

Mis à jour le 23/07/2025

S'inscrire

Formation Redshift Data Warehousing

3 jours (21 heures)

Présentation

Maîtrisez Amazon Redshift dans sa dimension experte grâce à cette formation avancée, conçue pour les architectes et ingénieurs data exigeants. Vous apprendrez à exploiter tout le potentiel de l'entrepôt de données Redshift, en optimisant les performances, en industrialisant vos pipelines, et en garantissant sécurité, gouvernance et scalabilité dans des architectures cloud de haut niveau.

La formation commence par un deep dive dans l'architecture MPP de Redshift, les types de nœuds, les mécanismes de distribution, de tri, de compression, et les meilleures pratiques pour booster vos requêtes SQL et exploiter les plans d'exécution.

Vous apprendrez à construire des pipelines d'ingestion efficaces, en combinant COPY, S3, Kinesis, Glue et dbt, tout en appliquant les stratégies de sécurité avancées (IAM, RLS, CLS) et d'audit via les vues système.

L'industrialisation est au cœur du programme : CI/CD avec Terraform, modélisation avec dbt, promotion entre environnements, intégration CloudWatch et diagnostic de performance deviennent vos outils du quotidien.

Comme pour toutes nos formations, celle-ci vous sera présentée avec les toutes dernières actualisations de [redshift data](#)

Objectifs

- Comprendre l'architecture avancée de Redshift et ses spécificités MPP, stockage colonne et scaling
- Savoir concevoir, optimiser et auditer des modèles de données performants dans Redshift
- Maîtriser les stratégies d'ingestion de données massives via COPY, Kinesis, Glue ou Spectrum
- Être capable de sécuriser, gouverner et tracer les accès et traitements sur des environnements Redshift critiques

- Savoir industrialiser les déploiements Redshift avec Terraform et automatiser les traitements via dbt et CI/CD
- Comprendre et exploiter Redshift Data Sharing, le mode Lakehouse et les architectures analytiques hybrides multicloud

Public visé

- Data Engineers
- Administrateurs Redshift

Pré-requis

- Maîtriser les fondamentaux du cloud AWS

Programme de la formation Redshift Data Warehousing

Rappels et évolutions avancées de Redshift

- Architecture MPP Redshift : deep dive
- Redshift RA3 vs DS2 : impact sur la performance
- Redshift Spectrum, AQUA, Concurrency Scaling, Data Sharing
- Intégration avec S3, Glue, Athena

Gestion avancée du stockage

- Stockage columnar et compression
- Tri (SORTKEY) et distribution (DISTKEY) optimaux
- Gestion du vacuum et du Analyze
- Monitoring du stockage

Optimisation des performances

- Analyse des plans d'exécution
- Tuning SQL et anti-patterns
- Matérialisation des vues : stratégies et cas d'usage

- Partitionnement logique via UNLOAD/EXTERNAL TABLES
- Atelier pratique : Optimisation d'un cluster existant, refactoring de requêtes lentes

Chargement massif et pipelines de données

- COPY optimisé depuis S3, Kinesis ou Glue
- Gestion des erreurs et de la reprise
- Redshift Streaming Ingestion
- ELT avec Redshift + dbt

Gouvernance, sécurité et accès

- Sécurité au niveau ligne et colonne (RLS/CLS)
- IAM vs authentification native
- Audit avec STL/ SVL / SVV tables
- Intégration avec Lake Formation et AWS KMS

Monitoring, diagnostic et audit

- Utilisation de STL_ALERT_EVENT_LOG, SVL_S3LOG
- Intégration CloudWatch, EventBridge, Lambda
- Audit des connexions, requêtes lentes, contention
- Atelier pratique : Création d'un pipeline d'ingestion sécurisé + dashboards de monitoring

Redshift Data Sharing et architecture multi-cluster

- Redshift Data Sharing entre comptes
- Cas d'usages multi-équipes, multi-environnements
- Sécurité et isolation des datasets

CI/CD et industrialisation

- Déploiement Redshift avec Terraform / CloudFormation
- dbt pour la modélisation, la documentation et les tests
- Versioning, revue de code, promotion entre environnements

Cas d'usage avancés et architecture hybride

- Lakehouse avec Redshift + Iceberg
- Redshift + Apache Hudi / Delta Lake
- Fédérations de requêtes avec sources externes
- Atelier final : Mise en place d'un projet complet Data Lake + Redshift + CI/CD + Monitoring

Sociétés concernées

Cette formation s'adresse à la fois aux particuliers ainsi qu'aux entreprises, petites ou grandes, souhaitant former ses équipes à une nouvelle technologie informatique avancée ou bien à acquérir des connaissances métiers spécifiques ou des méthodes modernes.

Positionnement à l'entrée en formation

Le positionnement à l'entrée en formation respecte les critères qualité Qualiopi. Dès son inscription définitive, l'apprenant reçoit un questionnaire d'auto-évaluation nous permettant d'apprécier son niveau estimé sur différents types de technologies, ses attentes et objectifs personnels quant à la formation à venir, dans les limites imposées par le format sélectionné. Ce questionnaire nous permet également d'anticiper certaines difficultés de connexion ou de sécurité interne en entreprise (intraentreprise ou classe virtuelle) qui pourraient être problématiques pour le suivi et le bon déroulement de la session de formation.

Méthodes pédagogiques

Stage Pratique : 60% Pratique, 40% Théorie. Support de la formation distribué au format numérique à tous les participants.

Organisation

Le cours alterne les apports théoriques du formateur soutenus par des exemples et des séances de réflexions, et de travail en groupe.

Validation

À la fin de la session, un questionnaire à choix multiples permet de vérifier l'acquisition correcte des compétences.

Sanction

Une attestation sera remise à chaque stagiaire qui aura suivi la totalité de la formation.