personnel, ce qui est plus sécurisé et robuste.

C. Le jeton d'accès personnel ne permet pas de déclencher des jobs de CI.

D. C'est plus rapide à configurer.  
**Réponse correcte : B**

#### **Thème 3 : Création et maintenance des environnements**

1. **Quelle est la principale distinction entre un environnement de "Développement" et un environnement de "Déploiement" ?**

A. Les environnements de développement ne peuvent pas exécuter de jobs.

B. Les environnements de développement sont liés aux informations d'identification personnelles du développeur, tandis que les environnements de déploiement utilisent des identifiants fixes.

C. Seuls les environnements de déploiement peuvent générer de la documentation.

D. Les environnements de développement ne peuvent pas utiliser de variables d'environnement.  
**Réponse correcte : B**

1. **Un administrateur souhaite empêcher les développeurs de lancer des exécutions sur l'environnement de Production. Comment peut-il configurer cela ?**

A. En ne leur donnant pas le mot de passe de l'entrepôt de production.

B. En utilisant le contrôle d'accès basé sur les rôles (RBAC) pour créer un ensemble de permissions qui n'inclut pas l'accès à cet environnement.

C. En cachant l'environnement de la liste.

D. En mettant la branche main en mode protégé sur GitHub.  
**Réponse correcte : B**

1. **Un projet a besoin d'une clé API tierce qui est différente pour la production et le développement. Quelle est la meilleure façon de gérer cela ?**

A. Écrire la clé directement dans le code du modèle avec une instruction if.

B. Définir une variable d'environnement dans dbt Cloud avec une valeur différente pour l'environnement de développement et de production.

C. Stocker les deux clés dans le dépôt Git.

D. Demander aux développeurs de la saisir manuellement à chaque fois.  
**Réponse correcte : B**

#### **Thème 4 : Création et maintenance des définitions de jobs**

1. **Qu'est-ce que le "self deferral" (ou "defer to self") pour un job ?**

A. C'est lorsqu'un job se réfère à l'état de sa propre exécution précédente réussie.

B. C'est un job qui ne peut pas être déclenché par un autre job.

C. C'est une fonctionnalité pour reporter l'exécution d'un job s'il est déjà en cours.

D. C'est la même chose que de ne pas utiliser le "deferral".  
**Réponse correcte : A**

1. **À quoi sert le paramètre threads dans la configuration d'un job ?**

A. À définir le nombre de jobs pouvant s'exécuter en parallèle.

B. À spécifier le nombre de modèles que dbt peut construire en parallèle au sein d'une même exécution, afin d'optimiser la vitesse.

C. À contrôler le nombre de développeurs qui peuvent travailler sur un projet simultanément.

D. À définir le nombre de notifications par e-mail à envoyer.  
**Réponse correcte : B**

1. **Quelle est la principale utilité des paramètres avancés de CI ("Advanced CI settings") ?**

A. Ils permettent de choisir la couleur des notifications Slack.

B. Ils permettent de déclencher automatiquement un job de CI uniquement lorsque des fichiers dans des chemins spécifiques ont été modifiés.

C. Ils permettent d'exécuter la CI sur des machines plus puissantes.

D. Ils activent le "self deferral".  
**Réponse correcte : B**

#### **Thème 5 : Configuration de la sécurité et des licences**

1. **Un utilisateur quitte l'entreprise. Lorsqu'un administrateur supprime son compte de dbt Cloud, que deviennent les jobs qu'il a créés ?**

A. Ils sont automatiquement supprimés.

B. Ils sont automatiquement réassignés à l'administrateur du compte.

C. Ils deviennent orphelins et ne peuvent plus être modifiés jusqu'à ce qu'un nouveau propriétaire soit assigné.

D. Ils continuent de fonctionner mais ne peuvent plus être modifiés.  
**Réponse correcte : B**

1. **Quelle est la différence entre l'assignation de permissions à un utilisateur individuel et à un groupe ?**

A. Les permissions de groupe ont toujours priorité sur les permissions individuelles.

B. L'assignation par groupe simplifie la gestion à grande échelle, assurant que tous les membres d'une équipe ont les mêmes droits de manière cohérente.

C. Les permissions individuelles sont temporaires, tandis que les permissions de groupe sont permanentes.

D. Il n'y en a aucune ; c'est juste une question de préférence visuelle.  
**Réponse correcte : B**

1. **Pour mettre en œuvre le RBAC, quelle est la première étape logique ?**

A. Assigner des utilisateurs à des projets.

B. Créer des groupes qui reflètent les équipes ou les rôles (ex: "Data Analysts", "Analytics Engineers").

C. Créer les jetons de service.

D. Configurer la connexion à l'entrepôt.  
**Réponse correcte : B**

#### **Thème 6 : Mise en place du monitoring et des alertes**

1. **Que se passe-t-il si l'URL d'un webhook configuré dans un job n'est plus valide ou que le service récepteur est en panne ?**

A. Le job dbt Cloud échoue avec une erreur de notification.

B. Le job dbt Cloud termine son exécution normalement, mais la notification est perdue.

C. dbt Cloud essaie de renvoyer la notification toutes les heures pendant 24 heures.

D. Une alerte par e-mail est envoyée à l'administrateur du compte pour l'informer de l'échec du webhook.  
**Réponse correcte : B**

1. **Pour déboguer rapidement un échec de job à partir d'une notification Slack (envoyée via webhook), quelle information est la plus cruciale à inclure dans le message personnalisé ?**

A. La version de dbt utilisée.

B. Le nom de la personne qui a écrit le code.

C. Un lien direct vers la page de l'exécution (Run URL) dans dbt Cloud.

D. Le coût estimé de l'exécution.  
**Réponse correcte : C**

1. **Un job contient 100 modèles. Est-il possible de configurer une alerte qui se déclenche si un modèle spécifique échoue, mais pas les autres ?**

A. Oui, via les paramètres de notification avancés au niveau du modèle.

B. Non, les notifications de dbt Cloud fonctionnent au niveau du job (réussite ou échec global).

C. Oui, en utilisant l'API dbt pour surveiller les logs en temps réel.

D. Oui, si le modèle est dans un projet séparé.  
**Réponse correcte : B**

#### **Thème 7 : Mise en place d'un dbt Mesh et références inter-projets**

1. **Le Projet A et le Projet B définissent tous deux un modèle public nommé dim\_customers. Dans le Projet C, qui dépend des deux, que se passe-t-il lors de l'appel à {{ ref('dim\_customers') }} ?**

A. dbt utilise le modèle du premier projet déclaré dans dbt\_project.yml.

B. dbt demande à l'utilisateur de choisir lequel utiliser lors de l'exécution.

C. L'exécution échoue en raison d'une référence ambiguë, car dbt ne peut pas choisir entre les deux.

D. dbt fusionne les deux modèles en un seul.  
**Réponse correcte : C**

1. **Un modèle est défini comme private. Peut-il être référencé par un autre modèle au sein du MÊME projet ?**

A. Non, private signifie qu'il ne peut être référencé par aucun autre modèle.

B. Oui, le modificateur d'accès (public, private, etc.) ne régit que les références inter-projets.

C. Uniquement si les modèles sont dans le même sous-dossier. D. Uniquement si l'on utilise la fonction source() au lieu de ref().  
**Réponse correcte : B**

1. **Comment le type d'environnement (développement vs. déploiement) affecte-t-il les références inter-projets ?**

A. Elles ne fonctionnent que dans les environnements de déploiement.

B. En développement, les références se résolvent aux modèles construits dans votre schéma de développement, tandis qu'en déploiement, elles dépendent du "deferral" vers un état de production.

C. Le type d'environnement n'a aucun impact sur leur fonctionnement.

D. En développement, toutes les références inter-projets sont automatiquement ignorées.  
**Réponse correcte : B**

#### **Thème 8 : Configuration et utilisation de dbt Explorer**

1. **Comment dbt Explorer indique-t-il qu'un modèle est matérialisé en tant que table incrémentale ?**

A. Le nom du nœud est en italique.

B. Il affiche une icône ou une couleur spécifique sur le nœud pour indiquer la stratégie de matérialisation.

C. Il n'est pas possible de voir la stratégie de matérialisation dans Explorer.

D. Il faut cliquer sur le nœud pour voir son code SQL.  
**Réponse correcte : B**

1. **Un nouvel analyste rejoint l'équipe. Comment peut-il utiliser dbt Explorer pour comprendre rapidement la structure du projet ?**

A. En visualisant le graphe pour identifier les sources, les modèles de "staging", intermédiaires et "marts", et comprendre le flux de données.

B. En exécutant toutes les commandes dbt depuis Explorer.

C. En lisant les logs de tous les jobs passés.

D. En regardant la liste des fichiers dans l'IDE.  
**Réponse correcte : A**

1. **Comment les tests sont-ils affichés dans dbt Explorer ?**

A. Ils ne sont pas visibles dans Explorer.

B. Ils apparaissent comme des nœuds dans le graphe, connectés aux modèles qu'ils testent, souvent avec une couleur distincte (ex: vert, jaune, rouge) pour indiquer leur statut.

C. Ils sont listés dans un tableau séparé, sans lien visuel avec les modèles.

D. Uniquement les tests qui ont échoué sont affichés.  
**Réponse correcte : B**

### **QCM de Préparation - dbt Cloud Architect (Partie 4)**

#### **Thème 1 : Configuration des connexions à l'entrepôt de données**

1. **Un job s'exécute correctement en production, mais un développeur ne peut pas exécuter de requêtes dans l'IDE avec le même projet. Quelle est la cause la plus probable ?**

A. Le dépôt Git n'est pas synchronisé.

B. La configuration OAuth pour l'authentification de l'utilisateur dans l'IDE est incorrecte ou ses permissions sur l'entrepôt sont manquantes, même si les identifiants du job sont corrects.

C. La version de dbt de l'environnement est obsolète.

D. Le développeur n'a pas de licence dbt Cloud.  
**Réponse correcte : B**

1. **Quel est le principal risque de sécurité si vous n'utilisez PAS la liste blanche d'IP (IP Whitelisting) pour un entrepôt qui le permet ?**

A. Les performances des jobs seront dégradées.

B. Les jobs dbt Cloud pourraient échouer de manière aléatoire.

C. Si les informations d'identification de votre entrepôt sont compromises, un attaquant pourrait s'y connecter depuis n'importe où sur Internet.

D. Vous ne pourrez pas utiliser l'authentification OAuth.  
**Réponse correcte : C**

1. **Lors de la création d'une nouvelle connexion de projet, est-il obligatoire de la lier immédiatement à un dépôt Git ?**

A. Oui, la connexion à l'entrepôt ne peut être testée sans une connexion Git active.

B. Non, vous pouvez configurer et tester la connexion à l'entrepôt avant de configurer l'intégration Git.

C. Oui, sinon les informations d'identification de l'entrepôt ne peuvent pas être enregistrées.

D. Non, mais vous ne pourrez pas utiliser l'IDE.  
**Réponse correcte : B**

#### **Thème 2 : Configuration des connexions Git**

1. **Un développeur commite accidentellement un très gros fichier (ex: 500 Mo) dans le dépôt. Quel impact cela peut-il avoir sur dbt Cloud ?**

A. Aucun, dbt Cloud ignore les fichiers non-.sql et non-.yml.

B. Le clonage du dépôt dans l'IDE et au début de chaque job sera beaucoup plus lent, augmentant les temps d'attente et d'exécution.

C. Le projet sera corrompu et devra être réimporté.

D. La connexion à l'entrepôt de données échouera.  
**Réponse correcte : B**

1. **Qu'est-ce que l'option "git clone" dans l'IDE de dbt Cloud permet de faire ?**

A. Créer une nouvelle branche.

B. Pousser les changements vers le dépôt distant.

C. Réinitialiser la branche locale à l'état du dépôt distant, effaçant les changements non commités.

D. Lancer un job de CI.  
**Réponse correcte : C**

1. **Si vous changez le nom de la branche par défaut de votre dépôt de master à main, que devez-vous mettre à jour dans dbt Cloud ?**

A. Rien, dbt Cloud détecte automatiquement le changement.

B. Vous devez mettre à jour la branche configurée pour votre environnement de production et potentiellement d'autres configurations qui référencent l'ancienne branche.

C. Uniquement la configuration de votre profil utilisateur.

D. Vous devez réimporter le projet.  
**Réponse correcte : B**

#### **Thème 3 : Création et maintenance des environnements**

1. **Comment pouvez-vous utiliser une variable d'environnement pour que les développeurs travaillent sur un échantillon de données plutôt que sur la totalité des données de production ?**

A. En créant une variable is\_development et en l'utilisant dans le code SQL pour ajouter une clause LIMIT ou WHERE.

B. En nommant l'environnement "sample\_data".

C. C'est impossible, les développeurs travaillent toujours sur la totalité des données.

D. En utilisant une variable pour changer le target.schema.  
**Réponse correcte : A**

1. **Quelle est la bonne pratique pour la gestion du schéma dans un environnement de développement ?**

A. Tous les développeurs écrivent dans le même schéma pour faciliter la collaboration.

B. Chaque développeur se voit attribuer un schéma unique pour éviter les conflits lors de la construction de modèles.

C. Les schémas de développement sont supprimés après chaque commit.

D. Les développeurs n'ont pas besoin d'écrire dans un schéma, ils lisent uniquement la production.  
**Réponse correcte : B**

1. **Un environnement de Staging est créé pour tester les Pull Requests avant de les fusionner dans main. Quelle configuration est la plus logique pour cet environnement ?**

A. Il devrait être lié à la branche main et utiliser les mêmes identifiants que la production.

B. Il devrait être un environnement de développement sans jobs associés.

C. Il devrait être un environnement de déploiement, déclenché par les jobs de CI, et écrire dans un schéma ou une base de données séparée (staging\_db).

D. Il devrait utiliser le "self deferral".  
**Réponse correcte : C**

#### **Thème 4 : Création et maintenance des définitions de jobs**

1. **Quelle est la différence entre dbt build et exécuter dbt run puis dbt test comme étapes séparées ?**

A. Il n'y a aucune différence, dbt build est juste un raccourci.

B. dbt build exécute les tests d'un modèle immédiatement après sa construction, ce qui peut être plus efficace pour identifier les erreurs plus tôt.

C. dbt build ne peut pas être utilisé dans les jobs de CI.

D. dbt build exécute également dbt seed et dbt snapshot.  
**Réponse correcte : B**

1. **Un job échoue de manière intermittente. Quelle est la première étape de débogage dans l'interface de dbt Cloud ?**

A. Augmenter le nombre de threads.

B. Changer la version de dbt.

C. Examiner attentivement les logs de l'exécution échouée pour identifier l'étape et le message d'erreur précis.

D. Supprimer et recréer le job.  
**Réponse correcte : C**

1. **Quelle est la fonction de la commande dbt source freshness dans un job ?**

A. Mettre à jour la documentation du projet.

B. Vérifier si les données des sources brutes sont à jour par rapport aux seuils définis, et générer des erreurs ou des avertissements si elles sont obsolètes.

C. Rafraîchir les tables sources en rechargeant les données.

D. Tester la connexion aux sources de données.  
**Réponse correcte : B**

#### **Thème 5 : Configuration de la sécurité et des licences**

1. **Un auditeur externe demande un rapport de toutes les permissions des utilisateurs pour un projet spécifique. Quelle fonctionnalité de dbt Cloud Enterprise est la plus adaptée pour répondre à cette demande ?**

A. L'historique des exécutions de jobs.

B. L'Audit Log, qui trace les modifications de permissions.

C. Les écrans de configuration des Groupes et des Ensembles de Permissions (Permission Sets).

D. Les webhooks de notification.  
**Réponse correcte : C** (combiné avec l'Audit Log pour l'historique).

1. **Pourquoi est-il important de configurer des groupes et des ensembles de permissions (Permission Sets) même pour une petite équipe ?**

A. Pour utiliser plus de licences dbt Cloud.

B. Pour préparer la scalabilité de l'équipe et s'assurer que les nouvelles personnes reçoivent les bonnes permissions de manière cohérente et sécurisée.

C. Pour améliorer les performances des jobs.

D. C'est une obligation pour utiliser l'intégration Git.  
**Réponse correcte : B**

1. **Quel type de permission est nécessaire pour pouvoir modifier les paramètres de facturation d'un compte dbt Cloud ?**

A. Job Admin

B. Account Admin

C. Billing Admin

D. Security Admin  
**Réponse correcte : C** et B, car Account Admin inclut généralement ces droits.

#### **Thème 6 : Mise en place du monitoring et des alertes**

1. **Votre équipe trouve que les notifications par e-mail sont trop nombreuses. Comment pourriez-vous les rendre moins "bruyantes" ?**

A. En désactivant toutes les notifications par e-mail.

B. En modifiant les jobs pour qu'ils n'envoient des notifications que pour les exécutions sur la branche main, et non pour les jobs de CI.

C. En créant une règle dans votre client de messagerie pour supprimer les e-mails de dbt.

D. En réduisant la fréquence d'exécution des jobs.  
**Réponse correcte : B**

1. **Quel est un cas d'usage pour déclencher un webhook à la fois au début (on\_start) et à la fin (on\_success/on\_failure) d'un job ?** A. Pour s'assurer que la notification est bien reçue. B. Pour mesurer précisément la durée d'un job dans un système de monitoring externe. C. C'est une mauvaise pratique qui crée des doublons. D. Pour tester le système de webhook.  
   **Réponse correcte : B**
2. **Les notifications dbt Cloud sont-elles capables de vous alerter si un job n'a PAS démarré comme prévu selon son calendrier ?**

A. Oui, il existe une alerte spécifique "Scheduled run did not start".

B. Non, les notifications de dbt Cloud ne se déclenchent que pour les événements liés à une exécution de job (début, succès, échec).

C. Uniquement si vous utilisez des webhooks.

D. Oui, mais seulement pour les comptes Enterprise.  
**Réponse correcte : B**

#### **Thème 7 : Mise en place d'un dbt Mesh et références inter-projets**

1. **Quel est l'objectif principal de la configuration de group sur un modèle dans le contexte de dbt Mesh ?**

A. Pour organiser visuellement les modèles dans la documentation.

B. Pour permettre à des modèles avec un accès protected d'être référencés par d'autres modèles du même groupe, même s'ils sont dans des projets différents.

C. Pour assigner des propriétaires de code à un ensemble de modèles.

D. Pour améliorer les performances d'exécution des modèles de ce groupe.  
**Réponse correcte : B** (combiné avec C).

1. **Un projet finance dépend d'un modèle public du projet core\_data. Si l'équipe de core\_data modifie ce modèle public d'une manière qui casse la dépendance, comment l'équipe finance sera-t-elle alertée ?**

A. dbt Cloud envoie une notification automatique aux propriétaires des projets dépendants.

B. Le job de CI du projet finance échouera lors de la prochaine exécution, signalant le problème.

C. La modification sera automatiquement bloquée si elle casse des dépendances en aval.

D. L'alerte se fait par communication manuelle ; dbt ne gère pas ce type de dépendance sociale.  
**Réponse correcte : B**

1. **Est-il possible d'utiliser dbt source freshness sur une table qui est en fait un modèle public d'un autre projet ?**

A. Non, la fraîcheur ne peut être vérifiée que sur les sources brutes définies.

B. Oui, en définissant ce modèle comme une source dans le projet aval, vous pouvez vérifier sa fraîcheur comme n'importe quelle autre source.

C. Non, car cela violerait les principes de dbt Mesh.

D. Uniquement si les deux projets sont dans le même dépôt Git.  
**Réponse correcte : B**

#### **Thème 8 : Configuration et utilisation de dbt Explorer**

1. **Comment un architecte peut-il utiliser dbt Explorer pour évaluer le "rayon d'impact" (blast radius) de la modification d'une source de données ?**

A. En simulant une exécution.

B. En sélectionnant le nœud de la source et en examinant tous ses enfants (dépendances en aval) pour voir combien de modèles et d'expositions seront affectés.

C. En regardant l'historique Git du fichier de configuration de la source.

D. En comptant le nombre de colonnes dans la source.  
**Réponse correcte : B**

1. **Dans dbt Explorer, que signifie un nœud de source avec un contour rouge ?**

A. La source n'est plus utilisée et peut être supprimée.

B. La dernière vérification de source freshness a échoué ou est obsolète ("error" state).

C. La source contient des données personnelles (PII).

D. La source est en cours de modification par un autre développeur.  
**Réponse correcte : B**

1. **dbt Explorer peut-il remplacer complètement la nécessité de lire le code SQL d'un projet ?**

A. Oui, c'est un outil entièrement visuel qui rend le code obsolète.

B. Non, c'est un outil de visualisation et de découverte qui complète la lecture du code ; il ne montre pas la logique métier complexe au sein du SQL.

C. Oui, car il contient un traducteur de SQL en langage naturel.

D. Non, car il n'affiche que les modèles qui ont réussi leurs tests.  
**Réponse correcte : B**

Voilà qui complète un QCM de 96 questions couvrant tous les aspects de la certification. Bonnes révisions et bonne chance pour votre examen ! 🚀